EMPRENDIMIENTO PREDICTIVO

USO DE MODELOS DE INTELIGENCIA ARTIFICIAL PARA ANTICIPAR EL ÉXITO EMPRESARIAL



Adriana Yamilet Herrera Granizo
Gema Viviana Paula Alarcón
Lenin Rómulo Briones Borja
Rafael Fernando Carrera Zurita

Editorial Andes Cognitio

EMPRENDIMIENTO PREDICTIVO: Uso de modelos de inteligencia artificial para anticipar el éxito empresarial

© Autores

Adriana Yamilet Herrera Granizo
Correo: ayherrera@espe.edu.ec
Orcid: https://orcid.org/0000-0001-5090-5981
Institución: Universidad de las Fuerzas Armadas ESPE. Sangolquí, Ecuador

Gema Viviana Paula Alarcón
Correo: gemapaula@unach.edu.ec
Orcid: https://orcid.org/0000-0001-8968-0740
Institución: Universidad Nacional de Chimborazo. Riobamba, Ecuador

Lenin Rómulo Briones Borja
Correo: lrbriones2@espe.edu.ec
Orcid: https://orcid.org/0000-0003-3184-492X
Institución: Universidad de las Fuerzas Armadas ESPE. Sangolquí, Ecuador

Rafael Fernando Carrera Zurita
Correo: rfcarrera1@espe.edu.ec
Orcid: https://orcid.org/0009-0000-6667-2123
Institución: Universidad de las Fuerzas Armadas ESPE. Sangolquí, Ecuador

Editorial "ANDES COGNITIO EDAC S.A.S."
DEPARTAMENTO DE EDICIÓN

Editado y Distribuido por:

Editorial:	Andes Cognitio
Sello Editorial:	978-9942-7408
Teléfono:	0995805659
Web:	https://andescognitio.org
ISBN:	978-9907-9566-8-9
DOI:	https://doi.org/10.64230/924mf502

© Primera Edición
© Noviembre 2026

Impreso en Ecuador

Revisión de Ortografía
Lcda. Cristina Paola Chamorro Ortega

Diseño de Portada
Ing. Pamela Rosa Taco Hernández Mgs

Diagramación
Ing. Yoselyn Andrea Rogel Gaibor

Director Editorial
Ec. Juan F. Villacis U. Mgs.

Aviso Legal
El contenido de este libro incluyendo textos, imágenes, gráficos, tablas, cuadros y referencias bibliográficas es de exclusiva responsabilidad del/ de los autor (es). Las opiniones, datos y criterios expresados no representan necesariamente la postura institucional ni el pensamiento de la Editorial Andes Cognitio.

Derechos de Autor ©

Este documento se publica bajo los términos y condiciones de la Licencia Creative Commons Reconocimiento – No Comercial – Compartirlgual 4.0 Internacional (CC BY-NC-SA 4.0).



Todos los derechos de autor y de propiedad intelectual e industrial relativos al contenido de esta publicación pertenecen exclusivamente a la "Editorial Andes Cognitio" y a sus respectivos autores. Queda expresamente prohibida, bajo las sanciones establecidas por la legislación vigente, la reproducción total o parcial de esta obra, su almacenamiento en sistemas informáticos, su tratamiento digital, así como cualquier forma de distribución, transmisión o comunicación pública por medios electrónicos, mecánicos, ópticos, químicos, de grabación o fotocopia sin la debida autorización previa y por escrito de los titulares del copyright.

Se exceptúan únicamente los usos con fines académicos o de investigación científica, siempre que no persigan propósitos comerciales y se realicen de forma gratuita, debiendo citarse en todo momento a la fuente editorial correspondiente. Las opiniones vertidas en los distintos capítulos son de exclusiva responsabilidad de los autores y no reflejan necesariamente la postura institucional de la editorial.

Comité Científico Académico

Dr. Jorge Gualberto Paredes Gavilanez PhD.
Universidad Técnica Estatal de Quevedo

Dr. Oscar Patricio López Solís PhD.
Universidad Técnica de Ambato

Ec. Carlos Roberto López Paredes PhD.
Escuela Superior Politécnica de Chimborazo Extensión Orellana

Dr. Héctor Enrique Hernández Altamirano PhD.
Universidad Técnica de Ambato

Dr. Carlos Arturo Jara Santillán PhD.
Escuela Superior Politécnica de Chimborazo

Dr. Guillermo Carrillo Espinosa PhD,
Universidad Autónoma de Chapingo - México

Dra. Doris Coromoto Pernía Barragán PhD,
Universidad de los Andes Tachira Venezuela

Ec. María Gabriela González Bautista PhD.
Universidad Nacional de Chimborazo

My. Efraín Arguello Arellano, Mgs.
Tecnológico Universitario ARGOS – Policía Nacional del Ecuador

Ing. Liliana Priscila Campos Llerena Mgs.
Universidad Técnica de Ambato

Dr. Mario Humberto Paguay Cuvi Mgs.
Escuela Superior Politécnica de Chimborazo

Ec. Oswaldo Javier Jácome Izurieta Mgs.
Universidad Técnica de Ambato

Ec. Juan Carlos Pèrez Briceño Mgs.
Instituto Superior Universitario Bolivariano

Dr. Luís Fernando Paz Villaeeoel Mgs.
Universidad Técnica Particular de Loja

Dra. Daysi Graciela Astudillo Condo Mgs.
Universidad Nacional de Chimborazo

Ec. Ligia Ximena Tapia Hermida Mgs.
Universidad Nacional de Chimborazo

Eco. Ángel Geovanny Carrión Gavilanes MBA.
Pontificia Universidad Católica del Ecuador sede Ambato

Ing. Paula Alejandra Abdo Peralta Mgs.
Escuela Superior Politécnica de Chimborazo

Ing. Milton Gabriel Del Hierro Mosquera Mgs.
Universidad Politécnica Estatal del Carchi

Ing. Catherine Gabriela Frey Erazo Mgs.
Escuela Superior Politécnica de Chimborazo

Ing. Juan Enrique Ureña Moreno Mgs.
Escuela Superior Politécnica de Chimborazo

Ing. José Fernando Esparza Parra Mgs.
Escuela Superior Politécnica de Chimborazo

Ing. Alexis Gabriel Reinoso Haro Mgs.
Universidad Estatal de Bolívar

Eco. Jaime Fernando Andrade Guamán Mgs.
Universidad Amawtay Wasi

Eco. Alejandro Carlos Lalama Freire Mgs.
Universidad Nacional de Chimborazo

Eco. Doménica Estefania Zabala Tapia
Universidad Nacional de Chimborazo

Constancia de Arbitraje

La Editorial Andes Cognitio, hace constar que este libro proviene de una investigación realizada por los autores, siendo sometido a un arbitraje bajo el sistema de doble ciego, de contenido y forma por jurados especialistas. Además, se realizó una revisión del enfoque, paradigma y método investigativo; desde la matriz epistémica asumida por los autores, aplicándose las normas APA, Séptima Edición, proceso de anti plagio en línea Compilatio, garantizándose así la cientificidad de la obra.

Comité Editorial

Eco. Juan Federico Villacis Uvidia Mgs.
Director de la Editorial Andes Cognitio

Lcda. Andrea Damaris Hernández Allauca PhD.
Editora de Andes Cognitio

PRÓLOGO

Emprender es, en esencia, tomar decisiones bajo incertidumbre por lo que quien abre un negocio debe estimar cuántos clientes llegarán, cuánto estarán dispuestos a pagar, cuánto costará operar y cuánto tiempo podrá sostenerse antes de alcanzar el equilibrio. Durante generaciones, esas estimaciones se hicieron con intuición, experiencia y unos pocos datos al alcance de la mano. En el Ecuador y en buena parte de América Latina, donde emprender es para muchas familias la principal fuente de ingresos, el resultado es conocido: abundan las iniciativas, pero solo una parte logra consolidarse y crecer. Mejorar la calidad de esas decisiones es, por tanto, una tarea que interesa tanto a quienes emprenden como a la economía y a la sociedad en su conjunto.

Este libro aparece en un momento en que esa tarea se ha vuelto, por primera vez, alcanzable a gran escala, hoy en día la inteligencia artificial ha reducido de manera notable el costo de predecir y herramientas que hace pocos años estaban reservadas a las grandes corporaciones y el pronóstico de la demanda, la segmentación de clientes, la evaluación del riesgo financiero o la simulación de escenarios se encuentran al alcance de los pequeños negocios. Sin embargo, la obra no cae en el entusiasmo fácil más bien advierte que un modelo puede acertar con los datos del pasado y equivocarse con los casos nuevos, que una recomendación que nadie sabe explicar difícilmente merece confianza y que, si las herramientas quedan en manos de unos pocos, la brecha entre los emprendimientos que sobreviven y los que cierran puede ampliarse en lugar de reducirse.

El mérito principal del texto consiste en ofrecer una visión integral del emprendimiento predictivo, que va desde sus fundamentos económicos hasta su aplicación práctica. A lo largo de sus capítulos se recorren los modelos econométricos clásicos y los algoritmos de aprendizaje automático, la inteligencia artificial explicable, el análisis de mercados y consumidores, las finanzas predictivas, los gemelos digitales y la simulación del riesgo, para culminar en la propuesta de un modelo integral con su propio sistema de puntuación, sus mecanismos de gobernanza y su estrategia de validación. Todo ello se presenta con un lenguaje accesible, sin renunciar al rigor, y con una atención constante a la realidad ecuatoriana: la informalidad, el emprendimiento por necesidad, la escasez de datos sistemáticos y los estudios que ya se han realizado en el país sobre la supervivencia de las empresas.

Por su amplitud, la obra está llamada a ser útil para lectores diversos y las personas emprendedoras encontrarán en ella orientaciones concretas para anticipar sus ventas, conocer mejor a sus

clientes y detectar a tiempo las señales de riesgo financiero; las entidades financieras, incubadoras y aceleradoras, criterios para evaluar con mayor precisión y equidad a los nuevos negocios; docentes y estudiantes, un material que integra la economía, la administración y la ciencia de datos; y quienes diseñan políticas públicas, elementos para construir programas de fomento sustentados en evidencia.

Recomiendo esta obra como una contribución valiosa al estudio y a la práctica del emprendimiento en la región ya que su enseñanza más importante quizá sea que la inteligencia artificial no sustituye a quien emprende, sino que amplía su capacidad de ver hacia adelante. El emprendimiento del futuro no será el que delegue sus decisiones en un algoritmo, sino el que sepa combinar la anticipación que ofrecen los datos con el juicio, la creatividad y la responsabilidad que siempre han definido a la actividad emprendedora.

ÍNDICE GENERAL

PRÓLOGO **2**

ÍNDICE GENERAL **4**

INTRODUCCIÓN **8**

CAPÍTULO I **12**

FUNDAMENTOS DEL EMPRENDIMIENTO PREDICTIVO Y LA NUEVA ARQUITECTURA DE LA INTELIGENCIA EMPRESARIAL **12**

1.1. Del emprendimiento tradicional al emprendimiento predictivo: evolución de los modelos de creación empresarial 12

1.2. La inteligencia artificial como nuevo factor estratégico en la economía emprendedora 16

1.3. Fundamentos económicos y administrativos de la predicción del desempeño empresarial 20

1.4. Ecosistemas emprendedores inteligentes: integración de datos, tecnología y decisiones estratégicas 24

1.5. Economía de los datos y construcción de ventajas competitivas en emprendimientos emergentes 29

1.6. Del análisis retrospectivo a la anticipación empresarial: modelos descriptivos, predictivos y prescriptivos 32

1.7. Arquitectura del emprendimiento predictivo 5.0: interacción entre inteligencia humana, algoritmos y sostenibilidad 36

1.8. Nuevos indicadores de viabilidad, resiliencia y potencial de crecimiento empresarial 40

CAPÍTULO II **45**

INTELIGENCIA ARTIFICIAL Y MODELOS ALGORÍTMICOS PARA PREDECIR EL ÉXITO O FRACASO DE LOS EMPRENDIMIENTOS **45**

2.1. Fundamentos del aprendizaje automático aplicados al análisis del éxito empresarial 45

2.2. Modelos econométricos Logit y Probit para estimar probabilidades de supervivencia y fracaso empresarial 49

2.3. Algoritmos Random Forest y árboles de decisión para identificar determinantes del desempeño emprendedor 54

2.4. Redes neuronales artificiales y Deep Learning para detectar patrones complejos de comportamiento empresarial 59

2.5. Modelos de Gradient Boosting y XGBoost para la clasificación predictiva de emprendimientos 63

2.6. Inteligencia artificial explicable (XAI): interpretación de resultados mediante SHAP y LIME 67

2.7. Validación de modelos predictivos: precisión, sensibilidad, especificidad, curva ROC y matriz de confusión 71

2.8. Construcción de un índice predictivo multidimensional de éxito empresarial mediante técnicas de inteligencia artificial 75

CAPÍTULO III 81

ECONOMÍA PREDICTIVA DEL EMPRENDIMIENTO: ANTICIPACIÓN DE MERCADOS, CLIENTES Y OPORTUNIDADES DE NEGOCIO 81

3.1. Inteligencia económica predictiva y detección anticipada de oportunidades emprendedoras 81

3.2. Modelos de pronóstico de demanda mediante series temporales, ARIMA y aprendizaje automático 85

3.3. Segmentación inteligente de consumidores mediante algoritmos de clustering y aprendizaje no supervisado 89

3.4. Procesamiento de lenguaje natural (NLP) para identificar necesidades insatisfechas y nuevas tendencias de consumo 93

3.5. Predicción del comportamiento del consumidor y estimación del valor del cliente a largo plazo 97

3.6. Análisis de sentimientos y minería de datos para anticipar la aceptación de productos innovadores 101

3.7. Sistemas de recomendación basados en inteligencia artificial para la personalización de modelos de negocio 105

3.8. Construcción de observatorios digitales de oportunidades empresariales mediante datos económicos y comerciales 110

CAPÍTULO IV 115

FINANZAS INTELIGENTES, GEMELOS DIGITALES Y SIMULACIÓN PREDICTIVA DEL RIESGO EMPRENDEDOR 115

4.1. Finanzas predictivas y transformación de los métodos tradicionales de evaluación de emprendimientos 115

4.2. Modelos de inteligencia artificial para pronosticar ingresos, costos, rentabilidad y flujos de efectivo 119

4.3. Predicción de insolvencia empresarial mediante modelos estadísticos y algoritmos de clasificación 122

4.4. Simulación Monte Carlo y análisis probabilístico de escenarios financieros emprendedores 127

4.5. Gemelos digitales empresariales: representación virtual del emprendimiento para experimentar decisiones estratégicas 131

4.6. Modelos predictivos de liquidez, endeudamiento y necesidades futuras de financiamiento 135

4.7. Optimización de inversiones emprendedoras mediante algoritmos evolutivos y aprendizaje por refuerzo 140

4.8. Diseño de sistemas de alerta temprana para detectar vulnerabilidades financieras y riesgos de cierre empresarial 144

CAPÍTULO V 150

EMPRENDIMIENTO AUTÓNOMO Y GOBERNANZA ALGORÍTMICA: HACIA UN MODELO INTEGRAL DE ÉXITO EMPRESARIAL PREDICTIVO 150

5.1. Arquitectura de un modelo integral de emprendimiento predictivo basado en inteligencia artificial 150

5.2. Sistemas inteligentes de apoyo a decisiones para emprendedores y gestores empresariales 154

5.3. Agentes de inteligencia artificial para la automatización supervisada de procesos estratégicos empresariales 158

5.4. Diseño de un tablero de control predictivo con indicadores de supervivencia, crecimiento y rentabilidad 162

5.5. Gobernanza algorítmica, protección de datos, transparencia y ética en las decisiones empresariales automatizadas 166

5.6. Evaluación de sesgos algorítmicos y mecanismos de supervisión humana en modelos de predicción empresarial 170

5.7. Desarrollo de un prototipo de sistema de puntuación predictiva del emprendimiento (Predictive Entrepreneurship Score) 175

5.8. Validación empírica del modelo integral mediante estudios de caso, datos empresariales y simulaciones de escenarios 179

REFERENCIAS BIBLIOGRÁFICAS 184

INTRODUCCIÓN

El emprendimiento ocupa un lugar central en las economías latinoamericanas, pero su vitalidad convive con una notable fragilidad. En el Ecuador, la actividad emprendedora temprana alcanzó en 2023 el 32,7 % de la población adulta, la tasa más alta entre las 46 economías comparadas por el Global Entrepreneurship Monitor; no obstante, el 90,8 % de quienes emprenden lo hace para ganarse la vida ante la escasez de empleo, la mayoría de los nuevos negocios se concentra en actividades orientadas al consumidor y la falta de rentabilidad y las dificultades de financiamiento encabezan las causas de cierre (Lasio et al., 2024). A esta situación se suma un cambio tecnológico de gran alcance, ya que la inteligencia artificial ha abaratado drásticamente la predicción y, con ello, ha modificado la manera en que se identifican oportunidades, se toman decisiones y se gestionan los riesgos (Agrawal et al., 2022). Sin embargo, el acceso a estas herramientas no es homogéneo, y la evidencia internacional advierte sobre una brecha de preparación entre quienes las aprovechan y quienes no lo hacen, que puede traducirse en diferencias crecientes en la probabilidad de supervivencia de los negocios (GEM, 2026).

La literatura sobre emprendimiento permite situar este momento en una trayectoria más amplia. Durante décadas, la creación de empresas se apoyó en la planificación rígida; posteriormente, el enfoque de la empresa emergente ajustada, o *lean startup*, propuso sustituir las proyecciones de largo plazo por la experimentación y el aprendizaje validado (Shepherd y Gruber, 2021). En la actualidad, la inteligencia artificial abre una tercera etapa, puesto que puede transformar la forma en que las nuevas empresas generan ideas, reconocen oportunidades, toman decisiones y escalan sus operaciones (Chalmers et al., 2021; Giuggioli y Pellegrini, 2023). Al mismo tiempo, la perspectiva de la Industria 5.0 recuerda que la transformación tecnológica debe orientarse hacia la centralidad del ser humano, la sostenibilidad y la resiliencia (Breque et al., 2021), y Raisch y Krakowski (2021) advierten que la automatización y el aumento de las capacidades humanas deben equilibrarse para evitar efectos negativos en las organizaciones. De esta convergencia surge la pregunta que guía la obra: ¿de qué manera pueden las personas emprendedoras, y las instituciones que las apoyan, utilizar la inteligencia artificial y los datos para anticipar el éxito o el fracaso de los negocios, sin renunciar a la transparencia, la equidad y la supervisión humana de las decisiones?

Para responder a esa pregunta, el libro desarrolla el concepto de emprendimiento predictivo, entendido como un enfoque de creación y gestión empresarial que integra la inteligencia artificial, el aprendizaje automático, la analítica de datos y la simulación en las decisiones de la persona emprendedora, con el propósito de anticipar riesgos y oportunidades y de favorecer la

supervivencia y el crecimiento sostenible de los negocios. En consecuencia, el objetivo general de la obra consiste en proponer un marco conceptual, metodológico y práctico para incorporar la inteligencia predictiva en los emprendimientos latinoamericanos, con especial atención al caso ecuatoriano. De este objetivo se derivan propósitos específicos, entre ellos explicar la evolución de los modelos de creación empresarial y los fundamentos económicos de la predicción; comparar los modelos econométricos y algorítmicos que permiten estimar el éxito o el fracaso de los emprendimientos; examinar las herramientas para anticipar mercados, clientes y oportunidades; analizar las técnicas de finanzas predictivas, gemelos digitales y simulación del riesgo; y diseñar un modelo integral de emprendimiento predictivo, acompañado de sus mecanismos de gobernanza y de validación.

Desde el punto de vista metodológico, la obra se sustenta en una revisión documental y analítica de literatura académica arbitrada, informes de organismos internacionales y estudios empíricos sobre la supervivencia empresarial en el Ecuador, publicados en su mayoría entre 2021 y 2026. Esta revisión se complementa con el análisis del marco normativo que regula el emprendimiento, la protección de datos y el uso de la inteligencia artificial, en el que destacan la Ley Orgánica de Emprendimiento e Innovación (Asamblea Nacional del Ecuador, 2020), la Ley Orgánica de Protección de Datos Personales (Asamblea Nacional del Ecuador, 2021) y la norma de la Superintendencia de Protección de Datos Personales sobre el uso de sistemas de inteligencia artificial (Superintendencia de Protección de Datos Personales [SPDP], 2026), junto con referentes internacionales como el reglamento europeo de inteligencia artificial (Parlamento Europeo y Consejo de la Unión Europea, 2024), la recomendación de la UNESCO sobre la ética de la inteligencia artificial (UNESCO, 2021) y el marco de gestión de riesgos del NIST (NIST, 2023). Las ideas de los autores consultados se presentan de forma parafraseada y con su correspondiente atribución, y las referencias de toda la obra se reúnen al final en una sola lista, ordenada alfabéticamente.

En cuanto a su alcance, el libro no ofrece recetas tecnológicas ni evalúa productos comerciales, sino que proporciona criterios para decidir qué modelos utilizar, con qué datos, bajo qué condiciones de validación y con qué salvaguardas. Con ese propósito, la obra se organiza en cinco capítulos que avanzan de lo conceptual a lo aplicado. El primero presenta los fundamentos del emprendimiento predictivo, desde la evolución de los modelos de creación empresarial hasta los nuevos indicadores de viabilidad y resiliencia. El segundo aborda el núcleo metodológico, con los modelos Logit y Probit, los árboles de decisión y bosques aleatorios, las

redes neuronales, la potenciación de gradiente, la inteligencia artificial explicable y la validación de los modelos. El tercero se orienta al mercado, con el pronóstico de la demanda, la segmentación de consumidores, el procesamiento del lenguaje natural, los sistemas de recomendación y los observatorios de oportunidades. El cuarto se ocupa de las finanzas inteligentes, la predicción de la insolvencia, la simulación Monte Carlo, los gemelos digitales y los sistemas de alerta temprana. Finalmente, el quinto propone un modelo integral de emprendimiento predictivo, con sus sistemas de apoyo a las decisiones, su tablero de control, su gobernanza algorítmica, un prototipo de puntuación predictiva y una estrategia de validación empírica. Cada subtema se cierra con una figura de elaboración propia que sintetiza sus ideas principales.

Por último, conviene subrayar que la adopción del emprendimiento predictivo no sigue un camino único. Cada negocio parte de condiciones diferentes en cuanto a datos, competencias y recursos, y avanzará a su propio ritmo, combinando la experiencia y la intuición con herramientas analíticas progresivamente más complejas. Esa diversidad no constituye un obstáculo, sino el punto de partida realista de cualquier estrategia, ya que el valor del emprendimiento predictivo no reside en reemplazar el juicio de quien emprende por un algoritmo, sino en ofrecerle mejores elementos para anticipar, decidir y aprender en un entorno de incertidumbre.

EDITORIAL ANDES COGNITIO

CAPÍTULO I

FUNDAMENTOS DEL EMPRENDIMIENTO PRDICTIVO Y LA NUEVA ARQUITECTURA DE LA INTELIGENCIA EMPRESARIAL



CAPÍTULO I

FUNDAMENTOS DEL EMPRENDIMIENTO PREDICTIVO Y LA NUEVA ARQUITECTURA DE LA INTELIGENCIA EMPRESARIAL

1.1. Del emprendimiento tradicional al emprendimiento predictivo: evolución de los modelos de creación empresarial

Durante buena parte del siglo pasado, crear una empresa se entendió como un proceso lineal y planificado. La persona emprendedora identificaba una idea, elaboraba un plan de negocios con proyecciones financieras a varios años, buscaba el capital necesario y ejecutaba lo previsto con la expectativa de que el mercado respondiera según lo estimado. Ese modelo aportó disciplina y un lenguaje común con bancos e inversionistas, pero descansaba en un supuesto frágil: que el futuro podía anticiparse con razonable precisión a partir de la intuición del fundador y de unos pocos datos de mercado. La realidad mostró con insistencia que la mayoría de las proyecciones no se cumplían y que muchas iniciativas fracasaban no por falta de esfuerzo, sino porque el producto, el precio o el cliente imaginados no coincidían con lo que el mercado estaba dispuesto a valorar. Asimismo, la investigación académica sobre la creación de nuevas empresas ha documentado la complejidad de ese proceso.

En una revisión de amplio alcance, Shepherd, Souitaris y Gruber (2021) identificaron diez subtemas que la literatura ha estudiado de manera recurrente —el fundador principal, el equipo fundador, las relaciones sociales, las cogniciones, la organización emergente, la estrategia de la nueva empresa, la emergencia organizacional, la legitimidad, la salida del fundador y el entorno emprendedor— y los agruparon en tres etapas: la cocreación de la oportunidad, la organización del emprendimiento y su desempeño. Esta mirada procesual desplaza la atención desde el plan como documento hacia la empresa como una realidad que se construye, se ajusta y se legitima con el tiempo, en interacción permanente con clientes, socios e instituciones.

Como respuesta práctica a las limitaciones de la planificación rígida surgió el enfoque de la empresa emergente ajustada, conocido internacionalmente como lean startup. Al respecto, Shepherd y Gruber (2021) examinaron sus cinco componentes centrales —el diseño del modelo de negocio, el aprendizaje validado y el desarrollo de clientes, el producto mínimo viable, la decisión entre perseverar o pivotar, y la navegación de la oportunidad de mercado— y advirtieron que, pese a su enorme difusión entre emprendedores, incubadoras y programas educativos, había recibido una atención académica limitada. Su análisis vincula cada componente con

hallazgos previos de la investigación y propone tender puentes entre la práctica y la teoría. El aporte decisivo de este enfoque fue convertir la creación de empresas en una secuencia de experimentos: en lugar de apostar todo a una predicción inicial, el emprendedor formula hipótesis, las somete a prueba con clientes reales y ajusta su propuesta según la evidencia que obtiene.

Además, la transformación digital añadió una nueva capa a esta evolución. La disponibilidad de plataformas en línea, servicios en la nube, herramientas de comercio electrónico y redes sociales redujo los costos de entrada, permitió llegar a mercados lejanos con inversiones modestas y, sobre todo, generó un flujo continuo de datos sobre el comportamiento de clientes, competidores y proveedores.

La evidencia empírica sugiere que esta transformación también modifica las oportunidades disponibles. En un análisis de datos de panel para los países de la zona euro entre 2009 y 2020, del Olmo-García et al. (2023) encontraron que una mayor robotización de la actividad industrial se asociaba con más oportunidades para el emprendimiento, mientras que la inversión privada en investigación y desarrollo mostraba una relación negativa con la actividad emprendedora, lo que llevó a los autores a recomendar políticas de formación, entornos favorables a la innovación digital y a la inteligencia artificial, y estímulos al intraemprendimiento.

La inteligencia artificial marca, sin embargo, un cambio de naturaleza distinta. En esa dirección, Chalmers, MacKenzie y Carter (2021) sostienen que estas tecnologías pueden transformar la manera en que las nuevas empresas generan ideas, venden y escalan, ya sea asumiendo determinadas tareas o asistiendo a las personas en su ejecución, y que ese cambio exige nuevas formas de organizar el trabajo emprendedor. Al mismo tiempo, advierten que el apalancamiento tecnológico trae consigo pasivos, entre ellos la posibilidad de que pequeñas empresas tradicionales queden desplazadas por intermediarios que dominan los datos y los algoritmos.

La inteligencia artificial, en este sentido, no es solo una herramienta adicional en el repertorio del emprendedor; altera las reglas de la competencia y redistribuye las ventajas entre quienes disponen de información y quienes carecen de ella. Con todo, sobre esta trayectoria puede definirse lo que en esta obra se denomina emprendimiento predictivo: un enfoque de creación y gestión de empresas que utiliza datos, modelos estadísticos y algoritmos de inteligencia artificial para estimar de manera sistemática la probabilidad de supervivencia, crecimiento y rentabilidad de una iniciativa, anticipar la evolución de la demanda y del riesgo, y orientar las decisiones antes de comprometer recursos de forma irreversible. El emprendimiento predictivo no

sustituye la experimentación del enfoque ajustado ni la creatividad del fundador; las complementa con una capacidad de anticipación basada en evidencia. Si el plan de negocios tradicional proyectaba un único futuro y el enfoque ajustado aprendía a partir de experimentos sucesivos, el enfoque predictivo trabaja con distribuciones de probabilidad, escenarios alternativos y señales tempranas que permiten corregir el rumbo con mayor oportunidad.

La lógica económica de este enfoque ha sido formulada con claridad por Agrawal, Gans y Goldfarb (2022), quienes caracterizan la inteligencia artificial como una tecnología que abarata drásticamente la predicción. Cuando predecir se vuelve barato, las organizaciones pueden incorporar estimaciones en decisiones que antes se tomaban por intuición, por regla fija o por costumbre; pero el valor de esa predicción depende de cómo se integra en el sistema de decisiones y del juicio humano que determina qué hacer con ella. Para un emprendimiento, esto significa que la ventaja no proviene únicamente de disponer de un modelo que estime la demanda o el riesgo de impago, sino de rediseñar los procesos de compra, producción, venta y financiamiento para aprovechar esas estimaciones.

En esta misma línea, el contexto latinoamericano y, en particular, el ecuatoriano, otorga a esta discusión una relevancia especial. Según el informe del Global Entrepreneurship Monitor para el Ecuador, en 2023 la tasa de actividad emprendedora temprana alcanzó el 32,7 % de la población adulta, la más alta entre las 46 economías comparadas, y la proporción de propietarios de negocios establecidos llegó al 24 % (Lasio et al., 2024).

Sin embargo, el mismo informe muestra los límites de ese dinamismo: el 90,8 % de quienes emprenden declara hacerlo para ganarse la vida ante la escasez de empleo, alrededor de tres cuartas partes de los nuevos negocios se concentra en actividades orientadas al consumidor, principalmente el comercio minorista, y la tasa de abandono entre los emprendedores en etapa temprana alcanzó el 17,33 %. Entre las causas de cierre, la falta de rentabilidad encabezó la lista con el 37,4 %, seguida por las dificultades de financiamiento, con el 21,4 % (Lasio et al., 2024).

Estas cifras describen una paradoja que atraviesa a buena parte de la región: abundan las iniciativas, pero escasean las que logran consolidarse y crecer. En un contexto así, el emprendimiento predictivo no es un lujo reservado a las empresas tecnológicas de alto crecimiento, sino una herramienta potencialmente útil para reducir la mortalidad de los pequeños negocios. Estimar con anticipación si un local comercial alcanzará el punto de equilibrio, si la demanda de un

barrio justifica una inversión o si la estructura de deuda resulta sostenible puede marcar la diferencia entre un cierre prematuro y una empresa viable.

La dificultad está en que muchos de estos emprendimientos carecen de datos organizados, de competencias analíticas y de acceso a herramientas especializadas, brechas que los capítulos siguientes examinan con detalle. Conviene precisar que el tránsito hacia el emprendimiento predictivo no es lineal ni homogéneo. En una misma ciudad latinoamericana conviven negocios que se gestionan con libretas de apuntes y empresas emergentes que ajustan su oferta a partir de métricas diarias.

Esta coexistencia obliga a pensar el cambio como un gradiente y no como un salto: para una tienda de barrio, registrar sistemáticamente sus ventas ya constituye un primer paso hacia la anticipación, mientras que para una empresa de base tecnológica el desafío puede consistir en integrar modelos de aprendizaje automático en su proceso de desarrollo de productos. Las metodologías de experimentación que se difundieron con la lógica de la empresa emergente ágil siguen siendo valiosas, porque la predicción no elimina la necesidad de probar hipótesis en el mercado; lo que cambia es que las pruebas pueden diseñarse mejor, priorizarse según su probabilidad de éxito y evaluarse con mayor rapidez. El emprendimiento predictivo, entendido de este modo, no reemplaza a la experimentación ni a la intuición, sino que les proporciona un marco de información más rico en el que desplegarse.

La evolución descrita puede leerse, por tanto, como una sucesión de respuestas a la incertidumbre. El modelo tradicional intentó dominarla mediante la planificación; el enfoque ajustado aprendió a convivir con ella mediante la experimentación; y el enfoque predictivo busca reducirla mediante la estimación sistemática de probabilidades. Estas etapas, sin embargo, no eliminan por completo a las anteriores. Un buen plan de negocios sigue siendo necesario para dialogar con financiadores, y la experimentación con clientes continúa siendo insustituible cuando no existen datos históricos.

Figura 1.1

Del emprendimiento tradicional al emprendimiento predictivo: evolución de los modelos de creación empresarial.



Fuente: Elaboración autores

1.2. La inteligencia artificial como nuevo factor estratégico en la economía emprendedora

La teoría económica clásica identificó la tierra, el trabajo y el capital como los factores de la producción, y más tarde incorporó el conocimiento y la tecnología como determinantes del crecimiento. En la economía contemporánea, la inteligencia artificial se perfila como un factor estratégico con características propias: no se agota con el uso, mejora a medida que se alimenta de más datos, puede replicarse a bajo costo y, combinada con la información adecuada, transforma la manera en que se toman las decisiones.

Para el emprendimiento, esta tecnología no representa solo una oportunidad de negocio —crear empresas que desarrollan o aplican inteligencia artificial—, sino también un insumo transversal que puede incorporarse en cualquier actividad, desde una panadería que ajusta su producción diaria hasta una plataforma de servicios financieros que evalúa el riesgo de sus clientes. De igual manera, la literatura académica ha comenzado a ordenar las formas en que la inteligencia artificial habilita la actividad emprendedora. Sobre este punto, Giuggioli y Pellegrini (2023), en una revisión sistemática de sesenta artículos, organizaron la evidencia en cuatro grupos.

El primero corresponde a la oportunidad, es decir, al uso de la inteligencia artificial y de los grandes volúmenes de datos para crear nuevos productos, servicios y modelos de negocio, incluidos aquellos que surgen de la combinación con sensores e internet de las cosas. El segundo, el más numeroso, se refiere a la toma de decisiones, donde estas tecnologías mejoran las

predicciones en ámbitos como el pronóstico de ventas, la segmentación de clientes, la selección de mercados de exportación o la anticipación del resultado de campañas de financiamiento colectivo. El tercero agrupa los efectos sobre el desempeño, como la productividad, la automatización y la mejora de los resultados comerciales y financieros. Y el cuarto reúne los aportes a la educación y la investigación sobre emprendimiento, que, según los autores, aceleran el desarrollo de los demás.

Esta clasificación permite advertir que la inteligencia artificial interviene en todas las fases del proceso emprendedor. Antes de crear la empresa, ayuda a identificar necesidades insatisfechas y a evaluar la viabilidad de una idea. Durante su puesta en marcha, apoya decisiones sobre precios, inventarios, canales de venta y contratación. En la etapa de crecimiento, facilita la personalización de la oferta, la gestión del riesgo y la búsqueda de financiamiento. Y, a lo largo de todo el ciclo, puede contribuir a que la persona emprendedora aprenda más rápido, siempre que sepa interpretar críticamente lo que los sistemas le sugieren. Por otra parte, la adopción efectiva, no obstante, avanza con lentitud y de manera desigual.

Un documento de discusión preparado por la OCDE para el Grupo de los Siete mostró que, en 2024, el 40 % de las grandes empresas de los países miembros utilizaba inteligencia artificial, frente al 20,4 % de las medianas y apenas el 11,9 % de las pequeñas; aunque el uso promedio había pasado de 5,6 % en 2020 a 14 % en 2024, las pequeñas empresas tenían menos de un tercio de la probabilidad de usar estas tecnologías que las grandes (Organización para la Cooperación y el Desarrollo Económicos [OCDE], 2025). El mismo estudio señala que la inteligencia artificial generativa es la modalidad más utilizada por las pequeñas y medianas empresas, por sus bajos costos de acceso, pero que la mayoría la emplea en tareas periféricas: solo el 29 % de las que la usan la aplican a sus actividades centrales.

Las expectativas de los propios emprendedores reflejan esa distancia. El informe mundial 2024/2025 del Global Entrepreneurship Monitor, basado en encuestas a más de ciento cincuenta mil personas, encontró que en 36 de las 49 economías analizadas menos del 30 % de los emprendedores en etapa temprana consideraba que la inteligencia artificial sería muy importante para su negocio en los tres años siguientes, y que muchos se mostraban inseguros o desconocían su posible papel (Global Entrepreneurship Monitor [GEM], 2025). El informe 2025/2026, que abarcó 53 economías, identificó una brecha de preparación entre quienes acceden a la inteligencia artificial y quienes no lo hacen, junto con una brecha de supervivencia que impide a muchas iniciativas convertirse en empresas establecidas (GEM, 2026). Las dos brechas no son

independientes: si las herramientas que mejoran las decisiones quedan concentradas en un grupo reducido de empresas, la distancia en la probabilidad de sobrevivir puede ampliarse.

Paralelamente, en el Ecuador, los datos disponibles muestran un uso creciente de la tecnología, aunque centrado en funciones básicas. Según el informe nacional del Global Entrepreneurship Monitor, el 56,5 % de los emprendedores en etapa temprana declaró utilizar más tecnología para vender sus productos o servicios (Lasio et al., 2024). Este dato es alentador, pero no equivale a una adopción estratégica de la inteligencia artificial. Vender por redes sociales o mediante aplicaciones de mensajería es un paso importante; utilizar los datos que esas interacciones generan para pronosticar la demanda, segmentar clientes o evaluar la rentabilidad de cada canal es un paso cualitativamente distinto que todavía pocos negocios han dado.

Las barreras que explican esta situación han sido identificadas con bastante precisión. El estudio de la OCDE (2025) destaca la falta de competencias, ya que cerca de la mitad de las pequeñas y medianas empresas de cuatro países del Grupo de los Siete señaló que su personal carecía de las habilidades necesarias para usar la inteligencia artificial generativa; las dificultades de financiamiento, que frenan las inversiones de largo plazo; las diferencias de conectividad, que alcanzan unos veinticinco puntos porcentuales entre pequeñas y grandes empresas en el acceso a banda ancha de alta velocidad; la escasez de datos de calidad y de capacidad de cómputo; y la preocupación por resultados inexactos, contenidos perjudiciales o problemas legales.

Estas barreras son, previsiblemente, más intensas en economías con alta informalidad y con predominio de microempresas, como las latinoamericanas. Sobre esta base, la incorporación de la inteligencia artificial plantea además una tensión de fondo sobre el papel de las personas en la empresa. En este sentido, Raisch y Krakowski (2021) describen esa tensión como una paradoja entre automatización y aumento: la primera supone que las máquinas asuman tareas humanas, mientras que el segundo implica una colaboración estrecha entre personas y sistemas. Los autores sostienen que ambas lógicas son interdependientes en el tiempo y en el espacio, y que privilegiar exclusivamente una de ellas puede desencadenar ciclos con efectos negativos para las organizaciones y la sociedad.

Para un emprendimiento, la lección es práctica: automatizar tareas repetitivas libera tiempo, pero la ventaja competitiva suele surgir cuando la inteligencia artificial amplía la capacidad de juicio del equipo, y no cuando lo reemplaza sin criterio.

Desde una perspectiva estratégica, la inteligencia artificial puede generar ventajas para los emprendimientos por varias vías. Puede reducir costos al automatizar procesos administrativos, de

atención al cliente o de control de inventarios. Puede aumentar los ingresos al mejorar la personalización de la oferta, la fijación de precios o la identificación de clientes potenciales. Puede reducir el riesgo al anticipar problemas de liquidez, impagos o caídas de demanda. Y puede acelerar el aprendizaje organizacional al convertir cada transacción en información útil para decisiones futuras.

Estas ventajas, con todo, no son automáticas, pues dependen de la calidad de los datos, de la pertinencia de los modelos y de la capacidad del equipo para actuar sobre las recomendaciones. Conviene también reconocer que la inteligencia artificial redistribuye el poder en los mercados. Al respecto, Chalmers, MacKenzie y Carter (2021) llaman la atención sobre el riesgo de que pequeñas empresas tradicionales queden excluidas por intermediarios digitales que concentran los datos y los algoritmos. Para muchos emprendimientos latinoamericanos, que venden a través de plataformas de comercio electrónico, de reparto a domicilio o de publicidad en línea, esa advertencia es concreta: los datos sobre sus clientes, sus ventas y su reputación quedan, en buena medida, en manos de terceros que deciden cómo se muestran sus productos y cuánto cobran por ello. Desarrollar capacidades propias de análisis y conservar el control sobre la información estratégica forma parte, por tanto, de la estrategia de cualquier negocio que aspire a sostenerse en el tiempo.

La consideración de la inteligencia artificial como factor estratégico tiene implicaciones también para la política pública. Si el costo de la predicción disminuye, pero el acceso a los datos, a las competencias y a la infraestructura sigue concentrado, los beneficios de esta tecnología tenderán a favorecer a las empresas que ya disponen de esos recursos. Para los países latinoamericanos, el riesgo es que la brecha entre grandes y pequeñas empresas se amplíe, y que los emprendimientos locales compitan en desventaja con plataformas globales que acumulan datos de millones de usuarios.

Las respuestas posibles incluyen la formación de competencias digitales en las personas emprendedoras, el acceso a datos públicos abiertos y de calidad, la oferta de servicios compartidos de analítica a través de universidades e incubadoras, y marcos regulatorios que equilibren la protección de los derechos con la promoción de la innovación (OCDE, 2025). Sin estas condiciones, la inteligencia artificial podría convertirse en un nuevo factor de desigualdad empresarial antes que en una palanca para el desarrollo de los nuevos negocios.

Desde esta perspectiva, la inteligencia artificial se ha convertido en un factor que condiciona la competitividad de los emprendimientos de manera creciente, pero su aprovechamiento no depende únicamente del acceso a la tecnología. Requiere datos, competencias, financiamiento, conectividad y una comprensión clara de para qué se quiere predecir y cómo se actuará con base en esa predicción.

Figura 1.2

La inteligencia artificial como nuevo factor estratégico en la economía emprendedora.



Fuente: Elaboración autores

1.3. Fundamentos económicos y administrativos de la predicción del desempeño empresarial

Toda decisión empresarial descansa, explícita o implícitamente, en una predicción. Contratar a un nuevo trabajador supone estimar que la demanda crecerá; abrir una sucursal implica prever que el barrio generará ventas suficientes; otorgar crédito a un cliente equivale a anticipar que pagará. En la mayoría de los pequeños negocios, estas predicciones se formulan de manera intuitiva, a partir de la experiencia del propietario y de la información que circula en su entorno inmediato.

La pregunta que motiva el emprendimiento predictivo es si esas estimaciones pueden hacerse de forma más sistemática, transparente y precisa, y qué consecuencias económicas y administrativas tiene hacerlo. Ahora bien, desde la economía, el punto de partida es la distinción entre riesgo e incertidumbre. En situaciones de riesgo, los posibles resultados y sus probabilidades pueden estimarse con base en la experiencia acumulada; en situaciones de incertidumbre, ni siquiera es posible enumerar con confianza todos los escenarios.

El emprendimiento se ubica tradicionalmente en el terreno de la incertidumbre, porque introduce productos, procesos o modelos de negocio sin antecedentes directos. La disponibilidad de datos y de modelos predictivos no elimina esa condición, pero permite trasladar una parte de las decisiones desde la incertidumbre hacia el riesgo cuantificable: lo que antes era una apuesta puede convertirse, al menos parcialmente, en una estimación con márgenes de error conocidos.

Agrawal, Gans y Goldfarb (2022) proponen un marco útil para comprender este desplazamiento. Según estos autores, la inteligencia artificial debe entenderse como una tecnología de predicción cuyo costo desciende de manera sostenida, y toda decisión combina dos elementos: la predicción, que estima lo que probablemente ocurrirá, y el juicio, que determina el valor que se asigna a cada resultado posible y, por tanto, qué conviene hacer.

A medida que la predicción se abarata, el juicio humano gana importancia relativa, y las organizaciones que más se benefician no son las que simplemente incorporan un algoritmo en un proceso existente, sino las que rediseñan sus sistemas de decisión para aprovechar mejores predicciones. Para un emprendimiento, esto implica que la ventaja no reside en el modelo aislado, sino en la manera en que las estimaciones se conectan con decisiones concretas sobre compras, precios, contratación o inversión. De forma complementaria, la economía de los datos complementa este razonamiento.

En esa dirección, Veldkamp y Chung (2024) revisan cómo se ha modelado la información digitalizada en la teoría macroeconómica, comparan sus propiedades con las de las ideas, las patentes o el aprendizaje por la práctica, y destacan tres aplicaciones principales: la innovación, la optimización de procesos y el emparejamiento entre oferta y demanda. Estas tres aplicaciones tienen correlatos directos en la vida de una empresa emergente. Los datos permiten innovar al revelar necesidades no atendidas, optimizar procesos al identificar ineficiencias, y emparejar mejor productos y clientes al conocer sus preferencias. A diferencia de otros activos, además, los datos pueden reutilizarse sin agotarse, lo que favorece a quienes los acumulan y organizan desde etapas tempranas.

Desde la administración, la predicción del desempeño empresarial se apoya en una larga tradición de estudios sobre los determinantes de la supervivencia y el fracaso de las empresas. Esta literatura ha examinado factores relacionados con el emprendedor —experiencia, formación, motivación—, con la empresa —tamaño, edad, estructura financiera, sector— y con el entorno —ciclo económico, competencia, acceso al financiamiento—. En el caso ecuatoriano, varios

estudios recientes han aplicado modelos estadísticos a datos administrativos para identificar esos determinantes. Sobre este punto, Zambrano, Sánchez y Valls Martínez (2021), mediante un modelo probit aplicado a las micro, pequeñas y medianas empresas registradas en la Superintendencia de Compañías, Valores y Seguros, encontraron que un mayor patrimonio, una rotación de cartera adecuada y un endeudamiento equilibrado aumentaban la esperanza de vida de las empresas, y que la región y el sector de actividad también influían en la probabilidad de supervivencia.

Vinculado con lo anterior, otros trabajos han utilizado técnicas de análisis de supervivencia, que permiten estudiar no solo si una empresa cierra, sino cuándo es más probable que lo haga. En este sentido, Chico y Altamirano (2023), al analizar 175 empresas manufactureras del cantón Ambato entre 2018 y 2022 mediante análisis discriminante, curvas de Kaplan-Meier y regresión de Cox, encontraron que la utilidad neta y los activos reducían el riesgo de cierre. En Manabí, González Cedeño (2026) aplicó un modelo de riesgos proporcionales de Cox a 174 empresas registradas en la Superintendencia y encontró que una mayor rentabilidad sobre los activos y un mayor volumen de ventas reducían el riesgo de cierre, mientras que el apalancamiento financiero lo incrementaba en alrededor de 11 % por cada desviación estándar; el estudio mostró, además, que el riesgo se concentraba en los primeros años de vida de las empresas.

Estos hallazgos ilustran un principio fundamental: el desempeño futuro de una empresa deja huellas observables en su información presente. La rentabilidad, el nivel de endeudamiento, la evolución de las ventas o la gestión de la cartera funcionan como señales que, adecuadamente analizadas, permiten estimar la probabilidad de supervivencia. La inteligencia artificial amplía este repertorio en dos direcciones. Por un lado, permite procesar un número mucho mayor de variables y capturar relaciones no lineales entre ellas. Por otro, incorpora fuentes de información no tradicionales, como las interacciones en redes sociales, las reseñas de clientes o los patrones de transacciones digitales.

Al respecto, Bidgoli, Vanani y Goodarzi (2024), por ejemplo, encontraron que variables como el número de seguidores y de empleados registrados en redes profesionales, los seguidores en redes sociales y el monto de la última ronda de financiamiento figuraban entre los predictores más influyentes del éxito de empresas emergentes. En concordancia con ello, la predicción del desempeño tiene también fundamentos administrativos vinculados con la teoría de la decisión y el control de gestión. Las empresas toman decisiones bajo restricciones de tiempo,

información y capacidad de procesamiento; por ello, recurren a reglas simplificadas, indicadores de alerta y rutinas que reducen la complejidad.

Los modelos predictivos pueden entenderse como una evolución de esas herramientas: en lugar de reglas fijas —por ejemplo, no otorgar crédito a clientes con más de treinta días de atraso—, permiten estimar probabilidades individualizadas y ajustar las decisiones a cada caso. En la práctica, esto exige integrar los modelos en los procesos de planificación, presupuesto y control, de modo que las predicciones se traduzcan en acciones y que los resultados observados retroalimenten los modelos.

Existen, sin embargo, límites que conviene reconocer desde el inicio. El primero es que las predicciones se basan en patrones del pasado y pueden fallar ante cambios estructurales, como una crisis sanitaria, una disrupción tecnológica o un cambio regulatorio. El segundo se refiere a la calidad y representatividad de los datos: los modelos construidos con información de empresas formales pueden no aplicarse a negocios informales, que en América Latina representan una proporción considerable del tejido productivo.

El tercero alude al riesgo de profecías autocumplidas, como cuando un modelo que predice el fracaso de un emprendimiento lleva a negarle financiamiento y provoca, precisamente, su cierre. Y el cuarto consiste en la tentación de confundir correlación con causalidad, atribuyendo a una variable un efecto que en realidad corresponde a factores no observados. A ello se suma que estos fundamentos permiten entender por qué la predicción del desempeño empresarial no puede reducirse a un ejercicio técnico. Cada modelo incorpora supuestos sobre qué significa el éxito, qué factores lo determinan y qué horizonte temporal resulta relevante. Una entidad financiera puede definir el éxito como la capacidad de pagar un crédito en un plazo de dos años; un fondo de capital de riesgo, como la posibilidad de multiplicar su inversión en una década; una persona emprendedora, como la posibilidad de sostener a su familia con estabilidad. Estas definiciones conducen a modelos distintos, que privilegian variables diferentes y producen recomendaciones que pueden incluso contradecirse.

La economía y la administración aportan los conceptos necesarios para hacer explícitos estos supuestos: la teoría de la decisión bajo incertidumbre, el análisis de los incentivos de los distintos actores y la comprensión de las capacidades organizacionales que convierten los recursos en resultados. Explicitar estos fundamentos es una condición para que los modelos predictivos

sean utilizados con discernimiento y no como cajas negras cuyas respuestas se aceptan sin examen.

