



**INVESTIGACIÓN**

**MODELO PREDICTIVO PARA RIESGO DE QUIEBRA EN EL SECTOR SALUD EN COLOMBIA PARA EL  
PERIODO 2023**

**JUAN CAMILO RAMÍREZ RÍOS**

**JUAN ESTEBAN VÉLEZ RESTREPO**

**PONTIFICIA UNIVERSIDAD JAVERIANA**

**FACULTAD DE CIENCIAS ECONÓMICAS Y ADMINISTRATIVAS**

**MESTRÍA EN FINANZAS**

**SANTIAGO DE CALI**

**2024**

**INVESTIGACIÓN**

**MODELO PREDICTIVO PARA RIESGO DE QUIEBRA EN EL SECTOR SALUD EN COLOMBIA PARA EL  
PERIODO 2023**

**JUAN CAMILO RAMÍREZ RÍOS**

**JUAN ESTEBAN VÉLEZ RESTREPO**

**Trabajo de grado presentado como requisito parcial para optar por el título  
de Magíster en Finanzas**

**Director del trabajo de grado: Víctor Alberto Peña Vargas**

**Profesor Departamento Contabilidad y Finanzas**

**PONTIFICIA UNIVERSIDAD JAVERIANA**

**FACULTAD DE CIENCIAS ECONÓMICAS Y ADMINISTRATIVAS**

**MAESTRÍA EN FINANZAS**

**SANTIAGO DE CALI**

**2024**

Santiago de Cali, 10 de mayo de 2024

Doctor

Fabian Fernando Osorio Tinoco

Decano

Facultad De Ciencias Económicas y Administrativas

Pontificia Universidad Javeriana

Cali

Por medio de la presente estamos entregando a usted el Trabajo de Grado cuyo título es “MODELO PREDICTIVO PARA RIESGO DE QUIEBRA EN EL SECTOR SALUD EN COLOMBIA PARA EL PERIODO 2023”.

Esperamos que este Trabajo cumpla con los requisitos académicos exigidos y que alcance el propósito para el cual fue elaborado.

Atentamente



Juan Camilo Ramírez Ríos

CC: 1144186862



Juan Esteban Vélez Restrepo

CC: 1037640449

Doctor

Fabian Fernando Osorio Tinoco

Decano

Facultad de Ciencias Económicas y Administrativas

Pontificia Universidad Javeriana

Cali

Por medio de la presente me permito comunicarle, que en mi calidad de director de trabajo de grado he leído detenidamente el informe final del estudio titulado **“MODELO PREDICTIVO PARA RIESGO DE QUIEBRA EN EL SECTOR SALUD EN COLOMBIA PARA EL PERIODO 2023”**, realizado por los estudiantes de Maestría en Finanzas de la Facultad de Ciencias Económicas y Administrativas de la Universidad Javeriana nombres: **Juan Camilo Ramírez Ríos C.C. 1144186862 & Juan Esteban Vélez Restrepo C.C. 1037640449**, y considero que cumple con todos los requisitos requeridos para ser presentada a evaluación.

Atentamente



---

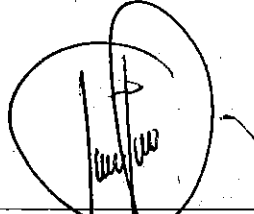
**Víctor Alberto Peña Vargas**

Director del Trabajo de Grado

ARTÍCULO 23 de la resolución N° 13 de julio 6 de 1946

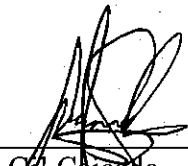
“La Universidad no se hace responsable por los conceptos emitidos por sus alumnos en sus trabajos de Tesis. Sólo velará porque no se publique nada contrario al dogma y a la moral católica y porque la Tesis no contenga ataques o polémicas puramente personales; antes bien, se vea en ellas al anhelo de buscar la Verdad y la Justicia”.

**“MODELO PREDICTIVO PARA RIESGO DE QUIEBRA EN EL SECTOR SALUD EN COLOMBIA PARA EL PERIODO 2023”**, Aprobado por el Comité de Trabajos de Grado en cumplimiento de los requisitos exigidos por la Pontificia Universidad Javeriana para optar por el título de Magíster en Finanzas.



