



Facultad de Ciencias de la Administración

Carrera de Economía

Insolvencia con modelos de regresión logística con regularización Elastic Net, MARS, Naive Bayes del Sector B - explotación de minas y canteras del 2015 al 2024

Trabajo de titulación previo a la obtención del grado de Economista

Autor:

Juan José Córdova Crespo

Director:

Luis Gabriel Pinos Luzuriaga

Cuenca – Ecuador

Año

2026

DEDICATORIA

Dedico este trabajo a mis padres, quienes con su amor, esfuerzo y apoyo incondicional han sido la base de todo lo que he logrado. Gracias por enseñarme a no rendirme y por acompañarme en cada paso de este camino.

A mis amigos, por su compañía, apoyo y por estar presentes en los momentos más importantes, brindándome ánimo y motivación para seguir adelante.

AGRADECIMIENTO

Quiero expresar mi más profundo agradecimiento a mis padres, quienes han sido el pilar fundamental en mi vida. Gracias por su apoyo incondicional, por sus sacrificios y por creer en mí incluso en los momentos más difíciles. Este logro también les pertenece a ustedes.

A mi hermana, por su compañía, paciencia y por estar siempre presente, brindándome ánimo y motivación a lo largo de este proceso.

A mi tutor, por su guía, conocimiento y dedicación. Sus enseñanzas y orientación fueron clave para el desarrollo de este trabajo y para mi formación académica.

Finalmente, quiero agradecer de manera muy especial a mi novia, por su apoyo constante, su comprensión y su cariño incondicional.

Índice de Contenidos

DEDICATORIA.....	i
AGRADECIMIENTO	ii
Índice de Contenidos	iii
Índice de Figuras	iv
Índice de Tablas.....	v
RESUMEN	vi
ABSTRACT	vi
1. Introducción	1
2. Marco Teórico y Estado del Arte	2
3. Métodos.....	8
4. Resultados	14
5. Discusión.....	22
6. Conclusiones	24
7. Referencias.....	27

Índice de Figuras

Figura 1 Distribución porcentual de la quiebra empresarial según la rentabilidad	16
Figura 2 Quiebra según tamaño de la empresa.....	16

Índice de Tablas

Tabla 1 Descripción de las variables	10
Tabla 2 Análisis cuantitativo de medianas las variables	14
Tabla 3 Tabla de correlación de las variables.....	17
Tabla 4 Resultados de los modelos.....	18
Tabla 5 Métricas de predictibilidad en datos del Test en el modelo de regresión logística con regularización Elastic Net, MARS y Naive Bayes	21

RESUMEN

Este estudio analiza la insolvencia empresarial del sector B de explotación de minas y canteras en Ecuador durante el período 2015–2024, con el objetivo de identificar las variables que influyen en la probabilidad de quiebra de las empresas del sector. La investigación se centra en la predicción de la insolvencia empresarial, ya que al combinar el análisis financiero con herramientas de aprendizaje automático se logra entender de manera más clara cuáles son los factores de riesgo empresarial. En este caso, se trabajó con una base de datos de 1427 empresas, las cuales mostraron que sus variables más relevantes son el tamaño de la empresa, el nivel de endeudamiento, la liquidez y la rentabilidad. Para poder desarrollar esta investigación, primero se utilizó un modelo de regresión logística, el cual permitió analizar el comportamiento de los indicadores financieros en relación con la probabilidad de quiebra. Luego se incorporaron modelos más avanzados como Elastic Net, MARS y Naive Bayes dentro de un enfoque de aprendizaje automático, con el fin de mejorar la capacidad predictiva y seleccionar automáticamente las variables más relevantes. Los resultados muestran que el endeudamiento, la liquidez y la rentabilidad son factores clave que explican el riesgo de insolvencia en las empresas. Por otro lado, se observa que el modelo MARS presenta un mejor desempeño frente a los modelos tradicionales, ya que es capaz de captar relaciones no lineales en las variables y ofrecer estimaciones más estables.

Palabras clave: condiciones económicas, datos estadísticos, deuda, finanzas, gestión financiera, investigación económica, riesgo financiero

ABSTRACT

This study analyzes corporate insolvency in Ecuador's mining and quarrying sector (Sector B) during the 2015–2024 period, with the aim of identifying the variables that influence the probability of bankruptcy among companies in the sector. The research focuses on predicting corporate insolvency; by combining financial analysis with machine learning tools, we gain a clearer understanding of the factors that constitute business risk. In this case, we worked with a database of 1,427 companies, which revealed that the most relevant variables are company size, debt level, liquidity, and profitability. To conduct this research, a logistic regression model was first used, which revealed how financial indicators relate to bankruptcy. More advanced models such as Elastic Net, MARS, and Naive Bayes were then incorporated into a machine learning framework to improve predictive capabilities and automatically select the most relevant variables. The results show that indebtedness, liquidity, and profitability are key factors explaining the risk of insolvency in companies. Furthermore, it is observed that the MARS model performs better than traditional models, as it is capable of capturing nonlinear relationships among the variables and provides more stable estimates.

Keywords: economic conditions, statistical data, debt, finance, financial management, economic research, financial risk



1. Introducción

En Ecuador, la predicción de la insolvencia constituye un tema de gran relevancia, debido a que en los últimos años se ha evidenciado el cierre de empresas, como consecuencia de crisis económicas, lo que ha evidenciado que existe una fragilidad del sistema financiero y la elevada exposición al riesgo de quiebra. Esta problemática se encuentra asociada principalmente a deficiencias en la planificación y a una gestión financiera inadecuada, y existen factores que limitan la capacidad de las organizaciones para enfrentar escenarios adversos. En este contexto, es de suma importancia utilizar herramientas que permiten medir y anticipar el riesgo de insolvencia, ya que es necesario adaptarlas al contexto ecuatoriano, con el fin de contribuir a una toma de decisiones más eficiente para poder mitigar impactos económicos. En este contexto, es de suma importancia tener en cuenta los sectores estratégicos como el sector B, que corresponde a minas y canteras. En 2022, aportó el 11,3 % al producto interno bruto según el Banco Central del Ecuador (2023), lo que nos muestra que tiene un gran peso dentro de la economía del país. Además, su importancia se ve en las exportaciones, ya que, en 2023, los minerales generaron alrededor de 3.277 millones de dólares. Asimismo, dentro del sector existen 1427 empresas, en las cuales predominan las micro, pequeñas y medianas, ya que estas son clave para la actividad económica del país, por lo que son más vulnerables frente a problemas financieros, lo que aumenta su exposición al riesgo de insolvencia.

La insolvencia empresarial no tiene una sola causa, sino que combina factores internos como externos. Los factores internos más relevantes se encuentran en problemas de liquidez, la baja rentabilidad, el elevado nivel de endeudamiento, la falta de reinversión y una gestión financiera deficiente; a estos se suman factores externos, como los cambios en los precios internacionales, la inestabilidad política, los efectos de la pandemia, los cuales han restringido el financiamiento y han afectado el desempeño de las empresas. Esta investigación tiene como objetivo analizar la insolvencia empresarial en el sector B de explotación de minas y canteras del Ecuador durante el periodo 2015-2024. De manera específica, se busca evaluar y contextualizar el desempeño financiero de las organizaciones del sector en el periodo de estudio, así como determinar las variables financieras que influyen en la probabilidad de insolvencia de las empresas, mediante el

uso de modelos econométricos y de aprendizaje automático que permitan una identificación eficaz del riesgo. La insolvencia empresarial genera impactos significativos a nivel financiero, a nivel social y a nivel macroeconómico, ya que la pérdida de empleo afecta a los acreedores, proveedores e inversionistas, ya que, a nivel agregado, el cierre de organizaciones puede afectar a la producción, el empleo y el crecimiento económico, debido a que es un sector estratégico del país. En este sentido, las micro, pequeñas y medianas empresas presentan una mayor vulnerabilidad al riesgo de insolvencia, especialmente aquellas con estructuras financieras débiles o con pocos años de operación. En cuanto a los principales resultados, se evidencia que el endeudamiento es el factor que más incrementa la probabilidad de insolvencia en las empresas del sector, mientras que un mayor tamaño empresarial y niveles adecuados de rentabilidad actúan como factores que reducen dicho riesgo. Asimismo, se identifica que los problemas de liquidez y las variaciones negativas en los ingresos están estrechamente asociados con un mayor deterioro financiero, lo que confirma la importancia de una gestión eficiente de los recursos para garantizar la sostenibilidad empresarial. Finalmente, el presente trabajo se estructura de la siguiente manera. En la primera parte se desarrolla el marco teórico y la revisión de la literatura sobre insolvencia empresarial y los principales modelos de predicción de quiebra. Posteriormente, se presenta la metodología de investigación, donde se describe la construcción de la base de datos, las variables utilizadas y la especificación de los modelos econométricos y de aprendizaje automático aplicados.

