



Implementación de un modelo de scoring de crédito para Mexichem Colombia S. A. S.

Credit score model development for Mexichem Colombia S. A. S.

Por

Daniel Camilo Vargas Izquierdo¹

Trabajo de grado presentado como requisito parcial para obtener el título de
Magíster en Administración Financiera

Asesor docente

Brayan Ricardo Rojas Ormaza

Universidad EAFIT
Escuela de Finanzas, Economía y Gobierno
Maestría en Administración Financiera
Bogotá
2024

¹ dcvargasi@eafit.edu.co

© 2024 por Daniel Vargas
Todos los Derechos Reservados

Resumen

Este proyecto buscó implementar un modelo de scoring de crédito para Mexichem Colombia S. A. S. utilizando técnicas de machine learning para predecir la probabilidad de impago de empresas. Se compararon cuatro algoritmos: árboles de decisión, random forest, gradient boosting y redes neuronales, cada uno con características únicas en cuanto a precisión y manejo de datos complejos. La investigación incluyó la selección y evaluación de variables relevantes mediante el índice de Gini y técnicas de eliminación recursiva para evitar el sobreajuste. Los resultados ayudaron a identificar el modelo más eficaz para predecir riesgos crediticios, optimizando las decisiones financieras.

Palabras claves: scoring de crédito, machine learning, impago, modelos predictivos, riesgo crediticio.

Abstract

This project sought to implement a credit scoring model for Mexichem Colombia SAS using machine learning techniques to predict the probability of default in companies. Four algorithms were compared: decision trees, random forest, gradient boosting and neural networks, each with unique characteristics in terms of accuracy and handling of complex data. The research included the selection and evaluation of relevant variables using the Gini index and recursive elimination techniques to avoid overfitting. The results helped to identify the most effective model to predict credit risks, optimizing financial decision-making.

Keywords: Credit scoring, machine learning, default, predictive models, credit risk.

Contenido

1.	Introducción	1
1.1	Descripción del problema.....	1
1.2	Objetivos.....	2
1.2.1	Objetivo general.....	2
1.2.2	Objetivos específicos	2
1.3	Justificación	2
2.	Marco teórico.....	4
2.1	Antecedentes.....	4
2.2	Marco legal	6
2.3	Marco contextual	6
2.4	Bases teóricas	7
2.4.1	Definición del riesgo de crédito.....	7
2.4.2	Clasificación del riesgo de crédito.....	7
2.4.3	Modelos de scoring de crédito	8
2.4.4	Tipos de modelos de scoring de crédito.....	8
2.4.5	Metodologías y técnicas.....	9
2.4.6	Indicadores financieros para evaluar el riesgo de crédito.....	10
3.	Metodología.....	12
3.1	Tipo y diseño de la investigación	12
3.2	Fases de la investigación	13
3.3	Desarrollo de la investigación	13
3.4	Proceso para la estimación de los modelos de maching learning.....	14
3.4.1	División del conjunto de datos.....	14
3.4.2	Método para optimización de los hiper-parámetros	15
3.4.3	Validación cruzada (cross-validation)	15
3.4.4	Selección de los hiper-parámetros óptimos	15
3.4.5	Estimación final del modelo	15
3.4.6	Evaluación del modelo final	16
4.	Resultados.....	17
4.1	Variables del estudio	17
4.2	Modelos de scoring implementados para el análisis de riesgo de crédito	17
4.2.1	Árboles de decisión.....	17
4.2.2	Random forest.....	18
4.2.3	Gradient boosting.....	19
4.2.4	Redes neuronales	20
4.3	Indicadores financieros relevantes para el cálculo de la PI.....	21
4.3.1	Variables con alto mean decrease Gini.....	22
4.4	Validación de los modelos de scoring mediante el software R para la selección del más confiable estadísticamente.....	24

4.4.1 Árboles de decisión.....	24
4.4.2 Random forest.....	24
4.4.3 Gradient boosting.....	25
4.4.4 Redes neuronales	25
4.5 Estimación de las probabilidades de incumplimiento (PI) y cálculo del scoring	27
4.6 Regla de decisión.....	29
5. Conclusiones y Recomendaciones.....	30
5.1 Conclusiones.....	30
5.2 Recomendaciones	31
Referencias	33
Apéndice.....	37
Política de crédito para Mexichem Colombia S. A. S.....	37

Índice de tablas

Tabla 1. Regla de decisión por número de empresas	29
---	----

Índice de figuras

Figura 1. Importancia de cada variable explicativa.....	21
Figura 2. Cantidad óptima de variables.....	23
Figura 3. Hiper-parámetros óptimos del modelo random forest	24
Figura 4. Hiper-parámetros óptimos del modelo gradient boosting.....	25
Figura 5. Hiper-parámetros óptimos del modelo de redes neuronales	26
Figura 6. Comparación de los modelos en la muestra de entrenamiento y testeo.....	26
Figura 7. Distribución de las probabilidades de incumplimiento (PI)	27
Figura 8. Distribución de las probabilidades de incumplimiento (PI) por impago	27
Figura 9. Distribución del scoring.....	28

1. Introducción

1.1 Descripción del problema

Las entidades que otorgan financiamiento a sus clientes deben llevar a cabo un análisis que les permita determinar los riesgos que se presentan al momento de conferirlo. Sin embargo, no contar con herramientas eficientes para una evaluación adecuada de los solicitantes de un crédito puede ser una barrera para ellas, lo que da paso a una inadecuada determinación de los riesgos y al incremento de las posibilidades de incumplimiento de pagos, que, en síntesis, se traducen en afectaciones financieras para ellas.

Queda sobreexpuesto que el acceso a la información sobre los riesgos de crédito de una solicitud cobra relevancia en la gestión financiera, puesto que permitirá determinar de forma acertada el otorgamiento de un crédito de forma segura. Frente a ello, es importante resaltar que adicionalmente se debe considerar el contexto donde se ejecutan las actividades para las cuales se solicita el crédito, ya que también puede incidir en la capacidad de pago del prestamista debido a factores asociados como rotaciones y niveles de inventarios, aspectos macroeconómicos y políticas públicas, entre otros, que pueden incrementar los riesgos de incumplimiento de sus obligaciones financieras.

Por ello, es evidente la necesidad de que las entidades que ofrecen crédito cuenten con modelos eficientes que les permitan determinar el score² de los solicitantes. Esta realidad se presenta en Mexichem Colombia S. A. S. —en adelante Mexichem o la Compañía—, una entidad que se encarga de financiar sus productos a compañías del sector de la construcción en Colombia y que actualmente viene realizando el análisis y la determinación del nivel de riesgo con un modelo general que no contempla las características particulares del sector que inciden en la capacidad de pago del prestamista potencial.

La implementación de dicho modelo general ha hecho que la Compañía otorgue créditos de forma ineficiente por causa de un cálculo inadecuado en el score de la entidad solicitante, que se está viendo reflejado en el porcentaje de incumplimiento de pagos por parte de sus clientes: en 2023, este se acercó al 13 %. Se comprende, entonces, que Mexichem requiere la implementación de un modelo de scoring de crédito que contemple todos los factores característicos del sector y los indicadores financieros de las compañías solicitantes para otorgar de forma eficiente el crédito de sus productos y mitigar los riesgos de incumplimiento de pagos que afectan sus finanzas y desempeño.

Según Yessenaman *et al.* (2023), un *modelo de scoring de crédito* es una herramienta analítica fundamental en el sector financiero, diseñada para estimar la probabilidad de que un solicitante de crédito incurra en incumplimiento de pago. Utilizando un enfoque basado en el análisis estadístico y matemático, este modelo integra datos históricos y variables financieras

² *Nota del autor.* En consideración al amplio uso de términos financieros en el idioma inglés, en este trabajo serán escritos sin cursivas.

empleando técnicas avanzadas como redes neuronales, bosques aleatorios y algoritmos de machine learning para generar un puntaje que categorice a los clientes según su nivel de riesgo crediticio.

La precisión de este modelo puede mejorarse incorporando fuentes de datos no convencionales como la información derivada del historial crediticio y su comportamiento frente a otros compromisos financieros adquiridos. Adicionalmente, los modelos híbridos y de ensamblaje, que combinan múltiples algoritmos, permiten generar predicciones más robustas y confiables (Simonov & Zboev, 2020). Sin embargo, su implementación enfrenta desafíos, entre ellos, la necesidad de personalización a mercados locales y la constante actualización del modelo para mantener su relevancia y efectividad en un entorno financiero en constante evolución.

Dicho lo anterior, esta investigación parte del hecho de la importancia del análisis de riesgo de crédito de las compañías que solicitan un crédito ante Mexichem; desde allí se formula la pregunta de investigación:

¿Cuál será el modelo de scoring que debe implementarse para la evaluación de las solicitudes de crédito de los clientes en Mexichem Colombia S. A. S.?

1.2 Objetivos

1.2.1 Objetivo general

Implementar un modelo de scoring para el análisis de riesgo de crédito que tienen las compañías del sector de la construcción que solicitan crédito con Mexichem Colombia S. A. S.

1.2.2 Objetivos específicos

- Describir los modelos de scoring implementados para el análisis de riesgo de crédito.
- Identificar los indicadores financieros relevantes para el cálculo de la probabilidad de incumplimiento (PI) de pago.
- Validar diferentes modelos de scoring mediante el software Rstudio para la selección del más confiable estadísticamente.
- Determinar, mediante el modelo de scoring validado, la probabilidad de incumplimiento (PI) de pago de las compañías del sector de la construcción que solicitan crédito con Mexichem Colombia S. A. S.

1.3 Justificación

La importancia del sector de la construcción está asociada al crecimiento y el desarrollo de las naciones y, en este sentido, su protección es importante para la sociedad en general. Esta realidad no difiere en Colombia y, por ello, las empresas que se encargan de ejecutar las actividades de esta industria deben ser eficientes en lo que respecta a la administración de sus recursos para garantizar su adecuado desempeño.

Para alcanzar este fin, las entidades que otorgan crédito deben utilizar herramientas adecuadas que permitan determinar el riesgo de crédito de las compañías solicitantes de un crédito y, por ello, la adaptación de modelos de scoring de crédito representan una alternativa eficaz para determinar su nivel de riesgo, ya que otorgan un puntaje a los clientes potenciales basado en el análisis estadístico y la captura de patrones del comportamiento y las características de clientes antiguos.

Con ello se puede comprender que la relevancia de la presente investigación se justifica desde el hecho de que el uso de estos modelos permite a las entidades de crédito asociarle al solicitante una probabilidad de incumplimiento —en adelante PI—, que puede ser expresada como un puntaje (score), de tal forma que, a mayor puntaje, menor PI. De este modo, aplicar este tipo de modelos en la Compañía representa una ventaja en lo concerniente a la precisión de las decisiones de otorgamiento de crédito, haciéndola más eficiente en términos de costo a través del ciclo de crédito y la reducción de los riesgos de afectaciones financieras por causas de incumplimiento de pagos.

Adicionalmente, la investigación es relevante porque incluye elementos sobre los aspectos e indicadores financieros que deben contemplarse para el análisis de otorgamiento de créditos —que están sujetos a las características propias de cada sector—, lo que le otorga a la academia un aporte en lo que respecta a la implementación adecuada de los modelos de scoring para mitigar los riesgos de sesgo de la información analizada y dar paso a errores en la asignación de créditos que generan incumplimientos que afectan la rentabilidad empresarial.

Por lo tanto, el desarrollo e implementación de un modelo de scoring de crédito adaptado a las necesidades específicas de Mexichem no solo contribuirá a mejorar la gestión del riesgo crediticio, sino que facilitará un entorno financiero más sólido, capaz de responder a las demandas de un sector en constante cambio, beneficiando tanto a la Compañía como a sus socios comerciales, y promoviendo un crecimiento sostenible y una mayor eficiencia en la utilización de los recursos financieros dentro del sector industrial colombiano.

2. Marco teórico

En este capítulo se abordan, en primer lugar, los conceptos fundamentales relacionados con el riesgo crediticio y la gestión de créditos, esenciales para comprender las prácticas y desafíos en el sector financiero. Aquí se exploran las metodologías y técnicas empleadas en la elaboración de modelos de scoring de crédito, con énfasis en las adaptaciones específicas para las cooperativas de ahorro y crédito, así como su aplicación práctica en contextos similares.

Posteriormente se analizan los modelos teóricos existentes sobre evaluación de riesgos y scoring crediticio, incluyendo sus ventajas y limitaciones. La revisión de la literatura permitirá identificar las mejores prácticas y enfoques teóricos que respaldarán el desarrollo del modelo propuesto; por lo tanto, se considerarán las normativas y directrices vigentes, a fin de asegurar que el modelo cumpla con los requisitos regulatorios y normativos del país.

En conjunto, este marco teórico proporciona el sustento necesario para la creación del modelo de scoring de crédito propuesto y permite una comprensión profunda de su impacto potencial en la gestión de riesgos y la mejora en la toma de decisiones crediticias, en tanto la integración de teorías, modelos y regulaciones relevantes garantiza la solidez y aplicabilidad del modelo.

2.1 Antecedentes

González López (2018) desarrolló un modelo estadístico para prever el incumplimiento en el pago de créditos en el ámbito del factoring, con el objetivo de construir un modelo predictivo que estimara la probabilidad de impago de documentos —mejorando así el proceso de admisión de operaciones— y una metodología que incluyera la comprensión del negocio, la obtención y preparación de datos y el modelamiento y la interpretación de los resultados, utilizando una técnica de scoring de crédito adaptada del sector bancario al factoring que permitió determinar el riesgo de crédito de los solicitantes de financiamiento.

