

UNIVERSIDAD TÉCNICA FEDERICO SANTA MARÍA

DEPARTAMENTO DE INGENIERÍA COMERCIAL

**COMPARACIÓN EMPÍRICA ENTRE EL MODELO Z-SCORE DE ALTMAN Y UN
MODELO DE RANDOM FOREST PARA LA PREDICCIÓN DE QUIEBRAS
EMPRESARIALES: DESEMPEÑO, INTERPRETABILIDAD Y CONVERGENCIA DE
VARIABLES FINANCIERAS**

Memoria para optar al título de Ingeniero Comercial

Autor: Gabriel Accatini González

Profesora Guía: Macarena Gatica Silva

Valparaíso – Chile

2026



CONSTANCIA DE VALIDACIÓN Y CONFIDENCIALIDAD DE MONOGRAFÍA A REPOSITORIO ACADÉMICO

1.- IDENTIFICACIÓN DEL TRABAJO ACADÉMICO

Tipo de monografía (marcar una opción): Memoria o trabajo de título Tesis de Postgrado

Título del trabajo: “Comparación empírica entre el modelo Z-Score de Altman y un modelo de Random Forest para la predicción de quiebras empresariales: desempeño, interpretabilidad y convergencia de variables financieras”

Nombre del candidato(a): Gabriel Ignacio Accatini González

Carrera / Grado: Ingeniería Comercial

Campus: Casa Central Valparaíso

Departamento: Ingeniería Comercial

2.- VALIDACIÓN DEL PROFESOR GUÍA/DIRECTOR DE TESIS

Yo, Macarena Gatica Silva, en mi calidad de profesor(a) guía/director(a) del trabajo académico mencionado anteriormente **DEJO CONSTANCIA** que:

- He revisado esta versión del documento y corresponde a la versión final aprobada del trabajo.
- El trabajo cumple con los requisitos académicos y de formato establecidos por la institución.

3.- EVALUACIÓN DE CONFIDENCIALIDAD POR PROPIEDAD INDUSTRIAL (marcar una opción)

El trabajo **NO contiene** información que amerite confidencialidad y puede ser publicado de inmediato en repositorio con acceso abierto.

El trabajo **CONTIENE** información con potenciales implicancias de propiedad industrial o intelectual y requiere un periodo de confidencialidad (**embargo**) por (**marcar una opción**):

6 meses 12 meses 2 años 3 años 5 años 10 años

Fundamentación de la necesidad de confidencialidad (obligatorio si se solicita embargo):

4.- FIRMAS

Profesor(a) guía o director(a) de memoria o tesis:

Fecha: 06/03/2026

Firma:

Estudiante o Candidato(a):

Fecha: 06/03/2026

Firma:

Este formulario debe ser insertado como página 2 de la memoria o tesis, completado y firmado por estudiante y profesor(a) antes de la entrega en portal PRISMA de Biblioteca USM.

Resumen

La predicción de quiebras empresariales constituye uno de los problemas centrales en las finanzas corporativas, debido a su relevancia para la gestión del riesgo financiero y la estabilidad de los sistemas económicos. Tradicionalmente, este fenómeno ha sido abordado mediante modelos estadísticos basados en ratios contables, entre los cuales destaca el modelo Z-score de Altman (1968) como principal referente metodológico. Sin embargo, el desarrollo reciente del aprendizaje automático ha abierto nuevas posibilidades para capturar relaciones no lineales y patrones complejos en datos financieros de alta dimensionalidad.

El presente estudio tiene como objetivo evaluar empíricamente el desempeño predictivo de un modelo de Random Forest en la detección de quiebras empresariales y compararlo con el modelo clásico Z'-score de Altman, utilizando el Taiwanese Bankruptcy Prediction Dataset. Adicionalmente, se propone un análisis de convergencia conceptual orientado a examinar si las variables más relevantes identificadas por el modelo de machine learning se corresponden con las dimensiones financieras tradicionales definidas por Altman, tales como liquidez, rentabilidad, solvencia y eficiencia.

Metodológicamente, se adopta un enfoque cuantitativo no experimental y transeccional. Se aplicaron técnicas de preprocesamiento, reducción de redundancia y tratamiento del desbalance de clases mediante métodos de remuestreo, integrados en un pipeline de modelamiento con validación cruzada y optimización de hiperparámetros. El desempeño de ambos modelos fue evaluado sobre un conjunto de prueba independiente utilizando métricas robustas para clasificación desbalanceada, incluyendo recall, F1-score y AUC-ROC.

Los resultados evidencian que el modelo Random Forest supera al Z-score de Altman en términos de capacidad predictiva global y detección de empresas en riesgo, mostrando un mayor poder discriminante frente a escenarios complejos. Asimismo, el análisis de importancia de variables revela una convergencia significativa entre los determinantes identificados por el modelo de aprendizaje automático y las dimensiones financieras clásicas, lo que sugiere que la superioridad empírica del enfoque moderno se sustenta en una representación más flexible y no lineal de constructos financieros tradicionales.

Estos hallazgos contribuyen a la literatura al integrar la precisión predictiva del machine learning con la interpretación económica, aportando evidencia empírica sobre el rol de la inteligencia artificial en la modernización de los modelos de predicción de quiebra empresarial.

Palabras clave: predicción de quiebras, Z-score de Altman, Random Forest, machine learning, riesgo financiero, interpretabilidad, insolvencia empresarial.

Abstract

Corporate bankruptcy prediction represents a central problem in corporate finance due to its relevance for financial risk management and economic stability. Traditionally, this phenomenon has been addressed through statistical models based on accounting ratios, among which Altman's Z-score (1968) stands as the main methodological benchmark. However, recent advances in machine learning have enabled the modeling of non-linear relationships and complex patterns in high-dimensional financial data.

The objective of this study is to empirically evaluate the predictive performance of a Random Forest model in detecting corporate bankruptcies and to compare it with the classical Altman Z'-score model, using the Taiwanese Bankruptcy Prediction Dataset. Additionally, the research proposes a conceptual convergence analysis aimed at examining whether the most relevant variables identified by the machine learning model correspond to the traditional financial dimensions defined by Altman, such as liquidity, profitability, solvency, and efficiency.

Methodologically, a quantitative, non-experimental, and cross-sectional design is adopted. Data preprocessing, redundancy reduction, and class imbalance treatment through resampling techniques were implemented within a modeling pipeline that integrates cross-validation and hyperparameter optimization. Both models were evaluated on an independent test set using robust metrics for imbalanced classification, including recall, F1-score, and AUC-ROC.

The results show that the Random Forest model outperforms the Altman Z-score in overall predictive performance and in the detection of financially distressed firms, exhibiting greater discriminative power in complex scenarios. Furthermore, the feature importance analysis reveals a significant convergence between the determinants identified by the machine learning model and classical financial dimensions, suggesting that the empirical superiority of the modern approach is grounded in a more flexible and non-linear representation of traditional financial constructs.

These findings contribute to the literature by integrating machine learning predictive power with financial interpretability, providing empirical evidence on the role of artificial intelligence in the modernization of corporate bankruptcy prediction models.

Keywords: bankruptcy prediction, Altman Z-score, Random Forest, machine learning, financial risk, interpretability, corporate insolvency.

ÍNDICE

ORIGEN Y PROPÓSITO DEL ESTUDIO	10
OBJETIVOS.....	12
OBJETIVO GENERAL.....	12
OBJETIVOS ESPECÍFICOS.....	12
ALCANCES DE LA INVESTIGACIÓN	13
LIMITACIONES DE LA INVESTIGACIÓN.....	14
ESTADO DEL ARTE: PREDICCIÓN DE QUIEBRAS EMPRESARIALES.....	16
DISTINCIÓN CONCEPTUAL ENTRE FINANCIERÍA DISTRESS E INSOLVENCIA/QUIEBRA	16
EL PARADIGMA CLÁSICO	16
<i>Antecedentes del paradigma clásico de predicción de quiebras</i>	16
<i>Fundamento y Evolución del Z-Score de Altman</i>	17
<i>Métodos, Datos y Limitaciones Críticas</i>	19
EL PARADIGMA MODERNO: APRENDIZAJE AUTOMÁTICO (MACHINE LEARNING).....	20
<i>Metodologías y Rendimiento Superior</i>	20
<i>Desafíos Metodológicos del ML</i>	21
BRECHAS DE INVESTIGACIÓN Y EL GAP DE CONVERGENCIA CONCEPTUAL	22
<i>Vacíos Fundamentales de la Literatura</i>	22
<i>El Gap de Investigación Conceptual (Brecha de Convergencia)</i>	23
POSICIONAMIENTO Y CONTRIBUCIÓN DEL PROYECTO	24
METODOLOGÍA	26
PREPARACIÓN Y PREPROCESAMIENTO DE LOS DATOS.....	27
<i>Selección y Adquisición del Dataset</i>	27
<i>Exploración, Limpieza y Transformación de Datos (EDA)</i>	27
<i>Mitigación del Desequilibrio de Clases (Resampling)</i>	28
CÁLCULO DEL MODELO Z-SCORE DE ALTMAN	28
ENTRENAMIENTO DEL MODELO RANDOM FOREST	29
COMPARACIÓN DEL DESEMPEÑO PREDICTIVO	29
ANÁLISIS DE IMPORTANCIA DE VARIABLES Y CONVERGENCIA CONCEPTUAL	30
INTEGRACIÓN, DISCUSIÓN Y REPRODUCIBILIDAD.....	31
APLICACIÓN METODOLÓGICA.....	32
DISEÑO COMPUTACIONAL Y FLUJO GENERAL DEL ESTUDIO	32
DESCRIPCIÓN DEL DATASET Y VALIDACIONES INICIALES.....	34
ANÁLISIS EXPLORATORIO DE DATOS (EDA)	36
PREPROCESAMIENTO Y PREPARACIÓN DE LOS DATOS	39
<i>Depuración inicial del dataset</i>	39
<i>Tratamiento de valores extremos (winsorización)</i>	39
<i>Reducción de redundancia mediante análisis de correlación</i>	40
<i>Consideraciones de escalamiento y definición del set final</i>	41
CONSTRUCCIÓN DEL Z-SCORE DE ALTMAN	41
<i>Selección y mapeo de variables financieras</i>	42

<i>Cálculo del Z''-score</i>	42
<i>Análisis de la distribución del Z''-score</i>	43
<i>Z''-score según estado de quiebra</i>	43
ALTMAN COMO MODELO DE CLASIFICACIÓN	44
<i>Clasificación por zonas clásicas y limitaciones</i>	45
<i>Orientación del score y análisis ROC</i>	45
<i>Determinación del umbral óptimo</i>	46
<i>Evaluación del desempeño clasificatorio</i>	47
<i>Síntesis metodológica</i>	48
PREPARACIÓN DEL SET DE ENTRENAMIENTO Y TEST	48
<i>Congelamiento de conjuntos y control de integridad</i>	49
<i>Clasificador base trivial</i>	49
<i>Implicancias metodológicas</i>	51
MANEJO DEL DESBALANCE DE CLASES	51
<i>Distribución original del conjunto de entrenamiento</i>	51
<i>Aplicación de SMOTE</i>	52
<i>Aplicación de SMOTETomek</i>	52
<i>Consideraciones metodológicas</i>	53
ENTRENAMIENTO DEL RANDOM FOREST BASE	53
<i>Random Forest base sin técnicas de balanceo</i>	54
<i>Random Forest base entrenado con SMOTE</i>	54
<i>Random Forest base entrenado con SMOTETomek</i>	55
<i>Comparación preliminar de configuraciones</i>	55
OPTIMIZACIÓN DEL MODELO RANDOM FOREST	56
<i>Definición del espacio de búsqueda</i>	56
<i>Estrategia de optimización</i>	57
<i>Selección del mejor modelo</i>	57
<i>Consideraciones metodológicas</i>	58
EVALUACIÓN FINAL DEL MODELO RANDOM FOREST	58
<i>Distribución de predicciones</i>	58
<i>Matriz de confusión</i>	59
<i>Métricas de desempeño</i>	59
<i>Curva ROC y capacidad discriminante</i>	60
<i>Análisis de resultados</i>	61
COMPARACIÓN EMPÍRICA ENTRE EL MODELO Z''-SCORE DE ALTMAN Y EL MODELO RANDOM FOREST	61
<i>Comparación de métricas de desempeño</i>	62
<i>Análisis de matrices de confusión</i>	63
<i>Comparación de curvas ROC</i>	63
<i>Discusión comparativa</i>	64
INTERPRETABILIDAD DEL RANDOM FOREST	65
<i>Importancia de variables en el Random Forest</i>	65
<i>Importancia acumulada y concentración explicativa</i>	66
<i>Importancia por permutación</i>	67
<i>Interpretabilidad global mediante valores SHAP</i>	68
<i>Interpretabilidad local: explicación de una predicción individual</i>	69
<i>Síntesis interpretativa</i>	69
ANÁLISIS FINANCIERO AGREGADO Y VINCULACIÓN TEÓRICA CON EL MODELO DE ALTMAN	70
<i>Reducción del espacio de variables según importancia predictiva</i>	70
<i>Agrupación conceptual en dimensiones financieras</i>	70

<i>Importancia agregada por dimensión financiera</i>	72
<i>Vinculación empírica con el modelo Z'' de Altman</i>	73
<i>Interpretación financiera de los resultados</i>	74
RESULTADO FINAL DE LA INVESTIGACIÓN	76
CONCLUSIONES Y RECOMENDACIONES	77
CONCLUSIONES.....	77
<i>Conclusiones metodológicas</i>	77
<i>Conclusiones empíricas</i>	77
<i>Conclusiones financieras y teóricas</i>	78
RECOMENDACIONES	78
<i>Recomendaciones para la práctica profesional</i>	78
<i>Recomendaciones para futuras investigaciones</i>	79
REFLEXIÓN FINAL.....	79
BIBLIOGRAFÍA	80
ANEXOS	82
ANEXO A - REPOSITORIO DE CÓDIGO Y REPRODUCIBILIDAD COMPUTACIONAL.....	82

Origen y propósito del estudio

La predicción de quiebras empresariales ha ocupado un lugar central en las finanzas corporativas desde mediados del siglo XX, debido a su relevancia para la estabilidad económica y la gestión del riesgo financiero. Anticipar oportunamente la probabilidad de insolvencia de una empresa permite a los inversionistas, acreedores, reguladores y directivos adoptar decisiones más informadas, reduciendo las pérdidas potenciales asociadas al incumplimiento y fortaleciendo la eficiencia del sistema financiero en su conjunto.

Dentro de este campo de estudio, el modelo Z-score de Altman (1968) se consolidó como el principal referente metodológico. Basado en el análisis discriminante multivariable, el modelo propone una combinación ponderada de razones financieras que permite clasificar a las empresas según su nivel de riesgo de quiebra. Su simplicidad, bajo requerimiento de datos y capacidad de interpretación facilitaron su adopción en contextos empresariales y académicos, convirtiéndolo en el modelo más citado y replicado en la literatura especializada. No obstante, su estructura lineal y dependiente de supuestos estadísticos restrictivos limita su capacidad de adaptación frente a las condiciones dinámicas y no lineales que caracterizan los mercados financieros actuales.

En las últimas dos décadas, el desarrollo de la ciencia de datos y el aprendizaje automático ha abierto nuevas posibilidades para abordar la predicción de insolvencias desde un enfoque más flexible y empírico. En particular, los modelos de machine learning (ML) permiten procesar grandes volúmenes de información, identificar patrones complejos y capturar interacciones entre variables que los modelos estadísticos tradicionales no logran representar adecuadamente. Entre ellos, el algoritmo Random Forest (RF) ha mostrado un desempeño sobresaliente por su capacidad de manejar datos de alta dimensionalidad, su resistencia al sobreajuste y su eficacia en conjuntos de datos desbalanceados, características habituales en la predicción de quiebras.

En este contexto, el presente estudio tiene como propósito entrenar un modelo de Random Forest utilizando el dataset Taiwanese Bankruptcy Prediction —una de las bases de datos más completas para entrenar modelos de ML en predicción de quiebras— y comparar sus resultados con las predicciones del modelo clásico Z-score de Altman. A través de esta

comparación, se busca determinar si los enfoques basados en aprendizaje automático logran mejorar la capacidad predictiva de los modelos tradicionales, así como evaluar su aplicabilidad y robustez en el contexto de la evaluación del riesgo financiero.

La contribución principal de esta investigación radica en contrastar dos paradigmas analíticos: por un lado, el enfoque estadístico clásico representado por Altman, y por otro, la metodología contemporánea basada en algoritmos ML. Este contraste no sólo permitirá evaluar diferencias en términos de precisión y sensibilidad, sino también reflexionar sobre el papel que desempeña la inteligencia artificial en la modernización de las herramientas financieras. De esta manera, el estudio pretende aportar evidencia empírica que enriquezca la discusión académica en torno a la evolución de los modelos de predicción de quiebras, destacando las oportunidades y desafíos que plantea la integración de técnicas de ciencia de datos en la gestión del riesgo empresarial.

Objetivos

Objetivo General

Evaluar el desempeño predictivo de un modelo de Random Forest en la detección de quiebras empresariales y compararlo con el modelo Z-score de Altman, analizando además la correspondencia entre las variables más relevantes en ambos enfoques.

Objetivos específicos

1. Calcular el Z-score de Altman para cada empresa del dataset *Taiwanese Bankruptcy Prediction* y clasificar su nivel de riesgo de quiebra.
2. Entrenar un modelo de Random Forest para predecir la variable *bankrupt = 1/0*.
3. Comparar la capacidad predictiva de ambos modelos mediante métricas de desempeño (Accuracy, Precision, Recall, F1-score y AUC).
4. Analizar la importancia de las variables (feature importance) en el modelo Random Forest.
5. Evaluar la correspondencia conceptual y empírica entre las variables relevantes del RF y los componentes financieros del Z-score (liquidez, rentabilidad, solvencia, eficiencia y actividad).

Alcances de la Investigación

El alcance de este estudio se centra en la evaluación empírica y la reflexión conceptual sobre la modernización de las herramientas de gestión del riesgo financiero. La investigación busca generar una comprensión profunda en los siguientes aspectos:

1. **Desempeño Predictivo y Contraste de Paradigmas:** Se establecerá una evaluación robusta y cuantitativa de la capacidad predictiva de dos metodologías fundamentales: el modelo estadístico clásico Z-Score de Altman y el enfoque contemporáneo de Machine Learning (ML) representado por el algoritmo Random Forest (RF). Comprenderemos si el enfoque basado en el aprendizaje automático logra mejorar significativamente la capacidad predictiva de los modelos tradicionales en un contexto de mercado específico (Taiwan). Esta evaluación se realizará mediante métricas estandarizadas como *Accuracy*, *Precision*, *Recall*, *F1-score* y el Área Bajo la Curva (AUC).
2. **Comprensión Estructural de la Superioridad de ML:** Se trascenderá la mera comparación de precisión para analizar la importancia de las variables (*feature importance*) dentro del modelo Random Forest. Esto nos permitirá reflexionar sobre el papel de la inteligencia artificial en la modernización de las herramientas financieras. El estudio abordará una de las brechas de investigación más críticas, que es la falta de análisis comparativo estructural entre los enfoques.
3. **Convergencia Conceptual y Validación Teórica (La Contribución Central):** El alcance más novedoso y principal de la investigación es el Análisis de Convergencia Conceptual. Este análisis permitirá delimitar si los determinantes de quiebra identificados por la metodología *data-driven* (RF) confirman, refinan o contradicen los constructos teóricos clásicos definidos por el modelo de Altman.
 - Si existe convergencia: Se validaría la teoría financiera clásica (basada en liquidez, solvencia, rentabilidad, apalancamiento y eficiencia operativa) desde una perspectiva moderna, sugiriendo que Altman identificó los determinantes esenciales a pesar de usar una metodología estadística no óptima.