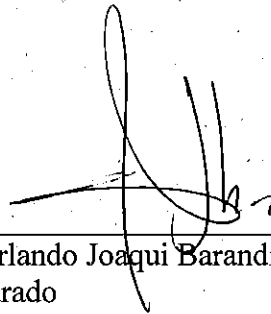
---

Fabian Fernando Osorio Tinoco  
Decano  
Facultad de Ciencias Económicas y Administrativas



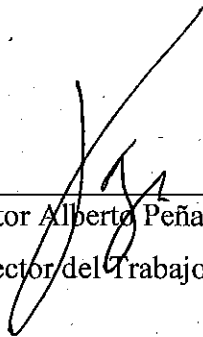
---

Jessica Gil Caicedo  
Directora de Maestría en Finanzas.



---

Orlando Joaquín Barandica  
Jurado



---

Victor Alberto Peña Vargas  
Director del Trabajo de Grado

Santiago de Cali, 11 de julio del 2024

## Contenido

<b>1.</b>	<b>RESUMEN.....</b>	<b>4</b>
<b>2.</b>	<b>ABSTRACT.....</b>	<b>5</b>
<b>3.</b>	<b>INTRODUCCIÓN .....</b>	<b>5</b>
<b>4.</b>	<b>PLANTEAMIENTO DEL PROBLEMA .....</b>	<b>10</b>
<b>6.</b>	<b>JUSTIFICACIÓN.....</b>	<b>13</b>
<b>7.</b>	<b>PREGUNTA DE INVESTIGACIÓN.....</b>	<b>13</b>
<b>8.</b>	<b>MARCO TEORICO .....</b>	<b>14</b>
<b>9.</b>	<b>METODOLOGÍA.....</b>	<b>20</b>
<b>9.1.</b>	<b>CARACTERÍSTICAS DE LA MUESTRA.....</b>	<b>20</b>
<b>10.</b>	<b>RESULTADOS .....</b>	<b>36</b>
<b>12.</b>	<b>REFERENCIAS.....</b>	<b>60</b>

## Índice de tablas

<b>Tabla 1.</b> División de datos.....	25
<b>Tabla 2.</b> Resumen de indicadores financieros en las empresas en quiebra.....	40
<b>Tabla 3.</b> Resumen indicadores financieros de empresas sanas.....	43
<b>Tabla 4.</b> Validación cruzada Modelo Logístico - Métricas .....	46
<b>Tabla 5.</b> Matriz de confusión Modelo Logístico .....	47
<b>Tabla 6.</b> Métricas datos validación cruzada Modelo Logístico.....	47
<b>Tabla 7.</b> Coeficientes Modelo Logístico.....	48
<b>Tabla 8.</b> Validación cruzada Modelo SVM - Métricas.....	50
<b>Tabla 9.</b> Matriz de confusión Modelo SVM.....	51
<b>Tabla 10.</b> Métricas datos de validación .....	51
<b>Tabla 11.</b> Validación cruzada Modelo Red Neuronal - Métricas .....	51
<b>Tabla 12.</b> Matriz de confusión Modelo Red Neuronal .....	52
<b>Tabla 13.</b> Métricas datos de validación Modelo Red Neuronal.....	53
<b>Tabla 14.</b> Resumen niveles de precisión de los tres modelos .....	53
<b>Tabla 15.</b> Métricas Modelo Logístico – Probabilidad muestra de empresas sanas.....	55

## Índice de figuras

<b>Figura 1. Validación cruzada .....</b>	<b>26</b>
<b>Figura 2. Matriz de correlación.....</b>	<b>36</b>
<b>Figura 3. Equilibrio de clases .....</b>	<b>38</b>
<b>Figura 4. Curva ROC.....</b>	<b>54</b>
<b>Figura 5. Distribución de las probabilidades quiebra de empresas sanas.....</b>	<b>55</b>

## 1. RESUMEN

Este trabajo de investigación presenta un análisis exhaustivo sobre la predicción de riesgos de quiebra en el sector salud en Colombia, centrandose en la evaluación de la solvencia financiera de las Instituciones Prestadoras de Salud. Frente a un escenario de crecientes desafíos económicos y financieros, exacerbados por la pandemia de COVID-19 y la integración de poblaciones migrantes, el sector salud colombiano enfrenta una crisis de insolvencia que amenaza su estabilidad y continuidad. La investigación se fundamenta en un robusto conjunto de datos proporcionado por la Supersalud, abarcando información financiera de 1335 empresas hasta el año 2022, con un enfoque particular en aquellas en proceso de liquidación.