Tulcanaza Aguilar (2021) buscó establecer directrices para desarrollar un modelo de scoring de originación dirigido a una cooperativa de ahorro y crédito. La propuesta tiene como objetivo identificar eficazmente a clientes potenciales y excluir a aquellos con alto riesgo, dado el crecimiento financiero, tecnológico e infraestructural reciente de estas entidades, con una metodología que incluye recopilación de datos, análisis estadístico y validación del modelo, y adaptando herramientas de scoring de crédito para ajustarse a las características específicas del segmento socioeconómico tres de Ecuador. Su estudio exploró variables claves para construir el modelo, evaluó la PI, examinó el marco normativo de la Superintendencia de Economía Popular y Solidaria (SEPS) de este país, y propuso una estructura que abarca una gestión más precisa del riesgo de crédito y la optimización en la toma de decisiones.

Leal Fica *et al.* (2018) abordaron el desarrollo de un modelo de evaluación crediticia para una empresa chilena dedicada a la producción y comercialización de productos derivados del asfalto, con el objetivo de proponer un modelo de scoring de crédito ajustado a las necesidades

específicas de la compañía Fantasía S. A. que redujera el riesgo de impagos y mejorara la gestión crediticia. El modelo propuesto muestra una efectividad significativa, con 81,82 % de los créditos otorgados a clientes, superando el nivel mínimo de evaluación establecido por la empresa, un hecho que indica una mejora en la gestión del riesgo crediticio.

Dinçer *et al.* (2020) evaluaron el desempeño del sector bancario turco mediante un enfoque basado en la herramienta *balanced scorecard*, utilizando el método *Analytic Network Process* (ANP). La metodología incluyó el análisis de treinta y tres bancos de depósito, para identificar qué perspectivas del *balanced scorecard* eran más apropiadas para los distintos tipos de entidades — estatales, privadas y extranjeras—. El estudio estimó las prioridades de los factores —que se normalizaron para cada perspectiva— y permitió obtener las prioridades finales y las clasificaciones para cada tipo de banco.

Los resultados de su trabajo mostraron que el factor financiero fue el más importante, con 65,7 %, seguido de la perspectiva del cliente, con 22,1 %. Los factores de aprendizaje y crecimiento ocuparon el tercer lugar, con 6,3 %, mientras que el factor interno fue el menos relevante, con 5,9 %. En cuanto a la propiedad de los bancos, los estatales mostraron el mejor desempeño, con 53,9 %, seguidos por los privados, con 36,1 %, mientras que los extranjeros quedaron en último lugar, con aproximadamente 10 %. Esta investigación proporcionó información valiosa para mejorar la gestión y las estrategias de los riesgos de crédito de los bancos turcos, así como para la formulación de políticas que promovieran un equilibrio más eficaz entre las diferentes perspectivas de desempeño financiero. (Dinçer *et al.*, 2020)

Rastegar y Faraji (2023) desarrollaron un modelo de *scoring* de crédito para clientes bancarios en países del continente asiático utilizando una red neuronal probabilística (*probabilistic neural network*, PNN) optimizada mediante un algoritmo genético, para estimar la PI y riesgo crediticio, particularmente en sistemas bancarios débiles. Para alcanzar este objetivo, los autores compararon ocho técnicas de *machine learning*, entre ellas árboles de decisión, *random forest*, *gradient boosting* y redes neuronales, con el propósito de evaluar su rendimiento en el contexto del *scoring* crediticio. Los resultados revelaron que el modelo optimizado no solo ofreció una alta precisión en la predicción de incumplimiento, sino que, además, su proceso de optimización presentó una mayor velocidad de convergencia.

En el ámbito nacional, Santiago y Urán (2021) desarrollaron un modelo de *scoring* de crédito para la empresa Grupo Factoring de Occidente S. A. S., con el objetivo de optimizar el proceso de evaluación del riesgo crediticio de sus clientes. El modelo se basa en una serie de variables relevantes que afectan la PI y están ajustadas específicamente a las características del negocio de factoring. Los resultados del análisis indican que el modelo propuesto es eficaz en la identificación de los clientes con mayor riesgo de impago, lo que le permitió a la empresa tomar decisiones más informadas sobre la concesión de créditos. Asimismo, la validación del modelo muestra una alta capacidad predictiva, mejorando significativamente la precisión en la evaluación del riesgo crediticio, un hecho que se traduce en una mayor eficiencia en la gestión de riesgos y en la reducción de pérdidas financieras asociadas a impagos.

Garzón Escobar (2020) desarrolló un modelo de *scoring* para evaluar el comportamiento de crédito en una empresa de servicios públicos domiciliarios que instala redes de gas y gestiona pagos

de medidores, con el objetivo de refinanciar deudas vencidas y mejorar los indicadores de morosidad, que presentaban un alto nivel de deterioro. Para lograrlo, aplicaron una regresión logística con un intervalo de confianza del 95 %, que les permitió identificar que las variables más influyentes en el scoring fueron el número de refinanciaciones y el número de veces en mora, y encontraron que un mayor número de refinanciaciones y períodos prolongados en mora incrementaba significativamente la PI, lo que daba a entender que estos usuarios no deberían recibir facilidades adicionales o nuevos créditos.

2.2 Marco legal

El marco legal para el desarrollo de un modelo de scoring de crédito en Colombia está fundamentado en una serie de normativas que regulan la protección al consumidor financiero, el manejo de datos personales y la gestión de riesgos crediticios.

La Superintendencia Financiera de Colombia (SFC) es la entidad encargada de regular y supervisar el sistema financiero del país. A través de circulares y resoluciones, la SFC establece lineamientos para la gestión del riesgo crediticio, incluyendo la implementación de sistemas de scoring de crédito. Al respecto, la Circular externa 011 de 2002 (Colombia, SFC, 2002) exige a las entidades financieras implementar el Sistema de Administración del Riesgo Crediticio (SARC), en el que deben incluir metodologías para evaluar la capacidad de pago de los solicitantes de crédito. (Ochoa *et al.*, 2010)

La Ley 1328 de 2009 (Colombia, Congreso de la República, 2009) establece las bases para la protección de los derechos de los consumidores de servicios financieros, asegurando transparencia y equidad en el proceso de concesión de crédito. Esta ley ha influido notablemente en el sistema financiero del país al impulsar una mayor transparencia y confianza entre los consumidores y las instituciones financieras; adicionalmente, al definir claramente los derechos de los consumidores, refuerza la estabilidad del sistema financiero y promueve una cultura de responsabilidad en la gestión del crédito.

La Ley 1676 de 2013 (Colombia, Congreso de la República, 2013) regula la gestión de cobros y el manejo de cuentas por cobrar, proporcionando directrices para reducir el riesgo de impago.

La Ley 1581 de 2012 (Colombia, Congreso de la República, 2012) y el Decreto 1377 de 2013 (Colombia, Presidencia de la República, 2013) aseguran la protección de datos personales, estableciendo principios y procedimientos para el manejo seguro de la información de los clientes.

Adicionalmente, las normativas de la SFC ofrecen directrices adicionales para la implementación de modelos predictivos y el uso de tecnologías financieras.

2.3 Marco contextual

En el marco del avance industrial en Colombia y la creciente importancia de gestionar de manera eficaz el riesgo crediticio, Mexichem, como parte del conglomerado global Orbia, desempeña un

importante papel como uno de los líderes en la producción y comercialización de productos químicos y petroquímicos en el país. Esta empresa, que se destaca por su enfoque innovador y su compromiso con la sostenibilidad en sus procesos productivos, enfrenta desafíos específicos en la gestión financiera debido a la volatilidad en los mercados de las materias primas y a las variaciones en la demanda industrial.

Para la Compañía, la adopción de modelos de scoring de crédito eficaces es fundamental para optimizar la asignación de créditos y gestionar de manera adecuada los riesgos financieros inherentes a sus operaciones, porque la implementación de estos modelos le permitirá mitigar los riesgos crediticios, mejorar el proceso de toma de decisiones en la concesión de créditos y mantener una estabilidad financiera esencial para su crecimiento continuo y su competitividad en el mercado.

En este contexto, resulta clave la colaboración con instituciones financieras que comprendan las particularidades de los sectores químico y petroquímico y las demandas del mercado global; de esta manera, Mexichem podrá obtener grandes beneficios al integrar prácticas avanzadas de evaluación crediticia para asegurar la solidez de su estructura financiera y fortalecer su posición dentro de la cadena de suministro global.

2.4 Bases teóricas

2.4.1 Definición del riesgo de crédito

Según Jarrow y Turnbull (1995), el *riesgo de crédito* es la posibilidad de que un solicitante de crédito no cumpla con sus obligaciones de pago según los términos acordados en el contrato de crédito, por lo que, este riesgo, fundamental en la gestión financiera y en la evaluación de la solidez de una institución crediticia, implica una evaluación detallada de la capacidad del solicitante para cumplir con sus obligaciones. En términos generales, el riesgo de crédito se disgrega en varios componentes, incluyendo el *riesgo de incumplimiento*: la probabilidad de que el solicitante no pague; el *riesgo de pérdida*: la cantidad que se perdería en caso de incumplimiento; y el *riesgo de recuperación*: la proporción del crédito recuperado después del incumplimiento. (Altman, 2000)

La medición del riesgo de crédito se realiza a través de diversas herramientas y técnicas que permiten cuantificar la PI y la pérdida esperada. Entre ellas se destaca el *credit scoring* —o scoring de crédito— como un método sistemático para evaluar la solvencia crediticia de un solicitante mediante el uso de modelos estadísticos que analizan una serie de variables financieras y cualitativas (Credit Suisse, 2013). Este modelo proporciona una puntuación que refleja la probabilidad de que el cliente al que se le otorga un crédito incumpla sus obligaciones, facilitando la toma de decisiones en la concesión del crédito.

2.4.2 Clasificación del riesgo de crédito

El riesgo de crédito se clasifica en categorías que individualmente abordan aspectos específicos del riesgo asociado con la actividad crediticia.

Riesgo de incumplimiento o riesgo de default

La probabilidad de que el solicitante de un crédito no cumpla con los pagos programados. Este riesgo se evalúa a través de modelos de scoring de crédito, que asignan una puntuación basada en el historial crediticio y otros factores relevantes. (Hu & Su, 2022)

Riesgo de pérdida

La cantidad que el prestamista perderá si el solicitante incumple, descontando cualquier recuperación esperada. (Sullivan, 2016)

Riesgo de contraparte

Es una categoría adicional que se basa en la posibilidad de que una de las partes, en un contrato financiero, no cumpla con sus obligaciones, afectando las transacciones financieras y los acuerdos pactados (Merton, 1974). Este riesgo se relaciona directamente con el scoring de crédito, ya que una evaluación adecuada de la contraparte ayuda a prevenir situaciones de incumplimiento.

La integración de estos riesgos en un modelo de scoring de crédito permite una evaluación holística del perfil crediticio del solicitante y facilita la gestión efectiva del riesgo crediticio en las instituciones financieras.

2.4.3 Modelos de scoring de crédito

La evolución de los modelos de scoring de crédito ha sido significativa desde sus inicios. En las décadas anteriores, los métodos tradicionales de evaluación crediticia se basaban principalmente en la experiencia subjetiva de los prestamistas y en el análisis manual de la información financiera del solicitante; sin embargo, con el avance tecnológico y el crecimiento exponencial de la disponibilidad de datos, estos modelos han experimentado una transformación considerable.

En sus primeras etapas, se centraban en parámetros simples como el historial de pagos y el nivel de endeudamiento, para evaluar la solvencia crediticia (Behr & Güttler, 2007). Con el tiempo, el desarrollo de técnicas estadísticas y la capacidad de procesamiento de datos permitieron la introducción de modelos más sofisticados.

En la década de 1980, el scoring de crédito comenzó a incorporar métodos estadísticos avanzados como la regresión logística, que ofreció una mayor precisión en la predicción del riesgo crediticio (Thomas, 2009). Hoy en día, los modelos han evolucionado para incluir técnicas basadas en el aprendizaje automático y la inteligencia artificial, que utilizan grandes volúmenes de datos no estructurados para mejorar la precisión y la capacidad predictiva de los modelos (Khandani *et al.*, 2010). Esta evolución refleja el constante avance en la capacidad de procesamiento de datos y el análisis predictivo, proporcionando a las instituciones financieras herramientas más robustas para la gestión del riesgo crediticio.

2.4.4 Tipos de modelos de scoring de crédito

Los modelos de scoring de crédito se pueden clasificar en las siguientes categorías según la metodología empleada:

Modelos de puntuación estadística

Similares al análisis discriminante y la regresión logística, que han sido utilizados durante mucho tiempo para evaluar la PI. Estos modelos clasifican a los solicitantes de un crédito en diferentes grupos basándose en variables financieras y no financieras para predecir su PI. (Altman, 1968)

Modelo de regresión logística

Modela la relación entre una variable dependiente binaria —incumplimiento o no incumplimiento— y un conjunto de variables independientes, ofreciendo probabilidades ajustadas de incumplimiento. (Cox, 1958).

Con la creciente disponibilidad de datos y el avance en técnicas computacionales, los modelos de scoring han incorporado métodos más sofisticados como los árboles de decisión y el aprendizaje automático.