2. Marco Teórico y Estado del Arte

Desde una perspectiva conceptual, la insolvencia puede definirse como la incapacidad de una empresa para cumplir con sus deudas y obligaciones financieras. En términos prácticos, implica la carencia de recursos suficientes para atender compromisos, lo que puede derivar en la desaparición de la organización, Cuberos Gómez (2005). La literatura coincide en que la insolvencia es producto de una combinación de factores internos, como problemas de liquidez, baja rentabilidad y elevado endeudamiento, y externos, como la inestabilidad política, caídas de precios internacionales y crisis sanitarias. El riesgo financiero se refiere a que una empresa experimente pérdidas económicas como consecuencia de decisiones financieras; pueden existir condiciones de mercado adversas o una mala gestión de recursos. De acuerdo con lo que comenta Ross et al. (2018), el riesgo financiero surge principalmente del uso de financiamiento externo y la incapacidad de la empresa para cumplir oportunamente con sus obligaciones financieras, lo cual afecta

a su estabilidad y su continuidad operativa. Desde otra perspectiva, Hull (2015) define el riesgo financiero como la incertidumbre en los flujos de efectivo futuros de una organización que se deriva de factores internos y externos que puedan alterar el desempeño económico, ya que este enfoque resalta que el riesgo financiero no solo depende de la estructura del capital, sino también de la exposición a variables macroeconómicas, tasas de interés y cambios en el entorno económico. En el ámbito empresarial Gitman & Zutter (2012) comentan que el riesgo financiero constituye un elemento clave en la solvencia y la sostenibilidad de las empresas, por lo que una gestión inadecuada incrementa el riesgo, y así puede conducir a problemas de liquidez, sobreendeudamiento y, en casos extremos, a la insolvencia empresarial. Existen diversos tipos de riesgos que afectan a las empresas. Ross et al. (2018) clasifican los riesgos financieros en varias categorías, entre las cuales destacan el riesgo de mercado, el riesgo de crédito, el riesgo de liquidez y el riesgo operativo financiero. El riesgo de mercado se relaciona con las fluctuaciones en variables como la tasa de interés, los tipos de cambio y los precios de los activos, por lo que son de gran importancia, ya que impactan negativamente en el valor de la empresa. Por otro lado, el riesgo de crédito se refiere a la posibilidad de que un contratante no cumpla con sus obligaciones financieras, afectando los flujos de efectivo esperados. El riesgo de liquidez está asociado con la incapacidad de la empresa para convertir sus activos en efectivo de manera oportuna (Van Greuning y Bratanovic, 2009), ya que es un tipo de riesgo crítico en empresas con estructuras financieras frágiles, pues una falta de liquidez puede provocar procesos de insolvencia. El riesgo operativo financiero comprende las pérdidas derivadas de fallas en procesos internos, sistemas o controles financieros; estos riesgos, aunque menos visibles (Van Greuning y Bratanovic, 2009), pueden mostrar efectos financieros y comprometer la estabilidad de la empresa en el largo plazo. El riesgo de insolvencia empresarial se define como la probabilidad de que una empresa no sea capaz de cumplir sus obligaciones financieras a medida que estas se presenten. Altman (1968) fue uno de los primeros en analizar la insolvencia de forma estructurada, ya que demostró que no se trata de un evento repentino, sino de un proceso que con el paso del tiempo se va desarrollando. Este deterioro se puede ver en indicadores como la liquidez, la rentabilidad y el nivel de endeudamiento. Posteriormente, Ohlson (1980) incorporó modelos de probabilidad para estudiar el riesgo de insolvencia, ya que, desde su punto de vista, la quiebra no ocurre repentinamente, sino que sigue un proceso, el cual puede anticiparse mediante el análisis de información contable y financiera, por lo que su enfoque permitió estimar la

probabilidad de insolvencia a partir de variables financieras observadas, sentando bases para el desarrollo de modelos predictivos. Desde un punto de vista bancario y regulatorio, Van Greuning y Bratanovic (2009) sostienen que el riesgo de la insolvencia representa una de las principales amenazas para la estabilidad financiera, ya que no solo afecta a la empresa individual, sino que genera efectos sistémicos en el sector económico al que pertenece. En conjunto, todo esto coincide en que el riesgo de insolvencia empresarial constituye una manifestación extrema del riesgo financiero, ya que es el resultado de una acumulación y una mala gestión de los diversos tipos de riesgos financieros que puede provocar insolvencia empresarial. Analizar este fenómeno es clave, ya que permite crear sistemas de alerta temprana que pueden ser de gran ayuda para que las empresas puedan tomar decisiones a tiempo. Esto es fundamental para evitar problemas mayores y así garantizar estabilidad y continuidad. Entender el riesgo financiero se vuelve necesario para comprender la insolvencia. En los últimos años se han desarrollado enfoques que permiten evaluar de manera más clara la probabilidad de quiebra a partir de información financiera, la cual ayuda a tomar decisiones más eficientes. Estos enfoques se han mostrado en modelos de predicción que utilizan ratios contables y variables explicativas para anticipar situaciones de deterioro financiero. Los planteamientos teóricos de Altman (1968) y posteriormente de Ohlson (1980) marcaron un punto de partida fundamental al establecer metodologías para la clasificación de empresas según su nivel de solvencia. Teniendo en cuenta estas contribuciones, la investigación empírica ha avanzado significativamente, ya que se ha logrado ampliar el análisis hacia distintos contextos económicos, de diferentes periodos, mediante el uso de distintas técnicas econométricas y de aprendizaje automático, con el objetivo de evaluar el desempeño, las fortalezas y las limitaciones de dichos modelos. En este sentido, el análisis de los siguientes estudios empíricos permite identificar los principales hallazgos, tendencias metodológicas y vacíos existentes sobre la predicción de la insolvencia empresarial. La predicción de la insolvencia empresarial constituye un tema de gran importancia a nivel internacional y particularmente en América Latina y Ecuador, debido a que en este último tiempo se ha incrementado el cierre de empresas asociado a crisis económicas, a inestabilidad financiera y a deficiencias en la gestión empresarial. En el caso ecuatoriano, estas problemáticas evidencian una alta exposición al riesgo de quiebra agravada por factores internos como la falta de planificación, la inadecuada administración y la limitada capacidad para adaptarse a escenarios adversos como los que vive el Ecuador. Desde una perspectiva teórica, los primeros modelos de predicción de insolvencia se basaron en el

análisis de ratios financieros. Altman (1968) propuso el modelo Z-Score fundamentado en una regresión múltiple, la cual combina indicadores de liquidez, rentabilidad, apalancamiento y solvencia; este modelo utiliza variables como el capital de trabajo sobre activos totales, las utilidades retenidas sobre activos totales, el resultado operativo sobre activos totales, el valor de mercado del patrimonio sobre el pasivo total y las ventas sobre activos totales, permitiendo clasificar a las empresas según su nivel de riesgo financiero. Ohlson (1980) desarrolló el modelo O-Score basado en una regresión logística, lo que mostró un avance importante frente al modelo de Altman, ya que dejó de lado el supuesto de normalidad de las variables. Este modelo considera variables financieras y contables como el tamaño de la empresa, el nivel de endeudamiento, la liquidez y la rentabilidad, y también toma en cuenta la estructura de capital y el desempeño financiero. El uso de modelos de predicción de insolvencia responde a la necesidad de anticipar quiebras y fortalecer la toma de decisiones empresariales y financieras, ya que son herramientas que permiten identificar vulnerabilidades a partir de información contable y financiera, clasificando a las empresas según su nivel de riesgo, lo que resulta fundamental tanto para los directivos, los inversionistas y los organismos reguladores.

En el contexto europeo, González et al. (2018) aplicaron el modelo de Altman al sector turístico español e identificaron que la insuficiencia de liquidez y el elevado apalancamiento fueron los principales factores asociados a la insolvencia. De manera complementaria, Mendaña et al. (2024) analizaron empresas turísticas españolas mediante análisis de conjuntos difusos combinados con variables financieras, concluyendo que la baja rentabilidad junto a altos niveles de apalancamiento explica la insolvencia empresarial. Asimismo, Sierpe et al. (2021) desarrollaron un modelo de alerta temprana enfocado en el sector de la construcción en España. El modelo que se utilizó fue el Random Forest, el cual mostró una capacidad predictiva cercana al 85 %, lo que demostró que herramientas de aprendizaje automático como Machine Learning son útiles para anticipar problemas de insolvencia. En el caso de Lituania, Kanapickiene et al. (2023) compararon distintos enfoques, como la regresión logística, redes neuronales y el modelo MARS, aplicados al sector de la construcción. Identificaron variables como la liquidez, la rentabilidad, el margen neto y el nivel de endeudamiento como variables significativas. El modelo MARS fue el que mostró el mejor desempeño, ya que superó el 94 % de precisión, lo que demuestra que es una herramienta útil para predecir la insolvencia. Por otro lado, Mehmood y De Luca (2025) desarrollaron un modelo

orientado a identificar dificultades financieras en empresas privadas europeas, combinando el modelo de Altman con regresión logística. Esta metodología alcanzó un 82 % de precisión y resaltó su utilidad para anticiparse a quiebras empresariales.

En el contexto turco, Aker y Karavardat (2023) evaluaron distintos modelos de aprendizaje automático, como la regresión logística, árboles de decisión, Random Forest, máquinas de vectores de soporte y Naive Bayes, aplicados a pymes, lo que mostró que las variables más relevantes fueron el pasivo total sobre activos totales, la liquidez corriente, la rentabilidad antes de impuestos sobre activos, el margen neto, la rotación de activos y el capital de trabajo. Al aplicarse varios modelos que mostraron altas precisiones, el modelo Naive Bayes destacó por su capacidad predictiva a tres años antes de la quiebra, con una exactitud del 97 %. En Indonesia, Prastyo (2025) comparó un modelo de valores extremos generalizados con modelos de Machine Learning tradicionales, como máquinas de vectores de soporte y regresión logística con regularización Elastic Net. Las variables más relevantes fueron el capital de trabajo sobre activos totales, destacando la estabilidad predictiva de los modelos hasta tres años antes de la insolvencia. Esta evidencia comparada en Europa y Asia nos muestra que tanto los modelos tradicionales como los enfoques modernos de aprendizaje automático tienen una elevada capacidad predictiva, ya que estudios como los de González et al. (2018), Mendaña et al. (2024), Kanapickiene et al. (2023), Mehmood y De Luca (2025), Aker y Karavardat (2023) y Prastyo et al. (2025) resaltan la liquidez, la rentabilidad y el apalancamiento como determinantes del riesgo financiero, así como su relevancia en modelos como MARS, Naive Bayes y la regresión logística con regularización Elastic Net.