- Si existe divergencia: Se obtendrá evidencia empírica que sugiere que las dinámicas de insolvencia han cambiado cualitativamente desde 1968, lo que podría invalidar supuestos teóricos clásicos.

Este análisis es crucial para mejorar la interpretabilidad del modelo RF, abordando parcialmente el problema de la "caja negra" al vincular los patrones complejos que detecta el algoritmo con la teoría financiera tradicional.

Limitaciones de la Investigación

Las limitaciones establecen los límites de la generalización y el alcance metodológico del estudio, definiendo dónde termina nuestra comprensión:

1. Limitación Contextual del Dataset: Los hallazgos predictivos estarán intrínsecamente ligados a las características del dataset de predicción de quiebras taiwanés.
 - Geografía y Sector: Los resultados de precisión obtenidos, tanto para el Z-Score como para el RF, se aplicarán al contexto empresarial de Taiwán y no son directamente generalizables a empresas en otros países o contextos (mercados emergentes, sectores no manufactureros) sin una revalidación.
 - Modelo de Altman: Aunque el modelo Z-score es robusto, su aplicación directa a empresas en un contexto diferente al estadounidense de los años 60 es metodológicamente incorrecta y produce resultados poco fiables.
2. Limitación en la Integración de Datos: El estudio se enfocará principalmente en datos basados en contabilidad (ratios financieros). El estudio no abordará la integración racional de fuentes de datos no financieros o macroeconómicos. Por lo tanto, no se comprenderá la contribución de factores externos o cualitativos (como la calidad de la gestión o el entorno regulatorio) a la superioridad predictiva de ML.
3. Enfoque en la Convergencia y no en la Causalidad: La investigación se centra en evaluar la correspondencia empírica y conceptual entre los determinantes de los dos modelos. Si bien esto aborda una brecha crítica, el estudio no se enfoca en determinar las causas estructurales subyacentes y exhaustivas de la superioridad de ML, ni investiga

profundamente la evolución de las relaciones no lineales que el RF capta con respecto a los supuestos lineales del Z-Score.

Estado del Arte: Predicción de Quiebras Empresariales

La predicción de quiebras es uno de los desafíos más críticos en la gestión financiera moderna, trascendiendo el ámbito académico para impactar en las decisiones de acreedores, inversionistas, reguladores financieros y directivos corporativos. Este problema representa una oportunidad para la Ingeniería Comercial de aplicar rigor científico en la construcción de sistemas de alerta temprana con impacto en la gestión de riesgo sistémico.

Distinción conceptual entre *financial distress* e insolvencia/quiebra

En la literatura financiera, resulta fundamental distinguir entre los conceptos de *financial distress* e *insolvencia o quiebra empresarial*. El *financial distress* se refiere a un estado progresivo de deterioro financiero, caracterizado por dificultades persistentes para cumplir obligaciones operativas y financieras, reducción de rentabilidad, tensiones de liquidez y pérdida de solvencia económica. Este proceso no implica necesariamente la quiebra legal, sino que representa una fase previa en la que la empresa aún puede continuar operando, reestructurarse o recuperarse.

En contraste, la quiebra o insolvencia corresponde a un evento formal, generalmente definido desde el ámbito legal o institucional, en el cual la empresa declara su incapacidad para cumplir con sus obligaciones financieras, activando procedimientos de liquidación o reorganización.

Los modelos clásicos y modernos de predicción de quiebras no buscan únicamente anticipar el evento legal de quiebra, sino aproximar empíricamente la probabilidad de que una empresa se encuentre en un estado avanzado de *financial distress* que pueda derivar en insolvencia. En este sentido, las variables contables y financieras utilizadas capturan trayectorias de deterioro económico subyacentes, más que el acto jurídico puntual de quiebra.

El Paradigma Clásico

Antecedentes del paradigma clásico de predicción de quiebras

Los primeros esfuerzos sistemáticos por anticipar la quiebra empresarial se remontan a los trabajos pioneros de Beaver (1966), quien demostró empíricamente que ciertos ratios financieros individuales, particularmente aquellos asociados a liquidez y apalancamiento,

presentaban capacidad discriminante entre empresas sanas y empresas en quiebra varios años antes del evento. Su enfoque univariante constituyó la base empírica inicial del campo.

Posteriormente, Altman (1968) amplió este enfoque mediante la incorporación de técnicas multivariantes, dando origen al modelo Z-score, que integró simultáneamente múltiples dimensiones financieras en una única función discriminante. Este avance marcó un punto de inflexión al permitir una evaluación más estructural del riesgo de insolvencia.

En décadas posteriores, surgieron modelos estadísticos alternativos que buscaron superar algunas limitaciones del análisis discriminante. Ohlson (1980) introdujo el uso de modelos logit para la predicción de quiebras, eliminando la dependencia de supuestos estrictos de normalidad y homocedasticidad, mientras que Zmijewski (1984) propuso un enfoque probit orientado a corregir sesgos de selección muestral. Estos modelos consolidaron el paradigma estadístico clásico, estableciendo que la quiebra empresarial puede ser abordada como un problema de clasificación probabilística basado en información contable.

En este contexto, el modelo de Altman se mantiene como el referente más influyente, no solo por su desempeño empírico, sino por haber estructurado conceptualmente las principales dimensiones financieras asociadas al deterioro empresarial.

Fundamento y Evolución del Z-Score de Altman

El modelo Z-Score de Altman (1968) se consolidó como uno de los principales referentes metodológicos en la predicción de quiebras empresariales y es ampliamente reconocido como uno de los primeros intentos sistemáticos de anticipar la insolvencia corporativa mediante herramientas cuantitativas. Este modelo introdujo un enfoque multivariante basado en el Análisis Discriminante Múltiple (Multiple Discriminant Analysis, MDA), técnica estadística que permite combinar simultáneamente diversos ratios financieros en una única función discriminante, maximizando la separación entre empresas sanas y empresas en quiebra.

La propuesta de Altman representó un quiebre respecto de los estudios previos, que se basaban principalmente en el análisis univariante de indicadores financieros aislados (Beaver, 1966). En contraste, el Z-Score integra dimensiones fundamentales del desempeño empresarial

—tales como liquidez, rentabilidad, solvencia y eficiencia— en un solo índice sintético, interpretado como una medida del riesgo de insolvencia.

Una de las principales fortalezas del modelo radica en su carácter operativo y explicativo, ya que no solo entrega una puntuación continua, sino que además clasifica a las empresas en distintas zonas financieras: zona segura, zona gris y zona de riesgo (distress zone), facilitando su interpretación económica y su aplicación práctica en contextos de evaluación financiera, crediticia y de inversión.

Debido a su simplicidad, robustez empírica y capacidad interpretativa, el modelo Z-Score ha sido extensamente utilizado y adaptado en distintos contextos, industrias y mercados, manteniéndose hasta hoy como un punto de referencia obligado en la literatura sobre predicción de quiebras. En este estudio, el Z''-score de Altman se adopta como modelo base, no solo con fines comparativos en términos predictivos, sino también como marco conceptual para analizar la convergencia entre los determinantes financieros tradicionales y los identificados por modelos de aprendizaje automático.

El modelo de Altman dominó durante décadas debido a su:

- Interpretabilidad Directa: Cada variable se corresponde con dimensiones claramente definidas de la salud financiera, como liquidez (X1), solvencia (X2), rentabilidad (X3), apalancamiento (X4) y eficiencia operativa (X5). Esto permite a los analistas financieros comprender *por qué* el modelo predice un resultado.
- Precisión Empírica Robusta: En su validación inicial, el modelo alcanzó una precisión del 72% en predicciones dos años antes del evento, y estudios posteriores confirmaron precisiones del 80-90% un año antes de la insolvencia.
- Simplicidad Computacional: Requiere únicamente información financiera estándar, lo que facilitó su adopción masiva.
- Evolución: A lo largo de más de cinco décadas, Altman desarrolló adaptaciones como el Z'-Score y el Z''-Score (1983) para empresas privadas (manufactureras y no

manufactureras), eliminando el ratio Ventas a Activos Totales (X5) para reducir sesgos específicos de la industria.

Métodos, Datos y Limitaciones Críticas

Los estudios clásicos emplean principalmente datos basados en contabilidad (ratios financieros, balances, estados de resultados). Otros modelos clásicos populares son el de Ohlson (1980), que aplica regresión logística (Logit), y el de Zmijewski (1984), que utiliza un enfoque probit.

Las limitaciones del Z-Score residen principalmente en sus supuestos estadísticos restrictivos:

- **Supuestos Estadísticos Restrictivos:** El MDA subyacente requiere supuestos rigurosos como la multinormalidad, la homocedasticidad y la linealidad.
- **Supuesto de Linealidad:** El modelo asume una relación lineal entre las variables independientes (ratios financieros) y la probabilidad de quiebra. La evidencia empírica sugiere que las relaciones financieras exhiben comportamientos no lineales, donde el impacto marginal de un ratio puede variar significativamente según el nivel inicial de apalancamiento.
- **Incapacidad para Capturar Interacciones Complejas:** El Z-Score procesa cada ratio de forma aislada, sin capturar cómo la interacción entre variables puede modificar su impacto conjunto (por ejemplo, cómo una baja rentabilidad puede ser compensada por una liquidez excepcional).
- **Generalización Limitada y Sensibilidad Contextual:** La aplicación directa del Z-Score (desarrollado con una muestra estadounidense de los años 60) a empresas en diferentes países o contextos (sectores no manufactureros, mercados emergentes) es metodológicamente incorrecta y produce resultados poco fiables. La precisión puede mejorar con la estimación específica por país, tamaño de la empresa e industria.

A pesar de que Altman (1983) recomendó utilizar datos recientes, un estudio de 2017 encontró que la mejora en el rendimiento predictivo del Z"-Score reestimado con datos

internacionales fue solo marginal, lo que indica que los coeficientes originales son extremadamente robustos a lo largo del tiempo y entre países.

El Paradigma Moderno: Aprendizaje Automático (Machine Learning)

El *machine learning* (ML) y la inteligencia artificial (IA) surgen como una respuesta natural a las limitaciones del paradigma clásico. El ML ofrece una flexibilidad no paramétrica, aprendiendo la estructura relacional directamente desde los datos sin requerir supuestos previos de linealidad o normalidad. Estos algoritmos están diseñados para capturar patrones multidimensionales e interacciones no lineales que escapan a los modelos estadísticos convencionales.

Metodologías y Rendimiento Superior

La investigación ha evolucionado desde técnicas estadísticas hacia modelos de ML.

- Modelos Clásicos de IA: Las Redes Neuronales (NN) han sido ampliamente estudiadas y alcanzan precisiones superiores (84-94%) al modelo de Altman. Las Máquinas de Soporte Vectorial (SVM) también son populares, logrando precisiones comparables (80-88%), aunque ambos presentan limitaciones de interpretabilidad.
- Modelos de Conjunto (Ensemble) e Híbridos: Existe una tendencia fuerte hacia el desarrollo de modelos híbridos y de conjunto.
 - Random Forest (RF): Este algoritmo ha mostrado un desempeño sobresaliente, alcanzando precisiones del 95-97% en la predicción de quiebras. Se valora su capacidad para manejar datos de alta dimensionalidad y su robustez contra el sobreajuste y la inestabilidad de los conjuntos de datos desbalanceados.
 - El sobreajuste (o overfitting) ocurre cuando un modelo de aprendizaje automático aprende con tanta precisión los detalles y el ruido de los datos de entrenamiento que pierde su capacidad de generalizar ante información nueva. En lugar de detectar el patrón subyacente, el modelo simplemente "memoriza" los ejemplos, volviéndose extremadamente eficiente con sus propios datos pero fallando

drásticamente al realizar predicciones en el mundo real. Es el equivalente a que un estudiante memorice las respuestas exactas de un examen de práctica, pero no entienda los conceptos necesarios para resolver preguntas diferentes en el examen final.

- Extreme Gradient Boosting (XGBoost): Demuestra ser muy competitivo, reportando precisiones de 92-95%. Es valorado por su capacidad de capturar interacciones no lineales complejas entre variables.

El rendimiento superior es consistente: un estudio de 2017 encontró que los modelos de ML tenían una precisión (71%-87%) superior a los modelos tradicionales (52%-77%).

Desafíos Metodológicos del ML

El campo de la predicción de quiebras presenta desafíos importantes que requieren ajustes metodológicos:

- Desequilibrio de Clases (Class Imbalance): Los conjuntos de datos son inherentemente desequilibrados, con una gran mayoría de empresas no quebradas. Por ejemplo, el dataset taiwanés de referencia muestra un desequilibrio masivo (96.774% no quiebra, 3.226% quiebra).
 - Soluciones: Para mitigar este problema, se utilizan técnicas de remuestreo (*resampling*), como el sobremuestreo (SMOTE) y el submuestreo (Tomek Links, ENN), que han demostrado mejorar notablemente el rendimiento.
 - SMOTE: Es una técnica de sobremuestreo diseñada para equilibrar conjuntos de datos donde una clase es mucho más frecuente que otra. En lugar de simplemente duplicar los registros existentes, SMOTE genera ejemplos sintéticos creando nuevos puntos de datos situados matemáticamente entre los casos reales de la clase minoritaria y sus vecinos más cercanos. Esto permite que el modelo de Inteligencia Artificial aprenda fronteras de decisión más robustas y evite el error de ignorar la clase menos común.

- Tomek Links: Es una técnica de limpieza de datos utilizada para eliminar el ruido y definir mejor la frontera entre clases en conjuntos de datos desequilibrados. A diferencia de SMOTE, esta técnica realiza un submuestreo (undersampling). Identifica pares de puntos muy cercanos pero de clases opuestas: si el punto A y el punto B son los vecinos más cercanos el uno del otro y pertenecen a distintas categorías, forman un "vínculo de Tomek". Al eliminar estos pares (o al menos el ejemplo de la clase mayoritaria), se aumenta el espacio entre las clases, lo que facilita que el modelo identifique correctamente los límites de decisión.
- Evaluación de Modelos: Debido al desequilibrio, la precisión (*Accuracy*) puede ser engañosa. Se prefieren métricas más robustas como el Área Bajo la Curva ROC (AUC), que mide el poder discriminativo, y el Área Bajo la Curva de Precisión-Recall (AUPR). Además, algunos estudios utilizan la métrica F2-measure para penalizar en mayor medida los falsos negativos (FN), que representan el error más costoso (no predecir una quiebra cuando esta sí ocurre).
- Selección de Características (*Feature Selection* - FS): Es fundamental para reducir la dimensionalidad y evitar el sobreajuste. Se emplean métodos *Wrapper* (como BPSO) y la eliminación de variables altamente correlacionadas.

Brechas de Investigación y el Gap de Convergencia Conceptual

A pesar del claro avance en la precisión predictiva, la literatura revela varios vacíos críticos. La principal crítica a la investigación reciente en ML es que ha caído en la paradoja de la precisión hueca: optimización obsesiva de métricas de rendimiento (AUC, Accuracy) a costa de la comprensión teórica.

Vacíos Fundamentales de la Literatura

- Interpretabilidad de ML ("Cajas Negras"): Los modelos más avanzados (NN, RF) son vistos como "cajas negras". La interpretabilidad es crucial en el sector financiero,

donde la regulación exige alta transparencia y justificación racional para las decisiones de riesgo crediticio.

- Integración Híbrida de Datos: La mayoría de los estudios siguen basándose en datos contables, y la integración racional de fuentes de datos no financieros o macroeconómicos sigue siendo un desafío.
- Falta de Análisis Comparativo Estructural: Numerosos estudios comparan la precisión (Altman vs. RF vs. XGBoost), pero prácticamente ninguno examina en profundidad las causas estructurales de la superioridad de ML. A menudo, las comparaciones son sesgadas porque no se permite que los modelos Z-Score "aprendan" de los datos de entrenamiento.

El Gap de Investigación Conceptual (Brecha de Convergencia)

El vacío más sólido y conceptual en la literatura, específicamente entre las herramientas de ML y la disciplina de las finanzas, radica en la falta de diálogo entre las variables predictivas modernas y la teoría financiera clásica.

El gap de investigación actual consiste en la falta de estudios rigurosos que realicen un análisis comparativo profundo de la convergencia conceptual entre (a) los determinantes de quiebra identificados por machine learning mediante técnicas de interpretabilidad (feature importance, SHAP values), y (b) los constructos teóricos financieros subyacentes en el modelo de Altman (liquidez, solvencia, rentabilidad, eficiencia operativa).

La literatura adolece de estudios que pregunten explícitamente si los determinantes identificados por Random Forest confirman, refinan o contradicen los constructos teóricos de Altman.

Implicaciones Teóricas del Gap:

- Si existe convergencia, se validaría la teoría financiera clásica desde una perspectiva *data-driven* moderna, sugiriendo que Altman identificó los determinantes esenciales con una metodología estadística no óptima.

- Si existe divergencia, sugeriría que las dinámicas modernas de insolvencia se han transformado cualitativamente desde 1968, invalidando supuestos teóricos clásicos.

Recientemente, herramientas como SHAP (SHapley Additive exPlanations) han permitido descomponer las predicciones de ML en contribuciones por variable, manteniendo la precisión del modelo original. Esta capacidad abre la posibilidad de un análisis integrado para determinar si los patrones subyacentes de insolvencia convergen o divergen de la teoría de Altman.

En síntesis, si bien la literatura reciente ha demostrado de manera consistente la superioridad predictiva de los modelos de aprendizaje automático frente a los enfoques clásicos, persiste una brecha relevante entre desempeño algorítmico y comprensión financiera. La mayoría de los estudios se concentran en optimizar métricas de clasificación, sin examinar sistemáticamente si los determinantes identificados por modelos data-driven convergen, amplían o contradicen los constructos teóricos que históricamente han estructurado el análisis financiero de la quiebra.

Este vacío limita la integración efectiva de los modelos de machine learning en contextos financieros, donde la interpretabilidad y la coherencia conceptual resultan tan relevantes como la precisión. En consecuencia, el presente estudio se posiciona explícitamente en esta brecha, proponiendo un análisis de convergencia conceptual orientado a evaluar si los patrones detectados por un modelo Random Forest reconstruyen empíricamente las dimensiones financieras clásicas —liquidez, rentabilidad, solvencia y eficiencia— o si revelan nuevas estructuras subyacentes del fenómeno de insolvencia empresarial.