La perspectiva administrativa recuerda, además, que el desempeño empresarial es el resultado de un proceso de gestión y no solo de condiciones iniciales. Dos emprendimientos con recursos semejantes pueden obtener resultados muy distintos según la calidad de sus decisiones cotidianas sobre precios, inventarios, cobranza, contratación o relación con los clientes. Esta constatación tiene una consecuencia importante para la predicción: los modelos que se limitan a las características observadas al inicio del negocio capturan solo una parte de los determinantes del éxito, y su precisión mejora cuando se incorporan indicadores que reflejan la evolución de la gestión a lo largo del tiempo.

De ahí la conveniencia de concebir la predicción como un proceso continuo, que se actualiza a medida que el negocio genera información sobre su funcionamiento, y no como un diagnóstico único realizado antes de su creación. No obstante, estos límites no invalidan el enfoque predictivo, pero obligan a utilizarlo con prudencia. Una predicción bien fundamentada es un insumo valioso para la decisión, no un veredicto. Su utilidad depende de que se conozcan su margen de error, las variables que la sustentan y las condiciones en las que fue construida, y de que el juicio humano intervenga para interpretarla a la luz del contexto.

Figura 1.3

Fundamentos económicos y administrativos de la predicción del desempeño empresarial.



Fuente: Elaboración autores

1.4. Ecosistemas emprendedores inteligentes: integración de datos, tecnología y decisiones estratégicas

Las empresas no nacen en el vacío. Su probabilidad de sobrevivir y crecer depende, en buena medida, del entorno en el que se desenvuelve: la disponibilidad de financiamiento, la calidad

del talento, la cultura empresarial, la infraestructura, la normativa, la presencia de universidades y centros de investigación, el acceso a mercados y la densidad de las redes entre actores. La noción de ecosistema emprendedor intenta capturar esa trama de relaciones. Stam y van de Ven (2021) lo definen como el conjunto de actores y factores interdependientes que habilitan y restringen el emprendimiento en un territorio, y su análisis empírico para las regiones de los Países Bajos mostró que la presencia de empresas de alto crecimiento se relaciona fuertemente con la calidad del ecosistema.

Los autores observaron, además, que los distintos elementos del ecosistema están estrechamente vinculados entre sí, lo que obliga a analizarlos como un sistema y no de manera aislada. Asimismo, el concepto ha ganado popularidad entre gobiernos, organismos de cooperación y universidades, pero su desarrollo teórico sigue siendo incompleto. En esa dirección, Wurth, Stam y Spigel (2022) lo describen como una noción paradójica: tiene raíces intelectuales profundas y la capacidad de articular distintas corrientes de investigación, pero permanece insuficientemente teorizada, y los procesos que explican la evolución de los ecosistemas son todavía poco comprendidos.

Los autores proponen una agenda de investigación transdisciplinaria que combine la evidencia sobre los mecanismos causales con una comprensión más amplia del contexto. Para quienes diseñan políticas o dirigen organizaciones de apoyo al emprendimiento, esta advertencia es importante: no basta con acumular incubadoras, fondos y eventos; hace falta entender cómo interactúan los elementos y qué resultados producen.

La transformación digital y la inteligencia artificial añaden una dimensión nueva a esta discusión. Un ecosistema emprendedor inteligente puede entenderse como aquel en el que los actores no solo se relacionan, sino que generan, comparten y utilizan datos para mejorar sus decisiones. En un ecosistema de este tipo, las entidades financieras utilizan modelos para evaluar con mayor precisión el riesgo de los nuevos negocios y ampliar el crédito a quienes antes quedaban excluidos; las incubadoras y aceleradoras emplean información sobre el desempeño de sus empresas para identificar tempranamente a las que necesitan apoyo; las universidades transfieren capacidades analíticas a las empresas de su entorno; y los gobiernos utilizan estadísticas actualizadas para orientar los programas de fomento hacia los sectores y territorios con mayor potencial.

Además, la integración de datos es, por tanto, el rasgo distintivo de estos ecosistemas. En la mayoría de los territorios latinoamericanos, la información sobre las empresas se encuentra dispersa entre registros mercantiles, administraciones tributarias, entidades de seguridad social, superintendencias, bancos, cámaras de comercio y plataformas privadas, sin mecanismos que permitan combinarla de manera segura y útil. Cada actor dispone de una pieza del rompecabezas, pero nadie observa el conjunto. Avanzar hacia ecosistemas inteligentes implica crear espacios de colaboración en los que esa información pueda compartirse con garantías de protección de datos, estándares comunes y reglas claras sobre su uso, de modo que los emprendedores, los financiadores y las instituciones de apoyo tomen decisiones mejor informadas.

La tecnología es el segundo componente. Las plataformas digitales de comercio, pago, contabilidad y gestión de clientes generan datos de manera continua y ofrecen, cada vez con mayor frecuencia, funciones analíticas integradas. Los servicios en la nube permiten acceder a capacidades de cómputo y a modelos de inteligencia artificial sin grandes inversiones iniciales. Sin embargo, la disponibilidad de estas herramientas no garantiza su aprovechamiento.

El estudio de la OCDE (2025) sobre la adopción de la inteligencia artificial en las pequeñas y medianas empresas recomienda fortalecer la conectividad, facilitar el acceso a datos y capacidad de cómputo, sensibilizar sobre los usos, beneficios y riesgos de estas tecnologías, desarrollar las capacidades de la fuerza laboral, mejorar la preparación para la inversión y adaptar las intervenciones a los distintos perfiles de empresas. Estas recomendaciones apuntan precisamente a los componentes de un ecosistema inteligente. En esta misma línea, el tercer componente son las decisiones estratégicas.

Un ecosistema produce valor cuando la información y la tecnología se traducen en mejores decisiones de los actores que lo integran. Para un emprendedor, esto puede significar elegir con mayor fundamento el mercado al que dirigirse, el momento de solicitar financiamiento o la conveniencia de asociarse con otra empresa. Para un financiador, implica asignar recursos con criterios más precisos. Para un gobierno, supone diseñar políticas basadas en evidencia y evaluar sus resultados. La calidad de las decisiones depende, a su vez, de las competencias de quienes las toman, lo que convierte la formación en análisis de datos y en uso responsable de la inteligencia artificial en un componente central del ecosistema.

El caso ecuatoriano ofrece elementos para valorar el punto de partida. Según el informe nacional del Global Entrepreneurship Monitor, las condiciones del entorno emprendedor del país obtuvieron en 2023 una calificación promedio de 3,9, ligeramente inferior al promedio de

América Latina y el Caribe, de 4; la infraestructura física y de servicios apareció como la condición más favorable, mientras que el apoyo financiero y la transferencia de investigación y desarrollo figuraron entre las más débiles, y la mayoría de las condiciones se había deteriorado respecto de años anteriores (Lasio et al., 2024).

El mismo informe señala que el 67,9 % de los emprendedores en etapa temprana consideraba que iniciar un negocio era más difícil que un año antes. Estas cifras sugieren que el ecosistema ecuatoriano cuenta con bases de infraestructura, pero enfrenta debilidades precisamente en los ámbitos que más dependen de la información y del conocimiento: el financiamiento y la transferencia tecnológica. De igual manera, en el plano normativo, la Ley Orgánica de Emprendimiento e Innovación, publicada en febrero de 2020, estableció un marco para fomentar el emprendimiento, la innovación y el desarrollo tecnológico, promover una cultura emprendedora e introducir nuevas formas societarias y de financiamiento para fortalecer el ecosistema, con un alcance que abarca a las economías pública, privada, mixta, popular y solidaria, cooperativa, asociativa, comunitaria y artesanal (Asamblea Nacional del Ecuador, 2020). Aunque se trata de una norma anterior al período que cubre esta obra, su inclusión resulta indispensable porque constituye el principal referente legal del ecosistema emprendedor ecuatoriano. Su aprovechamiento para construir un ecosistema inteligente dependerá de que sus instrumentos se articulen con la infraestructura de datos y con las capacidades analíticas que el emprendimiento predictivo requiere.

Del análisis anterior se desprenden algunas condiciones para que un ecosistema emprendedor evolucione hacia un modelo inteligente. Una de ellas es la existencia de infraestructura de datos compartida, con registros empresariales actualizados, interoperables y accesibles bajo reglas de protección de la información. Otra es la presencia de intermediarios capaces de transformar esos datos en servicios útiles para los emprendedores, como observatorios, centros de transferencia o plataformas de diagnóstico.

Una tercera es la disponibilidad de financiamiento que utilice información para evaluar el potencial de los negocios y no solo sus garantías. Una cuarta es la formación continua en competencias digitales y analíticas. Y una quinta es una gobernanza que coordine a los actores públicos, privados, académicos y sociales en torno a objetivos comunes, evaluando periódicamente los resultados. Por otra parte, un ecosistema inteligente se reconoce, en la práctica, por la circulación de información entre sus actores.

Cuando una incubadora comparte con una entidad financiera, con el consentimiento de los emprendedores, los indicadores de desempeño de los negocios que acompaña, el financiamiento puede asignarse con mejor criterio. Cuando una universidad pone a disposición de los emprendedores sus capacidades de análisis de datos, los pequeños negocios acceden a herramientas que no podrían costear por sí solos. Cuando un gobierno local publica en formatos abiertos la información sobre permisos de funcionamiento, flujos de visitantes o proyectos de infraestructura, las personas emprendedoras pueden identificar oportunidades con mayor anticipación. Estos intercambios requieren confianza, reglas claras sobre el uso de los datos y plataformas que faciliten la interoperabilidad. La literatura sobre ecosistemas emprendedores subraya que la calidad de las interacciones entre los elementos del sistema importa tanto como la presencia de cada uno de ellos (Stam y van de Ven, 2021), lo que en el contexto digital se traduce en la capacidad de compartir información útil de manera segura.

Un ecosistema emprendedor inteligente no se construye por decreto ni mediante la adquisición de tecnología. Surge de la interacción sostenida entre actores que aprenden a compartir información, a confiar unos en otros y a tomar decisiones basadas en evidencia. La evidencia de Stam y van de Ven (2021) sobre la relación entre la calidad del ecosistema y la presencia de empresas de alto crecimiento sugiere que esta inversión colectiva tiene recompensas tangibles. Para países como el Ecuador, con una base emprendedora amplia pero frágil, fortalecer el ecosistema puede ser tan importante como apoyar a cada emprendimiento por separado.

Figura 1.4

Ecosistemas emprendedores inteligentes: integración de datos, tecnología y decisiones estratégicas.



Fuente: Elaboración autores

1.5. Economía de los datos y construcción de ventajas competitivas en emprendimientos emergentes

En la economía contemporánea, los datos se han convertido en un recurso estratégico comparable, en algunos sectores, al capital físico o financiero. A diferencia de estos, sin embargo, los datos tienen propiedades económicas peculiares: pueden ser utilizados simultáneamente por varias personas sin agotarse, su valor aumenta cuando se combinan con otras fuentes, su costo de reproducción es prácticamente nulo y su utilidad depende en gran medida de la capacidad para analizarlos.

Estas propiedades explican por qué las empresas que lograron acumular y procesar grandes volúmenes de información alcanzaron posiciones dominantes en numerosos mercados, y por qué los emprendimientos emergentes necesitan pensar desde el inicio en qué datos generan, cómo los conservan y de qué manera pueden convertirlos en ventaja competitiva. Paralelamente, la investigación económica ha comenzado a sistematizar estas propiedades. Sobre este punto, Veldkamp y Chung (2024) revisan cómo se ha incorporado la información digitalizada en los modelos macroeconómicos y comparan sus características con las de las ideas, las patentes o el aprendizaje por la práctica.

Su revisión destaca que los datos permiten innovar, optimizar procesos y mejorar el emparejamiento entre oferentes y demandantes, y que muchas de sus consecuencias para el crecimiento y la distribución de los beneficios siguen siendo objeto de debate. En el plano del desarrollo, el Banco Mundial (2021) sostiene que los datos pueden mejorar significativamente la vida de las personas cuando se utilizan para diseñar mejores políticas, prestar servicios más eficaces y aumentar la rendición de cuentas, pero que ese potencial requiere un contrato social que genere confianza en su uso y distribuya equitativamente sus beneficios.

Para un emprendimiento, la economía de los datos se traduce en una pregunta concreta: ¿qué información genera el negocio en su operación cotidiana y cómo puede aprovecharla? Una tienda en línea registra cada visita, cada búsqueda, cada producto agregado al carrito y cada compra; un servicio de reparto conoce las rutas, los horarios y las preferencias de sus clientes; una empresa de software observa cómo los usuarios interactúan con su aplicación. Incluso un negocio físico pequeño genera datos sobre ventas, inventarios, pagos y proveedores. La mayoría de esta información se pierde o permanece en registros desordenados, sin que nadie la analice.

El primer paso para construir ventajas basadas en datos consiste, precisamente, en registrarla de manera sistemática y conservarla en formatos que permitan su análisis. Sobre esta base, las ventajas competitivas que pueden derivarse de los datos son de varios tipos. Una de ellas es la ventaja de conocimiento: un emprendimiento que comprende mejor que sus competidores las necesidades, los hábitos y la disposición a pagar de sus clientes puede diseñar ofertas más atractivas y fijar precios más adecuados. Otra es la ventaja de eficiencia: los datos sobre procesos permiten reducir desperdicios, optimizar inventarios y asignar mejor los recursos. A ello se suma la ventaja de aprendizaje: cada interacción con un cliente mejora los modelos predictivos, lo que a su vez mejora el producto y atrae más clientes, en un ciclo que se refuerza a sí mismo. Y la cuarta es la ventaja de red: en algunos modelos de negocio, el valor del servicio aumenta a medida que crece el número de usuarios y los datos que estos aportan, lo que dificulta la entrada de competidores.

Estas ventajas, sin embargo, no son automáticas ni permanentes. Los datos por sí solos rara vez constituyen una barrera insuperable, porque los competidores pueden obtener información similar, comprarla a terceros o generar la propia. La ventaja sostenible suele surgir de la combinación entre datos exclusivos, capacidades analíticas, conocimiento del negocio y una cultura organizacional que utiliza la evidencia para decidir. Además, los datos pierden valor con el tiempo cuando las preferencias de los clientes o las condiciones del mercado cambian, lo que obliga a actualizarlos de manera continua.

En la práctica, la investigación sobre la predicción del éxito de empresas emergentes ilustra el valor de las huellas digitales. En este sentido, Bidgoli, Vanani y Goodarzi (2024) encontraron que la presencia en redes profesionales y sociales —medida por el número de seguidores y de empleados registrados— figuraba entre los predictores más importantes del éxito, junto con el monto de la última ronda de financiamiento. Este hallazgo sugiere que la información pública generada por la actividad digital de una empresa puede ser utilizada por inversionistas y competidores para evaluar su potencial. Para los emprendedores, esto implica que su presencia digital no solo cumple funciones comerciales, sino que también comunica señales sobre la solidez del negocio.

La economía de los datos presenta, no obstante, desafíos específicos para los emprendimientos de menor tamaño. El estudio de la OCDE (2025) sobre la adopción de la inteligencia artificial en las pequeñas y medianas empresas identifica la escasez de datos de calidad y las dificultades para acceder a capacidad de cómputo como barreras relevantes. Muchas pequeñas empresas no

generan volúmenes suficientes de información para entrenar modelos propios, y dependen de plataformas de terceros que concentran los datos de miles de negocios. Esta dependencia puede ser útil, porque permite acceder a análisis que de otro modo serían inalcanzables, pero también implica ceder el control sobre un recurso estratégico.

Una alternativa intermedia consiste en la colaboración entre empresas, mediante asociaciones gremiales, cooperativas o consorcios que agreguen datos de manera segura para obtener análisis sectoriales. Desde esta perspectiva, el tratamiento de los datos implica, además, obligaciones jurídicas. Para el Ecuador, la Ley Orgánica de Protección de Datos Personales establece principios y deberes para todos los responsables del tratamiento de información personal, incluidas las empresas privadas de cualquier tamaño (Asamblea Nacional del Ecuador, 2021).

Un emprendimiento que recoge datos de sus clientes —nombres, correos electrónicos, historiales de compra, ubicaciones— debe informarles sobre la finalidad del tratamiento, obtener su consentimiento cuando corresponda, proteger la información frente a accesos indebidos y respetar sus derechos de acceso, rectificación y eliminación. Lejos de constituir solo una carga, el cumplimiento de estas obligaciones puede convertirse en una fuente de confianza y, por tanto, de ventaja competitiva frente a competidores menos cuidadosos.

Con estas consideraciones, una estrategia de datos para un emprendimiento emergente podría organizarse en torno a algunas decisiones clave. Conviene, en primer término, identificar qué datos son estratégicos para el modelo de negocio y cuáles pueden obtenerse de fuentes externas. Es necesario, además, establecer desde el inicio sistemas de registro ordenados, aunque sean sencillos, que permitan conservar la información de ventas, clientes, costos y operaciones. Resulta clave también definir qué preguntas se quieren responder con esos datos, para evitar acumular información sin propósito.

Hace falta, asimismo, desarrollar capacidades analíticas básicas en el equipo o mediante alianzas externas. Y, por último, debe asegurarse el cumplimiento de la normativa de protección de datos y la transparencia con los clientes sobre el uso de su información. De forma complementaria, la construcción de ventajas basadas en datos plantea a los emprendimientos un dilema práctico: qué datos recolectar desde el inicio. Recolectar demasiado genera costos, riesgos de seguridad y obligaciones legales; recolectar demasiado poco impide aprender.

Una orientación razonable consiste en partir de las decisiones más importantes del negocio y preguntarse qué información ayudaría a tomarlas mejor. Si la decisión crítica es cuánto

inventario comprar, los datos de ventas por producto y por día resultan esenciales; si es a qué clientes dirigir una promoción, lo son el historial de compras y los canales de contacto. Este enfoque, centrado en las decisiones, evita la acumulación indiscriminada de información y es coherente con el principio de minimización que la normativa de protección de datos personales exige (Asamblea Nacional del Ecuador, 2021). A medida que el negocio crece, el conjunto de datos puede ampliarse de manera ordenada, documentando su origen, su calidad y las condiciones de su uso, de modo que los datos se conviertan progresivamente en un activo y no en un pasivo.

La economía de los datos ofrece a los emprendimientos emergentes una oportunidad inédita para competir con empresas más grandes sobre la base del conocimiento y no solo del capital. Pero esa oportunidad está condicionada por la capacidad de generar, organizar, proteger y analizar información de manera sistemática. En contextos como el latinoamericano, donde muchas empresas operan con registros incompletos o informales, dar ese primer paso —ordenar los datos propios— puede ser tan transformador como incorporar el algoritmo más avanzado.

Figura 1.5

Economía de los datos y construcción de ventajas competitivas en emprendimientos emergentes.



Fuente: Elaboración autores

1.6. Del análisis retrospectivo a la anticipación empresarial: modelos descriptivos, predictivos y prescriptivos

La mayoría de las empresas utiliza la información principalmente para mirar hacia atrás. Los estados financieros muestran lo que ocurrió en el último ejercicio; los reportes de ventas

resumen el desempeño del mes anterior; los indicadores de gestión comparan lo alcanzado con lo planificado. Este análisis retrospectivo es indispensable, porque permite rendir cuentas, detectar desviaciones y aprender de la experiencia. Sin embargo, tiene una limitación evidente: cuando un problema aparece en los reportes, ya se ha producido.

El tránsito hacia la anticipación empresarial supone complementar la mirada retrospectiva con herramientas capaces de estimar lo que probablemente ocurrirá y de recomendar qué hacer al respecto. Vinculado con lo anterior, la literatura sobre analítica de negocios suele distinguir varios niveles de madurez en el uso de los datos. El análisis descriptivo responde a la pregunta de qué ocurrió, mediante estadísticas, tablas, gráficos y tableros que resumen la información histórica. El análisis diagnóstico indaga por qué ocurrió, al explorar relaciones entre variables y descomponer los resultados según sus causas probables.

El análisis predictivo estima qué es probable que ocurra, utilizando modelos estadísticos o algoritmos de aprendizaje automático que aprenden patrones de los datos pasados para proyectarlos al futuro. Y el análisis prescriptivo busca determinar qué conviene hacer, combinando predicciones con técnicas de optimización, simulación o reglas de decisión para recomendar cursos de acción concretos. Cada nivel se apoya en el anterior: no es posible predecir con fundamento sin una buena descripción de los datos, ni prescribir sin predicciones razonablemente confiables.

El análisis prescriptivo constituye el nivel más exigente y, a la vez, el menos desarrollado. En una revisión de 104 trabajos sobre aplicaciones de analítica prescriptiva basada en datos, Moesmann y Pedersen (2024) la caracterizan como un campo emergente de la analítica de negocios que sugiere opciones concretas para resolver problemas empresariales. Su desarrollo enfrenta desafíos importantes, porque las recomendaciones deben considerar las restricciones reales de la empresa, los objetivos a veces contradictorios de sus distintos actores y la incertidumbre de las predicciones en que se basan.

Para un emprendimiento, un sistema prescriptivo podría, por ejemplo, recomendar cuánto inventario comprar cada semana, qué precio fijar para cada producto según la demanda esperada o qué clientes priorizar en una campaña de cobranza. En concordancia con ello, la distinción entre predicción y prescripción conecta con el marco propuesto por Agrawal, Gans y Goldfarb (2022), según el cual toda decisión combina una predicción sobre lo que ocurrirá con un juicio sobre el valor de cada resultado.

Los modelos predictivos aportan la primera parte; la prescripción exige incorporar la segunda, es decir, las preferencias, los objetivos y las restricciones de quien decide. Por ello, la analítica prescriptiva no puede reducirse a un problema técnico: requiere que la empresa explicite sus prioridades —¿se prefiere maximizar las ventas o el margen?, ¿cuánto riesgo de quiebre de inventario se tolera?— y que las personas responsables comprendan y validen las recomendaciones del sistema.

La evidencia sobre la capacidad predictiva de los modelos aplicados al desempeño empresarial es alentadora, aunque heterogénea. En un estudio con datos de más de un millón trescientas mil empresas entre 2009 y 2022, Pekin y Şengül (2024) compararon diversos algoritmos para identificar empresas de alto crecimiento y encontraron que el área bajo la curva ROC —una medida de la capacidad de discriminación de los modelos— variaba entre 0,53 y 0,91 según el método y la definición de crecimiento utilizada; los algoritmos basados en potenciación de gradiente obtuvieron los mejores resultados, seguidos por los bosques aleatorios y la regresión logística, y las oportunidades de crecimiento aparecieron como el factor más relevante.

En Canadá, Houle y Macdonald (2023) aplicaron métodos de red elástica, bosques aleatorios y redes neuronales a datos administrativos tributarios para anticipar qué empresas se convertirían en empresas de alto crecimiento en un horizonte de tres años, y lograron identificar un subgrupo que contenía a la mayoría de las que efectivamente alcanzaron ese estatus, aunque también incluía empresas de características similares que no lo lograron.

En la práctica, estos resultados ilustran dos lecciones importantes. En primer lugar, cabe señalar que los modelos pueden mejorar sustancialmente la capacidad de anticipar el desempeño respecto de la intuición o de reglas simples, pero rara vez alcanzan una precisión perfecta. En segundo término, conviene advertir que su utilidad depende del propósito: un modelo que identifica un grupo amplio de empresas con potencial de crecimiento puede ser muy útil para focalizar un programa de apoyo, aunque incluya algunas que no crecerán, mientras que el mismo modelo podría resultar insuficiente para decidir una inversión de alto riesgo en una sola empresa. La elección del nivel de exigencia de un modelo debe considerar, por tanto, el costo de los distintos tipos de error, tema que se desarrolla con detalle en el segundo capítulo.

Para los emprendimientos, el tránsito del análisis retrospectivo a la anticipación puede organizarse de manera gradual. En una primera etapa, el objetivo es disponer de información descriptiva confiable: registros ordenados de ventas, costos, clientes e inventarios, y tableros sencillos que permitan observar su evolución. En una segunda etapa, el negocio puede incorporar análisis

diagnósticos que identifiquen, por ejemplo, qué productos generan mayor margen, qué clientes compran con más frecuencia o en qué días se concentran las ventas.

En una tercera etapa, pueden introducirse modelos predictivos para pronosticar la demanda, estimar el riesgo de impago o anticipar necesidades de liquidez. Y en una cuarta, más avanzada, el emprendimiento puede utilizar herramientas prescriptivas que recomienden decisiones operativas o estratégicas. A ello se suma que este recorrido no requiere necesariamente grandes inversiones. Muchas herramientas de hoja de cálculo, programas contables y plataformas de comercio electrónico incluyen hoy funciones de pronóstico y de visualización accesibles para pequeñas empresas.

Lo que sí requiere es disciplina en el registro de la información, claridad sobre las preguntas que se quieren responder y disposición a utilizar la evidencia para decidir. En economías con alta informalidad, como las latinoamericanas, el primer escalón —contar con datos descriptivos confiables— representa ya un avance significativo para una proporción considerable de los negocios.

La anticipación empresarial tiene también una dimensión cultural. Una organización acostumbrada a reaccionar ante los problemas debe aprender a actuar con base en probabilidades, lo que implica aceptar que algunas decisiones tomadas a partir de predicciones resultarán equivocadas. Si un modelo estima un 70 % de probabilidad de que la demanda aumente y la empresa amplía su inventario, habrá ocasiones en que el aumento no se produzca.

Evaluar las decisiones por su calidad —es decir, por la información y el razonamiento con que se tomaron— y no solo por sus resultados es una condición para que la anticipación se consolide como práctica organizacional. Asimismo, la transición entre estos tipos de analítica tiene también una dimensión organizacional. Una empresa que solo produce informes descriptivos suele organizar su información en torno a los cierres contables y a las reuniones periódicas de revisión. Incorporar la predicción exige cambiar los ritmos: los pronósticos deben actualizarse con frecuencia, contrastarse con lo que efectivamente ocurre y comunicarse a quienes toman decisiones operativas.

Pasar a la prescripción supone, además, definir con claridad los objetivos que se optimizan y los límites dentro de los cuales los sistemas pueden recomendar o ejecutar acciones. En los emprendimientos pequeños, donde una sola persona concentra muchas funciones, este cambio puede ser más sencillo en términos de coordinación, pero más difícil en términos de tiempo y

competencias. Por ello, conviene introducir la analítica avanzada en las áreas donde el beneficio es más evidente —el inventario, la cobranza o la fijación de precios— y extenderla progresivamente a medida que la organización aprende a utilizarla.

Visto en conjunto, el paso del análisis retrospectivo a la anticipación transforma la relación de la empresa con el tiempo.

Figura 1.6

Del análisis retrospectivo a la anticipación empresarial: modelos descriptivos, predictivos y prescriptivos.



Fuente: Elaboración autores

1.7. Arquitectura del emprendimiento predictivo 5.0: interacción entre inteligencia humana, algoritmos y sostenibilidad

La expresión 5.0 se ha difundido en los últimos años para designar una etapa del desarrollo tecnológico en la que la eficiencia deja de ser el único criterio de éxito. La Comisión Europea formuló el concepto de Industria 5.0 como una evolución de la Industria 4.0 que sitúa a las personas en el centro de la producción y que incorpora la sostenibilidad ambiental y la resiliencia como objetivos explícitos de la innovación; el punto de partida de la reflexión deja de ser lo que la tecnología permite hacer para convertirse en las necesidades humanas y sociales que debe atender (Breque et al., 2021).

En la misma línea, la noción de Sociedad 5.0 propone que la innovación se valore por su contribución a las personas y que el crecimiento económico se acompañe de respuestas a los desafíos sociales y ecológicos (Carayannis y Morawska-Jancelewicz, 2022). Además, trasladado al

emprendimiento, el enfoque 5.0 supone que la predicción del éxito empresarial no puede limitarse a maximizar la rentabilidad de corto plazo. Un emprendimiento predictivo 5.0 utiliza datos e inteligencia artificial para anticipar su desempeño, pero define ese desempeño de manera amplia, incluyendo la supervivencia en el tiempo, la calidad del empleo que genera, su impacto ambiental, su capacidad para resistir crisis y la confianza que construye con clientes y comunidades. La arquitectura de este enfoque puede describirse a partir de la interacción entre tres componentes: la inteligencia humana, los algoritmos y la sostenibilidad.

La inteligencia humana es el componente que define los propósitos, interpreta los resultados y asume la responsabilidad por las decisiones. Al respecto, Raisch y Krakowski (2021) advierten que la automatización y el aumento de las capacidades humanas son interdependientes, y que privilegiar exclusivamente la primera puede generar ciclos negativos para las organizaciones y la sociedad. En un emprendimiento, esto significa que los algoritmos deben diseñarse para ampliar el juicio del equipo y no para reemplazarlo sin criterio.

Las personas aportan conocimiento del contexto, sensibilidad hacia los valores y las relaciones, creatividad para imaginar alternativas que los datos históricos no contienen y capacidad para reconocer cuándo un modelo deja de ser válido porque las condiciones han cambiado. En esta misma línea, los algoritmos constituyen el segundo componente. Su función es procesar grandes volúmenes de información, identificar patrones, estimar probabilidades y, en los niveles más avanzados, recomendar decisiones.

En esa dirección, Chalmers, MacKenzie y Carter (2021) señalan que la inteligencia artificial puede transformar la generación de ideas, las ventas y el escalamiento de las nuevas empresas, pero también recuerdan que el apalancamiento tecnológico genera pasivos. Entre esos pasivos figuran la dependencia de proveedores tecnológicos, la opacidad de algunos modelos, la posibilidad de reproducir sesgos presentes en los datos y el riesgo de que las decisiones automatizadas afecten derechos de clientes o trabajadores. Por ello, los algoritmos deben estar sujetos a reglas de transparencia, validación y supervisión, temas que se desarrollan en el quinto capítulo.

La sostenibilidad es el tercer componente y el que distingue con mayor claridad al enfoque 5.0. Incluye la sostenibilidad económica, entendida como la capacidad del negocio para generar ingresos suficientes y mantener una estructura financiera sana; la sostenibilidad social, que abarca la calidad del empleo, la relación con la comunidad y el respeto de los derechos; y la sostenibilidad ambiental, que considera el uso de recursos, las emisiones y los residuos.

Incorporar estas dimensiones en los modelos predictivos implica ampliar las variables que se miden y los resultados que se intentan anticipar. Un modelo que solo predice la rentabilidad podría recomendar estrategias que comprometen la reputación o el cumplimiento ambiental de la empresa; uno que integra la sostenibilidad permite identificar esos riesgos con anticipación. De igual manera, la resiliencia ocupa un lugar especial dentro de este componente. Sobre este punto, Saad et al. (2021), en una revisión de 118 estudios sobre la resiliencia de las pequeñas y medianas empresas, encontraron que la literatura presenta definiciones y mediciones muy variadas, que los resultados sobre los factores que la favorecen no son concluyentes y que la investigación sobre empresas de países en desarrollo es escasa, pese a su enorme importancia económica en esos contextos.

Para el emprendimiento predictivo, la resiliencia implica no solo anticipar el escenario más probable, sino prepararse para escenarios adversos: simular el efecto de una caída abrupta de la demanda, de un aumento de costos o de una interrupción en el suministro, y diseñar respuestas antes de que esos eventos ocurran.

A partir de estos tres componentes puede describirse una arquitectura en capas. En la base se ubica la capa de datos, que reúne la información interna del negocio —ventas, costos, clientes, operaciones— y la información externa —mercado, competencia, entorno económico, normativa—, organizada con criterios de calidad y protección. Sobre ella se sitúa la capa analítica, donde los modelos descriptivos, predictivos y prescriptivos transforman los datos en estimaciones y recomendaciones.

La tercera capa es la de decisión, en la que las personas interpretan esas estimaciones, las contrastan con su conocimiento del contexto y eligen los cursos de acción. La cuarta capa corresponde a la gobernanza, que establece reglas sobre el uso de los datos y de los algoritmos, define responsabilidades y asegura la supervisión humana. Y la quinta, transversal a todas, es la de sostenibilidad, que orienta los objetivos del sistema hacia la creación de valor económico, social y ambiental. Por otra parte, esta arquitectura funciona como un ciclo de aprendizaje.

Las decisiones tomadas producen resultados que se registran como nuevos datos; esos datos permiten evaluar la calidad de las predicciones y ajustar los modelos; y los modelos mejorados orientan nuevas decisiones. La intervención humana resulta decisiva en cada vuelta del ciclo, porque es la que determina si los resultados son satisfactorios, si los modelos siguen siendo pertinentes y si los objetivos deben revisarse. Un sistema que funciona sin esa supervisión corre el riesgo de optimizar con creciente precisión objetivos equivocados.

En el contexto ecuatoriano y latinoamericano, la arquitectura del emprendimiento predictivo 5.0 debe adaptarse a la realidad de negocios mayoritariamente pequeños, con recursos limitados y, en muchos casos, en condiciones de informalidad. Para ellos, la capa de datos puede comenzar con registros sencillos en hojas de cálculo o aplicaciones contables; la capa analítica, con indicadores básicos y pronósticos simples; la capa de decisión, con reuniones periódicas en las que el equipo revise la información; y la capa de gobernanza, con reglas claras sobre el manejo de la información de los clientes.

La sostenibilidad, por su parte, no es un objetivo reservado a las grandes empresas: un pequeño negocio que reduce sus desperdicios, trata con justicia a sus trabajadores y construye relaciones de confianza con su comunidad aumenta también sus probabilidades de permanecer en el tiempo. Paralelamente, la arquitectura propuesta adquiere sentido cuando se observa su funcionamiento en un caso concreto. Considérese una pequeña empresa de alimentos procesados que vende en tiendas de barrio y supermercados de una ciudad intermedia.

En la capa de datos, registra sus ventas por cliente y producto mediante la facturación electrónica y obtiene información de precios de insumos. En la capa analítica, un modelo sencillo pronostica la demanda semanal y otro estima el riesgo de retraso en los pagos de cada cliente. En la capa de decisión, la persona emprendedora utiliza estos pronósticos para planificar la producción y negociar plazos, apoyándose en su conocimiento de los clientes y de la temporada. En la capa de sostenibilidad, se vigila el desperdicio de alimentos y el consumo de energía. Y en la capa de gobernanza, se establecen reglas sobre el uso de los datos de los clientes. Aunque estos componentes no requieren grandes inversiones, su articulación transforma la manera en que la empresa anticipa y decide.

Lo expuesto permite afirmar que la arquitectura del emprendimiento predictivo 5.0 propone una integración equilibrada entre capacidades humanas y algorítmicas al servicio de objetivos amplios de desempeño. Su valor no reside en la sofisticación de la tecnología, sino en la coherencia entre los datos que se recogen, los modelos que se aplican, las decisiones que se toman y los propósitos que guían a la empresa. Esta coherencia es la que permite que la predicción contribuya no solo a la supervivencia del negocio, sino también a su contribución al desarrollo de su entorno.

Figura 1.7

Arquitectura del emprendimiento predictivo 5.0: interacción entre inteligencia humana, algoritmos y sostenibilidad.



Fuente: Elaboración autores

1.8. Nuevos indicadores de viabilidad, resiliencia y potencial de crecimiento empresarial

Los indicadores tradicionales con que se evalúa a una empresa —rentabilidad, liquidez, endeudamiento, crecimiento de las ventas— fueron diseñados para organizaciones con trayectoria, estados financieros auditados y operaciones relativamente estables. Aplicados a emprendimientos nacientes, estos indicadores presentan limitaciones evidentes: muchas empresas emergentes registran pérdidas durante sus primeros años sin que ello signifique que no sean viables; otras muestran crecimientos muy elevados a partir de bases pequeñas que no garantizan su consolidación; y una proporción considerable carece de información contable completa.

El emprendimiento predictivo requiere, por ello, un conjunto renovado de indicadores que permita evaluar tres dimensiones complementarias: la viabilidad, la resiliencia y el potencial de crecimiento. Sobre esta base, la viabilidad se refiere a la capacidad de un emprendimiento para sostenerse económicamente en el tiempo.

Los estudios ecuatorianos sobre supervivencia empresarial ofrecen pistas sobre las variables más relevantes. En Manabí, González Cedeño (2026) encontró que la rentabilidad sobre los activos y el volumen de ventas reducían el riesgo de cierre, mientras que el apalancamiento financiero lo aumentaba, y que el riesgo se concentraba en los primeros años de vida de las

empresas. En el ámbito nacional, Zambrano et al. (2021) identificaron el patrimonio, la rotación de cartera y el equilibrio del endeudamiento como factores que prolongan la vida de las micro, pequeñas y medianas empresas. Estas variables pueden traducirse en indicadores de viabilidad adaptados a emprendimientos: la distancia al punto de equilibrio, los meses de operación que el efectivo disponible permite cubrir, la proporción de ingresos recurrentes, el margen de contribución por producto o la relación entre la deuda y la capacidad de generación de efectivo.

La resiliencia constituye la segunda dimensión y, probablemente, la más difícil de medir. En este sentido, Saad et al. (2021) constataron, en su revisión de la literatura sobre la resiliencia de las pequeñas y medianas empresas, que no existe consenso sobre su definición ni sobre su medición, y que la evidencia sobre los factores que la determinan es inconclusa. A pesar de ello, es posible identificar algunos rasgos que la literatura asocia con la capacidad de resistir y recuperarse frente a crisis: la diversificación de clientes, productos y proveedores; la disponibilidad de reservas financieras; la flexibilidad de la estructura de costos; las redes de apoyo con otras empresas e instituciones; y la capacidad de aprendizaje y adaptación del equipo. Indicadores como la concentración de las ventas en los principales clientes, la proporción de costos fijos, el número de proveedores alternativos o la velocidad de respuesta ante cambios en la demanda pueden servir como aproximaciones a esta dimensión.

Desde esta perspectiva, el potencial de crecimiento es la tercera dimensión. La literatura ha desarrollado definiciones operativas de las empresas de alto crecimiento basadas en el aumento del empleo o de la facturación durante un período determinado. Pekin y Şengül (2024) utilizaron ambas definiciones en su estudio con datos de más de un millón de empresas, y encontraron que los modelos de aprendizaje automático lograban anticipar el alto crecimiento con una capacidad de discriminación que, en los mejores casos, superaba el 0,9 en el área bajo la curva ROC; las oportunidades de crecimiento aparecieron como el factor más importante.

Al respecto, Houle y Macdonald (2023), por su parte, se propusieron identificar empresas que se convertirían en empresas de alto crecimiento en los tres años siguientes, a partir de información financiera, laboral, sectorial y geográfica. Estos trabajos sugieren que el potencial de crecimiento puede estimarse con indicadores que combinan la evolución reciente de las ventas y del empleo con características del sector, del mercado y de la propia organización.

Junto a estas tres dimensiones, el emprendimiento predictivo incorpora indicadores basados en fuentes de información no tradicionales. Los datos generados por la actividad digital —

seguidores en redes sociales, interacciones con clientes, reseñas, visitas a la página web, transacciones en plataformas de comercio electrónico— ofrecen señales tempranas sobre la tracción de un negocio.

En esa dirección, Bidgoli, Vanani y Goodarzi (2024) mostraron que la presencia en redes profesionales y sociales figuraba entre los predictores más relevantes del éxito de empresas emergentes. Estos indicadores tienen la ventaja de estar disponibles antes que la información contable y de actualizarse con mayor frecuencia, aunque también presentan riesgos: pueden ser manipulados, reflejan con frecuencia la popularidad más que la solidez económica y no están disponibles para negocios sin presencia digital. De forma complementaria, las estadísticas oficiales permiten situar estos indicadores en su contexto.

Según el registro estadístico de empresas del Instituto Nacional de Estadística y Censos, en 2023 el Ecuador contaba con 1.173.985 empresas, lo que representó una disminución de 5,3 % respecto del año anterior, aunque las ventas totales en términos reales y las plazas de empleo registrado crecieron en 3,1 % (Instituto Nacional de Estadística y Censos [INEC], 2024). Paralelamente, el informe nacional del Global Entrepreneurship Monitor registró una tasa de abandono de 17,33 % entre los emprendedores en etapa temprana (Lasio et al., 2024). Estas cifras revelan una dinámica de creación y destrucción de empresas intensa, en la que disponer de indicadores tempranos de viabilidad y resiliencia podría contribuir a reducir los cierres evitables.

La construcción de un sistema de indicadores para el emprendimiento predictivo debe seguir algunos principios. El primero es la pertinencia: cada indicador debe estar vinculado con una dimensión del desempeño que se pretende anticipar y con decisiones concretas que la empresa o sus financiadores pueden tomar. El segundo es la parsimonia: es preferible un conjunto reducido de indicadores relevantes que una lista extensa que nadie utiliza. El tercero es la oportunidad: los indicadores más valiosos son aquellos que anticipan problemas con tiempo suficiente para actuar.

El cuarto es la comparabilidad, que exige definiciones estables y, cuando sea posible, referencias sectoriales. Y el quinto es la equidad, que obliga a revisar si los indicadores favorecen sistemáticamente a determinados tipos de empresas —por ejemplo, las que tienen presencia digital o acceso previo al crédito— en detrimento de otras con potencial similar. Vinculado con lo anterior, estos indicadores pueden integrarse en índices compuestos que resuman la situación de un emprendimiento en una sola medida, tema que se desarrolla en el segundo capítulo al

abordar la construcción de un índice predictivo multidimensional. Conviene, no obstante, advertir que la síntesis en un único número facilita la comunicación, pero puede ocultar información importante. Un índice de éxito empresarial debe acompañarse siempre de la desagregación de sus componentes, para que quien lo utiliza pueda identificar en qué dimensión se encuentran las fortalezas y las debilidades del negocio.

La adopción de estos indicadores enfrenta, sin embargo, el problema de su medición en contextos de información limitada. Muchos emprendimientos latinoamericanos no llevan contabilidad formal, y los indicadores de resiliencia o de potencial de crecimiento requieren información que rara vez se registra de manera sistemática, como la diversificación de clientes y proveedores, la existencia de reservas de liquidez o la capacidad de adaptar la oferta ante cambios en la demanda.

Una alternativa consiste en combinar fuentes: los registros de facturación electrónica y de pagos digitales permiten construir indicadores financieros aproximados; las encuestas breves y estructuradas recogen información sobre capacidades y prácticas de gestión; y los datos del entorno, provenientes de estadísticas oficiales y de los informes sobre la actividad emprendedora (Lasio et al., 2024), sitúan a cada negocio en relación con su sector y territorio por lo que la construcción de estos indicadores debe seguir procedimientos rigurosos de normalización, ponderación y análisis de robustez para que sus resultados sean comparables y confiables.

Figura 1.8

Nuevos indicadores de viabilidad, resiliencia y potencial de crecimiento empresarial.



Fuente: Elaboración autores

EDITORIAL ANDES COGNITIO

CAPÍTULO II

INTELIGENCIA ARTIFICIAL Y MODELOS ALGORÍTMICOS PARA PREDECIR EL ÉXITO O FRACASO DE LOS EMPRENDIMIENTOS



CAPÍTULO II

INTELIGENCIA ARTIFICIAL Y MODELOS ALGORÍTMICOS PARA PREDECIR EL ÉXITO O FRACASO DE LOS EMPRENDIMIENTOS

2.1. Fundamentos del aprendizaje automático aplicados al análisis del éxito empresarial

El aprendizaje automático es la rama de la inteligencia artificial que permite a los sistemas informáticos aprender patrones a partir de datos, sin que cada regla deba programarse de manera explícita. En lugar de que una persona establezca que un negocio con determinadas características tiene alta probabilidad de fracasar, el algoritmo examina miles de casos históricos —empresas que sobrevivieron y empresas que cerraron, con sus respectivos atributos— e identifica por sí mismo las combinaciones de variables que mejor distinguen a unas de otras.

Una vez entrenado, el modelo puede aplicarse a empresas nuevas para estimar su probabilidad de éxito o de fracaso. Esta capacidad convierte al aprendizaje automático en la columna vertebral técnica del emprendimiento predictivo. En concordancia con ello, James, Witten, Hastie y Tibshirani (2021), en una de las obras de referencia más utilizadas en la enseñanza de esta disciplina, distinguen dos grandes familias de problemas. En el aprendizaje supervisado, cada observación de los datos de entrenamiento incluye tanto las variables explicativas como el resultado que se quiere predecir; el objetivo es aprender la relación entre ambos para estimar el resultado en casos nuevos.

Predecir si una empresa sobrevivirá cinco años, cuánto venderá el próximo trimestre o si un cliente pagará su crédito son problemas supervisados. En el aprendizaje no supervisado, en cambio, los datos no contienen un resultado conocido, y el propósito es descubrir estructuras ocultas, como grupos de empresas o de clientes con características similares. Ambas familias tienen aplicaciones en el análisis del éxito empresarial: la primera para anticipar resultados y la segunda para comprender la diversidad de trayectorias.

Dentro del aprendizaje supervisado, los problemas se dividen a su vez según el tipo de resultado. Cuando la variable de interés es categórica —por ejemplo, éxito o fracaso, supervivencia o cierre—, se habla de clasificación. Cuando es numérica —ventas, utilidades, número de empleados—, se habla de regresión. Muchos estudios sobre el éxito empresarial se plantean como problemas de clasificación binaria, porque las definiciones de éxito suelen reducirse a dos

estados: la empresa alcanzó o no un determinado hito, como sobrevivir un período, recibir una ronda de financiamiento, ser adquirida o salir a bolsa. Esta simplificación facilita el modelado, pero obliga a reflexionar cuidadosamente sobre qué significa el éxito en cada contexto.

A ello se suma que la definición del resultado es, de hecho, una de las decisiones más importantes en cualquier proyecto de predicción del desempeño empresarial. En los estudios sobre empresas emergentes de base tecnológica, el éxito suele asociarse con eventos como la adquisición por otra compañía, la salida a bolsa o la obtención de rondas sucesivas de inversión. En los estudios sobre micro, pequeñas y medianas empresas tradicionales, en cambio, el criterio más frecuente es la supervivencia o la ausencia de cierre durante un período determinado.

La literatura sobre empresas de alto crecimiento, por su parte, utiliza umbrales de aumento del empleo o de las ventas. Pekin y Şengül (2024) mostraron que la capacidad predictiva de los modelos varía según la definición adoptada: con los mismos datos y algoritmos, los resultados difieren cuando el alto crecimiento se mide por el empleo o por la facturación. Elegir la definición adecuada depende de la pregunta que se quiere responder y de quién utilizará la predicción.

El proceso de construcción de un modelo de aprendizaje automático sigue una secuencia que conviene conocer. Comienza con la recolección y preparación de los datos, que incluye la depuración de errores, el tratamiento de valores faltantes y la transformación de variables. Continúa con la selección de las variables explicativas, que pueden ser financieras, operativas, sectoriales, territoriales o relacionadas con las personas fundadoras. Sigue la división de los datos en un conjunto de entrenamiento, utilizado para que el algoritmo aprenda, y un conjunto de prueba, reservado para evaluar su desempeño en casos que no ha visto. Luego se entrenan uno o varios algoritmos, se ajustan sus parámetros y se comparan sus resultados.

Finalmente, el modelo seleccionado se valida, se interpreta y, si resulta satisfactorio, se pone en operación con mecanismos de seguimiento. Con todo, una advertencia metodológica central en este proceso es el riesgo de utilizar información que no estaría disponible en el momento en que se quiere hacer la predicción. Sobre este punto, Żbikowski y Antosiuk (2021) llamaron la atención sobre el sesgo de anticipación en los modelos que predicen el éxito de empresas emergentes a partir de bases de datos como Crunchbase: si se incluyen variables que solo se conocen después de que la empresa ha avanzado, el modelo parecerá muy preciso en la evaluación, pero fallará cuando se aplique a empresas nuevas. Su propuesta de un enfoque libre de este sesgo, con algoritmos de potenciación de gradiente, buscaba precisamente que el modelo pudiera utilizarse como sistema de apoyo a decisiones reales.

En la misma línea, Park, Choi y Feng (2024) identificaron dos sesgos frecuentes en estos modelos: el uso de predictores que no están disponibles en la etapa temprana y el desequilibrio entre clases, que se produce cuando los casos de éxito son muy pocos en comparación con los de fracaso.

El desequilibrio de clases merece una atención especial en el análisis del éxito empresarial. En muchas bases de datos, las empresas que alcanzan un resultado excepcional representan una fracción pequeña del total. En este sentido, Park et al. (2024) trabajaron con datos en los que la proporción entre casos de fracaso y de éxito era de aproximadamente nueve a uno, y observaron que algunos algoritmos tradicionales, como la regresión logística y las máquinas de vectores de soporte, identificaban muy pocos casos exitosos cuando se entrenaban con los datos originales.

Al generar observaciones sintéticas de la clase minoritaria mediante redes generativas adversariales, el desempeño de todos los clasificadores mejoró de manera sustancial. Este hallazgo ilustra que la calidad de un modelo no depende solo del algoritmo, sino también del tratamiento previo de los datos. Asimismo, los estudios empíricos recientes ofrecen una idea de lo que pueden lograr estas técnicas. Al respecto, Bidgoli, Vanani y Goodarzi (2024) compararon bosques aleatorios, potenciación de gradiente, perceptrones multicapa, regresión logística y máquinas de vectores de soporte para predecir el éxito de empresas emergentes, y obtuvieron la mayor exactitud con los bosques aleatorios, con 82 %, seguidos por la potenciación de gradiente, con 80 %.

Además, utilizaron técnicas de agrupamiento para describir perfiles de empresas y métodos de interpretación para identificar las variables más influyentes. Pekin y Şengül (2024), con una base de más de un millón de empresas, encontraron que los algoritmos de potenciación de gradiente superaban a los bosques aleatorios y a la regresión logística, mientras que los vecinos más cercanos y el clasificador bayesiano ingenuo mostraban un desempeño inferior.

Estos resultados no deben interpretarse como una jerarquía universal de algoritmos. El desempeño relativo de cada técnica depende del tamaño y la naturaleza de los datos, del tipo de variables disponibles, de la definición del resultado y de la manera en que se evalúa el modelo. Un principio ampliamente aceptado en la disciplina es que no existe un algoritmo superior en todos los problemas, por lo que la práctica recomendada consiste en comparar varias alternativas con criterios de evaluación adecuados y elegir la que mejor se ajusta al propósito.