En el contexto estadounidense, Shamara y Shinde (2025) evaluaron modelos como la regresión logística, Random Forest, redes neuronales y Naive Bayes, utilizando una base de 8200 empresas, las cuales mostraron que las variables más significativas son el capital de trabajo sobre activos totales y las utilidades retenidas sobre activos totales. El modelo Naive Bayes fue empleado como referencia comparativa, mostrando un buen nivel de precisión. En América Latina se han realizado varios estudios sobre la predicción de la insolvencia desde distintos puntos de vista. Por ejemplo, Moranga et al. (2019) aplicaron el modelo de Altman en empresas industriales chilenas; los resultados mostraron que muchas de ellas estaban en zona gris, lo que evidencia un nivel de crisis financiera donde

las variables determinantes son la liquidez y la rentabilidad. En Perú, Burneo et al. (2016) validaron el Z-Score de Altman, mostrando que el bajo capital de trabajo y el endeudamiento son factores críticos. Asimismo, Mongrut et al. (2011) analizaron empresas peruanas en quiebra mediante regresión logística, encontrando que el apalancamiento, el ROA y el PIB sectorial influyen en la insolvencia. En México, Cruz et al. (2017) compararon un modelo Logit y Poisson en el sector minero, concluyendo que el modelo Logit tiene un mayor poder predictivo. Por otro lado, Lozada et al. (2024) compararon los modelos de Altman y Ohlson en empresas de telecomunicaciones, mostrando que cada modelo clasifica el riesgo de forma distinta. En Colombia, Roque y Caicedo (2023) confirmaron que el modelo de Altman es consistente para identificar empresas en riesgo, mientras que Hernández (2022) compara modelos tradicionales con técnicas de Machine Learning, resaltando que los modelos más actuales tienen un mejor desempeño. En Ecuador se han realizado estudios como el de Orellana et al. (2023), quienes aplicaron modelos Logit y Probit en el sector de la construcción, encontrando que las microempresas son las más propensas a enfrentar la quiebra. De manera complementaria, Sigüenza y Orellana (2020) analizaron el sector de fabricación de productos minerales no metálicos con el modelo de Altman, Ohlson y CAMP. Sus principales resultados mostraron que la micro y pequeñas empresas enfrentan mayor probabilidad de insolvencia y problemas de liquidez. Bermeo y Armijos (2021) analizaron empresas de la construcción en el Azuay y evidenciaron una alta vulnerabilidad financiera. De la misma forma, Aymacaña et al. (2024) estudiaron empresas del Cotopaxi durante la crisis del COVID-19, mostrando que factores clave como la liquidez, el endeudamiento y el tamaño empresarial son variables importantes para la insolvencia empresarial; Así mismo, Haro et al. (2023) analizaron las empresas pymes del Ecuador específicamente Chimborazo, Cotopaxi, Pastaza y Tungurahua se aplicó el modelo de árbol de clasificación y sus resultados principales fueron que existe una relación directa entre la estructura de capital y la insolvencia empresarial. Estudios complementan estos hallazgos: Molina et al. (2022) aplicaron el modelo de Altman en empresas inmobiliarias de Pichincha y encontraron que muchas se ubicaban en zona segura. Girón et al. (2017) señalaron que la liquidez, el apalancamiento y el ROA son variables determinantes de la quiebra empresarial. Por otro lado, Haro (2021) analizó empresas del sector pesquero, encontrando que la regresión logística resulta significativa y que alrededor del 38 % de las empresas se encuentran en zona segura. Orellana et al. (2020) evaluaron el sector textil utilizando modelos de Altman, Ohlson y CAPM,

evidenciando alta vulnerabilidad en microempresas. Así mismo, Tustón y Macías (2025) aplicaron modelos de aprendizaje automático para la detección de fraudes financieros y modelos con Redes Neuronales supero el 93% de precisión, Radom Forest alcanzó un 97.7% de precisión, SVM con un 91% de presión y Naive Bayes alcanzó un 89% siendo eficiente con grandes bases de datos, Molina et al. (2024) desarrollaron un modelo Logit para cooperativas de ahorro y crédito; este modelo alcanzó una precisión del 97,8 %. En conjunto, estudios como los de Mejía et al. (2020), Villa et al. (2024), Puente et al. (2025), Molina et al. (2023) y Rivadeneira et al. (2022) confirman que tanto los modelos tradicionales como las técnicas de aprendizaje automático son herramientas aplicables para analizar el riesgo de insolvencia en distintos sectores del país. Los modelos de predicción de insolvencia empresarial son herramientas fundamentales para la gestión del riesgo financiero; aunque no existe un modelo único, la combinación de enfoques tradicionales y modernos ha demostrado mejorar la capacidad predictiva, ofreciendo una base sólida para su aplicación en el contexto ecuatoriano, especialmente en sectores vulnerables como el sector B, explotación de minas y canteras.

3. Métodos

La presente investigación adopta un enfoque cuantitativo, dado que es de suma importancia el análisis estadístico y econométrico de información financiera, con el objetivo de medir y predecir la probabilidad de insolvencia. El estudio utiliza modelos econométricos y aprendizaje automático para la identificación temprana del riesgo de quiebra de empresas del sector B: minas y canteras, aportando herramientas útiles para la toma de decisiones. El modelo utilizado es hipotético-deductivo, ya que parte de supuestos teóricos relacionados con la insolvencia empresarial y los indicadores financieros, los cuales son contrastados empíricamente mediante modelos de clasificación. Esta investigación plantea como hipótesis general que los modelos de regresión logística con regularización Elastic Net, MARS y Naive Bayes presentan una alta capacidad para predecir e identificar empresas en riesgo de quiebra. Se plantea que el alcance de la investigación es explicativo y predictivo; explicativo, porque analiza la relación entre indicadores financieros y la probabilidad de insolvencia empresarial, y predictivo, ya que estima la probabilidad de quiebra futura a partir de indicadores financieros. El diseño de la investigación es no experimental y longitudinal, debido a que las variables no son manipuladas, por lo que son observadas a lo largo del tiempo, utilizando información financiera anual correspondiente al periodo 2015–2024.

La investigación se apoya en el uso de fuentes secundarias oficiales, lo que proporciona información confiable. En este caso, la fuente principal de los datos proviene de la página web de la Superintendencia de Compañías, Valores y Seguros del Ecuador, específicamente de la sección de estados financieros y ranking empresarial, de donde se obtienen los datos de las empresas del sector B. A partir de la información contable reportada por las empresas, se construyen los indicadores financieros de liquidez, rentabilidad, endeudamiento, eficiencia y solvencia, los cuales formarán parte de las variables explicativas del modelo. De manera complementaria, se utilizan datos macroeconómicos y sectoriales provenientes del Banco Central del Ecuador y del Instituto Nacional de Estadística y Censos, así como información de la encuesta estructural empresarial, las cuales permiten contextualizar el desempeño financiero de las empresas dentro del entorno económico nacional y sectorial. Esto permite observar cómo ha cambiado la situación financiera de las empresas a lo largo del tiempo. De esta forma no solo se analiza un momento en específico, sino que también se puede observar de mejor manera el comportamiento del sector y los cambios que se han dado en la estructura a lo largo del tiempo. Posteriormente, se calcularon las nueve variables propuestas por Ohlson (1980) utilizando indicadores financieros obtenidos a partir de los estados financieros reportados a la Superintendencia de Compañías, Valores y Seguros del Ecuador. Las variables empleadas en el presente estudio corresponden a indicadores financieros, los cuales permiten evaluar la situación económica y financiera de las empresas del sector B, así como su probabilidad de insolvencia. La selección de estas variables se fundamenta en los trabajos realizados por Altman (1968) y Ohlson (1980) y en estudios aplicados a Latinoamérica. La variable dependiente representa el estado de insolvencia empresarial; esta definición binaria permite aplicar modelos de clasificación y estimar la probabilidad de insolvencia en función de los indicadores financieros y se expresa de la siguiente manera:

Y = 1: Empresa que presenta insolvencia, caracterizada por la no continuidad operativa, patrimonio negativo y pérdidas en un año.

Y = 0: Empresa solvente, tenga patrimonio positivo y que no presente pérdidas en un año .

Variables independientes.

Las variables independientes corresponden a indicadores financieros calculados a partir de los estados financieros reportados a la Superintendencia de Compañías, Valores y

Seguros del Ecuador. La tabla 1 muestra la descripción de las variables independientes que se utilizarán para predecir la insolvencia empresarial.