Posicionamiento y Contribución del Proyecto

- **Contraste de Paradigmas:** El objetivo general es evaluar el desempeño predictivo del modelo Random Forest y compararlo con el Z-Score de Altman, reflejando sobre el papel de la IA en la modernización de las herramientas financieras.
- **Análisis de Convergencia:** La contribución principal se centra en el análisis de la correspondencia conceptual y empírica entre las variables relevantes identificadas por el RF (*Feature Importance*) y los componentes financieros clásicos del Z-Score (liquidez, rentabilidad, solvencia, eficiencia y actividad).

Este enfoque es crucial porque ayuda a mejorar la interpretabilidad del RF, aportando evidencia empírica que enriquece la discusión académica sobre la evolución de los modelos de predicción de quiebras. Su trabajo no es solo una aplicación de modelos, sino una integración crítica que busca validar si el rendimiento mejorado de RF se basa en una mayor sofisticación para capturar interacciones complejas o si existe una alineación conceptual significativa con las dimensiones financieras definidas por Altman hace más de 50 años. Al llenar este vacío crítico, su investigación busca avanzar la epistemología de la ciencia financiera mediante la integración de paradigmas estadísticos clásicos y *data-driven* modernos.

Metáfora de Comprensión: Si la predicción de quiebras fuera como la navegación marítima, el Z-Score de Altman es como el sextante y el mapa de papel: una herramienta clásica, robusta y probada que funciona bien si se ajusta al contexto geográfico, pero que se basa en un cálculo lineal con supuestos estrictos. El Random Forest (RF) es el moderno sistema de navegación por satélite (GPS) potenciado con inteligencia artificial: es más rápido y preciso en condiciones complejas (datos no lineales), pero su mecanismo interno (el algoritmo) es menos obvio. Este trabajo busca validar si este GPS moderno (RF), al seleccionar qué variables son cruciales, sigue valorando las mismas "estrellas" y "puntos de referencia" (los ratios financieros) que valoraba el sextante clásico.

Metodología

La presente investigación se enmarca en un enfoque cuantitativo, dado que utiliza información financiera expresada en variables numéricas y evalúa el desempeño de los modelos mediante métricas objetivas de clasificación, tales como *accuracy*, *precision*, *recall* y *F1-score*. Este enfoque permite una comparación sistemática y replicable entre los modelos analizados.

El diseño del estudio es no experimental, ya que no se manipulan las variables independientes ni se interviene sobre el fenómeno de estudio, sino que se analizan datos secundarios previamente registrados en el *Taiwanese Bankruptcy Prediction Dataset*. En este contexto, la investigación se limita a observar las relaciones empíricas existentes entre las variables financieras y la condición de quiebra empresarial.

Asimismo, el estudio adopta un diseño transeccional (transversal), puesto que el análisis se realiza a partir de un corte específico de información financiera, sin considerar la evolución temporal de las empresas. Esto permite evaluar y comparar el desempeño predictivo de los modelos en un contexto estático común.

El alcance de la investigación es descriptivo–comparativo, en la medida en que busca caracterizar y contrastar el comportamiento predictivo del modelo Z-Score de Altman y de un modelo de *Random Forest*, sin pretender establecer relaciones de causalidad, sino describir diferencias y similitudes en términos de desempeño.

Finalmente, la investigación incorpora un componente exploratorio de carácter conceptual, centrado en el análisis de la importancia de las variables del modelo de *Random Forest*, con el objetivo de examinar su coherencia y convergencia con las dimensiones financieras tradicionales propuestas por Altman, contribuyendo así a la interpretación económica de modelos de *machine learning* aplicados a la predicción de quiebra.

La metodología se estructura en cinco fases principales, siguiendo una secuencia lógica diseñada para cumplir progresivamente el objetivo general y los objetivos específicos de la investigación: Preparación y preprocesamiento de los datos, cálculo de modelo z-score Altman, entrenamiento del modelo random forest, comparación del desempeño predictivo, análisis de importancia de variables y convergencia conceptual, y discusión y reproducibilidad.

Preparación y Preprocesamiento de los Datos

Esta fase garantiza un conjunto de datos robusto y limpio para la aplicación rigurosa de ambos modelos.

Selección y Adquisición del Dataset

Se utilizará el Taiwanese Bankruptcy Prediction Dataset, reconocido por su alta dimensionalidad (aproximadamente 6,819 instancias y 96 atributos financieros) y su fuerte desbalance de clases, con cerca del 96% de empresas clasificadas como no-quiebra y el 4% como quiebra.

Exploración, Limpieza y Transformación de Datos (EDA)

Se realizarán actividades de Estadística Descriptiva para comprender y limpiar el conjunto de datos:

1. Exploración de Distribución y Valores Atípicos: Se revisará la distribución de cada variable financiera, y se identificarán y gestionarán datos faltantes (si los hay) y valores atípicos (*outliers*).
2. Deduplicación y Filtrado de Características: Se identificarán y eliminarán las características redundantes con valores duplicados o altamente correlacionadas (con coeficientes de correlación absoluta superiores a un umbral) para evitar la multicolinealidad y la redundancia, lo que reducirá el conjunto inicial de variables (por ejemplo, a 75 variables).
3. Normalización: Se aplicará la estandarización (Z-Score normalization) para que todas las características contribuyan equitativamente al modelo. Adicionalmente, se puede aplicar la transformación logarítmica para reducir el sesgo (*skewness*) de las variables.
4. División de Datos: El dataset se dividirá en un conjunto de entrenamiento (70-80%) y un conjunto de prueba (20-30%). Se empleará el stratified split para preservar la proporción original de las clases (quiebra/no quiebra) en ambos subconjuntos.

Mitigación del Desequilibrio de Clases (Resampling)

Para abordar el desequilibrio crítico, se emplearán técnicas de remuestreo híbridas sobre el *training set*. Estos métodos combinan sobremuestreo de la minoritaria y submuestreo de la mayoritaria:

- SMOTE-Tomek y/o SMOTE-ENN: Estas técnicas generarán ejemplos sintéticos (SMOTE) y luego aplicarán algoritmos de limpieza (Tomek Links o ENN) para eliminar instancias ruidosas o superpuestas, evitando así que el modelo se sesgue hacia la clase mayoritaria.

Cálculo del Modelo Z-score de Altman

Para efectos del cálculo de Z-score se utilizará la versión de 1983 (Z'' -score) ya que se ajusta mejor a las limitaciones y el contexto del dataset trabajado al encontrarse diseñado para abordar empresas que no cotizan en bolsa y reducir el sesgo de la industria.

1. Selección y Construcción de Ratios: Se identificarán las cuatro razones financieras clave necesarias para el modelo Z'' – score:

X_1 : Capital de trabajo / Activos totales

X_2 : Utilidades retenidas / Activos totales

X_3 : Rentabilidad sobre activos (ROA) antes de intereses y después de impuestos

X_4 : Patrimonio / Pasivos totales

Utilizando la mejor aproximación disponible en el dataset taiwanés.

2. Cálculo del Z-score: Se aplicará la fórmula discriminante lineal de Altman, adaptada en su versión Z'' -score para empresas privadas y no manufactureras.

La formulación del modelo Z'' es la siguiente:

$$Z'' = 6,56 \cdot X_1 + 3,26 \cdot X_2 + 6,72 \cdot X_3 + 1,05 \cdot X_4$$

3. Clasificación de Riesgo: Cada empresa será clasificada según los umbrales estándar en zona segura, zona gris o zona de riesgo (*distress*).

- Zonas de clasificación: Zona Segura ($Z > 2,6$)

- Zona Gris ($1,1 < Z < 2,6$)
 - Zona de Riesgo ($Z < 1,1$)
4. Generación de Predicción: Se creará una variable binaria de predicción de quiebra (1/0) a partir de los umbrales de Altman para la comparación directa con el Random Forest.

Entrenamiento del Modelo Random Forest

Se entrenará el modelo de *Machine Learning* utilizando el conjunto de datos optimizado en la fase de EDA.

1. Selección del Modelo: Se utilizará Random Forest (RF) debido a su robustez inherente como algoritmo de conjunto (*ensemble*) y su eficacia probada en conjuntos de datos desbalanceados.
2. Entrenamiento: El RF se entrenará utilizando el conjunto de entrenamiento balanceado y con las características óptimas seleccionadas. Se empleará la Validación Cruzada (*k-fold*) para garantizar una evaluación robusta del rendimiento durante el entrenamiento.
3. Optimización de Hiperparámetros (HPO): Para encontrar la configuración más robusta del modelo, se empleará el método de Random Search (RS). La optimización se centrará en maximizar una métrica robusta para datos desequilibrados, como el AUC. Los parámetros claves a optimizar incluyen: `n_estimators`, `max_depth`, `min_samples_split`, `min_samples_leaf` y `max_features`.

Comparación del Desempeño Predictivo

Se comparará la capacidad predictiva de ambos modelos (Z-score y RF) utilizando el conjunto de prueba (*test set*) que nunca fue visto por los modelos durante el entrenamiento.

1. Generación de Predicciones: Se obtendrán las predicciones binarias para el Z-score y el RF sobre el *test set*.
2. Métricas de Evaluación: Se calcularán las métricas de desempeño estándar, priorizando aquellas que manejan mejor el desequilibrio de clases:
 - Accuracy, Precision, Recall (Sensibilidad), F1-score.

- AUC (Area Under the ROC Curve), que mide el poder discriminativo en todos los umbrales de clasificación.
3. Análisis Cuantitativo Comparativo: La comparación se basará en estos valores para determinar rigurosamente qué enfoque ofrece mejor capacidad predictiva en el dataset y adicionalmente, se utilizarán pruebas de Estadística Inferencial, para comparar AUCs (test de DeLong) y tasas de error (test de McNemar), y determinar si existen diferencias significativas.

Análisis de Importancia de Variables y Convergencia Conceptual

En esta fase contratamos el conocimiento *data-driven* del ML con la teoría financiera clásica.

1. Cálculo de Importancia de Variables
 - Se extraerá el Feature Importance del modelo Random Forest optimizado, utilizando métodos como el Mean Decrease in Impurity (MDI) para cuantificar la contribución de cada variable.
 - Se identificarán y ordenarán los *ratios* financieros que poseen el mayor poder predictivo dentro del modelo no lineal del RF.
2. Análisis de Convergencia Conceptual
 - Se categorizarán las variables más influyentes del RF según las cinco dimensiones financieras clásicas: Solvencia/Endeudamiento, Rentabilidad/Performance, Liquidez, Cobertura/Costo financiero, Crecimiento/Valor.
 - Se evaluará la correspondencia empírica y conceptual entre las variables relevantes del RF y los constructos del Z-score.
 - Este contraste determinará si el modelo RF confirma, refina o contradice los determinantes clásicos de la quiebra, contribuyendo a la interpretabilidad del modelo ML y a la discusión teórico-financiera.

Integración, Discusión y Reproducibilidad

Finalmente, se procederá a la síntesis de los resultados, la discusión del rol del ML en la predicción de insolvencias, y la interpretación de los hallazgos desde una perspectiva teórico-financiera. Para garantizar la confiabilidad y la replicabilidad del estudio, se fijará un *random seed* y se documentará detalladamente el código y los procedimientos de preprocesamiento.

Aplicación Metodológica

Diseño computacional y flujo general del estudio

El presente estudio adopta un enfoque cuantitativo aplicado, apoyado en técnicas de ciencia de datos y aprendizaje automático, con el objetivo de evaluar empíricamente la capacidad predictiva de un modelo de Random Forest en la detección de quiebras empresariales y analizar su convergencia conceptual con el modelo clásico Z''-score de Altman.

Desde el punto de vista computacional, el trabajo fue desarrollado íntegramente en lenguaje Python, utilizando un entorno de notebooks para la exploración, procesamiento y modelamiento de los datos. El diseño metodológico se estructuró como un flujo secuencial de etapas, donde cada una cumple un rol específico dentro del proceso de construcción, evaluación e interpretación del modelo.

De manera general, el flujo del estudio se organizó en las siguientes fases:

1. Adquisición y comprensión del conjunto de datos: Se cargó el conjunto de datos de empresas, verificando su estructura, dimensionalidad y tipo de variables. En esta etapa se analizó la variable objetivo asociada a quiebra empresarial, evaluando su distribución y detectando desde un inicio un fuerte desbalance entre clases, condición que influye directamente en las decisiones metodológicas posteriores.

Esta fase permitió caracterizar el problema como un caso de clasificación binaria desbalanceada, lo que justificó el uso de métricas específicas y técnicas de remuestreo.

2. Preprocesamiento y preparación de los datos: Posteriormente se efectuaron las tareas de limpieza y preparación del conjunto de datos, incluyendo:
 - verificación de valores faltantes y atípicos,
 - separación entre variables explicativas y variable objetivo,
 - partición del dataset en conjuntos de entrenamiento y prueba mediante muestreo estratificado, con el fin de preservar la proporción original de clases.

Esta etapa tuvo como objetivo garantizar la coherencia estadística de los subconjuntos y evitar filtraciones de información entre entrenamiento y evaluación.

3. Análisis descriptivo y línea base: Previo al entrenamiento de modelos complejos, se realizó un análisis descriptivo inicial que permitió contextualizar el problema, destacando el grado de desbalance de clases y calculando indicadores asociados al Z''-score de Altman. En paralelo, se construyó una línea base mediante la aplicación directa del modelo clásico, la cual sirve como punto de comparación metodológica y conceptual para los modelos de aprendizaje automático.
4. Construcción del pipeline de modelamiento: Se diseñó un pipeline de modelamiento que integró:
 - técnicas de remuestreo para el tratamiento del desbalance de clases,
 - el algoritmo Random Forest como modelo principal,
 - procedimientos de validación cruzada y optimización de hiperparámetros.

Esta estructura permitió encapsular de forma coherente las distintas transformaciones del flujo, asegurando reproducibilidad, control experimental y correcta evaluación fuera de muestra.

5. Optimización y evaluación del modelo: El modelo fue entrenado utilizando validación cruzada estratificada, maximizando el área bajo la curva ROC (AUC) como métrica principal de optimización. Posteriormente, el desempeño final se evaluó sobre el conjunto de prueba, empleando métricas robustas tales como recall, precisión, F1-score y AUC, junto con matrices de confusión, permitiendo un análisis detallado del comportamiento predictivo, especialmente sobre la clase minoritaria.
6. Interpretabilidad y análisis conceptual: Finalmente, se desarrolló un bloque específico orientado a la interpretabilidad del modelo, que incluyó:
 - análisis de importancia de variables,
 - estudio de importancia acumulada,

- evaluación mediante permutation importance,
- y clasificación de las variables según dimensiones financieras.

Este proceso permitió agrupar los determinantes del modelo en ejes financieros (rentabilidad, solvencia, liquidez, cobertura y crecimiento), facilitando la comparación estructural con los componentes del modelo de Altman y el análisis de convergencia conceptual entre enfoques clásicos y modernos.

7. Síntesis de resultados: Como etapa final, se integraron los resultados predictivos y conceptuales, construyendo tablas y figuras de síntesis que permiten evaluar tanto el desempeño empírico de los modelos como la coherencia financiera de los determinantes identificados.
8. Cierre del diseño: En conjunto, este diseño computacional permitió abordar el problema de predicción de quiebras desde una doble perspectiva: por un lado, la eficiencia predictiva de modelos de aprendizaje automático; y por otro, la interpretación económica de sus resultados, estableciendo un puente metodológico entre la modelación algorítmica y la teoría financiera clásica.

Descripción del dataset y validaciones iniciales

Se utilizó el *Taiwanese Bankruptcy Prediction Dataset*, compuesto por 6.819 observaciones (filas) y 96 variables financieras (columnas), estructuradas principalmente como indicadores contables y ratios. La variable objetivo corresponde a Bankrupt?, codificada de forma binaria: 0 = empresa financieramente estable y 1 = empresa financieramente inestable (quiebra/insolvencia).

La distribución de clases evidencia un fuerte desbalance, con 6.599 observaciones (96,77%) en la clase estable (0) y 220 observaciones (3,23%) en la clase inestable (1).

Estado financiero	Código	Frecuencia	Porcentaje (%)
Empresa estable	0	6.599	96,77
Empresa en quiebra	1	220	3,23
Total	—	6.819	100

Tabla 1. Distribución de la variable objetivo (Bankrupt?)
Fuente: *Elaboración propia.*

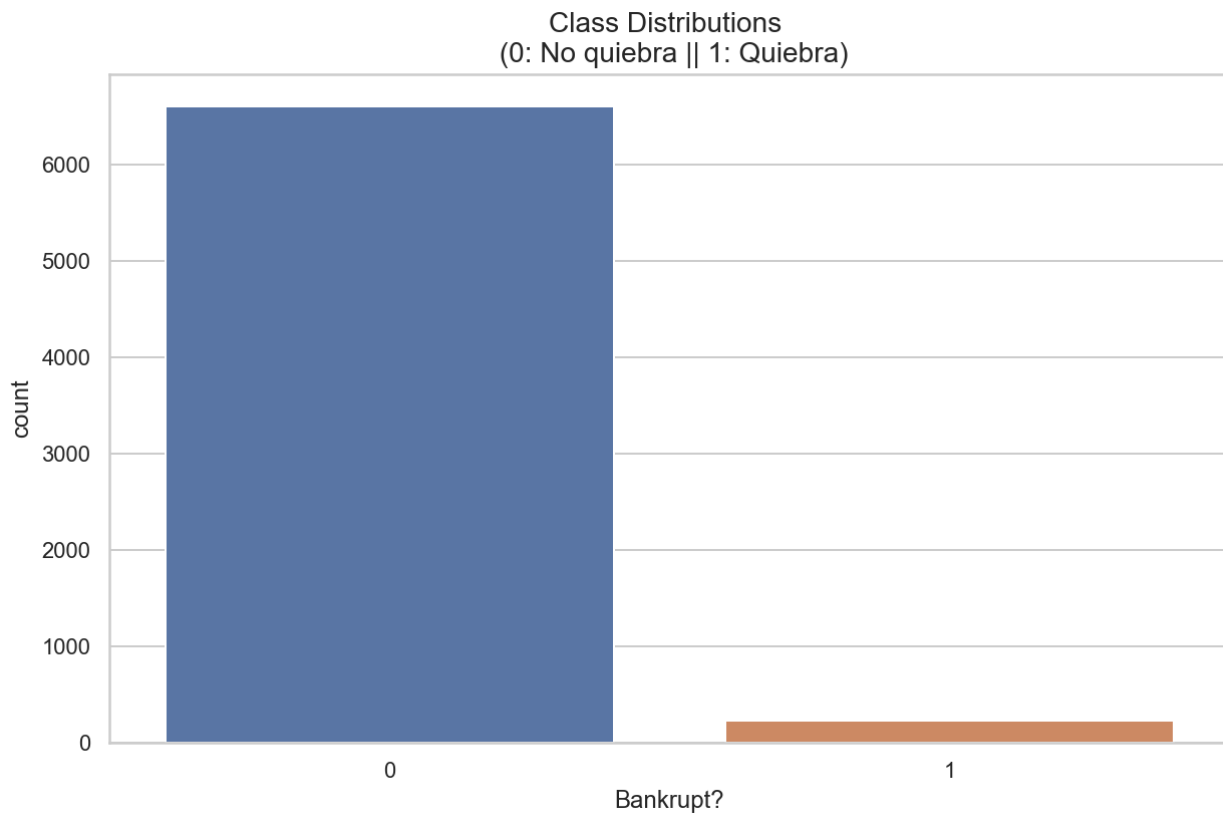


Figura 1. Distribución de clases del dataset (0: estable; 1: inestable)
Fuente: *Elaboración propia.*

Este desbalance es consistente con la naturaleza del fenómeno de insolvencia, donde los eventos de quiebra suelen ser relativamente infrecuentes en comparación con la población de empresas operativas. En consecuencia, este rasgo condiciona la estrategia metodológica del estudio, ya que métricas globales como *accuracy* pueden resultar engañosas si el modelo se sesga hacia la clase mayoritaria. Por lo anterior, en etapas posteriores se incorporan técnicas

específicas de manejo del desbalance y se prioriza el uso de métricas robustas para clasificación desbalanceada (p. ej., *recall*, *F1* y *AUC*).