Para los emprendimientos y las instituciones que los apoyan, esto significa que no basta con adoptar la técnica de moda: es necesario evaluar cuál funciona mejor con los datos y las preguntas propias. Además, un aspecto que distingue la aplicación del aprendizaje automático al análisis empresarial es la necesidad de interpretar los resultados. A diferencia de otros ámbitos, en los que basta con que un modelo acierte, en la gestión empresarial y financiera es necesario comprender por qué el modelo predice lo que predice, tanto para generar confianza como para identificar acciones de mejora. Si un modelo estima que un emprendimiento tiene alta probabilidad de cierre, la persona emprendedora necesita saber qué factores explican esa estimación para poder actuar sobre ellos. Por ello, la interpretabilidad, que se desarrolla en detalle más adelante en este capítulo, no es un complemento opcional, sino un requisito para que el aprendizaje automático contribuya efectivamente a mejorar las decisiones.

La aplicación de estas técnicas en el contexto ecuatoriano y latinoamericano enfrenta desafíos particulares. Las bases de datos más utilizadas en la literatura internacional, como las de empresas emergentes tecnológicas, representan un universo muy distinto del tejido empresarial de la región, dominado por microempresas tradicionales. Los registros administrativos disponibles, como los de la Superintendencia de Compañías, Valores y Seguros, cubren principalmente a las sociedades formales, mientras que una gran parte de los emprendimientos opera como personas naturales o en la informalidad.

Construir modelos pertinentes para la realidad regional exige, por tanto, combinar fuentes administrativas, encuestas y datos generados por las propias empresas, y validar cuidadosamente que los patrones aprendidos se apliquen a los negocios que se pretende evaluar. En esta misma línea, la aplicación del aprendizaje automático al estudio del éxito empresarial exige también atender a la calidad de la variable que se pretende predecir. En muchos estudios, el éxito se aproxima mediante eventos observables, como la obtención de una ronda de financiamiento, una adquisición o la permanencia en un registro empresarial.

Estas aproximaciones son útiles, pero imperfectas: una empresa puede seguir registrada sin actividad real, o puede cerrar voluntariamente tras un resultado satisfactorio para sus fundadores. Los errores en la variable objetivo se trasladan a los modelos, que aprenden a predecir la aproximación y no el fenómeno de interés. Por ello, la definición operativa del éxito debe discutirse explícitamente y, cuando sea posible, contrastarse con varias medidas alternativas.

En esa dirección, James et al. (2021) recuerdan que ningún método de aprendizaje estadístico domina a los demás en todos los conjuntos de datos, y que la elección del enfoque depende

tanto de la estructura de los datos como de la pregunta que se desea responder, lo que vale también para la definición de la variable de resultado.

Todo ello indica que el aprendizaje automático ofrece un conjunto poderoso de herramientas para analizar el éxito empresarial, siempre que se aplique con rigor metodológico. La definición clara del resultado, la preparación cuidadosa de los datos, la prevención del sesgo de anticipación, el tratamiento del desequilibrio de clases, la comparación entre algoritmos y la interpretación de los resultados son las condiciones que separan un modelo útil de uno que solo parece preciso.

Figura 2.1

Fundamentos del aprendizaje automático aplicados al análisis del éxito empresarial



Fuente: Elaboración autores

2.2. Modelos econométricos Logit y Probit para estimar probabilidades de supervivencia y fracaso empresarial

Antes de la expansión del aprendizaje automático, la econometría ya había desarrollado herramientas rigurosas para estimar la probabilidad de que ocurra un evento binario, como el cierre o la supervivencia de una empresa. Los modelos Logit y Probit son las más conocidas y siguen siendo ampliamente utilizados en la investigación y en la práctica financiera. Su vigencia no se explica por inercia: ofrecen resultados interpretables, permiten evaluar la significancia

estadística de cada variable, requieren menos datos que muchos algoritmos complejos y, en numerosos problemas, alcanzan una capacidad predictiva comparable a la de técnicas más sofisticadas.

Para el emprendimiento predictivo, constituyen un punto de partida natural y una referencia obligada contra la cual comparar cualquier modelo más complejo. De igual manera, el problema que resuelven estos modelos es el siguiente. Cuando la variable que se quiere explicar solo puede tomar dos valores —por ejemplo, 1 si la empresa cerró y 0 si sobrevivió—, una regresión lineal convencional resulta inadecuada, porque puede producir probabilidades negativas o superiores a uno. Los modelos Logit y Probit resuelven este inconveniente transformando la combinación lineal de las variables explicativas mediante una función que la convierte en una probabilidad comprendida entre cero y uno. En el modelo Logit, esa transformación utiliza la función logística; en el Probit, la función de distribución acumulada de la normal estándar. En la práctica, ambas producen resultados muy similares, y la elección entre ellas suele responder a convenciones disciplinarias o a la facilidad de interpretación.

La regresión logística ofrece una ventaja interpretativa particular. Sus coeficientes pueden expresarse como razones de probabilidades, que indican cuánto se multiplica la relación entre la probabilidad de que ocurra el evento y la de que no ocurra cuando una variable aumenta en una unidad, manteniendo constantes las demás. Sobre este punto, James et al. (2021) presentan este modelo como una de las técnicas fundamentales de clasificación y destacan que, además de estimar probabilidades, permite comprender la dirección y la magnitud del efecto de cada variable. Para un emprendedor o un analista de crédito, saber que un aumento del endeudamiento eleva en determinada proporción las probabilidades de cierre resulta más útil que una predicción sin explicación.

Por otra parte, la aplicación de estos modelos al estudio de la supervivencia empresarial en el Ecuador ha producido hallazgos relevantes. En este sentido, Zambrano et al. (2021) estimaron un modelo Probit con datos de micro, pequeñas y medianas empresas registradas en la Superintendencia de Compañías, Valores y Seguros, y encontraron que un mayor patrimonio, una rotación de cartera adecuada y un endeudamiento equilibrado aumentaban la esperanza de vida de las empresas; observaron también diferencias según la región, con mejores perspectivas de supervivencia en la Sierra, y según el sector, con resultados favorables en transporte y construcción. Estos resultados ilustran cómo un modelo relativamente sencillo puede orientar tanto la gestión de las empresas como el diseño de políticas de apoyo.

La regresión logística se ha aplicado también al análisis del éxito de microempresas. Al respecto, Cueva et al. (2025) estudiaron 213 microempresas de Machala durante la pandemia y encontraron que el género de la persona propietaria y las actividades de mercadeo resultaban significativas para explicar el éxito de sus modelos de negocio, con las segundas como el predictor de mayor peso; el modelo alcanzó una exactitud global de clasificación de 95,8 %. Aunque el tamaño de la muestra y su selección no probabilística limitan la generalización de los resultados, el estudio muestra la aplicabilidad de estas herramientas a negocios de pequeña escala, en los que la información disponible es limitada.

Paralelamente, la comparación entre los modelos econométricos y los algoritmos de aprendizaje automático no siempre favorece a estos últimos. En esa dirección, Zizi et al. (2021), al analizar el riesgo de dificultades financieras en pequeñas y medianas empresas marroquíes, compararon la regresión logística con redes neuronales a partir de razones financieras seleccionadas mediante técnicas de regularización y de selección paso a paso. El mejor modelo fue la regresión logística con selección paso a paso, que alcanzó una exactitud global de 93,33 % dos años antes de la dificultad financiera y de 95 % un año antes; la relación entre los intereses y las ventas y la rentabilidad sobre los activos aparecieron como los predictores más relevantes. Este resultado recuerda que la complejidad del algoritmo no garantiza mejores predicciones, especialmente cuando los datos son limitados.

Los modelos Logit y Probit presentan, sin embargo, limitaciones que conviene conocer. En primer lugar, cabe señalar que estos modelos suponen una relación de forma específica entre las variables explicativas y la probabilidad del evento, lo que puede resultar restrictivo cuando las relaciones son no lineales o cuando existen interacciones complejas entre variables. Es posible incorporar términos cuadráticos o interacciones, pero ello exige que el analista los especifique previamente.

La segunda limitación se refiere a la multicolinealidad: cuando varias variables explicativas están muy correlacionadas, como ocurre con frecuencia con las razones financieras, los coeficientes pueden volverse inestables y difíciles de interpretar. La tercera tiene que ver con los eventos poco frecuentes: cuando la proporción de empresas que cierran es muy baja, los modelos pueden subestimar sistemáticamente la probabilidad del evento, lo que ha llevado a desarrollar ajustes específicos y funciones de enlace alternativas. Ahora bien, una cuarta limitación es de naturaleza temporal. Los modelos Logit y Probit estiman la probabilidad de que el evento ocurra en un período determinado, pero no consideran cuándo ocurre.

Para estudiar la duración de las empresas, la literatura recurre a los modelos de supervivencia, como el de riesgos proporcionales de Cox, que analizan el tiempo hasta el cierre y permiten incorporar empresas que todavía no han cerrado al final del período de observación. Sobre este punto, Chico y Altamirano (2023) y González Cedeño (2026) aplicaron este tipo de modelos a empresas de Ambato y Manabí, respectivamente, y mostraron que la rentabilidad, las ventas y los activos reducen el riesgo de cierre, mientras que el apalancamiento financiero lo aumenta. La complementariedad entre los modelos de probabilidad y los de supervivencia permite responder no solo si una empresa cerrará, sino también en qué momento es más vulnerable.

La construcción de un buen modelo Logit o Probit para el emprendimiento predictivo exige atender a varias decisiones. La selección de variables debe basarse en la teoría y en la evidencia previa, evitando incluir indiscriminadamente todas las disponibles. Las razones financieras deben calcularse con información anterior al evento que se quiere predecir, para no incurrir en el sesgo de anticipación.

La muestra debe ser representativa de las empresas a las que se aplicará el modelo y la evaluación debe realizarse con datos distintos de los utilizados para estimarlo, mediante las métricas que se examinan más adelante en este capítulo. Cuando se cumplen estas condiciones, los modelos econométricos ofrecen una herramienta robusta, transparente y relativamente fácil de comunicar, sobre esta base, en el ecosistema emprendedor, estos modelos tienen aplicaciones concretas.

Las entidades financieras pueden utilizarlos para evaluar solicitudes de crédito de pequeños negocios; las incubadoras, para identificar a las empresas que requieren acompañamiento intensivo; los gobiernos, para focalizar programas de apoyo en los sectores o territorios con mayor riesgo de mortalidad empresarial; y los propios emprendedores, para comprender qué variables de su gestión influyen más en su probabilidad de supervivencia. En todos los casos, la transparencia de los coeficientes facilita la rendición de cuentas y permite a las personas afectadas comprender los criterios con que se las evalúa.

Una virtud adicional de los modelos Logit y Probit reside en su capacidad para comunicar resultados a públicos no especializados. Los efectos marginales pueden expresarse en términos intuitivos: por ejemplo, cuántos puntos porcentuales aumenta la probabilidad de supervivencia cuando la empresa incrementa su rentabilidad o reduce su endeudamiento, manteniendo constantes las demás condiciones. Esta forma de presentar los resultados facilita el diálogo entre

analistas, emprendedores y responsables de programas de apoyo, y permite traducir los hallazgos en recomendaciones concretas.

En el Ecuador, los estudios que han aplicado estos modelos a las empresas registradas en la Superintendencia de Compañías, Valores y Seguros ofrecen una referencia sobre los factores asociados a la permanencia de las mipymes (Zambrano et al., 2021), que puede servir como punto de partida para modelos más específicos por sector o por territorio. La combinación de estos modelos interpretables con algoritmos más flexibles, en una estrategia de comparación sistemática, permite aprovechar las ventajas de ambos enfoques.

Desde esta perspectiva, las extensiones de estos modelos permiten abordar preguntas más específicas. Los modelos multinomiales, por ejemplo, distinguen entre varios desenlaces posibles —continuar operando, cerrar, ser adquirido o transformarse—, en lugar de reducir el resultado a una dicotomía. Los modelos ordenados resultan adecuados cuando el resultado tiene categorías jerárquicas, como niveles de crecimiento bajo, medio y alto.

En cuanto a los modelos con datos de panel permiten aprovechar la información de las mismas empresas observadas en varios períodos, controlando las características que no cambian en el tiempo. Estas variantes conservan la interpretabilidad que caracteriza a la familia Logit y Probit, al tiempo que amplían su capacidad para representar la diversidad de trayectorias que siguen los emprendimientos.

Los modelos Logit y Probit conservan un lugar central en la caja de herramientas del emprendimiento predictivo. Su capacidad para estimar probabilidades interpretables con datos moderados, su sólida fundamentación estadística y su desempeño competitivo en muchos problemas los convierten en una referencia indispensable.

Los algoritmos de aprendizaje automático pueden mejorar sus resultados cuando las relaciones son complejas y los datos abundantes, pero una práctica rigurosa exige comparar siempre cualquier modelo sofisticado con estos modelos de base, para verificar que la complejidad adicional se traduce en una mejora real.

Figura 2.2

Modelos econométricos Logit y Probit para estimar probabilidades de supervivencia y fracaso empresarial



Fuente: Elaboración autores

2.3. Algoritmos Random Forest y árboles de decisión para identificar determinantes del desempeño emprendedor

Los árboles de decisión figuran entre las técnicas de aprendizaje automático más intuitivas. Su funcionamiento imita la manera en que una persona experimentada podría razonar al evaluar un negocio: primero pregunta si la empresa genera utilidades; si la respuesta es negativa, pregunta si cuenta con reservas de efectivo; si las tiene, pregunta por su nivel de endeudamiento, y así sucesivamente, hasta llegar a una conclusión sobre su probabilidad de supervivencia.

Cada pregunta divide a las empresas en grupos más homogéneos, y el resultado final es una estructura ramificada que puede representarse gráficamente y leerse como un conjunto de reglas. Esta transparencia convierte a los árboles en una herramienta atractiva para comunicar los determinantes del desempeño emprendedor a personas sin formación técnica.

De forma complementaria, James et al. (2021) describen cómo se construye un árbol de decisión. El algoritmo examina todas las variables disponibles y todos los posibles puntos de corte, y elige en cada paso la división que produce los grupos más homogéneos respecto del resultado que se quiere predecir. El proceso se repite en cada grupo resultante hasta que se cumple algún criterio de detención, como un número mínimo de observaciones por grupo o una profundidad máxima del árbol.

Para evitar que el árbol se ajuste en exceso a las particularidades de los datos de entrenamiento, se aplican técnicas de poda que eliminan las ramas que no mejoran la capacidad predictiva en datos nuevos.

Los árboles de decisión presentan ventajas importantes para el análisis empresarial. Capturan de manera natural las relaciones no lineales y las interacciones entre variables, sin que el analista deba especificarlas de antemano. Por ejemplo, pueden revelar que el endeudamiento solo aumenta el riesgo de cierre cuando la rentabilidad es baja, o que la antigüedad protege a las empresas de un sector pero no a las de otro. Manejan variables numéricas y categóricas sin necesidad de transformaciones complejas. Son relativamente robustos frente a valores atípicos. Y, sobre todo, producen reglas comprensibles que pueden discutirse con los emprendedores y con los responsables de las decisiones.

Vinculado con lo anterior, su principal debilidad es la inestabilidad. Pequeños cambios en los datos de entrenamiento pueden producir árboles muy diferentes, y un único árbol suele tener una capacidad predictiva limitada en comparación con otros métodos. Para superar este problema surgieron los métodos de conjunto, que combinan muchos árboles para obtener predicciones más estables y precisas. El bosque aleatorio, conocido en la literatura internacional como random forest, es el más difundido de ellos.

Su lógica consiste en construir cientos de árboles, cada uno entrenado con una muestra aleatoria de las observaciones y considerando en cada división solo un subconjunto aleatorio de las variables; la predicción final se obtiene promediando o sometiendo a votación los resultados de todos los árboles. Esta doble aleatorización reduce la correlación entre los árboles y mejora la capacidad de generalización del conjunto (James et al., 2021).

La evidencia empírica sobre la predicción del éxito empresarial muestra que los bosques aleatorios se encuentran entre los algoritmos con mejor desempeño. En este sentido, Bidgoli, Vanani y Goodarzi (2024) obtuvieron con esta técnica la mayor exactitud, de 82 %, en la predicción del éxito de empresas emergentes, por encima de la potenciación de gradiente, las redes neuronales multicapa, la regresión logística y las máquinas de vectores de soporte. Pekin y Şengül (2024), en su estudio sobre empresas de alto crecimiento, ubicaron a los bosques aleatorios en segundo lugar, detrás de los algoritmos de potenciación de gradiente y por delante de la regresión logística.

Al respecto, Houle y Macdonald (2023) incluyeron esta técnica entre los tres métodos utilizados para identificar empresas canadienses con potencial de alto crecimiento a partir de datos administrativos. En concordancia con ello, una de las contribuciones más valiosas de los bosques aleatorios al análisis del desempeño emprendedor es su capacidad para estimar la importancia de las variables.

Como cada árbol utiliza distintas variables en distintas divisiones, es posible medir cuánto contribuye cada una a mejorar la precisión de las predicciones. Esta medida permite ordenar los factores según su relevancia y orientar la atención de los emprendedores y de los responsables de políticas hacia aquellos con mayor influencia. En esa dirección, Bidgoli et al. (2024) combinaron este tipo de medidas con técnicas de interpretación basadas en valores de Shapley y encontraron que la presencia en redes profesionales y sociales, el número de empleados y el monto de la última ronda de financiamiento figuraban entre los determinantes más influyentes del éxito.

La interpretación de la importancia de las variables exige, no obstante, cautela. Las medidas de importancia indican cuánto contribuye una variable a la capacidad predictiva del modelo, pero no establecen una relación causal. Si el número de seguidores en redes sociales aparece como un predictor importante del éxito, ello no significa que aumentar artificialmente los seguidores mejore las perspectivas de la empresa: es probable que esa variable refleje otros factores, como la calidad del producto, el esfuerzo comercial o la reputación. Además, cuando varias variables están correlacionadas, la importancia puede repartirse entre ellas de maneras que dificultan la interpretación.

Por ello, los resultados de un bosque aleatorio deben complementarse con conocimiento del contexto y, cuando sea posible, con métodos de interpretación más finos, como los que se examinan más adelante en este capítulo. A ello se suma que los árboles de decisión y los bosques aleatorios tienen aplicaciones concretas en el apoyo a los emprendimientos. Un árbol sencillo puede utilizarse como herramienta de diagnóstico en programas de incubación, para clasificar a los emprendimientos según su nivel de riesgo y asignarles distintos tipos de acompañamiento y un bosque aleatorio entrenado con datos históricos de una entidad financiera puede apoyar la evaluación de solicitudes de crédito de pequeños negocios, complementando la evaluación tradicional basada en garantías.

En el contexto ecuatoriano, estas técnicas podrían aplicarse a los registros de la Superintendencia de Compañías, Valores y Seguros, que han sido utilizados en estudios econométricos sobre

supervivencia empresarial (Zambrano et al., 2021; González Cedeño, 2026). La ventaja de los bosques aleatorios frente a los modelos Probit o de Cox sería su capacidad para capturar interacciones y no linealidades sin especificarlas de antemano; su desventaja, una menor transparencia en la relación entre cada variable y el resultado.

Una estrategia razonable consiste en utilizar ambos tipos de modelos de manera complementaria: los econométricos para comprender los efectos promedio de las variables y los bosques aleatorios para mejorar la capacidad predictiva y descubrir patrones que la especificación lineal no captura. Asimismo, la aplicación de estos algoritmos requiere algunas decisiones técnicas que influyen en sus resultados. El número de árboles debe ser suficiente para estabilizar las predicciones; el número de variables consideradas en cada división controla el grado de aleatoriedad del bosque; y la profundidad máxima de los árboles o el número mínimo de observaciones por hoja regulan su complejidad.

Estos parámetros se ajustan habitualmente mediante validación cruzada, una técnica que divide los datos de entrenamiento en varias partes para estimar el desempeño del modelo en datos no utilizados durante el ajuste. Cuando las clases están desequilibradas, como ocurre con frecuencia en los estudios de éxito empresarial, conviene además ponderar las observaciones o aplicar técnicas de remuestreo para que el modelo no ignore la clase minoritaria (Park et al., 2024).

La aplicación de estos algoritmos a datos de emprendimientos latinoamericanos debe considerar algunas particularidades. Las bases de datos suelen combinar variables financieras, provenientes de registros administrativos, con variables cualitativas sobre la persona emprendedora y su negocio, recogidas mediante encuestas. Los árboles y bosques manejan bien esta mezcla de tipos de variables y son relativamente robustos ante valores atípicos, lo que los hace adecuados para este tipo de información.

Sin embargo, cuando ciertas categorías tienen pocos casos, como los emprendimientos de sectores poco frecuentes o de territorios pequeños, los árboles pueden generar divisiones poco confiables. Una práctica recomendable consiste en agrupar las categorías escasas, establecer un número mínimo de observaciones por hoja y validar la estabilidad de las variables importantes mediante remuestreo. La importancia de las variables obtenida con estos algoritmos debe interpretarse, además, como una medida de su contribución a la predicción y no como una prueba de causalidad (Molnar, 2022).

Además, una variante útil de estos algoritmos, para el estudio de la supervivencia empresarial, son los bosques aleatorios de supervivencia, que adaptan la lógica de los árboles a datos en los que interesa no solo si una empresa cierra, sino cuándo lo hace, y en los que algunas empresas todavía no han experimentado el evento al terminar el período de observación. Esta adaptación permite estimar curvas de supervivencia individuales, que muestran cómo evoluciona la probabilidad de permanecer en actividad a lo largo del tiempo para cada emprendimiento según sus características. En cuanto a los modelos de riesgos proporcionales tradicionales, los bosques de supervivencia no requieren suponer que el efecto de cada variable se mantiene constante en el tiempo, lo que puede resultar más realista en los primeros años de vida de las empresas, cuando los factores de riesgo cambian con rapidez.

Desde la perspectiva del emprendimiento predictivo, los árboles de decisión y los bosques aleatorios combinan capacidad predictiva, flexibilidad para capturar relaciones complejas y herramientas para identificar los determinantes más relevantes del desempeño. Su uso en el emprendimiento predictivo permite pasar de la intuición sobre los factores de éxito a una evidencia sistemática sobre su importancia relativa.

Figura 2.3

Algoritmos Random Forest y árboles de decisión para identificar determinantes del desempeño emprendedor



Fuente: Elaboración autores

2.4. Redes neuronales artificiales y Deep Learning para detectar patrones complejos de comportamiento empresarial

Las redes neuronales artificiales se inspiran, de manera muy simplificada, en el funcionamiento del cerebro. Están formadas por unidades de procesamiento organizadas en capas: una capa de entrada que recibe las variables, una o varias capas ocultas que transforman esa información mediante combinaciones ponderadas y funciones no lineales, y una capa de salida que produce la predicción. Durante el entrenamiento, la red ajusta los pesos de las conexiones entre unidades para reducir el error entre sus predicciones y los resultados observados. Cuando la red tiene muchas capas ocultas, se habla de aprendizaje profundo, o deep learning, una familia de métodos que ha transformado campos como el reconocimiento de imágenes, la traducción automática y el procesamiento del lenguaje natural.

En esta misma línea, James et al. (2021) presentan el aprendizaje profundo como una de las áreas de mayor crecimiento del aprendizaje estadístico y destacan sus principales variantes. Las redes densas o completamente conectadas procesan conjuntos de variables numéricas. Las redes convolucionales están diseñadas para datos con estructura espacial, como las imágenes. Las redes recurrentes y sus versiones mejoradas procesan secuencias, como series de tiempo o textos, porque conservan una memoria de los elementos anteriores. Cada arquitectura se adapta a un tipo de información, lo que abre posibilidades diversas para el análisis del comportamiento empresarial.

La principal fortaleza de las redes neuronales es su capacidad para aprender relaciones complejas y no lineales entre muchas variables, así como para procesar información no estructurada. En el análisis del emprendimiento, esto permite trabajar con fuentes de datos que los modelos tradicionales no pueden aprovechar con facilidad: las descripciones textuales de los modelos de negocio, las reseñas de los clientes, las publicaciones en redes sociales, las imágenes de los productos o las secuencias de transacciones a lo largo del tiempo. Una red recurrente podría, por ejemplo, analizar la evolución mensual de las ventas, los pagos y los saldos bancarios de un pequeño negocio para detectar patrones que anticipan dificultades financieras antes de que se reflejen en los estados contables.

De igual manera, Houle y Macdonald (2023) incluyeron las redes neuronales entre los métodos utilizados para identificar empresas canadienses con potencial de alto crecimiento, junto con la red elástica y los bosques aleatorios. Sobre este punto, Bidgoli, Vanani y Goodarzi (2024)

también evaluaron perceptrones multicapa en la predicción del éxito de empresas emergentes. En ambos casos, las redes formaron parte de comparaciones más amplias y no se impusieron de manera clara sobre los métodos basados en árboles. Este patrón es coherente con una línea de investigación que ha examinado el desempeño relativo de las redes neuronales en datos tabulares, es decir, en datos organizados en filas y columnas como los que predominan en la información financiera y empresarial.

Grinsztajn, Oyallon y Varoquaux (2022), a partir de un análisis comparativo con 45 conjuntos de datos, concluyeron que los modelos basados en árboles, como los bosques aleatorios y los de potenciación de gradiente, seguían siendo los de mejor desempeño en datos tabulares de tamaño mediano, incluso sin considerar su menor costo de entrenamiento.

Los autores atribuyeron esta ventaja a diferencias en los sesgos inductivos de ambos tipos de modelos y señalaron tres desafíos para las redes neuronales en este tipo de datos: la robustez frente a variables no informativas, la preservación de la orientación de los datos y la capacidad para aprender funciones irregulares. En este sentido, Shwartz-Ziv y Armon (2022) llegaron a conclusiones similares al comparar modelos de aprendizaje profundo propuestos para datos tabulares con algoritmos de potenciación de gradiente, y encontraron que estos últimos mantenían, en general, un mejor desempeño y requerían menos ajuste. Por otra parte, la evidencia aplicada al ámbito empresarial apunta en la misma dirección.

Al respecto, Zizi et al. (2021), al comparar redes neuronales con regresión logística para predecir dificultades financieras en pequeñas y medianas empresas de Marruecos, obtuvieron los mejores resultados con la regresión logística, que alcanzó una exactitud de 95 % un año antes del evento. Este hallazgo sugiere que, cuando los datos son limitados y las variables corresponden a razones financieras bien seleccionadas, la complejidad de las redes neuronales no se traduce necesariamente en mejores predicciones y puede, en cambio, aumentar el riesgo de sobreajuste.

Estas comparaciones no significan que las redes neuronales carezcan de utilidad para el emprendimiento predictivo. Su ventaja se manifiesta principalmente cuando se trabaja con grandes volúmenes de datos no estructurados, cuando las relaciones entre variables son especialmente complejas o cuando se dispone de secuencias temporales extensas. El análisis de textos —descripciones de proyectos, documentos de planes de negocio, comentarios de clientes— es uno de los ámbitos en los que el aprendizaje profundo ofrece capacidades difícilmente alcanzables

con otros métodos, como se examina en el tercer capítulo a propósito del procesamiento del lenguaje natural.

También resulta prometedor en la detección de patrones en transacciones financieras de alta frecuencia y en la combinación de múltiples tipos de información en un solo modelo. No obstante, la aplicación de redes neuronales plantea, sin embargo, desafíos prácticos considerables. El primero es la necesidad de datos abundantes: estos modelos tienen muchos parámetros y requieren grandes cantidades de observaciones para aprender sin sobreajustarse. El segundo es el costo computacional, tanto en el entrenamiento como, en algunos casos, en la operación.

El tercero es la opacidad: las redes profundas funcionan como cajas negras cuyos mecanismos internos son difíciles de interpretar, lo que complica su uso en decisiones que afectan a personas y que exigen explicación, como la concesión de crédito. El cuarto es la sensibilidad a las decisiones de diseño, como el número de capas, el número de unidades, las funciones de activación o la tasa de aprendizaje, que deben ajustarse cuidadosamente.

Frente a estos desafíos, la práctica recomendada en el emprendimiento predictivo consiste en considerar las redes neuronales como una opción dentro de un conjunto más amplio de técnicas, y no como la solución por defecto. Para problemas con datos tabulares de tamaño moderado, como los que suelen tener las entidades de apoyo al emprendimiento o los registros administrativos, conviene comenzar con modelos más sencillos —regresión logística, árboles, bosques aleatorios, potenciación de gradiente— y recurrir a las redes neuronales cuando existan razones específicas, como la disponibilidad de datos no estructurados o de secuencias temporales extensas.

En todos los casos, la comparación con modelos de base y la evaluación con datos independientes son indispensables. Paralelamente, en el contexto latinoamericano, el aprendizaje profundo podría encontrar aplicaciones relevantes en el análisis de la información generada por las plataformas digitales que utilizan los pequeños negocios: billeteras móviles, sistemas de pago, comercio electrónico y redes sociales. Estas fuentes producen secuencias de datos que reflejan con gran detalle la actividad cotidiana de las empresas y que podrían utilizarse para evaluar su desempeño cuando no existen estados financieros formales. Sin embargo, su aprovechamiento exige resolver cuestiones de acceso a los datos, de protección de la información personal y de transparencia en el uso de los modelos, en el marco de normas como la Ley Orgánica de Protección de Datos Personales (Asamblea Nacional del Ecuador, 2021).

El uso de redes neuronales en la predicción empresarial plantea también exigencias de infraestructura y de competencias que no deben subestimarse. Su entrenamiento requiere ajustar numerosos hiperparámetros, como el número de capas, el número de neuronas, la tasa de aprendizaje o las técnicas de regularización, y su desempeño puede variar considerablemente según estas decisiones. Para un equipo pequeño, el tiempo dedicado a este ajuste puede ser mayor que el beneficio obtenido, especialmente cuando los datos son tabulares y de tamaño moderado.

En cambio, cuando el emprendimiento dispone de datos no estructurados —imágenes de productos, textos de reseñas, registros de navegación—, las redes profundas ofrecen capacidades que los modelos tradicionales no tienen. La recomendación práctica es seleccionar la herramienta según el tipo de datos y la pregunta, comparando sistemáticamente su desempeño con alternativas más sencillas, tal como sugiere la evidencia comparativa sobre datos tabulares (Grinsztajn et al., 2022).

Puede sostenerse, así, que las redes neuronales y el aprendizaje profundo amplían el alcance del emprendimiento predictivo hacia fuentes de información y patrones que otras técnicas no pueden capturar. Su potencial es mayor en los problemas que involucran datos no estructurados o secuenciales, mientras que en los datos tabulares convencionales la evidencia sugiere que los modelos basados en árboles suelen ofrecer resultados iguales o mejores con menor costo y mayor facilidad de interpretación.

Figura 2.4

Redes neuronales artificiales y Deep Learning para detectar patrones complejos de comportamiento empresarial.



Fuente: Elaboración autores

2.5. Modelos de Gradient Boosting y XGBoost para la clasificación predictiva de emprendimientos

La potenciación de gradiente, conocida en la literatura internacional como gradient boosting, es una técnica de aprendizaje automático que construye un modelo predictivo de manera secuencial. A diferencia de los bosques aleatorios, que entrenan muchos árboles de forma independiente y luego promedian sus resultados, la potenciación de gradiente entrena los árboles uno después de otro, de modo que cada nuevo árbol se concentra en corregir los errores cometidos por los anteriores.

El resultado es un conjunto de árboles pequeños que, sumados, producen predicciones muy precisas. Esta lógica de aprendizaje progresivo, en la que cada paso mejora lo que el anterior no logró captar, ha convertido a estos algoritmos en una de las herramientas más eficaces para problemas de clasificación con datos tabulares. Sobre esta base, James et al. (2021) explican que la potenciación funciona ajustando árboles de decisión a los residuos del modelo vigente, es decir, a la diferencia entre lo observado y lo predicho, y agregando cada nuevo árbol con un peso reducido que controla la velocidad de aprendizaje.

Esta tasa de aprendizaje es un parámetro clave: valores pequeños requieren más árboles pero suelen producir modelos con mejor capacidad de generalización. Otros parámetros importantes son el número total de árboles, que debe limitarse para evitar el sobreajuste, y la profundidad de cada árbol, que determina el nivel de interacción entre variables que el modelo puede capturar.

XGBoost es una implementación de la potenciación de gradiente que se popularizó por su eficiencia computacional y por incorporar mecanismos de regularización que reducen el riesgo de sobreajuste. Junto con otras implementaciones similares, se ha convertido en una referencia en competencias de ciencia de datos y en aplicaciones empresariales. Su popularidad no es casual: la evidencia comparativa muestra que los métodos de potenciación de gradiente suelen ubicarse entre los de mejor desempeño en datos tabulares.

En esa dirección, Grinsztajn, Oyallon y Varoquaux (2022) encontraron que los modelos basados en árboles, entre ellos XGBoost, seguían superando a las redes neuronales en conjuntos de datos tabulares de tamaño mediano, y Shwartz-Ziv y Armon (2022) concluyeron que XGBoost mantenía, en general, un desempeño superior al de los modelos de aprendizaje profundo diseñados

para este tipo de datos, con menores necesidades de ajuste. Desde esta perspectiva, en la predicción del desempeño empresarial, los resultados son consistentes con este panorama.

Pekin y Şengül (2024) compararon siete algoritmos para identificar empresas de alto crecimiento en una base de más de un millón trescientas mil empresas, y encontraron que XGBoost ofrecía el mejor desempeño, seguido por los bosques aleatorios y la regresión logística; el área bajo la curva ROC alcanzó valores de hasta 0,91 en la definición de crecimiento basada en la facturación.

Sobre este punto, Żbikowski y Antosiuk (2021), en su modelo libre de sesgo de anticipación para predecir el éxito de empresas emergentes con datos de Crunchbase, también obtuvieron resultados prometedores con un clasificador de potenciación de gradiente. En este sentido, Park, Choi y Feng (2024), por su parte, observaron que XGBoost era el algoritmo con mejor desempeño tanto con los datos originales, desequilibrados, como después de equilibrarlos mediante redes generativas adversariales, con un promedio de alrededor de 96,5 % en sus métricas de evaluación tras el equilibrio.

Estos resultados no deben llevar a concluir que XGBoost sea siempre la mejor opción. Al respecto, Bidgoli, Vanani y Goodarzi (2024) obtuvieron un desempeño ligeramente superior con los bosques aleatorios que con la potenciación de gradiente en su estudio sobre empresas emergentes, con exactitudes de 82 % y 80 %, respectivamente. Las diferencias entre los métodos basados en árboles suelen ser pequeñas, y dependen de las características de los datos, de la definición del resultado y del cuidado con que se ajustan los parámetros. Lo que la evidencia sí sugiere con bastante consistencia es que, en problemas de clasificación con datos financieros y empresariales estructurados, los métodos de conjuntos de árboles constituyen un punto de partida sólido.

De forma complementaria, la aplicación de la potenciación de gradiente a la clasificación de emprendimientos ofrece varias ventajas prácticas. Estos algoritmos manejan variables de distinta naturaleza sin requerir transformaciones complejas; incorporan mecanismos para tratar valores faltantes, frecuentes en la información de pequeñas empresas; capturan relaciones no lineales e interacciones entre variables; y permiten ponderar las observaciones para compensar el desequilibrio entre clases. Además, producen medidas de importancia de las variables y son compatibles con métodos de interpretación como los basados en valores de Shapley, que permiten explicar cada predicción individual, aspecto que se analiza más adelante.

Sus limitaciones también deben considerarse. Una de ellas es el riesgo de sobreajuste cuando se utilizan demasiados árboles, árboles demasiado profundos o tasas de aprendizaje elevadas, especialmente con conjuntos de datos pequeños. Otra es la necesidad de ajustar varios parámetros que interactúan entre sí, lo que exige procedimientos sistemáticos de validación cruzada.

A ello se suma una interpretabilidad menor que la de los modelos lineales o los árboles individuales, que debe compensarse con herramientas de explicación. Y la cuarta, compartida con todos los métodos basados en datos históricos, es la dificultad para anticipar situaciones que no tienen precedentes en los datos de entrenamiento, como una crisis económica de características inéditas. Vinculado con lo anterior, para un proyecto de clasificación predictiva de emprendimientos, una secuencia de trabajo razonable podría organizarse de la siguiente manera. Se comienza definiendo con precisión el resultado que se quiere predecir y el horizonte temporal, por ejemplo, el cierre de la empresa en los tres años siguientes.

Se construye luego un conjunto de variables disponibles en el momento de la predicción, evitando el sesgo de anticipación. Se estima un modelo de base, como una regresión logística, que servirá de referencia. Se entrenan después modelos de conjuntos de árboles, ajustando sus parámetros mediante validación cruzada y tratando el desequilibrio de clases. Se evalúan todos los modelos con un conjunto de datos independiente y con métricas adecuadas. Y, finalmente, se interpreta el modelo seleccionado para verificar que sus predicciones respondan a patrones razonables y no a artefactos de los datos.

En el contexto de los ecosistemas emprendedores latinoamericanos, la potenciación de gradiente podría aplicarse en diversos ámbitos. Las entidades de microfinanzas y las cooperativas de ahorro y crédito podrían utilizarla para mejorar la evaluación del riesgo de sus clientes emprendedores, incorporando información sobre la actividad del negocio además de las garantías.

Los programas públicos de apoyo podrían emplearla para identificar a los emprendimientos con mayor probabilidad de crecimiento y concentrar en ellos ciertos instrumentos, sin descuidar el apoyo a los más vulnerables. Y las plataformas de comercio electrónico o de pagos podrían ofrecer a sus usuarios diagnósticos basados en estos modelos para alertarlos sobre riesgos y oportunidades. En concordancia con ello, cada una de estas aplicaciones plantea, sin embargo, exigencias de gobernanza.

Un modelo de clasificación que influye en el acceso al crédito o a programas públicos debe ser evaluado en cuanto a su precisión, pero también en cuanto a su equidad, para evitar que

reproduzca desventajas históricas de determinados grupos o territorios. Debe ser explicable para las personas afectadas y estar sujeto a supervisión humana, en línea con los marcos normativos que se examinan en el quinto capítulo. La eficacia técnica de XGBoost y de los métodos de potenciación de gradiente es una condición necesaria, pero no suficiente, para su uso responsable.

La popularidad de XGBoost y de sus variantes ha favorecido la disponibilidad de implementaciones eficientes y de documentación abundante, lo que reduce las barreras de entrada para equipos con formación intermedia en ciencia de datos. Esta accesibilidad, sin embargo, puede conducir a utilizar el algoritmo con su configuración por defecto, sin una búsqueda cuidadosa de hiperparámetros ni una validación adecuada. Un procedimiento recomendable incluye la división de los datos en conjuntos de entrenamiento, validación y prueba; la búsqueda de hiperparámetros mediante validación cruzada; la detención temprana del entrenamiento cuando el error de validación deja de disminuir; y la evaluación final con datos no utilizados durante el ajuste.

En problemas de predicción del fracaso empresarial, donde las clases suelen estar desequilibradas, conviene además ponderar los casos o ajustar el umbral de decisión, y evaluar el desempeño con métricas que no se dejen engañar por la clase mayoritaria (Chicco y Jurman, 2023). A ello se suma que la comparación entre las distintas implementaciones de la potenciación de gradiente muestra diferencias de velocidad y de manejo de variables categóricas, pero su desempeño predictivo suele ser similar cuando se ajustan con cuidado.

Para el emprendimiento predictivo, la elección entre ellas puede guiarse por criterios prácticos: la disponibilidad de documentación en el lenguaje de programación que domina el equipo, la facilidad para integrar el modelo en los sistemas existentes y la compatibilidad con las herramientas de explicabilidad. Más importante que la implementación concreta es la disciplina en el proceso de modelado: definir con claridad la variable que se predice, preparar los datos con rigor, validar el modelo con procedimientos adecuados y documentar todas las decisiones tomadas, de modo que los resultados puedan reproducirse y revisarse.

Figura 2.5

Modelos de Gradient Boosting y XGBoost para la clasificación predictiva de emprendimientos



Fuente: Elaboración autores

2.6. Inteligencia artificial explicable (XAI): interpretación de resultados mediante SHAP y LIME

A medida que los modelos de aprendizaje automático ganan capacidad predictiva, suelen perder transparencia. Una regresión logística permite leer directamente cómo influye cada variable en la probabilidad estimada; un bosque aleatorio con quinientos árboles o un modelo de potenciación de gradiente con miles de divisiones no ofrece esa lectura inmediata. Esta tensión entre precisión e interpretabilidad tiene consecuencias prácticas en el emprendimiento predictivo. Una persona emprendedora a la que un sistema le niega un crédito necesita saber por qué; un analista que recomienda invertir en una empresa emergente debe poder justificar su decisión ante un comité; y un programa público que selecciona beneficiarios con ayuda de un algoritmo está obligado a rendir cuentas sobre sus criterios.

La inteligencia artificial explicable, conocida por sus siglas en inglés XAI, reúne los métodos que buscan resolver esta tensión. Asimismo, Dwivedi et al. (2023) describen la inteligencia artificial explicable como el conjunto de métodos de aprendizaje automático que ayudan a las personas a comprender y a confiar de manera apropiada en los modelos, y que facilitan la construcción de modelos más explicables. Los autores sostienen que la capacidad de explicar un

modelo se ha convertido en una expectativa clave para generar confianza y desplegar la inteligencia artificial en ámbitos de alto impacto, proponen una taxonomía para clasificar los distintos enfoques y los vinculan con las etapas del desarrollo de un sistema de aprendizaje automático. Esta perspectiva subraya que la explicabilidad no es un añadido final, sino una consideración que debe estar presente desde el diseño.

Los métodos de explicación pueden clasificarse según varios criterios. Una primera distinción separa los modelos intrínsecamente interpretables, como la regresión lineal o los árboles de decisión pequeños, de los métodos de explicación posterior, que se aplican a modelos complejos ya entrenados para comprender su funcionamiento. Una segunda distinción diferencia las explicaciones globales, que describen el comportamiento general del modelo —por ejemplo, qué variables son más importantes en promedio—, de las explicaciones locales, que explican una predicción concreta, como la probabilidad de cierre estimada para una empresa en particular. Una tercera separa los métodos específicos de un tipo de modelo de los métodos agnósticos, aplicables a cualquier algoritmo.

Sobre este punto, Molnar (2022), en una obra de referencia sobre el aprendizaje automático interpretable, desarrolla en detalle estas categorías y los métodos más utilizados en cada una. Además, entre los métodos agnósticos de explicación local, dos se han convertido en estándares de la práctica: SHAP y LIME. SHAP, acrónimo de SHapley Additive exPlanations, se basa en un concepto de la teoría de juegos cooperativos, los valores de Shapley, que asignan a cada jugador una parte del resultado colectivo según su contribución promedio en todas las coaliciones posibles.

Aplicado a la predicción, cada variable se trata como un jugador, y su valor de Shapley indica cuánto contribuyó a que la predicción para un caso específico se aleje del promedio. Si un modelo estima que una empresa tiene un 40 % de probabilidad de cierre, frente a un promedio de 15 %, SHAP permite descomponer esa diferencia y mostrar, por ejemplo, que el alto endeudamiento añadió quince puntos, la baja rentabilidad diez y la antigüedad redujo cinco (Molnar, 2022).

Las explicaciones basadas en SHAP tienen propiedades atractivas: son consistentes, suman exactamente la diferencia entre la predicción y el promedio, y pueden agregarse para obtener explicaciones globales. Por ello se han difundido ampliamente en aplicaciones financieras. En este sentido, Bussmann et al. (2021) aplicaron valores de Shapley a un modelo de riesgo crediticio para 15.000 pequeñas y medianas empresas que solicitaron financiamiento en plataformas

de préstamos entre pares, y mostraron que, al agrupar a los prestatarios según la similitud de sus explicaciones, era posible identificar conjuntos de características financieras comunes tanto entre los riesgosos como entre los no riesgosos; esas agrupaciones ayudaban a explicar las calificaciones crediticias y a anticipar el comportamiento futuro.

Al respecto, Bidgoli, Vanani y Goodarzi (2024) utilizaron también valores de Shapley para identificar los factores más influyentes en la predicción del éxito de empresas emergentes, entre los que destacaron la presencia en redes profesionales y sociales y el monto de la última ronda de financiamiento. En esta misma línea, LIME, acrónimo de Local Interpretable Model-agnostic Explanations, sigue una lógica distinta. Para explicar una predicción concreta, genera variaciones artificiales del caso analizado, observa cómo cambian las predicciones del modelo complejo y ajusta un modelo sencillo, generalmente lineal, que aproxima ese comportamiento en la vecindad del caso. Los coeficientes de este modelo local se interpretan como la explicación de la predicción. LIME es flexible y aplicable a datos tabulares, textos e imágenes, pero sus explicaciones pueden variar según la forma en que se generan las variaciones y la definición de la vecindad, lo que exige cautela en su interpretación (Molnar, 2022).

La elección entre estos métodos depende del propósito. SHAP ofrece una base teórica más sólida y explicaciones consistentes, aunque su cálculo puede ser costoso en modelos complejos con muchas variables; existen, sin embargo, versiones optimizadas para modelos basados en árboles que lo hacen eficiente. LIME es más rápido y fácil de aplicar a distintos tipos de datos, pero sus resultados son menos estables.

En la práctica, muchos equipos utilizan ambos métodos de manera complementaria y contrastan sus resultados con el conocimiento del dominio para verificar que las explicaciones tengan sentido. De igual manera, la explicabilidad tiene, además, una dimensión normativa creciente. El reglamento europeo de inteligencia artificial clasifica como de alto riesgo los sistemas destinados a evaluar la solvencia crediticia de personas físicas y establece para ellos exigencias de transparencia, documentación y supervisión humana (Parlamento Europeo y Consejo de la Unión Europea, 2024).

En el ámbito ecuatoriano, la norma de la Superintendencia de Protección de Datos Personales sobre el uso de sistemas de inteligencia artificial reconoce el derecho de las personas a ser informadas sobre el carácter automatizado de un proceso y a solicitar una explicación sobre la lógica del algoritmo cuando exista un impacto directo, además de exigir revisión humana para

las decisiones que produzcan efectos jurídicos o afecten significativamente sus derechos (Superintendencia de Protección de Datos Personales [SPDP], 2026). Cuando un modelo predictivo influye en el acceso de un emprendedor al financiamiento, estas exigencias convierten la explicabilidad en una obligación y no solo en una buena práctica.

Para el emprendimiento predictivo, la inteligencia artificial explicable cumple al menos cuatro funciones. Una de ellas es generar confianza en los usuarios de los modelos, al mostrar que las predicciones responden a factores razonables. Otra es orientar la acción, porque una explicación que identifica los factores que elevan el riesgo indica también dónde intervenir para reducirlo. Una tercera función consiste en detectar errores y sesgos, ya que las explicaciones pueden revelar que el modelo se apoya en variables espurias o discriminatorias. Y la cuarta es cumplir con las exigencias legales de transparencia y de derecho a la explicación. Conviene advertir, sin embargo, sobre los límites de las explicaciones.

Los métodos como SHAP y LIME describen cómo se comporta el modelo, no cómo funciona la realidad: si el modelo ha aprendido una correlación espuria, la explicación reproducirá esa correlación sin advertir que carece de fundamento causal. Además, las explicaciones pueden ser malinterpretadas por personas sin formación técnica, que podrían atribuirles una certeza que no tienen. Por ello, las explicaciones deben comunicarse con un lenguaje claro, acompañadas de advertencias sobre su alcance, y complementarse con el juicio de personas que conozcan el contexto de cada empresa.

La explicabilidad tiene, además, un valor pedagógico para las personas emprendedoras. Cuando un modelo indica que la probabilidad de supervivencia de un negocio es baja y explica que ello se debe principalmente a su elevado endeudamiento y a la concentración de sus ventas en pocos clientes, ofrece una lección de gestión que la persona puede aplicar de inmediato. Las explicaciones contrafactuales, que muestran qué cambios en las variables modificarían la predicción, resultan especialmente útiles en este sentido: indican, por ejemplo, cuánto debería reducirse el endeudamiento para que el negocio pasara a un nivel de riesgo moderado.

Estas explicaciones deben presentarse con cautela, porque señalan asociaciones aprendidas por el modelo y no garantizan que el cambio sugerido produzca efectivamente el resultado esperado. Aun así, combinadas con el acompañamiento de asesores, pueden convertir los modelos predictivos en herramientas de aprendizaje y no solo de evaluación (Dwivedi et al., 2023). Por otra parte, la explicabilidad debe adaptarse, por otra parte, a su destinatario. Un analista de datos necesita explicaciones detalladas que le permitan depurar el modelo y detectar errores; un

responsable de crédito, explicaciones que le ayuden a justificar sus decisiones ante sus superiores y ante los clientes; una persona emprendedora, explicaciones sencillas que le indiquen qué aspectos de su negocio requieren atención; y una autoridad supervisora, información que le permita verificar que el sistema cumple la normativa. Una misma predicción puede, por tanto, requerir varias explicaciones distintas, presentadas con distinto nivel de detalle y en distinto lenguaje. Diseñar estas explicaciones es una tarea que combina competencias técnicas, de comunicación y de conocimiento del dominio, y que conviene planificar desde el inicio del desarrollo de los modelos.

Figura 2.6

Inteligencia artificial explicable (XAI): interpretación de resultados mediante SHAP y LIME



Fuente: Elaboración autores

2.7. Validación de modelos predictivos: precisión, sensibilidad, especificidad, curva ROC y matriz de confusión

Un modelo predictivo solo es útil si sus estimaciones resultan confiables cuando se aplican a casos nuevos. Esta afirmación, aparentemente obvia, encierra uno de los desafíos más importantes del emprendimiento predictivo. Con datos suficientes y algoritmos flexibles, es relativamente fácil construir un modelo que reproduzca casi a la perfección los resultados de las empresas utilizadas para entrenarlo; lo difícil es que ese modelo acierte con empresas que no ha visto.