Tabla 1

Descripción de las variables

Variable	Explicación	Formula
X1: Tamaño de la empresa	Mide la dimensión económica de la empresa, asociada a su capacidad, operativa y financiera, empresas de mayor tamaño, suelen tener mayor estabilidad y acceso financiero.	$X1 = \ln(\text{Activos Totales})$
X2: Índice de endeudamiento	Refleja el grado de apalancamiento financiero y la dependencia de recursos externos, ya que tener un alto nivel de endeudamiento, incrementa el riesgo de insolvencia.	$X2 = \frac{\text{Activo Total}}{\text{Pasivo Total}}$
X3: Capital de trabajo	Mide la capacidad de la empresa para financiar sus operaciones de corto plazo, ya que al tener un capital del trabajo negativo indica problemas de liquidez, lo que puede llevar a la insolvencia empresarial.	$X3 = \text{Activo Corriente} - \text{Pasivo Corriente}$
X4: Índice de liquidez corriente	Evalúa la capacidad de la empresa para cumplir sus obligaciones de corto plazo, utilizando sus activos corrientes	$X4 = \frac{\text{Pasivo Corriente}}{\text{Activo Corriente}}$
X5: Dummy de solvencia	Variable que identifica si la empresa presenta una situación de solvencia, mínima aceptable.	1: <i>Empresa solvente</i> 0: <i>Empresa insolvente</i>
X6: Rentabilidad sobre activos (ROA)	Mide la eficiencia de la empresa para generar utilidades a partir de sus activos totales.	$X6 = \frac{\text{Activo Total}}{\text{Utilidad Neta}}$
X7: Índice de rentabilidad	Evalúa desempeño financiero general de la empresa, reflejando su capacidad de generar beneficios a la empresa.	$X7 = \frac{\text{Ingresos Totales}}{\text{Resultado Operativo}}$

X8: Dummy de rentabilidad	Variable que identifica si la empresa presenta resultados positivos	1: <i>Rentabilidad positiva</i> 0: <i>Rentabilidad negativa</i>
X9: Cambio de utilidad neta	Captura la variación anual del desempeño financiero, teniendo en cuenta los cambios negativos, ya que esos cambios son señales tempranas de insolvencia.	$X9 = \frac{Utilidad_t - Utilidad_{t-1}}{Utilidad_{t-1}}$

Fuente: Basado en Ohlson (1980)

Una vez definidas las variables que se ocuparán para realizar el siguiente estudio, se desarrolló un proceso de tratamiento y procesamiento de datos, el cual tiene como objetivo garantizar la calidad y la consistencia de la información utilizada para la estimación de los modelos. Lo primero que se realizó fue la eliminación de las empresas que no presentan ventas por actividades ordinarias, ya que la ausencia de ingresos operativos puede generar problemas en la estimación de los modelos, y así causar una mala interpretación de los indicadores financieros. En segundo lugar, se realizó la división de los datos en entrenamiento y prueba, a los que se les destinó un 80 % a los datos de entrenamiento, que dio como resultado 5948 datos, y un 20 % a los datos de prueba, que dio como resultado 1487 datos, seleccionados de forma aleatoria. En tercer lugar, se estableció un esquema de remuestreo con los indicadores de desempeño utilizados para la evaluación de los modelos. Para el procesamiento se empleó un procedimiento de validación cruzada de diez particiones (10 fold cross validation). El conjunto de datos destinado al entrenamiento fue dividido en diez subconjuntos. Se ocuparon métricas que evalúan en cada remuestreo, las cuales son exactitud (accuracy), la sensibilidad, la especificidad y la exactitud balanceada; finalmente, en el cuarto paso se procedió a ajustar los modelos planteados, como: regresión logística con regularización Elastic Net, MARS, Naive Bayes.

Por otro lado, se llevó a cabo un proceso de identificación y eliminación de valores atípicos con el fin de evitar distorsiones en la estimación de los coeficientes para mejorar la estabilidad de los datos. Se empleó el diagrama de caja y bigotes, según Gujarati & Porter (2010), utilizando criterios basados en percentiles, metodología recomendada en análisis exploratorio de datos para detección de observaciones extremas, ya que la exclusión mejora la capacidad predictiva de los modelos y reduce la influencia de observaciones extremas que podrían sesgar los resultados.

La variable dependiente de la insolvencia fue utilizada como criterio de estratificación, lo que nos ayudó a mantener una proporción constante de empresas solventes e insolventes. Debido a la presencia de desbalance entre las clases, se aplicó una técnica de sobremuestreo con el fin de equilibrar las observaciones antes del entrenamiento de los modelos. Posteriormente se refinó la estrategia de muestreo y las métricas de evaluación, a las cuales se aplicó un esquema de validación cruzada de 10 particiones, el cual permitió evaluar de una manera más robusta el desempeño de los modelos para poder así reducir el riesgo de sobreajuste; la exclusión de los valores atípicos permite mejorar la estabilidad de los coeficientes estimados y la capacidad predictiva de los modelos. Finalmente, el conjunto de datos fue sometido a procesos de verificación de consistencia, asegurando que las variables cumplan los supuestos básicos y requeridos por los modelos de clasificación, lo que nos permite trabajar con una base homogénea y estadísticamente adecuada para la estimación de modelos de regresión logística con regularización Elastic Net, MARS y Naive Bayes. En conjunto, estas variables permiten medir la liquidez, rentabilidad, solvencia, endeudamiento y eficiencia para poder evaluar la estabilidad financiera de las empresas; su utilización es de suma importancia en los modelos clásicos y con modelos modernos que nos van a ayudar a predecir la quiebra y es adecuada para el análisis del sector minero ecuatoriano. Con el objetivo de analizar y predecir la probabilidad de insolvencia empresarial del sector B, se emplean modelos econométricos y de aprendizaje automático orientados a la clasificación de empresas, según su situación financiera. Estos modelos nos permiten identificar patrones financieros asociados con el riesgo de insolvencia y estiman la probabilidad de quiebra a partir de los indicadores seleccionados.

La regresión logística se utiliza como modelo base para estimar la probabilidad de que una empresa se encuentra en una situación de insolvencia, ya que considera una variable dependiente binaria y su fórmula se expresa de la siguiente manera.

$$P(Y = 1 | X) = \frac{1}{1 + e^{-(\beta^0 + \beta^1 X^1 + \beta^2 X^2 + \beta^3 X^3 + \beta^4 X^4 + \beta^5 X^5 + \beta^6 X^6 + \beta^7 X^7 + \beta^8 X^8 + \beta^9 X^9)}} \quad (1)$$

, donde Y representa la insolvencia empresarial, Xi son los indicadores financieros y Bi los parámetros a estimar. Con el fin de mejorar la capacidad predictiva y reducir la multicolinealidad, se aplica la regresión logística con regularización Elastic Net, que

combina penalizaciones L1 y L2 para seleccionar variables relevantes y atenuar la multicolinealidad, cuya fórmula se expresa de la siguiente manera:

$$\min B \frac{1}{n} \sum_{i=1}^n [-y_i \log(p_i) - (1 - y_i) \log(1 - p_i)] + \lambda((1 - \alpha) \frac{1}{2} \|B_2^2\| + \alpha \|B\|_1) \quad (2)$$

Este modelo ha demostrado alta eficacia en estudios de predicción de quiebra, como los realizados por Cruz et al. (2017) y Lambreton (2023). Asimismo, en Indonesia, Prastyo et al. (2025), al aplicar el modelo de regresión logística con la regularización Elastic Net, demostraron que fue eficaz para predecir la insolvencia empresarial, ya que mantuvieron la estabilidad hasta tres años.

El modelo MARS es una técnica no paramétrica, la cual capta relaciones no lineales y efectos de interacción entre los indicadores financieros. Este modelo divide el espacio de datos en distintas regiones y ajusta modelos lineales por tramos, adaptándose de forma flexible a la estructura de los datos, donde su fórmula se expresa de la siguiente manera:

$$Y = \beta_0 + \sum_{m=1}^M B_m H_m(X) \quad (3)$$

Cada función base $h_m(X)$ representa combinaciones de las variables independientes mediante funciones por tramos, lo que resulta útil para identificar umbrales financieros críticos asociados al riesgo de insolvencia. En Lituania, Kanapickiene et al. (2023) aplicaron el modelo MARS a datos financieros. El estudio mostró que el modelo MARS tuvo la mejor precisión y eficacia, ya que sus valores fueron superiores al 94 %, lo que lleva a la conclusión de que es un modelo robusto para predecir la insolvencia de las empresas.

El modelo Naive Bayes es un clasificador probabilístico basado en el teorema de Bayes, ampliamente utilizado para la predicción de eventos binarios. Este modelo asume independencia condicional entre las variables explicativas, lo que ayuda a que el cálculo sea más simple y permite procesar grandes volúmenes de información financiera de manera eficiente. Por otro lado, en Turquía, Aker y Karavardat (2023) realizaron el modelo Naive Bayes, el cual se aplicó a ratios financieros, ya que presentó el mejor desempeño hasta tres años antes de la quiebra, con un 97 % de exactitud, permitiendo diferenciar entre empresas en riesgo y empresas solventes dentro del sector B. Por otro lado, Shamara y Shinde (2025), en Estados Unidos, aplicaron el modelo Naive Bayes para

comparar resultados, mostrando un nivel bastante alto, debido a que este modelo permite estimar de forma sencilla la probabilidad de que una empresa entre en quiebra..

4. Resultados

La base de datos analizada está conformada por 1427 empresas pertenecientes al sector B, correspondientes al periodo 2015 al 2024, lo que nos permite tener un panorama del comportamiento financiero del sector. Esta base de datos incluye variables estructurales y financieras, las cuales ayudan a describir la situación económica y operativa de las empresas del sector. El análisis descriptivo comienza con el análisis cuantitativo de las variables utilizadas para realizar el modelo, que se muestra en la tabla 2.

Tabla 2

Análisis cuantitativo de medianas las variables

Medianas	X1	X2	X3	X4	X6	X7
Quiebra	3,17	0,49	0,23	0,51	0,00	0,00
No quiebra	2,87	1,46	-0,55	2,16	-0,23	-0,13

Las cuales muestran evidencia de alta dispersión en la mayoría de variables financieras, así como la presencia de valores atípicos, particularmente en el total de activos, total de pasivos e indicadores financieros. Esto muestra que existe una heterogeneidad en las empresas del sector, donde hay empresas grandes con estructuras financieras amplias, así como micro y pequeñas empresas con estructuras financieras mucho más débiles. Asimismo, se logró identificar valores negativos en variables como la rentabilidad, el apalancamiento y otros ratios financieros, lo que evidencia que existen empresas con pérdidas operativas o con bajo patrimonio. Teniendo en cuenta este comportamiento, es característico de sectores en los cuales se intensifica la inversión de capital, donde los resultados pueden verse fuertemente afectados por variaciones de precios internacionales o costos operacionales elevados. Adicionalmente, se logró identificar que las empresas micro y pequeñas del sector B se caracterizan por una estructura empresarial de menor escala, lo que implica mayores restricciones de acceso financiero y menor capacidad de absorción ante cambios en los precios internacionales.