Como parte del control de calidad, se realizaron validaciones iniciales orientadas a asegurar consistencia del conjunto de datos: revisión de estructura general, verificación de presencia de valores faltantes y detección de filas duplicadas. Estas comprobaciones permiten establecer una línea base de integridad del dataset antes de aplicar transformaciones, construir el Z-score y entrenar modelos de aprendizaje automático.

Análisis Exploratorio de Datos (EDA)

El análisis exploratorio se orientó a caracterizar (i) la magnitud del desbalance de la variable objetivo, (ii) señales preliminares de diferenciación entre empresas estables e inestables en variables financieras representativas, y (iii) relaciones monotónicas y potencial redundancia entre variables. En primer lugar, la distribución observada confirma un escenario de clasificación altamente asimétrico, en el cual un modelo trivial que prediga siempre la clase mayoritaria podría alcanzar una exactitud elevada sin aportar capacidad real para detectar quiebras. Por este motivo, el estudio considera desde su etapa exploratoria la necesidad de utilizar estrategias de entrenamiento y evaluación que privilegien la detección de la clase minoritaria.

En segundo lugar, se evaluaron variables seleccionadas mediante diagramas de caja segmentados por condición de quiebra. De manera consistente con la intuición financiera clásica, se observan diferencias sistemáticas entre los grupos en variables asociadas a liquidez, rentabilidad y estructura de capital/apalancamiento.

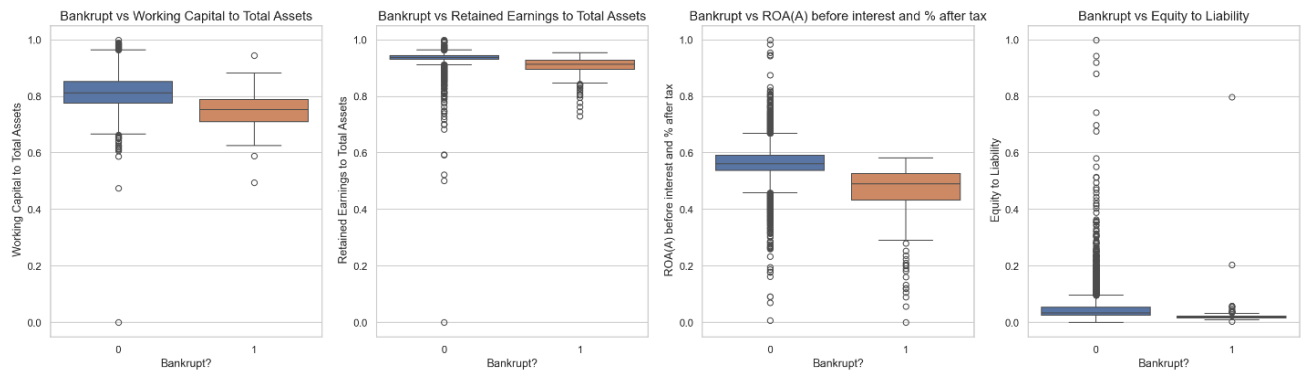


Figura 2. Boxplots variables financieras presentes en Z-Score Altman según condición de quiebra

Fuente: *Elaboración propia.*

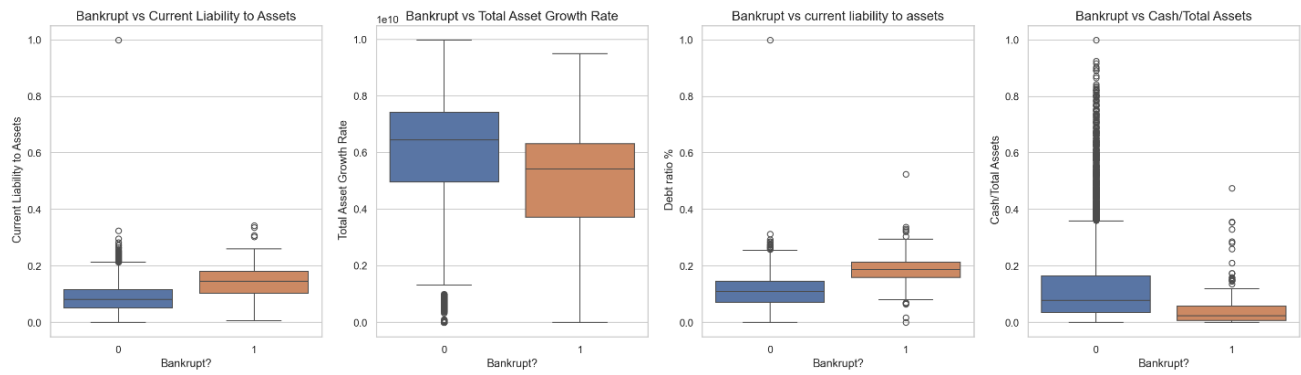


Figura 3. Boxplots variables financieras no presentes en Z''-Score según condición de quiebra
Fuente: *Elaboración propia.*

Las variables ilustran diferencias preliminares en dimensiones de liquidez, rentabilidad y apalancamiento. En términos generales, las empresas clasificadas como inestables tienden a exhibir niveles inferiores en indicadores de rentabilidad (p. ej., medidas tipo ROA) y patrones compatibles con una estructura financiera más frágil. Estos resultados se interpretan como evidencia exploratoria de separación potencial entre clases, sin implicar relaciones causales.

Finalmente, se analizó la dependencia entre variables mediante una matriz de correlación (Spearman).



Figura 4. Matriz de correlación (Spearman) entre variables financieras
Fuente: Elaboración propia.

El mapa de calor evidencia bloques de correlación entre múltiples variables, lo que sugiere redundancia informativa en parte del set de características. Este hallazgo sustenta metodológicamente la posterior aplicación de criterios de preprocesamiento orientados a reducir colinealidad/redundancia, favoreciendo un conjunto de variables más parsimonioso y estable para el entrenamiento, además de facilitar el análisis de interpretabilidad del modelo Random Forest.

Preprocesamiento y preparación de los datos

El preprocesamiento tuvo como objetivo garantizar un conjunto de datos consistente, robusto y adecuado para la aplicación tanto del modelo clásico Z''-score de Altman como del modelo de aprendizaje automático Random Forest. Las transformaciones realizadas se orientaron a controlar la influencia de valores extremos, reducir redundancia informativa y preparar un conjunto de características que favorezca la estabilidad del entrenamiento y la posterior interpretación de resultados.

Depuración inicial del dataset

Como primera etapa, se verificó la integridad del conjunto de datos. No se detectaron valores faltantes en ninguna de las variables financieras, ni la existencia de filas duplicadas, por lo que no fue necesario aplicar técnicas de imputación ni eliminación de observaciones. En consecuencia, se conservó la totalidad de las 6.819 empresas, evitando la pérdida de información relevante, especialmente crítica dada la baja proporción de casos de quiebra.

Asimismo, se confirmó que la variable objetivo Bankrupt? se encontraba correctamente definida como variable binaria, separándose explícitamente del conjunto de características para evitar cualquier tipo de fuga de información (*data leakage*) durante el preprocesamiento.

Tratamiento de valores extremos (winsorización)

El análisis exploratorio evidenció la presencia de valores extremos en múltiples variables financieras, fenómeno habitual en bases contables de alta dimensionalidad. Dado el fuerte desbalance de clases y la baja frecuencia de observaciones correspondientes a quiebras, se optó por no eliminar registros, ya que dicha estrategia podría suprimir información crítica asociada precisamente a empresas en condiciones financieras extremas.

En su lugar, se aplicó un procedimiento de winsorización exclusivamente sobre el conjunto de variables explicativas, manteniendo intacta la variable objetivo. Este enfoque permite limitar la influencia de outliers mediante el recorte de los valores en los percentiles extremos, reduciendo la distorsión en la distribución de las variables sin alterar el tamaño muestral ni la estructura de clases.

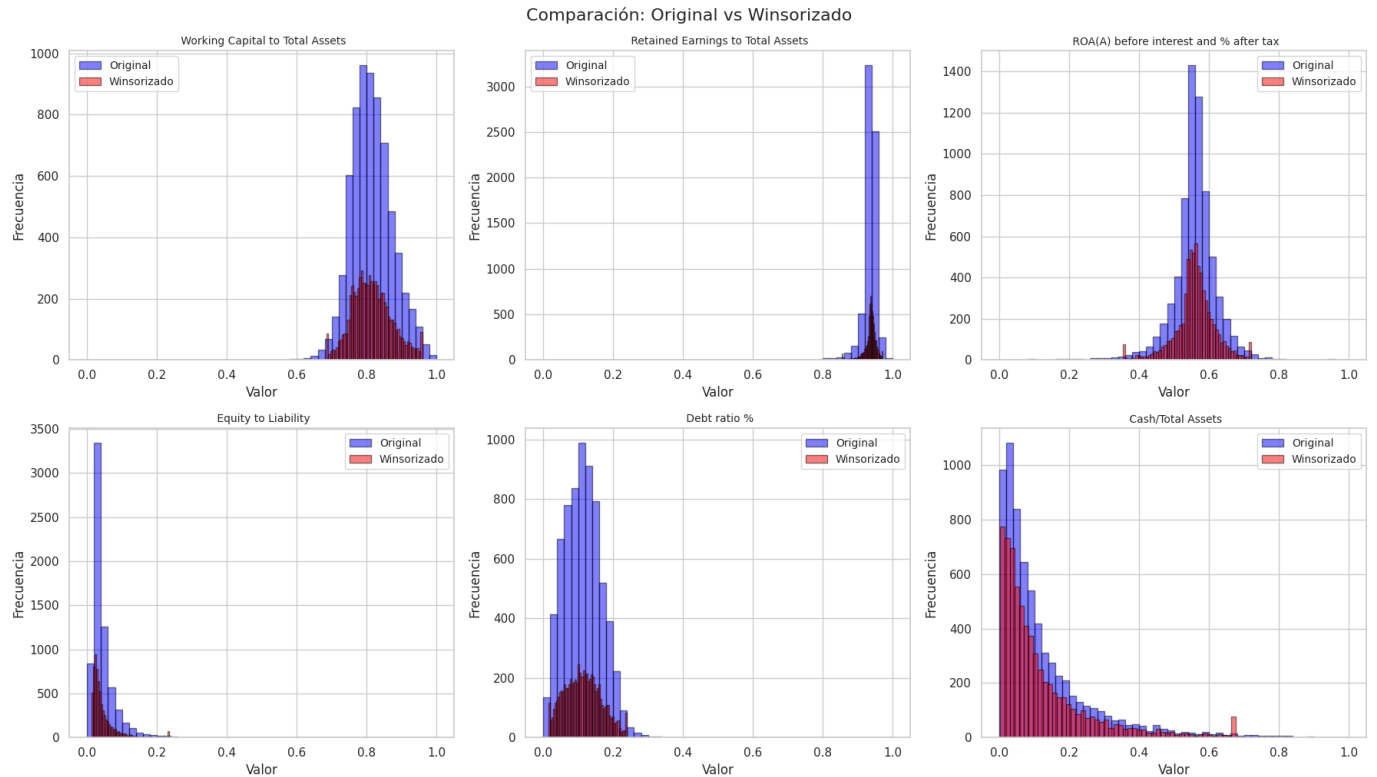


Figura 5. Comparación de distribuciones antes y después de la winsorización para variables financieras seleccionadas.

Fuente: Elaboración propia.

La comparación entre las distribuciones originales y winsorizadas confirma una reducción significativa de colas extremas, preservando al mismo tiempo la forma general de las distribuciones.

Reducción de redundancia mediante análisis de correlación

Con el fin de disminuir la redundancia informativa y mitigar problemas derivados de alta colinealidad, se aplicó un análisis de correlación entre las variables explicativas utilizando el coeficiente de Spearman, adecuado para relaciones monótonicas no necesariamente lineales. Se estableció un umbral de $|\rho| > 0,95$ como criterio para identificar pares de variables altamente correlacionadas.

Como resultado de este procedimiento, se eliminaron 17 variables, reduciendo el conjunto inicial desde 95 variables financieras a 78, manteniendo las 6.819 observaciones originales. Esta reducción busca favorecer un conjunto de características más parsimonioso,

facilitar el análisis de importancia de variables en el Random Forest y disminuir la probabilidad de que múltiples variables representen esencialmente la misma dimensión financiera.

Consideraciones de escalamiento y definición del set final

No se aplicaron técnicas de escalamiento o normalización a las variables financieras. Esta decisión se fundamenta en dos razones metodológicas. En primer lugar, el algoritmo Random Forest se basa en particiones de árboles de decisión, por lo que no es sensible a la escala de las variables, a diferencia de modelos basados en distancias o gradientes. En segundo lugar, el cálculo del Z''-score de Altman se basa en ratios financieros interpretables en su escala original, por lo que estandarizarlos alteraría su significado económico.

De este modo, el conjunto final de datos preprocesados quedó conformado por 6.819 observaciones y 78 variables explicativas, constituyendo la base sobre la cual se realizarán posteriormente la separación en conjuntos de entrenamiento y prueba, el manejo del desbalance de clases y el entrenamiento de los modelos predictivos.

Construcción del Z-score de Altman

El modelo Z-score de Altman se utiliza en este estudio como referente clásico para la predicción de quiebras y como punto de contraste conceptual y empírico frente al modelo de Random Forest. En particular, se adopta la versión Z''-score, propuesta por Altman (1983), diseñada para empresas no manufactureras y privadas, dado que elimina el componente de rotación de activos presente en la versión original y reduce el sesgo asociado a diferencias sectoriales, haciéndola más adecuada para contextos empresariales heterogéneos como el del dataset taiwanés.

La formulación del modelo Z'' es la siguiente:

$$Z'' = 6,56 \cdot X_1 + 3,26 \cdot X_2 + 6,72 \cdot X_3 + 1,05 \cdot X_4$$

donde:

X_1 : Capital de trabajo / Activos totales

X_2 : Utilidades retenidas / Activos totales

X_3 : Rentabilidad sobre activos (ROA) antes de intereses y después de impuestos

X_4 : Patrimonio / Pasivos totales

Selección y mapeo de variables financieras

A partir del conjunto de datos preprocesado, se identificaron las variables que mejor aproximan los componentes financieros definidos por Altman. El mapeo realizado fue el siguiente:

X_1 : Working Capital to Total Assets

X_2 : Retained Earnings to Total Assets

X_3 : ROA(A) before interest and % after tax

X_4 : Equity to Liability

Este procedimiento permitió construir un subconjunto específico de variables para el cálculo del Z''-score, asegurando coherencia conceptual entre los ratios originales del modelo y los indicadores financieros disponibles en el dataset.

Componente Altman	Definición financiera	Variable en el dataset
X1	Working Capital / Total Assets	Working Capital to Total Assets
X2	Retained Earnings / Total Assets	Retained Earnings to Total Assets
X3	ROA before interest and after tax	ROA(A) before interest and % after tax
X4	Equity / Total Liabilities	Equity to Liability

Tabla 2. Correspondencia entre variables del dataset y componentes del modelo Z'' de Altman.

Fuente: Elaboración propia.

Cálculo del Z''-score

Con base en las variables seleccionadas, se construyó un dataframe específico para el modelo Altman y se aplicó la función discriminante correspondiente a cada observación. El cálculo se realizó para la totalidad de las 6.819 empresas, generando una nueva variable continua Z''-Score, que resume la condición financiera relativa de cada firma según el enfoque clásico.

Los estadísticos descriptivos del Z''-score evidencian una distribución altamente concentrada, con valores promedio elevados, lo cual es coherente con la predominancia de empresas clasificadas como financieramente estables en el dataset.

Estadístico	Z''-Score
Count	6.819
Mean	12,19
Std	0,74
Min	3,84
25%	11,82
50%	12,18
75%	12,59
Max	15,91

Tabla 3. Estadísticos descriptivos del Z''-score de Altman.
Fuente: Elaboración propia.

Análisis de la distribución del Z''-score

La distribución empírica del Z''-score muestra una forma aproximadamente unimodal y concentrada en torno a valores elevados, reflejando que la mayoría de las empresas se ubican, según el criterio de Altman, en zonas de relativa estabilidad financiera. La inspección visual permite identificar asimetrías y colas que representan observaciones con condiciones financieras significativamente más frágiles.

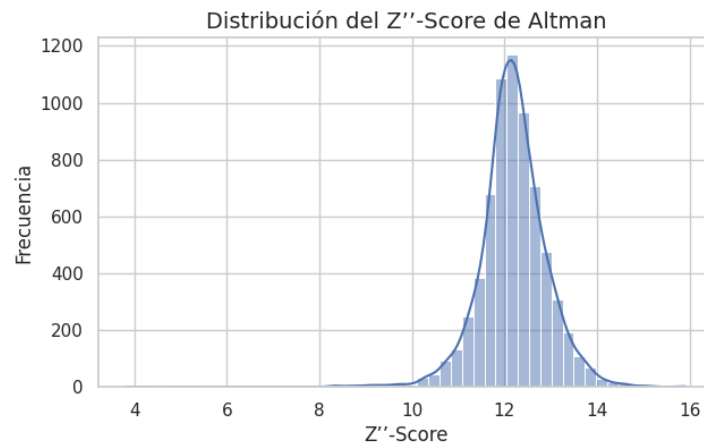


Figura 6. Distribución del Z''-score de Altman en el conjunto de datos.
Fuente: Elaboración propia.

Z''-score según estado de quiebra

Con el fin de evaluar la capacidad discriminante preliminar del modelo, se analizó el comportamiento del Z''-score en función del estado real de quiebra. El diagrama de cajas muestra una diferenciación sistemática entre empresas estables e inestables, observándose que, en

promedio, las empresas clasificadas como no quebradas presentan valores de Z'' superiores a aquellas que efectivamente experimentaron quiebra.