La validación es el conjunto de procedimientos y métricas que permiten estimar esa capacidad de generalización y elegir, entre varios modelos, el que mejor sirve al propósito para el que se construye. Paralelamente, el primer principio de la validación es separar los datos utilizados

para entrenar el modelo de los datos utilizados para evaluarlo. En esa dirección, James et al. (2021) explican que el error de un modelo en sus propios datos de entrenamiento tiende a subestimar sistemáticamente su error en datos nuevos, porque el algoritmo puede haber aprendido particularidades irrepetibles de esa muestra.

Por ello, la práctica habitual consiste en reservar un conjunto de prueba que el modelo no ve durante el entrenamiento, o en aplicar la validación cruzada, que divide los datos en varias partes, entrena el modelo con todas menos una y lo evalúa con la restante, repitiendo el proceso hasta que cada parte haya servido como conjunto de prueba. En problemas con dimensión temporal, como la predicción del cierre de empresas, conviene además respetar el orden del tiempo: entrenar con datos de años anteriores y evaluar con datos de años posteriores, para simular las condiciones reales de uso.

En un problema de clasificación binaria, como distinguir entre empresas que cierran y empresas que sobreviven, esta matriz organiza los resultados en cuatro casillas. Los verdaderos positivos son las empresas que el modelo predijo que cerrarían y efectivamente cerraron; los verdaderos negativos, las que predijo que sobrevivirían y sobrevivieron; los falsos positivos, las que predijo que cerrarían pero sobrevivieron; y los falsos negativos, las que predijo que sobrevivirían, pero cerraron. A partir de estas cuatro cifras se calculan las métricas de desempeño más utilizadas. Sobre esta base, la exactitud es la proporción de predicciones correctas sobre el total.

Aunque es la métrica más intuitiva, puede resultar engañosa cuando las clases están desequilibradas. Si solo el 5 % de las empresas cierra, un modelo que prediga que ninguna cerrará tendrá una exactitud del 95 %, pese a ser completamente inútil para identificar el riesgo.

La sensibilidad, también llamada exhaustividad, mide la proporción de casos positivos que el modelo identifica correctamente: de todas las empresas que cerraron, cuántas detectó. La especificidad mide la proporción de casos negativos correctamente identificados: de todas las empresas que sobrevivieron, cuántas clasificó como tales. La precisión, por su parte, indica qué proporción de las predicciones positivas resultó correcta: de todas las empresas señaladas como en riesgo, cuántas cerraron efectivamente. La puntuación F1 combina precisión y sensibilidad en una sola medida.

La curva ROC ofrece una perspectiva complementaria. Muchos modelos no producen directamente una clasificación, sino una probabilidad; la clasificación surge al establecer un umbral, por ejemplo, considerar en riesgo a toda empresa con probabilidad de cierre superior al 50 %. La curva ROC representa la sensibilidad frente a la tasa de falsos positivos para todos los

umbrales posibles, y el área bajo esa curva resume la capacidad del modelo para ordenar los casos: un valor de 0,5 equivale a una clasificación al azar y un valor de 1 a una discriminación perfecta.

Esta medida se ha convertido en un estándar en la evaluación de modelos de riesgo. Pekin y Şengül (2024), por ejemplo, informaron áreas bajo la curva entre 0,53 y 0,91 para distintos algoritmos de predicción de empresas de alto crecimiento. Sin embargo, el área bajo la curva ROC tiene limitaciones que la literatura reciente ha subrayado. Sobre este punto, Chicco y Jurman (2023) sostienen que esta medida puede premiar modelos con sensibilidad y especificidad débiles, produciendo resultados excesivamente optimistas, y que no informa sobre la precisión ni sobre el valor predictivo negativo.

Los autores proponen sustituirla por el coeficiente de correlación de Matthews, que solo alcanza valores altos cuando las cuatro tasas básicas —sensibilidad, especificidad, precisión y valor predictivo negativo— son simultáneamente altas. Esta métrica, que considera las cuatro casillas de la matriz de confusión, resulta especialmente adecuada para problemas con clases desequilibradas, como la predicción del cierre o del éxito excepcional de empresas. No es casual que Pekin y Şengül (2024) la incluyeran entre sus criterios de evaluación junto con el área bajo la curva, la exactitud, la precisión, la sensibilidad y la puntuación F1.

La elección de la métrica más adecuada depende de los costos asociados a cada tipo de error. En la evaluación del riesgo crediticio de un emprendimiento, un falso negativo —otorgar crédito a una empresa que luego no puede pagarlo— implica una pérdida para la entidad financiera, mientras que un falso positivo —negar crédito a una empresa que habría pagado— implica una oportunidad perdida para la entidad y, sobre todo, para el emprendedor. En un programa público de alerta temprana destinado a ofrecer acompañamiento a empresas en riesgo, los falsos positivos pueden ser relativamente baratos, porque el apoyo adicional a una empresa que no lo necesitaba tiene un costo moderado, mientras que los falsos negativos, que dejan sin apoyo a empresas que cerrarán, resultan más costosos.

Ajustar el umbral de clasificación según estos costos es una decisión de gestión que no puede delegarse por completo en el algoritmo. Desde esta perspectiva, la calibración es otro aspecto relevante de la validación. Un modelo está bien calibrado cuando sus probabilidades estimadas corresponden a las frecuencias observadas: si asigna un 20 % de probabilidad de cierre a un grupo de empresas, alrededor del 20 % de ellas debería cerrar efectivamente.

Un modelo puede ordenar correctamente a las empresas según su riesgo, con una buena área bajo la curva, y estar sin embargo mal calibrado, sobrestimando o subestimando sistemáticamente las probabilidades. Cuando las predicciones se utilizan para fijar tasas de interés, calcular provisiones o asignar recursos, la calibración resulta tan importante como la capacidad de discriminación.

Un modelo entrenado con datos de un período de crecimiento económico puede deteriorarse cuando las condiciones cambian; por ello, conviene monitorear su desempeño de manera continua y reentrenarlo cuando sea necesario. Evaluar las métricas por separado para distintos grupos permite detectar estas diferencias y, en su caso, identificar sesgos que afecten la equidad de las decisiones, tema que se aborda en el quinto capítulo.

En la práctica, los estudios ecuatorianos sobre supervivencia empresarial ilustran la importancia de interpretar las métricas con prudencia. En este sentido, Cueva et al. (2025) informaron una exactitud global de clasificación de 95,8 % en su modelo logístico para microempresas de Machala; sin embargo, el tamaño reducido de la muestra y su selección no probabilística aconsejan no extrapolar ese resultado. En general, cifras de exactitud muy elevadas obtenidas con muestras pequeñas o sin validación en datos independientes deben tomarse como indicios y no como garantías del desempeño en la práctica.

La validación debe considerar, por último, la estabilidad de los modelos a lo largo del tiempo por lo que un modelo entrenado con datos de un período de crecimiento económico puede perder precisión cuando las condiciones cambian, como ocurrió con muchos modelos de riesgo durante la pandemia; la validación fuera del tiempo, que entrena el modelo con datos de años anteriores y lo evalúa con datos posteriores, ofreciendo una estimación más realista de su desempeño futuro.

En cuanto, al seguimiento continuo de las métricas una vez que el modelo está en uso permite detectar su deterioro y activar su recalibración para los emprendimientos y para las entidades que los apoyan, esta vigilancia no requiere herramientas complejas: basta con registrar las predicciones realizadas, comparar periódicamente con los resultados observados y calcular las métricas de discriminación y calibración en ventanas sucesivas.

Figura 2.7

Validación de modelos predictivos: precisión, sensibilidad, especificidad, curva ROC y matriz de confusión



Fuente: Elaboración autores

2.8. Construcción de un índice predictivo multidimensional de éxito empresarial mediante técnicas de inteligencia artificial

Los modelos examinados en este capítulo producen, por lo general, una estimación sobre un único resultado: la probabilidad de que una empresa cierre, crezca o alcance un hito determinado. Sin embargo, el éxito empresarial es un fenómeno multidimensional. Una empresa puede ser rentable pero frágil ante cambios del mercado; puede crecer rápidamente, pero con un endeudamiento insostenible; puede sobrevivir durante años sin generar empleo de calidad ni capacidad de innovación.

Un índice predictivo multidimensional busca integrar estas facetas en una medida sintética que permita evaluar el potencial de un emprendimiento de manera más completa, sin perder la posibilidad de examinar cada una de sus dimensiones por separado. Vinculado con lo anterior, la construcción de indicadores compuestos cuenta con una metodología consolidada.

El Centro Común de Investigación de la Comisión Europea propone una guía de diez pasos que comienza con la definición de un marco teórico y continúa con la selección de los datos, la imputación de valores faltantes, el análisis multivariante, la normalización, la ponderación, la agregación, el análisis de sensibilidad, la vinculación con otras medidas y la visualización de los resultados (Centro Común de Investigación [JRC], s. f.).

Esta secuencia, desarrollada originalmente para índices de política pública, ofrece un marco riguroso que puede adaptarse a la construcción de un índice de éxito empresarial y enriquecerse con técnicas de inteligencia artificial en varias de sus etapas.

El marco teórico constituye el punto de partida. Con base en lo desarrollado en el primer capítulo, un índice de éxito empresarial podría organizarse en torno a cinco dimensiones. La viabilidad financiera recoge la capacidad del negocio para generar ingresos suficientes, cubrir sus costos y mantener una estructura de deuda sostenible. La tracción comercial refleja la evolución de las ventas, la base de clientes y la aceptación del producto en el mercado.

La resiliencia captura la capacidad de resistir y adaptarse frente a crisis, mediante la diversificación, las reservas y la flexibilidad. La capacidad organizacional y de innovación incluye las competencias del equipo, la adopción de tecnologías y la capacidad de aprendizaje. Y la sostenibilidad abarca la calidad del empleo, el impacto ambiental y la relación con la comunidad. Cada dimensión debe estar respaldada por la literatura y por la evidencia empírica sobre los determinantes del desempeño.

En concordancia con ello, la selección de variables es el segundo paso. Para cada dimensión se identifican indicadores medibles, disponibles en el momento de la evaluación y relevantes para el tipo de empresas que se quiere analizar. Los estudios ecuatorianos sobre supervivencia ofrecen orientación en la dimensión financiera: la rentabilidad sobre los activos, el volumen de ventas, el apalancamiento financiero, el patrimonio y la rotación de cartera han mostrado relación con el riesgo de cierre (Zambrano et al., 2021; González Cedeño, 2026).

En la dimensión de tracción, pueden utilizarse la tasa de crecimiento de las ventas, el número de clientes recurrentes y, cuando existan, indicadores de presencia digital, que la literatura ha identificado como predictores relevantes del éxito de empresas emergentes (Bidgoli et al., 2024). Las técnicas de aprendizaje automático, en particular las medidas de importancia de las variables de los bosques aleatorios y de los modelos de potenciación de gradiente, pueden ayudar a seleccionar los indicadores con mayor capacidad predictiva dentro de cada dimensión.

La imputación de valores faltantes es especialmente relevante en el caso de los emprendimientos, cuya información suele ser incompleta. Además de los métodos estadísticos tradicionales, existen técnicas basadas en aprendizaje automático que estiman los valores faltantes a partir de las relaciones observadas entre las variables. Su aplicación debe documentarse cuidadosamente, porque las imputaciones pueden influir en los resultados del índice.

El análisis multivariante, por su parte, permite examinar la estructura de los datos, identificar redundancias entre indicadores y verificar que las dimensiones definidas teóricamente se correspondan con patrones observables; las técnicas de agrupamiento y de reducción de dimensionalidad resultan útiles en esta etapa. A ello se suma que la normalización transforma los indicadores, medidos en unidades distintas, a una escala común que permita combinarlos. La ponderación determina la importancia relativa de cada indicador y de cada dimensión.

Aquí la inteligencia artificial ofrece una alternativa a los métodos tradicionales: en lugar de asignar los pesos por juicio de expertos o de forma igualitaria, pueden derivarse de un modelo predictivo entrenado para anticipar un resultado observable, como la supervivencia a cinco años o el crecimiento de las ventas. Los valores de Shapley promedio de cada variable, por ejemplo, ofrecen una medida de su contribución a la predicción que puede utilizarse como base para la ponderación (Molnar, 2022). Este enfoque tiene la ventaja de anclar el índice en la evidencia empírica, aunque exige cautela para no reducir el éxito empresarial exclusivamente a aquello que el modelo predice.

La agregación combina los indicadores ponderados en un valor por dimensión y las dimensiones en un valor global. La elección del método de agregación tiene implicaciones conceptuales: una suma ponderada permite que un buen desempeño en una dimensión compense uno malo en otra, mientras que una media geométrica penaliza los desequilibrios. Para un índice de éxito empresarial, una agregación que limite la compensación puede resultar más adecuada, porque una empresa muy rentable pero sin ninguna resiliencia no debería obtener la misma calificación que una con desempeño equilibrado.

Asimismo, el análisis de sensibilidad evalúa cuánto cambian los resultados cuando se modifican las decisiones metodológicas, como la normalización, los pesos o el método de agregación. Un índice robusto debe producir clasificaciones relativamente estables ante cambios razonables en estas decisiones. La vinculación con otras medidas consiste en verificar que el índice se relacione con resultados observables: si es un índice predictivo, las empresas con mejores puntuaciones deberían mostrar, en el seguimiento posterior, mayores tasas de supervivencia y crecimiento.

Esta validación externa puede realizarse con las métricas descritas en este capítulo, como el área bajo la curva ROC o el coeficiente de correlación de Matthews (Chicco y Jurman, 2023).

Por último, la visualización debe comunicar el resultado global junto con el perfil por dimensiones, para que los usuarios puedan identificar fortalezas y debilidades específicas.

Un índice predictivo multidimensional construido de esta manera tiene aplicaciones diversas. Para los emprendedores, ofrece un diagnóstico integral que señala en qué dimensiones conviene concentrar los esfuerzos de mejora. Para las entidades financieras, complementa la evaluación crediticia con información sobre el potencial y la resiliencia del negocio. Para las incubadoras y aceleradoras, permite seguir la evolución de las empresas acompañadas y evaluar el impacto de sus programas. Y para los responsables de políticas, facilita la identificación de sectores y territorios con mayor potencial o con mayores riesgos, así como el seguimiento de la evolución del tejido empresarial.

Con todo, su construcción exige, no obstante, atender a varios riesgos. El primero es la reducción de una realidad compleja a un solo número, que puede ocultar información relevante si no se presenta junto con sus componentes. El segundo es el riesgo de sesgo: si los indicadores o los pesos favorecen sistemáticamente a determinados tipos de empresas, el índice puede reproducir desigualdades existentes. El tercero es la manipulación: cuando un índice se utiliza para asignar recursos, las empresas pueden intentar mejorar artificialmente los indicadores sin modificar su desempeño real. Y el cuarto es la obsolescencia, porque los factores que predicen el éxito pueden cambiar con el tiempo, lo que obliga a revisar periódicamente el índice.

La construcción de un índice de este tipo debe prever también su comunicación a los distintos usuarios. Un número único facilita la comparación y el seguimiento, pero oculta la diversidad de situaciones que pueden conducir al mismo resultado: dos emprendimientos con igual puntuación pueden tener fortalezas y debilidades muy distintas. Por ello, el índice debe presentarse siempre junto con sus dimensiones y con la explicación de los factores que más lo afectan.

La guía metodológica del Centro Común de Investigación recomienda, en este sentido, acompañar los indicadores compuestos con la información desagregada que permita entender su origen y con análisis de sensibilidad que muestren cómo cambiarían los resultados bajo distintos supuestos de ponderación (JRC, s. f.). Esta transparencia no solo mejora la comprensión, sino que reduce el riesgo de que el índice se utilice de manera mecánica para decisiones que requieren un examen más detallado de cada caso. Además, el índice puede cumplir, además, una función de seguimiento en el tiempo.

Calculado periódicamente para un mismo emprendimiento, permite observar si su situación mejora o se deteriora, y relacionar esos cambios con las decisiones tomadas o con los

acontecimientos del entorno. Para los programas de apoyo, la evolución del índice de los emprendimientos acompañados ofrece una medida del efecto de sus intervenciones, más rica que el simple recuento de negocios que permanecen abiertos. Para que esta función de seguimiento sea válida, la metodología del índice debe mantenerse estable entre mediciones o, cuando se modifique, recalcularse retrospectivamente para preservar la comparabilidad de la serie.

Al término de este recorrido, conviene retener que el índice predictivo multidimensional de éxito empresarial representa una síntesis de los métodos examinados en este capítulo. Combina la rigurosidad metodológica de los indicadores compuestos con la capacidad de las técnicas de inteligencia artificial para seleccionar variables, ponderarlas y validar su poder predictivo. Bien construido, ofrece una herramienta de diagnóstico y de decisión que integra las distintas facetas del desempeño emprendedor.

Figura 2.8

Construcción de un índice predictivo multidimensional de éxito empresarial mediante técnicas de inteligencia artificial.



Fuente: Elaboración autores

EDITORIAL ANDES COGNITIO

CAPÍTULO III

ECONOMÍA PREDICTIVA DEL EMPRENDIMIENTO: ANTICIPACIÓN DE MERCADOS , CLIENTES Y OPORTUNIDADES DE NEGOCIO



CAPÍTULO III

ECONOMÍA PREDICTIVA DEL EMPRENDIMIENTO: ANTICIPACIÓN DE MERCADOS, CLIENTES Y OPORTUNIDADES DE NEGOCIO

3.1. Inteligencia económica predictiva y detección anticipada de oportunidades emprendedoras

Toda empresa nace de una oportunidad: una necesidad no atendida, un cambio en los hábitos de consumo, una tecnología que abarata un proceso o una regulación que abre un mercado. La literatura sobre emprendimiento ha discutido durante décadas si estas oportunidades se descubren, porque existen en el entorno a la espera de que alguien las identifique, o si se crean a partir de la acción y la imaginación de las personas emprendedoras. Cualquiera que sea la posición que se adopte, hay un consenso en que la capacidad de reconocer oportunidades antes que otros constituyen una ventaja decisiva.

La inteligencia económica predictiva busca fortalecer esa capacidad mediante el uso sistemático de datos y modelos que anticipen tendencias de mercado, cambios en la demanda y nichos emergentes. En esta misma línea, la inteligencia económica puede entenderse como el proceso organizado de recolección, análisis y difusión de información sobre el entorno económico, competitivo, tecnológico y regulatorio, con el propósito de apoyar la toma de decisiones estratégicas. Tradicionalmente, esta función se apoyaba en fuentes como estudios de mercado, estadísticas oficiales, informes sectoriales y redes de contactos.

La transformación digital ha multiplicado las fuentes disponibles: búsquedas en internet, publicaciones en redes sociales, reseñas de consumidores, precios en plataformas de comercio electrónico, registros de patentes y marcas, ofertas de empleo, flujos de pagos digitales. La inteligencia artificial permite procesar estas fuentes a gran escala y extraer de ellas señales tempranas de cambio.

La revisión de Giuggioli y Pellegrini (2023) sobre la inteligencia artificial como habilitadora del emprendimiento identifica precisamente la oportunidad como uno de sus ámbitos de aplicación: la combinación de inteligencia artificial y grandes volúmenes de datos puede facilitar la identificación de nuevos productos, servicios y modelos de negocio. Este potencial se

manifiesta en varias formas. Los modelos de series temporales pueden detectar categorías de productos cuya demanda crece de manera sostenida. El análisis de búsquedas en internet puede revelar intereses emergentes antes de que se traduzcan en compras.

El procesamiento del lenguaje natural aplicado a reseñas y conversaciones puede identificar necesidades insatisfechas. Y el análisis de redes puede mostrar cómo se difunden las innovaciones entre grupos de consumidores. En la práctica, los datos de búsquedas en internet ilustran bien las posibilidades y los límites de estas fuentes. Al respecto, Bantis, Clements y Urquhart (2023) evaluaron si los datos de búsquedas de Google mejoraban la estimación del crecimiento del producto interno bruto en los Estados Unidos y en Brasil, y encontraron que los modelos que combinaban indicadores económicos tradicionales con datos de categorías de búsqueda superaban a los que no los incluían, con beneficios similares en ambos países; sin embargo, el aporte dependía de la manera en que se seleccionaban las variables, y la utilización de datos de búsqueda más detallados no mejoraba los resultados.

Para el emprendimiento, estos hallazgos sugieren que la información de búsquedas puede anticipar cambios en la actividad económica, pero que su valor depende de un procesamiento metodológico cuidadoso y de su combinación con otras fuentes.

La detección anticipada de oportunidades puede organizarse como un ciclo de inteligencia con varias etapas. En primer lugar, destaca la definición de las preguntas estratégicas: qué mercados, segmentos o tecnologías interesan a la persona emprendedora y qué decisiones dependen de la información. En segundo término, sobresale la identificación de fuentes pertinentes, combinando datos públicos, privados y propios. A ello se suma la recolección y el procesamiento, que incluye la depuración, la integración y, cuando corresponde, la anonimización de los datos.

Asimismo, resulta relevante el análisis, mediante técnicas descriptivas, predictivas y de minería de textos que permitan identificar tendencias, anomalías y patrones emergentes. Por último, cabe mencionar la interpretación, en la que las personas con conocimiento del mercado evalúan las señales detectadas y las convierten en hipótesis de oportunidad. Y la sexta es la validación, en la que esas hipótesis se contrastan con clientes reales, en la línea del aprendizaje validado que propone el enfoque de la empresa emergente ajustada (Shepherd y Gruber, 2021). De igual manera, esta última etapa merece énfasis.

La inteligencia económica predictiva no sustituye el contacto directo con los clientes ni la experimentación; los orienta. Una señal detectada en los datos —por ejemplo, un aumento sostenido de las búsquedas de productos alimenticios sin gluten en una ciudad— constituye una

hipótesis que debe verificarse: ¿quiénes buscan esos productos?, ¿cuánto están dispuestos a pagar?, ¿dónde los compran actualmente?, ¿qué insatisfacciones expresan? Las respuestas a estas preguntas requieren entrevistas, pruebas de concepto y productos mínimos viables. La combinación entre el análisis de datos a gran escala y la validación cualitativa en pequeña escala permite reducir el riesgo de invertir en oportunidades que solo existen en las cifras.

En el contexto ecuatoriano, la inteligencia económica predictiva podría contribuir a modificar un rasgo estructural del emprendimiento nacional. El informe del Global Entrepreneurship Monitor para el Ecuador muestra que alrededor de tres cuartas partes de los emprendedores en etapa temprana se concentran en actividades orientadas al consumidor, principalmente el comercio minorista, mientras que los servicios a empresas representan apenas el 3,7 %; además, solo el 1 % de los negocios en etapa temprana obtiene un cuarto o más de sus ingresos del exterior, la tasa de internacionalización más baja entre las economías comparadas (Lasio et al., 2024).

Esta concentración en actividades de baja diferenciación y en el mercado interno sugiere que muchas personas emprenden en los sectores más visibles, con alta competencia y márgenes estrechos, sin identificar oportunidades en nichos menos saturados. Por otra parte, herramientas de inteligencia económica accesibles podrían ayudar a cambiar esa dinámica. Un observatorio que analice la evolución de la demanda por sectores y territorios, las importaciones de productos que podrían sustituirse con producción local, las búsquedas de servicios no disponibles en determinadas ciudades o las tendencias de consumo en mercados de exportación podría orientar a los emprendedores hacia oportunidades con mayor potencial.

Estas herramientas cobran especial relevancia en un contexto en el que la mayoría emprende por necesidad: si quien emprende para ganarse la vida dispone de información sobre los sectores con menor saturación y mejores perspectivas, puede elegir con mayor fundamento dónde invertir sus escasos recursos.

La detección anticipada de oportunidades enfrenta, sin embargo, varios riesgos. El primero es el de las tendencias pasajeras, que pueden confundirse con cambios estructurales de la demanda. El segundo es el de los sesgos de las fuentes digitales, que reflejan el comportamiento de las personas conectadas y pueden subrepresentar a grupos de menores ingresos, de mayor edad o residentes en zonas rurales. El tercero es el de la convergencia: si muchos emprendedores utilizan las mismas herramientas y detectan las mismas señales, la oportunidad puede saturarse

rápidamente. Y el cuarto es el de la sobreinterpretación, que lleva a atribuir significado a correlaciones casuales.

Mitigar estos riesgos exige combinar fuentes, contrastar las señales con conocimiento del mercado y validar las hipótesis antes de comprometer recursos significativos. Paralelamente, la dimensión ética también está presente. La recolección de datos sobre el comportamiento de los consumidores debe respetar la normativa de protección de datos personales, que en el Ecuador establece obligaciones para todos los responsables del tratamiento (Asamblea Nacional del Ecuador, 2021).

El uso de datos agregados y anonimizados, de fuentes públicas o de información obtenida con el consentimiento de las personas reduce los riesgos jurídicos y reputacionales. La inteligencia económica legítima se distingue del espionaje o de la explotación indebida de la información precisamente por el respeto de estos límites.

En la práctica, un sistema de inteligencia económica para emprendedores puede comenzar de forma modesta. Una persona que planea abrir un negocio de alimentos saludables, por ejemplo, puede combinar la información del registro estadístico de empresas sobre el número de establecimientos similares en su ciudad, la evolución de las búsquedas en internet relacionadas con este tipo de productos, las reseñas de los negocios existentes y la información sobre la población y el ingreso del sector donde desea ubicarse. Si bien estas fuentes no son concluyentes por sí solas, su combinación ofrece una visión más completa de la oportunidad.

El valor de la inteligencia económica reside precisamente en esa integración y en la regularidad con que se actualiza: una vigilancia periódica permite detectar cambios en la competencia o en las preferencias antes de que afecten las ventas. Las fuentes oficiales, como las estadísticas del Instituto Nacional de Estadística y Censos (INEC, 2024), ofrecen la base sobre la cual se construyen estos análisis. Sobre esta base, la inteligencia económica predictiva exige también una disciplina de verificación.

Las señales que se captan en internet y en las redes sociales pueden ser ruidosas, manipuladas o poco representativas, y una tendencia aparente puede responder a un fenómeno pasajero o a la actividad de un grupo reducido de usuarios. Por ello, conviene contrastar las señales provenientes de distintas fuentes, comprobar su persistencia en el tiempo y, cuando la decisión lo justifique, verificarlas mediante experimentos de bajo costo, como la oferta preliminar de un producto o la consulta directa a clientes potenciales.

Esta combinación entre la vigilancia de datos y la validación en el terreno reduce el riesgo de invertir en oportunidades ilusorias.

Figura 3.1

Construcción de un índice predictivo multidimensional de éxito empresarial mediante técnicas de inteligencia artificial.



Fuente: Elaboración autores

3.2. Modelos de pronóstico de demanda mediante series temporales, ARIMA y aprendizaje automático

Pocas decisiones empresariales dependen tanto de una predicción como las relacionadas con la demanda. Cuánto producir, cuánto inventario mantener, cuántas personas contratar, cuándo lanzar una promoción o cuánto financiamiento solicitar son decisiones que descansan, explícita o implícitamente, en una estimación de cuánto venderá la empresa en el futuro. Un pronóstico demasiado optimista conduce a inventarios excesivos, costos financieros innecesarios y, en el caso de productos perecederos, pérdidas directas; uno demasiado pesimista provoca quiebres de inventario, clientes insatisfechos y oportunidades perdidas.

Para los emprendimientos, que operan con márgenes estrechos y reservas limitadas, la calidad del pronóstico de demanda puede marcar la diferencia entre la supervivencia y el cierre. Desde esta perspectiva, Hyndman y Athanasopoulos (2021), en una obra de referencia sobre la disciplina, definen el pronóstico como la predicción del futuro de la manera más precisa posible a partir de toda la información disponible, incluidos los datos históricos y el conocimiento de los eventos futuros que pueden afectar los resultados.

Los autores subrayan que la facilidad para pronosticar un fenómeno depende de varios factores: cuánto se comprenden los elementos que lo determinan, cuántos datos se dispone, cuán similar

será el futuro al pasado y si los propios pronósticos pueden influir en lo que se intenta predecir. La demanda de un producto básico con ventas estables durante años puede pronosticarse con bastante precisión; la de un producto nuevo, sin historia, presenta desafíos mucho mayores.

Los métodos de series temporales parten de la idea de que el comportamiento pasado de una variable contiene información sobre su comportamiento futuro. Una serie temporal de ventas suele combinar varios componentes: una tendencia de largo plazo, que refleja el crecimiento o la contracción del negocio; una estacionalidad, que recoge patrones que se repiten en períodos fijos, como el aumento de las ventas en diciembre o los fines de semana; ciclos de duración variable, asociados a la situación económica; y una variación irregular, que no sigue patrones identificables.

Descomponer la serie en estos elementos constituye, con frecuencia, el primer paso del análisis (Hyndman y Athanasopoulos, 2021). De forma complementaria, entre los métodos clásicos, dos familias han dominado la práctica durante décadas. El suavizamiento exponencial produce pronósticos como promedios ponderados de las observaciones pasadas, en los que los pesos disminuyen a medida que las observaciones se alejan en el tiempo; sus variantes permiten incorporar tendencia y estacionalidad.

Los modelos ARIMA, acrónimo de autorregresivos integrados de medias móviles, describen la serie a partir de sus propios valores pasados y de los errores de pronóstico anteriores, previa transformación para que sus propiedades estadísticas sean estables en el tiempo. Ambos enfoques son relativamente sencillos de aplicar, están disponibles en programas estadísticos de uso común y ofrecen, junto con el pronóstico puntual, intervalos que expresan la incertidumbre de la predicción (Hyndman y Athanasopoulos, 2021).

El aprendizaje automático ha ampliado considerablemente las posibilidades del pronóstico de demanda. Las competencias internacionales de pronóstico, conocidas como competencias M, ofrecen una evidencia valiosa sobre el desempeño comparado de los distintos enfoques. En esa dirección, Makridakis, Spiliotis y Assimakopoulos (2022) analizaron los resultados de la competencia M5, que consistió en pronosticar 42.840 series temporales jerárquicas de ventas unitarias de una cadena minorista, y constataron que, por primera vez, todos los métodos de mejor desempeño fueron de aprendizaje automático puro y superaron claramente a los métodos estadísticos de referencia.

La mayoría de los métodos destacados utilizó algoritmos de potenciación de gradiente, y el equipo ganador mejoró en 22,4 % la precisión del mejor método de referencia en la medida

global de error, aunque las ganancias fueron menores en los niveles más desagregados. Vinculado con lo anterior, las lecciones de esta competencia son especialmente útiles para el emprendimiento predictivo. Una de ellas es el valor del aprendizaje transversal: entrenar un solo modelo con muchas series relacionadas, en lugar de un modelo para cada producto, produjo mejores resultados, porque permite aprovechar patrones comunes entre productos.

Otra es la importancia de las variables externas, como precios, promociones y eventos especiales, que mejoraron la precisión de manera significativa. A ello se suma el beneficio de combinar varios modelos, ya que los promedios de distintos pronósticos resultaron más precisos que los modelos individuales. Y la cuarta es la relevancia de una estrategia de validación sólida, puesto que muchos equipos no lograron seleccionar su mejor propuesta por basarse en evaluaciones engañosas (Makridakis et al., 2022).

Para un emprendimiento, estas lecciones pueden traducirse en recomendaciones prácticas. Cuando el negocio vende muchos productos con historial limitado, conviene utilizar modelos que aprendan de todos los productos simultáneamente, en lugar de intentar pronosticar cada uno por separado. Resulta útil incorporar en los modelos la información sobre precios, promociones, días festivos, condiciones climáticas o eventos locales que puedan influir en las ventas.

Combinar pronósticos de métodos distintos —por ejemplo, un suavizamiento exponencial y un modelo de potenciación de gradiente— suele reducir el error. Y es indispensable evaluar los pronósticos con datos que el modelo no utilizó para su ajuste, respetando el orden temporal de las observaciones. Ahora bien, el pronóstico de la demanda de productos nuevos representa un desafío particular, porque no existe historia de ventas sobre la cual construir el modelo.

En estos casos, pueden utilizarse varias estrategias complementarias: tomar como referencia la trayectoria de productos similares lanzados anteriormente, por la propia empresa o por otras; utilizar señales tempranas, como las búsquedas en internet, las preinscripciones o las interacciones en redes sociales; aplicar modelos de difusión que describen cómo se adoptan las innovaciones en una población; o recurrir al juicio estructurado de personas expertas. A medida que el producto acumula historia de ventas, los modelos pueden recalibrarse y ganar precisión.

La incertidumbre debe formar parte de todo pronóstico. Un pronóstico puntual —por ejemplo, que se venderán 500 unidades el próximo mes— resulta mucho menos útil que uno que indique un rango probable, como entre 420 y 590 unidades con un 80 % de probabilidad.

Los intervalos de pronóstico permiten tomar decisiones que consideran el riesgo: si el costo de quedarse sin inventario es alto, convendrá planificar para el extremo superior del rango; si el costo de un inventario excedente es mayor, para el extremo inferior. Sobre este punto, Hyndman y Athanasopoulos (2021) insisten en que comunicar la incertidumbre es tan importante como el pronóstico mismo. No obstante, en América Latina, el pronóstico de demanda enfrenta algunas dificultades específicas. Muchas pequeñas empresas no registran sus ventas de manera sistemática, lo que impide construir series temporales confiables.

La volatilidad macroeconómica, las variaciones de precios y los choques externos dificultan la extrapolación de patrones pasados. La informalidad de una parte del mercado limita la disponibilidad de datos sectoriales. Frente a estas limitaciones, el primer paso consiste en establecer registros diarios o semanales de ventas por producto, aunque sea en una hoja de cálculo; a partir de pocos meses de información, es posible aplicar métodos sencillos que ya mejoran las decisiones respecto de la intuición pura.

La elección del método de pronóstico debe considerar también los recursos disponibles para mantenerlo. Un modelo de aprendizaje automático que requiere reentrenarse cada semana con datos externos puede ofrecer pronósticos algo más precisos, pero su mantenimiento exige competencias y tiempo que muchos emprendimientos no tienen. Un método de suavizamiento exponencial, en cambio, puede implementarse en una hoja de cálculo y actualizarse automáticamente con cada nueva observación.

La evaluación de los pronósticos debe comparar siempre los métodos complejos con referencias sencillas, como el valor del mismo período del año anterior o el promedio de las últimas semanas: si un modelo sofisticado no supera claramente a estas referencias, su complejidad no se justifica (Hyndman y Athanasopoulos, 2021). Para los emprendimientos que comienzan, una estrategia razonable consiste en empezar con métodos sencillos, medir su error de manera sistemática y avanzar hacia técnicas más complejas solo cuando los datos acumulados y el beneficio esperado lo justifiquen.

En concordancia con ello, el pronóstico de demanda tiene, además, consecuencias directas sobre la sostenibilidad. Un emprendimiento que produce o compra más de lo que vende genera desperdicio, inmoviliza recursos y, en el caso de los alimentos o de los productos perecibles, contribuye a pérdidas que tienen un costo ambiental. Uno que produce menos de lo que podría vender pierde ingresos y deteriora la relación con sus clientes.

Mejorar la precisión de los pronósticos permite reducir ambos tipos de error y ajustar la producción y las compras a la demanda efectiva. Esta relación entre la anticipación y el uso eficiente de los recursos ilustra cómo el emprendimiento predictivo puede alinearse con los objetivos de sostenibilidad que forman parte de su arquitectura.

La evidencia de las competencias internacionales muestra que los métodos de aprendizaje automático pueden superar a los enfoques estadísticos tradicionales cuando se dispone de muchas series relacionadas y de información externa, pero también que los métodos clásicos siguen siendo valiosos como referencia y como componentes de combinaciones.

Figura 3.2

Modelos de pronóstico de demanda mediante series temporales, ARIMA y aprendizaje automático



Fuente: Elaboración autores

3.3. Segmentación inteligente de consumidores mediante algoritmos de clustering y aprendizaje no supervisado

No todos los clientes de un emprendimiento son iguales. Algunos compran con frecuencia y en pequeñas cantidades; otros, de manera esporádica, pero en volúmenes elevados. Algunos son sensibles al precio; otros valoran la calidad, la rapidez o la atención personalizada. Algunos llegan por recomendación; otros, a través de redes sociales o de plataformas digitales. Reconocer esta diversidad y organizarla en grupos con características comunes —segmentos— permite diseñar ofertas, mensajes y canales adecuados a cada uno. La segmentación de mercados es una práctica clásica del mercadeo, pero el aprendizaje automático la ha transformado al permitir descubrir grupos a partir de los datos, sin necesidad de definirlos de antemano.

A ello se suma que las técnicas de agrupamiento, conocidas como clustering, pertenecen al aprendizaje no supervisado. A diferencia de los modelos de clasificación, que aprenden a predecir un resultado conocido, el agrupamiento busca estructuras en datos que no contienen etiquetas: su objetivo es organizar las observaciones en grupos de modo que los elementos de un mismo grupo sean lo más parecidos posible entre sí y lo más distintos posible de los elementos de otros grupos (James et al., 2021). Aplicado a los consumidores, esto significa encontrar conjuntos de clientes con patrones de compra, preferencias o características similares, que pueden ser invisibles a simple vista.

El algoritmo de k-medias es el más utilizado en la segmentación de clientes. Su funcionamiento es sencillo: el analista define el número de grupos que desea obtener; el algoritmo ubica inicialmente ese número de centros en el espacio de las variables, asigna cada cliente al centro más cercano, recalcula los centros como el promedio de los clientes asignados y repite el proceso hasta que las asignaciones se estabilizan.

El agrupamiento jerárquico, por su parte, construye una estructura en forma de árbol que muestra cómo se fusionan progresivamente los clientes en grupos cada vez mayores, lo que permite elegir el número de segmentos a posteriori. Existen también métodos basados en densidad, como DBSCAN, que identifican grupos de forma irregular y detectan observaciones atípicas que no pertenecen a ningún grupo (James et al., 2021). Asimismo, la literatura reciente confirma la centralidad de estas técnicas.

En este sentido, Salminen et al. (2023), en una revisión sistemática de 172 artículos sobre segmentación algorítmica de clientes, identificaron 46 algoritmos distintos y constataron que k-medias era el más empleado. La revisión encontró, además, que los estudios utilizaban 14 métricas de evaluación, que rara vez recurrían a especialistas del dominio para valorar los resultados, que más de cuatro de cada cinco utilizaban el número de segmentos como único parámetro ajustado y que, típicamente, generaban alrededor de cuatro segmentos. Estos hallazgos sugieren que la práctica de la segmentación algorítmica todavía tiene margen de mejora, especialmente en la validación de los segmentos con criterios de negocio.

La segmentación también se ha aplicado al análisis de las propias empresas. Al respecto, Bidgoli, Vanani y Goodarzi (2024) compararon el agrupamiento jerárquico, k-medias y DBSCAN para describir perfiles de empresas emergentes, y obtuvieron los mejores resultados con k-medias, con un coeficiente de silueta de 72 %, una medida de la calidad de la separación entre grupos. Este tipo de análisis puede resultar útil para incubadoras, aceleradoras o programas de

apoyo que necesitan comprender la diversidad de los emprendimientos que atienden y diseñar intervenciones diferenciadas.

Además, en la práctica de los emprendimientos, uno de los enfoques más difundidos combina el agrupamiento con el análisis de recencia, frecuencia y valor monetario. Para cada cliente se calcula cuánto tiempo ha pasado desde su última compra, cuántas veces ha comprado en un período y cuánto ha gastado en total. Estos tres indicadores, sencillos de obtener a partir de los registros de ventas, capturan aspectos esenciales del comportamiento de compra y permiten identificar, mediante agrupamiento, segmentos como los clientes más valiosos y fieles, los clientes nuevos con potencial, los clientes que compraban con frecuencia pero han dejado de hacerlo y los clientes ocasionales de bajo valor. Cada segmento sugiere una acción distinta: fidelizar a los primeros, desarrollar a los segundos, recuperar a los terceros y gestionar con eficiencia a los últimos.

La calidad de una segmentación no se mide únicamente por criterios estadísticos. Un buen segmento debe ser identificable, es decir, sus miembros deben poder reconocerse; suficientemente grande como para justificar una estrategia específica; accesible mediante canales concretos; estable en el tiempo; y, sobre todo, accionable, en el sentido de que debe sugerir decisiones diferenciadas. En esa dirección, Salminen et al. (2023) advierten precisamente sobre la escasa participación de especialistas del dominio en la evaluación de los segmentos: un agrupamiento estadísticamente impecable puede carecer de utilidad si no se traduce en estrategias comerciales viables.

Por ello, el análisis algorítmico debe complementarse con el conocimiento del negocio y, cuando sea posible, con investigación cualitativa sobre los clientes de cada segmento. En esta misma línea, la aplicación del agrupamiento exige varias decisiones técnicas que influyen en los resultados. Las variables deben seleccionarse con criterio, porque el algoritmo agrupará a los clientes según las dimensiones que se le proporcionen. Las variables medidas en escalas distintas deben normalizarse, para evitar que las de mayor magnitud dominen el resultado.

El número de segmentos debe elegirse considerando tanto indicadores estadísticos como la utilidad práctica; un número excesivo de segmentos dificulta la gestión, mientras que uno insuficiente oculta diferencias relevantes. Y los resultados deben verificarse con distintas inicializaciones o métodos, porque algunos algoritmos pueden producir agrupamientos diferentes según su punto de partida.

Para los emprendimientos latinoamericanos, la segmentación inteligente ofrece oportunidades concretas. Un pequeño comercio que registra sus ventas con un sistema de punto de venta o una tienda en línea que conserva el historial de compras de sus clientes dispone ya de la información necesaria para un análisis de recencia, frecuencia y valor monetario. Plataformas de comercio electrónico, de mensajería empresarial y de gestión de clientes incorporan cada vez más funciones de segmentación automatizada. El desafío principal no es tecnológico, sino de disciplina en el registro de los datos y de capacidad para traducir los segmentos en acciones comerciales.

De igual manera, la segmentación plantea también consideraciones éticas. Agrupar a los clientes según su comportamiento puede derivar en prácticas discriminatorias si se utilizan variables sensibles o si se ofrecen condiciones desfavorables a determinados grupos. La personalización de precios, por ejemplo, puede resultar legítima cuando refleja diferencias en costos o en servicios, pero problemática cuando explota la vulnerabilidad de ciertos consumidores. El tratamiento de los datos utilizados para la segmentación debe respetar, además, la normativa de protección de datos personales, informando a los clientes sobre el uso de su información (Asamblea Nacional del Ecuador, 2021).

La segmentación tiene también una dimensión territorial que resulta especialmente relevante en los países latinoamericanos, caracterizados por fuertes diferencias entre regiones, entre zonas urbanas y rurales, y entre barrios de una misma ciudad. Un emprendimiento que vende por canales digitales puede descubrir, al segmentar a sus clientes, que una parte importante de sus compradores proviene de ciudades intermedias donde la oferta es escasa, lo que podría orientar decisiones de logística o de apertura de puntos de venta.

A la inversa, un negocio físico puede identificar que sus clientes más valiosos residen en zonas alejadas y que su fidelidad depende de la facilidad de acceso. La incorporación de variables geográficas en los algoritmos de agrupamiento, combinada con la información de las estadísticas oficiales sobre población e ingresos, permite construir segmentos que reflejen esta diversidad (Salminen et al., 2023). El uso de estos datos debe respetar, en todo caso, las normas de protección de la información personal de los clientes.

Hay razones, por consiguiente, para afirmar que la segmentación inteligente de consumidores permite que los emprendimientos comprendan la diversidad de sus clientes con un nivel de detalle antes reservado a las grandes empresas. Las técnicas de agrupamiento, combinadas con indicadores sencillos de comportamiento de compra y con el conocimiento del negocio, ofrecen una base para diseñar ofertas más pertinentes, asignar mejor los recursos comerciales y anticipar

los cambios en el comportamiento de cada grupo de clientes. Su valor final depende, sin embargo, de que los segmentos identificados se traduzcan en decisiones concretas y respetuosas de los derechos de las personas.

Figura 3.3

Segmentación inteligente de consumidores mediante algoritmos de clustering y aprendizaje no supervisado



Fuente: Elaboración autores

3.4. Procesamiento de lenguaje natural (NLP) para identificar necesidades insatisfechas y nuevas tendencias de consumo

Una parte considerable de la información sobre lo que los consumidores necesitan, valoran o rechazan se expresa en palabras. Las reseñas en plataformas de comercio electrónico, los comentarios en redes sociales, las preguntas en foros, las conversaciones con servicios de atención al cliente, las búsquedas en internet y las respuestas abiertas de encuestas contienen descripciones espontáneas de experiencias, frustraciones y deseos.

Durante mucho tiempo, esta información resultó difícil de aprovechar a gran escala, porque su análisis requería que personas leyeran y clasificaran manualmente miles de textos. El procesamiento del lenguaje natural, conocido por sus siglas en inglés NLP, ha cambiado esta situación al permitir que los sistemas informáticos analicen textos de manera automatizada. Por otra

parte, el procesamiento del lenguaje natural reúne un conjunto de técnicas que permiten a las computadoras interpretar, clasificar y generar lenguaje humano.

Sus aplicaciones van desde tareas relativamente sencillas, como contar la frecuencia de palabras o identificar términos clave, hasta tareas complejas, como comprender el significado de una oración, detectar el tono emocional de un comentario o resumir un documento extenso. Los avances recientes del aprendizaje profundo, y en particular de los modelos de aprendizaje por transferencia preentrenados con grandes volúmenes de texto, han mejorado de manera notable la capacidad de estos sistemas para capturar el contexto y los matices del lenguaje, hasta superar con claridad a los métodos basados en diccionarios en tareas como la clasificación del sentimiento (Hartmann et al., 2023).

Para el emprendimiento predictivo, una de las aplicaciones más valiosas del procesamiento del lenguaje natural es la identificación de necesidades insatisfechas. Sobre este punto, Spada et al. (2023) desarrollaron un enfoque para descubrir necesidades de los usuarios a partir de reseñas en línea y lo aplicaron a más de trece mil reseñas de productos de software de dos pequeñas empresas italianas.

Mediante diccionarios especializados de grupos de usuarios, ocupaciones y atributos de calidad de los sistemas, identificaron qué tipos de usuarios mencionaban los productos y qué necesidades expresaban con mayor frecuencia. Los resultados mostraron, por ejemplo, que la simplicidad de uso, la seguridad, la accesibilidad y el uso móvil figuraban entre las necesidades más importantes, y que algunos grupos de usuarios valoraban especialmente la flexibilidad y la personalización.

Los gerentes de ambas empresas utilizaron estos hallazgos para elegir sus segmentos objetivo y para orientar decisiones sobre sus modelos de negocio. En la práctica, este estudio ilustra tanto el potencial como las limitaciones de las técnicas basadas en diccionarios. Su principal ventaja es la precisión: cuando un término del diccionario aparece en un texto, la identificación es confiable. Su principal limitación es la cobertura: los resultados dependen de la exhaustividad de los diccionarios, por lo que pueden omitir necesidades expresadas con palabras no previstas. Los autores sugieren combinar este enfoque con métodos basados en reglas o en aprendizaje automático para ampliar la capacidad de detección (Spada et al., 2023).

Las técnicas de procesamiento del lenguaje natural aplicables a la detección de necesidades y tendencias son diversas. El modelado de temas permite descubrir, en grandes colecciones de textos, los asuntos que aparecen de manera recurrente, sin necesidad de definirlos de antemano;

aplicado a las reseñas de una categoría de productos, puede revelar los aspectos que más preocupan a los consumidores.

La extracción de aspectos identifica las características específicas de un producto o servicio que se mencionan en los comentarios —precio, calidad, entrega, atención— y las asocia con valoraciones positivas o negativas. La clasificación de textos permite organizar automáticamente los comentarios según categorías definidas, como quejas, sugerencias o preguntas. Y el análisis de la evolución temporal de estos elementos permite detectar tendencias emergentes, como el aumento de las menciones a un atributo que antes no se consideraba relevante. Paralelamente, la identificación de necesidades insatisfechas resulta especialmente valiosa en las etapas tempranas del emprendimiento.

Una persona que evalúa la posibilidad de crear un negocio en un sector puede analizar las reseñas de los productos y servicios existentes para identificar las quejas más frecuentes; esas quejas señalan oportunidades de diferenciación. Si los clientes de los servicios de reparto de una ciudad se quejan reiteradamente de los tiempos de entrega, una nueva empresa podría construir su propuesta de valor en torno a la rapidez. Si los usuarios de aplicaciones de gestión financiera para pequeños negocios expresan dificultades con la complejidad de las herramientas, existe una oportunidad para soluciones más simples. El procesamiento del lenguaje natural permite realizar este tipo de análisis a una escala y con una sistematicidad imposibles de alcanzar manualmente.

Las tendencias de consumo pueden detectarse de manera similar. El análisis de publicaciones en redes sociales, de artículos de blogs especializados o de búsquedas en internet puede revelar el surgimiento de nuevos intereses, valores o prácticas de consumo, como la preocupación por la sostenibilidad, la preferencia por productos locales o la adopción de nuevas dietas. Detectar estas tendencias en sus etapas iniciales permite a los emprendedores posicionarse antes de que el mercado se sature.

Sin embargo, la interpretación de estas señales requiere cautela, porque las conversaciones en línea no siempre representan el comportamiento de la población general y pueden estar influidas por campañas organizadas o por grupos particularmente activos. Con todo, el procesamiento del lenguaje natural presenta desafíos específicos en el contexto latinoamericano. Muchos de los modelos y recursos más avanzados se desarrollaron inicialmente para el inglés, y su

desempeño en español puede ser inferior, especialmente con las variantes regionales, el lenguaje coloquial, los modismos y la ironía frecuentes en las redes sociales.

En el Ecuador, como en otros países de la región, el uso de expresiones locales, la mezcla de lenguas y las particularidades del lenguaje escrito en canales informales dificultan el análisis automatizado. Superar estas limitaciones exige adaptar los modelos con datos locales, construir diccionarios que incorporen el vocabulario regional y validar los resultados con personas que conozcan el contexto.