Los *árboles de decisión* dividen el conjunto de datos en segmentos basados en decisiones sucesivas, proporcionando una representación clara de cómo diferentes variables influyen en la PI. (Breiman *et al.*, 1986)

Los modelos de *aprendizaje automático*, como las máquinas de soporte vectorial (support vector machines, SVM) y las redes neuronales, han demostrado ser efectivos en la captura de patrones complejos y no lineales en grandes volúmenes de datos, superando a los métodos tradicionales en términos de precisión y adaptabilidad. (Hastie *et al.*, 2009)

2.4.5 Metodologías y técnicas

Las metodologías para construir modelos de scoring de crédito han evolucionado con el tiempo incorporando técnicas estadísticas y de aprendizaje automático como la *regresión logística*, que sigue siendo una técnica fundamental, ya que modela la PI en función de un conjunto de variables predictoras proporcionando interpretaciones directas de los efectos de cada variable en el riesgo crediticio. (Hosmer & Lemeshow, 2000)

El *análisis discriminante*, aunque menos común en la actualidad, sigue siendo relevante para entender cómo las características financieras distinguen entre solicitantes buenos y solicitantes malos. (Feldman & Nagle, 2002)

Las técnicas más recientes, como el *aprendizaje automático*, han revolucionado la construcción de modelos de scoring al permitir la identificación de patrones complejos y la integración de múltiples fuentes de datos. Entre ellas se destacan las máquinas de soporte vectorial (SVM), que son especialmente útiles en la clasificación de solicitantes, al encontrar el hiper-plano óptimo que separa a los buenos de los malos en un espacio de características multi-dimensionales (Cortés & Vapnik, 1995). Y las redes neuronales, que imitan el funcionamiento del cerebro humano para capturar relaciones no lineales entre variables, mejorando la capacidad del modelo para predecir el riesgo crediticio con alta precisión. (Rumelhart *et al.*, 1986)

Nuevas técnicas de scoring

Entre las nuevas técnicas de scoring se encuentran innovaciones como el uso de algoritmos de aprendizaje automático, que permiten procesar grandes volúmenes de datos y descubrir patrones complejos que podrían ser pasados por alto por las técnicas tradicionales. Asimismo, modelos avanzados como random forest, gradient boosting y redes neuronales han demostrado una notable eficacia al integrar una amplia gama de variables y ajustar sus predicciones en función de datos actualizados.

La incorporación de datos no convencionales como los provenientes de las redes sociales, las transacciones digitales y los comportamientos en línea está enriqueciendo las evaluaciones crediticias y proporcionando una visión más detallada del perfil del solicitante. Además, las técnicas de modelado híbrido y ensamblaje, que combinan diversos algoritmos para optimizar la precisión y minimizar el riesgo de sobreajuste, están ganando terreno. También están surgiendo metodologías de scoring probabilístico como las basadas en redes neuronales probabilísticas optimizadas mediante algoritmos genéticos, que ofrecen herramientas poderosas para calcular la PI y gestionar el riesgo crediticio de manera más efectiva.

Estas metodologías avanzadas no solo mejoran la precisión de las predicciones, sino que facilitan la adaptación a un entorno financiero en constante cambio, promoviendo una gestión de riesgos más robusta y eficiente.

2.4.6 Indicadores financieros para evaluar el riesgo de crédito

Como fue mencionado, la determinación de la adjudicación de un préstamo debe llevarse a cabo a través de una evaluación profunda que debe tener en consideración indicadores financieros que proporcionen una visión integral de la salud financiera de la empresa solicitante, ya que son esenciales para una evaluación precisa del riesgo de crédito. Dichos indicadores se describen a continuación.

Liquidez

La liquidez es esencial para evaluar la capacidad de una empresa para cumplir con sus obligaciones a corto plazo. En esta área son claves la razón corriente y la prueba ácida.

La razón corriente, que se calcula dividiendo los activos corrientes entre los pasivos corrientes, proporciona una visión general de la capacidad de una empresa para enfrentar sus obligaciones de corto plazo con sus activos líquidos (Brigham & Ehrhardt, 2016). Un valor superior a 1 indica una posición de liquidez adecuada, aunque valores excesivamente altos pueden señalar una utilización ineficiente de los recursos. (White *et al.*, 2003)

La prueba ácida, que excluye los inventarios del cálculo de activos corrientes, ofrece una evaluación más estricta de la liquidez. Este indicador es particularmente útil en industrias donde los inventarios no son fácilmente convertibles en efectivo. Un valor alto en la prueba ácida indica una sólida capacidad de pago y reduce el riesgo crediticio.

Solvencia

La solvencia se refiere a la capacidad de una empresa para cumplir con sus obligaciones financieras a largo plazo, ya que la razón de endeudamiento y la cobertura de intereses son indicadores claves en esta área.

La razón de endeudamiento, calculada como la proporción de deuda total respecto al capital total, muestra el grado de apalancamiento de una empresa (Ross *et al.*, 2016). Una ratio elevada puede señalar un mayor riesgo de incumplimiento, pero también puede reflejar una estrategia de financiamiento agresiva.

La cobertura de intereses mide la capacidad de una empresa para cubrir los intereses de su deuda con sus ganancias antes de intereses e impuestos (Koller *et al.*, 2010). Una alta cobertura sugiere que la empresa solicitante tiene una base sólida para manejar su deuda, disminuyendo el riesgo de incumplimiento y mejorando su perfil crediticio.

Rentabilidad

Los indicadores de rentabilidad como el margen de utilidad neta y el retorno sobre los activos (ROA) son cruciales para evaluar la eficiencia de una empresa en generar ganancias. Este margen, que se calcula dividiendo el beneficio neto entre las ventas totales, mide la proporción de ingresos convertida en ganancias netas (Brigham & Ehrhardt, 2016). Un margen alto indica eficiencia en la gestión de costos y la generación de beneficios.

El ROA, calculado como el beneficio neto dividido por el total de activos, evalúa la eficacia de una empresa en utilizar sus activos para generar ganancias. Un ROA alto refleja una buena gestión de los recursos y una mayor capacidad para cumplir con las obligaciones financieras y reduce el riesgo crediticio.

Actividad y Eficiencia

Los indicadores de actividad y eficiencia como la rotación de inventarios y la rotación de cuentas por cobrar proporcionan información sobre la gestión operativa de una empresa.

La rotación de inventarios, calculada como el costo de las ventas dividido por el inventario promedio, indica la velocidad con la que esta convierte inventarios en ventas (White *et al.*, 2003). Una alta rotación de inventarios sugiere una gestión eficiente de ellos.

La rotación de cuentas por cobrar, calculada como las ventas netas a crédito divididas por el saldo promedio de cuentas por cobrar, mide la eficiencia en la cobranza. Una alta rotación de cuentas por cobrar indica una eficaz gestión de crédito, mejorando el flujo de efectivo y reduciendo el riesgo de morosidad.

3. Metodología

El objetivo de esta investigación es desarrollar un modelo predictivo para estimar la probabilidad de impago de empresas mediante técnicas avanzadas de machine learning, creando un índice de scoring para cuantificar su capacidad de pago y apoyar la toma de decisiones crediticias. La investigación emplea un enfoque comparativo que analiza cuatro algoritmos prominentes: gradient boosting, árboles de decisión, random forest y redes neuronales.

El algoritmo gradient boosting adopta un enfoque secuencial en el cual los errores de los modelos anteriores son corregidos en cada iteración, mejorando progresivamente la precisión del modelo (Bai *et al.*, 2022). Este enfoque ha demostrado ser eficaz, sobre todo en situaciones donde las relaciones entre las variables son altamente no lineales.

El algoritmo árboles de decisión se destaca por su simplicidad e interpretabilidad, aunque puede ser propenso al sobreajuste en conjuntos de datos complejos. (Quinlan, 1986)

El algoritmo random forest mitiga esta limitación al promediar múltiples árboles, lo que genera predicciones más estables y robustas (Breiman, 2001).

Finalmente, el algoritmo redes neuronales, aunque menos interpretable, es muy potente para capturar patrones complejos y no lineales en los datos. (LeCun *et al.*, 2015)

Este enfoque permitirá evaluar la precisión y adaptabilidad de cada modelo frente a datos heterogéneos, proporcionando una visión integral para identificar el modelo más adecuado para predecir el riesgo de impago.

3.1 Tipo y diseño de investigación

Este estudio se clasifica como una investigación cuantitativa en el área de las finanzas corporativas, aplicada al análisis y la medición del riesgo de crédito a través de modelos de scoring con un enfoque comparativo. Su objetivo es desarrollar un modelo predictivo para estimar la probabilidad de impago de empresas utilizando técnicas avanzadas de machine learning, específicamente mediante la creación de un índice de scoring para facilitar la toma de decisiones crediticias.

La investigación se centra en desarrollar, mediante el software estadístico R, una herramienta práctica para la evaluación del riesgo crediticio utilizando técnicas avanzadas para mejorar la precisión en la predicción de impagos. Adicionalmente, la investigación evalúa y compara cuatro algoritmos prominentes en machine learning —gradient boosting, árboles de decisión, random forest y redes neuronales— para determinar cuál es el más adecuado para estimar la probabilidad de impago.

El alcance de la investigación abarca la evaluación y la comparación de diferentes técnicas de machine learning en términos de precisión y adaptabilidad frente a datos heterogéneos. Se busca así proporcionar una visión integral sobre cuál de los modelos es más eficaz para predecir el riesgo de impago de empresas, considerando las características específicas de cada algoritmo.

3.2 Fases de la investigación

- 1) Definición de variables. Selección y definición de las variables relevantes para la estimación de la PI utilizando indicadores financieros y no financieros. En esta fase se cumple el segundo objetivo específico de la investigación.
- 2) Desarrollo del modelo predictivo. Descripción e implementación de los algoritmos seleccionados —gradient boosting, árboles de decisión, random forest y redes neuronales— para estimar la probabilidad de impago. En esta fase se cumple el primer objetivo específico de la investigación.
- 3) Comparación de algoritmos. Evaluación comparativa de la precisión y adaptabilidad de cada modelo frente a datos heterogéneos, identificando sus fortalezas y limitaciones. En esta fase se cumple el tercer objetivo específico de la investigación.
- 4) Estimación de la probabilidad de impago por empresa. Una vez se seleccione el mejor modelo, este se utiliza para predecir la probabilidad de impago de cada una de las empresas en la muestra.
- 5) Creación del índice de scoring. Generación de un índice de scoring basado en las probabilidades de impago, para cuantificar la capacidad de pago y facilitar la toma de decisiones crediticias. El scoring se calculará utilizando la siguiente fórmula: $\text{Scoring} = (1 - \text{PI}) * 100$, donde PI es la probabilidad de impago. Este cálculo transformará las probabilidades estimadas en una escala de 0 a 100, donde un mayor puntaje indicará una menor PI y, por tanto, una mayor capacidad de pago.

Este enfoque permitirá una evaluación estandarizada y fácilmente interpretable de la capacidad crediticia de las empresas, facilitando la toma de decisiones basada en un índice claro y cuantificable. En esta fase se cumple el último objetivo específico de la investigación, con lo cual se alcanza el objetivo general.

3.3 Desarrollo de la investigación

El objetivo de esta investigación es construir un modelo predictivo que permita estimar la probabilidad de impago de empresas utilizando técnicas de machine learning. A partir de allí, se desarrolla un índice de scoring para cuantificar la capacidad de pago de cada empresa, que puede servir como herramienta para la toma de decisiones crediticias.

El modelo predictivo se estima con el software estadístico Rstudio, un entorno de desarrollo integrado (integrated development environment, IDE), diseñado para trabajar con el lenguaje de programación R, ampliamente utilizado en estadística y ciencia de datos, que proporciona una interfaz amigable y una serie de herramientas que facilitan el análisis de datos, la visualización y la implementación de modelos predictivos (RStudio Team, 2020). Además, cuenta con una amplia gama de paquetes que permiten realizar análisis que van desde procedimientos estadísticos tradicionales hasta técnicas avanzadas de machine learning. (Wickham *et al.*, 2019)

Para la estimación de la PI, se optó por un enfoque comparativo de modelos de machine learning eligiendo cuatro algoritmos populares y probados en la literatura crediticia: gradient boosting, árboles de decisión, random forest y redes neuronales. Cada uno de estos algoritmos ofrece diferentes características en términos de precisión, interpretabilidad y capacidad para manejar datos complejos, lo que los convierte en buenos candidatos para esta tarea.

El gradient boosting adopta un enfoque secuencial en el que los errores de los modelos anteriores son corregidos en cada iteración mejorando progresivamente su precisión (Bai *et al.*, 2022). Este enfoque ha demostrado ser eficaz, sobre todo en situaciones donde las relaciones entre las variables son altamente no lineales.

Los árboles de decisión representan una metodología sencilla e interpretable que permite descomponer el problema de clasificación en una serie de reglas secuenciales que facilitan la comprensión de las relaciones entre las variables (Chang *et al.*, 2016). No obstante, los árboles por sí solos tienden a ser propensos al sobreajuste, lo que lleva a considerar alternativas más robustas.

En este sentido, el random forest se presenta como una extensión más compleja que reduce el riesgo de sobreajuste al promediar los resultados de múltiples árboles, proporcionando así mayor estabilidad en las predicciones. (Behr & Weinblat, 2017)

Finalmente, las redes neuronales aportan una mayor flexibilidad, pues permiten capturar patrones complejos y relaciones no obvias entre las variables financieras (Sariev & Germano, 2020). Aunque menos interpretables, su capacidad de ajuste en situaciones de alta complejidad las convierte en un componente clave del análisis comparativo.

Este enfoque metodológico permitirá comparar estos cuatro algoritmos tanto en términos de precisión predictiva como en su capacidad para manejar datos heterogéneos y complejos, y ofrecerá una visión integral de cuál de ellos es más adecuado para predecir el riesgo de impago.