Asimismo, la distribución de la variable dependiente asociada con la quiebra empresarial muestra un desbalance entre empresas solventes e insolventes, ya que las empresas sin quiebra constituyen la categoría predominante, lo que justifica la aplicación de técnicas de balanceo y muestreo para la etapa de modelación, con el objetivo de mejorar la

capacidad predictiva de los modelos y evitar sesgos de especificación. Por otro lado, las medianas de las variables X1, X2, X3, X4, X6 y X7 difieren entre empresas que quiebran y aquellas que no. El uso de la mediana permite una mejor comparación entre ambos grupos, ya que reduce la influencia de valores extremos en los datos.

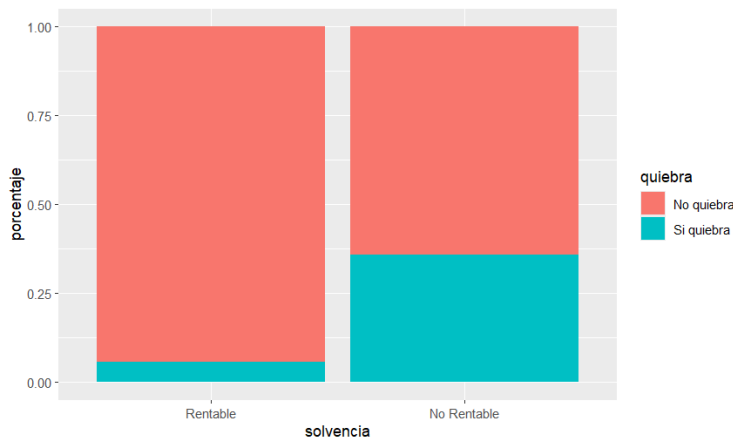
En el caso de X1, las empresas en quiebra muestran una mediana mayor en comparación con las empresas no quebradas, lo que sugiere que en la muestra analizada la quiebra no se concentra únicamente en empresas de menor tamaño, sino que también afecta a empresas de mayor escala. Por otro lado, en X2 se observa una diferencia marcada entre ambos grupos, ya que las empresas no quebradas presentan una mediana significativamente mayor frente a las empresas quebradas, lo que indica que un mayor nivel de endeudamiento no determina necesariamente que una empresa esté en quiebra. En cuanto a X3, las empresas en quiebra muestran una mediana positiva, mientras que las empresas no quebradas presentan una mediana negativa. Esto evidencia que ambos grupos tienen estructuras financieras distintas y que el capital de trabajo por sí solo no es un factor determinante de la quiebra, sino que puede depender de la interacción con otras variables financieras.

Asimismo, en X4 la mediana es considerablemente menor en las empresas en quiebra respecto a las empresas no quebradas, lo que indica que las empresas insolventes tienen menor capacidad para cubrir sus obligaciones de corto plazo, confirmando la relevancia de la liquidez como un indicador clave de estabilidad financiera. En relación con X6, la mediana de las empresas en quiebra es prácticamente cero, mientras que en las empresas no quebradas es negativa, lo que muestra diferencias en el comportamiento de esta variable. Por otro lado, en X7 se observa que el grupo de empresas no quebradas tiene una mediana negativa, lo que indica una menor relevancia de esta variable para explicar la quiebra de las empresas insolventes.

Así mismo, se analizará de forma cualitativa las variables; la primera que se analizará es la figura 1.

Figura 1

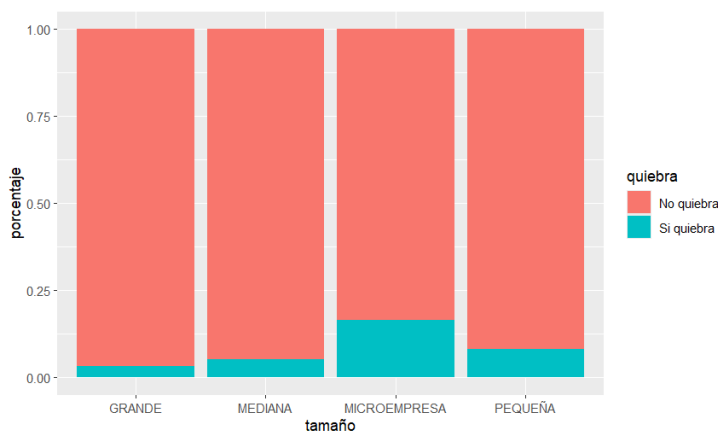
Distribución porcentual de la quiebra empresarial según la rentabilidad



Éste gráfico nos muestra el grupo de rentable, el cual muestra que la prioridad de quiebras es muy baja, predominando a las empresas que no tienen la propiedad de quiebra; en el grupo de no rentable, se observa un incremento sustancial del porcentaje de empresas quebradas, lo que indica una relación directa entre las pérdidas y el riesgo de insolvencia y por último, la categoría NA mantiene una porción reducida de quiebra, pero no debe interpretarse con cautela por tratarse de información incompleta. De forma complementaria se evalúa la figura 2.

Figura 2

Quiebra según tamaño de la empresa



Observamos que las microempresas tienen mayor probabilidad de quiebra, lo que evidencia una mayor vulnerabilidad financiera. Por otro lado, las pequeñas y medianas empresas presentan una probabilidad de quiebra menor que las microempresas, pero no superior a las grandes y, por último, las empresas grandes muestran menor incidencia de

quiebra, reflejando mayor estabilidad operativa y financiera. Esta tendencia es coherente con la teoría financiera, ya que la evidencia empírica muestra que las empresas de mayor tamaño suelen tener mayor acceso a financiamiento, diversificación y capacidad de absorber problemas financieros.

Por otro lado, tenemos la tabla 3 de correlación de las variables financieras consideradas en el modelo, la cual permite identificar la relación lineal entre cada indicador.

Tabla 3

Tabla de correlación de las variables

	X2	X4	X3	X9	X7	X1	X6
X2	1.00	0.03	-0.70	-0.01	0.01	-0.06	-0.34
X4	0.03	1.00	-0.04	0.01	0.01	0.01	0.02
X3	-0.70	-0.04	1.00	0.01	0.03	0.03	-0.08
X9	-0.01	0.01	0.01	1.00	0.01	0.01	0.03
X7	0.01	0.01	0.03	0.01	1.00	0.02	0.05
X1	-0.06	0.01	0.03	0.01	0.02	1.00	0.05
X6	-0.34	0.02	-0.08	0.03	0.05	0.05	1.00

La correlación muestra la intensidad y el sentido de la relación lineal entre cada variable, ya que los coeficientes de correlación se encuentran en el intervalo de -1 a 1, donde un valor cercano a uno indica una relación positiva fuerte, valores cercanos a -1 una relación negativa fuerte y valores próximos a cero sugieren una relación débil o inexistente. En términos generales, el gráfico muestra que la mayoría de las correlaciones entre las variables financieras son bajas o moderadas, lo que sugiere ausencia de relaciones fuertes. Se observa una relación inversa entre las variables asociadas a la liquidez y el endeudamiento, lo cual es consistente con la teoría financiera, dado que mayores niveles de apalancamiento suelen estar acompañados de una menor cobertura de las obligaciones de corto plazo. Asimismo, las variables relacionadas con el tamaño de la empresa presentan correlaciones moderadas con algunos indicadores financieros, reflejando que la dimensión de la empresa influye parcialmente en su estructura financiera. La ausencia de correlaciones altas entre variables confirma que no existe un problema severo de multicolinealidad, lo que respalda la pertinencia de utilizar estos indicadores de manera conjunta en modelos econométricos y de aprendizaje automático.

Modelos

La tabla permite observar tanto las variables que fueron incluidas en cada modelo como aquellas que no fueron utilizadas, en el caso de los modelos Elastic Net, MARS y Naive Bayes. No obstante, se presentan los coeficientes principales de los indicadores

financieros (X1, X2, X3, X4, X6, X7 y X9), junto con las variables categóricas relacionadas con la rentabilidad, el tamaño y la ubicación. Al comparar estos resultados entre modelos tradicionales y enfoques de Machine Learning, se puede observar qué variables tienen mayor peso en cada caso, lo que facilita identificar cuáles factores se mantienen constantes en la predicción de la insolvencia.

La tabla 4 muestra los resultados obtenidos de los modelos aplicados para la predicción de insolvencia, donde se presentan los coeficientes y variables relevantes.