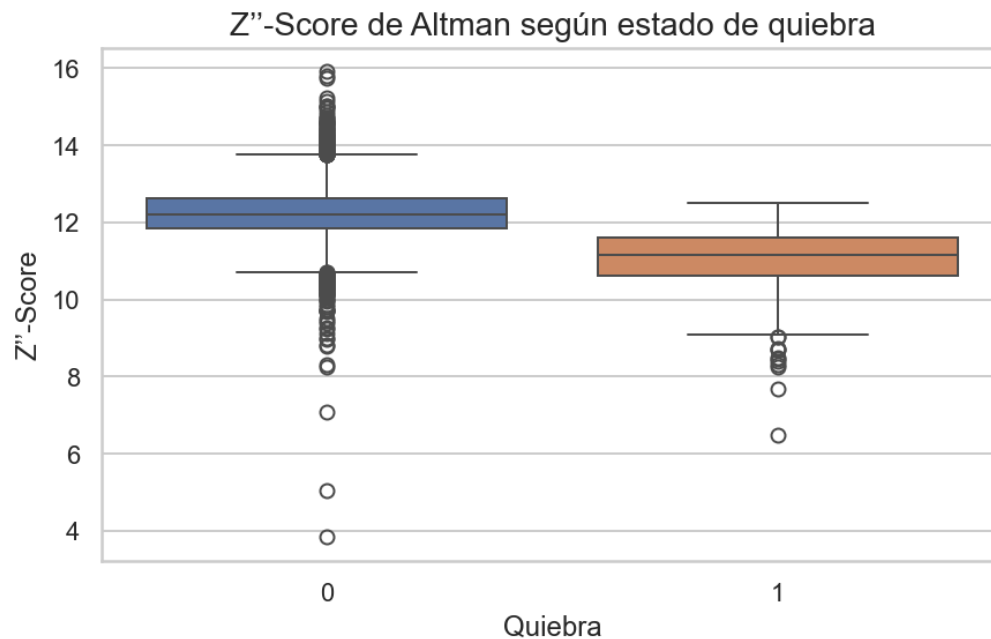


Figura X. Distribución del Z'' -score según estado de quiebra.
Fuente: Elaboración propia.

Si bien existe solapamiento entre distribuciones —fenómeno esperable dada la complejidad del proceso de insolvencia— este resultado exploratorio sugiere que el Z'' -score captura información relevante asociada al riesgo financiero, validando su utilización como modelo de referencia para la comparación empírica posterior.

Altman como modelo de clasificación

Si bien el Z'' -score de Altman se construye originalmente como un índice continuo de salud financiera, su aplicación práctica más habitual consiste en clasificar empresas en categorías de riesgo. En este estudio, el modelo se utiliza explícitamente como un clasificador binario de quiebra, con el fin de establecer un punto de comparación empírico directo frente al modelo de Random Forest.

Clasificación por zonas clásicas y limitaciones

En una primera aproximación, se aplicaron los umbrales tradicionales propuestos por Altman, que segmentan las empresas en zonas de riesgo financiero (zona segura, zona gris y zona de quiebra). Sin embargo, al trasladar estas zonas a una clasificación binaria, se observó que la totalidad de las empresas quedaban clasificadas como no quebradas, impidiendo la discriminación efectiva de los eventos de quiebra.

Este resultado sugiere que los puntos de corte originales, calibrados sobre empresas estadounidenses en décadas anteriores, no resultan directamente transferibles al contexto del dataset taiwanés, caracterizado por escalas distintas y una estructura financiera específica.

Orientación del score y análisis ROC

Con el objetivo de evaluar la capacidad discriminante real del Z'' -score, se construyó la curva ROC, analizando la relación entre tasa de verdaderos positivos y falsos positivos para distintos umbrales. Este análisis evidenció que el score se encontraba inversamente relacionado con la probabilidad de quiebra, es decir, a menor Z'' -score, mayor riesgo de quiebra, coherente con la interpretación financiera del modelo.

Por esta razón, se ajustó la orientación del score al construir la función discriminante, de modo que los valores más bajos del Z'' -score representaran mayor probabilidad del evento “quiebra”. Tras este ajuste, la curva ROC mostró un desempeño consistente, con un AUC cercano a 0,89, confirmando que el índice contiene señal predictiva relevante.

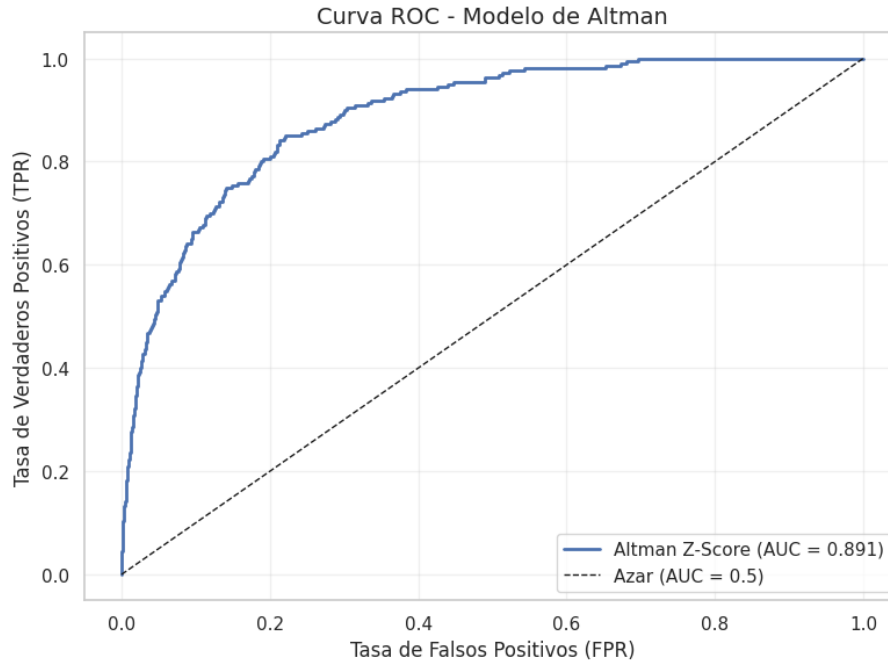


Figura 8. Curva ROC del modelo Z'' -score de Altman como clasificador binario.
 Fuente: Elaboración propia.

Determinación del umbral óptimo

A partir de la curva ROC corregida, se seleccionó un umbral óptimo utilizando el criterio de maximización del índice de Youden (TPR – FPR), obteniéndose un punto de corte aproximado de:

$$Z'' \approx 11,80$$

Este valor define la regla de clasificación utilizada en adelante: empresas con Z'' inferior al umbral se clasifican como en riesgo de quiebra, mientras que aquellas con valores superiores se clasifican como financieramente estables.

La aplicación de este umbral generó una distribución de predicciones considerablemente distinta a la proporción real de quiebras, clasificando un número elevado de empresas como potencialmente inestables, lo que anticipa un comportamiento conservador del modelo.

Predicción Altman	Frecuencia
0 (No quiebra)	5180
1 (Quiebra)	1639
Total	6819

Tabla 4. Distribución de empresas clasificadas por el modelo de Altman.
 Fuente: Elaboración propia.

Evaluación del desempeño clasificatorio

La capacidad predictiva del modelo se evaluó mediante una matriz de confusión y métricas estándar de clasificación. Los resultados muestran que el modelo de Altman logra capturar una alta proporción de las quiebras reales (recall elevado), a costa de una baja precisión, clasificando como en riesgo a un número significativo de empresas que no experimentaron quiebra.



Figura 9. Matriz de confusión del modelo Z''-score de Altman.
Fuente: Elaboración propia.

Clase	Precision	Recall	F1-score	Soporte
0 (No quiebra)	0.99	0.78	0.87	6599
1 (Quiebra)	0.11	0.85	0.20	220
Accuracy			0.78	6819
Macro avg	0.55	0.81	0.54	6819
Weighted avg	0.97	0.78	0.85	6819

Tabla 5. Métricas de desempeño del modelo Z''-score de Altman.
Fuente: Elaboración propia.

Este comportamiento refleja una lógica coherente con la naturaleza del modelo: prioriza la detección de empresas potencialmente frágiles, aun cuando ello implique un aumento considerable en falsos positivos.

Síntesis metodológica

En conjunto, estos resultados evidencian que el modelo de Altman contiene información discriminante relevante, pero presenta limitaciones importantes al ser aplicado como clasificador binario en un entorno altamente desbalanceado y contemporáneo. En particular, se observa un trade-off marcado entre sensibilidad y precisión, lo que justifica la necesidad de explorar modelos de aprendizaje automático capaces de capturar relaciones no lineales y de adaptarse al desbalance de clases.

Este comportamiento servirá como línea base empírica para la comparación posterior con el modelo de Random Forest.

Preparación del set de entrenamiento y test

Con el fin de evaluar de manera objetiva el desempeño de los modelos predictivos, el conjunto de datos preprocesado se dividió en subconjuntos de entrenamiento y prueba, utilizando una partición 80/20. Dado el fuerte desbalance de clases presente en la variable objetivo, la división se realizó de forma estratificada, asegurando que la proporción de empresas quebradas y no quebradas se mantuviera consistente en ambos conjuntos.

Conjunto	Clase	Frecuencia	Proporción
Completo	No quiebra	6.599	96,77%
Completo	Quiebra	220	3,23%
Entrenamiento	No quiebra	5.279	96,77%
Entrenamiento	Quiebra	176	3,23%
Prueba	No quiebra	1.320	96,77%
Prueba	Quiebra	44	3,23%

Tabla 6. Distribución de clases en el dataset completo y en los conjuntos de entrenamiento y prueba.

Fuente: Elaboración propia.

De este modo, el conjunto de entrenamiento quedó conformado por 5.455 observaciones, mientras que el conjunto de prueba quedó compuesto por 1.364 empresas, conservando en ambos casos una proporción aproximada de 3,23 % de quiebras.

Congelamiento de conjuntos y control de integridad

Una vez realizada la partición, se crearon copias explícitas de los conjuntos de entrenamiento y prueba, las cuales fueron utilizadas de manera independiente a lo largo del flujo experimental. Este procedimiento tiene como objetivo evitar modificaciones accidentales de los datos originales y reducir el riesgo de fugas de información (data leakage) entre las distintas etapas de preprocesamiento, remuestreo y entrenamiento de modelos.

Esta decisión metodológica busca fortalecer la reproducibilidad del estudio y garantizar que las evaluaciones posteriores se realicen exclusivamente sobre datos no utilizados durante el entrenamiento.

Clasificador base trivial

Antes de entrenar modelos predictivos, se construyó un clasificador base que asigna todas las observaciones a la clase mayoritaria (empresa no quebrada). Este modelo ingenuo permite establecer un punto de referencia mínimo, particularmente relevante en contextos altamente desbalanceados, donde métricas como la exactitud pueden resultar engañosas.

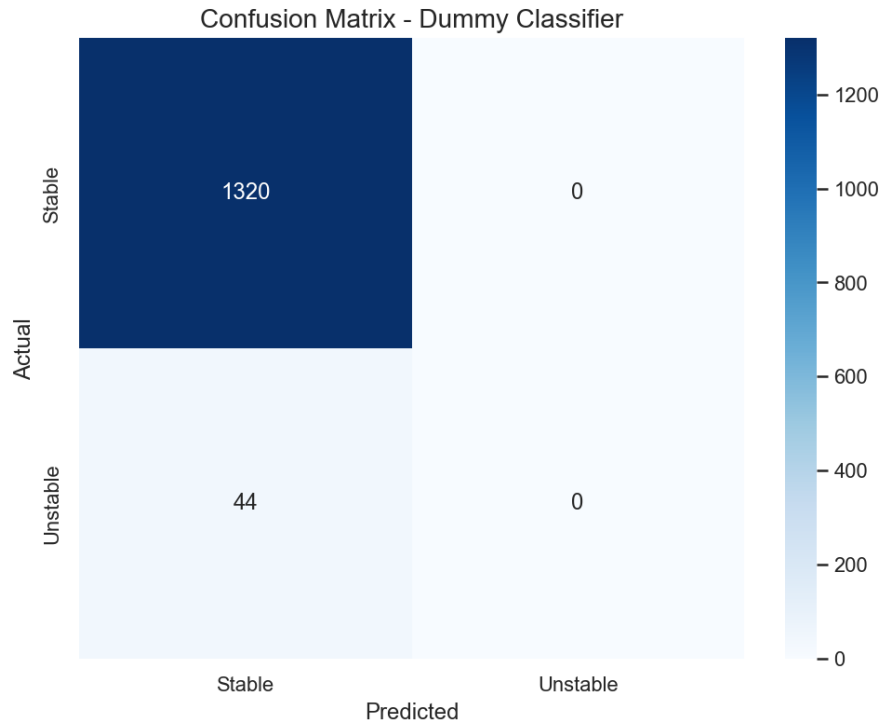


Figura 10. Matriz de confusión del clasificador base trivial en el conjunto de prueba.
Fuente: *Elaboración propia.*

Clase	Precision	Recall	F1-score	Soporte
0 (No quiebra)	0,97	1	0,98	1.320
1 (Quiebra)	0	0	0	44
Accuracy			0,97	1.364
Macro avg	0,48	0,5	0,49	1.364
Weighted avg	0,94	0,97	0,95	1.364

Tabla 7. Métricas de desempeño del clasificador base trivial.
Fuente: *Elaboración propia.*

La evaluación del clasificador trivial sobre el conjunto de prueba muestra una exactitud elevada, producto del predominio de empresas sanas; sin embargo, presenta incapacidad total para detectar quiebras, obteniendo recall nulo para la clase minoritaria. Este resultado evidencia que cualquier modelo útil debe superar significativamente este desempeño base, especialmente en términos de sensibilidad frente a empresas en riesgo.

Implicancias metodológicas

La marcada diferencia entre la alta exactitud del clasificador trivial y su nula capacidad para detectar quiebras confirma que el problema abordado no puede evaluarse adecuadamente mediante métricas agregadas. En consecuencia, los modelos posteriores se evaluarán priorizando indicadores como recall, precisión, F1-score y AUC-ROC, orientados a capturar de manera más fiel el desempeño sobre la clase minoritaria.

Este enfoque justifica, además, la incorporación de técnicas específicas de manejo del desbalance de clases, abordadas en el apartado siguiente.

Manejo del desbalance de clases

El conjunto de entrenamiento presenta un fuerte desbalance de clases, con aproximadamente un 3,23 % de empresas en quiebra frente a un 96,77 % de empresas financieramente estables. Este escenario plantea un desafío relevante para el entrenamiento de modelos de clasificación, ya que los algoritmos tienden a sesgarse hacia la clase mayoritaria, reduciendo significativamente su capacidad para detectar eventos poco frecuentes pero críticos, como la quiebra empresarial.

Por este motivo, se implementaron técnicas específicas de remuestreo, aplicadas exclusivamente sobre el conjunto de entrenamiento, preservando intacto el conjunto de prueba con el fin de garantizar evaluaciones realistas y libres de fuga de información.

Distribución original del conjunto de entrenamiento

Previo a cualquier intervención, se analizó la distribución de clases en el conjunto de entrenamiento, confirmándose la marcada asimetría entre empresas quebradas y no quebradas.

Conjunto	Clase	Frecuencia	Proporción
Entrenamiento (original)	No quiebra	5.279	96,77%
Entrenamiento (original)	Quiebra	176	3,23%
Total	—	5.455	100%

Tabla 8. Distribución de clases en el conjunto de entrenamiento original.

Fuente: Elaboración propia.

Aplicación de SMOTE

Como primera estrategia de balanceo se utilizó la técnica SMOTE (Synthetic Minority Over-sampling Technique), la cual genera observaciones sintéticas de la clase minoritaria a partir de combinaciones lineales entre un punto y sus vecinos más cercanos en el espacio de características. Este procedimiento permite aumentar la representación de la clase de quiebra sin replicar directamente observaciones existentes, reduciendo el riesgo de sobreajuste.

Tras aplicar SMOTE, el conjunto de entrenamiento quedó perfectamente balanceado, alcanzando una proporción aproximada de 50 % de empresas quebradas y 50 % de no quebradas.

Conjunto	Clase	Frecuencia	Proporción
Entrenamiento SMOTE	No quiebra	5.279	50,00%
Entrenamiento SMOTE	Quiebra	5.279	50,00%
Total	—	10.558	100%

Tabla 9. Distribución de clases en el conjunto de entrenamiento tras aplicar SMOTE.

Fuente: Elaboración propia.

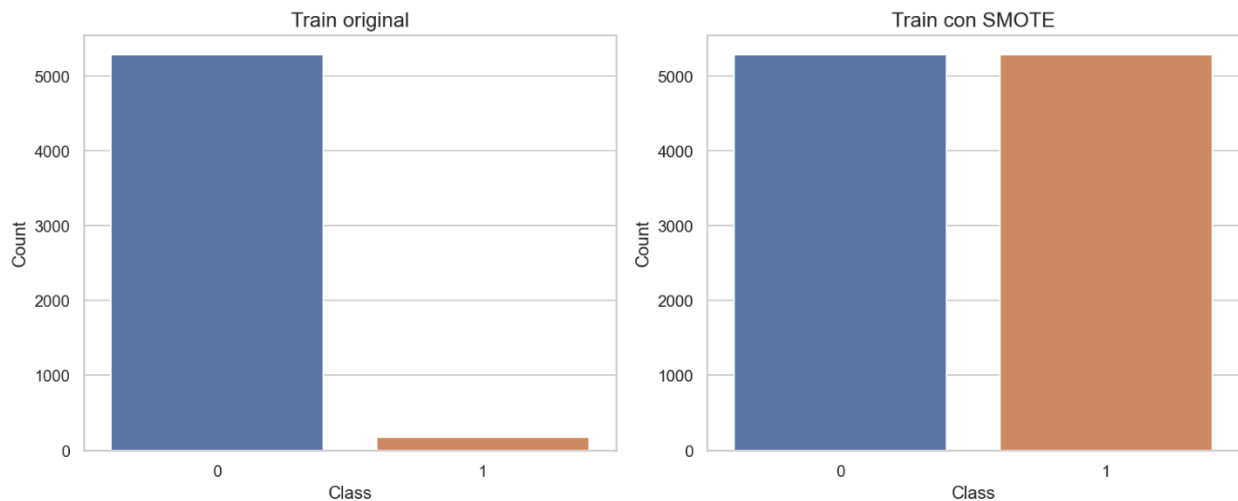


Figura 11. Comparación de la distribución de clases antes y después de aplicar SMOTE.

Fuente: Elaboración propia.

Aplicación de SMOTETomek

Adicionalmente, se evaluó la técnica SMOTETomek, la cual combina el oversampling sintético de SMOTE con la eliminación de enlaces de Tomek, que corresponden a pares de

observaciones de clases opuestas ubicadas muy próximas entre sí. Este método busca no solo balancear el conjunto de datos, sino también limpiar la frontera de decisión, removiendo observaciones potencialmente ambiguas o ruidosas.

La aplicación de SMOTETomek produjo un conjunto de entrenamiento igualmente balanceado, con una leve reducción en el número total de observaciones producto de la eliminación de ejemplos cercanos a la frontera entre clases.

Conjunto	Clase	Frecuencia	Proporción
Entrenamiento SMOTE - Tomek	No quiebra	5.250	50,00%
Entrenamiento SMOTE - Tomek	Quiebra	5.250	50,00%
Total	—	10.500	100%

Tabla 10. Distribución de clases en el conjunto de entrenamiento tras aplicar SMOTE - Tomek.
Fuente: Elaboración propia.