3.4 Proceso para la estimación de los modelos de machine learning

Para la estimación de los modelos de machine learning se deben seguir varios pasos de manera tal que el modelo estimado sea el mejor posible. Este proceso, que se sigue para cada uno de los cuatro algoritmos mencionados, se describe a continuación:

3.4.1 División del conjunto de datos

El primer paso es dividir los datos en tres subconjuntos:

- Conjunto de entrenamiento: para entrenar el modelo.
- Conjunto de validación: para evaluar el desempeño del modelo durante la búsqueda de hiper-parámetros.
- Conjunto de prueba: para evaluar el modelo final una vez seleccionados los hiper-parámetros óptimos.

3.4.2 Método para optimización de los hiper-parámetros

Los hiper-parámetros son configuraciones específicas del modelo que no son aprendidas directamente durante el entrenamiento. Algunos métodos comunes para optimizarlos son los siguientes:

- *Búsqueda en malla (grid search)*

Se define un conjunto finito de valores posibles para cada hiper-parámetro y se prueban todas las combinaciones posibles. Es exhaustivo, pero puede ser costoso en tiempo.

- *Búsqueda aleatoria (random search)*

En lugar de probar todas las combinaciones posibles, se seleccionan combinaciones aleatorias de los hiper-parámetros de manera controlada. Es más eficiente que el grid search, especialmente cuando el espacio de búsqueda es grande.

- *Optimización bayesiana*

Utiliza un enfoque probabilístico para modelar la relación entre los hiper-parámetros y el desempeño del modelo. Este método explora de manera más eficiente el espacio de los hiper-parámetros y ajusta el siguiente conjunto de pruebas basándose en los resultados previos.

En esta investigación se utiliza el primer método: búsqueda en malla (grid search), debido a su mayor sencillez y menor costo computacional.

3.4.3 Validación cruzada (cross-validation)

En este paso, para cada combinación de hiper-parámetros se aplica validación cruzada, en la que el conjunto de entrenamiento se divide en varias partes. El modelo se entrena en algunas de ellas y se valida en las otras. Este proceso se repite para asegurar que no se sobreajuste a una división específica del conjunto de datos, y el desempeño promedio de las diferentes divisiones se utiliza para evaluar la calidad de los hiper-parámetros.

3.4.4 Selección de los hiper-parámetros óptimos

Tras evaluar todas las combinaciones de hiper-parámetros —según el método de optimización elegido—, se seleccionan aquellos que maximicen el desempeño del modelo en el conjunto de validación. Esto puede ser en función de la métrica de evaluación elegida —como precisión, error cuadrático medio o área bajo la curva (area under the curve, AUC), dependiendo del tipo de problema: clasificación o regresión.

3.4.5 Estimación final del modelo

Una vez seleccionados los hiper-parámetros óptimos, se entrena el modelo final utilizando todo el conjunto de entrenamiento con ellos. Este modelo se puede evaluar con el conjunto de prueba para obtener una estimación confiable de su desempeño real en datos no vistos.

3.4.6 Evaluación del modelo final

Finalmente, el modelo con los hiper-parámetros optimizados se ensaya con el conjunto de prueba. Este paso proporciona una evaluación objetiva y final del desempeño del modelo en términos de la métrica seleccionada, asegurando que no está sobreajustado.

4. Resultados

Este capítulo presenta los resultados de esta investigación. En primer lugar, se muestra la descripción de los modelos de *maching learning* utilizados y, posteriormente, los resultados de la estimación de estos.

4.1 Variables del estudio

En esta investigación, cada variable representa un indicador financiero clave para evaluar la situación de una entidad o empresa.

Endeudamiento, que mide el nivel de obligaciones financieras respecto a los recursos propios, es útil para evaluar el riesgo de insolvencia. ROA (retorno sobre los activos) muestra la rentabilidad generada por los activos, reflejando la eficiencia operativa. RAZÓN_CORRIENTE indica la capacidad de la entidad o empresa para cubrir obligaciones a corto plazo. PRUEBA ÁCIDA es una versión más estricta de la liquidez, excluyendo inventarios. ROE (retorno sobre el patrimonio) mide la rentabilidad para los accionistas. MARGEN y MARGEN_NETO reflejan la eficiencia en la generación de beneficios operativos y finales, respectivamente. RECUPERACIÓN evalúa la capacidad para recuperar cartera vencida. CONCENTRACIÓN_CP mide la proporción de pasivos a corto plazo sobre los totales, indicando vulnerabilidad en la gestión del capital. RENTABILIDAD OPERACIONAL muestra los beneficios derivados de la actividad principal de la entidad o empresa antes de impuestos e intereses. MARGEN_OPERACIONAL evalúa qué tan rentables son las operaciones en relación con las ventas, Y CENTRAL_RIESGO refleja el historial crediticio o riesgo acumulado, relevante para evaluar el comportamiento financiero del deudor.

4.2 Modelos de scoring implementados para el análisis de riesgo de crédito

4.2.1 Árboles de decisión

Los árboles de decisión son un modelo de aprendizaje supervisado que se utiliza tanto para tareas de clasificación como de regresión (Chang *et al.*, 2016). En el contexto de la predicción de impago, permiten dividir el conjunto de datos en subconjuntos más homogéneos basados en una serie de reglas de decisión que generan una estructura jerárquica en forma de árbol.

El funcionamiento de un árbol de decisión se basa en seleccionar variables explicativas que mejor separan las observaciones en términos de la variable objetivo, en este caso, el impago. Este proceso de selección se realiza utilizando métricas como la Gini impurity o la entropía —para problemas de clasificación—, o el error cuadrático medio —para problemas de regresión—. Estas métricas ayudan a identificar las divisiones (*splits*) que maximizan la homogeneidad dentro de los subconjuntos resultantes.

Cada nodo del árbol representa una condición sobre una variable explicativa, y las ramas que salen él dividen el conjunto de datos en dos grupos. Este proceso continúa de manera recursiva, creando nuevas divisiones en los subconjuntos a medida que se avanza por el árbol (Chang *et al.*, 2016). El objetivo final es que las hojas del árbol —los nodos terminales— contengan observaciones que sean lo más homogéneas posible con respecto a la variable dependiente, es decir, que en cada hoja haya un grupo de empresas con características similares en términos de riesgo de impago.

Una ventaja significativa de los árboles de decisión es su interpretabilidad. Es fácil visualizar cómo cada variable contribuye a la predicción y entender las reglas que llevan a una clasificación específica. Sin embargo, una desventaja importante es que tienden a sobreajustarse a los datos de entrenamiento, especialmente cuando el árbol es profundo. Esto significa que pueden aprender demasiado bien las particularidades del conjunto de datos, reduciendo su capacidad para generalizar a datos nuevos. Para evitar esto, se suele aplicar un proceso de poda del árbol, eliminando las ramas que no aportan significativamente a la predicción final.

4.2.2 Random forest

Random forest es una técnica de aprendizaje supervisado basada en la combinación de múltiples árboles de decisión para mejorar la precisión predictiva y la capacidad de generalización del modelo (Behr & Weinblat, 2017). A diferencia de un solo árbol de decisión, que puede ser propenso al sobreajuste, random forest utiliza un enfoque de ensamble que genera un conjunto de árboles a partir de diferentes muestras de los datos y promedia sus predicciones, lo que reduce la varianza y mejora la robustez del modelo.

El proceso comienza con la creación de varios árboles de decisión a partir de subconjuntos aleatorios de los datos, mediante un proceso conocido como bootstrap sampling. En este, cada árbol se entrena con una muestra aleatoria del conjunto de datos que se genera con remplazo —algunas observaciones pueden aparecer en varios árboles, mientras que otras pueden ser omitidas—. Este enfoque permite que cada árbol se especialice en una parte distinta de los datos, lo que mitiga la tendencia al sobreajuste de los árboles individuales.

Además, en cada nodo de un árbol de decisión, random forest selecciona un subconjunto aleatorio de las variables explicativas para determinar las divisiones, lo que introduce más diversidad entre los árboles. Esta aleatorización adicional tiene el efecto de reducir la correlación entre ellos, lo que resulta en un modelo de ensamble más robusto.

Una vez que todos los árboles de decisión han sido entrenados, el modelo combina las predicciones de cada uno para generar la predicción final (Behr & Weinblat, 2017). En problemas de clasificación como el análisis de impago, se suele utilizar el voto mayoritario: la clase que recibe más votos de los diferentes árboles es la predicción final del random forest; y en problemas de regresión, se toma el promedio de las predicciones de los árboles.

Una de las principales ventajas de random forest es su robustez frente al sobreajuste. Al combinar múltiples árboles de decisión y promediar sus resultados, el modelo es menos sensible a las variaciones en los datos de entrenamiento y ofrece una mejor capacidad de generalización en

comparación con un solo árbol (Behr & Weinblat, 2017). Además, es capaz de manejar conjuntos de datos con un gran número de variables y detectar interacciones complejas entre ellas.

Random forest también proporciona una medida de importancia de las variables que puede ser útil para entender qué factores tienen un mayor impacto en la predicción de impago. Esto se calcula observando cuánto mejora la precisión del modelo, al utilizar una variable específica para hacer las divisiones en los nodos.

Una desventaja del modelo, en comparación con un único árbol de decisión, es la pérdida de interpretabilidad. Si bien es posible calcular la importancia de las variables, es mucho más difícil explicar cómo se llega a una predicción específica, ya que esta es el resultado de la combinación de muchos árboles.

4.2.3 Gradient boosting

Gradient boosting es un enfoque de ensamble que, al igual que random forest, combina múltiples modelos para mejorar la capacidad predictiva. Sin embargo, a diferencia de este último, donde los árboles son independientes entre sí, el proceso de gradient boosting es secuencial y aditivo: los modelos se entrenan uno tras otro, y cada nuevo modelo intenta corregir los errores de los modelos anteriores (Bai *et al.*, 2022). Esto convierte a gradient boosting en una técnica muy poderosa para manejar problemas complejos como la predicción del riesgo de impago.

El concepto central detrás de este modelo es *boosting*, que consiste en transformar múltiples modelos débiles —normalmente árboles de decisión simples y poco profundos— en un modelo fuerte mediante la combinación de estos (Bai *et al.*, 2022). El modelo funciona iterativamente ajustando un árbol a los residuos (errores) de los árboles anteriores, y en cada iteración se entrena un nuevo árbol que intenta corregir las predicciones erróneas del conjunto anterior, minimizando la función de pérdida a través del gradiente descendente.

El proceso comienza con un modelo inicial, que puede ser un simple promedio en problemas de regresión o una predicción uniforme en clasificación. Luego, en cada etapa, se calcula el error residual de las predicciones actuales con respecto a los datos reales y se entrena un nuevo árbol para ajustar esos errores. Este proceso se repite hasta que se haya agregado un número predeterminado de árboles o el modelo alcance un umbral de error satisfactorio.

Uno de los puntos fuertes de gradient boosting es su capacidad para manejar relaciones complejas y no lineales entre las variables (Bai *et al.*, 2022). Debido a que los modelos se ajustan iterativamente para corregir errores, esta herramienta es muy eficaz para identificar patrones ocultos en los datos que los modelos más simples podrían pasar por alto.

Otra ventaja es su flexibilidad. El modelo puede adaptarse a una amplia variedad de funciones de pérdida, lo que le permite ser útil en múltiples contextos como la clasificación de impago, la regresión y otros problemas.

Adicionalmente, permite un control preciso del sobreajuste mediante la regularización. Parámetros como el learning rate (la tasa de aprendizaje) permiten controlar la contribución de cada nuevo árbol evitando que el modelo se ajuste demasiado a los datos de entrenamiento. También se pueden limitar la profundidad de los árboles y el número total de iteraciones, lo que

mejora la capacidad de generalización; sin embargo, esta potencia y flexibilidad trae algunos costos.

En comparación con métodos como random forest, gradient boosting tiende a ser más susceptible al sobreajuste, especialmente si no se ajustan cuidadosamente hiper-parámetros como tasa de aprendizaje y número de iteraciones. Un modelo de gradient boosting mal configurado puede ajustarse demasiado a los datos de entrenamiento y fallar en su capacidad de generalizar a datos nuevos.

4.2.4 Redes neuronales

Las redes neuronales son modelos de aprendizaje profundo que se inspiran en el funcionamiento del cerebro humano para realizar predicciones y clasificación de datos. Estos modelos son particularmente potentes para capturar patrones complejos y no lineales en grandes volúmenes de datos (Sariev & Germano, 2020), lo que los convierte en una herramienta valiosa para la predicción del riesgo de impago.

Una red neuronal está compuesta por varias capas de nodos o neuronas, cada una conectada a las neuronas de la capa siguiente. Las redes neuronales típicas se dividen en tres tipos de capas, así:

Capa de entrada

Recibe los datos de entrada, en este caso, las variables financieras de las empresas. Aquí, cada nodo representa una característica o variable del conjunto de datos.

Capas ocultas

Las redes neuronales pueden tener una o más capas ocultas entre las capas de entrada y salida. Cada nodo, en una capa oculta, realiza cálculos sobre las entradas que recibe, pasando el resultado a la siguiente capa. Los nodos en estas capas aplican funciones de activación no lineales como la función ReLU (rectified linear unit) o la sigmoide, para introducir no linealidades en el modelo.

Capa de salida

Produce la predicción final del modelo. En el caso de la predicción del riesgo de impago, suele tener un solo nodo con una función de activación sigmoide para obtener una probabilidad entre 0 y 1.