La utilización de textos generados por los consumidores plantea también cuestiones éticas y jurídicas. Aunque muchas reseñas y publicaciones son públicas, su recolección y análisis masivo puede afectar la privacidad de las personas, especialmente cuando se combinan con otros datos que permiten identificarlas. La normativa de protección de datos personales exige considerar la finalidad del tratamiento, la proporcionalidad y la protección de la información (Asamblea Nacional del Ecuador, 2021).

El análisis de textos agregados y anonimizados, orientado a identificar patrones generales y no a perfilar a individuos, reduce estos riesgos y resulta suficiente para la mayoría de los propósitos de inteligencia de mercado. Ahora bien, el procesamiento del lenguaje natural en español presenta desafíos propios. Aunque los modelos multilingües han mejorado considerablemente, el desempeño en español sigue siendo, en muchas tareas, inferior al obtenido en inglés, y las variantes regionales, el uso de expresiones coloquiales, las abreviaturas propias de las redes sociales y la mezcla de idiomas pueden dificultar el análisis.

Estas limitaciones aconsejan validar los resultados de los modelos con muestras revisadas por personas que conozcan el contexto, y ajustar los modelos con ejemplos locales cuando sea posible. En este sentido, Hartmann et al. (2023) mostraron que el ajuste de los modelos de lenguaje con datos del dominio de aplicación mejora sustancialmente su precisión, lo que sugiere el valor de construir conjuntos de entrenamiento con textos de consumidores latinoamericanos.

Figura 3.4

Procesamiento de lenguaje natural (NLP) para identificar necesidades insatisfechas y nuevas tendencias de consumo



Fuente: Elaboración autores

3.5. Predicción del comportamiento del consumidor y estimación del valor del cliente a largo plazo

Atraer a un nuevo cliente suele resultar costoso: requiere inversión en publicidad, promociones, tiempo de venta y, con frecuencia, descuentos iniciales. Para que esa inversión se justifique, el cliente debe generar, a lo largo de su relación con la empresa, ingresos que superen su costo de adquisición y de atención. El valor del cliente a largo plazo, conocido en la literatura como customer lifetime value, es la estimación del valor económico que un cliente aportará durante toda su relación con el negocio. Predecir este valor, junto con otros aspectos del comportamiento futuro de los consumidores, permite a los emprendimientos asignar sus recursos comerciales de manera más eficiente y construir relaciones sostenibles con los clientes que más contribuyen a su viabilidad.

Sobre esta base, la predicción del comportamiento del consumidor abarca varias preguntas interrelacionadas. ¿Volverá a comprar este cliente? ¿Cuándo lo hará? ¿Cuánto gastará? ¿Qué productos le interesarán? ¿Existe riesgo de que abandone la relación con la empresa? Cada una de estas preguntas puede abordarse con técnicas específicas. La probabilidad de recompra y el riesgo de abandono se modelan habitualmente como problemas de clasificación, con las técnicas examinadas en el segundo capítulo, como la regresión logística, los bosques aleatorios o la potenciación de gradiente.

El momento de la próxima compra puede analizarse con modelos de supervivencia, que estiman el tiempo hasta un evento. El monto del gasto se aborda con modelos de regresión. Y las preferencias por productos se estudian con los sistemas de recomendación que se examinan más adelante en este capítulo.

El valor del cliente a largo plazo integra estas predicciones en una sola medida. En su formulación más simple, se calcula como la suma de los márgenes que el cliente generará en cada período futuro, multiplicados por la probabilidad de que siga siendo cliente en ese período y descontados al valor presente. Esta formulación pone de relieve sus tres componentes esenciales: el margen por período, que depende de cuánto compra el cliente y de la rentabilidad de lo que compra; la retención, que expresa la probabilidad de que la relación continúe; y el descuento, que refleja que un ingreso futuro vale menos que uno presente.

Los modelos más avanzados estiman estos componentes de manera individualizada, a partir del historial de cada cliente y de sus características. No obstante, Firmansyah, Machado y Moreira (2024), en una revisión sistemática sobre el uso de la inteligencia artificial para gestionar el valor del cliente a largo plazo, destacan la creciente utilización del aprendizaje automático para estimar un valor ajustado por riesgo, es decir, una medida que no solo considera el valor esperado de cada cliente, sino también la incertidumbre asociada a su comportamiento.

La revisión identifica el uso de análisis de media y varianza para optimizar carteras de clientes, en analogía con la gestión de carteras de inversión, y señala las fluctuaciones del ingreso de los clientes como un factor de riesgo importante. Los autores sostienen que estos enfoques pueden ayudar a las empresas a refinar su gestión del riesgo, identificar clientes de alto riesgo y construir modelos de valor más dinámicos y basados en datos.

Esta perspectiva de cartera resulta especialmente útil para los emprendimientos. Una empresa cuyos ingresos dependen de unos pocos clientes de alto valor puede mostrar un valor promedio elevado, pero enfrenta un riesgo considerable si alguno de ellos abandona la relación. Otra con muchos clientes de valor moderado y comportamiento estable puede tener un valor promedio menor pero una base de ingresos más resiliente.

Considerar tanto el valor esperado como su variabilidad permite tomar decisiones más equilibradas sobre a qué clientes atraer, cómo diversificar la cartera y cuánto invertir en la retención de cada grupo. Desde esta perspectiva, las aplicaciones prácticas de la predicción del valor del cliente son numerosas. En la adquisición de clientes, permite estimar cuánto se puede invertir

para atraer a un cliente de determinado perfil sin perder dinero, y orientar las campañas hacia los canales y segmentos que atraen clientes de mayor valor.

En la retención, permite identificar a los clientes valiosos con riesgo de abandono y dirigirles acciones específicas antes de que se vayan. En el desarrollo de la relación, orienta las ofertas de productos complementarios hacia los clientes con mayor potencial de crecimiento. Y en la planificación financiera, aporta una estimación de los ingresos futuros que puede complementar los pronósticos de demanda examinados anteriormente.

La predicción del abandono merece una atención particular. En los negocios basados en suscripciones o en relaciones recurrentes —servicios de software, gimnasios, servicios de telecomunicaciones, plataformas de contenidos, servicios profesionales—, la tasa de abandono es uno de los determinantes más importantes de la viabilidad. Un modelo que identifique tempranamente a los clientes con señales de desinterés —menor frecuencia de uso, reducción del gasto, quejas recientes, falta de respuesta a las comunicaciones— permite intervenir antes de que la relación se pierda.

La eficacia de estas intervenciones debe evaluarse, sin embargo, con cuidado: no todos los clientes en riesgo responden a las acciones de retención, y algunas acciones pueden resultar contraproducentes si se perciben como invasivas. Con todo, la aplicación de estos modelos en emprendimientos pequeños enfrenta algunas limitaciones. En primer lugar, destaca la cantidad de datos: estimar con precisión el comportamiento futuro de los clientes requiere historiales suficientemente largos, que las empresas nuevas no tienen.

En estos casos, pueden utilizarse inicialmente modelos simples, basados en promedios por segmento, y refinarlos a medida que se acumula información. En segundo término, sobresale la identificación de los clientes: muchos negocios no registran quién compra, sino solo qué se vende, lo que impide construir historiales individuales. Programas de fidelización, registros de clientes en las ventas en línea o sistemas de facturación electrónica pueden resolver parcialmente este problema. A ello se suma la estabilidad del entorno: en contextos de alta volatilidad económica, los patrones de comportamiento pueden cambiar con rapidez, lo que obliga a actualizar los modelos con frecuencia.

La predicción del comportamiento de los consumidores plantea también exigencias éticas. Los modelos pueden utilizar información sensible o inferir características que las personas no han revelado. Las decisiones basadas en el valor estimado de cada cliente pueden derivar en tratos

diferenciados que, en determinados casos, resulten injustos, como ofrecer peores condiciones a quienes el modelo considera menos valiosos.

La normativa de protección de datos personales exige transparencia sobre el uso de la información y, cuando se toman decisiones automatizadas con efectos significativos, el derecho a obtener una explicación y una revisión humana (Asamblea Nacional del Ecuador, 2021; SPDP, 2026). Una gestión responsable del valor del cliente combina la eficiencia comercial con el respeto de estos derechos y con la construcción de relaciones de confianza.

De forma complementaria, para la realidad latinoamericana, la difusión de los pagos digitales, del comercio electrónico y de la facturación electrónica está generando un volumen creciente de información sobre el comportamiento de compra, incluso en negocios pequeños. Este cambio crea condiciones para que los emprendimientos de la región incorporen la predicción del valor del cliente en su gestión, una práctica antes reservada a las grandes empresas con departamentos de análisis. Las plataformas que ofrecen estos servicios integran progresivamente funciones analíticas que facilitan su aplicación.

La estimación del valor del cliente a largo plazo permite, además, orientar las decisiones de inversión en adquisición de clientes. Si un emprendimiento conoce cuánto valor genera en promedio un cliente de cada segmento a lo largo de su relación con la empresa, puede determinar cuánto le conviene invertir para atraer a nuevos clientes de ese segmento sin comprometer su rentabilidad. Esta comparación entre el costo de adquisición y el valor esperado del cliente constituye uno de los indicadores más utilizados por las empresas digitales para evaluar la sostenibilidad de su crecimiento. Un crecimiento rápido basado en la adquisición de clientes cuyo valor esperado es inferior a su costo de adquisición puede conducir a pérdidas crecientes, aunque las ventas aumenten.

Al respecto, Firmansyah et al. (2024) destacan que la inteligencia artificial permite estimar este valor con mayor precisión y gestionarlo de manera dinámica, ajustando las acciones de retención y de desarrollo según la evolución de cada cliente. Vinculado con lo anterior, la predicción del abandono de clientes merece una consideración específica. Para muchos emprendimientos, retener a un cliente existente resulta menos costoso que atraer a uno nuevo, y una pequeña reducción de la tasa de abandono puede tener un efecto considerable sobre los ingresos a largo plazo.

Los modelos de clasificación pueden estimar, a partir del comportamiento reciente de cada cliente —la frecuencia de sus compras, el tiempo transcurrido desde la última, sus

reclamaciones o su respuesta a las comunicaciones—, la probabilidad de que deje de comprar en los próximos meses. Esta estimación permite dirigir las acciones de retención hacia los clientes con mayor riesgo y mayor valor, en lugar de dispersar los esfuerzos de manera indiscriminada.

La predicción del comportamiento del consumidor y la estimación del valor del cliente a largo plazo permiten que los emprendimientos pasen de una gestión comercial centrada en las transacciones a una centrada en las relaciones. Conocer qué clientes generan más valor, cuáles están en riesgo de abandonar y cuánto conviene invertir en cada uno mejora la eficiencia del gasto comercial y contribuye a la estabilidad de los ingresos. Incorporar el riesgo en estas estimaciones, como propone la literatura reciente, añade una dimensión de prudencia indispensable para negocios que no pueden permitirse depender de unos pocos clientes.

Figura 3.5

Predicción del comportamiento del consumidor y estimación del valor del cliente a largo plazo



Fuente: Elaboración autores

3.6. Análisis de sentimientos y minería de datos para anticipar la aceptación de productos innovadores

El lanzamiento de un producto innovador es una de las decisiones más arriesgadas que puede tomar un emprendimiento. A diferencia de los productos conocidos, cuya demanda puede estimarse a partir de la historia, los productos nuevos carecen de referentes directos, y la reacción de los consumidores resulta difícil de prever. Muchas innovaciones fracasan no por defectos técnicos, sino porque no responden a lo que los clientes valoran, porque se comunican de manera inadecuada o porque llegan al mercado en un momento inoportuno. El análisis de

sentimientos y la minería de datos ofrecen herramientas para reducir esta incertidumbre, al permitir conocer cómo perciben los consumidores un producto o una idea antes, durante y después de su lanzamiento.

En concordancia con ello, el análisis de sentimientos es una rama del procesamiento del lenguaje natural que busca identificar automáticamente la valoración emocional expresada en un texto: si un comentario es positivo, negativo o neutral, y, en sus versiones más refinadas, qué emociones específicas expresa o qué aspectos del producto se valoran de cada manera.

En esa dirección, Wankhade, Rao y Kulkarni (2022), en una revisión amplia del campo, describen los principales enfoques: los basados en diccionarios, que asignan valores positivos o negativos a las palabras y los combinan para calificar el texto; los basados en aprendizaje automático, que aprenden a clasificar los textos a partir de ejemplos etiquetados; y los enfoques híbridos, que combinan ambos. La revisión destaca también las aplicaciones del análisis de sentimientos en ámbitos como el mercadeo, la política o los servicios, y los desafíos que persisten, entre ellos la detección del sarcasmo, el manejo de las negaciones y la adaptación a distintos dominios e idiomas.

La precisión de los distintos métodos varía considerablemente. Sobre este punto, Hartmann et al. (2023) realizaron un metaanálisis de 272 conjuntos de datos con doce millones de documentos etiquetados y encontraron que los modelos de aprendizaje por transferencia recientemente propuestos ofrecían la mayor precisión, al clasificar en promedio más de veinte puntos porcentuales más documentos correctamente que los diccionarios de uso extendido.

Los autores advierten, no obstante, que estos modelos pueden desempeñarse peor de lo que sugieren las clasificaciones de referencia, que los diccionarios conservan la ventaja de ser más fáciles de interpretar y que las expectativas realistas de desempeño dependen de factores como el número de categorías de sentimiento y la longitud de los textos. Para los emprendedores, esta evidencia sugiere que la elección del método debe equilibrar la precisión con la interpretabilidad y los recursos disponibles.

A ello se suma que la minería de datos amplía el análisis más allá de los textos. Se refiere al conjunto de técnicas que permiten descubrir patrones, relaciones y anomalías en grandes volúmenes de información, combinando métodos estadísticos y de aprendizaje automático. En el contexto de la aceptación de productos innovadores, la minería de datos puede aplicarse a los registros de navegación en sitios web, a las interacciones con publicaciones en redes sociales, a los resultados de pruebas de concepto, a las preinscripciones o a las ventas iniciales en

mercados de prueba. Al combinar estas fuentes con el análisis de sentimientos, es posible construir una imagen más completa de la reacción del mercado.

Antes del lanzamiento, el análisis de sentimientos puede aplicarse a las conversaciones sobre productos similares o sobre la necesidad que la innovación pretende atender. Si los consumidores expresan frustración con las soluciones existentes, existe un espacio para la innovación; si expresan satisfacción, la propuesta deberá ofrecer ventajas muy claras para desplazar a los competidores.

También puede utilizarse para evaluar la recepción de prototipos, conceptos o campañas de prueba: presentar una idea a un grupo de potenciales clientes, recoger sus comentarios y analizar su tono permite estimar el interés y detectar objeciones antes de invertir en la producción a escala. Asimismo, durante el lanzamiento, el seguimiento de los sentimientos en tiempo real permite detectar problemas con rapidez. Un aumento de comentarios negativos sobre un aspecto específico —el precio, la entrega, una funcionalidad— puede revelar un defecto que requiere corrección inmediata.

Un patrón de comentarios positivos en un segmento inesperado puede señalar una oportunidad de mercado no prevista. El análisis basado en aspectos, que asocia los sentimientos con características concretas del producto, resulta particularmente útil en esta etapa, porque transforma la valoración general en información accionable sobre qué mejorar.

Después del lanzamiento, el análisis sistemático de reseñas y comentarios permite evaluar la aceptación del producto, comparar su recepción con la de los competidores y orientar las versiones siguientes. Las tendencias en el sentimiento a lo largo del tiempo pueden servir como indicadores tempranos de cambios en las ventas: un deterioro progresivo de las valoraciones puede anticipar una caída de la demanda antes de que se refleje en los ingresos.

Integrar estos indicadores en los modelos de pronóstico de demanda y en los tableros de control del negocio permite actuar con mayor anticipación. Ahora bien, el análisis de sentimientos presenta limitaciones que conviene reconocer. Una de ellas es la representatividad: las personas que escriben comentarios en línea no representan necesariamente al conjunto de los consumidores, y suelen estar sobrerrepresentadas las experiencias muy positivas o muy negativas. Otra es la autenticidad: las reseñas falsas, los comentarios pagados y las campañas coordinadas pueden distorsionar los resultados.

A ello se suma la complejidad del lenguaje: la ironía, el sarcasmo, las expresiones regionales y las referencias culturales dificultan la clasificación automática, especialmente en idiomas y variantes con menos recursos que el inglés, como el español de los distintos países latinoamericanos (Wankhade et al., 2022). Por último, cabe mencionar la distancia entre la opinión expresada y el comportamiento: un comentario positivo no garantiza una compra, ni uno negativo un abandono.

Para los emprendimientos de la región, estas herramientas resultan cada vez más accesibles. Existen modelos preentrenados de análisis de sentimientos disponibles de manera abierta, incluidos algunos que pueden aplicarse a textos en español; Hartmann et al. (2023), por ejemplo, pusieron a disposición un modelo de aprendizaje por transferencia que puede utilizarse con una facilidad comparable a la de un diccionario. Las plataformas de gestión de redes sociales y de atención al cliente incorporan funciones de análisis de sentimientos integradas. El desafío principal reside en definir qué preguntas se quieren responder, recolectar los datos de manera ética, validar los resultados con muestras revisadas por personas y traducir los hallazgos en decisiones sobre el producto y su comunicación.

Además, la recolección y el análisis de opiniones deben respetar, además, la privacidad de las personas. Aunque los comentarios públicos pueden analizarse para identificar patrones generales, su vinculación con perfiles individuales, su almacenamiento prolongado o su uso con finalidades distintas de las previstas pueden contravenir la normativa de protección de datos personales (Asamblea Nacional del Ecuador, 2021). El análisis agregado, centrado en las tendencias y no en las personas, permite aprovechar el potencial de estas herramientas con menores riesgos.

El análisis de sentimientos puede integrarse también en el proceso de desarrollo de productos. Antes del lanzamiento, el análisis de las conversaciones en redes sociales y de las reseñas de productos similares permite identificar los atributos más valorados y los aspectos que generan insatisfacción, información que puede orientar el diseño. Durante las pruebas con usuarios, el análisis de sus comentarios abiertos complementa las métricas cuantitativas y ayuda a comprender las razones de su aceptación o rechazo.

Después del lanzamiento, el seguimiento del sentimiento en las reseñas permite detectar problemas de calidad o de servicio con rapidez y medir el efecto de las mejoras introducidas. Esta integración convierte la voz de los consumidores en un insumo continuo del proceso de innovación (Wankhade et al., 2022).

Para los emprendimientos con recursos limitados, este análisis puede comenzar con la lectura sistemática y la clasificación manual de los comentarios, y automatizarse a medida que su volumen crece. Visto en conjunto, el análisis de sentimientos y la minería de datos ofrecen a los emprendimientos una forma de escuchar al mercado de manera continua y sistemática. Aplicados antes, durante y después del lanzamiento de un producto innovador, permiten anticipar su aceptación, detectar problemas con rapidez y orientar las mejoras con base en la experiencia de los consumidores. Su valor depende de la elección adecuada de los métodos, de la validación de los resultados y de su integración con otras fuentes de información, de modo que las opiniones expresadas en línea se conviertan en un insumo más, y no en el único, para decidir sobre la innovación.

Figura 3.6

Análisis de sentimientos y minería de datos para anticipar la aceptación de productos innovadores



Fuente: Elaboración autores

3.7. Sistemas de recomendación basados en inteligencia artificial para la personalización de modelos de negocio

Cuando una plataforma de comercio electrónico sugiere un producto que complementa la compra anterior, cuando un servicio de música propone canciones afines a los gustos del usuario o cuando una aplicación de reparto recomienda un restaurante cercano, detrás de esas sugerencias opera un sistema de recomendación. Estos sistemas utilizan datos sobre las preferencias y los comportamientos de los usuarios para predecir qué productos, servicios o contenidos les resultarán más relevantes. Se han convertido en uno de los componentes más rentables de la

economía digital, porque aumentan las ventas, mejoran la experiencia de los clientes y facilitan el descubrimiento de productos en catálogos extensos.

Para los emprendimientos, representan una herramienta para personalizar su propuesta de valor y competir con empresas de mayor tamaño. En esta misma línea, Roy y Dutta (2022), en una revisión sistemática de los sistemas de recomendación, analizan los algoritmos que los sustentan, construyen una taxonomía de los componentes necesarios para desarrollar un sistema eficaz y examinan sus aplicaciones en ámbitos como los libros, las películas y los productos de consumo. La revisión identifica también las principales limitaciones de estos sistemas, entre ellas la escalabilidad, el problema del arranque en frío y la dispersión de los datos, y subraya que la diversidad de técnicas disponibles hace compleja la elección del enfoque adecuado para cada aplicación.

Los sistemas de recomendación suelen clasificarse en tres grandes enfoques. El filtrado colaborativo se basa en la idea de que las personas con gustos similares en el pasado tendrán gustos similares en el futuro: si dos clientes compraron productos parecidos, lo que uno compró y el otro no constituye una recomendación potencial. Este enfoque no necesita conocer las características de los productos, sino únicamente el historial de interacciones entre usuarios y productos.

El filtrado basado en contenido, en cambio, recomienda productos similares a los que el usuario ya ha valorado positivamente, a partir de las características de esos productos, como la categoría, el precio, la marca o la descripción. Y los sistemas híbridos combinan ambos enfoques para aprovechar sus fortalezas y compensar sus debilidades. De igual manera, los problemas identificados por Roy y Dutta (2022) tienen especial relevancia para los emprendimientos. El arranque en frío se produce cuando un usuario o un producto nuevo carece de historial suficiente para generar recomendaciones confiables; para una empresa que recién inicia operaciones, todos sus clientes y productos están, en algún sentido, en esta situación.

La dispersión de los datos se refiere a que la mayoría de los usuarios interactúa solo con una pequeña fracción de los productos, lo que dificulta encontrar patrones. La escalabilidad, por su parte, se vuelve un desafío cuando el número de usuarios y productos crece rápidamente. Las estrategias para mitigar estos problemas incluyen el uso de información sobre las características de los productos y de los usuarios, la recomendación inicial de productos populares, la solicitud explícita de preferencias al registrarse y la combinación de distintos enfoques.

Más allá de la recomendación de productos, la lógica de estos sistemas puede extenderse a la personalización del modelo de negocio en su conjunto. Un emprendimiento puede utilizar los datos de comportamiento de sus clientes para personalizar no solo qué productos ofrece, sino también cómo los presenta, a qué precio, a través de qué canal y en qué momento. La personalización de la comunicación —el contenido de un correo, el horario de un mensaje, la oferta destacada en una página— puede aumentar significativamente la respuesta de los clientes.

La personalización del servicio, adaptando la atención a las preferencias y necesidades de cada cliente, puede mejorar su satisfacción y fidelidad. Y la personalización del producto, mediante configuraciones a medida, puede convertirse en una fuente de diferenciación frente a competidores que ofrecen soluciones estandarizadas. Por otra parte, los sistemas de recomendación también pueden aplicarse en contextos de negocio a negocio.

Una distribuidora mayorista puede recomendar a sus clientes minoristas los productos que mejor se venden en tiendas con características similares; una plataforma de servicios para pequeñas empresas puede sugerir herramientas o proveedores según el perfil de cada negocio; un programa de apoyo al emprendimiento puede recomendar a cada emprendedor los cursos, mentores o fuentes de financiamiento más adecuados a su situación. En este último caso, la lógica del filtrado colaborativo resulta especialmente prometedora: los emprendimientos con trayectorias similares que se beneficiaron de determinados apoyos ofrecen información valiosa para orientar a otros.

La implementación de un sistema de recomendación no requiere necesariamente grandes inversiones. Muchas plataformas de comercio electrónico incluyen funciones básicas de recomendación, como productos frecuentemente comprados juntos o productos vistos por otros clientes. Para emprendimientos con capacidades técnicas, existen bibliotecas de código abierto que facilitan el desarrollo de sistemas propios. Lo que sí resulta indispensable es disponer de datos sobre las interacciones de los clientes, organizados de manera que puedan analizarse, y definir con claridad qué se quiere lograr con las recomendaciones: aumentar el valor de cada compra, mejorar la retención, facilitar el descubrimiento de productos o mejorar la experiencia.

Paralelamente, la evaluación de los sistemas de recomendación debe considerar varios criterios. La precisión mide si las recomendaciones coinciden con lo que los usuarios efectivamente eligen. La cobertura indica qué proporción del catálogo llega a recomendarse, lo que importa para evitar que el sistema se concentre solo en los productos más populares. La diversidad y la

novedad valoran que las recomendaciones no sean siempre las mismas ni demasiado obvias. Y, sobre todo, el impacto en el negocio, medido mediante experimentos controlados que comparan el comportamiento de los clientes que reciben recomendaciones con el de quienes no las reciben, permite saber si el sistema genera valor real.

Los sistemas de recomendación plantean también riesgos que los emprendimientos deben gestionar. El primero es el de las burbujas de filtro, en las que el sistema refuerza continuamente las preferencias existentes y limita la exposición de los usuarios a alternativas. El segundo es el de la manipulación, cuando las recomendaciones se diseñan para favorecer los intereses de la empresa en detrimento de los del cliente, por ejemplo, promoviendo productos de mayor margen aunque no sean los más adecuados.

El tercero es el de la privacidad, porque estos sistemas requieren recopilar y analizar información detallada sobre el comportamiento de las personas. La normativa de protección de datos personales exige transparencia sobre estos tratamientos y el respeto de los derechos de los titulares (Asamblea Nacional del Ecuador, 2021). Una personalización que respeta estos límites fortalece la confianza; una que los ignora puede generar rechazo. Sobre esta base, en el contexto latinoamericano, donde muchos emprendimientos venden a través de plataformas de terceros, conviene considerar la tensión advertida por Chalmers, MacKenzie y Carter (2021) sobre el poder de los intermediarios digitales. Cuando los sistemas de recomendación pertenecen a las plataformas y no a los emprendimientos, son aquellas las que deciden qué productos se muestran a los consumidores y en qué orden, lo que puede favorecer a determinados vendedores o a los productos de la propia plataforma.

La implementación de sistemas de recomendación en emprendimientos pequeños enfrenta el problema del arranque en frío: al inicio, el negocio no dispone de suficiente información sobre las preferencias de sus clientes ni sobre las relaciones entre sus productos. Para superar esta limitación, pueden utilizarse inicialmente recomendaciones basadas en reglas sencillas, como los productos más vendidos, los que se compran juntos con mayor frecuencia o los que corresponden a la temporada, e incorporar progresivamente algoritmos de filtrado colaborativo a medida que se acumulan datos.

Las plataformas de comercio electrónico disponibles para pequeñas empresas suelen incluir funciones de recomendación básicas que pueden aprovecharse sin desarrollo propio. En este sentido, Roy y Dutta (2022) señalan que los enfoques híbridos, que combinan información

sobre los productos con información sobre el comportamiento de los usuarios, ayudan a mitigar el problema del arranque en frío y a mejorar la diversidad de las recomendaciones.

Al seleccionar qué productos o servicios se muestran a cada cliente, influyen en sus decisiones de compra y pueden orientarlas hacia los intereses del negocio en detrimento de los del consumidor, por ejemplo, privilegiando productos de mayor margen frente a los más adecuados para sus necesidades. También pueden encerrar a los usuarios en un círculo de recomendaciones similares, reduciendo su exposición a alternativas diversas. Un uso responsable de estos sistemas implica informar a los clientes sobre el carácter personalizado de las recomendaciones, ofrecerles la posibilidad de desactivarlas y orientarlas genuinamente a su satisfacción.

Lo expuesto permite afirmar que los sistemas de recomendación basados en inteligencia artificial permiten a los emprendimientos personalizar su oferta y su relación con los clientes a una escala antes reservada a las grandes empresas. Su aprovechamiento exige datos sobre las interacciones de los clientes, una elección adecuada del enfoque según la etapa del negocio, una evaluación rigurosa de su impacto y una gestión responsable de los riesgos de manipulación y privacidad. Bien utilizados, contribuyen a que cada cliente encuentre lo que necesita y a que el emprendimiento construya relaciones más duraderas y rentables.

Figura 3.7

Sistemas de recomendación basados en inteligencia artificial para la personalización de modelos de negocio



Fuente: Elaboración autores

3.8. Construcción de observatorios digitales de oportunidades empresariales mediante datos económicos y comerciales

Las herramientas examinadas en este capítulo —inteligencia económica, pronóstico de demanda, segmentación, procesamiento del lenguaje natural, análisis de sentimientos y sistemas de recomendación— ofrecen a los emprendimientos capacidades poderosas para anticipar el comportamiento de los mercados. Sin embargo, la mayoría de las pequeñas empresas no dispone de los datos, las competencias ni los recursos necesarios para aplicarlas por sí sola. Un observatorio digital de oportunidades empresariales puede resolver parcialmente esta limitación, al concentrar la recolección, el análisis y la difusión de información económica y comercial en una institución que pone sus resultados a disposición del conjunto del ecosistema emprendedor.

No obstante, un observatorio de este tipo puede definirse como una plataforma institucional que recopila de manera sistemática datos sobre la economía, los mercados, los consumidores y las empresas de un territorio o sector; los procesa mediante técnicas analíticas y de inteligencia artificial; y genera productos de información orientados a identificar oportunidades de negocio, anticipar riesgos y apoyar decisiones de emprendedores, financiadores y responsables de políticas. A diferencia de los sistemas estadísticos tradicionales, que publican información con periodicidad fija y orientada principalmente a la medición, un observatorio de oportunidades se orienta a la acción: su propósito es responder preguntas concretas sobre dónde y en qué conviene emprender.

La construcción de un observatorio de oportunidades requiere integrar diversas fuentes de información. Las estadísticas oficiales constituyen la base: en el Ecuador, el registro estadístico de empresas del Instituto Nacional de Estadística y Censos ofrece información sobre el número de empresas, sus ventas, su empleo y su distribución sectorial y territorial (INEC, 2024), y el informe del Global Entrepreneurship Monitor aporta datos sobre las características, motivaciones y dificultades de los emprendedores (Lasio et al., 2024).

Los registros administrativos de superintendencias, administraciones tributarias y entidades de seguridad social complementan esta información con mayor detalle. A estas fuentes se pueden sumar datos de comercio exterior, precios, búsquedas en internet, ofertas de empleo, reseñas de consumidores y publicaciones en redes sociales, que aportan señales más inmediatas sobre la evolución de los mercados. De forma complementaria, la disponibilidad de datos abiertos facilita esta integración. En el caso ecuatoriano, la Ley Orgánica de Transparencia y Acceso a la

Información Pública de 2023 exige que la información de publicación obligatoria se difunda en formato de datos abiertos, definidos como datos digitales que pueden ser utilizados, reutilizados y redistribuidos libremente (Asamblea Nacional del Ecuador, 2023). Esta obligación crea condiciones para que observatorios públicos, académicos o privados aprovechen la información generada por el Estado para producir análisis útiles para los emprendedores.

Un mapa de saturación sectorial y territorial permitiría a los emprendedores identificar dónde existen demasiados negocios de un mismo tipo y dónde, en cambio, la oferta es insuficiente en relación con la población o la demanda. Un tablero de tendencias de demanda mostraría las categorías de productos y servicios con crecimiento sostenido, a partir de datos de ventas, búsquedas o importaciones. Un sistema de alertas sobre cambios regulatorios, tributarios o de mercado ayudaría a anticipar amenazas y oportunidades. Y un catálogo de necesidades insatisfechas, construido mediante el análisis de reseñas y conversaciones de consumidores, señalaría nichos para la innovación. Vinculado con lo anterior, la experiencia internacional muestra que las fuentes de datos alternativas pueden mejorar la anticipación de la actividad económica.

Al respecto, Bantis, Clements y Urquhart (2023) encontraron que la incorporación de datos de búsquedas en internet mejoraba los pronósticos del crecimiento económico en los Estados Unidos y en Brasil, siempre que se seleccionaran adecuadamente las variables. Las lecciones de la competencia de pronóstico M5 sobre el valor de las variables externas y del aprendizaje transversal entre series relacionadas (Makridakis et al., 2022) sugieren que un observatorio que integre múltiples fuentes podría producir pronósticos sectoriales más precisos que los que cada empresa obtendría por separado.

La arquitectura técnica de un observatorio puede organizarse en capas, en línea con la arquitectura del emprendimiento predictivo descrita en el primer capítulo. La capa de captura recoge los datos de las distintas fuentes mediante descargas periódicas, conexiones automatizadas o acuerdos de intercambio. La capa de almacenamiento y gobernanza organiza la información, verifica su calidad, garantiza la protección de los datos personales y documenta su origen. Por otro lado, la capa analítica aplica los modelos de pronóstico, segmentación, procesamiento del lenguaje natural y detección de tendencias.

Además, la capa de difusión presenta los resultados mediante tableros interactivos, informes, alertas y servicios de consulta y la capa de retroalimentación recoge la valoración de los usuarios y los resultados de las decisiones tomadas, para mejorar continuamente los productos del

observatorio. Debe definirse quién lo administra —una universidad, una entidad pública, una cámara empresarial, una alianza entre varias de ellas—, cómo se financia de manera sostenible, qué datos se comparten y con qué garantías, y cómo se asegura la independencia y la calidad de sus análisis. La colaboración entre actores del ecosistema emprendedor, destacada por la literatura sobre ecosistemas como condición de su buen funcionamiento (Stam y van de Ven, 2021), resulta indispensable: ningún actor dispone por sí solo de todos los datos ni de todas las capacidades necesarias.

Un observatorio de oportunidades debe atender también a la equidad en el acceso a la información. Si sus productos solo están al alcance de quienes tienen formación técnica o conectividad de calidad, podría ampliar las brechas entre emprendedores. Por ello, conviene ofrecer productos con distintos niveles de complejidad, desde tableros interactivos para usuarios avanzados hasta resúmenes sencillos, boletines y servicios de orientación personalizada para quienes tienen menos experiencia.

La articulación con incubadoras, centros de desarrollo empresarial, gobiernos locales y organizaciones de la economía popular y solidaria puede facilitar que la información llegue a quienes más la necesitan. A ello se suma que en el contexto ecuatoriano, un observatorio de oportunidades podría contribuir a enfrentar algunos de los rasgos estructurales del emprendimiento nacional. La concentración de los nuevos negocios en actividades de comercio orientadas al consumidor y la muy baja proporción de emprendimientos con ingresos del exterior (Lasio et al., 2024) sugieren que muchas personas emprenden sin información suficiente sobre oportunidades alternativas. Proporcionar análisis sobre sectores con menor saturación, sobre mercados de exportación accesibles o sobre necesidades locales no atendidas podría orientar a una parte de los emprendedores hacia actividades con mejores perspectivas de supervivencia y crecimiento.

La evaluación del impacto de un observatorio constituye otro aspecto que debe preverse desde su diseño. No basta con medir el número de consultas o descargas de sus productos; interesa saber si la información proporcionada influye en las decisiones de los emprendedores y si esas decisiones mejoran sus resultados. Para ello, el observatorio puede realizar encuestas de seguimiento a sus usuarios, registrar las decisiones tomadas a partir de sus análisis y, cuando sea posible, comparar la supervivencia y el crecimiento de los emprendimientos que utilizaron sus servicios con los de negocios similares que no lo hicieron.

Esta evaluación permite ajustar los productos a las necesidades reales de los usuarios, justificar su financiamiento ante las instituciones que lo sostienen y contribuir al conocimiento sobre el

papel de la información en el éxito emprendedor. La retroalimentación sistemática es, además, coherente con la lógica de aprendizaje continuo que caracteriza a todo el enfoque predictivo.

Figura 3.8

Construcción de observatorios digitales de oportunidades empresariales mediante datos económicos y comerciales



Fuente: Elaboración autores

EDITORIAL ANDES COGNITIO

CAPÍTULO IV

FINANZAS INTELIGENTES, GEMELOS DIGITALES Y SIMULACIÓN PREDICTIVA DEL RIESGO EMPRENDEDOR



CAPÍTULO IV

FINANZAS INTELIGENTES, GEMELOS DIGITALES Y SIMULACIÓN PREDICTIVA DEL RIESGO EMPRENDEDOR

4.1. Finanzas predictivas y transformación de los métodos tradicionales de evaluación de emprendimientos

La evaluación financiera de los proyectos empresariales se ha apoyado tradicionalmente en un conjunto de herramientas consolidadas: el flujo de caja proyectado, el valor actual neto, la tasa interna de retorno, el período de recuperación de la inversión y el análisis del punto de equilibrio. Estas herramientas comparten una lógica común: a partir de supuestos sobre ventas, costos, inversiones y financiamiento, se proyectan los flujos de efectivo futuros y se descuentan a una tasa que refleja el costo del capital y el riesgo del proyecto. Si el valor resultante es positivo, el proyecto se considera viable.

Esta metodología ha sido, y sigue siendo, el lenguaje común entre emprendedores, bancos, inversionistas y organismos de fomento. Con todo, su aplicación a los emprendimientos tropieza, sin embargo, con limitaciones importantes. En un negocio nuevo, sin historia, las proyecciones de ventas y costos descansan en estimaciones frágiles que pueden alejarse considerablemente de la realidad.

El riesgo suele reducirse a una tasa de descuento elevada o a unos pocos escenarios —optimista, esperado y pesimista— que no capturan la diversidad de trayectorias posibles. El análisis, por otra parte, se realiza una vez, al inicio, y rara vez se actualiza cuando llega nueva información. A ello se suma su desconexión de los datos operativos, pues las proyecciones se construyen en hojas de cálculo aisladas, sin vínculo con la información real que el negocio genera día a día.

Las finanzas predictivas proponen transformar estos métodos mediante la incorporación sistemática de datos y modelos de inteligencia artificial. La investigación sobre la inteligencia artificial y el aprendizaje automático en las finanzas ha crecido de manera notable. En esa dirección, Goodell, Kumar, Lim y Pattnaik (2021), en un análisis bibliométrico de la literatura publicada entre 1986 y 2021, identificaron tres grandes áreas de investigación: la construcción de carteras, la valoración y el comportamiento de los inversionistas; el fraude y las dificultades financieras; y la inferencia de sentimientos, el pronóstico y la planificación.

Las tres tienen aplicaciones directas en la evaluación de emprendimientos: estimar el valor de una empresa emergente, anticipar sus dificultades financieras y pronosticar sus resultados para planificar su desarrollo. Asimismo, la transformación que proponen las finanzas predictivas puede describirse en varias dimensiones. En lugar de supuestos fijos, utiliza distribuciones de probabilidad para las variables clave, estimadas a partir de datos históricos del propio negocio, de empresas comparables o del sector.

En lugar de unos pocos escenarios, genera miles de trayectorias posibles mediante simulación, lo que permite estimar no solo el valor esperado del proyecto, sino también la probabilidad de que resulte negativo, la magnitud de las pérdidas posibles y los factores que más influyen en el resultado. En lugar de un análisis estático, propone una evaluación continua, que se actualiza a medida que el negocio genera datos sobre sus ventas, costos y flujos de efectivo. Y en lugar de hojas de cálculo aisladas, integra la evaluación con los sistemas contables, de ventas y bancarios del emprendimiento.

Estas transformaciones tienen particular relevancia en el contexto de los emprendimientos latinoamericanos, para los que el acceso al financiamiento constituye uno de los principales obstáculos. Según el informe del Global Entrepreneurship Monitor para el Ecuador, las dificultades para obtener financiamiento figuraron como la segunda causa de cierre de los negocios, con el 21,4 %, después de la falta de rentabilidad, y el apoyo financiero apareció entre las condiciones más débiles del entorno emprendedor del país (Lasio et al., 2024). El estudio de la OCDE (2025) sobre la adopción de la inteligencia artificial en las pequeñas y medianas empresas señala, en la misma línea, que estas empresas enfrentan con frecuencia dificultades de acceso al crédito y que el endurecimiento de las condiciones crediticias ha frenado sus inversiones de largo plazo.

Ahora bien, una de las razones de estas dificultades es la asimetría de información entre emprendedores y financiadores. Las entidades financieras no disponen de información suficiente para evaluar el riesgo de los nuevos negocios, que carecen de historial crediticio y, con frecuencia, de garantías. Ante esta incertidumbre, tienden a racionar el crédito o a exigir condiciones que muchos emprendimientos no pueden cumplir.

Las finanzas predictivas pueden contribuir a reducir esta asimetría al incorporar en la evaluación información adicional sobre el desempeño del negocio: ventas registradas en plataformas digitales, pagos electrónicos recibidos, comportamiento de los clientes, cumplimiento de obligaciones con proveedores. Sobre este punto, Bussmann et al. (2021) mostraron, en un estudio sobre 15.000 pequeñas y medianas empresas que solicitaron financiamiento en plataformas de

préstamos entre pares, que los modelos de aprendizaje automático acompañados de técnicas de explicación podían agrupar a los prestatarios según características financieras comunes y ayudar a anticipar su comportamiento futuro.

La aplicación de las finanzas predictivas no implica abandonar las herramientas tradicionales, sino enriquecerlas. El valor actual neto sigue siendo un criterio útil, pero se calcula ahora como una distribución y no como un único número. El punto de equilibrio sigue siendo una referencia, pero se acompaña de una estimación de la probabilidad de alcanzarlo en distintos plazos. El flujo de caja sigue siendo el centro del análisis, pero se pronostica con modelos que aprenden de los datos reales del negocio. Esta combinación permite conservar el lenguaje común con los financiadores, al tiempo que se ofrece una visión más completa de la incertidumbre. Además, la evaluación predictiva de emprendimientos puede organizarse en varias etapas. En la etapa previa a la creación, se utilizan datos de empresas comparables y del sector para estimar las distribuciones de las variables clave y evaluar la viabilidad del proyecto.

En los primeros meses de operación, los datos reales del negocio comienzan a sustituir progresivamente a los supuestos iniciales, y las proyecciones se recalibran. En la etapa de consolidación, los modelos se alimentan de un historial suficiente para producir pronósticos más precisos y para detectar tempranamente desviaciones. Y en la etapa de crecimiento, la evaluación se extiende a decisiones de inversión, expansión o búsqueda de financiamiento adicional. En cada etapa, la incertidumbre se reduce a medida que aumenta la información disponible.

Las finanzas predictivas plantean también desafíos que conviene reconocer. Los modelos dependen de la calidad de los datos, que en muchos emprendimientos es deficiente. La sofisticación de las técnicas puede generar una falsa sensación de precisión, cuando en realidad las estimaciones siguen sujetas a una incertidumbre considerable.

La utilización de datos alternativos para evaluar el riesgo crediticio puede generar problemas de privacidad y de discriminación si no se gestiona adecuadamente, tema que se aborda en el quinto capítulo. Y la adopción de estas herramientas requiere competencias financieras y analíticas que muchas personas emprendedoras todavía no tienen, lo que convierte la formación en un componente indispensable de cualquier estrategia de finanzas predictivas.

En esta misma línea, la transformación de la evaluación financiera alcanza también a la relación entre emprendedores e inversionistas. En las rondas de financiamiento de empresas emergentes,

la valoración se ha basado tradicionalmente en comparaciones con transacciones similares y en proyecciones optimistas elaboradas por los propios fundadores.

La incorporación de modelos predictivos permite contrastar esas proyecciones con la evidencia de empresas comparables, estimar la probabilidad de alcanzar los hitos previstos y negociar condiciones que reflejen de manera más equilibrada el riesgo asumido por ambas partes. Para las personas emprendedoras, disponer de proyecciones probabilísticas sustentadas en datos fortalece su posición en la negociación y facilita la comunicación con inversionistas que exigen cada vez más rigor analítico. Para los inversionistas, reduce la dependencia de la intuición y permite construir carteras más diversificadas. En este sentido, Goodell et al. (2021) identifican la valoración y el comportamiento de los inversionistas como uno de los grandes campos de aplicación de la inteligencia artificial en las finanzas.

Al sustituir los supuestos fijos por distribuciones de probabilidad, los escenarios aislados por simulaciones, el análisis estático por la evaluación continua y las hojas de cálculo aisladas por sistemas integrados, ofrecen a los emprendedores y a sus financiadores una visión más realista de la viabilidad de los proyectos. En economías donde el acceso al financiamiento limita la supervivencia de los nuevos negocios, esta capacidad de evaluar mejor el riesgo puede contribuir a ampliar el crédito hacia emprendimientos viables que hoy quedan excluidos.

Figura 4.1

Finanzas predictivas y transformación de los métodos tradicionales de evaluación de emprendimientos



Fuente: Elaboración autores

4.2. Modelos de inteligencia artificial para pronosticar ingresos, costos, rentabilidad y flujos de efectivo

La salud financiera de un emprendimiento depende de cuatro magnitudes estrechamente relacionadas: los ingresos que genera, los costos en que incurre, la rentabilidad que resulta de la diferencia entre ambos y los flujos de efectivo que determinan su capacidad para pagar sus obligaciones en el momento oportuno. Una empresa puede ser rentable en términos contables y quebrar por falta de liquidez, si sus clientes pagan tarde mientras sus proveedores exigen pagos inmediatos; o puede disponer de efectivo abundante durante un tiempo mientras acumula pérdidas que comprometen su futuro. Pronosticar estas cuatro magnitudes de manera integrada permite anticipar los problemas antes de que se manifiesten y tomar decisiones con mayor fundamento.

De igual manera, el pronóstico de ingresos constituye el punto de partida, porque de él dependen buena parte de los costos variables y de los flujos de efectivo. Las técnicas de pronóstico de demanda examinadas en el tercer capítulo se aplican directamente a esta tarea. Al respecto, Hyndman y Athanasopoulos (2021) describen cómo los métodos de series temporales, como el suavizamiento exponencial y los modelos ARIMA, permiten proyectar las ventas a partir de su comportamiento histórico, incorporando tendencia y estacionalidad, y cómo los modelos de regresión dinámica pueden incluir variables explicativas como precios, promociones o condiciones económicas.

La evidencia de la competencia M5 muestra que los algoritmos de aprendizaje automático, en particular los de potenciación de gradiente, pueden superar a los métodos estadísticos cuando se dispone de muchas series relacionadas y de información externa relevante, y que la combinación de modelos suele mejorar la precisión (Makridakis et al., 2022).

El pronóstico de costos presenta características distintas. Una parte de los costos es fija y relativamente previsible, como los arriendos, los salarios del personal permanente o las cuotas de los préstamos. Otra parte es variable y depende del volumen de actividad, como las materias primas, las comisiones de venta o los costos de envío. Y otra parte está sujeta a variaciones de precios que el emprendimiento no controla, como los insumos importados, la energía o el transporte.

Los modelos de inteligencia artificial pueden contribuir a pronosticar los costos variables a partir de los pronósticos de ventas, y a anticipar las variaciones de precios de los insumos a

partir de series históricas e indicadores externos. La integración de ambos pronósticos permite estimar los márgenes futuros y su sensibilidad frente a cambios en las condiciones del mercado.

Por otra parte, la rentabilidad pronosticada resulta de combinar los pronósticos de ingresos y costos. Sin embargo, conviene ir más allá de una estimación puntual y considerar su distribución. Si las ventas pueden variar en un rango determinado y los costos de los insumos en otro, la rentabilidad resultante puede ubicarse en un rango más amplio, que incluye escenarios de pérdida.

Los modelos que generan pronósticos probabilísticos, con intervalos que expresan la incertidumbre, permiten estimar la probabilidad de que el negocio alcance determinados niveles de rentabilidad o de que incurra en pérdidas (Hyndman y Athanasopoulos, 2021). Esta información resulta más útil para la toma de decisiones que una cifra única, porque permite evaluar el riesgo asumido.

El pronóstico de los flujos de efectivo es, probablemente, el más importante para la supervivencia de un emprendimiento. A diferencia de la rentabilidad, que se registra cuando se producen las ventas y los gastos, los flujos de efectivo reflejan el momento en que el dinero entra y sale efectivamente de la empresa. Su pronóstico requiere modelar no solo cuánto se venderá, sino también cuándo pagarán los clientes; no solo cuánto se gastará, sino también cuándo deberán pagarse los proveedores, los impuestos, los salarios y las deudas.

Los modelos de aprendizaje automático pueden estimar, a partir del historial de cada cliente, la probabilidad de que pague a tiempo, con retraso o no pague, lo que permite proyectar los cobros con mayor realismo que el simple supuesto de que todos pagarán en el plazo acordado. Paralelamente, la evidencia sobre los determinantes del cierre de empresas en el Ecuador subraya la importancia de estos pronósticos.

Los estudios basados en datos de la Superintendencia de Compañías, Valores y Seguros han mostrado que la rentabilidad sobre los activos y el volumen de ventas reducen el riesgo de cierre, mientras que el apalancamiento financiero lo aumenta (González Cedeño, 2026), y que la rotación de cartera y el equilibrio del endeudamiento prolongan la vida de las empresas (Zambrano et al., 2021). Pronosticar la evolución de estas variables permite anticipar el deterioro de la situación financiera antes de que se vuelva irreversible.

Del mismo modo, Hidalgo-Díaz, Alaminos y Delgado-Gómez (2025), al analizar 8.400 empresas europeas, encontraron que el margen de beneficio neto y la rentabilidad sobre los activos

mantenían su capacidad de anticipar la insolvencia hasta tres años antes del evento, mientras que otros indicadores solo resultaban significativos en horizontes más cortos.

La construcción de un sistema integrado de pronóstico financiero para un emprendimiento puede organizarse en varios módulos. Un módulo de ventas pronostica los ingresos por producto, canal y período, utilizando los datos históricos y las variables externas disponibles. Un módulo de costos estima los costos variables a partir de las ventas pronosticadas y proyecta los costos fijos según los compromisos existentes.

Un módulo de cobros y pagos estima los plazos efectivos en que se recibirán y realizarán los pagos, a partir del comportamiento histórico de clientes y proveedores. Un módulo de resultados integra los anteriores para proyectar la rentabilidad y su distribución. Y un módulo de tesorería proyecta el saldo de efectivo período a período, señalando los momentos en que podría resultar insuficiente. Sobre esta base, para que este sistema funcione, debe alimentarse con datos actualizados. La integración con los sistemas contables, de facturación electrónica, de ventas en línea y bancarios permite que los pronósticos se actualicen automáticamente a medida que se registran nuevas transacciones. En muchos países latinoamericanos, la generalización de la facturación electrónica ha creado una fuente valiosa de información estructurada sobre ventas y compras, que puede aprovecharse para estos propósitos con el consentimiento de las empresas.