Consideraciones metodológicas

Las técnicas de remuestreo fueron utilizadas exclusivamente durante la etapa de entrenamiento, con el objetivo de mejorar la capacidad de aprendizaje del modelo sobre la clase minoritaria, manteniendo el conjunto de prueba como una aproximación fiel al entorno real. Este enfoque permite evaluar el desempeño de los modelos bajo condiciones de fuerte desbalance, coherentes con el contexto práctico de predicción de quiebras.

Los conjuntos balanceados obtenidos mediante SMOTE y SMOTETomek se utilizarán en los apartados siguientes para entrenar y optimizar los modelos de Random Forest, permitiendo comparar empíricamente el impacto de las distintas estrategias de balanceo.

Entrenamiento del Random Forest base

Con el objetivo de establecer un punto de referencia para los procesos posteriores de optimización, se entrenó inicialmente un modelo Random Forest base, utilizando hiperparámetros estándar y evaluando su desempeño mediante validación cruzada. Este procedimiento permitió analizar tanto la capacidad predictiva inicial del algoritmo como el impacto del desbalance de clases sobre su rendimiento.

Se evaluaron tres configuraciones:

- Sin técnicas de balanceo
- Con conjunto de entrenamiento balanceado mediante SMOTE
- Con conjunto de entrenamiento balanceado mediante SMOTETomek.

Random Forest base sin técnicas de balanceo

En una primera etapa, se entrenó un modelo Random Forest sobre el conjunto de entrenamiento original, altamente desbalanceado. Los resultados evidencian una alta exactitud global, coherente con la predominancia de empresas no quebradas, pero un bajo recall y F1-score para la clase minoritaria, reflejando la limitada capacidad del modelo para identificar quiebras bajo estas condiciones.

Métrica	Valor
Accuracy	0,97
Precision (quiebra)	0,70
Recall (quiebra)	0,16
F1-score (quiebra)	0,26
AUC	0,93

Tabla 11. Desempeño del Random Forest base sin balanceo de clases.

Fuente: Elaboración propia.

A pesar de ello, el valor elevado del AUC-ROC indica que el modelo logra ordenar adecuadamente las observaciones según su riesgo relativo, aun cuando el umbral de clasificación favorece fuertemente a la clase mayoritaria.

Random Forest base entrenado con SMOTE

Posteriormente, se entrenó el mismo modelo utilizando el conjunto de entrenamiento balanceado mediante SMOTE. En esta configuración se observa una mejora sustancial en la capacidad de detección de quiebras, reflejada en un aumento significativo del recall y del F1-score.

Métrica	Valor
Accuracy	0,95
Precision (quiebra)	0,36
Recall (quiebra)	0,50

F1-score (quiebra)	0,41
AUC	0,94

Tabla 12. Desempeño del Random Forest base entrenado con SMOTE.

Fuente: Elaboración propia.

Este resultado confirma que el remuestreo sintético permite al modelo aprender patrones asociados a la clase minoritaria, reduciendo el sesgo hacia la clase mayoritaria. Como contrapartida, se registra una disminución moderada de la exactitud global, consistente con el aumento de falsos positivos, fenómeno esperable en contextos de fuerte desbalance.

Random Forest base entrenado con SMOTETomek

Finalmente, se entrenó el modelo utilizando el conjunto de entrenamiento procesado mediante SMOTETomek, combinando generación sintética de observaciones y limpieza de la frontera de decisión. Esta variante muestra un ligero incremento adicional en recall y F1-score, junto con un desempeño AUC estable, sugiriendo una mejora marginal en el equilibrio entre sensibilidad y precisión.

Métrica	Valor
Accuracy	0,96
Precision (quiebra)	0,37
Recall (quiebra)	0,52
F1-score (quiebra)	0,43
AUC	0,94

Tabla 13. Desempeño del Random Forest base entrenado con SMOTETomek.

Fuente: Elaboración propia.

Este comportamiento es coherente con el objetivo de SMOTETomek, que no solo balancea las clases, sino que también reduce la presencia de observaciones ambiguas cercanas a la frontera entre clases.

Comparación preliminar de configuraciones

La comparación de los tres enfoques evidencia que el Random Forest base entrenado sin técnicas de balanceo resulta inadecuado para un problema de predicción de quiebra, debido a su baja capacidad para identificar empresas en riesgo. En contraste, las versiones entrenadas con

SMOTE y SMOTETomek presentan mejoras sustantivas en métricas orientadas a la clase minoritaria, siendo SMOTETomek la que ofrece el mejor desempeño preliminar global.

Estos resultados justifican la incorporación de técnicas de remuestreo en las etapas posteriores de optimización de hiperparámetros.

Métrica	RF Base	RF SMOTE	RF ST
Accuracy	0,97	0,95	0,96
Precision (quiebra)	0,70	0,36	0,37
Recall (quiebra)	0,16	0,50	0,52
F1-score (quiebra)	0,26	0,41	0,43
AUC	0,93	0,94	0,94

Tabla 14. Comparación del desempeño del Random Forest base, SMOTE y SMOTETomek.

Fuente: Elaboración propia.

Optimización del modelo Random Forest

Con el objetivo de maximizar el desempeño predictivo del modelo, se procedió a optimizar los hiperparámetros del Random Forest. Dado el mejor comportamiento observado en la etapa preliminar, se seleccionó como base de optimización la configuración que incorpora SMOTETomek dentro del pipeline, permitiendo abordar simultáneamente el desbalance de clases y el entrenamiento del clasificador.

Definición del espacio de búsqueda

Se definió un conjunto acotado de hiperparámetros relevantes para el algoritmo Random Forest, incluyendo tanto parámetros estructurales del bosque como restricciones orientadas a controlar la complejidad del modelo. Entre ellos se consideraron:

- número de árboles (*n_estimators*),
- profundidad máxima de los árboles (*max_depth*),
- tamaño mínimo de muestra requerido para dividir un nodo (*min_samples_split*),
- tamaño mínimo de muestra por hoja (*min_samples_leaf*),
- número de variables evaluadas en cada partición (*max_features*).

Este espacio de búsqueda fue incorporado en un pipeline que integra las etapas de remuestreo mediante SMOTETomek y clasificación mediante Random Forest, garantizando que cada combinación de hiperparámetros fuera evaluada bajo validación cruzada sin fuga de información.

Estrategia de optimización

La búsqueda de hiperparámetros se realizó utilizando Randomized Search con validación cruzada, técnica que permite explorar eficientemente espacios de alta dimensionalidad mediante muestreo aleatorio de combinaciones de parámetros. Se utilizó como métrica central de optimización el AUC-ROC, debido a su robustez frente al desbalance de clases y su capacidad para evaluar el poder discriminante del modelo de manera independiente del umbral de clasificación.

Cada iteración del proceso consistió en:

1. aplicación de SMOTETomek sobre el subconjunto de entrenamiento,
2. entrenamiento del Random Forest con una combinación específica de hiperparámetros,
3. evaluación del desempeño mediante validación cruzada.

De este modo, el criterio de selección del modelo final se basó en el promedio del AUC-ROC obtenido a lo largo de los folds.

Selección del mejor modelo

Como resultado del proceso de optimización, se obtuvo una combinación de hiperparámetros que maximizó el desempeño medio en validación cruzada. La mejor configuración encontrada fue:

Hiperparámetro	Valor óptimo
Número de árboles (n_estimators)	100
Profundidad máxima (max_depth)	10
Mín. muestras para dividir nodo (min_samples_split)	2
Mín. muestras por hoja (min_samples_leaf)	4
Variables consideradas por split (max_features)	sqrt
AUC promedio (CV)	0,94

Tabla 15. Mejores hiperparámetros del modelo Random Forest optimizado y desempeño en validación cruzada.

Fuente: Elaboración propia.

Esta configuración alcanzó un AUC-ROC promedio cercano a 0,94, confirmando la capacidad del modelo optimizado para discriminar entre empresas quebradas y no quebradas de manera robusta.

Consideraciones metodológicas

La optimización se realizó exclusivamente sobre el conjunto de entrenamiento mediante validación cruzada, manteniendo el conjunto de prueba completamente aislado para su utilización en la evaluación final. Esta estrategia permite evitar sobreajuste por selección de modelo y garantiza que el desempeño reportado posteriormente corresponda a una estimación honesta de capacidad predictiva fuera de muestra.

El modelo resultante será utilizado en el apartado siguiente para efectuar la evaluación final sobre el conjunto de prueba y su posterior comparación empírica con el modelo Z''-score de Altman.

Evaluación final del modelo Random Forest

Una vez seleccionado el modelo Random Forest optimizado, se procedió a evaluar su desempeño final utilizando exclusivamente el conjunto de prueba, el cual no fue utilizado en ninguna etapa de entrenamiento, balanceo ni optimización. Esta evaluación permite estimar de manera objetiva la capacidad de generalización del modelo.

En primer lugar, se calcularon las predicciones de clase y las probabilidades asociadas a la quiebra para cada empresa del conjunto de prueba.

Distribución de predicciones

El modelo clasificó un total de 104 empresas como en quiebra dentro del conjunto de prueba, reflejando un comportamiento significativamente distinto al clasificador trivial y a los modelos no balanceados.

Clasificación predicha RF	Frecuencia
No quiebra (0)	1.260
Quiebra (1)	104

Total **1.364**

Tabla 16. Distribución de clases predichas por el modelo Random Forest en el conjunto de prueba.
Fuente: *Elaboración propia.*

Matriz de confusión

La matriz de confusión evidencia el comportamiento detallado del modelo frente a ambos tipos de error. En particular, se observa una alta capacidad para identificar empresas sanas y una detección efectiva de una proporción relevante de empresas en quiebra, aunque acompañada de falsos positivos, fenómeno esperable en un contexto de fuerte desbalance.

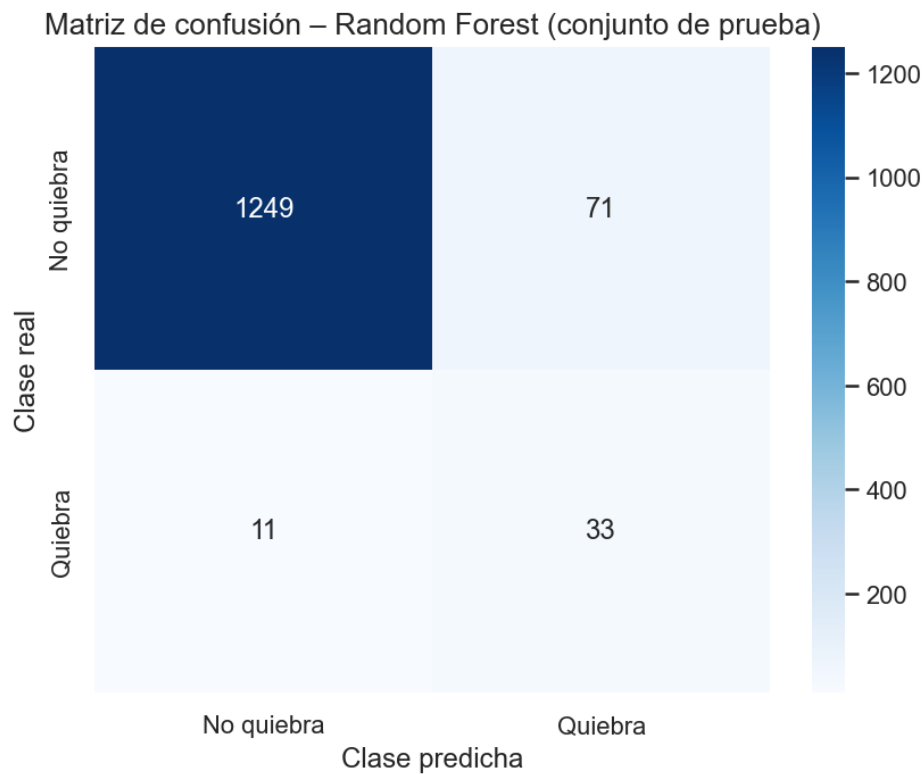


Figura 12. Matriz de confusión del modelo Random Forest optimizado en el conjunto de prueba.
Fuente: *Elaboración propia.*

Métricas de desempeño

Clase	Precision	Recall	F1-score	Soporte
0 (No quiebra)	0,99	0,95	0,97	1.320
1 (Quiebra)	0,32	0,75	0,45	44
Accuracy			0,94	1.364
Macro avg	0,65	0,85	0,71	1.364
Weighted avg	0,97	0,94	0,95	1.364

Tabla 17. Métricas de desempeño del modelo Random Forest optimizado en el conjunto de prueba.

Fuente: Elaboración propia.

El desempeño del modelo en el conjunto de prueba se resume en las siguientes métricas:

- Accuracy $\approx 0,94$
- Precision (quiebra) $\approx 0,32$
- Recall (quiebra) $\approx 0,75$
- F1-score (quiebra) $\approx 0,45$
- AUC-ROC $\approx 0,95$

Estos resultados indican que el modelo logra identificar aproximadamente tres de cada cuatro empresas en quiebra, manteniendo al mismo tiempo una elevada capacidad discriminante global, medida a través del AUC.

Curva ROC y capacidad discriminante

La curva ROC obtenida sobre el conjunto de prueba muestra un comportamiento claramente superior al azar, con un área bajo la curva cercana a 0,95, lo que confirma la alta capacidad del modelo para ordenar empresas según su riesgo de quiebra, independientemente del umbral de clasificación utilizado.

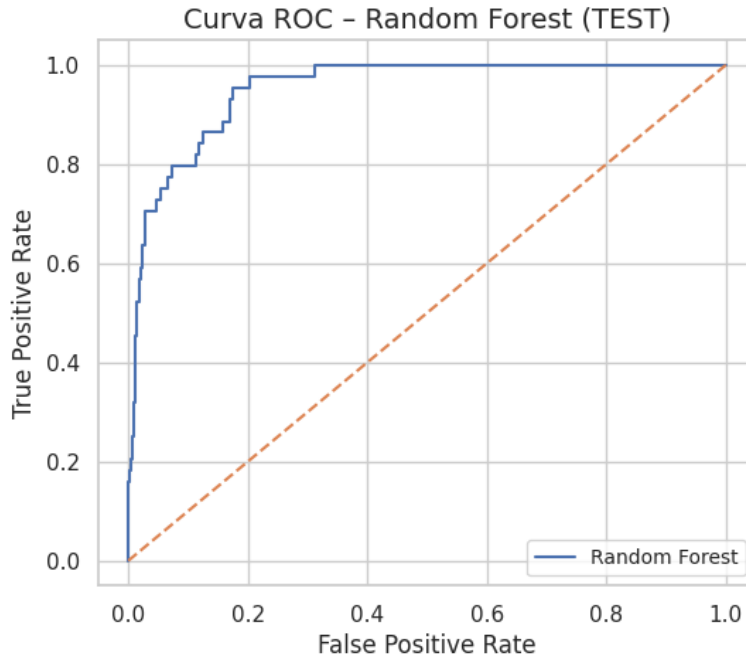


Figura X. Curva ROC del modelo Random Forest optimizado en el conjunto de prueba.
Fuente: Elaboración propia.

Análisis de resultados

Desde una perspectiva aplicada, el modelo presenta un alto nivel de sensibilidad frente a quiebras, lo que resulta particularmente relevante en contextos financieros, donde el costo de no detectar una empresa en riesgo suele ser superior al de emitir una alerta incorrecta. Si bien la precisión asociada a la clase minoritaria es moderada, esta se interpreta como el resultado natural del compromiso entre detección temprana y control de falsos positivos en escenarios altamente desbalanceados.

En conjunto, los resultados confirman que el modelo Random Forest optimizado logra un equilibrio adecuado entre capacidad predictiva y sensibilidad al riesgo, constituyendo una base sólida para su comparación empírica con el modelo Z'' -score de Altman.

Comparación empírica entre el modelo Z'' -score de Altman y el modelo Random Forest

Con el fin de evaluar empíricamente el desempeño del enfoque clásico de predicción financiera frente a un modelo de machine learning, se compararon los resultados obtenidos por

el modelo Z''-score de Altman y el Random Forest optimizado, utilizando el mismo conjunto de prueba.

En primer lugar, se calcularon las predicciones del Z''-score sobre el conjunto de prueba, clasificando las empresas según el umbral óptimo previamente determinado. Este procedimiento identificó un total de 315 empresas como en quiebra, número considerablemente superior al estimado por el modelo Random Forest.

Predicción Altman	Frecuencia
No quiebra (0)	1.049
Quiebra (1)	315
Total	1.364

Tabla 18. Distribución de clases predichas por el modelo Z''-score de Altman en el conjunto de prueba.

Fuente: Elaboración propia.

Comparación de métricas de desempeño

El desempeño de ambos modelos se resume en la Tabla X, donde se presentan métricas globales y orientadas a la clase minoritaria.

Métrica	Altman Z''-score	Random Forest
Accuracy	0,79	0,94
Precision (quiebra)	0,12	0,32
Recall (quiebra)	0,89	0,75
F1-score (quiebra)	0,22	0,45
AUC	0,90	0,95

Tabla X. Comparación de métricas de desempeño: Altman vs Random Forest.

Fuente: Elaboración propia.

Los resultados muestran que el modelo de Altman alcanza un recall considerablemente mayor, identificando cerca del 89 % de las quiebras. Sin embargo, esta elevada sensibilidad se obtiene a costa de una muy baja precisión, reflejando un alto número de falsos positivos.

Por su parte, el modelo Random Forest logra un AUC superior y un F1-score aproximadamente doble, indicando una mejor capacidad de discriminación global y un equilibrio más eficiente entre detección de quiebras y control de falsas alarmas.

Análisis de matrices de confusión

El análisis conjunto de las matrices de confusión permite profundizar en la naturaleza de los errores cometidos por cada modelo.

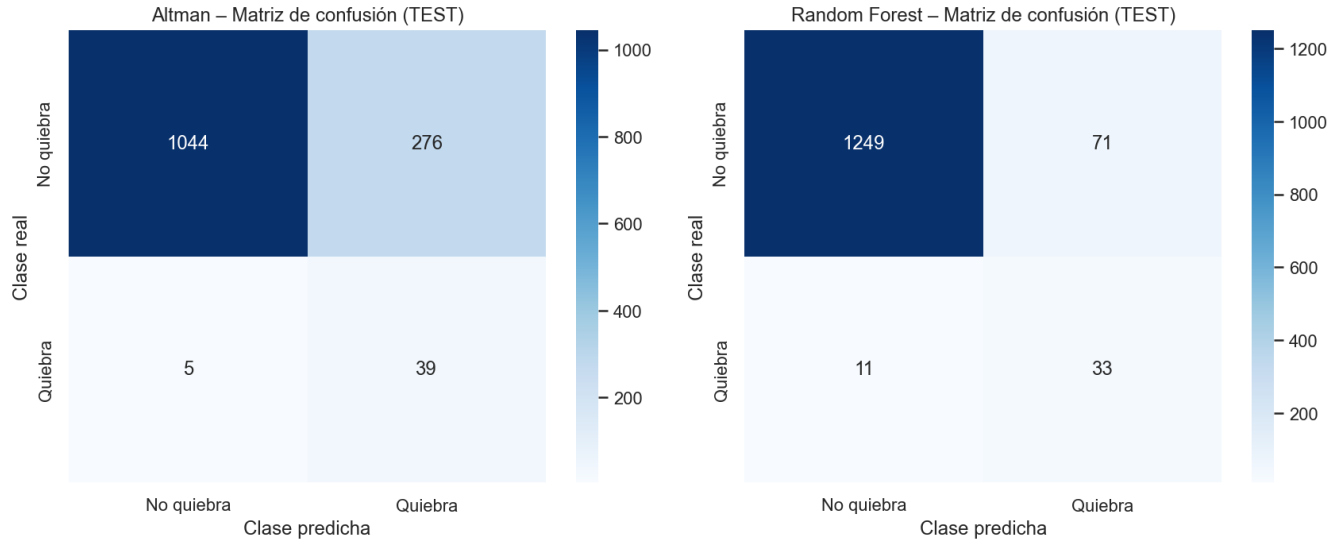


Figura 13. Matrices de confusión comparativas en el conjunto de prueba: Altman vs Random Forest.