El entrenamiento de una red neuronal se realiza mediante un proceso llamado propagación hacia adelante y retro-propagación (Sariev & Germano, 2020). Durante el primero, los datos de entrada se pasan a través de la red y las predicciones se comparan con las etiquetas reales usando una función de pérdida como la entropía cruzada para problemas de clasificación. Y en la retro-propagación se ajustan los pesos de las conexiones entre las neuronas mediante un algoritmo de optimización como el gradiente descendente, con el objetivo de minimizar la función de pérdida. Este proceso se repite durante múltiples iteraciones (o épocas) hasta que el modelo converge a un conjunto de pesos que proporciona buenas predicciones.

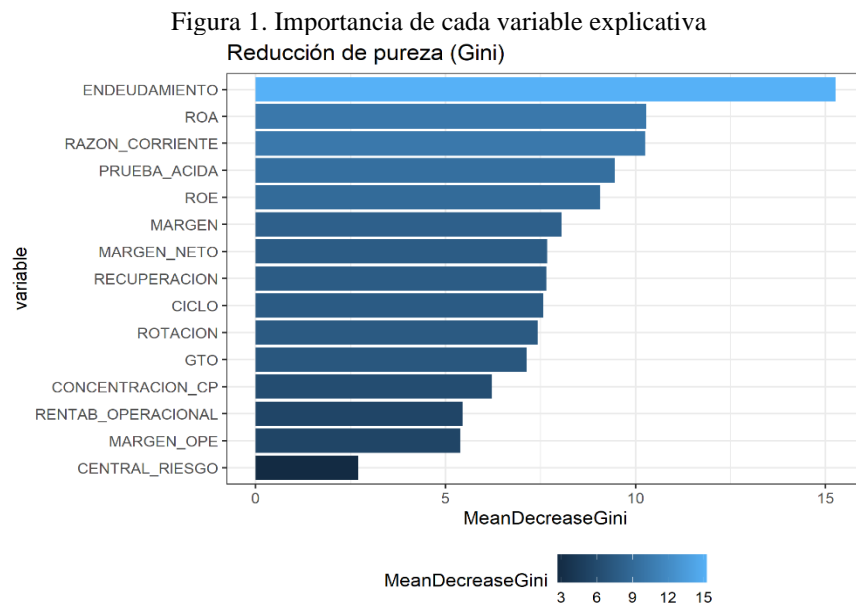
Las redes neuronales ofrecen gran capacidad para modelar relaciones complejas y no lineales entre las variables (Sariev & Germano, 2020). Esto es particularmente útil en la predicción del riesgo de impago, donde las relaciones entre las características financieras pueden ser altamente complicadas y no lineales. Además, son escalables y pueden manejar grandes conjuntos de datos y características, lo que permite capturar patrones en los datos que otros modelos pueden pasar por alto.

Una de las principales desventajas de las redes neuronales es su falta de interpretabilidad. A diferencia de los árboles de decisión o del random forest, es más difícil entender, debido a su estructura compleja, cómo se toman las decisiones. Esto puede ser un inconveniente cuando se requiere una explicación clara de cómo se determinan las predicciones.

Asimismo, el entrenamiento de las redes neuronales también puede ser computacionalmente intensivo y requerir mucho tiempo y recursos, especialmente cuando se trabaja con redes profundas y grandes volúmenes de datos. Además, la configuración de los hiper-parámetros —como el número de capas ocultas, el número de neuronas en cada capa y la tasa de aprendizaje— puede ser complicada y a menudo requiere de una búsqueda exhaustiva para encontrar la combinación óptima.

4.3 Indicadores financieros relevantes para el cálculo de la PI

Luego de hacer el análisis de la importancia relativa de los indicadores financieros de las empresas, se evidenciaron los resultados que se muestran en la Figura 1.



Fuente: elaboración del autor.

La Figura 1 presenta la importancia de las variables en un modelo de random forest estimado con toda la muestra, utilizando la métrica de *mean decrease Gini*. El índice de Gini es una métrica utilizada para medir la impureza en un nodo de un árbol de decisión. En términos simples, la pureza se refiere a qué tan “puros” u homogéneos son los datos en un nodo. Si todos ellos pertenecen a una sola clase, se dice que el nodo es puro. Este índice evalúa la probabilidad de clasificar incorrectamente un elemento si se selecciona al azar en un nodo. Un valor de Gini de 0 significa pureza completa: todos los elementos son de la misma clase. Un valor de Gini más alto indica mayor impureza.

El mean decrease Gini mide la disminución promedio en el índice causada por una variable en particular en el modelo random forest. En cada división o nodo, el modelo evalúa si la inclusión de una variable mejora la pureza del nodo, es decir, hace que los elementos en un nodo sean más homogéneos en términos de su clase.

4.3.1 Variables con alto mean decrease Gini

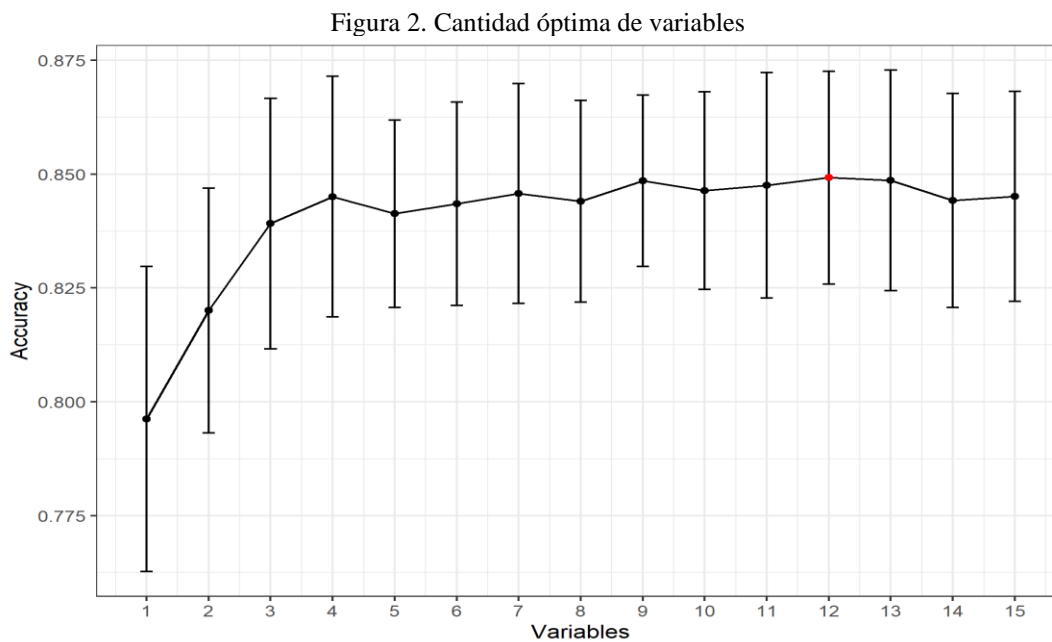
Estas variables son más importantes, ya que contribuyen significativamente a mejorar la pureza de los nodos en el árbol, lo que incrementa la capacidad del modelo para clasificar correctamente. Variables con un bajo índice de mean decrease Gini o cercano a cero (0) son menos importantes, ya que no mejoran mucho la pureza y, por tanto, no afectan significativamente el rendimiento del modelo. Los resultados de la Figura 1 sugieren lo siguiente:

- Endeudamiento es la variable más importante según el índice de Gini. Esta variable contribuye significativamente a mejorar la pureza de los nodos, lo que significa que, al usarla, el modelo mejora notablemente su capacidad para clasificar correctamente.
- ROA y RAZÓN_CORRIENTE también son variables claves, ya que su inclusión en las divisiones de los árboles de decisión mejora la pureza del modelo.
- PRUEBA_ACIDA, ROE, MARGEN, MARGEN_NETO y RECUPERACION siguen en importancia. Estas variables también ayudan a mejorar la clasificación del modelo, pero su impacto es menor en comparación con las primeras tres.
- CONCENTRACION_CP, RENTAB_OPERACIONAL, MARGEN_OPE y CENTRAL_RIESGO tienen un menor impacto en la pureza de los nodos, lo que sugiere que su contribución a mejorar la clasificación es relativamente baja.

Es importante hacer hincapié en que la prueba anterior no sugiere qué variables deberían excluirse del modelo: solo determina la importancia relativa de cada variable. Para determinar si algunas de ellas deberían eliminarse, se utiliza una técnica de eliminación recursiva que se utiliza en machine learning para la selección de características (features), y cuyo objetivo es identificar el subconjunto de indicadores más relevantes para el modelo eliminando aquellos que no aportan o tienen un impacto negativo en su rendimiento, de manera tal que se evite el overfitting.

Con base en este método se infiere el número óptimo de variables por incluir en el modelo. Este método sigue los siguientes pasos:

- 1) Entrenamiento del modelo. El algoritmo random forest se entrena con todas las variables disponibles.
- 2) Evaluación de importancia. Se mide la importancia de cada variable. En random forest, esto se hace normalmente a través de métricas como la impureza de Gini o la pérdida de precisión al permutar los valores de una variable, lo que permite estimar cuánto contribuye cada variable a las predicciones del modelo.
- 3) Eliminación de la variable menos importante. La variable con la menor importancia se elimina del conjunto de datos.
- 4) Repetición del proceso. El modelo se vuelve a entrenar sin la variable eliminada y se recalcula la importancia de las variables restantes. Este proceso se repite de forma recursiva hasta que se alcanza un número determinado de variables o un umbral de desempeño predefinido.
- 5) Selección del mejor subconjunto de variables. Finalmente se selecciona el subconjunto de variables que proporciona el mejor desempeño predictivo [Figura 2].



Fuente: elaboración del autor.

Según los resultados, la cantidad óptima de variables es 12, es decir, se deben eliminar tres: RENTAB_OPERACIONAL, CONCENTRACION_CP y CENTRAL_RIESGO, que son precisamente las menos importantes de acuerdo con la disminución promedio de la pureza —v. la Figura 1—.

Se procede ahora con la estimación de los modelos, ajustando el valor de los hiper-parámetros correspondientes para encontrar sus valores óptimos. Cabe aclarar que la estimación se realiza con el 90 % de la muestra, reservando el 10 % para el testeo. A su vez, con la muestra de entrenamiento se realiza validación cruzada con dos particiones y 15 repeticiones, por lo cual cada modelo se ajusta y evalúa $2 * 15 = 30$ veces, equivalentes a 30 re-muestras.

4.4 Validación de los modelos de scoring mediante el software R para la selección del más confiable estadísticamente

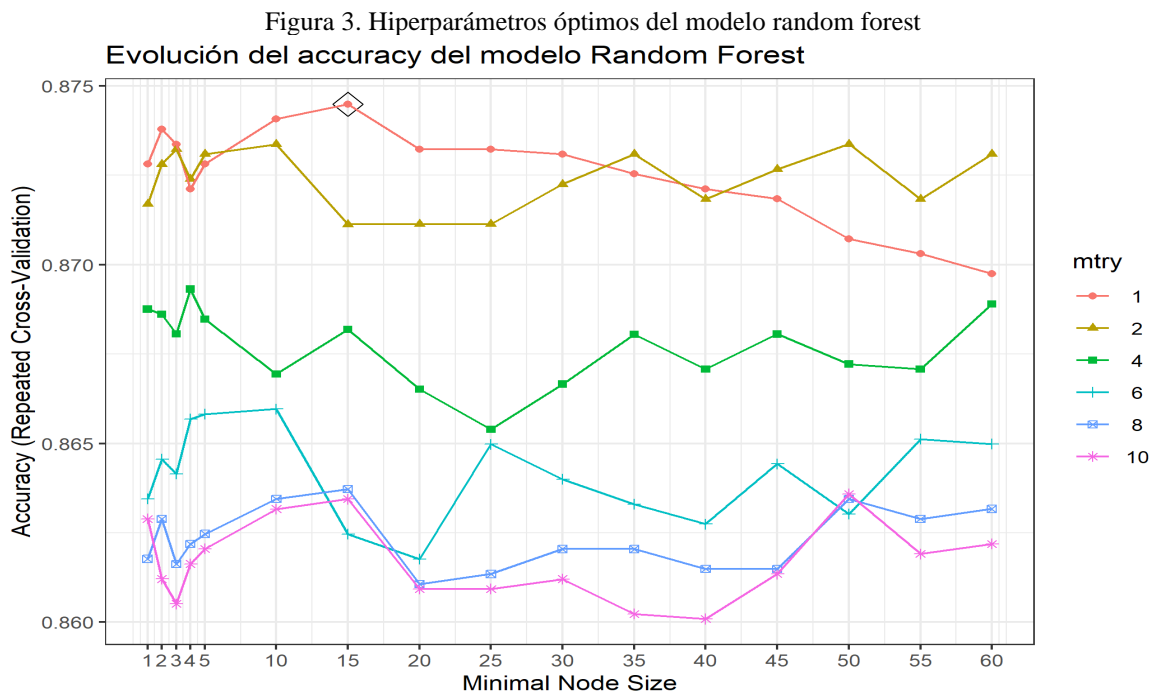
A continuación, se presenta el proceso de estimación y validación de cada modelo siguiendo los pasos señalados en la sección 3.4 de este trabajo: “Proceso para la estimación de los modelos de machine learning”.

4.4.1 Árboles de decisión

Este modelo no contiene hiper-parámetros, por lo cual, su estimación es simple. En el total de remuestros, el modelo de árboles tiene una precisión del 85 %.