El pronóstico financiero mediante inteligencia artificial enfrenta, no obstante, limitaciones. Los modelos aprenden de los patrones pasados y pueden fallar ante cambios abruptos, como la pérdida de un cliente importante, una crisis económica o una disrupción en la cadena de suministro. Los emprendimientos nuevos carecen de historial suficiente para entrenar modelos complejos, por lo que conviene comenzar con métodos sencillos y referencias sectoriales. Y la precisión de los pronósticos disminuye a medida que se alejan en el tiempo: la evidencia sobre la predicción de la insolvencia muestra que la exactitud de los modelos se reduce cuanto más lejos del evento se realiza la estimación (Hidalgo-Díaz et al., 2025). Por ello, los pronósticos deben actualizarse con frecuencia y complementarse con el juicio de quienes conocen el negocio.

Desde esta perspectiva, un aspecto práctico de estos sistemas es la frecuencia con que se presentan los pronósticos. Para la gestión de la tesorería, resultan útiles las proyecciones semanales a tres meses, que permiten anticipar los períodos de estrechez y programar pagos y cobros. Para la planificación del negocio, conviene disponer de proyecciones mensuales a un año, que

orienten las decisiones de compras, contratación e inversión. Y para las decisiones estratégicas, las proyecciones anuales a tres o cinco años, aunque más inciertas, permiten evaluar la viabilidad de largo plazo. Cada horizonte requiere un nivel de detalle distinto y se apoya en métodos diferentes: los pronósticos de corto plazo pueden basarse en el comportamiento reciente de las series, mientras que los de largo plazo dependen más de supuestos sobre el mercado y la estrategia. Presentar los tres horizontes de manera coherente, con sus respectivos márgenes de incertidumbre, ofrece a la persona emprendedora una visión completa de la trayectoria financiera de su negocio.

Los modelos de inteligencia artificial para pronosticar ingresos, costos, rentabilidad y flujos de efectivo permiten que los emprendimientos gestionen sus finanzas con una mirada hacia adelante. Al integrar los distintos pronósticos en un sistema coherente, alimentado con datos actualizados y que expresa la incertidumbre de manera explícita, ofrecen a las personas emprendedoras una herramienta para anticipar problemas de rentabilidad y de liquidez, planificar sus necesidades de financiamiento y tomar decisiones operativas con mayor fundamento.

Figura 4.2

Modelos de inteligencia artificial para pronosticar ingresos, costos, rentabilidad y flujos de efectivo.



Fuente: Elaboración autores

4.3. Predicción de insolvencia empresarial mediante modelos estadísticos y algoritmos de clasificación

La predicción de la insolvencia empresarial es una de las áreas más antiguas y desarrolladas de la aplicación de modelos cuantitativos a las finanzas. Desde los trabajos pioneros de la segunda mitad del siglo pasado, que combinaron razones financieras mediante análisis discriminante

para distinguir entre empresas solventes y empresas en quiebra, la disciplina ha evolucionado hacia modelos de probabilidad, como los Logit y Probit, y más recientemente hacia algoritmos de aprendizaje automático.

Esta trayectoria refleja una búsqueda constante de mayor precisión y anticipación, motivada por las consecuencias de la insolvencia para acreedores, inversionistas, trabajadores y, en el caso de los emprendimientos, para las propias personas emprendedoras y sus familias. De forma complementaria, la insolvencia puede definirse como la incapacidad de una empresa para cumplir con sus obligaciones financieras a su vencimiento.

En la práctica, los estudios utilizan distintas operacionalizaciones: la declaración formal de quiebra, el inicio de procesos de liquidación o disolución, el incumplimiento de pagos, la clasificación regulatoria como empresa en dificultades o la presencia de indicadores contables que reflejan una situación crítica, como un patrimonio negativo. La elección de la definición influye en los resultados y en la utilidad de los modelos, porque algunas definiciones capturan el problema cuando ya es irreversible, mientras que otras permiten detectarlo en etapas más tempranas.

Los modelos estadísticos tradicionales siguen ocupando un lugar relevante en la predicción de la insolvencia. En esa dirección, Zizi et al. (2021), al comparar la regresión logística con redes neuronales en pequeñas y medianas empresas de Marruecos, obtuvieron los mejores resultados con la regresión logística con selección paso a paso, que alcanzó una exactitud de 93,33 % dos años antes de la dificultad financiera y de 95 % un año antes.

Los autores identificaron la relación entre los gastos por intereses y las ventas, así como la rentabilidad sobre los activos, como los predictores más relevantes, y observaron que los modelos clasificaban con mayor precisión a las empresas en dificultades que a las sanas. Para el Ecuador, los modelos Probit y de riesgos proporcionales de Cox aplicados a datos de la Superintendencia de Compañías, Valores y Seguros han identificado la rentabilidad, las ventas, el patrimonio y el apalancamiento como determinantes del riesgo de cierre (Zambrano et al., 2021; González Cedeño, 2026).

Vinculado con lo anterior, los algoritmos de aprendizaje automático han ampliado las posibilidades de este campo. Sobre este punto, Zhao, Li y Dai (2023) desarrollaron un sistema de alerta temprana de crisis financieras para pequeñas y medianas empresas que, tras seleccionar los indicadores más significativos mediante métodos estadísticos, combinaba varios modelos

mediante una técnica de fusión por apilamiento. XGBoost superó a los modelos de referencia, como las redes neuronales de retropropagación y la regresión logística, y el modelo combinado alcanzó una exactitud de 85,8 % en la identificación de las empresas en situación especial y de 81,9 % en las demás, con un desempeño superior al de los métodos de votación y promedio. Los autores concluyeron que el aprendizaje automático resultaba más práctico, eficiente y preciso que los modelos econométricos tradicionales para esta tarea.

La evidencia más reciente incorpora también arquitecturas de aprendizaje profundo. En este sentido, Hidalgo-Díaz, Alaminos y Delgado-Gómez (2025) analizaron 8.400 empresas de Alemania, Francia, Italia y España entre 2017 y 2022, y encontraron que un modelo que combinaba redes convolucionales y redes recurrentes de memoria de corto y largo plazo ofrecía el mejor desempeño, con una exactitud cercana al 98 % un año antes de la insolvencia, de alrededor de 97 % dos años antes y de cerca de 95 % tres años antes.

El estudio mostró, además, que la liquidez, el margen de beneficio neto, la rentabilidad sobre los activos y la capacidad de pago de la deuda eran los indicadores más significativos en el horizonte de un año, mientras que en el horizonte de tres años solo los indicadores de rentabilidad mantenían su significancia.

No obstante, estos resultados deben interpretarse con cautela. Las exactitudes muy elevadas obtenidas en algunos estudios pueden depender de las características de las muestras, de la definición de insolvencia, del equilibrio entre empresas solventes e insolventes en los datos de evaluación y de las métricas utilizadas. Como se examinó en el segundo capítulo, la exactitud puede resultar engañosa cuando las clases están desequilibradas, y medidas como el coeficiente de correlación de Matthews ofrecen una evaluación más exigente (Chicco y Jurman, 2023). Además, los modelos entrenados con datos de empresas medianas y grandes de economías europeas o asiáticas pueden no aplicarse directamente a los emprendimientos latinoamericanos, que difieren en tamaño, estructura financiera, disponibilidad de información y entorno económico.

Una lección recurrente de la literatura es que la capacidad predictiva disminuye a medida que se amplía el horizonte de predicción. Predecir la insolvencia un año antes es más fácil que predecirla tres años antes, porque las señales de deterioro se vuelven más evidentes a medida que el evento se aproxima. Sin embargo, la utilidad de la predicción aumenta con la anticipación, porque deja más tiempo para actuar. Existe, por tanto, una tensión entre precisión y anticipación que debe resolverse según el propósito: un sistema de alerta destinado a ofrecer apoyo

preventivo puede aceptar una menor precisión a cambio de una mayor anticipación, mientras que un sistema de provisiones crediticias puede privilegiar la precisión en horizontes cortos.

En concordancia con ello, los indicadores que la literatura identifica de manera consistente como predictores de la insolvencia pueden agruparse en varias familias. Los de rentabilidad, como la rentabilidad sobre los activos y el margen de beneficio neto, reflejan la capacidad del negocio para generar excedentes y suelen mantener su poder predictivo en horizontes más largos. Los de liquidez, como la razón corriente o la prueba ácida, indican la capacidad para cubrir las obligaciones de corto plazo.

Los de endeudamiento y cobertura, como el apalancamiento, la relación entre intereses y ventas o la capacidad de pago de la deuda, miden la carga financiera. Y los de actividad, como la rotación de cartera o de inventarios, reflejan la eficiencia en la gestión del capital de trabajo. La combinación de estas familias en un modelo permite capturar las distintas vías por las que una empresa puede llegar a la insolvencia.

Para los emprendimientos, la predicción de la insolvencia enfrenta una dificultad adicional: la escasez de información financiera confiable. Muchos pequeños negocios no elaboran estados financieros formales, o los elaboran solo con fines tributarios y con retraso. En estos casos, las fuentes alternativas de información adquieren relevancia: los movimientos de las cuentas bancarias, las transacciones con medios de pago electrónicos, los registros de facturación electrónica o los datos de las plataformas de comercio.

Estas fuentes permiten construir indicadores de flujo de efectivo, de estabilidad de los ingresos y de cumplimiento de pagos que pueden sustituir parcialmente a las razones contables tradicionales. A ello se suma que la predicción de la insolvencia tiene aplicaciones para distintos actores. Para las entidades financieras, permite evaluar el riesgo de los créditos y fijar provisiones adecuadas. Para los proveedores, ayuda a decidir las condiciones comerciales que ofrecen a sus clientes. Para los inversionistas, contribuye a evaluar el riesgo de sus participaciones. Para los gobiernos y los organismos de apoyo, permite identificar a las empresas en riesgo y ofrecerles acompañamiento antes de que la situación sea irreversible. Y para los propios emprendedores, ofrece una herramienta de autodiagnóstico que señala las áreas de su gestión financiera que requieren atención.

La interpretación de los modelos de insolvencia exige distinguir entre las señales de deterioro y sus causas. Un modelo puede detectar que la liquidez de una empresa está disminuyendo y

que, por ello, su probabilidad de insolvencia aumenta; pero la caída de la liquidez puede obedecer a causas muy distintas, como la pérdida de un cliente importante, un aumento de los costos, una inversión mal planificada o simplemente el crecimiento de las ventas a crédito.

Las técnicas de explicabilidad ayudan a identificar qué indicadores contribuyen al riesgo, pero el diagnóstico de las causas y la definición de las medidas correctivas requieren el conocimiento del negocio que solo tienen sus responsables o sus asesores. Por ello, los modelos de insolvencia funcionan mejor como herramientas de alerta y de orientación del análisis que como sustitutos del diagnóstico. Al respecto, Zizi et al. (2021) observaron, además, que los modelos tendían a clasificar mejor a las empresas en dificultades que a las sanas, lo que refuerza la necesidad de interpretar sus resultados con criterio.

Asimismo, la evaluación del costo de los errores resulta especialmente relevante en este campo. Clasificar como solvente a una empresa que terminará en insolvencia puede conducir a pérdidas para sus acreedores; clasificar como insolvente a una empresa viable puede privarla del crédito que necesitaba para sobrevivir y convertir la predicción en una profecía autocumplida. El peso relativo de ambos errores depende de quién utiliza el modelo y con qué fin, y debe reflejarse en la elección del umbral de decisión. Para los programas de apoyo preventivo, cuyo propósito es acompañar a las empresas en riesgo, el costo de una falsa alarma es relativamente bajo, mientras que el de una alerta omitida puede ser muy alto, lo que justifica umbrales más sensibles.

Desde la perspectiva del emprendimiento predictivo, la predicción de la insolvencia empresarial combina una tradición estadística consolidada con las capacidades más recientes del aprendizaje automático y del aprendizaje profundo. La evidencia muestra que estos modelos pueden anticipar con notable precisión el deterioro financiero de las empresas, especialmente en horizontes cortos, y que los indicadores de rentabilidad conservan su poder predictivo en horizontes más largos. Su aplicación a los emprendimientos latinoamericanos exige adaptar los modelos a la información disponible, validar su desempeño con métricas adecuadas y utilizarlos como herramientas de prevención más que como instrumentos de exclusión.

Figura 4.3

Predicción de insolvencia empresarial mediante modelos estadísticos y algoritmos de clasificación.



Fuente: Elaboración autores

4.4. Simulación Monte Carlo y análisis probabilístico de escenarios financieros emprendedores

Cuando una persona emprendedora prepara la proyección financiera de su negocio, suele construir un escenario único: venderá tantas unidades a tal precio, con tales costos, y obtendrá tal utilidad. En el mejor de los casos, añade un escenario optimista y otro pesimista. Este enfoque, aunque útil, ofrece una visión muy limitada de la incertidumbre. En la realidad, las ventas, los precios, los costos y los plazos de cobro pueden tomar una infinidad de valores, que se combinan de maneras diversas para producir resultados muy distintos. La simulación Monte Carlo permite explorar esa diversidad de manera sistemática, generando miles de escenarios posibles y analizando la distribución de los resultados.

Además, la simulación Monte Carlo es un método computacional que utiliza números aleatorios para modelar sistemas con incertidumbre. Su lógica es sencilla. En primer lugar, se identifican las variables inciertas del modelo financiero —por ejemplo, el volumen de ventas, el precio, el costo de los insumos, la tasa de crecimiento— y se asigna a cada una una distribución de probabilidad que describe los valores que puede tomar y su probabilidad. En segundo lugar, se generan aleatoriamente valores para cada variable según su distribución, y se calcula el resultado del modelo, como el valor actual neto o el saldo de caja al final del año. En tercer lugar, se repite este proceso miles de veces, obteniendo miles de resultados posibles.

Finalmente, se analiza la distribución de esos resultados para estimar su valor esperado, su variabilidad y la probabilidad de que ocurran determinados eventos, como pérdidas o déficits de liquidez.

Fabianová, Janeková, Fedorko y Molnár (2023) propusieron una metodología integral para evaluar proyectos de inversión mediante simulación Monte Carlo. Su enfoque comienza con el pronóstico de los factores inciertos, la selección de los factores de riesgo y la definición de su incertidumbre; continúa con la construcción de un modelo matemático que calcula el criterio de eficiencia económica del proyecto; utiliza la simulación para pronosticar ese criterio; evalúa el riesgo a partir de los resultados; y, si el riesgo resulta inaceptable, propone formas de optimizar el proyecto.

Esta secuencia ofrece un marco aplicable a la evaluación de emprendimientos, en los que la incertidumbre sobre las variables clave es especialmente elevada. En esta misma línea, la simulación Monte Carlo aporta varias ventajas frente al análisis de escenarios tradicional. Permite considerar simultáneamente la incertidumbre de muchas variables, en lugar de modificarlas de una en una. Captura las correlaciones entre variables, como la relación entre el precio y el volumen de ventas, o entre los costos de distintos insumos.

Produce una distribución completa de los resultados, que permite responder preguntas como cuál es la probabilidad de que el proyecto pierda dinero, cuál es la pérdida máxima esperada con un 95 % de confianza o cuántos meses de operación son necesarios para que la probabilidad de alcanzar el punto de equilibrio supere el 80 %. Y facilita el análisis de sensibilidad, al identificar qué variables contribuyen más a la variabilidad de los resultados.

La definición de las distribuciones de probabilidad es el paso más delicado de la simulación. Cuando existen datos históricos, las distribuciones pueden estimarse a partir de ellos; cuando no existen, como ocurre con frecuencia en los emprendimientos, pueden construirse a partir de datos de empresas comparables, de información sectorial o del juicio estructurado de personas expertas.

Las distribuciones triangulares, definidas por un valor mínimo, uno más probable y uno máximo, resultan útiles cuando la información es limitada, porque son fáciles de comunicar y de estimar. Las distribuciones normales, lognormales o uniformes se utilizan cuando se dispone de más información sobre el comportamiento de las variables. De igual manera, la combinación de la simulación Monte Carlo con los modelos de inteligencia artificial amplía sus posibilidades.

Los modelos de pronóstico examinados en el capítulo anterior producen no solo estimaciones puntuales, sino también distribuciones de los valores futuros, que pueden utilizarse directamente como insumos de la simulación. En esa dirección, Hyndman y Athanasopoulos (2021) explican cómo pueden generarse trayectorias futuras simuladas de una serie temporal a partir de un modelo ajustado, lo que permite obtener intervalos de pronóstico incluso cuando las distribuciones de los errores no son normales. Integrar estas trayectorias simuladas de ventas en un modelo financiero completo permite propagar la incertidumbre de la demanda hacia la rentabilidad y los flujos de efectivo.

Un ejemplo ilustra la utilidad del método. Supóngase que una persona evalúa la apertura de una cafetería. Sus ventas mensuales dependen del número de clientes y del gasto promedio por cliente; sus costos, del precio de los insumos, del arriendo y de los salarios. Con base en información de cafeterías similares y en su conocimiento del barrio, define distribuciones para cada variable.

La simulación de diez mil escenarios podría mostrar, por ejemplo, que el valor actual neto esperado es positivo, pero que existe una probabilidad considerable de que el negocio no alcance el punto de equilibrio durante el primer año, y que la variable que más influye en el resultado es el número de clientes diarios. Esta información permite tomar decisiones más informadas: negociar un período de gracia en el arriendo, prever un fondo de reserva para los primeros meses o invertir en actividades que aumenten la afluencia de clientes.

Con todo, la simulación Monte Carlo presenta también limitaciones. Sus resultados dependen de la calidad de las distribuciones asignadas: si estas no reflejan adecuadamente la incertidumbre real, la simulación producirá una falsa sensación de precisión. El modelo financiero subyacente debe ser correcto, porque la simulación no corrige sus errores lógicos. Las correlaciones entre variables deben especificarse adecuadamente, porque ignorarlas puede subestimar o sobrestimar el riesgo. Y los eventos extremos, poco frecuentes pero de gran impacto, como una pandemia o una crisis financiera, pueden quedar fuera de las distribuciones estimadas a partir de datos de períodos normales. Por ello, la simulación debe complementarse con el análisis de escenarios extremos específicos.

La difusión de herramientas informáticas ha hecho la simulación Monte Carlo accesible para los emprendimientos. Existen complementos para hojas de cálculo y bibliotecas de programación de código abierto que permiten implementar simulaciones sin grandes inversiones. Los

programas de apoyo al emprendimiento, las incubadoras y las entidades financieras podrían incorporar estas herramientas en sus procesos de evaluación, ofreciendo a los emprendedores un análisis más completo del riesgo de sus proyectos.

La comunicación de los resultados merece atención especial: gráficos que muestren la distribución de los resultados posibles y probabilidades expresadas en términos sencillos facilitan la comprensión por parte de personas sin formación estadística. Ahora bien, en el contexto de los emprendimientos ecuatorianos, donde la falta de rentabilidad y las dificultades de financiamiento figuran entre las principales causas de cierre (Lasio et al., 2024), la simulación Monte Carlo podría contribuir a que las personas emprendedoras evalúen con mayor realismo la viabilidad de sus proyectos antes de comprometer sus ahorros o endeudarse.

Conocer la probabilidad de no alcanzar el punto de equilibrio en los primeros meses, o la magnitud del capital de trabajo necesario para resistir un período de ventas bajas, permite preparar respuestas antes de que los problemas se presenten.

Una extensión natural de la simulación consiste en utilizarla para comparar decisiones. En lugar de evaluar un único proyecto, la persona emprendedora puede simular varias alternativas —por ejemplo, alquilar un local más grande o uno más pequeño, contratar personal a tiempo completo o por horas, financiarse con deuda o con aportes propios— y comparar las distribuciones de resultados de cada una. Esta comparación permite elegir no solo la alternativa con mayor valor esperado, sino la que ofrece el mejor equilibrio entre rentabilidad y riesgo de acuerdo con la situación y las preferencias de quien decide.

Una persona con pocos ahorros y responsabilidades familiares puede preferir una alternativa con menor rentabilidad esperada, pero con menor probabilidad de pérdidas graves, mientras que otra con mayor respaldo puede asumir más riesgo. La simulación hace visible este intercambio y permite tomar decisiones coherentes con la tolerancia al riesgo de cada persona (Fabianová et al., 2023). Puede sostenerse, así, que la simulación Monte Carlo y el análisis probabilístico de escenarios transforman la manera en que se evalúa el riesgo de los emprendimientos. En lugar de un escenario único o de unos pocos escenarios arbitrarios, ofrecen una visión completa de los resultados posibles y de sus probabilidades.

Figura 4.4

Simulación Monte Carlo y análisis probabilístico de escenarios financieros emprendedores.



Fuente: Elaboración autores

4.5. Gemelos digitales empresariales: representación virtual del emprendimiento para experimentar decisiones estratégicas

El concepto de gemelo digital nació en la ingeniería y la manufactura, donde designa una réplica virtual de un objeto físico —una turbina, una máquina, una línea de producción— que se alimenta en tiempo real con datos de sensores y permite monitorear su estado, anticipar fallas y probar modificaciones sin intervenir el objeto real. En los últimos años, la idea se ha extendido más allá de los objetos físicos hacia los procesos de negocio y las organizaciones completas.

Esta extensión abre una posibilidad atractiva para el emprendimiento predictivo: construir una representación virtual de la empresa en la que puedan ensayarse decisiones estratégicas antes de ejecutarlas en la realidad. Por otra parte, Wurm et al. (2023) describen los gemelos digitales de organizaciones como modelos computacionales que se alimentan de flujos de datos operativos para seguir, evaluar y mejorar el desempeño organizacional.

Los autores advierten que, pese al interés creciente de la industria y de las firmas de consultoría, la investigación académica sobre este tipo de gemelos es todavía escasa, y proponen una agenda que vincula el concepto con la teoría organizacional y con los métodos de simulación de sistemas sociales. Sobre este punto, Fornari et al. (2024), en un manifiesto de investigación sobre los gemelos digitales de procesos de negocio, los definen como réplicas virtuales de los procesos que se actualizan en tiempo real a partir de los datos generados por su ejecución, y señalan

entre sus usos la supervisión continua, la predicción del comportamiento futuro y la evaluación de alternativas de rediseño.

Un gemelo digital empresarial se distingue de un modelo financiero convencional en varios aspectos. Mientras la hoja de cálculo tradicional se construye una vez a partir de supuestos y se actualiza esporádicamente, el gemelo se conecta de manera continua con los sistemas de información de la empresa —ventas, inventarios, compras, nómina, cuentas bancarias— y refleja su estado actual. Mientras el modelo financiero suele representar la empresa de manera agregada, el gemelo puede reproducir con mayor detalle sus procesos, recursos y relaciones: cuántos pedidos ingresan, cuánto tarda su preparación, qué capacidad tiene cada recurso, cómo responden los clientes a los cambios de precio. Y mientras el modelo convencional proyecta un futuro, el gemelo permite simular muchos futuros alternativos en función de las decisiones que se tomen.

Paralelamente, la arquitectura de un gemelo digital empresarial comprende varios componentes. El primero de ellos, la capa de datos, conecta el gemelo con las fuentes de información de la empresa y del entorno. Sobre ella se sitúa un modelo de la organización que representa sus procesos, recursos, restricciones y relaciones causales; puede construirse mediante técnicas de simulación de eventos discretos, de dinámica de sistemas o de modelos basados en agentes, según la naturaleza del negocio.

Los modelos predictivos examinados en este libro —de demanda, de comportamiento del cliente, de riesgo financiero— se integran en el gemelo para estimar las respuestas del entorno a las decisiones. Una interfaz de experimentación permite a las personas usuarias plantear preguntas del tipo «qué pasaría si» y observar sus consecuencias simuladas. Por último, un mecanismo de calibración compara periódicamente las predicciones del gemelo con lo que efectivamente ocurre y ajusta el modelo para mantener su fidelidad.

Las decisiones que pueden ensayarse en un gemelo digital empresarial son de distinto alcance. En el plano operativo, el gemelo permite evaluar el efecto de contratar a una persona adicional, ampliar el horario de atención o modificar la política de inventarios. En el comercial, puede simular las consecuencias de un cambio de precios, del lanzamiento de un nuevo producto o de la apertura de un canal de venta en línea.

En el financiero, permite anticipar cómo afectaría a la liquidez un aumento de las ventas a crédito, la contratación de un préstamo o el retraso en los pagos de un cliente importante. Y en el estratégico, puede ayudar a evaluar la apertura de una nueva sucursal, la entrada a otro

mercado o la incorporación de un socio. Sobre esta base, la combinación del gemelo digital con la simulación Monte Carlo resulta especialmente valiosa.

Cada decisión ensayada en el gemelo no produce un único resultado, sino una distribución de resultados posibles que refleja la incertidumbre de la demanda, los precios y los comportamientos. Así, la comparación entre alternativas puede basarse no solo en el resultado esperado, sino también en el riesgo de cada una: una decisión puede ofrecer un beneficio esperado algo menor, pero con una probabilidad mucho más baja de generar una crisis de liquidez. Esta capacidad de comparar alternativas en términos de riesgo y rentabilidad resulta difícil de lograr con los métodos tradicionales.

Para los emprendimientos, la construcción de un gemelo digital completo puede parecer fuera de alcance. Las aplicaciones industriales más conocidas se han desarrollado en grandes empresas, con inversiones considerables en sensores, plataformas y personal especializado. Sin embargo, la lógica del gemelo puede aplicarse de manera escalonada. Un pequeño negocio puede comenzar con un modelo sencillo que replique sus flujos principales —ventas, costos, cobros y pagos— conectado a su sistema de facturación electrónica y a su cuenta bancaria, y enriquecerlo progresivamente con modelos de demanda, de inventarios o de comportamiento de clientes a medida que crece su madurez digital.

Las plataformas de gestión empresarial en la nube, cada vez más accesibles, facilitan esta integración. Desde esta perspectiva, los gemelos digitales tienen también aplicaciones en el ecosistema emprendedor. Las incubadoras y aceleradoras podrían utilizar gemelos simplificados para acompañar a sus emprendimientos, ensayando con ellos distintas estrategias antes de su implementación. Las entidades financieras podrían emplearlos para evaluar cómo responderían los flujos de efectivo de un negocio ante distintos esquemas de financiamiento. Y los programas públicos de apoyo podrían construir gemelos sectoriales que representen el funcionamiento típico de las pequeñas empresas de un sector, útiles para evaluar el posible efecto de políticas o de choques externos sobre su supervivencia.

La utilización de gemelos digitales exige, no obstante, reconocer sus limitaciones. Todo gemelo es una simplificación de la realidad, y su utilidad depende de que capture los aspectos relevantes para las decisiones que se pretende ensayar. Las relaciones causales incorporadas en el modelo —por ejemplo, cuánto aumentan las ventas ante una reducción de precios— suelen ser inciertas

y deben estimarse con cuidado, porque un supuesto erróneo puede conducir a recomendaciones equivocadas.

El comportamiento humano de clientes, competidores y colaboradores es difícil de modelar, y los cambios estructurales del entorno pueden volver obsoleto un gemelo calibrado con datos del pasado. En este sentido, Wurm et al. (2023) subrayan, precisamente, la necesidad de mayor investigación sobre la validez y los límites de estos modelos en contextos organizacionales. De forma complementaria, la gestión de los datos que alimentan el gemelo plantea, además, exigencias de seguridad y de protección de la información.

Un gemelo digital concentra información sensible sobre clientes, trabajadores, proveedores y finanzas de la empresa, por lo que su diseño debe incorporar medidas de seguridad adecuadas y cumplir con la normativa de protección de datos personales vigente, como la Ley Orgánica de Protección de Datos Personales en el Ecuador (Asamblea Nacional del Ecuador, 2021).

La relación entre los gemelos digitales y la experimentación emprendedora merece un comentario adicional. Las metodologías de emprendimiento ágil recomiendan validar las hipótesis del modelo de negocio mediante experimentos rápidos y de bajo costo con clientes reales. El gemelo digital no sustituye esa validación, porque sus resultados dependen de supuestos que solo el mercado puede confirmar; pero puede ayudar a decidir qué experimentos realizar, al identificar las hipótesis cuya incertidumbre tiene mayor efecto sobre los resultados.

Si la simulación muestra que la viabilidad del negocio depende críticamente de la tasa de recompra de los clientes, el experimento más valioso será el que permita estimar esa tasa con mayor precisión. De este modo, el gemelo y la experimentación se complementan: el primero orienta la búsqueda de información, y la segunda aporta los datos que permiten calibrar el gemelo. Al respecto, Fornari et al. (2024) destacan, precisamente, la sincronización continua entre la réplica virtual y el proceso real como rasgo distintivo de estos sistemas.

De lo anterior se desprende que los gemelos digitales empresariales ofrecen un laboratorio en el que las personas emprendedoras pueden experimentar con sus decisiones sin arriesgar su negocio real. Al combinar datos actualizados, modelos de los procesos de la empresa, predicciones sobre el entorno y simulación de la incertidumbre, permiten comparar alternativas en términos de rentabilidad y riesgo antes de comprometer recursos. Su adopción progresiva, desde modelos sencillos hasta representaciones más completas, puede convertir la experimentación virtual en una práctica habitual de la gestión emprendedora.

Figura 4.5

Gemelos digitales empresariales: representación virtual del emprendimiento para experimentar decisiones estratégicas.



Fuente: Elaboración autores

4.6. Modelos predictivos de liquidez, endeudamiento y necesidades futuras de financiamiento

La liquidez es la condición más inmediata de la supervivencia empresarial. Un negocio puede sobrevivir durante un tiempo sin utilidades, pero no puede hacerlo sin efectivo para pagar a sus trabajadores, proveedores y acreedores. Por ello, una parte considerable de los cierres de emprendimientos no responde a la inviabilidad del modelo de negocio, sino a crisis de liquidez que podrían haberse anticipado y gestionado con mayor tiempo. Los modelos predictivos de liquidez, endeudamiento y necesidades de financiamiento buscan precisamente ofrecer esa anticipación: estimar con antelación cuándo el efectivo será insuficiente, cuánto endeudamiento puede soportar la empresa y qué financiamiento necesitará para sostener su operación y su crecimiento.

Vinculado con lo anterior, la predicción de la liquidez parte del pronóstico de los flujos de efectivo examinado anteriormente, pero se concentra en el saldo disponible y en su evolución período a período. Su resultado central es una proyección del saldo de caja, acompañada de una estimación de su incertidumbre.

A partir de ella pueden calcularse indicadores como la probabilidad de que el saldo caiga por debajo de un umbral mínimo en las próximas semanas o meses, el número de días de operación que el efectivo disponible permitiría cubrir sin nuevos ingresos o el monto de financiamiento

de corto plazo que sería necesario para atravesar un período de estrechez. Estos indicadores resultan más útiles para la gestión diaria que las razones de liquidez tradicionales, que describen la situación en un momento dado sin proyectarla.

La evidencia sobre la predicción de dificultades financieras confirma la importancia de la liquidez como señal de alerta. En esa dirección, Hidalgo-Díaz, Alaminos y Delgado-Gómez (2025) encontraron que la liquidez figuraba entre los indicadores más significativos para anticipar la insolvencia en el horizonte de un año, junto con el margen de beneficio neto, la rentabilidad sobre los activos y la capacidad de pago de la deuda.

Su pérdida de significancia en horizontes más largos sugiere que la liquidez actúa como una señal de deterioro inmediato, mientras que la rentabilidad refleja las condiciones estructurales que, con el tiempo, terminan afectando la disponibilidad de efectivo. Un sistema de gestión financiera predictiva debe, por tanto, vigilar ambas dimensiones: la liquidez para el corto plazo y la rentabilidad para el mediano. En concordancia con ello, el endeudamiento es la otra cara de la gestión financiera.

Los emprendimientos recurren a la deuda para financiar su inversión inicial, su capital de trabajo o su crecimiento, y una deuda bien utilizada puede acelerar su desarrollo. Sin embargo, el exceso de endeudamiento aumenta el riesgo de cierre. En el ámbito ecuatoriano, el análisis de supervivencia aplicado a empresas registradas en la Superintendencia de Compañías, Valores y Seguros mostró que el apalancamiento financiero elevaba el riesgo de cierre, con una razón de riesgo superior a uno, mientras que la rentabilidad sobre los activos lo reducía (González Cedeño, 2026). Sobre este punto, Zambrano et al. (2021) encontraron, en la misma línea, que el equilibrio del endeudamiento contribuía a prolongar la vida de las empresas. La relación entre los gastos por intereses y las ventas, identificada por Zizi et al. (2021) como uno de los predictores más relevantes de las dificultades financieras, ofrece un indicador sencillo de la carga que la deuda impone sobre la actividad del negocio.

Los modelos predictivos del endeudamiento permiten estimar la capacidad de deuda de un emprendimiento, es decir, el monto máximo de obligaciones que puede asumir sin comprometer su liquidez ni su solvencia. Esta capacidad no es fija, sino que depende de los flujos de efectivo esperados y de su variabilidad. Un negocio con ingresos estables puede soportar una carga de deuda mayor que otro con ingresos volátiles, aunque ambos tengan el mismo nivel promedio de ventas.

La simulación de los flujos de efectivo bajo distintos escenarios permite estimar la probabilidad de que, con un determinado nivel de deuda, la empresa no pueda cumplir con sus pagos en algún momento, y ajustar el monto y el plazo del financiamiento para mantener esa probabilidad dentro de límites aceptables. A ello se suma que la estimación de las necesidades futuras de financiamiento integra los pronósticos de liquidez y la planificación del crecimiento.

Un emprendimiento que crece necesita financiar el aumento de su capital de trabajo —más inventarios, más cuentas por cobrar— y, con frecuencia, nuevas inversiones en equipos, locales o tecnología. Paradójicamente, el crecimiento rápido puede generar crisis de liquidez en negocios rentables, si el efectivo necesario para financiarlo llega después de que deba pagarse. Los modelos predictivos permiten anticipar estas necesidades con meses de antelación, lo que facilita la búsqueda de financiamiento en condiciones más favorables que las que se obtienen cuando la necesidad es urgente.

La anticipación de las necesidades de financiamiento tiene particular importancia en contextos donde el acceso al crédito es limitado. El informe del Global Entrepreneurship Monitor para el Ecuador situó las dificultades de financiamiento como la segunda causa de cierre de los negocios y el apoyo financiero entre las condiciones más débiles del entorno emprendedor (Lasio et al., 2024). En estas condiciones, la capacidad de prever las necesidades con tiempo y de presentar a los financiadores proyecciones sustentadas en datos puede marcar la diferencia entre obtener el crédito o no.

Una solicitud acompañada de pronósticos de flujos de efectivo, de una estimación del riesgo de impago y de un análisis de la capacidad de deuda ofrece a la entidad financiera una base más sólida para su decisión. Asimismo, la diversidad de fuentes de financiamiento disponibles para los emprendimientos amplía las decisiones que deben tomarse. Además del crédito bancario tradicional, las personas emprendedoras pueden recurrir al financiamiento de proveedores, a las cooperativas de ahorro y crédito, a las plataformas de financiamiento colaborativo, a inversionistas ángeles o a fondos de capital de riesgo, cada uno con costos, plazos, garantías y condiciones distintas.

Los modelos predictivos pueden ayudar a evaluar qué combinación de fuentes resulta más adecuada en cada momento, considerando no solo su costo, sino también su efecto sobre el riesgo de liquidez y la flexibilidad de la empresa. La Ley Orgánica de Emprendimiento e Innovación

reconoce varias de estas alternativas, entre ellas los mecanismos de financiamiento colaborativo (Asamblea Nacional del Ecuador, 2020).

El aprendizaje automático puede contribuir a mejorar cada componente de estos modelos. Los algoritmos de clasificación permiten estimar la probabilidad de que cada cliente pague con retraso, lo que mejora el pronóstico de los cobros. Los modelos de series temporales y de potenciación de gradiente permiten pronosticar las ventas con mayor precisión.

Los modelos de supervivencia permiten estimar el riesgo de cierre asociado a distintos niveles de endeudamiento. Y la explicabilidad de los modelos, mediante técnicas como los valores SHAP, permite identificar qué factores impulsan las proyecciones de liquidez y qué palancas podrían utilizarse para mejorarlas (Bussmann et al., 2021). No obstante, estos modelos deben utilizarse con prudencia. Las proyecciones de liquidez dependen de supuestos sobre el comportamiento de clientes y proveedores que pueden cambiar bruscamente en situaciones de crisis, cuando los clientes retrasan sus pagos y los proveedores endurecen sus condiciones al mismo tiempo.

La experiencia de crisis recientes muestra que los choques sistémicos pueden afectar simultáneamente a muchas empresas y desbordar las previsiones basadas en condiciones normales. Por ello, los modelos de liquidez deben complementarse con pruebas de estrés que evalúen la resistencia del negocio ante escenarios adversos extremos, y con políticas prudentes de reservas de efectivo.

Las cooperativas de ahorro y crédito y las entidades de microfinanzas desempeñan un papel relevante en el financiamiento de los emprendimientos en el Ecuador y en otros países de la región, especialmente para los negocios de menor tamaño y para la economía popular y solidaria. Estas entidades suelen tener un conocimiento cercano de sus socios y clientes, pero disponen de menos recursos para desarrollar modelos sofisticados de evaluación del riesgo.

La incorporación de modelos predictivos de liquidez y capacidad de pago, construidos con los datos de las propias entidades y validados cuidadosamente, podría mejorar la asignación de sus recursos y reducir la morosidad, sin perder la relación de proximidad que constituye su ventaja. La cooperación entre estas entidades, las universidades y los organismos de supervisión para desarrollar herramientas compartidas permitiría distribuir los costos y aprovechar mejor la información disponible, siempre con las garantías de protección de datos que exige la normativa (Asamblea Nacional del Ecuador, 2021).

Además, las pruebas de estrés constituyen una herramienta complementaria de especial utilidad. Consisten en evaluar cómo se comportaría la liquidez del negocio ante escenarios adversos específicos y plausibles, como una caída abrupta de las ventas durante varios meses, el retraso simultáneo de los pagos de los principales clientes, un aumento significativo del costo de los insumos o el encarecimiento del crédito. A diferencia de la simulación de escenarios aleatorios, las pruebas de estrés se concentran en situaciones extremas que, aunque poco probables, tendrían consecuencias graves. Sus resultados permiten dimensionar las reservas de liquidez necesarias, negociar líneas de crédito contingentes y preparar planes de acción para el caso de que esas situaciones se presenten.

Considerado en su conjunto, el panorama muestra que los modelos predictivos de liquidez, endeudamiento y necesidades de financiamiento permiten que los emprendimientos gestionen su supervivencia financiera con anticipación. Al proyectar el saldo de caja y su incertidumbre, estimar la capacidad de deuda y prever las necesidades futuras de financiamiento, ofrecen a las personas emprendedoras tiempo para actuar y a los financiadores información para decidir. En economías donde la falta de liquidez y las dificultades de financiamiento figuran entre las principales causas de cierre de los nuevos negocios, esta capacidad de anticipación constituye un componente central del emprendimiento predictivo.

Figura 4.6

Modelos predictivos de liquidez, endeudamiento y necesidades futuras de financiamiento.



Fuente: Elaboración autores

4.7. Optimización de inversiones emprendedoras mediante algoritmos evolutivos y aprendizaje por refuerzo

Las decisiones de inversión son, para cualquier emprendimiento, decisiones de asignación de recursos escasos entre alternativas competidoras. Una persona emprendedora con un capital limitado debe decidir cuánto destinar a inventario, a equipos, a publicidad, a la contratación de personal o a una reserva de liquidez. Un inversionista que apoya a varios emprendimientos debe decidir cómo distribuir sus fondos entre ellos. Un fondo público de fomento debe elegir qué proyectos financiar entre muchas solicitudes. En todos estos casos, el número de combinaciones posibles crece rápidamente con el número de alternativas, y las restricciones presupuestarias, de riesgo o de diversificación vuelven el problema difícil de resolver con los métodos tradicionales.

Los algoritmos evolutivos y el aprendizaje por refuerzo ofrecen herramientas para abordar esta complejidad. Con todo, el problema de la optimización de carteras de inversión ha sido estudiado ampliamente desde la teoría moderna de carteras, que propone elegir la combinación de activos que maximice el rendimiento esperado para un nivel dado de riesgo. En su formulación básica, este problema puede resolverse con métodos matemáticos convencionales. Sin embargo, cuando se incorporan restricciones realistas, como límites al monto invertido en cada activo o al número de activos de la cartera, el problema se vuelve computacionalmente muy difícil. En este sentido, Erwin y Engelbrecht (2023), en una revisión de más de ciento cuarenta trabajos sobre metaheurísticas aplicadas a la optimización de carteras, explican que las restricciones de cardinalidad y de límites transforman el problema en uno de complejidad no polinómica, para el cual los algoritmos evolutivos y de inteligencia de enjambre se han consolidado como alternativas eficaces.

Los algoritmos evolutivos se inspiran en los mecanismos de la evolución biológica. Parten de una población de soluciones candidatas —por ejemplo, distintas formas de distribuir el presupuesto entre las alternativas de inversión— y las evalúan según un criterio de calidad, como la rentabilidad ajustada por riesgo. Las mejores soluciones se seleccionan y se combinan entre sí para generar nuevas soluciones, a las que se introducen pequeñas variaciones aleatorias. Tras muchas generaciones, la población evoluciona hacia soluciones de alta calidad.

Los algoritmos de inteligencia de enjambre, como la optimización por enjambre de partículas, siguen una lógica similar, inspirada en el comportamiento colectivo de bandadas de aves o colonias de insectos. Estos métodos no garantizan encontrar la solución óptima, pero suelen hallar

soluciones muy buenas en tiempos razonables para problemas donde los métodos exactos resultan impracticables.

En esta misma línea, una ventaja importante de los algoritmos evolutivos es su capacidad para abordar problemas con múltiples objetivos. En la realidad, las decisiones de inversión rara vez persiguen un único fin: se busca rentabilidad, pero también liquidez, diversificación, impacto social o ambiental y coherencia con la estrategia de la empresa. Los algoritmos evolutivos multiobjetivo generan un conjunto de soluciones que representan distintos equilibrios entre los objetivos, conocido como frente de Pareto, y permiten a quienes deciden elegir entre ellas según sus preferencias. Al respecto, Erwin y Engelbrecht (2023) identifican los enfoques multiobjetivo, las carteras dinámicas y la incorporación de precios pronosticados como direcciones prometedoras de la investigación en este campo.

El aprendizaje por refuerzo ofrece una perspectiva complementaria. A diferencia de los métodos que buscan la mejor decisión en un momento dado, el aprendizaje por refuerzo aborda problemas de decisión secuencial, en los que un agente toma decisiones repetidas en un entorno cambiante y aprende, mediante la experiencia, qué acciones conducen a mejores resultados a largo plazo.

El agente observa el estado del entorno, elige una acción, recibe una recompensa y observa el nuevo estado; a partir de muchas interacciones, aprende una política que indica qué acción tomar en cada situación. En esa dirección, Hambly, Xu y Yang (2023), en una revisión de los desarrollos recientes del aprendizaje por refuerzo en las finanzas, señalan que estos métodos dependen de menos supuestos sobre el modelo del entorno que los enfoques tradicionales, y documentan aplicaciones en la ejecución óptima de órdenes, la optimización de carteras, la valoración y cobertura de opciones, la creación de mercado, el enrutamiento de órdenes y el asesoramiento financiero automatizado.

De igual manera, en el ámbito del emprendimiento, el aprendizaje por refuerzo puede aplicarse a decisiones secuenciales como la gestión de inventarios, la fijación dinámica de precios, la asignación del presupuesto de publicidad entre canales o la gestión de la tesorería. Por ejemplo, un agente podría aprender, a partir de datos simulados en un gemelo digital del negocio, cuánto inventario pedir cada semana para minimizar los costos de almacenamiento y de faltantes, o cómo distribuir la inversión publicitaria entre redes sociales, buscadores y otros medios según la respuesta observada de los clientes.

La combinación con los gemelos digitales es especialmente útil, porque permite que el agente aprenda en un entorno simulado sin arriesgar el negocio real.

Para los inversionistas y fondos que apoyan emprendimientos, estas herramientas pueden contribuir a construir carteras más equilibradas. La inversión en empresas emergentes se caracteriza por una alta incertidumbre y por distribuciones de resultados muy asimétricas, en las que unos pocos éxitos compensan muchos fracasos. Los modelos predictivos de éxito examinados en el segundo capítulo pueden estimar la probabilidad de éxito de cada emprendimiento, y los algoritmos de optimización pueden utilizar esas estimaciones para seleccionar una cartera que equilibre el rendimiento esperado, el riesgo y la diversificación por sector, etapa o territorio. Los fondos públicos de fomento podrían emplear enfoques similares para distribuir sus recursos de manera que maximicen su impacto económico y social sujeto a restricciones presupuestarias.

Ahora bien, la aplicación de estas técnicas requiere cautela. Los algoritmos de optimización son tan buenos como los datos y los pronósticos que los alimentan: si las estimaciones de rentabilidad o de riesgo son erróneas, la cartera optimizada puede resultar peor que una elegida con criterios sencillos. El aprendizaje por refuerzo necesita grandes cantidades de experiencia, que en la práctica suelen obtenerse mediante simulación, por lo que su desempeño depende de la fidelidad del entorno simulado. Sobre este punto, Hambly, Xu y Yang (2023) advierten sobre las dificultades que plantean los datos financieros ruidosos y no estacionarios para estos métodos. Y la complejidad de los algoritmos puede dificultar su comprensión por parte de quienes deben tomar las decisiones, lo que refuerza la necesidad de técnicas de explicabilidad y de supervisión humana.

En América Latina, estas herramientas pueden parecer distantes de la realidad de la mayoría de los emprendimientos. Sin embargo, su lógica puede aplicarse a problemas de pequeña escala con herramientas accesibles. Las bibliotecas de código abierto para algoritmos genéticos y aprendizaje por refuerzo, las plataformas de gestión empresarial con funciones de optimización incorporadas y los servicios de asesoría de incubadoras y universidades pueden acercar estas capacidades a emprendedores que no disponen de personal especializado.

La formación en métodos cuantitativos y la colaboración con instituciones académicas pueden contribuir a ampliar su adopción. En la práctica, un ejemplo sencillo ayuda a ilustrar la utilidad de estas técnicas en un emprendimiento. Supóngase que una persona dispone de un presupuesto limitado para invertir en cinco alternativas: ampliar su inventario, renovar un equipo, contratar publicidad en redes sociales, capacitar a su personal y constituir una reserva de liquidez.

Cada alternativa tiene un rendimiento esperado incierto y un efecto distinto sobre el riesgo del negocio, y algunas se complementan entre sí, como la publicidad y el inventario. Un algoritmo genético puede explorar miles de combinaciones de montos asignados a cada alternativa, evaluarlas mediante simulación en un modelo financiero del negocio y proponer un conjunto de asignaciones que equilibren rentabilidad y riesgo.

La persona emprendedora elige luego entre estas propuestas según sus prioridades. Este procedimiento, que hace unos años habría requerido programas costosos, puede implementarse hoy con herramientas de código abierto y un modelo financiero bien construido (Erwin y Engelbrecht, 2023).

Conviene advertir, asimismo, sobre la tentación de optimizar en exceso. Una cartera o una asignación de recursos que resulta óptima según los pronósticos disponibles puede ser muy sensible a pequeños errores en esos pronósticos, y comportarse mal cuando la realidad se aparta de lo previsto. Las técnicas de optimización robusta, que buscan soluciones que funcionen razonablemente bien en una variedad de escenarios en lugar de soluciones óptimas para un escenario único, ofrecen una alternativa más prudente en contextos de alta incertidumbre como el emprendedor.

Combinar la optimización con la simulación de escenarios permite evaluar esta robustez antes de tomar la decisión. Por otra parte, la optimización de inversiones emprendedoras mediante algoritmos evolutivos y aprendizaje por refuerzo permite abordar decisiones de asignación de recursos cuya complejidad desborda los métodos tradicionales. Los algoritmos evolutivos ayudan a encontrar buenas combinaciones de inversión bajo múltiples restricciones y objetivos, y el aprendizaje por refuerzo permite aprender políticas de decisión secuencial en entornos cambiantes.

Figura 4.7

Optimización de inversiones emprendedoras mediante algoritmos evolutivos y aprendizaje por refuerzo.



Fuente: Elaboración autores

4.8. Diseño de sistemas de alerta temprana para detectar vulnerabilidades financieras y riesgos de cierre empresarial

Un sistema de alerta temprana es un mecanismo que vigila de manera continua un conjunto de indicadores, compara su evolución con umbrales o patrones de riesgo y emite señales cuando detecta condiciones que anticipan un evento adverso. Estos sistemas se utilizan desde hace décadas en ámbitos como la meteorología, la salud pública o la supervisión bancaria, y su lógica se aplica de manera natural a la gestión de las empresas. En el emprendimiento, un sistema de alerta temprana busca detectar las vulnerabilidades financieras que podrían conducir al cierre del negocio con suficiente antelación para que sea posible corregirlas.

Las herramientas examinadas a lo largo de este capítulo —pronósticos financieros, modelos de insolvencia, simulación, gemelos digitales y modelos de liquidez— encuentran en estos sistemas su aplicación más directa. Paralelamente, la evidencia reciente muestra que los sistemas de alerta temprana basados en aprendizaje automático pueden alcanzar un desempeño considerable.

En este sentido, Zhao, Li y Dai (2023) diseñaron un sistema para pequeñas y medianas empresas que, tras seleccionar los indicadores financieros más significativos, combinaba varios modelos de aprendizaje automático mediante una técnica de apilamiento; XGBoost superó a los modelos de referencia y el sistema combinado obtuvo los mejores resultados en la identificación de las empresas en situación de crisis.

Al respecto, Hidalgo-Díaz, Alaminos y Delgado-Gómez (2025), al desarrollar nuevas herramientas de alerta temprana para la prevención de la insolvencia en el contexto europeo, obtuvieron exactitudes superiores al 94 % hasta tres años antes del evento con un modelo que combinaba redes convolucionales y recurrentes, y destacaron la utilidad de estas herramientas para orientar las políticas de prevención.