Fuente: Elaboración propia.

Se observa que el modelo de Altman clasifica un volumen considerable de empresas sanas como en quiebra, priorizando la minimización de falsos negativos. En contraste, el modelo Random Forest reduce significativamente los falsos positivos, aunque permite escapar una mayor proporción de quiebras reales.

Este comportamiento refleja enfoques conceptuales distintos: un modelo financiero diseñado para alerta temprana frente a un modelo algorítmico orientado a optimizar la eficiencia predictiva.

Comparación de curvas ROC

Finalmente, se compararon las curvas ROC de note ambos modelos sobre el conjunto de prueba.

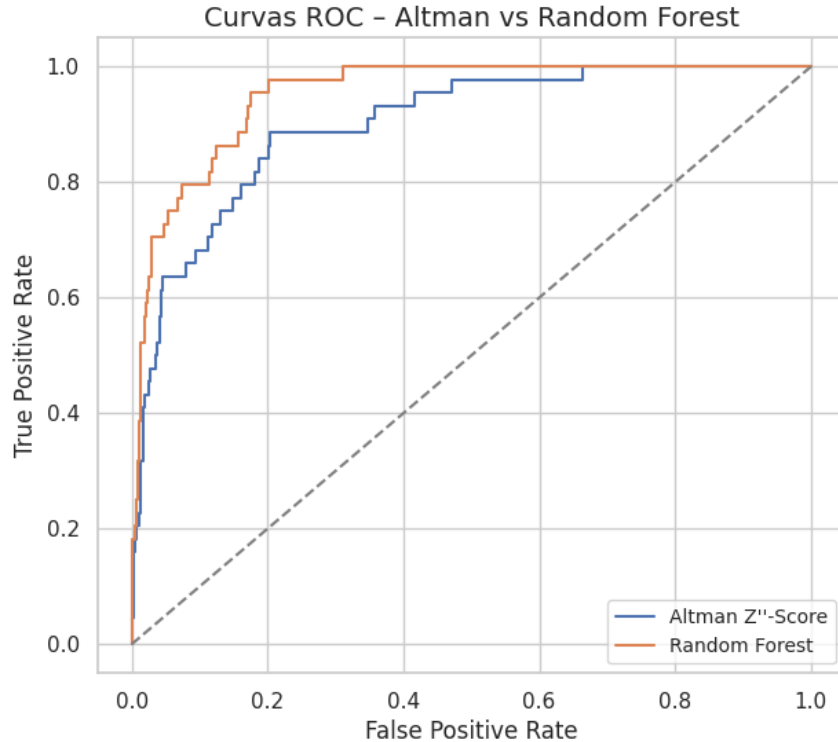


Figura 14. Curvas ROC comparativas: Altman Z''-score vs Random Forest.
Fuente: Elaboración propia.

El Random Forest presenta una curva consistentemente superior, con un AUC mayor, confirmando su mayor capacidad para discriminar empresas según su riesgo de quiebra a lo largo de todos los umbrales de decisión.

Discusión comparativa

Los resultados evidencian que el modelo Z''-score de Altman y el modelo Random Forest no solo difieren en desempeño cuantitativo, sino también en su lógica de clasificación. Mientras Altman privilegia la detección de prácticamente todas las quiebras potenciales, aun a costa de una alta tasa de falsas alarmas, el modelo Random Forest logra un mejor compromiso entre sensibilidad, precisión y poder discriminante.

En este sentido, el Z''-score se perfila como una herramienta eficaz para procesos de filtrado inicial o monitoreo de riesgo, mientras que el modelo Random Forest ofrece ventajas relevantes para escenarios donde se requiere una evaluación más fina y económicamente eficiente del riesgo de quiebra.

Interpretabilidad del Random Forest

Con el objetivo de analizar los patrones financieros aprendidos por el modelo Random Forest y evaluar su coherencia con la teoría económica, se aplicaron distintas técnicas de interpretabilidad. Estas incluyeron medidas de importancia estructural, análisis por permutación y métodos basados en valores SHAP, abordando tanto interpretaciones globales como locales.

Importancia de variables en el Random Forest

En una primera etapa, se extrajeron las importancias internas del modelo, asociadas a la contribución promedio de cada variable en la construcción de los árboles de decisión.

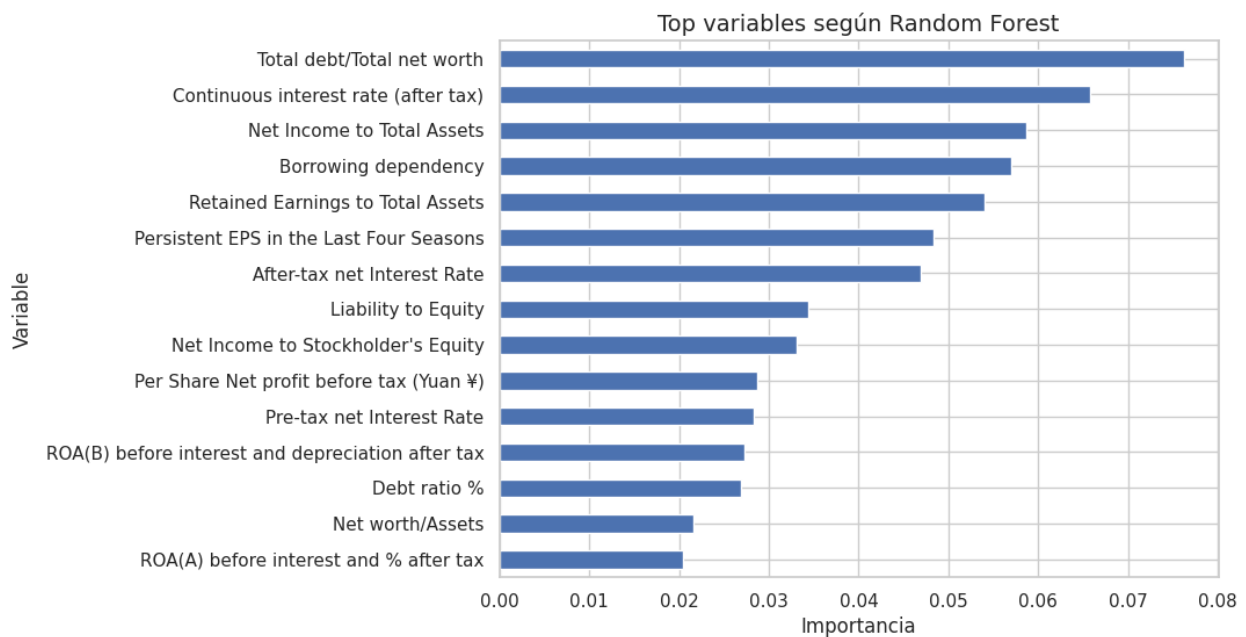


Figura 15. Variables más relevantes según el modelo Random Forest optimizado.

Fuente: Elaboración propia.

Las variables más influyentes corresponden principalmente a indicadores de apalancamiento financiero, rentabilidad y dependencia del financiamiento externo, destacando ratios como deuda sobre patrimonio, ingresos netos sobre activos, dependencia de préstamos y utilidades retenidas.

Este resultado sugiere que el modelo basa sus predicciones en dimensiones financieras fundamentales, coherentes con la literatura de predicción de quiebra.

Importancia acumulada y concentración explicativa

Para evaluar la concentración de la información, se analizó la importancia acumulada de las variables.

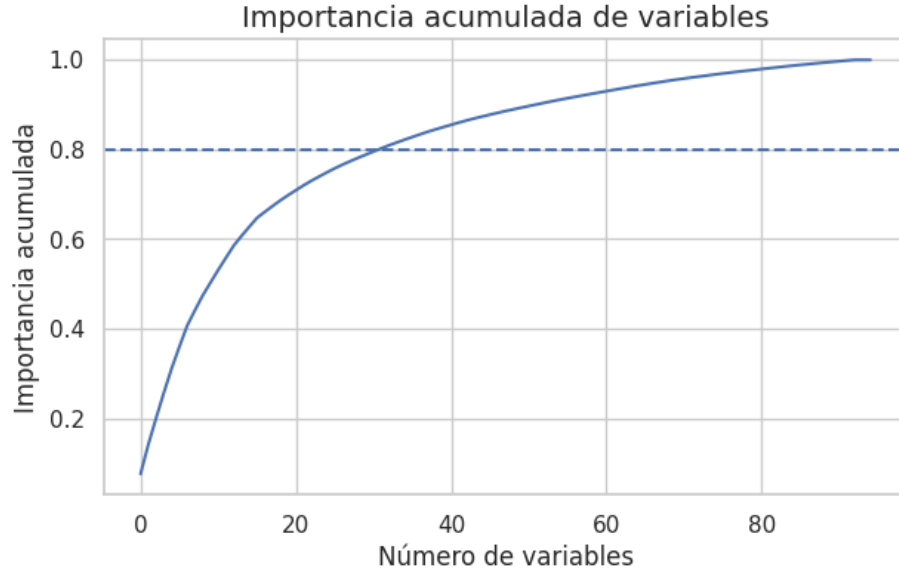


Figura 16. Importancia acumulada de las variables del modelo Random Forest.

Fuente: Elaboración propia.

Se observa que aproximadamente 37 variables concentran cerca del 80 % de la importancia total, evidenciando la existencia de un núcleo reducido de factores financieros dominantes en el proceso de clasificación.

Ranking	Variable	Importancia acumulada
1	Total debt/Total net worth	0,0762
2	Continuous interest rate (after tax)	0,1420
3	Net Income to Total Assets	0,2007
4	Borrowing dependency	0,2576
5	Retained Earnings to Total Assets	0,3116
6	Persistent EPS in the Last Four Seasons	0,3599
7	After-tax net Interest Rate	0,4069
8	Liability to Equity	0,4412
9	Net Income to Stockholder's Equity	0,4743
10	Per Share Net profit before tax (Yuan ¥)	0,503
11	Pre-tax net Interest Rate	0,5314
12	ROA(B) before interest and depreciation after tax	0,5586
13	Debt ratio %	0,5856
14	Net worth/Assets	0,6072

15	ROA(A) before interest and % after tax	0,6276
16	Net profit before tax/Paid-in capital	0,6475
17	Net Value Growth Rate	0,6609
18	Current Liability to Equity	0,6738
19	Current Liabilities/Equity	0,6864
20	Total income/Total expense	0,6981
21	Quick Ratio	0,7091
22	ROA(C) before interest and depreciation before interest	0,7201
23	Cash Turnover Rate	0,7304
24	Interest-bearing debt interest rate	0,7397
25	Operating Profit Rate	0,749
26	Interest Expense Ratio	0,7579
27	Equity to Liability	0,7662
28	Degree of Financial Leverage (DFL)	0,7739
29	Interest Coverage Ratio (Interest expense to EBIT)	0,7813
30	Current Liability to Current Assets	0,7885
31	Current Liability to Assets	0,7955
32	Allocation rate per person	0,8026

Tabla 20. Ranking de variables que concentran 80% del poder explicativo.

Fuente: Elaboración propia.

Importancia por permutación

Adicionalmente, se aplicó permutation importance, técnica que evalúa el deterioro del desempeño del modelo ante la alteración aleatoria de cada variable.

Ranking	Variable	Importancia por permutación
1	Degree of Financial Leverage (DFL)	0,0037
2	Total Asset Growth Rate	0,0028
3	Equity to Liability	0,0022
4	Long-term fund suitability ratio (A)	0,0021
5	Per Share Net profit before tax (Yuan ¥)	0,0018
6	Current Liability to Liability	0,0016
7	Working Capital to Total Assets	0,0015
8	Pre-tax net Interest Rate	0,0015
9	ROA(A) before interest and % after tax	0,0015
10	Interest Coverage Ratio (Interest expense to EBIT)	0,0015
11	Net worth/Assets	0,0013
12	Fixed Assets Turnover Frequency	0,0013
13	Research and development expense rate	0,0013
14	Current Liabilities/Liability	0,0013
15	Net Value Growth Rate	0,0012

Tabla 21. Variables más relevantes según importancia por permutación.

Fuente: *Elaboración propia.*

Los resultados confirman la relevancia de los mismos grupos de indicadores financieros, reforzando la robustez de las conclusiones obtenidas a partir de las importancias internas del modelo.

Interpretabilidad global mediante valores SHAP

Con el fin de analizar no solo la relevancia, sino también el sentido del impacto de cada variable, se calcularon valores SHAP.

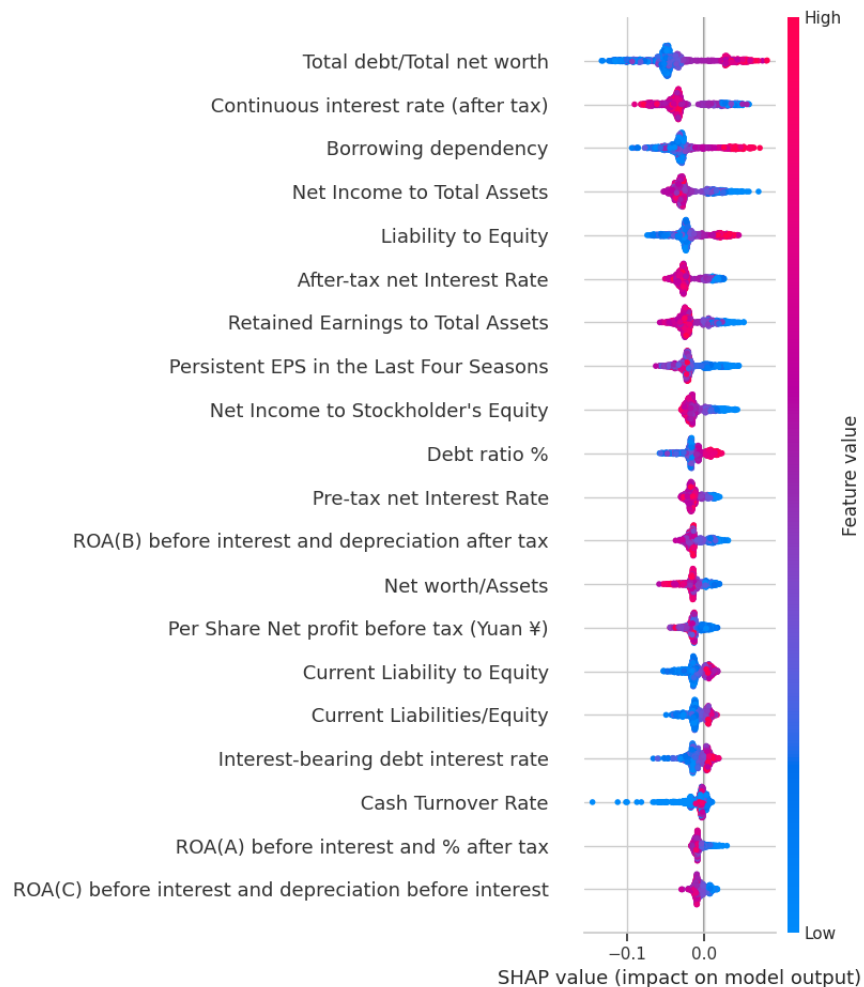


Figura 17. Gráfico resumen de valores SHAP para el modelo Random Forest.

Fuente: *Elaboración propia.*

Este gráfico muestra cómo valores altos o bajos de cada ratio financiero contribuyen positiva o negativamente a la probabilidad de quiebra. Se observa, por ejemplo, que niveles

elevados de endeudamiento y dependencia crediticia tienden a incrementar el riesgo estimado, mientras que mayores niveles de rentabilidad y acumulación de utilidades lo reducen.

Estos patrones refuerzan la coherencia económica del modelo.

Interpretabilidad local: explicación de una predicción individual

Finalmente, se analizó un caso individual mediante un gráfico SHAP local, descomponiendo la predicción en los aportes específicos de cada variable.

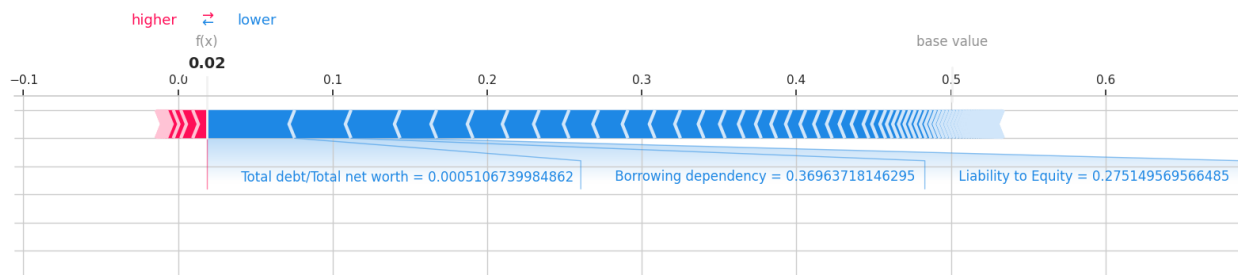


Figura 18. Explicación individual de una predicción del modelo Random Forest mediante valores SHAP.

Fuente: Elaboración propia.

Este análisis permite visualizar cómo distintos indicadores financieros empujan la predicción hacia quiebra o no quiebra, ilustrando el proceso interno del modelo y habilitando una interpretación financiera directa de decisiones algorítmicas.

Síntesis interpretativa

En conjunto, los resultados muestran que el modelo Random Forest no actúa como una caja negra, sino que fundamenta sus predicciones principalmente en variables asociadas a estructura de capital, rentabilidad, presión financiera y sostenibilidad de resultados, dimensiones centrales en la literatura de quiebras empresariales.

Este hallazgo resulta clave para establecer el vínculo conceptual entre el enfoque de machine learning y los modelos financieros tradicionales, aspecto que se desarrolla en el apartado siguiente.

Análisis financiero agregado y vinculación teórica con el modelo de Altman

Reducción del espacio de variables según importancia predictiva

Con el objetivo de interpretar financieramente el modelo Random Forest optimizado, se extrajeron las importancias de las variables estimadas por el algoritmo. A partir de estas, se ordenaron las características de mayor a menor contribución al desempeño del modelo y se calculó su importancia acumulada.

Como criterio de reducción, se seleccionaron aquellas variables que, en conjunto, explican aproximadamente el 80% de la importancia total, lo que permitió reducir el conjunto original a un subconjunto de 31 variables financieramente relevantes, manteniendo la mayor parte del poder explicativo del modelo.