4.4.2 Random forest

Para este modelo se ajustan los hiper-parámetros “mínimo tamaño de los nodos” y “mtry”. Este último puede tomar los valores {1, 2, 4, 6, 8, 10}, y el tamaño mínimo de los nodos es {1, 2, 3, 4, 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60}. La combinación de estos valores produce un total de 96 posibles modelos [Figura 3].



Fuente: elaboración del autor.

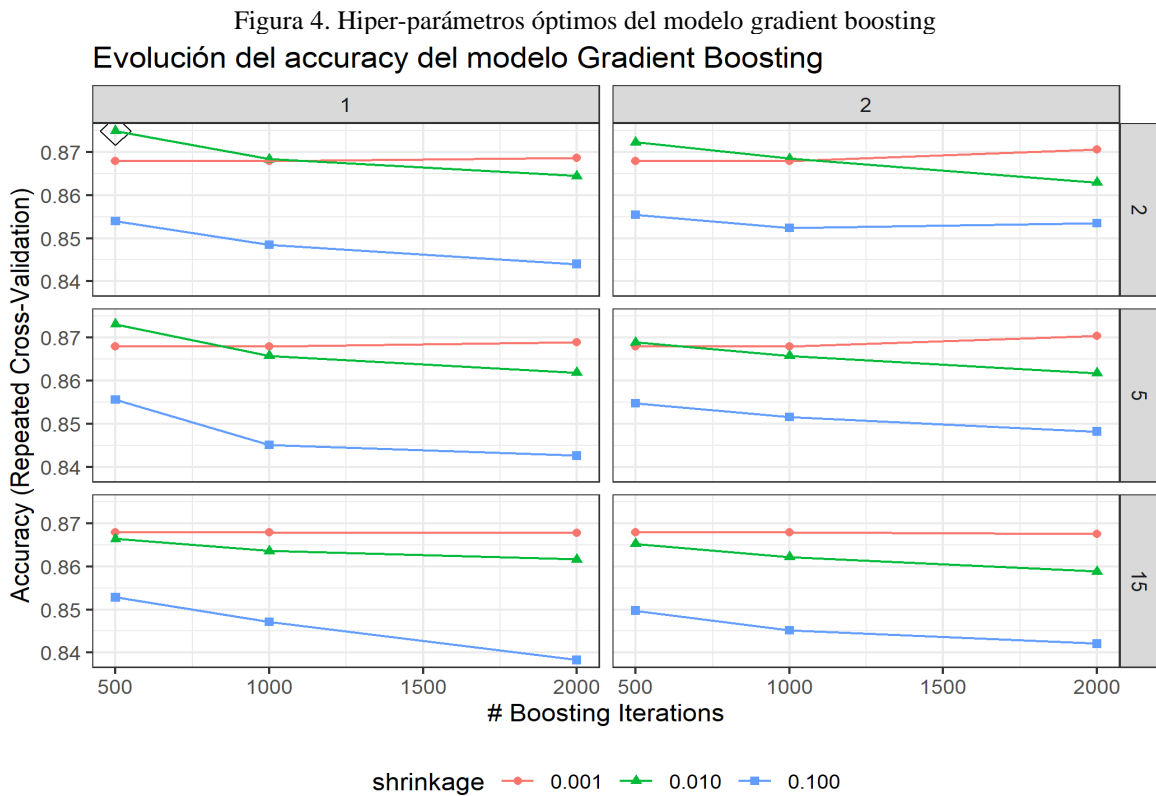
La combinación óptima de parámetros corresponde a {mtry = 1, min.node.size = 15}.

4.4.3 Gradient boosting

Para este modelo se tunean los hiper-parámetros “interaction.depth” —el número total de divisiones que tiene el árbol—, “n.trees” —el número de modelos que forman el ensamble—, “shrinkage” —la tasa de aprendizaje— y “n.minobsinnode” —el número mínimo de observaciones que debe tener un nodo para poder ser dividido—.

Estos parámetros pueden tomar los siguientes valores: interaction.depth = {1, 2}; n.trees = c(500, 1000, 2000); shrinkage = c(0,001, 0,01, 0,1); y n.minobsinnode = {2, 5, 15}. Su combinación produce un total de 54 posibles modelos.

La estimación arroja los resultados que se muestran en la Figura 4. La combinación óptima de parámetros corresponde a interaction.depth = 1, n.trees = 1000, n.minobsinnode = 2 y shrinkage = 0.01.



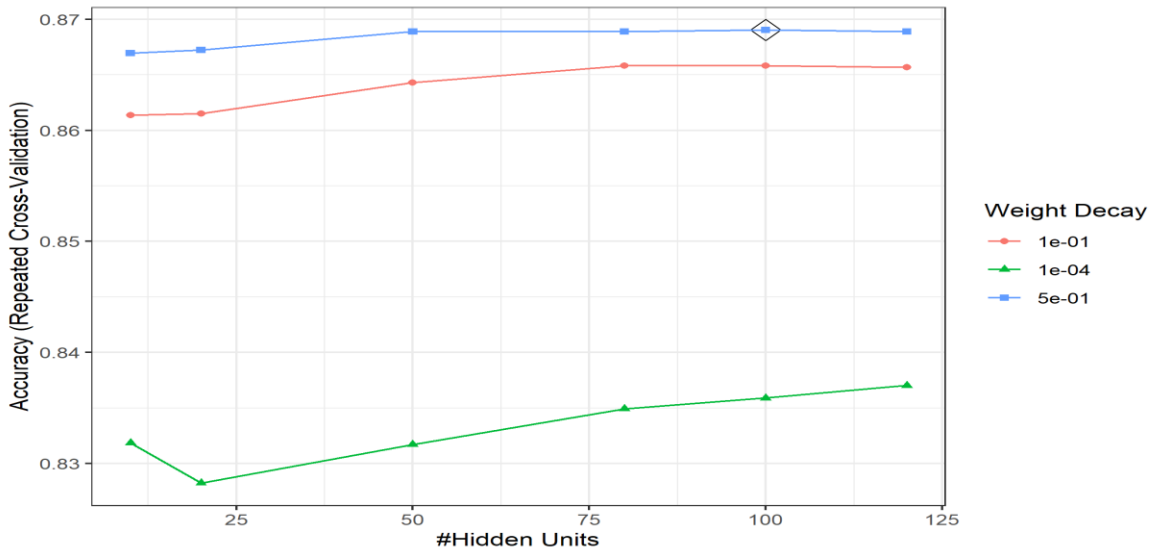
Fuente: elaboración del autor.

4.4.4 Redes neuronales

Para este modelo se tunean los hiper-parámetros “size” —el número de neuronas en la capa oculta— y “decay” —que controla la regularización durante el entrenamiento de la red—. Estos parámetros pueden tomar los siguientes valores: size = {10, 20, 50, 80, 100, 120}; y decay = c(0,0001, 0,1, 0,5). Su combinación produce un total de 54 posibles modelos.

La estimación arroja resultados que se muestran en la Figura 5. La combinación óptima de parámetros corresponde a size = 100 y decay = 0,5.

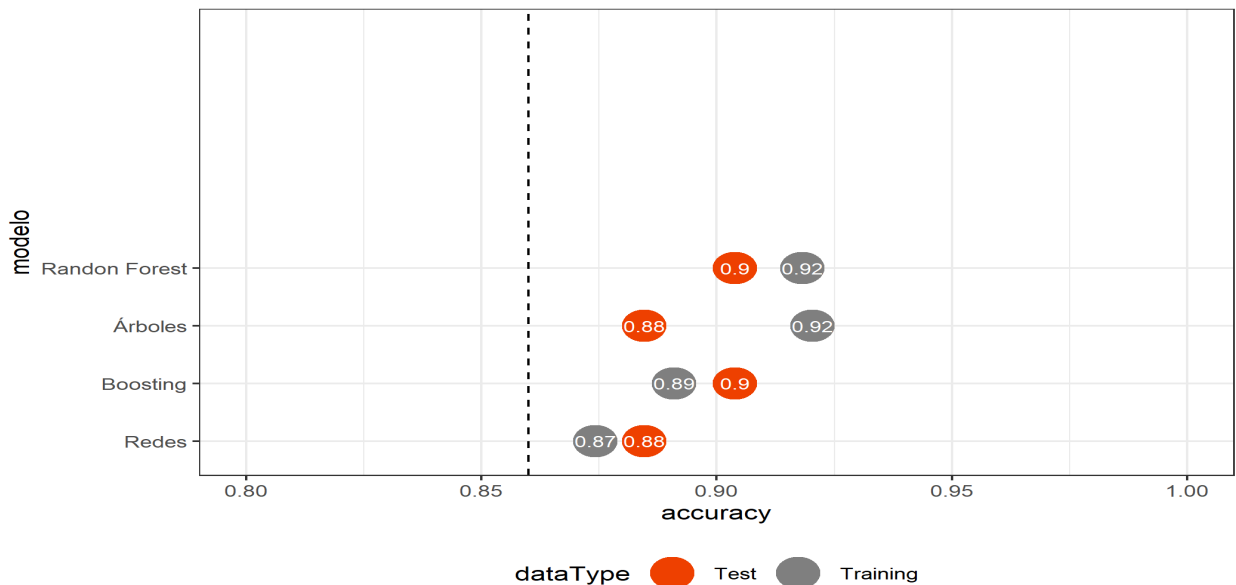
Figura 5. Hiper-parámetros óptimos del modelo de redes neuronales
Evolución del accuracy del modelo NNET



Fuente: elaboración del autor.

Una vez ajustados cada uno de los modelos, se compara su precisión tanto en la muestra de entrenamiento como en la muestra de testeo. Según los resultados, en la primera, los mejores modelos son árboles de decisión y random forest, con una precisión de 92 %. Sin embargo, la precisión del modelo de árboles de decisión en la muestra de testeo disminuye a 88 %, mientras que, en el caso de random forest, disminuye a 90 %. Por tanto, el modelo escogido para estimar la PI es este último [Figura 6].

Figura 6. Comparación de los modelos en la muestra de entrenamiento y testeo
Accuracy de entrenamiento y test

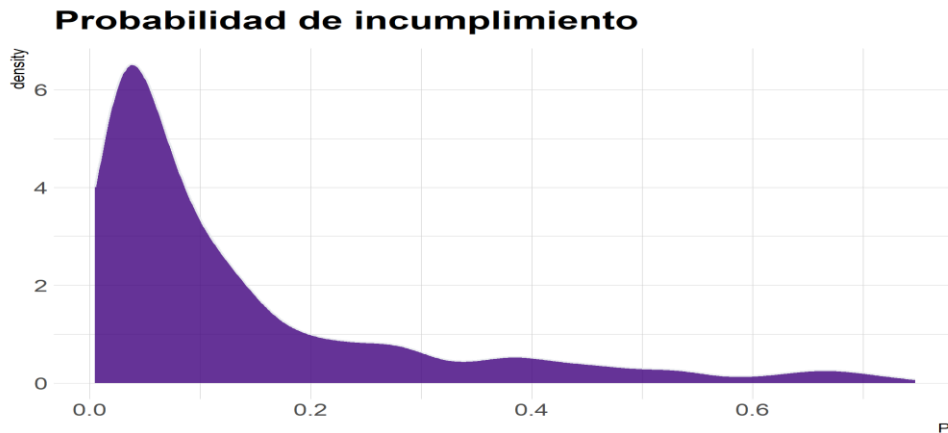


Fuente: elaboración del autor.

4.5 Estimación de las probabilidades de incumplimiento (PI) y cálculo del scoring

Una vez identificado el modelo más confiable, se procedió a realizar la determinación de las PI y el scoring. La Figura 7 muestra la distribución de las PI.

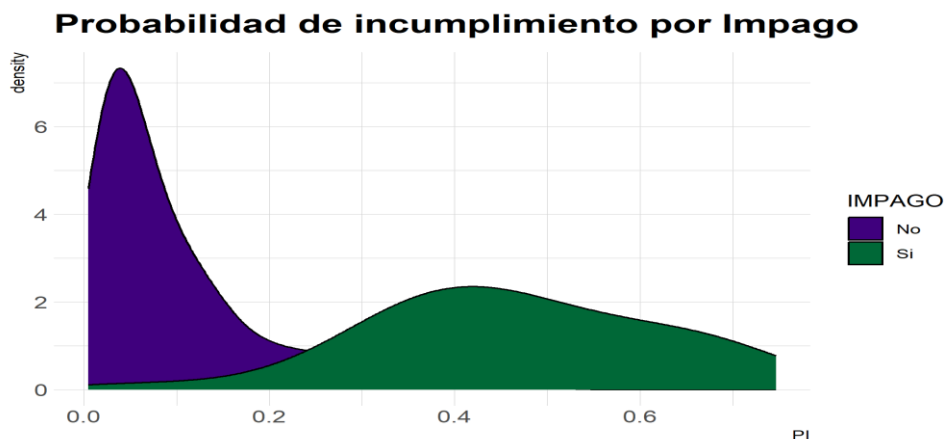
Figura 7. Distribución de las probabilidades de incumplimiento (PI)



Fuente: elaboración del autor.

En la Figura 7 se observa la distribución de las probabilidades de incumplimiento por empresa estimadas con el modelo random forest. La mayor densidad de las empresas tiene probabilidades de incumplimiento cercanas a cero (0). Esto indica que la mayoría de ellas en la muestra tienen una baja probabilidad de incumplir con sus préstamos. La curva desciende de forma rápida a medida que la PI aumenta, lo que sugiere que pocas empresas tienen una PI superior al 20 %. Aunque pocas, en todo caso no se pueden perder de vista a las empresas en la cola derecha de la distribución, ya que tienen mayores probabilidades de incumplimiento y representan un mayor riesgo para el portafolio de créditos [Figura 8].