Tabla 4

Resultados de los modelos

Variables	Regresión Logística Clásica (1)	Elastic Net (2)	MARS (3)	Naive Bayes (4)
Intercepto	1.60	(-)	(-)	-0.11
X1	0.37	0.00	No usada	-0.21
X2	162.00	20.70	Usada	0.27
X3	73.30	-12.00	Usada	-0.20
X4	0.17	0.00	No usada	0.01
X6	-1.40	0.00	No usada	-0.13
X7	0.05	0.00	No usada	0.02
X9	-0.66	0.00	No usada	-0.31
Rentabilidad (No Rentable)	81.40	2.58	Usada	1.36
X2 × Rentabilidad	(-)	16.60	No usada	0.18
Tamaño	0.79 (mediana)	1.76 (micro) / 1.05 (pequeña)	No usada	Micro (0.36) / mediana (-0.86) / pequeña (-0.32)
Provincia	(-)	Guayas (-1.27) / el oro (0.99)	No usada	Guayas (-1.04) / el Oro (0.51)
Ciudad	(-)	Machala (-1.11) / quito (-1.03)	No usada	(-)

El primer modelo que se ocupó fue una regresión logística, a la cual se le realizó un procesamiento de datos, ya que se estratificó la base de datos dividiéndola en un 80 % para la base de entrenamiento y un 20 % para la base de prueba; sobre los datos de entrenamiento se estandarizaron todas las variables para pasar a valores Z, además se ajustaron de tal forma que, al realizar el entrenamiento, exista una misma proporción que

en el test, de manera que los ceros y los unos estén bien distribuidos en la base de datos, ya que es el modelo base de este estudio y es ampliamente utilizado en la literatura para la predicción de quiebra empresarial. Este modelo permite estimar la probabilidad de ocurrencia de un evento binario.

La regresión logística evidencia que el endeudamiento (X2) es el principal determinante del riesgo de quiebra, ya que su coeficiente positivo elevado indica que mayores niveles de deuda incrementan significativamente la probabilidad de insolvencia. Por otro lado, el capital de trabajo (X3) presenta un efecto positivo, lo que muestra que determinadas estructuras de liquidez podrían asociarse con mayor riesgo. La falta de rentabilidad incrementa la probabilidad de quiebra, lo que confirma que la ausencia de utilidades afecta de forma negativa la sostenibilidad financiera. En comparación con las variables X6 y X9, las cuales presentan coeficientes negativos, actúan como factores que reducen el riesgo. Por su parte, Elastic Net fue estimado mediante validación cruzada, lo que permite optimizar parámetros de regularización y mejorar el desempeño predictivo. Como resultado, el modelo penalizó el conjunto inicial de variables, ya que concentra el poder explicativo en aquellas variables con mayor capacidad para discriminar entre empresas quebradas y no quebradas. Los principales resultados muestran que las variables con mayor peso dentro del modelo son el índice de endeudamiento (X2), que presenta un coeficiente positivo, lo que indica que a mayores niveles de deuda incrementa la probabilidad de quiebra, consolidándose como el principal determinante del riesgo. Asimismo, la interacción entre el endeudamiento y la no rentabilidad también muestra un signo positivo, evidenciando el efecto del apalancamiento cuando la empresa no genera utilidades. Por el contrario, el capital de trabajo (X3) muestra un coeficiente negativo, actuando como variable protectora al reducir el riesgo de insolvencia; la condición de no rentabilidad mantiene un signo positivo, reforzando que las empresas sin utilidades enfrentan una mayor probabilidad de quiebra.

El modelo MARS fue estimado con el objetivo de capturar relaciones no lineales y efectos por tramos entre las variables financieras a partir de un proceso de selección automática. El modelo captó tres predictores principales, lo que evidencia una estructura parsimoniosa con alto poder predictivo. De acuerdo con la gráfica, el capital de trabajo (X3) se identifica como el factor más relevante en la explicación del riesgo de quiebra, seguido por la condición de no rentabilidad y, en tercer lugar, el nivel de endeudamiento (X2). El resto de variables financieras presentan una contribución nula dentro del modelo, lo que

confirma que el riesgo de insolvencia se concentra principalmente en indicadores como liquidez, rentabilidad y estructura financiera.

Cuando se analiza el modelo MARS, se observa que el capital de trabajo tiene un comportamiento no lineal. Esto significa que, en niveles bajos de capital de trabajo, el modelo aumenta la probabilidad de quiebra, lo que deja claro que no solo importa cuánto capital de trabajo tenga la empresa, sino también cómo lo utiliza para mantener su operación en marcha. Asimismo, la rentabilidad, al dividirse en categorías, muestra que las empresas no rentables tienen un impacto positivo en la probabilidad de quiebra, ya que una empresa que presenta pérdidas aumenta considerablemente su riesgo de insolvencia. Este hallazgo es coherente con los resultados obtenidos tanto en la regresión logística tradicional como en el modelo Elastic Net, reforzando el papel de la rentabilidad como determinante del riesgo empresarial. En cuanto a la variable endeudamiento (X2), el modelo MARS identifica un comportamiento no lineal donde niveles bajos de endeudamiento se asocian con una menor probabilidad de quiebra, mientras que valores más elevados muestran una probabilidad de quiebra mayor. Esto resalta la importancia de una menor dependencia del financiamiento externo para la estabilidad financiera.

El modelo Naive Bayes fue estimado bajo el supuesto de independencia entre las variables explicativas, lo que permite calcular la probabilidad de quiebra empresarial. Los resultados principales muestran que el capital de trabajo (X3) presenta un coeficiente positivo de 0,27, lo que indica que determinados niveles de liquidez podrían estar asociados con mayor probabilidad de quiebra. En contraste, el endeudamiento (X2) presentó un coeficiente negativo de 0,21, a diferencia de los modelos Elastic Net y MARS, donde el apalancamiento incrementó el riesgo. Este efecto puede explicarse por la estructura probabilística del modelo, donde el supuesto de independencia puede reducir el peso explicativo entre variables correlacionadas. El cambio de utilidad neta mostró un coeficiente negativo de 0,31, actuando como variable protectora, lo que implica que mejoras en el desempeño financiero reducen la probabilidad de quiebra, lo cual es coherente con la teoría financiera. Por su parte, la condición de no rentabilidad (X8) presentó un coeficiente positivo de 0,02, aunque en menor magnitud que en otros modelos, lo que confirma que las empresas con pérdidas enfrentan mayor riesgo de insolvencia. En cuanto a las variables estructurales, como el tamaño empresarial, se observa que las microempresas presentan un coeficiente positivo de 0,36, indicando una mayor probabilidad de quiebra, mientras que las empresas medianas presentan un coeficiente negativo de 0,86, mostrando un menor riesgo.

Estos cuatro modelos coinciden en que las variables más significativas son el capital de trabajo, el nivel de endeudamiento y la rentabilidad, ya que aparecen de manera consistente como variables relevantes en todos los enfoques, lo que muestra una robustez empírica de estos indicadores y confirma su importancia para predecir la insolvencia. La regresión logística proporciona una base de interpretación clara; Elastic Net mejora dicha base al seleccionar las variables más relevantes y mejorar la capacidad predictiva; mientras que el modelo MARS ofrece una visión más profunda del comportamiento no lineal de las variables financieras. Al realizar esta comparación, la evidencia muestra que los modelos no se contradicen, sino que se complementan, ya que cada uno aporta distintos niveles de análisis. Finalmente, el modelo Naive Bayes aporta una perspectiva probabilística basada en la contribución de cada variable explicativa, destacando patrones relevantes en el riesgo de insolvencia y mostrando una alta capacidad para detectar empresas en situación de quiebra. En conjunto, estos modelos fortalecen la validez de los resultados y proporcionan una visión integral del riesgo de insolvencia empresarial del sector B.

En la tabla 5 se presentan las métricas de desempeño predictivo de cada modelo utilizado, lo que permite comparar la capacidad de clasificación de Elastic Net, MARS y Naive Bayes.

Tabla 5

Métricas de predictibilidad en datos del Test en el modelo de regresión logística con regularización Elastic Net, MARS y Naive Bayes

Métrica	Elastic Net	MARS	Naive Bayes
Accuracy	0.83	0.93	0.91
Kappa	0.40	0.72	0.42
Sensibilidad (Recall)	0.84	0.93	0.99
Especificidad	0.70	0.92	0.32
PPV (Precisión Positiva)	0.95	0.99	0.92
NPV	0.38	0.64	0.79
MCC	0.42	0.73	0.47
J-index (Youden)	0.54	0.85	0.31
Balanced Accuracy	0.77	0.93	0.66
Detection Prevalence	0.78	0.83	0.95

Precision	0.95	0.99	0.92
Recall	0.84	0.93	0.99
F-measure (F1-score)	0.89	0.96	0.95

En el caso del modelo Elastic Net, las métricas en conjunto muestran un desempeño aceptable, con un nivel de accuracy moderado, lo que indica una buena proporción de clasificaciones correctas. La sensibilidad es relativamente alta, lo que implica que el modelo identifica a la mayoría de las empresas quebradas; por el contrario, la especificidad es más baja en comparación con MARS, lo que sugiere mayores dificultades para clasificar correctamente las empresas no quebradas. El valor de precisión es elevado, lo que indica que, cuando el modelo predice la quiebra, suele acertar. A pesar de ello, el NPV evidencia limitaciones en predecir empresas sanas. En conjunto, Elastic Net ofrece estabilidad y parsimonia, aunque con menor poder predictivo.

Por su parte, el modelo MARS mostró un desempeño superior en el conjunto de prueba, alcanzando mayores niveles de accuracy, sensibilidad y especificidad, lo que evidencia una mejor capacidad para distinguir entre empresas quebradas y no quebradas. La alta precisión muestra que las predicciones de quiebra son altamente confiables, mientras que este modelo presenta un equilibrio sólido entre clasificación positiva y negativa; la balanced accuracy elevada indica que el modelo no se ve afectado por el desbalance de clases.

El modelo Naive Bayes presenta una alta capacidad para identificar las empresas en situación de riesgo de quiebra, evidenciada por una elevada sensibilidad. Es de suma importancia tener en cuenta que su baja especificidad indica dificultades para clasificar correctamente empresas no quebradas, generando un número considerable de falsos positivos. En general, se observa que es un modelo eficaz para la detección temprana del riesgo, aunque muestra un desempeño menos equilibrado en comparación con Elastic Net y MARS.