El diseño de un sistema de alerta temprana para emprendimientos comprende varias decisiones. La primera se refiere al evento que se busca anticipar: puede tratarse del cierre del negocio, de la insolvencia, del incumplimiento de pagos, de una caída sostenida de las ventas o de una crisis de liquidez. Cada definición implica distintos horizontes y distintos indicadores. Luego deben elegirse los indicadores que se vigilarán, que pueden combinar las familias identificadas en la literatura: rentabilidad, liquidez, endeudamiento, actividad y, en el caso de los emprendimientos, indicadores no financieros como la evolución del número de clientes, la tasa de recompra, las reseñas en línea o el cumplimiento de obligaciones tributarias.

También es necesario definir el modelo que transformará esos indicadores en una señal de riesgo, que puede ir desde reglas sencillas basadas en umbrales hasta modelos de aprendizaje automático. Por último, se establece el protocolo de respuesta: qué ocurre cuando se emite una alerta, quién la recibe y qué acciones se activan. No obstante, los niveles de alerta permiten graduar la respuesta según la gravedad del riesgo.

Un esquema habitual utiliza tres o cuatro niveles, a menudo representados con colores: un nivel verde indica una situación normal; uno amarillo señala un deterioro incipiente que requiere atención; uno naranja advierte un riesgo significativo que exige medidas correctivas; y uno rojo indica una situación crítica que requiere intervención inmediata. La asignación de cada nivel puede basarse en la probabilidad estimada de cierre en un horizonte determinado, y los umbrales entre niveles deben calibrarse considerando los costos de las falsas alarmas y de las alertas omitidas, como se examinó en el segundo capítulo al tratar la validación de modelos.

La elección de los horizontes de alerta debe responder a la tensión entre anticipación y precisión. Un sistema que solo detecta el riesgo un mes antes del cierre puede ser muy preciso, pero deja poco margen para actuar. Uno que intenta anticipar el riesgo con dos o tres años de antelación ofrece más tiempo, pero sus señales son menos confiables. La evidencia sugiere que los indicadores de rentabilidad son los que mantienen su capacidad predictiva en horizontes más

largos, mientras que la liquidez y la capacidad de pago de la deuda son más informativas en el corto plazo (Hidalgo-Díaz et al., 2025).

Un sistema bien diseñado puede combinar ambos horizontes: alertas estructurales de largo plazo basadas en la rentabilidad y alertas operativas de corto plazo basadas en la liquidez y los flujos de efectivo. Sobre esta base, la explicabilidad resulta indispensable en un sistema de alerta temprana. Una alerta que solo indica que el riesgo es alto, sin explicar por qué, tiene escasa utilidad para quien debe actuar.

Las técnicas de explicabilidad examinadas en el segundo capítulo, como los valores SHAP, permiten identificar qué indicadores contribuyen más al riesgo estimado en cada caso, de modo que la alerta pueda acompañarse de un diagnóstico: por ejemplo, que el riesgo aumentó por el deterioro de la rotación de cartera y por el incremento de la relación entre los gastos financieros y las ventas. En esa dirección, Bussmann et al. (2021) mostraron cómo la combinación de modelos de aprendizaje automático con explicaciones de este tipo permite que las evaluaciones de riesgo crediticio resulten comprensibles y supervisables.

Un sistema de alerta temprana no termina con la emisión de la señal. Su valor depende de las respuestas que activa. Para la propia persona emprendedora, una alerta puede desencadenar una revisión de los costos, una renegociación de plazos con proveedores, una campaña de cobranza o la búsqueda de financiamiento de corto plazo. Para una incubadora o un programa de apoyo, puede activar un acompañamiento especializado en gestión financiera.

Para una entidad financiera, puede dar lugar a una reestructuración preventiva de la deuda antes de que se produzca el incumplimiento. Y para los responsables de políticas públicas, la agregación de alertas por sector o territorio puede revelar vulnerabilidades sistémicas que requieren medidas de alcance más amplio. Desde esta perspectiva, en el Ecuador, la construcción de sistemas de alerta temprana podría aprovechar la información disponible en los registros administrativos.

La Superintendencia de Compañías, Valores y Seguros recopila los estados financieros de las sociedades, que han servido de base para estudios sobre los determinantes del cierre empresarial (Zambrano et al., 2021; González Cedeño, 2026). El registro estadístico de empresas del Instituto Nacional de Estadística y Censos ofrece información sobre la dinámica empresarial (INEC, 2024). Y la facturación electrónica genera datos sobre las transacciones de las empresas con mayor frecuencia que los estados financieros anuales.

La integración de estas fuentes, con las debidas garantías de protección de datos y confidencialidad previstas en la normativa vigente (Asamblea Nacional del Ecuador, 2021), permitiría construir sistemas de alerta para las sociedades registradas, y su extensión a los negocios de personas naturales y de la economía popular y solidaria requeriría fuentes complementarias.

El diseño de estos sistemas debe considerar también sus posibles efectos no deseados. Una alerta compartida con terceros, como proveedores o entidades financieras, podría precipitar el cierre que se pretende evitar, si conduce a la retirada del crédito o al endurecimiento de las condiciones comerciales.

Por ello, la definición de quién accede a las alertas y con qué propósito es una decisión de gobernanza tan importante como la elección del modelo. Las alertas deben concebirse como instrumentos de apoyo y prevención, no como mecanismos de exclusión. Del mismo modo, los modelos deben evaluarse periódicamente para detectar sesgos que pudieran perjudicar sistemáticamente a determinados grupos de emprendedores, tema que se examina en el capítulo siguiente.

De forma complementaria, la implementación puede seguir una secuencia gradual. En una fase inicial, el sistema puede basarse en pocos indicadores con umbrales definidos a partir de la experiencia y de referencias sectoriales. A medida que se acumulan datos sobre las empresas y sus desenlaces, pueden entrenarse modelos estadísticos y de aprendizaje automático que estimen probabilidades de cierre. Posteriormente, el sistema puede integrarse con los pronósticos financieros, los gemelos digitales y los tableros de control de la empresa, para que las alertas se acompañen de simulaciones de las posibles medidas correctivas. En cada fase, la validación del desempeño del sistema con datos reales es indispensable para mantener la confianza de sus usuarios.

La participación de las propias personas emprendedoras en el diseño de los sistemas de alerta contribuye a su eficacia. Cuando los indicadores, los umbrales y los protocolos de respuesta se definen sin conocer la realidad de los negocios, las alertas pueden resultar irrelevantes o llegar en momentos en que no es posible actuar. La consulta con emprendedores de distintos sectores permite identificar qué señales consideran más reveladoras, qué información están dispuestos a compartir y qué tipo de apoyo les resultaría útil al recibir una alerta.

Esta participación favorece, además, la confianza en el sistema y su adopción. Las experiencias de sistemas de alerta en otros ámbitos, como la gestión de riesgos de desastres, muestran que

su eficacia depende tanto de la calidad técnica de la detección como de la existencia de capacidades y mecanismos para responder. Sobre este punto, Hidalgo-Díaz et al. (2025) subrayan que el valor de estas herramientas reside en su capacidad para orientar políticas de prevención de la insolvencia.

Vinculado con lo anterior, la comunicación de las alertas a las personas emprendedoras requiere un cuidado particular. Una alerta formulada en un lenguaje técnico o alarmista puede generar angustia sin orientar la acción, mientras que una formulada de manera clara, con la explicación de sus causas y con sugerencias concretas, puede movilizar respuestas oportunas. Conviene que cada alerta indique qué indicador ha cambiado, en qué medida, qué consecuencias podría tener si la tendencia continúa y qué acciones se recomiendan, así como los recursos de apoyo disponibles. La frecuencia de las alertas debe calibrarse para evitar la fatiga que producen los avisos demasiado frecuentes, que terminan siendo ignorados.

Al término de este recorrido, conviene retener que los sistemas de alerta temprana constituyen la síntesis práctica de las finanzas inteligentes aplicadas al emprendimiento.

Figura 4.8

Diseño de sistemas de alerta temprana para detectar vulnerabilidades financieras y riesgos de cierre empresarial.



Fuente: Elaboración autores

.

EDITORIAL ANDES COGNITIO

CAPÍTULO V

EMPRENDIMIENTO AUTÓNOMO Y GOBERNANZA ALGORÍTIMICA: HACIA UN MODELO INTEGRAL DE ÉXITO EMPRESARIAL PREDICTIVO



CAPÍTULO V

EMPRENDIMIENTO AUTÓNOMO Y GOBERNANZA ALGORÍTMICA: HACIA UN MODELO INTEGRAL DE ÉXITO EMPRESARIAL PREDICTIVO

5.1. Arquitectura de un modelo integral de emprendimiento predictivo basado en inteligencia artificial

A lo largo de los capítulos anteriores se han examinado por separado distintas piezas del emprendimiento predictivo: los fundamentos conceptuales y económicos de la anticipación, los modelos algorítmicos para estimar la probabilidad de éxito o fracaso, las herramientas para anticipar mercados y clientes, y los instrumentos financieros para evaluar el riesgo y la sostenibilidad.

Cada una de estas piezas tiene valor propio, pero su potencial se multiplica cuando se articulan en un modelo integral, en el que los datos fluyen entre los distintos componentes, las predicciones de un ámbito alimentan las decisiones de otro y el conjunto se somete a reglas comunes de gobernanza. Proponer esa articulación es el propósito de este capítulo. En concordancia con ello, un modelo integral de emprendimiento predictivo puede entenderse como un sistema sociotécnico.

Es técnico porque descansa en datos, algoritmos, plataformas informáticas y procedimientos analíticos. Pero es también social, porque su funcionamiento depende de las personas que lo diseñan, lo alimentan, interpretan sus resultados y toman las decisiones, así como de las instituciones que regulan su uso y de los valores que orientan sus fines.

En este sentido, Raisch y Krakowski (2021) advierten que la automatización y la ampliación de las capacidades humanas mediante inteligencia artificial no son alternativas excluyentes, sino dimensiones que deben combinarse: los sistemas que automatizan tareas en un horizonte corto pueden generar consecuencias organizacionales no previstas si no se acompañan de procesos que mantengan a las personas involucradas en la comprensión y el control de las decisiones. Esta perspectiva atraviesa la arquitectura que aquí se propone.

La arquitectura puede describirse en cinco capas interrelacionadas. La base está constituida por la capa de datos, que integra la información interna del emprendimiento —ventas, costos, clientes, inventarios, cuentas bancarias, facturación electrónica— con información externa del

mercado, del sector y de la economía: estadísticas oficiales, registros administrativos, datos abiertos, búsquedas en internet, reseñas de consumidores y otras fuentes alternativas.

En esta capa se resuelven los problemas de captura, calidad, integración, almacenamiento y protección de la información, y se documenta el origen de cada dato y las condiciones de su uso. A ello se suma que sobre ella se sitúa la capa analítica, donde residen los modelos predictivos examinados en los capítulos precedentes. Pueden distinguirse en ella cuatro módulos. El módulo de viabilidad estima la probabilidad de supervivencia y de éxito del emprendimiento mediante modelos de clasificación y de supervivencia, como los Logit, Probit, Random Forest o XGBoost.

El módulo de mercado pronostica la demanda, segmenta a los clientes, estima su valor a largo plazo y detecta tendencias mediante series temporales, algoritmos de agrupamiento y procesamiento del lenguaje natural. El módulo financiero proyecta ingresos, costos, rentabilidad y flujos de efectivo, simula escenarios mediante el método Monte Carlo y estima la probabilidad de insolvencia. Y el módulo de explicabilidad acompaña las predicciones de los demás módulos con información sobre los factores que las determinan, mediante técnicas como SHAP y LIME.

La tercera capa es la de decisión, en la que las predicciones se transforman en recomendaciones y alternativas de acción. Aquí operan los sistemas de apoyo a decisiones, los gemelos digitales que permiten ensayar estrategias, los algoritmos de optimización que proponen asignaciones de recursos y los agentes de inteligencia artificial que pueden ejecutar determinadas tareas bajo supervisión. Esta capa no sustituye el juicio de la persona emprendedora, sino que le ofrece un abanico de opciones evaluadas en términos de rentabilidad, riesgo y sostenibilidad.

Al respecto, Agrawal, Gans y Goldfarb (2022) recuerdan que la predicción y el juicio son complementarios: la máquina abarata la estimación de lo que probablemente ocurrirá, pero la valoración de las consecuencias y la elección entre alternativas siguen requiriendo criterio humano. Por su parte, la cuarta capa corresponde a la interfaz, el punto de contacto entre el sistema y las personas usuarias.

Incluye el tablero de control predictivo con los indicadores clave, el sistema de alertas tempranas, los informes periódicos y, eventualmente, asistentes conversacionales que permiten consultar el sistema en lenguaje natural. Su diseño debe adaptarse a usuarios con distintos niveles de formación técnica: una persona emprendedora de un pequeño comercio necesita información

sencilla y accionable, mientras que un analista de una entidad financiera puede requerir mayor detalle.

La quinta capa, transversal a las anteriores, es la de gobernanza. Define las reglas sobre el uso de los datos, la protección de la información personal, la transparencia de los modelos, la evaluación de sesgos, la supervisión humana y la rendición de cuentas. En el caso ecuatoriano, esta capa debe garantizar el cumplimiento de la Ley Orgánica de Protección de Datos Personales (Asamblea Nacional del Ecuador, 2021) y de la normativa específica sobre el tratamiento de datos en sistemas de inteligencia artificial emitida por la autoridad de protección de datos (Superintendencia de Protección de Datos Personales [SPDP], 2026).

En el plano internacional, el Reglamento europeo de inteligencia artificial ofrece una referencia sobre los requisitos aplicables a los sistemas de alto riesgo, entre los que incluye aquellos destinados a evaluar la solvencia crediticia de las personas físicas (Parlamento Europeo y Consejo de la Unión Europea, 2024).

Además, la interacción entre las capas sigue un ciclo continuo. Los datos alimentan los modelos; los modelos producen predicciones; las predicciones sustentan decisiones; las decisiones generan resultados que se registran como nuevos datos; y esos datos permiten evaluar el acierto de las predicciones y recalibrar los modelos. Este ciclo de retroalimentación es lo que distingue un modelo integral de un conjunto de herramientas aisladas: el sistema aprende de su propia experiencia y mejora con el tiempo.

En esa dirección, Gupta, Modgil, Bhattacharyya y Bose (2022), en su revisión sobre la inteligencia artificial en los sistemas de apoyo a decisiones, destacan precisamente la capacidad de aprendizaje y la validación dentro del marco de decisión como criterios para evaluar estos sistemas.

El modelo integral puede adoptar distintas configuraciones según quién lo utilice. Para una persona emprendedora individual, puede materializarse en una plataforma de gestión que integre su contabilidad, sus ventas y sus pronósticos, con un tablero de indicadores y alertas. Para una incubadora o aceleradora, puede funcionar como un sistema de acompañamiento de su cartera de emprendimientos, que identifique cuáles requieren apoyo y en qué áreas.

Para una entidad financiera, puede integrarse en sus procesos de evaluación y seguimiento crediticio. Y para un gobierno local o nacional, puede servir como infraestructura de inteligencia económica que articule observatorios de oportunidades, sistemas de alerta y programas de

fomento. En esta misma línea, la implementación del modelo integral debe ser gradual y adaptada a la madurez digital de cada actor. Un emprendimiento que apenas registra sus ventas en un cuaderno no puede saltar de inmediato a un sistema con gemelos digitales y agentes autónomos.

La trayectoria razonable comienza por la digitalización de los registros básicos, continúa con indicadores descriptivos y pronósticos sencillos, incorpora progresivamente modelos predictivos más complejos y, solo cuando existen datos, competencias y gobernanza suficientes, avanza hacia la automatización supervisada de decisiones. Esta gradualidad es coherente con la evidencia sobre la adopción de inteligencia artificial en pequeñas y medianas empresas, que identifica la falta de competencias, de datos y de recursos como barreras principales (OCDE, 2025).

La perspectiva de la Industria 5.0, que coloca a las personas, la resiliencia y la sostenibilidad en el centro de la transformación tecnológica (Breque et al., 2021), ofrece un marco para orientar el diseño del modelo hacia esos fines. Ello implica, por ejemplo, incorporar indicadores ambientales y sociales en el tablero de control, evaluar el efecto de las decisiones sobre trabajadores y comunidades, y asegurar que los beneficios de la predicción lleguen también a los emprendedores con menos recursos.

De igual manera, la viabilidad de esta arquitectura en el contexto ecuatoriano depende de varias condiciones habilitantes. Destaca, ante todo, la conectividad: aunque el acceso a internet ha crecido, persisten diferencias importantes entre zonas urbanas y rurales que limitan el uso de plataformas en línea. Otra condición es la formalización: los emprendimientos informales no generan los registros de facturación y contabilidad que alimentan buena parte de los modelos. Se requiere también la disponibilidad de datos públicos de calidad y en formatos abiertos, que la normativa de transparencia promueve (Asamblea Nacional del Ecuador, 2023). Y resulta indispensable la formación de competencias digitales y analíticas en las personas emprendedoras y en los profesionales que las acompañan.

Por lo que estas condiciones no dependen exclusivamente de los emprendedores, de modo que su desarrollo exige la acción coordinada de gobiernos, universidades, gremios y entidades financieras, lo que vuelve a situar el emprendimiento predictivo en el marco de los ecosistemas emprendedores.

Las evidencias reunidas sugieren que la arquitectura propuesta ofrece una visión de conjunto que articula datos, modelos, decisiones, interfaces y gobernanza en un sistema que aprende de

su experiencia. Su valor no reside en la sofisticación de cada componente, sino en la integración entre ellos y en la presencia de personas que comprenden, supervisan y orientan el sistema hacia fines legítimos.

Figura 5.1

Arquitectura de un modelo integral de emprendimiento predictivo basado en inteligencia artificial.



Fuente: Elaboración autores

5.2. Sistemas inteligentes de apoyo a decisiones para emprendedores y gestores empresariales

Los sistemas de apoyo a decisiones tienen una larga historia en la administración. Desde las primeras aplicaciones informáticas que ayudaban a los gerentes a analizar alternativas mediante modelos de optimización y simulación, estos sistemas se han propuesto un objetivo constante: mejorar la calidad de las decisiones humanas sin sustituirlas, ofreciendo información, análisis y evaluación de opciones a quienes deben elegir. La incorporación de la inteligencia artificial ha transformado su alcance.

Los sistemas actuales no solo procesan los datos que se les proporcionan, sino que aprenden de ellos, generan predicciones, detectan patrones que escapan a la observación humana y proponen alternativas de acción que se actualizan a medida que cambian las circunstancias. Por otra parte, Gupta, Modgil, Bhattacharyya y Bose (2022) examinaron la integración de la inteligencia artificial en los sistemas de apoyo a decisiones en el campo de la investigación de operaciones, y mapearon las sinergias, diferencias y superposiciones entre estos tres ámbitos.

Los autores evaluaron los estudios previos según su capacidad de pronóstico, el uso de grandes volúmenes de datos, el número de factores considerados, la capacidad de aprendizaje y la

validación dentro del marco de decisión, e identificaron como direcciones futuras de investigación la predicción y el aprendizaje, la toma de decisiones y la optimización orientadas a decisiones inteligentes en contextos de incertidumbre. Estas dimensiones ofrecen una guía para diseñar sistemas de apoyo dirigidos a emprendedores.

Un sistema inteligente de apoyo a decisiones para emprendimientos puede concebirse como la capa de decisión de la arquitectura integral descrita en el apartado precedente. Recibe las predicciones de los módulos analíticos —probabilidad de supervivencia, pronósticos de demanda, proyecciones financieras, riesgo de insolvencia— y las transforma en información útil para decidir. Su funcionamiento puede descomponerse en varias operaciones: identificar las decisiones que la persona emprendedora debe tomar, generar alternativas para cada una, evaluar sus consecuencias probables en términos de los objetivos del negocio, comparar las alternativas considerando el riesgo y presentar los resultados de manera comprensible.

Paralelamente, las decisiones que estos sistemas pueden apoyar son variadas. Antes de crear la empresa, ayudan a evaluar la viabilidad de la idea, a elegir el segmento de clientes al que dirigirse, a decidir la ubicación del negocio o a estimar el capital inicial necesario. Durante la operación, apoyan decisiones de fijación de precios, de compras e inventarios, de contratación de personal, de inversión publicitaria o de gestión del crédito a clientes. En las etapas de crecimiento, orientan decisiones sobre la apertura de nuevos mercados, el lanzamiento de productos, la búsqueda de financiamiento o la incorporación de socios. Y en situaciones de crisis, ayudan a identificar las medidas que con mayor probabilidad permitirán preservar la viabilidad del negocio.

El valor de un sistema de apoyo a decisiones no depende solo de la precisión de sus predicciones, sino también de la manera en que se integra con el juicio humano. Las personas tienden a desarrollar dos actitudes problemáticas frente a las recomendaciones algorítmicas. En primer lugar, destaca la aversión, es decir, el rechazo a utilizar las recomendaciones de un algoritmo después de observar que se equivoca, aunque sus errores sean menos frecuentes que los humanos. La opuesta es la confianza excesiva o sesgo de automatización, que lleva a aceptar las recomendaciones sin examinarlas críticamente.

Un buen diseño busca una confianza calibrada, en la que la persona usuaria comprenda las fortalezas y limitaciones del sistema y sepa cuándo seguir sus recomendaciones y cuándo apartarse de ellas. Sobre esta base, la explicabilidad contribuye de manera decisiva a esa calibración.

Un sistema que acompaña sus recomendaciones con información sobre los factores que las sustentan permite a la persona usuaria contrastarlas con su propio conocimiento del negocio. Si el sistema recomienda reducir el inventario de un producto porque pronostica una caída de la demanda, y explica que esa predicción se basa en la tendencia de las ventas recientes y en la estacionalidad observada en años anteriores, la persona emprendedora puede evaluar si existen circunstancias que el sistema no conoce, como un pedido importante comprometido o una campaña promocional prevista. Sobre este punto, Molnar (2022) subraya que la interpretabilidad no solo facilita la confianza, sino que permite detectar errores, sesgos y relaciones espurias en los modelos.

La forma de presentar la incertidumbre merece también atención. Las predicciones nunca son certezas, y un sistema que presenta sus pronósticos como cifras exactas induce a una falsa sensación de seguridad. Resulta preferible comunicar rangos, probabilidades y escenarios: en lugar de afirmar que las ventas del próximo mes serán de diez mil dólares, indicar que con un 80 % de probabilidad se ubicarán entre ocho mil y doce mil.

Esta información, aunque menos precisa en apariencia, es más honesta y más útil para decidir, porque permite preparar respuestas para los distintos escenarios. Los métodos de pronóstico probabilístico y de simulación examinados en capítulos anteriores ofrecen las herramientas para generar esta información (Hyndman y Athanasopoulos, 2021). Desde esta perspectiva, los sistemas de apoyo a decisiones deben adaptarse a las características de sus usuarios.

Una persona emprendedora con poca formación técnica y escaso tiempo necesita recomendaciones concretas, expresadas en lenguaje sencillo y vinculadas a acciones inmediatas. Un gestor empresarial con experiencia analítica puede preferir explorar los datos, modificar los supuestos y comparar escenarios por sí mismo. Los asesores de incubadoras o entidades financieras pueden requerir vistas agregadas de varios emprendimientos. Un diseño flexible, que ofrezca distintos niveles de detalle según el perfil del usuario, permite que el mismo sistema sirva a públicos diversos.

La incorporación de interfaces conversacionales basadas en modelos de lenguaje abre nuevas posibilidades. Una persona emprendedora podría consultar al sistema en lenguaje natural —por ejemplo, preguntar qué ocurriría con su flujo de caja si contrata a un vendedor adicional— y recibir una respuesta que integre las predicciones de los distintos módulos. Esta forma de interacción reduce las barreras técnicas de acceso, pero plantea también riesgos: los modelos de

lenguaje pueden generar respuestas plausibles pero erróneas, por lo que deben conectarse a los datos y modelos verificados del sistema y no responder solo a partir de su conocimiento general.

En este sentido, Dwivedi et al. (2023) advierten, en su revisión sobre la inteligencia artificial explicable, que la confianza en los sistemas depende de la posibilidad de comprender y verificar sus resultados. De forma complementaria, para la realidad latinoamericana, los sistemas de apoyo a decisiones pueden contribuir a reducir algunas de las brechas que limitan la supervivencia de los emprendimientos.

Muchas personas emprenden sin formación en gestión ni acceso a asesoría profesional, y toman decisiones críticas basándose en la intuición o en la experiencia de familiares y conocidos. Un sistema accesible, que ofrezca análisis y recomendaciones fundamentados en datos, podría cumplir parcialmente el papel de un asesor. Para ello, debe ser asequible, funcionar en dispositivos móviles, adaptarse a las condiciones de conectividad y expresarse en un lenguaje cercano a sus usuarios. La articulación con programas de formación, incubadoras y centros de desarrollo empresarial puede facilitar su adopción y su uso adecuado.

El diseño de estos sistemas puede beneficiarse de la participación de sus futuros usuarios. Los métodos de diseño centrado en las personas proponen observar cómo toman decisiones los emprendedores en su contexto real, identificar los momentos en que necesitan información y probar prototipos de las interfaces antes de su desarrollo completo.

Este proceso permite descubrir, por ejemplo, que una persona que gestiona un pequeño comercio revisa su información en el teléfono móvil durante pocos minutos al final del día, lo que exige presentar las recomendaciones de manera muy sintética; o que prefiere recibir alertas por mensajería instantánea antes que consultar un tablero. Atender a estos hábitos puede marcar la diferencia entre un sistema que se utiliza y otro que se abandona tras las primeras semanas.

Al respecto, Gupta et al. (2022) señalan la validación dentro del marco de decisión como un criterio central para evaluar estos sistemas, lo que incluye su utilidad efectiva para quienes deciden. Vinculado con lo anterior, los sistemas inteligentes de apoyo a decisiones representan el punto en que las capacidades predictivas del modelo integral se convierten en acción.

Evidentemente su valor depende de la calidad de las predicciones que integran, pero también de su capacidad para presentar alternativas evaluadas, comunicar la incertidumbre de manera honesta, explicar sus recomendaciones y adaptarse a las características de quienes los utilizan.

Figura 5.2

Sistemas inteligentes de apoyo a decisiones para emprendedores y gestores empresariales.



Fuente: Elaboración autores

5.3. Agentes de inteligencia artificial para la automatización supervisada de procesos estratégicos empresariales

La evolución reciente de la inteligencia artificial ha dado lugar a una nueva categoría de sistemas: los agentes. A diferencia de los modelos predictivos, que responden a una consulta con una estimación, y de los sistemas de apoyo a decisiones, que presentan alternativas para que una persona elija, los agentes son capaces de planificar una secuencia de acciones para alcanzar un objetivo, ejecutarlas utilizando herramientas externas, observar los resultados y ajustar su conducta en función de ellos.

El desarrollo de los grandes modelos de lenguaje ha impulsado notablemente esta capacidad, al dotar a los agentes de una comprensión del lenguaje natural y de una capacidad de razonamiento que les permiten abordar tareas abiertas y poco estructuradas. En concordancia con ello, Wang et al. (2024), en una revisión sistemática de los agentes autónomos basados en grandes modelos de lenguaje, propusieron un marco unificado para describir su construcción, compuesto por cuatro módulos: un módulo de perfil, que define el rol y las características del agente; un módulo de memoria, que almacena la información de interacciones pasadas y le permite acumular experiencia; un módulo de planificación, que descompone los objetivos en tareas y decide el orden en que deben realizarse; y un módulo de acción, que traduce las decisiones en operaciones concretas, como consultar una base de datos, ejecutar un cálculo, enviar un mensaje o utilizar una aplicación.

Los autores revisaron también las aplicaciones de estos agentes en las ciencias sociales, las ciencias naturales y la ingeniería, así como las estrategias para evaluarlos y los desafíos pendientes.

En el ámbito empresarial, los agentes de inteligencia artificial pueden automatizar procesos que hasta ahora requerían la intervención continua de personas. Un agente de compras podría vigilar los niveles de inventario, consultar los pronósticos de demanda, comparar precios de proveedores y preparar órdenes de compra. Un agente de cobranza podría identificar las facturas vencidas, priorizar a los clientes según su riesgo de impago, enviar recordatorios personalizados y registrar las respuestas.

Un agente de inteligencia comercial podría rastrear las reseñas de clientes, las publicaciones en redes sociales y las noticias del sector, y elaborar informes periódicos sobre tendencias y oportunidades. Un agente financiero podría actualizar diariamente las proyecciones de flujo de caja y emitir alertas cuando detecte riesgos de liquidez. Con todo, la automatización de procesos estratégicos, sin embargo, exige cautela. A diferencia de las tareas operativas rutinarias, las decisiones estratégicas tienen consecuencias significativas, a menudo difíciles de revertir, y dependen de valoraciones que involucran objetivos, valores y conocimiento del contexto que los agentes no poseen plenamente.

Por ello, el concepto central de este apartado es el de automatización supervisada: los agentes ejecutan tareas de preparación, análisis y ejecución dentro de límites definidos, pero las decisiones relevantes requieren la aprobación de una persona responsable. En esa dirección, Raisch y Krakowski (2021) advierten que la automatización sin la participación humana puede generar efectos indeseados en el largo plazo, mientras que la combinación de automatización y ampliación de capacidades humanas permite aprovechar las ventajas de ambas.

El diseño de la supervisión puede organizarse según el nivel de riesgo de cada tarea. Para tareas de bajo riesgo y fácilmente reversibles, como la elaboración de un informe o el envío de un recordatorio de pago, el agente puede actuar de manera autónoma, con una revisión posterior de sus resultados. Para tareas de riesgo moderado, como la emisión de una orden de compra dentro de un presupuesto preestablecido, el agente puede actuar dentro de límites definidos y solicitar aprobación cuando los supere.

Para tareas de alto riesgo, como la contratación de un préstamo, la modificación de precios de toda la línea de productos o la terminación de la relación con un cliente, el agente debe limitarse

a preparar la información y la recomendación, dejando la decisión en manos humanas. Este esquema graduado permite aprovechar la eficiencia de la automatización sin perder el control de las decisiones importantes. Ahora bien, los agentes de inteligencia artificial plantean riesgos específicos que deben gestionarse. Los grandes modelos de lenguaje pueden generar información inexacta presentada con apariencia de certeza, lo que puede conducir a acciones equivocadas si no se verifica.

Los agentes que interactúan con sistemas externos pueden ser manipulados mediante instrucciones maliciosas incluidas en los datos que procesan, como un correo electrónico o una página web. Las acciones encadenadas pueden amplificar errores iniciales. Y la delegación de tareas a agentes puede diluir la responsabilidad sobre sus consecuencias. Sobre este punto, Wang et al. (2024) identifican, entre los desafíos pendientes de estos sistemas, la fiabilidad de su comportamiento, la alineación con los objetivos humanos y la evaluación de su desempeño en entornos reales.

Para mitigar estos riesgos, la implementación de agentes en los emprendimientos debería observar varios principios. La delimitación clara de permisos, de modo que cada agente acceda solo a los datos y herramientas que necesita para su tarea. El registro completo de sus acciones, que permita reconstruir qué hizo, por qué y con qué resultados. La verificación de la información crítica contra fuentes confiables antes de actuar sobre ella.

Los mecanismos de interrupción, que permitan a una persona detener al agente en cualquier momento. Y la evaluación periódica de su desempeño, comparando sus resultados con los objetivos previstos. Estos principios se corresponden con las exigencias de supervisión humana, registro y transparencia que el Reglamento europeo de inteligencia artificial establece para los sistemas de alto riesgo (Parlamento Europeo y Consejo de la Unión Europea, 2024). A ello se suma que la protección de los datos personales adquiere especial relevancia cuando los agentes procesan información de clientes, trabajadores o proveedores.

Para el Ecuador, la Ley Orgánica de Protección de Datos Personales reconoce a los titulares el derecho a no ser objeto de decisiones basadas únicamente en valoraciones automatizadas que produzcan efectos jurídicos o los afecten de manera significativa (Asamblea Nacional del Ecuador, 2021). Un agente de cobranza que decida por sí solo restringir el crédito a un cliente, o un agente de selección de personal que descarte automáticamente a candidatos, podría entrar en conflicto con este derecho. La supervisión humana de las decisiones con efectos significativos sobre las personas no es, por tanto, solo una buena práctica, sino una exigencia normativa.

Para los emprendimientos de menor tamaño, los agentes de inteligencia artificial ofrecen la posibilidad de acceder a capacidades que antes requerían personal especializado. Una persona que gestiona sola su negocio podría delegar en agentes tareas de contabilidad básica, atención de consultas de clientes, publicación en redes sociales o seguimiento de cobros, liberando tiempo para las actividades que requieren su atención directa.

La difusión de plataformas que ofrecen agentes configurables sin necesidad de programación facilita esta adopción. No obstante, el aprovechamiento seguro de estas herramientas requiere competencias digitales y comprensión de sus limitaciones, lo que refuerza la importancia de la formación de las personas emprendedoras. No obstante, la experiencia acumulada con los agentes sugiere también la conveniencia de comenzar por tareas acotadas y de bajo riesgo, en las que sus errores puedan detectarse y corregirse con facilidad, y ampliar progresivamente su ámbito de acción a medida que se gana confianza en su funcionamiento.

Un emprendimiento puede iniciar, por ejemplo, con un agente que redacte borradores de respuestas a consultas frecuentes de los clientes, revisados por una persona antes de su envío; continuar con un agente que prepare informes semanales de ventas y alertas de inventario; y solo después considerar agentes que ejecuten acciones con efectos externos, como pedidos a proveedores. En cada etapa, la comparación entre los resultados del agente y los que se habrían obtenido sin él permite evaluar su aporte real. Este enfoque gradual reduce los riesgos y facilita el aprendizaje organizacional necesario para integrar a los agentes en los procesos de la empresa (Wang et al., 2024).

Hay razones, por consiguiente, para afirmar que los agentes de inteligencia artificial representan un avance significativo en la capacidad de los emprendimientos para automatizar procesos complejos.

Figura 5.3

Agentes de inteligencia artificial para la automatización supervisada de procesos estratégicos empresariales.



Fuente: Elaboración autores

5.4. Diseño de un tablero de control predictivo con indicadores de supervivencia, crecimiento y rentabilidad

El tablero de control es una herramienta de gestión que presenta de manera visual y sintética los indicadores clave de una organización, de modo que quienes la dirigen puedan conocer su situación de un vistazo y orientar sus decisiones. Los tableros tradicionales muestran principalmente indicadores retrospectivos: las ventas del mes anterior, el margen del último trimestre, el saldo actual de caja. Un tablero de control predictivo añade a esa mirada hacia atrás una mirada hacia adelante, al incorporar pronósticos, probabilidades y alertas que anticipan la evolución del negocio.

Su propósito no es solo informar sobre lo ocurrido, sino ayudar a prever lo que podría ocurrir y a actuar a tiempo. Asimismo, el diseño de un tablero predictivo para emprendimientos parte de una pregunta fundamental: qué necesita saber la persona emprendedora para tomar mejores decisiones. La respuesta puede organizarse en torno a tres dimensiones que este libro ha examinado de manera recurrente: la supervivencia, el crecimiento y la rentabilidad. Cada una responde a una inquietud distinta.

La supervivencia se refiere a la probabilidad de que el negocio continúe operando en los próximos meses y años. El crecimiento, a la capacidad de ampliar sus ventas, sus clientes y su presencia en el mercado. La rentabilidad, a su capacidad para generar excedentes que remuneren el esfuerzo y la inversión. Un emprendimiento puede crecer sin ser rentable, o ser rentable sin

crecer, y ambas situaciones pueden comprometer su supervivencia; por ello, el tablero debe mostrar las tres dimensiones de manera equilibrada.

Los indicadores de supervivencia ocupan un lugar central. El principal de ellos es la probabilidad estimada de supervivencia en distintos horizontes —por ejemplo, seis, doce y veinticuatro meses—, calculada mediante los modelos de clasificación y de supervivencia examinados en el segundo capítulo. Se acompaña de indicadores de liquidez proyectada, como los días de operación que el efectivo disponible permitiría cubrir o la probabilidad de que el saldo de caja caiga por debajo de un umbral mínimo.

La relación entre los gastos financieros y las ventas, identificada por Zizi et al. (2021) como predictor relevante de las dificultades financieras, y el apalancamiento, cuyo efecto sobre el riesgo de cierre ha sido documentado en empresas ecuatorianas (González Cedeño, 2026), completan esta dimensión. Además, los indicadores de crecimiento muestran la trayectoria del negocio y sus perspectivas. Incluyen las ventas pronosticadas para los próximos períodos con sus intervalos de confianza, la tasa de crecimiento esperada, el número de clientes activos y su evolución, la tasa de adquisición de nuevos clientes y la tasa de retención.

El valor estimado de los clientes a largo plazo, examinado en el tercer capítulo, ofrece una medida de la base de clientes como activo del negocio (Firmansyah et al., 2024). Las señales de mercado, como la evolución de las búsquedas en internet relacionadas con los productos del negocio o el sentimiento expresado en las reseñas de los clientes, pueden anticipar cambios en la demanda antes de que se reflejen en las ventas.

Los indicadores de rentabilidad reflejan la capacidad del negocio para generar excedentes. Incluyen el margen bruto y el margen neto proyectados, la rentabilidad sobre los activos, el punto de equilibrio y la probabilidad de alcanzarlo en distintos plazos. La evidencia sobre la predicción de la insolvencia muestra que los indicadores de rentabilidad, como el margen de beneficio neto y la rentabilidad sobre los activos, mantienen su capacidad predictiva en horizontes más largos que los de liquidez (Hidalgo-Díaz et al., 2025), lo que justifica su presencia destacada en el tablero como señales estructurales de la salud del negocio.

En esta misma línea, a estas tres dimensiones conviene añadir indicadores de sostenibilidad y de impacto, coherentes con la orientación del emprendimiento predictivo hacia la Industria 5.0. Pueden incluir el consumo de energía o materiales por unidad producida, la generación de empleo y su calidad, la proporción de proveedores locales o la satisfacción de los trabajadores.

Estos indicadores no solo responden a consideraciones éticas, sino que reflejan factores que influyen cada vez más en la preferencia de los consumidores, en el acceso al financiamiento y en la resiliencia de largo plazo de las empresas.

La construcción de los indicadores compuestos que sintetizan varias dimensiones debe seguir procedimientos rigurosos. La guía del Centro Común de Investigación de la Comisión Europea para la construcción de indicadores compuestos propone diez pasos que incluyen la definición del marco conceptual, la selección de los datos, la imputación de valores faltantes, el análisis multivariante, la normalización, la ponderación y agregación, el análisis de robustez y sensibilidad, y la presentación de los resultados (JRC, s. f.).

Seguir estos pasos permite que los indicadores sintéticos del tablero, como un índice global de salud del emprendimiento, sean coherentes, transparentes y robustos. De igual manera, el diseño visual del tablero influye decisivamente en su utilidad. Los principios de una buena visualización recomiendan priorizar la información más importante en la parte superior y en tamaño destacado, utilizar colores de manera consistente y con significado —por ejemplo, verde, amarillo y rojo para los niveles de alerta—, evitar el exceso de información que dificulte la lectura y permitir que el usuario profundice en los detalles cuando lo necesite.

Los pronósticos deben presentarse con sus rangos de incertidumbre, mediante bandas o abanicos que muestren los valores posibles, y no como líneas únicas que sugieran una precisión inexistente. Las alertas deben acompañarse de explicaciones que indiquen sus causas y de sugerencias de acción.

Un tablero predictivo para un pequeño emprendimiento podría organizarse en una sola pantalla con cuatro bloques. En la parte superior, un indicador sintético de salud del negocio, con su nivel de alerta y la probabilidad de supervivencia a doce meses. Debajo, tres bloques correspondientes a supervivencia, crecimiento y rentabilidad, cada uno con tres o cuatro indicadores clave, su tendencia reciente y su pronóstico.

Al pie, una sección de alertas y recomendaciones que señale las situaciones que requieren atención y las acciones sugeridas. Desde cada indicador, el usuario podría acceder a un análisis más detallado y a la explicación de los factores que lo determinan. Por otra parte, la actualización de los datos determina la utilidad del tablero. Un tablero alimentado manualmente una vez al mes ofrece una visión limitada; uno conectado a los sistemas de facturación electrónica, de ventas y bancarios del emprendimiento puede actualizarse diariamente o incluso en tiempo real.

La integración con las fuentes de datos externas —estadísticas sectoriales, datos de mercado, observatorios de oportunidades— permite situar el desempeño del negocio en relación con su entorno y con empresas comparables.

El tablero predictivo debe concebirse, por último, como una herramienta de aprendizaje. Al registrar las predicciones realizadas y compararlas con los resultados efectivos, permite evaluar la precisión de los modelos y recalibrarlos. Al registrar las decisiones tomadas en respuesta a las alertas y sus consecuencias, permite aprender qué acciones resultan más eficaces en cada situación. Este ciclo de retroalimentación convierte el tablero en un componente activo del modelo integral de emprendimiento predictivo.

Paralelamente, la selección de los indicadores debe adaptarse, además, al tipo de emprendimiento. Un negocio de comercio minorista necesita vigilar la rotación de inventarios y el margen por producto; uno de servicios profesionales, la utilización de las horas disponibles y la concentración de clientes; una empresa digital por suscripción, la tasa de cancelación y el costo de adquisición de clientes; un emprendimiento agrícola, los precios de sus productos y las condiciones climáticas. Un tablero genérico, con los mismos indicadores para todos los negocios, corre el riesgo de mostrar información poco relevante para unos y de omitir la esencial para otros.

Una solución práctica consiste en definir un núcleo común de indicadores de supervivencia, crecimiento y rentabilidad, y complementarlo con módulos sectoriales que incorporen las métricas propias de cada actividad. Esta estructura modular facilita también la comparación de cada negocio con otros de su mismo sector y la agregación de información para los programas de apoyo.

Como balance, el tablero de control predictivo constituye la ventana a través de la cual la persona emprendedora observa el presente y el futuro probable de su negocio. Al integrar indicadores de supervivencia, crecimiento, rentabilidad y sostenibilidad, con pronósticos que expresan su incertidumbre y alertas que explican sus causas, ofrece una visión de conjunto que facilita la anticipación y la acción oportuna.

Figura 5.4

Diseño de un tablero de control predictivo con indicadores de supervivencia, crecimiento y rentabilidad.



Fuente: Elaboración autores

5.5. Gobernanza algorítmica, protección de datos, transparencia y ética en las decisiones empresariales automatizadas

La creciente presencia de algoritmos en las decisiones empresariales plantea una pregunta que trasciende lo técnico: quién define las reglas bajo las cuales operan estos sistemas, cómo se asegura que respeten los derechos de las personas y a quién se puede exigir cuentas cuando producen daños. La gobernanza algorítmica responde a esa pregunta. Puede entenderse como el conjunto de principios, normas, procesos y estructuras que orientan el diseño, el desarrollo, la utilización y la supervisión de los sistemas algorítmicos, con el fin de que sus beneficios se obtengan de manera legítima y sus riesgos se gestionen de forma responsable.

En el emprendimiento predictivo, la gobernanza no es un añadido posterior a la tecnología, sino una condición para su adopción sostenible. Sobre esta base, el marco normativo internacional sobre la inteligencia artificial ha avanzado de manera significativa en los últimos años. La Recomendación sobre la Ética de la Inteligencia Artificial, adoptada por la Conferencia General de la UNESCO en noviembre de 2021, estableció el primer instrumento normativo de alcance mundial en esta materia, con principios como la proporcionalidad, la seguridad, la equidad y la no discriminación, la privacidad, la supervisión y decisión humanas, la transparencia y la explicabilidad, y la responsabilidad y rendición de cuentas (Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura [UNESCO], 2021).

El Reglamento europeo de inteligencia artificial, por su parte, adoptó un enfoque basado en el riesgo, que prohíbe determinadas prácticas, impone requisitos estrictos a los sistemas de alto riesgo —entre ellos, los destinados a evaluar la solvencia de las personas físicas o a seleccionar personal— y establece obligaciones de transparencia para otros sistemas (Parlamento Europeo y Consejo de la Unión Europea, 2024).

En el ámbito de la gestión de riesgos, el Marco de Gestión de Riesgos de Inteligencia Artificial del Instituto Nacional de Estándares y Tecnología de los Estados Unidos propone organizar la gobernanza en torno a cuatro funciones: gobernar, que establece una cultura organizacional de gestión de riesgos; mapear, que identifica el contexto y los riesgos de cada sistema; medir, que evalúa y monitorea esos riesgos; y gestionar, que prioriza y atiende los riesgos identificados (National Institute of Standards and Technology [NIST], 2023). Aunque se trata de un marco voluntario, su estructura ofrece una guía práctica aplicable también a organizaciones pequeñas que desean adoptar la inteligencia artificial de manera responsable.

Desde esta perspectiva, en el ámbito ecuatoriano, la protección de datos personales constituye el pilar principal del marco normativo aplicable. La Ley Orgánica de Protección de Datos Personales establece los principios que rigen el tratamiento de la información personal, entre ellos la licitud, la transparencia, la finalidad, la minimización, la proporcionalidad, la confidencialidad y la seguridad; reconoce los derechos de los titulares, como el acceso, la rectificación, la eliminación, la oposición y la portabilidad; y consagra el derecho a no ser objeto de decisiones basadas únicamente en valoraciones automatizadas que produzcan efectos jurídicos o afecten significativamente a las personas (Asamblea Nacional del Ecuador, 2021).

La Superintendencia de Protección de Datos Personales emitió en 2026 una norma general destinada a garantizar este derecho en el uso de sistemas de inteligencia artificial, que precisa las obligaciones de quienes tratan datos personales mediante estos sistemas (SPDP, 2026).

Para un emprendimiento que utiliza modelos predictivos, estas normas tienen consecuencias concretas. Si el sistema procesa datos de clientes para segmentarlos, recomendarles productos o evaluar su riesgo crediticio, debe contar con una base legal para ese tratamiento, informar a los titulares sobre el uso de sus datos y la lógica de las decisiones automatizadas, limitar la recolección a los datos necesarios, protegerlos con medidas de seguridad adecuadas y permitir el ejercicio de los derechos reconocidos.

Si las decisiones automatizadas producen efectos significativos sobre las personas, como la negación de un crédito o de un empleo, deben existir mecanismos de intervención humana que permitan revisarlas. De forma complementaria, la transparencia constituye un principio central de la gobernanza algorítmica, pero su alcance debe precisarse. No se trata de publicar el código de los algoritmos, lo que en muchos casos resultaría inútil para las personas afectadas e inconveniente para las empresas, sino de ofrecer información comprensible sobre qué decisiones se toman con apoyo de algoritmos, qué datos se utilizan, cuáles son los principales factores que influyen en el resultado y cómo puede solicitarse una revisión.

Las técnicas de inteligencia artificial explicable examinadas en el segundo capítulo, como SHAP y LIME, permiten generar explicaciones individuales de las predicciones que facilitan esta transparencia (Molnar, 2022; Dwivedi et al., 2023).

La ética de las decisiones automatizadas va más allá del cumplimiento normativo. Un sistema puede respetar formalmente la ley y, sin embargo, producir resultados injustos, reforzar desigualdades existentes o deteriorar la confianza de clientes y trabajadores. Las preguntas éticas que deben plantearse incluyen si el uso del algoritmo es proporcional al fin perseguido, si sus beneficios se distribuyen de manera equitativa, si sus errores afectan de manera desproporcionada a determinados grupos, si las personas afectadas comprenden y aceptan su utilización, y si existen alternativas menos invasivas para alcanzar el mismo objetivo.

Estas preguntas no tienen respuestas automáticas, y requieren la deliberación de personas con conocimiento del contexto y de los valores en juego. Con todo, la gobernanza algorítmica en un emprendimiento puede organizarse mediante instrumentos proporcionales a su tamaño y al riesgo de sus sistemas. Un inventario de los sistemas algorítmicos utilizados, con la descripción de su finalidad, los datos que procesan y las decisiones que apoyan, permite conocer la exposición al riesgo.

Una evaluación de impacto previa a la implantación de los sistemas de mayor riesgo identifica los posibles efectos sobre los derechos de las personas y las medidas para mitigarlos. Una política de uso de datos y de inteligencia artificial, comunicada a trabajadores y clientes, establece las reglas aplicables. Un responsable designado asume la supervisión de los sistemas y la atención de las consultas y reclamaciones. Y una revisión periódica verifica que los sistemas siguen funcionando conforme a lo previsto.

Para los emprendimientos más pequeños, estas exigencias pueden parecer desproporcionadas. Sin embargo, buena parte de los sistemas que utilizan son provistos por terceros —plataformas

de comercio electrónico, de gestión empresarial, de publicidad digital, de pagos—, lo que traslada parte de la responsabilidad a los proveedores y convierte la selección de proveedores confiables en una decisión de gobernanza.

Las entidades de apoyo al emprendimiento, las cámaras empresariales y las universidades pueden contribuir con guías prácticas, modelos de políticas y formación que faciliten el cumplimiento sin cargas excesivas. La autoridad de protección de datos puede, a su vez, ofrecer orientaciones específicas para las pequeñas empresas. Vinculado con lo anterior, la gobernanza algorítmica representa también una oportunidad competitiva.

Los consumidores y los socios comerciales valoran cada vez más la protección de su información y el uso responsable de la tecnología. Un emprendimiento que demuestra prácticas transparentes y éticas en el uso de datos y algoritmos puede generar mayor confianza y diferenciarse de sus competidores. Las entidades financieras e inversionistas, por su parte, incorporan crecientemente criterios de gobernanza en sus evaluaciones, lo que puede facilitar el acceso al financiamiento para las empresas que cuentan con prácticas sólidas.

La gobernanza algorítmica se extiende también a la relación con los proveedores de tecnología. Cuando un emprendimiento contrata servicios de inteligencia artificial en la nube, entrega datos de sus clientes a terceros que pueden procesarlos en otros países y bajo otras normas. Los contratos con estos proveedores deben establecer con claridad las finalidades del tratamiento, las medidas de seguridad, la ubicación de los datos, las condiciones de transferencia internacional y la eliminación de la información al terminar la relación, conforme a lo que exige la normativa de protección de datos (Asamblea Nacional del Ecuador, 2021).