Figura 8. Distribución de la probabilidad de incumplimiento (PI) por impago



Fuente: elaboración del autor.

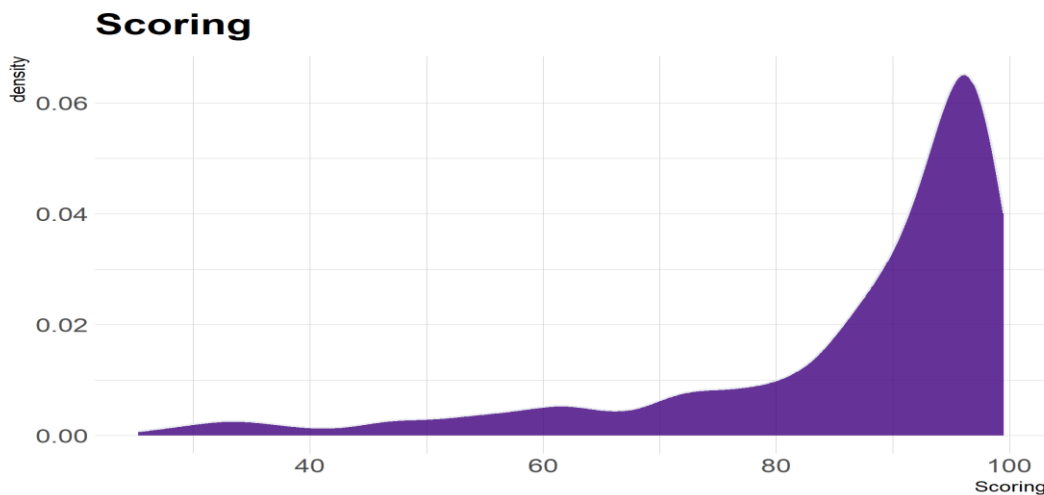
En la Figura 8 se observan las empresas que ha impagado y las que no lo han hecho. La curva sombreada de color morado corresponde a la densidad de probabilidades de incumplimiento para las segundas, y la curva sombreada de color verde corresponde a la densidad de probabilidades de incumplimiento para las primeras. Para los casos en los que no hubo impago (morado), la mayoría de las observaciones se concentran en valores de PI cercanos a 0. Para los casos con impago (verde), la mayor densidad de observaciones se encuentra en valores más altos de PI, entre 0,4 y 0,6.

Como es de esperarse, el modelo le asigna probabilidades de incumplimiento más altas a las empresas que ha impagado sus préstamos.

Finalmente, con base en la PI se crea el índice Scoring, el cual, por simplicidad, es el complemento de la PI: la probabilidad de cumplir [Figura 9].

$$\text{Scoring} = (1 - \text{PI}) * 100$$

Figura 9. Distribución del scoring



Fuente: elaboración del autor.

En la Figura 9 se observa la distribución de la variable Scoring, que representa el complemento de la PI escalada al rango de 0 a 100. En el eje X se ven los valores del scoring, que son calculados como $(1 - \text{PI}) * 100$, con valores que van aproximadamente desde 0 hasta 100. En el eje Y, la densidad de la distribución refleja la concentración de individuos o empresas en diferentes rangos de scoring.

La curva de densidad muestra que la mayor parte de los datos se agrupa en valores altos de scoring, especialmente entre 90 y 100, lo que indica que la mayoría de los casos tiene una alta probabilidad de cumplimiento y sugiere que es más probable que no incumplan con sus compromisos. La curva crece de manera pronunciada hacia estos valores altos, alcanzando un pico en el rango más cercano a 100, es decir, que la distribución muestra que una gran parte de los observados tiene una baja PI, lo que se traduce en altos valores de scoring.

Los resultados de esta investigación consolidan y amplían las aportaciones de estudios previos al validar la superioridad del algoritmo random forest en comparación con otras técnicas de machine learning, y al subrayar la relevancia de la selección de variables y la adecuada validación en la construcción de modelos predictivos para el scoring de crédito. Liaw y Wiener (2002), en su investigación pionera sobre random forests, demostraron cómo este algoritmo puede manejar grandes volúmenes de datos y seleccionar variables claves a través del índice de Gini.

De igual modo, la presente investigación corrobora sus hallazgos al confirmar que variables como Endeudamiento, ROA y RAZÓN CORRIENTE son determinantes para optimizar la precisión del modelo. Chen y Guestrin (2016), al explorar el algoritmo gradient boosting, destacaron su capacidad para afinar la precisión del modelo mediante la corrección iterativa de errores, y aunque encontraron que random forest ofreció un mejor equilibrio entre precisión y capacidad de generalización, era crucial reconocer que gradient boosting podía ser altamente efectivo en escenarios con relaciones no lineales complejas entre variables.

Finalmente, Kuhn y Johnson (2013), en su estudio sobre técnicas de machine learning, subrayaron la importancia de la validación cruzada y la selección adecuada de los modelos para mitigar el sobreajuste. En consonancia con estas recomendaciones, su investigación aplicó un riguroso enfoque de ajuste de hiper-parámetros y validación que reveló que random forest no solo se destacó en la muestra de entrenamiento, sino también en la de testeo, lo que enfatiza la necesidad de prácticas de validación y ajuste precisas para asegurar la robustez y generalización de los modelos predictivos.

4.6 Regla de decisión

La Tabla 1 muestra la regla de decisión por número de empresas.

Tabla 1. Regla de decisión por número de empresas

REGLA DE DECISIÓN	# EMPRESAS
Scoring < 65 %, en seguimiento	63
Scoring 65-80 %, a comité	56
Scoring > 80 %, aprobado	410

Fuente: elaboración del autor.

Esta regla de decisión asigna una categoría de crédito basada en la variable Scoring, que representa la puntuación crediticia de los solicitantes.

- Scoring < 65 %. El solicitante está en seguimiento debido a un bajo puntaje que sugiere un mayor riesgo de crédito.
- Scoring entre 65 y 80 %. El caso debe ser revisado por un comité para una evaluación más detallada antes de tomar una decisión final.
- Scoring > 80 %. El crédito está aprobado debido a un puntaje alto, lo que sugiere un bajo riesgo.

5. Conclusiones y Recomendaciones

5.1 Conclusiones

La evolución de los modelos de scoring de crédito ha transformado la forma en que las instituciones financieras evalúan el riesgo crediticio, pasando de enfoques estadísticos tradicionales a sofisticadas metodologías basadas en machine learning. Algoritmos como random forest, gradient boosting y redes neuronales han mejorado significativamente la precisión, al procesar grandes volúmenes de datos y detectar patrones complejos. Además, la integración de fuentes no tradicionales como datos de las redes sociales y las transacciones en línea ha enriquecido las evaluaciones crediticias. Estos avances permiten una mayor adaptabilidad a contextos financieros cambiantes, mejorando la gestión del riesgo y facilitando una asignación más eficiente de los recursos financieros.

La presente investigación desarrolló un modelo de scoring empresarial a partir de un enfoque comparativo exhaustivo entre cuatro modelos avanzados de machine learning para predecir la probabilidad de impago de empresas, con el objetivo de desarrollar un índice de scoring efectivo para la toma de decisiones crediticias.

Los árboles de decisión se destacan por su simplicidad y facilidad de interpretación, pero su susceptibilidad al sobreajuste resalta la importancia de métodos más robustos como el random forest, que mejora la estabilidad de las predicciones mediante la agregación de múltiples árboles.

El gradient boosting se distingue por su capacidad de corregir errores de manera iterativa, siendo especialmente útil en escenarios con relaciones no lineales complejas entre variables.

Por último, las redes neuronales, aunque menos interpretables, ofrecen una notable flexibilidad para capturar patrones complejos en los datos financieros.

El análisis comparativo de estos modelos proporcionó una comprensión integral de su rendimiento y precisión predictiva en el manejo de datos complejos y heterogéneos. Los resultados de esta comparación permitieron identificar el modelo más adecuado para optimizar la gestión del riesgo crediticio, lo que contribuirá a una mejor toma de decisiones en el ámbito financiero.

Los resultados obtenidos a partir del modelo random forest muestran que la mayor parte de las empresas analizadas presenta una baja PI, con una distribución concentrada en torno a valores cercanos a cero (0). Esto sugiere que la mayoría de las empresas en la muestra son consideradas solventes y tienen una baja propensión a incumplir sus compromisos financieros. Sin embargo, se observa una menor densidad de empresas en la cola derecha de la distribución, las cuales presentan probabilidades de incumplimiento superiores al 20 %. A pesar de su baja frecuencia, estas empresas representan un riesgo significativo para el portafolio crediticio y merecen una atención especial.

El modelo demostró ser eficaz al asignar probabilidades de incumplimiento más elevadas a las empresas que han incumplido en comparación con aquellas que no lo han hecho, lo que refuerza su capacidad predictiva. Finalmente, la creación del índice de scoring, calculado como el complemento de la PI, proporciona una medida directa y fácilmente interpretable de la capacidad de pago de cada empresa. La distribución del scoring refleja que la mayoría de ellas tienen altos valores en este índice, lo que confirma su baja PI. En resumen, el modelo random forest ofrece una estimación precisa y confiable del riesgo de crédito, permitiendo una diferenciación clara entre empresas con distintos niveles de riesgo.

Las oportunidades futuras de investigación en scoring de crédito son amplias y prometedoras, impulsadas por el rápido avance de las tecnologías de datos y la inteligencia artificial. En primer lugar, la integración de fuentes de datos no tradicionales como el comportamiento en las redes sociales, las transacciones en línea y los datos de dispositivos móviles presenta una oportunidad para mejorar la precisión de los modelos de scoring, proporcionando una visión más completa del riesgo crediticio. Además, el desarrollo de modelos de aprendizaje profundo y técnicas avanzadas como el procesamiento del lenguaje natural (natural language processing, NLP) y el análisis de sentimientos podría permitir capturar patrones aún más complejos y no lineales en los datos financieros.

Otra área con gran potencial es la investigación en métodos de fairness y ética en los modelos de scoring, enfocada en reducir sesgos algorítmicos y garantizar una evaluación justa de los solicitantes de crédito. Asimismo, la adaptabilidad de los modelos a contextos locales en mercados emergentes representa una línea de estudio importante, dado que la transferencia de modelos globales a estas regiones puede no ser eficiente sin personalización y ajustes contextuales.

Finalmente, el análisis de los efectos de la volatilidad económica y los ciclos financieros en la efectividad de los modelos de scoring de crédito abre una vía para explorar la resiliencia de estas herramientas ante cambios macroeconómicos, lo que sería clave para mejorar la gestión de riesgos en escenarios de incertidumbre económica.

5.2 Recomendaciones

Es necesario que la empresa objeto de estudio pueda validar la relevancia del proceso investigativo desarrollado; por ello, se presentarán los resultados del modelo a su dirección financiera, con el propósito de que cuente con una herramienta para el análisis de riesgos y la toma de decisiones adecuadas. En ella se recomienda que corra el modelo presentado, para evaluar su impacto en los resultados obtenidos para la determinación de un crédito.

Durante el mes de agosto de 2024, la Compañía realizó la refrendación de cupos de todos los clientes con crédito, revisando la información financiera reciente, por lo que se recomienda

determinar el score de todos ellos con el modelo propuesto, a fin de determinar a cuáles se considera que se les mantuvo crédito de manera errónea, para hacer seguimiento durante los próximos meses y confirmar la eficacia del modelo durante el próximo año.

En la actualidad, la Compañía concede cupo de crédito según el criterio de cada uno de los jefes de crédito, por lo que se recomienda incluir en el análisis el puntaje de score que arroje el modelo en las solicitudes nuevas, como una herramienta de apoyo.

Se recomienda a la Compañía establecer una política de crédito de tres categorías basadas en el scoring crediticio de los clientes. Aquellos con un puntaje superior a 80 puntos recibirán aprobación automática de crédito; aquellos con un puntaje entre 65 y 80 serán evaluados por un comité de crédito que determinará la viabilidad del préstamo tras una revisión detallada; finalmente, las solicitudes de clientes con un scoring inferior a 65 puntos serán rechazadas de forma automática debido al alto riesgo de incumplimiento.

Esta política busca equilibrar el crecimiento del negocio con la gestión responsable del riesgo crediticio.

Se recomienda implementar un seguimiento semestral al modelo de scoring de la Compañía. Este proceso permitirá evaluar su efectividad y precisión en la identificación de riesgos crediticios, así como realizar los ajustes necesarios para adaptarse a cambios en las condiciones del mercado y en los comportamientos de los clientes. El seguimiento consistirá en la revisión de los resultados históricos, la tasa de morosidad de los créditos aprobados y el análisis de posibles mejoras en los algoritmos utilizados, con el objetivo de garantizar que el modelo continúe siendo un instrumento confiable para la toma de decisiones crediticias.

Es importante aclarar que la implementación de este modelo se utilizó con una base de clientes que se encontraban sin reporte negativo en las centrales de riesgo, ya que esto representa un alto riesgo de crédito. Por ello se recomienda que la Compañía siga implementando la política de exclusión de estudios de crédito a aquellos clientes que tengan un reporte negativo, debido a que esto evidencia que es tipo de clientes presenta un mal hábito de pago confirmado.