5. Discusión

Los resultados de esta investigación nos muestran que la insolvencia en el sector B durante el periodo 2015–2024 no se da por casualidad, sino que está relacionada con factores financieros estructurales, especialmente el endeudamiento, la liquidez y la rentabilidad. Estos resultados coinciden con lo planteado por Altman (1968), quien comenta que variables como el apalancamiento y el capital de trabajo son de suma

importancia para anticipar problemas financieros en las empresas. De igual manera, los resultados también coinciden con el enfoque de Ohlson (1980), quien propuso la regresión logística como una herramienta más flexible para estimar la probabilidad de quiebra empresarial. En este estudio se tomó como base el modelo de Ohlson (1980), pero se decidió eliminar la variable X5, dado que esta causa problemas de endogeneidad. Esto se da ya que, en el modelo original, dicha variable representa una condición previa de insolvencia, pero en esta investigación la variable dependiente identifica directamente a las empresas insolventes a partir de criterios como el patrimonio negativo y la pérdida de continuidad operativa. En cuanto a los resultados, el endeudamiento (X2) aparece como uno de los factores más determinantes, mostrando un efecto positivo y significativo en la regresión logística con regularización Elastic Net. Por lo que, a medida que aumenta la deuda, también lo hace la insolvencia. Este resultado es consistente con estudios de Burneo et al. (2016), Mongrut et al. (2011), Moranga et al. (2019) y Roque y Caicedo (2023), quienes afirmaron que el pasivo total sobre activos es un factor clave del riesgo financiero. Asimismo, el modelo Elastic Net permite observar algo adicional, ya que el endeudamiento se combina con la falta de rentabilidad, incrementando el riesgo, por lo que este resultado refuerza lo encontrado por Prastyo (2025), quien destaca que este tipo de modelos ayuda a captar relaciones más complejas entre las variables. Por otro lado, el modelo MARS identificó al capital de trabajo (X3) como uno de los factores más importantes, pero con un comportamiento no lineal. Esto muestra que el capital de trabajo no siempre tiene un efecto positivo, ya que en ciertos niveles puede incrementar el riesgo de quiebra. Este resultado también se observa en Kanapickiene et al. (2023), quienes muestran que el modelo MARS superó a los demás modelos. Asimismo, estudios como el de Sierpe et al. (2021), utilizando el modelo Random Forest, refuerzan la idea de que la insolvencia empresarial responde a relaciones más complejas que no siempre se pueden explicar con modelos lineales.

En cuanto a la rentabilidad, su importancia también ha sido respaldada por estudios como los de Mehmood y De Luca (2025) y Hernández (2022), que destacan que indicadores como el ROA y el margen neto son significativos para anticipar la insolvencia. Desde un enfoque de modelos modernos como Machine Learning, investigaciones como las de Shamara y Shinde (2025) y Aker y Karavardat (2023) muestran cómo modelos como Naive Bayes y Random Forest pueden alcanzar altos niveles de precisión, incluso superando modelos tradicionales.

En esta investigación, el modelo MARS fue el que mostró el mejor desempeño, logrando un buen equilibrio entre sensibilidad y especificidad, lo que coincide con lo planteado por Kanapickiene et al. (2023) y Lozada et al. (2024). Por su parte, el modelo Naive Bayes destacó por su alta sensibilidad, lo que lo hace útil como herramienta de alerta temprana, aunque con menor especificidad, tal como señala Cruz et al. (2017).

En conjunto, los resultados muestran que no existe un solo modelo para detectar la insolvencia empresarial, sino que es necesario utilizar enfoques que se adapten al contexto económico de la investigación. Cada modelo tiene sus ventajas dependiendo de lo que se busque. El modelo MARS destaca por captar relaciones no lineales; Elastic Net por su estabilidad, la reducción de la multicolinealidad y su capacidad de selección automática de variables; y Naive Bayes por su capacidad de detectar riesgos de forma temprana. Al combinar estos enfoques tradicionales con técnicas de Machine Learning, se obtiene una estrategia sólida para predecir la insolvencia empresarial, lo que respalda estudios ecuatorianos como los de Orellana (2020, 2023), Bermeo y Armijos (2021), Molina (2022, 2024), Haro (2021), Girón (2017), Mejía (2020), Villa (2024), Puente (2025), Tuston y Macías (2025) y Rivadeneira (2022), quienes confirman que estos modelos son aplicables en distintos sectores del país.

6. Conclusiones

La presente investigación analiza la insolvencia del sector B, explotación de minas y canteras en el Ecuador durante el periodo 2015–2024. El Ecuador vive en un contexto en el cual existe volatilidad en la economía. El principal problema se basa en identificar los patrones existentes en las empresas para poder anticipar la quiebra, ya que este sector forma parte de un eje estratégico de la economía. En este estudio no solo se buscó analizar el comportamiento financiero del sector, sino también construir un modelo predictivo capaz de determinar las variables que inciden en la probabilidad de insolvencia, que aportan evidencia empírica útil para la gestión empresarial y la toma de decisiones. Se logró evaluar y contextualizar el desempeño financiero de las organizaciones del sector B entre el 2015–2024, evidenciando una marcada heterogeneidad en su estructura económica.

El análisis descriptivo muestra que hay una alta variación entre las empresas en aspectos importantes como activos, pasivos, el nivel de endeudamiento y la rentabilidad. Por otro

lado, se detectan valores extremos; en algunos casos, los resultados son negativos en ciertos indicadores financieros. Asimismo, existe una diferencia clara según el tamaño de las empresas, ya que las empresas más grandes tienen estructuras financieras más sólidas, mientras que las micro y pequeñas son más vulnerables desde el punto de vista financiero. Los resultados permiten observar que variables como el endeudamiento, la rentabilidad, la liquidez y el tamaño empresarial influyen directamente en la probabilidad de insolvencia en el modelo de regresión logística con regularización Elastic Net. Se evidencia que el endeudamiento es el principal factor de riesgo, ya que a mayor deuda existe mayor probabilidad de quiebra. En cambio, la rentabilidad desempeña un papel protector, ayudando a reducir el riesgo. También el tamaño de la empresa importa, ya que las micro y pequeñas enfrentan una mayor probabilidad de insolvencia en comparación con las empresas más grandes.

Estos hallazgos confirman que la insolvencia en el sector no es un fenómeno aleatorio, sino el resultado de desequilibrios financieros estructurales que pueden ser cuantificados y anticipados mediante herramientas econométricas y de aprendizaje automático.

En cuanto a la capacidad predictiva, los resultados muestran que cada uno de los modelos aporta ventajas distintas, pero complementarias. El modelo MARS obtuvo el mejor desempeño general, ya que logró capturar relaciones no lineales entre las variables financieras, especialmente en indicadores como el endeudamiento y el capital de trabajo, lo que mejoró su precisión al clasificar empresas con y sin riesgo de insolvencia. Por su parte, el modelo Elastic Net aportó mayor robustez en la selección de variables relevantes, permitiendo identificar los factores que más influyen en la probabilidad de quiebra. Además, su mecanismo de penalización ayudó a reducir los problemas de multicolinealidad entre los indicadores financieros, facilitando una interpretación más clara de los resultados. Por último, el modelo Naive Bayes destacó por su alta sensibilidad para detectar empresas en situación de riesgo, lo que lo convierte en una herramienta especialmente útil para sistemas de alerta temprana. En conjunto, estos resultados muestran que la combinación de distintos enfoques de modelización permite obtener una visión más completa y sólida para anticipar escenarios de insolvencia empresarial.

Finalmente, se concluye que la insolvencia empresarial responde a la combinación de debilidades financieras internas y condiciones estructurales del entorno económico. El modelo predictivo desarrollado permite estimar la probabilidad de quiebra con base en

variables financieras clave, constituyéndose en un instrumento técnico relevante para la gestión del riesgo empresarial y la supervisión sectorial. No obstante, el estudio presenta limitaciones relacionadas con la disponibilidad y calidad de la información contable, posibles sesgos en los datos reportados y la exclusión de variables macroeconómicas dinámicas o cualitativas. Como líneas futuras de investigación, se recomienda incorporar variables macroeconómicas, aplicar modelos no lineales adicionales y extender el análisis a otros sectores económicos del país, con el fin de fortalecer la capacidad predictiva y contribuir al diseño de políticas públicas orientadas a la prevención de insolvencias empresariales.