Debe prestarse atención, además, a si los datos del emprendimiento se utilizan para entrenar los modelos del proveedor, lo que podría comprometer información confidencial del negocio. La lectura cuidadosa de las condiciones de servicio y la preferencia por proveedores que ofrezcan garantías verificables constituyen prácticas sencillas que reducen considerablemente los riesgos de la adopción tecnológica. En concordancia con ello, la gobernanza algorítmica exige, por último, capacidades institucionales en el Estado.

La autoridad de protección de datos necesita recursos técnicos para supervisar sistemas complejos, atender reclamaciones y emitir orientaciones; los organismos de fomento del emprendimiento deben incorporar criterios de uso responsable de la inteligencia artificial en sus programas; y las universidades pueden contribuir con la formación de profesionales capaces de auditar

algoritmos y con la investigación sobre sus efectos. La construcción de estas capacidades es gradual, pero resulta indispensable para que el marco normativo se traduzca en una protección efectiva de los derechos y en un entorno de confianza para la innovación.

La gobernanza algorítmica, la protección de datos, la transparencia y la ética constituyen la base sobre la cual puede construirse un emprendimiento predictivo legítimo y sostenible. El marco normativo ecuatoriano, articulado en torno a la protección de datos personales y la norma sobre su tratamiento en sistemas de inteligencia artificial, junto con las referencias internacionales de la UNESCO, la Unión Europea y el NIST, ofrece orientaciones para que los emprendimientos aprovechen los beneficios de la predicción respetando los derechos de las personas. La adopción de instrumentos de gobernanza proporcionales al tamaño y al riesgo de cada organización permite convertir estas exigencias en prácticas cotidianas de gestión responsable.

Figura 5.5

Gobernanza algorítmica, protección de datos, transparencia y ética en las decisiones empresariales automatizadas.



Fuente: Elaboración autores

5.6. Evaluación de sesgos algorítmicos y mecanismos de supervisión humana en modelos de predicción empresarial

Los modelos predictivos aprenden de los datos, y los datos reflejan el mundo tal como ha sido, con sus desigualdades, exclusiones y prejuicios. Cuando un modelo se entrena con información histórica sobre qué emprendimientos obtuvieron financiamiento, cuáles sobrevivieron o qué clientes pagaron sus deudas, puede aprender a reproducir los patrones de discriminación presentes en esa historia. Si en el pasado las mujeres emprendedoras, las personas de determinados territorios o los negocios informales tuvieron menos acceso al crédito, un modelo entrenado

con esos datos podría asignarles sistemáticamente menores puntuaciones, no porque sean menos viables, sino porque el sistema aprendió de un mundo que las trató de manera desigual.

La evaluación de sesgos algorítmicos busca detectar y corregir estas distorsiones. A ello se suma que Mehrabi, Morstatter, Saxena, Lerman y Galstyan (2021), en una revisión exhaustiva sobre el sesgo y la equidad en el aprendizaje automático, identificaron múltiples fuentes de sesgo a lo largo del ciclo de vida de los sistemas. Algunas provienen de los datos: el sesgo de representación, cuando ciertos grupos están subrepresentados en la muestra; el sesgo de medición, cuando las variables se miden de manera distinta o con distinta calidad para diferentes grupos; el sesgo histórico, cuando los datos reflejan desigualdades sociales preexistentes.

Otras surgen del algoritmo y de las decisiones de diseño, como la elección de las variables o de la función que se optimiza. Y otras aparecen en la interacción con los usuarios, cuando las decisiones basadas en el modelo modifican el comportamiento de las personas y, con ello, los datos futuros. Los autores revisaron también las distintas definiciones de equidad propuestas en la literatura y los métodos para mitigar el sesgo en distintas etapas.

La definición de equidad no es única, y las distintas definiciones pueden resultar incompatibles entre sí. Una de ellas, la paridad demográfica, exige que la proporción de resultados favorables sea similar entre grupos; por ejemplo, que el porcentaje de emprendimientos con puntuación alta sea parecido entre negocios dirigidos por hombres y por mujeres. Otra, la igualdad de oportunidades exige que la tasa de verdaderos positivos sea similar, es decir, que los emprendimientos que efectivamente tendrán éxito reciban puntuaciones altas con la misma probabilidad en todos los grupos. Una tercera, la calibración por grupos, exige que una misma puntuación corresponda a la misma probabilidad real de éxito en cada grupo.

Cuando las tasas de éxito observadas difieren entre grupos, no es posible satisfacer simultáneamente todas estas condiciones, lo que obliga a elegir qué noción de equidad resulta más adecuada para cada aplicación (Mehrabi et al., 2021). Asimismo, la evaluación de sesgos en un modelo de predicción empresarial puede organizarse en varias etapas. En primer término, deben identificarse los atributos protegidos o sensibles que podrían dar lugar a discriminación, como el sexo, la edad, la etnia, la discapacidad o el territorio de origen, así como las variables que podrían actuar como sustitutas de esos atributos, como el código postal o el tipo de establecimiento educativo.

A continuación, se calculan las métricas de desempeño del modelo —exactitud, sensibilidad, especificidad, tasa de falsos positivos y de falsos negativos— de manera desagregada para cada grupo, y se comparan entre sí. Luego se analizan las explicaciones del modelo para determinar si los atributos sensibles o sus sustitutos influyen de manera relevante en las predicciones. Finalmente, se evalúan las consecuencias de las diferencias detectadas en términos de los derechos y oportunidades de las personas afectadas.

Las técnicas de inteligencia artificial explicable resultan útiles en esta evaluación. Los valores SHAP permiten cuantificar la contribución de cada variable a las predicciones, tanto de manera global como para cada caso individual, y pueden revelar si una variable aparentemente neutral está actuando como sustituta de un atributo protegido (Molnar, 2022). En este sentido, Bussmann et al. (2021) mostraron, en el ámbito de la evaluación del riesgo crediticio de pequeñas y medianas empresas, cómo estas explicaciones permiten examinar el comportamiento de los modelos y facilitar su supervisión por parte de las personas responsables. Además, cuando se detectan sesgos, existen distintas estrategias de mitigación.

Las de preprocesamiento actúan sobre los datos antes del entrenamiento, por ejemplo, reequilibrando la representación de los grupos o corrigiendo las variables con sesgo de medición. Las de procesamiento interno incorporan restricciones de equidad en el propio algoritmo de aprendizaje. Y las de posprocesamiento ajustan las predicciones del modelo ya entrenado, por ejemplo, utilizando umbrales de decisión diferenciados por grupo.

Al respecto, Park, Choi y Feng (2024), al abordar el desequilibrio de los datos en la predicción del éxito de empresas emergentes mediante redes generativas adversariales, ilustran cómo las técnicas de tratamiento de los datos pueden reducir los sesgos derivados de la subrepresentación de determinados casos. Dado que no existe una estrategia universalmente superior, su elección depende de la naturaleza del sesgo, de la definición de equidad adoptada y de las restricciones legales y prácticas.

La supervisión humana constituye el complemento indispensable de la evaluación técnica. Los procedimientos estadísticos, por sí solos, no pueden detectar todos los sesgos posibles ni resolver las cuestiones de valor que plantea la equidad. Las personas que supervisan los modelos aportan conocimiento del contexto, sensibilidad ante situaciones particulares y capacidad de juicio sobre la legitimidad de las decisiones.

El Reglamento europeo de inteligencia artificial exige que los sistemas de alto riesgo se diseñen de modo que puedan ser supervisados eficazmente por personas físicas durante su utilización,

con el fin de prevenir o reducir al mínimo los riesgos para la salud, la seguridad o los derechos fundamentales (Parlamento Europeo y Consejo de la Unión Europea, 2024).

La Recomendación de la UNESCO sobre la ética de la inteligencia artificial establece, en la misma línea, que la responsabilidad última de las decisiones debe recaer en las personas (UNESCO, 2021). En esta misma línea, la supervisión humana efectiva requiere más que la mera presencia de una persona que firme las decisiones del sistema. Exige que la persona supervisora comprenda el funcionamiento del modelo y sus limitaciones, disponga de la información necesaria para evaluar cada caso, tenga tiempo suficiente para hacerlo y cuente con la autoridad para apartarse de la recomendación algorítmica cuando lo considere justificado.

También requiere estar alerta frente al sesgo de automatización, la tendencia a aceptar acríticamente las recomendaciones del sistema. Los diseños que presentan la recomendación del algoritmo junto con su explicación, su nivel de confianza y los factores que podrían justificar una decisión distinta favorecen una supervisión más reflexiva.

En el contexto del emprendimiento ecuatoriano, la evaluación de sesgos tiene particular importancia. Los estudios sobre la actividad emprendedora del país muestran diferencias entre grupos de población en las motivaciones, los sectores y las condiciones en que se emprende (Lasio et al., 2024), y las desigualdades territoriales y sociales en el acceso a recursos son un rasgo persistente de la región.

Un modelo predictivo que no considere estas brechas podría reforzarlas, al asignar menores probabilidades de éxito a los grupos que históricamente han enfrentado mayores obstáculos. A la inversa, un modelo cuidadosamente evaluado y supervisado podría contribuir a identificar emprendimientos viables en esos grupos que los métodos tradicionales pasaban por alto. La Ley Orgánica de Protección de Datos Personales considera datos sensibles, entre otros, los relativos a la etnia, la salud y la orientación sexual, y establece restricciones específicas para su tratamiento (Asamblea Nacional del Ecuador, 2021), lo que debe tenerse en cuenta al diseñar las evaluaciones de sesgo.

De igual manera, la evaluación de sesgos y la supervisión humana deben ser procesos continuos, no ejercicios únicos realizados antes de la implantación. Los datos cambian, las poblaciones atendidas se modifican y los modelos pueden desarrollar nuevos sesgos con el tiempo. La vigilancia periódica de las métricas de equidad, la revisión de las reclamaciones de las

personas afectadas y la auditoría independiente de los sistemas de mayor riesgo permiten detectar a tiempo los problemas y corregirlos.

La documentación de los modelos facilita tanto la evaluación de sesgos como la supervisión. Una ficha descriptiva de cada modelo, que indique su finalidad, los datos con que fue entrenado, su desempeño general y por grupos, sus limitaciones conocidas y las condiciones en que no debe utilizarse, permite que las personas que lo supervisan y las que se ven afectadas por sus decisiones comprendan su alcance.

Del mismo modo, la documentación de los conjuntos de datos, con información sobre su origen, su período de cobertura, la población que representan y sus posibles sesgos, ayuda a anticipar problemas antes de que se manifiesten en las predicciones. Estas prácticas, que se han difundido en la comunidad de investigación sobre aprendizaje automático, resultan especialmente útiles cuando los modelos se transfieren entre organizaciones o se utilizan en contextos distintos de aquellos para los que fueron diseñados (Mehrabi et al., 2021), situación frecuente cuando se adaptan modelos desarrollados en otros países a la realidad latinoamericana.

Su aplicación exige identificar las fuentes de sesgo, medir el desempeño de los modelos de manera desagregada, elegir conscientemente la noción de equidad más adecuada, aplicar técnicas de mitigación y mantener una supervisión humana informada, con autoridad y atenta a los riesgos de la automatización.

Figura 5.6

Evaluación de sesgos algorítmicos y mecanismos de supervisión humana en modelos de predicción empresarial.



Fuente: Elaboración autores

5.7. Desarrollo de un prototipo de sistema de puntuación predictiva del emprendimiento (Predictive Entrepreneurship Score)

Las herramientas examinadas a lo largo de este libro pueden sintetizarse en un instrumento concreto: un sistema de puntuación que estime, de manera transparente y fundamentada, el potencial de éxito de un emprendimiento. Este instrumento, al que se denomina aquí Predictive Entrepreneurship Score o puntuación predictiva del emprendimiento, se concibe como un prototipo que articula los modelos de clasificación, los indicadores de viabilidad, los pronósticos de mercado y las proyecciones financieras en una medida única, acompañada de sus componentes y de una explicación de los factores que la determinan.

Su propósito no es emitir un veredicto sobre el futuro de un negocio, sino ofrecer a emprendedores, financiadores y programas de apoyo una referencia común para orientar decisiones y acompañamientos. Por otra parte, el diseño de la puntuación parte de un principio: ninguna dimensión aislada basta para anticipar el éxito empresarial.

La evidencia revisada en los capítulos anteriores muestra que la supervivencia y el crecimiento dependen de factores financieros, como la rentabilidad, la liquidez y el endeudamiento (Zizi et al., 2021; González Cedeño, 2026); de factores de mercado, como la demanda, la saturación del sector y el valor de los clientes; de factores relacionados con las capacidades de la persona emprendedora y su equipo; de factores tecnológicos, como el grado de digitalización; y de factores del entorno, como las condiciones del ecosistema y del territorio (Stam y van de Ven, 2021). La puntuación predictiva integra estas dimensiones en una estructura jerárquica, en la que cada dimensión se compone de varios indicadores y contribuye a una medida global.

La estructura propuesta comprende cinco dimensiones. La dimensión financiera incluye la rentabilidad sobre los activos, el margen neto, la liquidez proyectada, el apalancamiento y la relación entre los gastos financieros y las ventas. La dimensión de mercado reúne el crecimiento pronosticado de las ventas, la tasa de retención de clientes, el valor estimado de los clientes a largo plazo y un índice de saturación del sector en el territorio. La dimensión de capacidades considera la experiencia sectorial de la persona emprendedora, su formación en gestión, la composición del equipo y la existencia de redes de apoyo.

La dimensión tecnológica mide el grado de digitalización de los registros, la presencia en canales digitales y el uso de herramientas analíticas. Y la dimensión de entorno incorpora

indicadores del ecosistema local, como el acceso a financiamiento, la densidad de servicios de apoyo y las condiciones económicas del sector.

Paralelamente, para combinar estas dimensiones pueden seguirse dos vías complementarias. En primer lugar, destaca la vía del indicador compuesto, que sigue los pasos recomendados por el Centro Común de Investigación de la Comisión Europea: definición del marco conceptual, selección de los indicadores, tratamiento de los valores faltantes, análisis multivariante, normalización, ponderación, agregación y análisis de robustez (JRC, s. f.).

En esta vía, cada indicador se normaliza en una escala de cero a cien, se agregan los indicadores de cada dimensión y, a continuación, las dimensiones, utilizando pesos definidos por criterio experto o por métodos estadísticos. En segundo término, sobresale la vía del modelo predictivo, en la que un algoritmo de clasificación, como una regresión logística o un modelo de potenciación de gradiente, se entrena con datos históricos de emprendimientos para estimar la probabilidad de supervivencia a un horizonte determinado, y esa probabilidad, expresada en una escala de cero a cien, constituye la puntuación.

Cada vía tiene ventajas y limitaciones. El indicador compuesto es transparente, fácil de comunicar y aplicable, aunque no existan datos históricos suficientes para entrenar un modelo; sin embargo, sus pesos pueden ser arbitrarios y su relación con la probabilidad real de éxito no está garantizada. El modelo predictivo aprende de los datos la importancia relativa de cada factor y produce estimaciones calibradas de probabilidad; pero requiere datos históricos de calidad, puede reproducir sesgos y resulta menos intuitivo.

El prototipo propone combinarlas: utilizar el indicador compuesto como estructura de presentación, que permite a los usuarios comprender los componentes de la puntuación, y el modelo predictivo como motor de estimación, cuyos resultados se descomponen mediante valores SHAP en las contribuciones de cada dimensión. De este modo, la puntuación global corresponde a una probabilidad estimada, mientras que sus componentes muestran qué dimensiones la elevan o la reducen (Molnar, 2022).

Sobre esta base, la puntuación se acompaña de una escala interpretativa. Valores entre cero y treinta y cuatro podrían corresponder a un nivel de riesgo alto, que sugiere replantear aspectos centrales del modelo de negocio o de su financiamiento antes de continuar; valores entre treinta y cinco y cincuenta y nueve, a un nivel de riesgo moderado, que indica la necesidad de fortalecer determinadas dimensiones; entre sesenta y setenta y nueve, a un potencial favorable con aspectos por mejorar; y desde ochenta, a un potencial alto.

Estos umbrales son indicativos y deben calibrarse con datos reales, considerando los costos de los distintos tipos de error para cada aplicación, como se examinó en el segundo capítulo al tratar la validación de modelos (Chicco y Jurman, 2023).

La salida del sistema no se limita a un número. Para cada emprendimiento evaluado, el prototipo genera un informe que incluye la puntuación global con su intervalo de incertidumbre, la puntuación de cada dimensión, los tres factores que más contribuyen positivamente y los tres que más la reducen, una comparación con emprendimientos similares del mismo sector y territorio, y un conjunto de recomendaciones orientadas a mejorar las dimensiones más débiles.

El informe puede incluir, además, simulaciones del efecto que tendrían determinadas acciones sobre la puntuación: por ejemplo, cuánto aumentaría la probabilidad de supervivencia si se redujera el endeudamiento o se incorporaran canales de venta digitales. Desde esta perspectiva, la implementación técnica del prototipo puede realizarse con herramientas accesibles. Un flujo de trabajo sencillo comienza con un formulario de captura de información, complementado con la importación de datos de facturación electrónica y de registros contables cuando estén disponibles. Los datos se validan y normalizan en una base de datos estructurada.

El modelo de clasificación, entrenado previamente con datos históricos, calcula la probabilidad de supervivencia, y un módulo de explicación descompone la predicción por dimensiones. Un módulo de simulación genera el intervalo de incertidumbre y los escenarios alternativos. Por último, un generador de informes presenta los resultados en un tablero o en un documento descargable. Las bibliotecas de código abierto para aprendizaje automático y visualización permiten construir este flujo sin licencias costosas.

El uso de la puntuación debe someterse a las reglas de gobernanza examinadas en los apartados anteriores. Cuando se utilice para decisiones con efectos significativos sobre las personas, como la concesión de un crédito o el acceso a un programa público de apoyo, deben garantizarse la información a los titulares, la posibilidad de solicitar una revisión humana y la ausencia de discriminación, conforme a la normativa de protección de datos (Asamblea Nacional del Ecuador, 2021; SPDP, 2026).

La evaluación periódica de sesgos debe verificar que la puntuación no perjudica sistemáticamente a determinados grupos de emprendedores. Y su presentación debe evitar que se interprete como una sentencia: una puntuación baja no significa que el negocio vaya a fracasar, sino que presenta riesgos que conviene atender. De forma complementaria, la puntuación predictiva del

emprendimiento puede cumplir funciones distintas para cada actor del ecosistema. Para la persona emprendedora, es una herramienta de autodiagnóstico que señala sus fortalezas y debilidades y le permite seguir su evolución en el tiempo.

Para una incubadora o un programa de apoyo, ofrece un criterio para priorizar el acompañamiento y medir el efecto de sus intervenciones. Para una entidad financiera, complementa la evaluación crediticia tradicional con información sobre dimensiones que no aparecen en los estados financieros. Y para los responsables de políticas públicas, la agregación de puntuaciones por sector o territorio permite identificar dónde se concentran los emprendimientos más vulnerables y orientar los recursos de fomento.

El desarrollo del prototipo puede seguir una hoja de ruta en tres fases. En una fase piloto, la puntuación se construye como un indicador compuesto con pesos definidos por un panel de personas expertas y se aplica a un grupo reducido de emprendimientos acompañados por una incubadora o un programa de apoyo, recogiendo sus valoraciones sobre la utilidad y la comprensión de los resultados.

En una segunda fase, con los datos acumulados sobre la evolución de esos emprendimientos, se entrena un modelo predictivo y se compara su desempeño con el del indicador compuesto. En una tercera fase, el sistema se amplía a una población mayor, se integra con fuentes administrativas de datos y se somete a evaluaciones periódicas de precisión y equidad. Esta progresión permite aprender en cada etapa, corregir los errores antes de que afecten a muchas personas y construir la confianza de los usuarios, condición indispensable para que la puntuación sea adoptada.

Combina la estructura transparente de un indicador compuesto con la capacidad de estimación de un modelo predictivo, descompone sus resultados en dimensiones comprensibles, expresa su incertidumbre y se somete a reglas de gobernanza que protegen los derechos de las personas evaluadas. Su valor práctico dependerá de su validación con datos reales, de su calibración para los contextos en que se aplique y de su uso como instrumento de orientación y apoyo, nunca como mecanismo de exclusión.

Figura 5.7

Desarrollo de un prototipo de sistema de puntuación predictiva del emprendimiento (Predictive Entrepreneurship Score).



Fuente: Elaboración autores

5.8. Validación empírica del modelo integral mediante estudios de caso, datos empresariales y simulaciones de escenarios

Un modelo de predicción solo adquiere legitimidad cuando demuestra que sus estimaciones se corresponden con la realidad. La validación empírica es el proceso mediante el cual se verifica esa correspondencia: se contrastan las predicciones con los resultados observados, se evalúa su precisión con métricas adecuadas, se examina su comportamiento en situaciones diversas y se identifican sus limitaciones.

Para el modelo integral de emprendimiento predictivo y para la puntuación que lo sintetiza, la validación debe combinar tres estrategias complementarias: el análisis estadístico con datos empresariales, los estudios de caso que permiten comprender en profundidad el funcionamiento del modelo en situaciones concretas y las simulaciones de escenarios que exploran su respuesta ante condiciones diferentes de las observadas. Vinculado con lo anterior, la validación con datos empresariales constituye la base cuantitativa.

Requiere disponer de información sobre un conjunto amplio de emprendimientos, con los valores de los indicadores en un momento inicial y el resultado observado en un horizonte posterior: si sobrevivieron, crecieron o cerraron. En el Ecuador, los estados financieros que las sociedades presentan a la Superintendencia de Compañías, Valores y Seguros han permitido estudiar los determinantes de la supervivencia empresarial mediante modelos Probit y de riesgos proporcionales (Zambrano et al., 2021; González Cedeño, 2026), y el registro estadístico de

empresas ofrece información sobre la dinámica de creación y cierre de unidades productivas (INEC, 2024).

Estas fuentes, complementadas con datos de programas de apoyo, incubadoras y entidades financieras, permitirían construir una base de validación adecuada, con las garantías de anonimización y confidencialidad que exige la normativa.

El procedimiento de validación estadística debe seguir las buenas prácticas examinadas en el segundo capítulo. Los datos se dividen en un conjunto de entrenamiento y otro de prueba, o se utiliza validación cruzada, de modo que el desempeño se evalúe con casos que el modelo no ha visto durante su construcción (James et al., 2021). Cuando los datos tienen una dimensión temporal, conviene realizar una validación fuera del tiempo, entrenando el modelo con cohortes de emprendimientos de años anteriores y evaluándolo con cohortes posteriores, para verificar que sus predicciones se mantienen ante cambios en las condiciones económicas.

Las métricas deben incluir la capacidad de discriminación, medida por el área bajo la curva ROC; el coeficiente de correlación de Matthews, más exigente cuando las clases están desequilibradas (Chicco y Jurman, 2023); la sensibilidad y la especificidad en los umbrales de decisión elegidos; y la calibración, que verifica si las probabilidades estimadas corresponden a las frecuencias observadas. En la práctica, para ilustrar este procedimiento, se realizó un ejercicio de simulación con fines exclusivamente didácticos.

Se generó una base sintética de tres mil emprendimientos con siete variables —rentabilidad sobre los activos, apalancamiento, liquidez, crecimiento de las ventas, nivel de digitalización, años de experiencia de la persona emprendedora y saturación del sector—, cuya relación con la supervivencia a tres años se definió siguiendo la dirección de los efectos documentados en la literatura revisada. Los datos no corresponden a empresas reales y sus resultados no deben interpretarse como evidencia empírica sobre el emprendimiento ecuatoriano. Con el 70 % de los casos se entrenaron una regresión logística y un modelo de potenciación de gradiente, y con el 30 % restante se evaluó su desempeño.

La regresión logística obtuvo un área bajo la curva ROC de 0,76, un coeficiente de Matthews de 0,38, una sensibilidad de 0,62 y una especificidad de 0,76; el modelo de potenciación de gradiente alcanzó un área bajo la curva de 0,74 y un coeficiente de Matthews de 0,33. El resultado, en el que el modelo más sencillo iguala o supera al más complejo, es coherente con la observación de que la ventaja de los algoritmos sofisticados depende de la existencia de

relaciones no lineales y de interacciones complejas en los datos (Grinsztajn et al., 2022), y recuerda que la elección del modelo debe basarse en la validación y no en su sofisticación aparente.

Los estudios de caso complementan la validación estadística con una comprensión cualitativa. Consisten en aplicar el modelo a un número reducido de emprendimientos reales, seleccionados para representar situaciones diversas, y analizar en profundidad sus resultados junto con las personas emprendedoras. Este análisis permite verificar si las explicaciones del modelo son coherentes con el conocimiento que tienen los emprendedores de sus negocios, identificar factores relevantes que el modelo no captura, evaluar la utilidad práctica de las recomendaciones y detectar posibles problemas de comprensión o de aceptación. Los estudios de caso resultan especialmente valiosos en las primeras etapas de implantación, cuando todavía no se dispone de datos suficientes para una validación estadística robusta.

En concordancia con ello, en el ejercicio didáctico, se aplicó la puntuación a tres perfiles hipotéticos que representan situaciones frecuentes en el emprendimiento latinoamericano. El primero, un comercio barrial con baja rentabilidad, endeudamiento elevado, escasa digitalización y ubicado en un sector muy saturado obtuvo una puntuación de 13 sobre 100, en el nivel de riesgo alto. El segundo, un servicio digital con rentabilidad y crecimiento favorables, endeudamiento moderado y baja saturación, alcanzó 73 puntos, en el nivel de potencial favorable.

El tercero, una pequeña manufactura local con indicadores intermedios y una persona emprendedora con amplia experiencia, obtuvo 44 puntos, en el nivel de riesgo moderado. Las explicaciones de la puntuación mostraron que, en el primer caso, la saturación del sector y el apalancamiento eran los factores que más reducían la probabilidad de supervivencia, mientras que en el tercero la experiencia de la persona emprendedora compensaba parcialmente una rentabilidad modesta. Este tipo de análisis permite que la puntuación se traduzca en recomendaciones diferenciadas.

Las simulaciones de escenarios permiten evaluar la robustez del modelo ante condiciones distintas de las observadas. Mediante el método Monte Carlo, se generan miles de variaciones de los indicadores de cada emprendimiento bajo distintos supuestos sobre el entorno —un escenario base, uno adverso con caída de la rentabilidad y del crecimiento, y uno favorable— y se observa la distribución de las puntuaciones resultantes (Fabianová et al., 2023).

En el ejercicio didáctico, con cinco mil simulaciones por escenario, la puntuación del servicio digital se ubicó en torno a 73 en el escenario base, descendió a 63 en el adverso y ascendió a 78 en el favorable; la de la manufactura local pasó de 44 en el escenario base a 32 en el adverso, cruzando hacia el nivel de riesgo alto; y la del comercio barrial se mantuvo en el nivel de riesgo alto en los tres escenarios. Estos resultados ilustran cómo la simulación revela la vulnerabilidad diferenciada de cada negocio frente a los cambios del entorno, información que una puntuación puntual no muestra.

A ello se suma que la validación debe incluir también la evaluación de la equidad. Las métricas de desempeño deben calcularse de manera desagregada por sexo de la persona emprendedora, territorio, sector y otras características relevantes, para verificar que el modelo funciona con similar precisión en todos los grupos y que no asigna sistemáticamente puntuaciones menores a determinados colectivos sin que ello se justifique por diferencias reales en su probabilidad de éxito (Mehrabi et al., 2021). Si se detectan diferencias, deben investigarse sus causas y aplicarse las medidas de mitigación examinadas en el apartado sobre sesgos algorítmicos.

La validación no es un acto único, sino un proceso continuo. Una vez implantado, el modelo debe someterse a un seguimiento periódico que compare sus predicciones con los resultados observados, detecte la pérdida de precisión derivada de cambios en las condiciones económicas o en la población de emprendimientos evaluados, y active su recalibración cuando sea necesario.

El Marco de Gestión de Riesgos de Inteligencia Artificial del NIST sitúa la medición y el monitoreo continuo de los sistemas entre sus funciones centrales (NIST, 2023). Los registros de las decisiones tomadas a partir de la puntuación y de sus consecuencias permiten, además, evaluar si su uso mejora efectivamente los resultados de los emprendimientos, que es, en última instancia, la prueba de su valor.

Los datos disponibles suelen cubrir solo a las empresas formalmente constituidas, mientras que una proporción considerable de los emprendimientos latinoamericanos opera en la informalidad por lo que los resultados observados pueden estar influidos por las propias decisiones tomadas a partir del modelo: si los emprendimientos con puntuación baja reciben menos financiamiento, su menor supervivencia puede deberse a esa restricción y no a su viabilidad intrínseca. Esta base serviría no solo para validar el modelo integral, sino también para orientar las políticas públicas de fomento y para generar conocimiento académico sobre los determinantes del éxito

empresarial en la región, contribuyendo a cerrar la brecha de investigación aplicada que existe en este campo.

Figura 5.8

Validación empírica del modelo integral mediante estudios de caso, datos empresariales y simulaciones de escenarios.



Fuente: Elaboración autores

REFERENCIAS BIBLIOGRÁFICAS

Agrawal, A., Gans, J., y Goldfarb, A. (2022). *Power and prediction: The disruptive economics of artificial intelligence*. Harvard Business Review Press.

Asamblea Nacional del Ecuador. (2020, 28 de febrero). *Ley Orgánica de Emprendimiento e Innovación*. Registro Oficial, Suplemento n.º 151.

Asamblea Nacional del Ecuador. (2021, 26 de mayo). *Ley Orgánica de Protección de Datos Personales*. Registro Oficial, Quinto Suplemento n.º 459.

Asamblea Nacional del Ecuador. (2023, 7 de febrero). *Ley Orgánica de Transparencia y Acceso a la Información Pública*. Registro Oficial, Segundo Suplemento n.º 245.

Banco Mundial. (2021). *Informe sobre el desarrollo mundial 2021: Datos para una vida mejor*. Banco Mundial. https://wdr2021.worldbank.org/the-report

Bantis, E., Clements, M. P., y Urquhart, A. (2023). Forecasting GDP growth rates in the United States and Brazil using Google Trends. *International Journal of Forecasting, 39*(4), 1909–1924. https://doi.org/10.1016/j.ijforecast.2022.10.003

Bidgoli, M. R., Vanani, I. R., y Goodarzi, M. (2024). Predicting the success of startups using a machine learning approach. *Journal of Innovation and Entrepreneurship, 13*(1), 1–27. https://doi.org/10.1186/s13731-024-00436-x

Breque, M., De Nul, L., y Petridis, A. (2021). *Industry 5.0: Towards a sustainable, human-centric and resilient European industry*. Comisión Europea, Dirección General de Investigación e Innovación. https://research-and-innovation.ec.europa.eu/knowledge-publications-tools-and-data/publications/all-publications/industry-50_en

Bussmann, N., Giudici, P., Marinelli, D., y Papenbrock, J. (2021). Explainable machine learning in credit risk management. *Computational Economics, 57*(1), 203–216. https://doi.org/10.1007/s10614-020-10042-0

Carayannis, E. G., y Morawska-Jancelewicz, J. (2022). The futures of Europe: Society 5.0 and Industry 5.0 as driving forces of future universities. *Journal of the Knowledge Economy, 13*(4), 3445–3471. https://doi.org/10.1007/s13132-021-00854-2

Centro Común de Investigación [JRC]. (s. f.). *10-step guide: Composite indicators & scoreboards*. Comisión Europea, Knowledge for Policy. https://knowledge-for-policy.ec.europa.eu/composite-indicators/toolkit_en/navigation-page/10-step-guide_en

Chalmers, D., MacKenzie, N. G., y Carter, S. (2021). Artificial intelligence and entrepreneurship: Implications for venture creation in the fourth industrial revolution. *Entrepreneurship Theory and Practice, 45*(5), 1028–1053. https://doi.org/10.1177/1042258720934581

Chicco, D., y Jurman, G. (2023). The Matthews correlation coefficient (MCC) should replace the ROC AUC as the standard metric for assessing binary classification. *BioData Mining, 16*. https://doi.org/10.1186/s13040-023-00322-4

Chico, J. R., y Altamirano, D. J. (2023). Factores de riesgo en la mortalidad de las empresas productivas del cantón Ambato. *Journal of Science and Research, 8*(3), 122–141. https://revistas.utb.edu.ec/index.php/sr/article/view/2919

Cueva, A. A., Feijoo, I. M., Sarmiento, B. C., y Mite, M. T. (2025). Uso de la regresión logística para predecir el éxito de modelos de negocio en microempresas. *Revista Universidad de Guayaquil, 139*(1), 1–9. https://doi.org/10.53591/rug.v139i1.1296

del Olmo-García, F., Crecente-Romero, F. J., del Val-Núñez, M. T., y Sarabia-Alegría, M. (2023). Entrepreneurial activity in an environment of digital transformation: An analysis of relevant factors in the euro area. *Humanities and Social Sciences Communications, 10*, Artículo 755. https://doi.org/10.1057/s41599-023-02270-0

Dwivedi, R., Dave, D., Naik, H., Singhal, S., Rana, O., Patel, P., Qian, B., Wen, Z., Shah, T., Morgan, G., y Ranjan, R. (2023). Explainable AI (XAI): Core ideas, techniques, and solutions. *ACM Computing Surveys, 55*(9). https://doi.org/10.1145/3561048

Erwin, K., y Engelbrecht, A. (2023). Meta-heuristics for portfolio optimization. *Soft Computing, 27*(24), 19045–19073. https://doi.org/10.1007/s00500-023-08177-x

Fabianová, J., Janeková, J., Fedorko, G., y Molnár, V. (2023). A comprehensive methodology for investment project assessment based on Monte Carlo simulation. *Applied Sciences, 13*(10), Artículo 6103. https://doi.org/10.3390/app13106103

Firmansyah, E. B., Machado, M. R., y Moreira, J. L. R. (2024). How can artificial intelligence (AI) be used to manage customer lifetime value (CLV)? A systematic literature review. *International Journal of Information Management Data Insights, 4*(2), Artículo 100279.

Fornari, F., Compagnucci, I., Callisto De Donato, M., Bertrand, Y., Beyel, H. H., Carrión, E., Franceschetti, M., Groher, W., Grüger, J., Kilic, E., Koschmider, A., Leotta, F., Li, C.-

Y., Lugaresi, G., Malburg, L., Mangler, J., Mecella, M., Pastor, O., Riss, U., … Valderas, P. (2024). Digital twins of business processes: A research manifesto. *Internet of Things*, Artículo 101477. https://doi.org/10.1016/j.iot.2024.101477

Giuggioli, G., y Pellegrini, M. M. (2023). Artificial intelligence as an enabler for entrepreneurs: A systematic literature review and an agenda for future research. *International Journal of Entrepreneurial Behavior & Research, 29*(4), 816–837. https://doi.org/10.1108/IJEBR-05-2021-0426

Global Entrepreneurship Monitor [GEM]. (2025). *Global Entrepreneurship Monitor 2024/2025 global report: Entrepreneurship reality check*. GEM. https://gemconsortium.org/report/51621

Global Entrepreneurship Monitor [GEM]. (2026). *Global Entrepreneurship Monitor 2025/2026 global report: From uncertainty to opportunity*. GEM. https://www.gemconsortium.org/reports/latest-global-report

González Cedeño, G. (2026). Modelo de riesgos proporcionales de Cox aplicado a empresas de Manabí – Ecuador. *Revista Científica y Arbitrada de la Facultad de Ciencias Sociales, Derecho y Bienestar "Eloy Alfaro de Economía y Gestión"*, 83–102. https://revistas.uleam.edu.ec/index.php/economiaygestion/article/download/973/1226/4896

Goodell, J. W., Kumar, S., Lim, W. M., y Pattnaik, D. (2021). Artificial intelligence and machine learning in finance: Identifying foundations, themes, and research clusters from bibliometric analysis. *Journal of Behavioral and Experimental Finance, 32*, Artículo 100577. https://doi.org/10.1016/j.jbef.2021.100577

Grinsztajn, L., Oyallon, E., y Varoquaux, G. (2022). Why do tree-based models still outperform deep learning on typical tabular data? En *Advances in Neural Information Processing Systems 35 (NeurIPS 2022)*. https://arxiv.org/abs/2207.08815

Gupta, S., Modgil, S., Bhattacharyya, S., y Bose, I. (2022). Artificial intelligence for decision support systems in the field of operations research: Review and future scope of research. *Annals of Operations Research, 308*(1), 215–274. https://doi.org/10.1007/s10479-020-03856-6

Hambly, B., Xu, R., y Yang, H. (2023). Recent advances in reinforcement learning in finance. *Mathematical Finance, 33*(3), 437–503. https://doi.org/10.1111/mafi.12382

Hartmann, J., Heitmann, M., Siebert, C., y Schamp, C. (2023). More than a feeling: Accuracy and application of sentiment analysis. *International Journal of Research in Marketing, 40*(1), 75–87. https://doi.org/10.1016/j.ijresmar.2022.05.005

Hidalgo-Díaz, R., Alaminos, D., y Delgado-Gómez, J. (2025). New early warning tools for insolvency prevention in the European context. *Tourism & Management Studies, 21*(3), 53–70. https://doi.org/10.18089/tms.20250304

Houle, S., y Macdonald, R. (2023). *Identifying nascent high-growth firms using machine learning* (Staff Working Paper 2023-53). Bank of Canada. https://doi.org/10.34989/swp-2023-53

Hyndman, R. J., y Athanasopoulos, G. (2021). *Forecasting: Principles and practice* (3.ª ed.). OTexts. https://otexts.com/fpp3/

Instituto Nacional de Estadística y Censos [INEC]. (2024). *Registro Estadístico de Empresas 2023 (definitivo)*. INEC. https://www.ecuadorencifras.gob.ec/registro-estadistico-de-empresas-2023-definitivo/

James, G., Witten, D., Hastie, T., y Tibshirani, R. (2021). *An introduction to statistical learning: With applications in R* (2.ª ed.). Springer. https://doi.org/10.1007/978-1-0716-1418-1

Lasio, V., Amaya, A., Espinosa, M. P., Mahauad, M. D., y Sarango, P. (2024). *Global Entrepreneurship Monitor Ecuador 2023/2024*. ESPAE-ESPOL y UTPL. https://edes.utpl.edu.ec/wp-content/uploads/2024/07/Informe-GEM-Ecuador-2023.pdf

Makridakis, S., Spiliotis, E., y Assimakopoulos, V. (2022). M5 accuracy competition: Results, findings, and conclusions. *International Journal of Forecasting, 38*(4), 1346–1364. https://doi.org/10.1016/j.ijforecast.2021.11.013

Mehrabi, N., Morstatter, F., Saxena, N., Lerman, K., y Galstyan, A. (2021). A survey on bias and fairness in machine learning. *ACM Computing Surveys, 54*(6), Artículo 115. https://doi.org/10.1145/3457607

Moesmann, M., y Pedersen, T. B. (2024). *Data-driven prescriptive analytics applications: A comprehensive survey* (arXiv:2412.00034). arXiv. https://arxiv.org/abs/2412.00034

Molnar, C. (2022). *Interpretable machine learning: A guide for making black box models explainable* (2.ª ed.). https://christophm.github.io/interpretable-ml-book/

National Institute of Standards and Technology [NIST]. (2023). *Artificial intelligence risk management framework (AI RMF 1.0)* (NIST AI 100-1). Departamento de Comercio de los Estados Unidos. https://doi.org/10.6028/NIST.AI.100-1

Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura [UNESCO]. (2021). *Recomendación sobre la ética de la inteligencia artificial*. Adoptada por la Conferencia General en su 41.ª reunión, 23 de noviembre de 2021.

Organización para la Cooperación y el Desarrollo Económicos [OCDE]. (2025). *AI adoption by small and medium-sized enterprises* (OECD discussion paper for the G7). OECD Publishing.

Park, J., Choi, S., y Feng, Y. (2024). Predicting startup success using two bias-free machine learning: Resolving data imbalance using generative adversarial networks. *Journal of Big Data, 11*, Artículo 122. https://doi.org/10.1186/s40537-024-00993-8

Parlamento Europeo y Consejo de la Unión Europea. (2024). *Reglamento (UE) 2024/1689 por el que se establecen normas armonizadas en materia de inteligencia artificial (Reglamento de Inteligencia Artificial)*. Diario Oficial de la Unión Europea.

Pekin, S., y Şengül, A. (2024). The good, the better and the challenging: Insights into predicting high-growth firms using machine learning. *Borsa Istanbul Review, 24*, 47–60. https://doi.org/10.1016/j.bir.2024.12.001

Raisch, S., y Krakowski, S. (2021). Artificial intelligence and management: The automation–augmentation paradox. *Academy of Management Review, 46*(1), 192–210. https://doi.org/10.5465/amr.2018.0072

Roy, D., y Dutta, M. (2022). A systematic review and research perspective on recommender systems. *Journal of Big Data, 9*(1), 1–36. https://doi.org/10.1186/s40537-022-00592-5

Saad, M. H., Hagelaar, G., van der Velde, G., y Omta, S. W. F. (2021). Conceptualization of SMEs' business resilience: A systematic literature review. *Cogent Business & Management, 8*(1), Artículo 1938347. https://doi.org/10.1080/23311975.2021.1938347

Salminen, J., Mustak, M., Sufyan, M., y Jansen, B. J. (2023). How can algorithms help in segmenting users and customers? A systematic review and research agenda for algorithmic customer segmentation. *Journal of Marketing Analytics, 11*(4), 677–692. https://doi.org/10.1057/s41270-023-00235-5

Shepherd, D. A., Souitaris, V., y Gruber, M. (2021). Creating new ventures: A review and research agenda. *Journal of Management, 47*(1), 11–42. https://doi.org/10.1177/0149206319900537

Shepherd, D. A., y Gruber, M. (2021). The lean startup framework: Closing the academic–practitioner divide. *Entrepreneurship Theory and Practice, 45*(5), 967–998. https://doi.org/10.1177/1042258719899415

Shwartz-Ziv, R., y Armon, A. (2022). Tabular data: Deep learning is not all you need. *Information Fusion, 81*, 84–90. https://doi.org/10.1016/j.inffus.2021.11.011

Spada, I., Barandoni, S., Giordano, V., Chiarello, F., Fantoni, G., y Martini, A. (2023). What users want: A natural language processing approach to discover users' needs from online reviews. En *Proceedings of the International Conference on Engineering Design (ICED23)* (pp. 3879–3888). https://doi.org/10.1017/pds.2023.389

Stam, E., y van de Ven, A. (2021). Entrepreneurial ecosystem elements. *Small Business Economics, 56*(2), 809–832. https://doi.org/10.1007/s11187-019-00270-6

Superintendencia de Protección de Datos Personales [SPDP]. (2026, 10 de marzo). *Resolución n.º SPDP-SPD-2026-0009-R: Norma general para la garantía del derecho de protección de datos personales en el uso de sistemas de inteligencia artificial*. Registro Oficial n.º 240.

Veldkamp, L., y Chung, C. (2024). Data and the aggregate economy. *Journal of Economic Literature, 62*(2), 458–484. https://doi.org/10.1257/jel.20221580

Wang, L., Ma, C., Feng, X., Zhang, Z., Yang, H., Zhang, J., Chen, Z., Tang, J., Chen, X., Lin, Y., Zhao, W. X., Wei, Z., y Wen, J.-R. (2024). A survey on large language model based autonomous agents. *Frontiers of Computer Science, 18*(6), Artículo 186345. https://doi.org/10.1007/s11704-024-40231-1

Wankhade, M., Rao, A. C. S., y Kulkarni, C. (2022). A survey on sentiment analysis methods, applications, and challenges. *Artificial Intelligence Review, 55*(7), 5731–5780. https://doi.org/10.1007/s10462-022-10144-1

Wurm, B., Becker, M., Pentland, B. T., Lyytinen, K., Weber, B., Grisold, T., Mendling, J., y Kremser, W. (2023). Digital twins of organizations: A socio-technical view on

challenges and opportunities for future research. *Communications of the Association for Information Systems, 52*, 552–565. https://doi.org/10.17705/1CAIS.05223

Wurth, B., Stam, E., y Spigel, B. (2022). Toward an entrepreneurial ecosystem research program. *Entrepreneurship Theory and Practice, 46*(3), 729–778. https://doi.org/10.1177/1042258721998948

Zambrano, F. J., Sánchez, M. E., y Valls Martínez, M. del C. (2021). Factores explicativos de la supervivencia empresarial de la mipyme en Ecuador. *Estudios de Economía Aplicada, 39*(8). https://doi.org/10.25115/eea.v39i8.4061

Żbikowski, K., y Antosiuk, P. (2021). A machine learning, bias-free approach for predicting business success using Crunchbase data. *Information Processing & Management, 58*(4), Artículo 102555. https://doi.org/10.1016/j.ipm.2021.102555

Zhao, Q., Li, Y., y Dai, Z. (2023). Machine-learning-enabled intelligence computing for crisis management in small and medium-sized enterprises (SMEs). *Technological Forecasting and Social Change, 191*, Artículo 122492. https://doi.org/10.1016/j.techfore.2023.122492

Zizi, Y., Jamali-Alaoui, A., El Goumi, B., Oudgou, M., y El Moudden, A. (2021). An optimal model of financial distress prediction: A comparative study between neural networks and logistic regression. *Risks, 9*(11), Artículo 200. https://doi.org/10.3390/risks9110200

EDITORIAL ANDES COGNITIO



Adriana Yamilet Herrera Granizo

Economista (Universidad Nacional de Chimborazo), Máster en Economía Circular y Desarrollo Sostenible (Universidad Internacional de Valencia) y Máster en Gestión de Proyectos (Universidad Internacional de Valencia). Ha trabajado como Consultor Senior en INFODATOS), desarrollando actividades de consultoría especializada y asesoría empresarial en el sector privado. Jefe de Auditoría en Hiladera Guijarro, responsable de la planificación, ejecución y supervisión de procesos de auditoría interna. Editor en CEFRO, encargado de la revisión, edición y gestión de contenidos institucionales y editoriales. Editor en jefe del Rist-Dírham de los Emiratos Árabes Unidos y ScriptSpace Journals de la India, liderando la dirección editorial, coordinación de publicaciones y control de calidad de contenidos. Docente de la Universidad de las Fuerzas Armadas (ESPE), impartiendo asignaturas de formación profesional y contribuyendo al desarrollo académico de estudiantes universitarios. Autora de artículos sobre sostenibilidad, teoría de juegos y economía circular en revistas indexadas. Actualmente lidera proyectos sobre inversiones circulares en parques ecoindustriales y simbiosis hídrica.

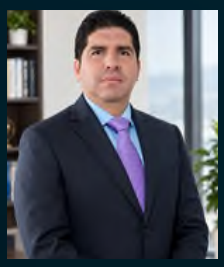


Gema Viviana Paula Alarcón

Profesional ecuatoriana con trayectoria en finanzas, contabilidad y auditoría, la autora es ingeniera en Banca y Finanzas e ingeniera en Contabilidad y Auditoría. Cuenta con dos maestrías, una en Contabilidad y Auditoría y otra en Auditoría Integral. Su formación le ha permitido especializarse en control interno, gestión de riesgos y auditoría aplicada.

Antes de incorporarse a la docencia universitaria, trabajó en el sector financiero cooperativo ecuatoriano. Su paso por la Cooperativa 4 de Octubre, Nueva Esperanza y UCACME le permitió profundizar en la gestión financiera y conocer de primera mano los procesos administrativos de estas instituciones. Actualmente es docente de la Universidad Nacional de Chimborazo, donde incorpora esa experiencia a la formación de estudiantes de Contabilidad, Auditoría y Finanzas.

Su labor investigadora se centra en la auditoría integral, la gestión financiera y el análisis administrativo. En la docencia y en sus trabajos académicos, conjuga el rigor técnico con la capacidad de análisis. La ética, el liderazgo y la vocación docente orientan su compromiso con la formación de profesionales en el Ecuador.



Lenin Rómulo Briones Borja

Magíster en Turismo con mención en Gestión Turística por la Universidad de Especialidades Turísticas e Ingeniero en Administración de Empresas Comercializadoras de Turismo. Profesional con más de dos décadas de experiencia en la industria turística y una sólida trayectoria docente. Actualmente se desempeña como docente en la Universidad de las Fuerzas Armadas ESPE, donde contribuye a la formación de estudiantes universitarios en el área de turismo. Su experiencia académica comprende áreas relacionadas con patrimonio, turismo sostenible, geografía, ecoturismo, circuitos y marketing turístico. En el ámbito profesional ha trabajado en agencias de viajes y empresas vinculadas con la operación y comercialización de servicios turísticos. Ha participado en proyectos de investigación y vinculación orientados a la planificación, aprovechamiento de recursos turísticos, desarrollo comunitario y marketing. Es coautor de publicaciones académicas sobre aprovechamiento turístico en la provincia de Imbabura y marketing aplicado al sector. Sus intereses académicos e investigativos se centran en la gestión y planificación turística, el patrimonio, la sostenibilidad y la puesta en valor de los recursos naturales y culturales.



Rafael Fernando Carrera Zurita

Docente universitario e investigador en turismo, es licenciado en Ecoturismo y Guía Nacional por la Pontificia Universidad Católica del Ecuador y magíster en Ecoturismo y Manejo de Áreas Naturales por la Universidad Tecnológica Equinoccial. Coordinó programas de intercambio cultural y trabajó en el Parque Nacional Cayambe-Coca como especialista provincial de áreas protegidas del Ministerio del Ambiente del Ecuador.

Ejerció la docencia en el Instituto Superior Tecnológico de Turismo y Patrimonio Yavirac y en la Escuela de Hotelería y Turismo de la Universidad de Las Américas. Actualmente es docente de la Universidad de las Fuerzas Armadas ESPE, donde cursa el Doctorado en Administración. Su investigación se centra en la gestión turística, la sostenibilidad de los destinos y el desempeño de las organizaciones del sector.

ISBN: 978-9907-9566-8-9