Referencias

- Abdullayev, N. (2021). Credit scoring models in risk assessment and the specifics of their application in banks of Uzbekistan. *Siberian Financial School*, 1, 67-69. <https://doi.org/10.34020/1993-4386-2021-1-67-69>
- Altman, E. I. (1968). Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. *Journal of Finance*, 23(4), 589-609. <https://doi.org/10.1111/j.1540-6261.1968.tb00843.x>
- Altman, E. I. (2000). *Predicting financial distress of companies: Revisiting the Z-Score and ZETA models*. Ciudad de Nueva York, NY: Stern School of Business, New York University. DOI:10.4337/9780857936097.00027
- Bai, M., Zheng, Y., & Shen, Y. (2022). Gradient boosting survival tree with applications in credit scoring. *Journal of the Operational Research Society*, 73(1), 39-55. <https://doi.org/10.1080/01605682.2021.1919035>
- Becerra Molina, E., Ojeda Orellana, R., Lituma Yascaribay, M., & Carrasco Ruano, T. (2023). Análisis de estados financieros como herramienta útil para la gestión económica tras la pandemia Covid-19. *Revista Universidad y Sociedad*, 15(3), 263-272. http://scielo.sld.cu/scielo.php?script=sci_arttext&pid=S2218-36202023000300263&lng=es&tlng=es
- Behr, A., & Weinblat, J. (2017). Default patterns in seven EU countries: a random forest approach. *International Journal of the Economics of Business*, 24(2), 181-222. DOI:10.1080/13571516.2016.1252532
- Behr, P., & Güttler, A. (2007). Credit risk assessment with discriminant analysis. *Journal of Credit Risk*, 3(3), 43-63. https://www.researchgate.net/publication/228229883_Credit_Risk_Assessment_and_Relationship_Lending_An_Empirical_Analysis_of_German_Small_and_Medium-Sized_Enterprises
- Breiman, L., Friedman, J., Olshen, R., & Stone, C. (1986). *Classification and regression trees*. Boston, MA: Wadsworth International Group. <https://rafalab.dfci.harvard.edu/pages/649/section-11.pdf>
- Brigham, E. F., & Ehrhardt, M. C. (2016). *Financial management: Theory & Practice*. Boston, MA: Cengage Learning.
- Chang, Y. C., Chang, K. H., Chu, H. H., & Tong, L. I. (2016). Establishing decision tree-based short-term default credit risk assessment models. *Communications in Statistics-Theory and Methods*, 45(23), 6803-6815. DOI:10.1080/03610926.2014.968730
- Chen, T., & Guestrin, C. (2016). *XGBoost: a scalable tree boosting system* [ponencia]. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (pp. 785-794). <https://doi.org/10.1145/2939672.2939785>

- Colombia, Congreso de la República. (2009). *Ley 1328 de 2009* (“por la cual se dictan normas de protección al consumidor financiero”). Bogotá: Diario Oficial 47 411. http://www.secretariassenado.gov.co/senado/basedoc/ley_1328_2009.html
- Colombia, Congreso de la República. (2012). *Ley 1581 de 2012* (“por la cual se dictan disposiciones generales para la protección de datos personales”). Bogotá: Diario Oficial 48 587. http://www.secretariassenado.gov.co/senado/basedoc/ley_1581_2012.html
- Colombia, Congreso de la República. (2013). *Ley 1676 de 2013* (“por la cual se regula el proceso de cobro de obligaciones”). Bogotá: Diario Oficial 48 888. http://www.secretariassenado.gov.co/senado/basedoc/ley_1676_2013.html
- Colombia, Presidencia de la República. *Decreto 1377 de 2013* (“por el cual se reglamenta parcialmente la Ley 1581 de 2012”). <https://www.funcionpublica.gov.co/eva/gestornormativo/norma.php?i=53646>
- Colombia, Superintendencia Financiera de Colombia, SFC. (2002). *Circular externa 011 del 5 de marzo de 2002*. Bogotá: SFC. <https://www.fasecolda.com/cms/wp-content/uploads/2019/08/ce11-2002.pdf>
- Copeland, T., Koller, T., & Murrin, J. (2000). *Valuation: Measuring and managing the value of companies*. Hoboken, NJ: John Wiley & Sons.
- Cortés, C., & Vapnik, V. (1995). Support-Vector networks. *Machine Learning*, 20(3), 273-297. <https://www.scirp.org/reference/referencespapers?referenceid=1150668>
- Cox, D. R. (1958). The regression analysis of binary sequences. *Journal of the Royal Statistical Society, Series B*, 20(2), 215-242. <https://academic.oup.com/jrsssb/article-abstract/20/2/215/7027376>
- Credit Suisse. (2013). *Credit risk management: an overview*. Zurich: Credit Suisse Research.
- Dinçer, H., Hacıoglu, U., & Yuksel, S. (2020). Balanced scorecard-based performance assessment of Turkish banking sector with the Analytic Network Process (ANP). *International Journal of Decision Sciences & Applications*, 1, 1-21. DOI: 10.20525/ijdsa.v1i1.1415
- Feldman, R., & Nagle, J. (2002). Discriminant analysis and risk management. *Financial Analysts Journal*, 58(1), 67-78. <https://www.tandfonline.com/toc/ufaj20/58/1>
- Garzón Escobar, H. F. (2020). *Modelo scoring de comportamiento de cartera de crédito para una empresa de servicios públicos domiciliarios* [tesis de maestría, Universidad de Santander, Bucaramanga]. <https://repositorio.udes.edu.co/entities/publication/c21409a5-8d8d-4b01-b70b-60743fa4943d>
- Gibson, C. H. (2013). *Financial reporting and analysis*. Boston, MA: Cengage Learning.
- González López, P. P. (2018). *Desarrollo de modelo credit scoring para admisión de facturas en factoring* [tesis de maestría, Universidad de Chile, Santiago]. <https://repositorio.uchile.cl/handle/2250/149160>
- Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The elements of statistical learning: Data mining, inference, and prediction*. Springer.
- Horngren, C. T., Sundem, G. L., & Elliott, J. A. (2013). *Introduction to financial accounting*. Londres: Pearson.

- Hosmer, D. W., & Lemeshow, S. (2000). *Applied logistic regression*. Hoboken, NJ: John Wiley & Sons. <https://doi.org/10.48550/arXiv.1908.03385>
- Hu, Y., & Su, J. (2022). Research on credit risk evaluation of commercial banks based on artificial neural network model. *Procedia Computer Science*, 199, 1168-1176. <https://doi.org/10.1016/j.procs.2022.01.148>
- Huacchillo Pardo, L. A., Ramos Farroñán, E. V., & Pulache Lozada, J. L. (2020). La gestión financiera y su incidencia en la toma de decisiones financieras. *Revista Universidad y Sociedad*, 12(2), 356-362. http://scielo.sld.cu/scielo.php?script=sci_arttext&pid=S2218-36202020000200356&lng=es&tlng=es
- Jarrow, R. A., y Turnbull, S. M. (1995). Pricing derivatives on financial securities subject to credit risk. *The Journal of Finance*, 50(1), 53-85. <https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1540-6261.1995.tb05167.x>
- Khandani, A. E., Kim, A. J., & Lo, A. W. (2010). Consumer credit-risk models via machine-learning algorithms. *Journal of Banking & Finance*, 34(11), 2767-2787. <https://doi.org/10.1016/j.jbankfin.2010.06.001>
- Koller, T., Goedhart, M., & Wessels, D. (2010). *Valuation: Measuring and managing the value of companies*. Hoboken, NJ: John Wiley & Sons.
- Kuhn, M., & Johnson, K. (2013). *Applied predictive modeling*. Springer. <https://link.springer.com/book/10.1007/978-1-4614-6849-3>
- Leal Fica, A. L., Aranguiz Casanova, M. A., & Gallegos Mardones, J. (2018). Análisis de riesgo crediticio: Propuesta del modelo credit scoring. *Revista de la Facultad de Ciencias Económicas: Investigación y Reflexión*, 26(1), 181-207. <https://www.redalyc.org/pdf/909/90953767010.pdf>
- LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep Learning. *Nature*, 521, 436-444. <https://doi.org/10.1038/nature14539>
- Liaw, A., & Wiener, M. (2002). Classification and regression by random forest. *R News*, 2(3), 18-22. <https://CRAN.R-project.org/doc/Rnews/>
- Merton, R. C. (1974). On the pricing of corporate debt: The risk structure of interest rates. *Journal of Finance*, 29(2), 449-470. <https://www.jstor.org/stable/2978814>
- Ochoa, J. C., Galeano, W., & Agudelo, L. G. (2010). Construcción de un modelo de scoring para el otorgamiento de crédito en una entidad financiera. *Perfil de Coyuntura Económica*, 16, 191-222. <http://www.scielo.org.co/pdf/pece/n16/n16a10.pdf>
- Quinlan, J. R. (1986). Induction of decision trees. *Machine Learning*, 1(1), 81-106. <https://doi.org/10.1007/BF00116251>
- Rastegar, M., & Faraji, M. (2023). Development of a hybrid credit scoring model for the banking system. *Scientia Iranica*. <https://doi.org/10.24200/sci.2023.60399.6778>
- Ross, S. A., Westerfield, R. W., & Jaffe, J. (2016). *Corporate finance: Core principles and applications*. Ciudad de Nueva York, NY: McGraw-Hill.
- RStudio Team. (2020). *RStudio: Integrated development for R*. RStudio, PBC. <https://rstudio.com>
- Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Learning representations by back-propagating errors. *Nature*, 323, 533-536. <https://www.nature.com/articles/323533a0.pdf>

- Santiago Sandoval, C., & Urán Vélez, A. (2021). Modelo de credit scoring para la empresa Grupo Factoring de Occidente S. A. S. [tesis de maestría, Universidad EAFIT, Medellín]. <https://repository.eafit.edu.co/server/api/core/bitstreams/52071f2b-dfd6-4c56-8c09-13fe284c171d/content>
- Sariev, E., & Germano, G. (2020). Bayesian regularized artificial neural networks for the estimation of the probability of default. *Quantitative Finance*, 20(2), 311-328. <https://doi.org/10.1080/14697688.2019.1633014>
- Simonov, P., & Zboev, E. (2020). *Scoring credit rating of individuals*. Makhachkala, Dagestan, Rusia: Herald of Dagestan State University. <https://doi.org/10.21779/2542-0321-2020-35-2-18-26>
- Sullivan, M. (2016). *Managing credit risk in financial institutions*. Londres: Financial Times Publishing. <https://www.ft.com/reports/risk-management-financial-institutions>
- Thomas, L. C. (2009). *Consumer credit models: Development, evaluation, and validation*. Cambridge: Oxford University Press.
- Tulcanaza Aguilar, M. Á. (2021). *Propuesta de un modelo de score de originación para la cartera de consumo de una cooperativa de ahorro y crédito del segmento 3 en el Ecuador* [tesis de maestría, Universidad Andina Simón Bolívar, Quito]. <https://repositorio.uasb.edu.ec/bitstream/10644/8115/1/T3536-MFGR-Tulcanaza-Propuesta.pdf>
- White, G. I., Sondhi, A. C., & Fried, D. (2003). *The analysis and use of financial statements*. Hoboken, NJ: John Wiley & Sons.
- Wickham, H., François, R., Henry, L., & Müller, K. (2019). *dplyr: A grammar of data manipulation. R package version 1.0.2*. <https://CRAN.R-project.org/package=dplyr>
- Yessenaman, D., Alimzhanova, L., & Sarbasova, K. (2023). Modeling software for an effective banking scoring system. *Bulletin of Toraihyrov University. Physics & Mathematics series*. <https://doi.org/10.48081/egvg4566>

Apéndice

Política de crédito para Mexichem Colombia S. A. S.

Objetivo

Establecer criterios claros para la autorización, revisión y rechazo de solicitudes de crédito por parte de los clientes de Mexichem, basados en el sistema de scoring crediticio de la Compañía, a fin de garantizar una adecuada gestión de riesgos financieros.

Política

1) Autorización automática de crédito

Se otorgará crédito de manera automática a los clientes que obtengan un scoring mayor de 80 puntos. Estos clientes han demostrado un historial de pago satisfactorio y tienen un perfil financiero sólido.

Condiciones. No requerirán revisión adicional y podrán acceder a los productos y servicios en las condiciones crediticias predeterminadas.

2) Revisión por el comité de crédito

Las solicitudes de crédito de clientes que obtengan un scoring entre 65 y 80 puntos serán sometidas a consideración de un comité de crédito.

Acciones del comité. El comité evaluará factores adicionales como historial de pagos, antigüedad en la relación comercial y otras variables cualitativas antes de decidir si aprobar o denegar el crédito. Adicionalmente, tendrá la facultad de otorgar condiciones especiales si considera que los riesgos son aceptables.

3) Rechazo de solicitudes

Se rechazarán automáticamente las solicitudes de crédito de los clientes cuyo scoring sea menor de 65 puntos.

Motivos. Este puntaje refleja un alto riesgo de incumplimiento, por lo que la Compañía no considerará la aprobación de financiamiento a estos clientes para evitar exposición a pérdidas.

Revisión periódica. Esta política estará sujeta a revisiones periódicas basadas en las condiciones del mercado y la evolución del comportamiento crediticio de los clientes. Mexichem se reserva el derecho de modificar los criterios de scoring según lo considere necesario para mantener una cartera de crédito sana; inicialmente se sugiere hacerlo cada seis meses.

Finalmente, esta política busca equilibrar el crecimiento del negocio con una sólida gestión del riesgo crediticio, promoviendo la eficiencia en la toma de decisiones y asegurando que los recursos de la Compañía se asignen de manera responsable.