Agrupación conceptual en dimensiones financieras

Las variables seleccionadas fueron clasificadas manualmente en cinco dimensiones financieras fundamentales, de acuerdo con su naturaleza económica y con la literatura clásica de análisis financiero y predicción de quiebras:

- Rentabilidad
- Solvencia / Endeudamiento
- Liquidez
- Cobertura financiera / Costo de la deuda
- Crecimiento / Valor

Esta clasificación se fundamenta en que la literatura especializada identifica estas dimensiones como los principales determinantes del fracaso empresarial (Beaver, 1966; Altman, 1968; Ohlson, 1980; Damodaran, 2015), ya que representan de manera complementaria la capacidad de la empresa para generar resultados, sostener su estructura financiera y cumplir sus compromisos en el corto y largo plazo.

Ranking	Variable	Importancia	Importancia acumulada	Dimensión financiera
1	Total debt/Total net worth	0,0762	0,0762	Solvencia / Endeudamiento

2	Continuous interest rate (after tax)	0,0658	0,1420	Rentabilidad
3	Net Income to Total Assets	0,0587	0,2007	Rentabilidad
4	Borrowing dependency	0,0569	0,2576	Solvencia / Endeudamiento
5	Retained Earnings to Total Assets	0,0540	0,3116	Rentabilidad
6	Persistent EPS in the Last Four Seasons	0,0484	0,3599	Rentabilidad
7	After-tax net Interest Rate	0,0470	0,4069	Rentabilidad
8	Liability to Equity	0,0343	0,4412	Solvencia / Endeudamiento
9	Net Income to Stockholder's Equity	0,0331	0,4743	Rentabilidad
10	Per Share Net profit before tax (Yuan ¥)	0,0287	0,5030	Rentabilidad
11	Pre-tax net Interest Rate	0,0284	0,5314	Rentabilidad
12	ROA(B) before interest and depreciation after tax	0,0273	0,5586	Rentabilidad
13	Debt ratio %	0,0269	0,5856	Solvencia / Endeudamiento
14	Net worth/Assets	0,0216	0,6072	Solvencia / Endeudamiento
15	ROA(A) before interest and % after tax	0,0204	0,6276	Rentabilidad
16	Net profit before tax/Paid-in capital	0,0199	0,6475	Rentabilidad
17	Net Value Growth Rate	0,0134	0,6609	Crecimiento / Valor
18	Current Liability to Equity	0,0129	0,6738	Solvencia / Endeudamiento
19	Current Liabilities/Equity	0,0126	0,6864	Solvencia / Endeudamiento
20	Total income/Total expense	0,0117	0,6981	Rentabilidad
21	Quick Ratio	0,0110	0,7091	Liquidez
22	ROA(C) before interest and depreciation before interest	0,0110	0,7201	Rentabilidad
23	Cash Turnover Rate	0,0103	0,7304	Liquidez
24	Interest-bearing debt interest rate	0,0093	0,7397	Cobertura financiera
25	Operating Profit Rate	0,0093	0,7490	Rentabilidad
26	Interest Expense Ratio	0,0090	0,7579	Cobertura financiera

27	Equity to Liability	0,0082	0,7662	Solvencia / Endeudamiento
28	Degree of Financial Leverage (DFL)	0,0078	0,7739	Solvencia / Endeudamiento
29	Interest Coverage Ratio (Interest expense to EBIT)	0,0074	0,7813	Cobertura financiera
30	Current Liability to Current Assets	0,0071	0,7885	Liquidez
31	Current Liability to Assets	0,0071	0,7955	Solvencia / Endeudamiento

Tabla 22. Categorización de las variables en dimensiones financieras generales.

Fuente: Elaboración propia.

Importancia agregada por dimensión financiera

Una vez clasificadas las variables, se sumaron sus importancias individuales para obtener la contribución total de cada dimensión financiera al poder predictivo del modelo.

Los resultados muestran una marcada concentración explicativa en dos dimensiones principales:

- Rentabilidad: 46,3%
- Solvencia / Endeudamiento: 26,5%

Mientras que las restantes dimensiones presentan una contribución considerablemente menor:

- Cobertura financiera: 2,6%
- Liquidez: 2,1%
- Crecimiento / Valor: 1,3%

Dimensión financiera	Importancia agregada	Porcentaje del total
Rentabilidad	0,46	58,78%
Solvencia / Endeudamiento	0,26	33,56%
Cobertura financiera	0,03	3,26%
Liquidez	0,02	2,70%
Crecimiento / Valor	0,01	1,70%
Total	0,79	100%

Tabla 23. Importancia y % del total agrupado por Dimensión Financiera.

Fuente: Elaboración propia.

Este resultado indica que el modelo Random Forest identifica a la capacidad de generación de resultados y a la estructura de capital como los principales factores asociados a la quiebra empresarial, lo cual es consistente con la teoría financiera y los modelos clásicos de predicción de insolvencia.

Vinculación empírica con el modelo Z'' de Altman

Con el fin de evaluar en qué medida el modelo de aprendizaje automático replica o amplía el enfoque clásico, las dimensiones financieras fueron asociadas a los componentes del modelo Z'' de Altman de la siguiente manera:

- Liquidez → X1 (Working Capital / Total Assets)
- Rentabilidad → X2 y X3 (Retained Earnings / Total Assets; EBIT / Total Assets)
- Solvencia / Endeudamiento → X4 (Equity / Total Liabilities)
- Cobertura financiera → No explícito en Altman
- Crecimiento / Valor → No explícito en Altman

Posteriormente, se agregaron las importancias según su relación con el modelo de Altman, obteniendo que aproximadamente:

- 58,8% del poder explicativo del modelo se asocia a dimensiones vinculadas a X2 y X3 (rentabilidad),
- 33,6% se relaciona con X4 (estructura de capital),
- 2,7% con X1 (liquidez),
- y cerca de 5% corresponde a dimensiones no explícitamente consideradas en Altman.

Relación con Altman	Importancia agregada	Porcentaje del total
X2 – X3 (Rentabilidad)	0,46	58,78%
X4 (Solvencia / Endeudamiento)	0,26	33,56%
No explícito en Altman	0,04	4,96%
X1 (Liquidez)	0,02	2,70%
Total	0,79	100,00%

Tabla 24. Importancia en % del total agrupada por relación con Altman.

Fuente: *Elaboración propia.*

Distribución del poder explicativo según dimensiones de Altman

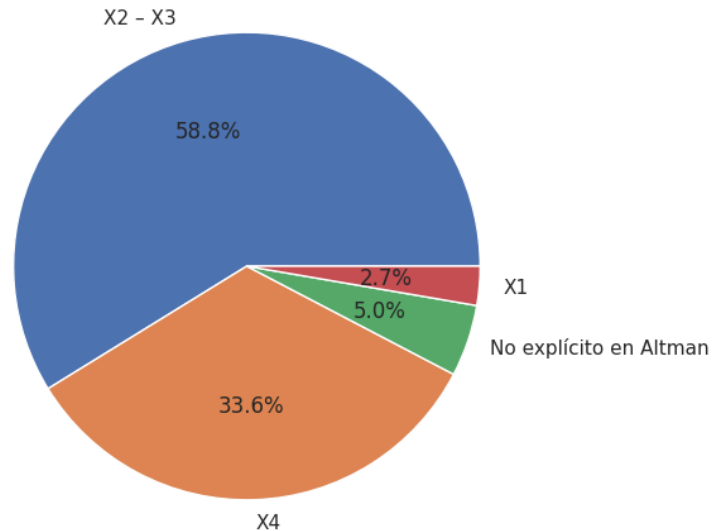


Figura 19. Distribución del poder explicativo según dimensiones de Altman.

Fuente: *Elaboración propia.*

Este hallazgo sugiere que, aun cuando el modelo Random Forest no incorpora supuestos financieros explícitos, tiende endógenamente a reorganizar la información en torno a las mismas dimensiones fundamentales propuestas por Altman, validando empíricamente la estructura conceptual del modelo Z''.

Interpretación financiera de los resultados

Los resultados obtenidos permiten concluir que el fenómeno de quiebra empresarial es explicado mayoritariamente por:

- Deterioro de la rentabilidad, reflejando incapacidad sostenida para generar resultados sobre los activos y el patrimonio.
- Fragilidad en la estructura financiera, caracterizada por altos niveles de endeudamiento relativo y dependencia de financiamiento externo.

Las dimensiones de liquidez y cobertura financiera, si bien relevantes, presentan un rol secundario, actuando principalmente como mecanismos de transmisión de un deterioro estructural previo más que como causas primarias.

Asimismo, la emergencia de variables asociadas a crecimiento y valor, ausentes en la formulación original de Altman, indica que los modelos de aprendizaje automático permiten ampliar el marco tradicional, incorporando elementos dinámicos que complementan el enfoque contable clásico.

En conjunto, este análisis sugiere que los modelos de aprendizaje automático no sustituyen el enfoque financiero tradicional, sino que lo validan empíricamente, lo refinan y lo extienden, proporcionando una representación más flexible de las múltiples dimensiones del riesgo de quiebra.

Resultado final de la investigación

El objetivo general de esta investigación fue evaluar empíricamente el desempeño del modelo Z''-Score de Altman frente a un modelo de aprendizaje automático (Random Forest) en la predicción de quiebra empresarial, utilizando información financiera contable y un enfoque metodológico reproducible.

A partir del desarrollo del estudio, se construyó un flujo completo de análisis que incluyó: exploración de datos, preprocesamiento, tratamiento del desbalance de clases, entrenamiento de modelos, optimización de hiperparámetros, evaluación fuera de muestra, comparación de modelos e interpretación financiera.

El resultado final de la investigación puede sintetizarse en tres hallazgos centrales:

1. El modelo Random Forest optimizado con SMOTE-Tomek supera sistemáticamente al modelo Z''-Score de Altman en capacidad discriminatoria global, obteniendo un AUC de 0,95 frente a 0,90 del modelo clásico, junto con una mejora sustantiva en métricas compuestas como F1-score.
2. El modelo de aprendizaje automático logra un mejor equilibrio entre detección de quiebras (recall) y reducción de falsas alarmas, mostrando un desempeño más adecuado en un contexto de clases altamente desbalanceadas.
3. Desde una perspectiva interpretativa, el modelo Random Forest reconstruye endógenamente las dimensiones financieras fundamentales del análisis clásico (rentabilidad, solvencia, liquidez), validando empíricamente la estructura conceptual del modelo de Altman, al tiempo que incorpora nuevas dimensiones no explícitas, como crecimiento y cobertura financiera.

En consecuencia, el resultado final del trabajo no se limita a demostrar una superioridad técnica, sino que propone un marco integrador, donde los modelos de aprendizaje automático extienden, refinan y robustecen la lógica financiera tradicional.

Conclusiones y recomendaciones

Conclusiones

Conclusiones metodológicas

Los resultados confirman que, en contextos de quiebra empresarial caracterizados por alta no linealidad y fuerte desbalance de clases, los modelos de aprendizaje automático basados en ensamblados ofrecen ventajas significativas frente a los enfoques estadísticos clásicos.

La incorporación de técnicas de remuestreo (SMOTE y SMOTE-Tomek) permitió mejorar sustantivamente la capacidad del modelo para aprender patrones asociados a la clase minoritaria, lo cual se reflejó en incrementos relevantes en recall, F1-score y AUC. Esto demuestra que, en problemas financieros reales, el tratamiento del desbalance es tan relevante como la elección del algoritmo.

Asimismo, la utilización de validación cruzada, optimización de hiperparámetros y evaluación fuera de muestra permitió construir un modelo robusto, reduciendo riesgos de sobreajuste y fortaleciendo la validez de los resultados.

Conclusiones empíricas

Desde el punto de vista empírico, el modelo Random Forest optimizado logró:

- Mantener una alta capacidad discriminadora global.
- Incrementar significativamente la detección efectiva de empresas en quiebra.
- Reducir la dependencia de un umbral fijo, permitiendo análisis probabilísticos más flexibles.

La comparación directa mostró que el modelo de Altman, si bien conserva una notable vigencia, presenta limitaciones relevantes en contextos actuales, particularmente en presencia de estructuras financieras complejas y distribuciones altamente asimétricas.

No obstante, se concluye que el Z''-Score no pierde valor como herramienta de screening inicial, sino que resulta insuficiente como sistema único de alerta temprana.

Conclusiones financieras y teóricas

Uno de los hallazgos más relevantes del estudio es que el modelo de aprendizaje automático, aun sin imponer estructura financiera alguna, identifica como principales determinantes de la quiebra:

- Rentabilidad sostenida
- Solvencia y estructura de capital

Estas dimensiones concentran más del 70% del poder explicativo del modelo, lo que es plenamente consistente con la literatura clásica (Altman, Beaver, Ohlson) y valida empíricamente los fundamentos del análisis financiero tradicional.

Adicionalmente, emergen variables asociadas a crecimiento, valor y cobertura financiera, lo que sugiere que los procesos de quiebra no son únicamente eventos contables estáticos, sino fenómenos dinámicos, vinculados a trayectorias de deterioro financiero.

En este sentido, el trabajo concluye que los modelos de aprendizaje automático no sustituyen el enfoque financiero, sino que permiten operacionalizarlo a gran escala, capturar relaciones no lineales y enriquecer su capacidad explicativa.

Recomendaciones

Recomendaciones para la práctica profesional

A partir de los resultados obtenidos, se proponen las siguientes recomendaciones:

- No utilizar modelos clásicos de forma aislada. El Z-score de Altman debe emplearse como herramienta preliminar, pero complementarse con modelos más flexibles cuando se trate de decisiones críticas (crédito, inversiones, gestión de riesgo).
- Incorporar modelos de machine learning como sistemas de alerta temprana. Modelos como Random Forest permiten construir sistemas de monitoreo continuo, basados en probabilidades de quiebra y no solo en clasificaciones rígidas.

- Priorizar métricas adecuadas al problema financiero. En contextos de quiebra, la accuracy resulta insuficiente. Se recomienda privilegiar métricas como recall, F1-score y AUC, alineadas con los costos reales del error.
- Integrar interpretabilidad al proceso decisional. El uso de técnicas como SHAP y agrupación financiera de variables permite transformar modelos complejos en herramientas útiles para analistas financieros, auditores y gestores de riesgo.

Recomendaciones para futuras investigaciones

Este estudio abre múltiples líneas de investigación futuras:

- Evaluar modelos adicionales (XGBoost, LightGBM, redes neuronales profundas).
- Incorporar variables macroeconómicas y sectoriales.
- Analizar modelos dinámicos, basados en series de tiempo.
- Replicar el estudio en datasets latinoamericanos.
- Construir modelos de predicción temporal de quiebra (horizonte 1–3 años).
- Explorar enfoques de cost-sensitive learning, alineados con pérdidas económicas reales.

Reflexión final

La evidencia obtenida demuestra que la predicción de quiebra empresarial se beneficia sustancialmente de enfoques híbridos, donde la teoría financiera aporta estructura conceptual y los modelos de aprendizaje automático proveen capacidad adaptativa.

El principal aporte de esta investigación radica en mostrar que es posible unir rigurosidad financiera, interpretabilidad y alto desempeño predictivo, construyendo modelos que no solo anticipan eventos de quiebra, sino que contribuyen a comprenderlos.

Bajo esta perspectiva, la quiebra deja de ser un fenómeno meramente contable para convertirse en un proceso multicausal, dinámico y medible, susceptible de ser analizado con herramientas avanzadas al servicio de la toma de decisiones financieras modernas.

Bibliografia

1. **Beaver, W. H.** (1966). Financial ratios as predictors of failure. *Journal of Accounting Research*, 4, 71–111.
2. **Altman, E. I.** (1968). Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. *The Journal of Finance*, 23(4), 589–609.
3. **Altman, E. I.** (1983). *Corporate financial distress: A complete guide to predicting, avoiding, and dealing with bankruptcy*. John Wiley & Sons.
4. **Ohlson, J. A.** (1980). Financial ratios and the probabilistic prediction of bankruptcy. *Journal of Accounting Research*, 18(1), 109–131.
4. 5. **Zmijewski, M. E.** (1984). Methodological issues related to the estimation of financial distress prediction models. *Journal of Accounting Research*, 22, 59–82.
6. **Altman, E. I., Iwanicz-Drozdowska, M., Laitinen, E. K., & Suvas, A.** (2017). Financial distress prediction in an international context: A review and empirical analysis of Altman's Z-score model. *Journal of international financial management & accounting*, 28(2), 131-171.
7. **Dasilas, A., & Rigoni, A.** (2024). Machine learning techniques in bankruptcy prediction: A systematic literature review. *Expert Systems with Applications*, 255, 124761. <https://doi.org/10.1016/j.eswa.2024.124761>.
8. **Insani, F., Pontoh, G. T., & Said, D.** (2024). Corporate bankruptcy prediction model: A systematic literature review. *International Journal of Business, Economics and Management*, 7(3), 143-159. <https://doi.org/10.21744/ijbem.v7n3.2222>.
9. **Rahman, M. J., & Zhu, H.** (2024). Predicting financial distress using ML approaches. *Journal of Contemporary Accounting & Economics*, 20, 100403. <https://doi.org/10.1016/j.jcae.2024.100403>.
10. **Rashid, F., Khan, R. A., & Qureshi, I. H.** (2023). A Comprehensive Review of the Altman Z-Score Model Across Industries. *The Business Review*, 27(2), 35-42.

11. **Shi, Y., & Li, X.** (2019). An overview of bankruptcy prediction models for corporate firms: A systematic literature review. *Intangible Capital*, 15(2), 114–127. <https://doi.org/10.3926/ic.1354>.

12. **Wang, H., & Liu, X.** (2021). Undersampling bankruptcy prediction: Taiwan bankruptcy data. *PLoS ONE*, 16(7): e0254030. <https://doi.org/10.1371/journal.pone.0254030>.

13. **Wang, X., Brorsson, M., & Kräusss, Z.** (2024). *Datasets for Advanced Bankruptcy Prediction: A survey and Taxonomy*. arXiv:2411.01928v1 [cs.CE]. Preprint submitted to Expert System with Applications.

Anexos

Anexo A - Repositorio de código y reproducibilidad computacional

Con el fin de garantizar la transparencia metodológica, la replicabilidad del estudio y el acceso al flujo completo de análisis, el código desarrollado para esta investigación se encuentra disponible en un repositorio digital.

El repositorio contiene:

- Notebook principal con el flujo completo del estudio (carga de datos, EDA, preprocesamiento, modelamiento y evaluación).
- Scripts auxiliares utilizados en el tratamiento de datos y visualización.
- Descripción de las librerías utilizadas y versión del entorno.
- Instrucciones básicas de ejecución.

El código puede ser consultado en el siguiente enlace:

<https://colab.research.google.com/drive/1KPwJSPBTDkPdENEquN8-eBf4E5ENBPI?usp=sharing>

El notebook permite reproducir íntegramente los resultados presentados en esta memoria, incluyendo el cálculo del Z''-score de Altman, el entrenamiento del modelo Random Forest, la evaluación predictiva y el análisis de importancia de variables.