7. Referencias

- Aker, C., & Karavardat, S. (2023). Corporate bankruptcy prediction using financial ratios and machine learning techniques. *Journal of Risk and Financial Management*, 16(4), 1-20. <https://doi.org/https://doi.org/10.3390/jrfm16040172>
- Altman, E. I. (1968). Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. *Journal of Finance*, 23(4), 589-609.
- Aymacaña, A., Cangui, Y., & Urbina, M. (2024). Desafíos de insolvencia financiera durante el covid-19: un análisis de la liquidez y endeudamiento del sector empresarial de la provincia de Cotopaxi en los periodos 2020 – 2021. *ResearchGate*, 8(3), 1-18. <https://doi.org/http://dx.doi.org/10.56048/MQR20225.8.3.2024.3112-3132>
- Banco Central del Ecuador . (2023). *Cuentas nacionales anuales*. Banco Central del Ecuador. <https://www.bce.fin.ec/>
- Bermeo, D., & Armijos, J. (2021). Predicción de quiebra bajo el modelo Z2 Altman en empresas de construcción de edificios residenciales de la provincia del Azuay. *Economía y Política*, 33, 48-63. <https://doi.org/https://doi.org/https://doi.org/10.25097/rep.n33.2021.03>
- Burneo, K., & Lizarzaburu, E. (2016). Gobierno corporativo en mercados emergentes: Evidencia de Lima – BVL. *Globalización, Competitividad y Gobernabilidad*, 23(40), 1-33. <https://doi.org/https://doi.org/10.3232/GCG.2016.V10.N2.03>
- Cruz, S., Gavira, N., & García, R. (2017). Eficiencia de los modelos Poisson y Logístico en la asignación de probabilidades de incumplimiento a empresas mineras mexicanas. *Revista Mexicana de Economía y Finanzas (REMEF)*, 12(1), 1-21. <https://doi.org/https://doi.org/10.21919/remef.v12i1.198>
- Giron, H., Villanueva, J., & Armas, R. (2017). Determinantes de la quiebra empresarial en las empresas ecuatorianas en el año 2016. *Revista UTPL*, 4(13), 108-126. <https://revistapublicando.org/revista/index.php/crv/article/view/780>
- Gitman, L. J., & Zutter, C. J. (2012). *Principios de administración financiera*. Mexico : Pearson Educación.
- Gómez, G. (2005). Insolvencia: evolución de un concepto. *Revista de Derecho Privado*, 27-33. <https://doi.org/repositorio.uniandes.edu.co>
- González, J., & Sánchez, A. (2018). Predicción de quiebra en empresas privadas sanitarias españolas mediante optimización de la Z-Score de Altman. *Gaceta Sanitaria*. <https://doi.org/https://doi.org/10.1016/j.gaceta.2018.01.001>
- Haro, A. (2021). La estructura financiera y el fracaso empresarial: una apreciación a las grandes empresas de pesca y acuicultura. *Revista Ciencia Multidisciplinaria CUNORI*, 5(1), 1-15. <https://doi.org/doi.org/10.36314/cunori.v5i1.148>
- Haro, A., Pico, J., Sánchez, D., Ramírez, S., & Ulloa, J. (2023). Estructura de capital óptima analizada mediante insolvencia empresarial. *Revista Latinoamericana de*

- Hernández, L. (2022). Crisis empresarial en Colombia Probabilidad de entrar de procesos de insolvencia 2016-2019. *CEDE*(14), 1-37. <https://economia.uniandes.edu.co/>
- Hull, J. C. (2015). *Risk Management and Financial Institutions*. Estados Unidos: Wiley.
- Kanapickienė, R., Kanapickas, T., & Nečiūnas, A. (2023). Bankruptcy Prediction for Micro and Small Enterprises: The Case of the Lithuanian Construction Sector. *Risks*. <https://doi.org/https://doi.org/10.3390/risks11050097>
- Lambertons, V. (2023). Predicción del fracaso empresarial utilizando. *REMEF*, 18(3), 1-22. <https://doi.org/doi.org/10.21919/remef.v18i3.729>
- Lozada, O., Sauza, B., Cruz, D., & Lechuga, C. (2024). Anticipando la quiebra: Decisiones informadas a través del análisis de los modelos Altman y Ohlson. *Ingenio y Conciencia. Boletín Científico de la Escuela Superior Ciudad Sahagún*, 11, 17-27. <https://doi.org/https://doi.org/10.29057/escs.v11iEspecial.13681>
- Mehmood, A., & De Luca, F. (2025). Financial distress prediction in private firms: Developing a model for troubled debt restructuring. *Journal of Applied Accounting Research*, 26(2), 206-222. <https://doi.org/https://doi.org/10.1108/JAAR-03-2023-0064>
- Mejía, M., & Flores, J. (2020). Aplicación del Modelo Z-Score de Altman para clasificar niveles de quiebra financiera en el sector comercial de la provincia de Manabí – Ecuador. *593 Digital Publisher CEIT*, 5(1), 26-39. <https://doi.org/https://doi.org/10.33386/593dp.2020.5-1.318>
- Mendaña, C., Remo, N., & Toral, M. (2024). Análisis combinado de factores del fracaso empresarial en el sector turístico español. *Revista de Estudios Empresariales*(2), 237-365. <https://doi.org/https://doi.org/10.17561/ree.n2.2024.8273>
- Molina, P., Flores, K., Flores, C., & Molina, D. (2023). Modelo de predicción de quiebra en empresas de comercio en Ecuador: Uso del modelo logístico de Ohlson. *Ecociencia*, 7(3), 66-81. <https://doi.org/https://doi.org/10.21855/ecociencia.103.812>
- Molina, P., Molina, D., & Flores, C. (2022). Modelo de predicción de quiebra Z2 de Altman de análisis multivariable en empresas del sector inmobiliario de la provincia de Pichincha. *Revista Científica ECOCIENCIA*, 9(2), 53-76. <https://doi.org/https://doi.org/10.21855/ecociencia.92.643>
- Molina, P., Ramírez, A., Flores, K., & Flores, C. (2024). Modelo Ohlson: Un estudio de bancarrota en tiempos de incertidumbre en las cooperativas de ahorro y crédito de Ecuador. *Revista Sigma*, 11(2), 153-168. <https://doi.org/https://doi.org/10.24133/e4925e41>
- Mongrut, S., Alberti, F., Fuenzalida, D., & Akamine, M. (2011). Determinantes de la insolvencia empresarial en el Perú. *Academia. Revista Latinoamericana de Administración*, 47, 126–139. <https://www.redalyc.org/articulo.oa?id=71618917009>

- Moranga, H., & Rossi, M. (2019). Gobierno corporativo y riesgo de quiebra en las empresas chilenas. *Globalización, Competitividad y Gobernabilidad*, 15(28), 52-71. <https://doi.org/https://doi.org/10.3232/GCG.2019.V13.N1.06>
- Ohlson, J. (1980). Financial ratios and the probabilistic prediction of bankruptcy. *Journal of Accounting Research*, 1, 109-131.
- Orellana, I., Pinos, L., Reyes, M., Cevallos, E., & Tonon, L. (2023). Probabilidad de quiebra empresarial en el sector de construcción del Ecuador: Periodo 2011 – 2020. *Economías*, 41, 71-90. <https://doi.org/https://doi.org/10.17981/econcu.44.2.2023.Econ.2>
- Orellana, I., Reyes, M., Tonon, L., & Pinos, L. (2020). Medición de los niveles de riesgos de liquidez, insolvencia y mercado en el sector textil del Ecuador. *Universidad Tecnica de Ambato*, 5(17), 1041-1057. <https://doi.org/https://doi.org/10.31243/bcoyu.27.2020.1014>
- Prastyo, D., Nugroho, A., & Valackienė, A. (2025). Corporate bankruptcy prediction using machine learning techniques and financial ratios. *International Journal of Financial Research*, 16(1), 45-60. <https://doi.org/https://doi.org/10.5430/ijfr.v16n1p45>
- Puente, M., Dávalos, E., Bardoscia, K., & Cáceres, L. (2025). Análisis de la solvencia financiera mediante el modelo Z-Score de las empresas del centro cerámico de Cuenca. *Perspectivas Sociales y Administrativas*, 7(1), 61-73. <https://doi.org/https://doi.org/10.61347/psa.v3i1.77>
- Rivadeneira, J., Saltos, R., Rivera, M., & Carpio, R. (2022). Predicción de la quiebra empresarial en el sector agroindustrial de la ciudad de Machala. *ACI Avances en Ciencias e Ingenierías*, 28(4), 45-60. <https://doi.org/https://doi.org/10.18272/aci.v14i2.2695>
- Roque, D., & Caicedo, A. (2023). Relación entre los indicadores financieros del modelo Altman Z y el puntaje Z. *Retos. Revista de Ciencias de la Administración y Economía*, 13(25), 129-148. <https://doi.org/https://doi.org/10.17163/ret.n25.2023.09>
- Ross, S., Westerfield, R., & Jordan, B. (2018). *Fundamentos de finanzas corporativas*. Estados Unidos : McGraw-Hill Education.
- Shamara, K., & Shinde, A. (2025). Bankruptcy Prediction Using Machine Learning and Data Preprocessing Techniques. *Analitics*, 4(21). <https://doi.org/10.3390/analytics4030022>
- Sierpe, J. I., Merino, M. D., Raposo, A. P., & Vitiello, V. (2021). Algoritmos de Random Forest como alerta temprana para la predicción de insolvencias en empresas constructoras = Random Forest algorithms as early warning tools for the prediction of insolvencies in construction companies. *Anales de Edificación*. <https://doi.org/https://doi.org/10.20868/ade.2021.4766>
- Sigüeza, A., & Orellana, I. (2020). Análisis de riesgo financiero en el sector de fabricación de otros productos minerales no metálicos del Ecuador. *Revista Economía y Política*, 7(32), 1-16. <https://doi.org/https://doi.org/10.25097/rep.n32.2020.05>

- Superintendencia de Compañías, Valores y Seguros. (2024). Información empresarial del Ecuador: <https://bit.ly/420o8II>
- Tuston, J., & Macías, E. (2025). Modelos de machine learning para la detección de fraudes financieros: Una revisión de la literatura. *Revista UNESUM-Ciencias*, 9(2), 220-224. <https://doi.org/doi.org/10.47230/unesum-ciencias.v9.n2.2025.220-234>
- Van Greuning, H., & Bratanovic, S. (2009). *Analyzing Banking Risk*. Estados Unidos: World Bank Publications.
- Villa, V., Velarde, R., Flores, M., & León, J. (2024). Evaluación de la solvencia de fábricas de calzado en Tungurahua: un análisis basado en el modelo Z-Score. *Esprint Investigación*, 3(3), 137-150. <https://doi.org/https://doi.org/10.61347/ei.v3i3.87>