Utilizando metodologías avanzadas de análisis, se evalúan y comparan tres modelos predictivos: Regresión Logística, Máquina de Vectores de Soporte (SVM) y Redes Neuronales. Cada modelo se sometió a un riguroso proceso de validación cruzada y se analizó mediante una serie de métricas de rendimiento, incluyendo precisión, sensibilidad y especificidad, para determinar su capacidad predictiva. Los resultados indican que, mientras cada modelo ofrece insights valiosos, ciertas técnicas proporcionan un mejor entendimiento y predicción de los riesgos financieros inminentes.

Este estudio contribuye significativamente a la literatura existente al ofrecer una metodología replicable y un modelo predictivo optimizado que puede ser utilizado por los reguladores y administradores del sector salud para realizar intervenciones proactivas y fundamentadas. Asimismo, el análisis proporciona una base empírica para la formulación de políticas públicas que aspiren a fortalecer la resiliencia financiera del sistema de salud en Colombia.

## **2. ABSTRACT**

The financial stability of the health sector in Colombia faces unprecedented challenges, exacerbated by socioeconomic and public health factors such as the COVID-19 pandemic and the growing migrant population. This study addresses the urgent need to predict insolvency in Health Service Entities (EPS) through the application of advanced predictive models. Utilizing data from Supersalud, which includes financial information from 1,335 companies up to 2022, three predictive models were evaluated: Logistic Regression, Support Vector Machine (SVM), and Neural Networks. Each model underwent rigorous cross-validation and was compared using performance metrics such as accuracy, sensitivity, and specificity. The results highlight the models' ability to identify early signs of financial risk, providing valuable tools for proactive decision-making. This study not only contributes to academic literature, offering a framework for future research, but also provides a critical resource for policymakers and health sector administrators in implementing strategies to ensure the continuity and effectiveness of medical care in Colombia.

## **3. INTRODUCCIÓN**

La vitalidad financiera de las empresas, esencial para el dinamismo económico, propulsa la generación de empleo y aporta significativamente al erario. En este contexto, el Banco Mundial resalta cómo el acceso universal a servicios de salud emerge como un pilar en la batalla contra la pobreza (Banco Mundial, 2024). Por ende, el sector salud de Colombia, fortalecido tras la implementación de la Ley 100 de 1993 que inauguró un sistema de seguridad social abarcador, ha logrado un hito con una cobertura de salud del 99% hacia 2022. Este avance, no solo contribuye con el 5.1% al PIB, sino que también subraya su papel indispensable en el tejido socioeconómico del país (Melo-Becerra et al., 2023a).

Gracias a su atractiva combinación de costos accesibles y servicios certificados de alta calidad, Colombia se ha consolidado como un competidor clave en el ámbito del turismo médico, respaldado por entidades de salud robustas y especializadas. La demanda por estos servicios ha mostrado un crecimiento sostenido, con la excepción del periodo de la COVID-19, donde las restricciones de viaje impactaron temporalmente el sector (López Rodríguez et al., 2023). Además, según datos de la OCDE a pesar de que Colombia tiene un gasto en sanidad por debajo de la media de países de la entidad, este representa un 8,1% del PIB para el año 2022, siendo este un rubro significativo dentro de la cartera de la Nación (*Health at a Glance 2023*, 2023). En donde se tienen dos grandes grupos de financiadores del sector, por un lado, quienes cotizan al régimen contributivo y por el otro los recursos fiscales del presupuesto general de la nación.

A pesar de sus contribuciones, el sector salud enfrenta retos, destacando la crisis del Covid-19, el desequilibrio entre regímenes contributivo y subsidiado —exacerbado por la inclusión de migrantes venezolanos—, el incremento en costos de servicios, la insolvencia de entidades sanitarias, y una mayor demanda ligada al envejecimiento poblacional, lo que presiona al Estado a aumentar su inversión en salud (Melo-Becerra et al., 2023b).

Las empresas que ingresan al proceso de reestructuración de la ley 1116 de 2006 de la SuperSociedades, lo hacen por la incapacidad de cumplir con sus obligaciones financieras y deuda con terceros, definido como la cesación de pago por más de noventa días de dos o más obligaciones a favor de dos o más acreedores, o que tenga mínimo dos demandas de ejecución presentadas por dos o más acreedores para el pago de deudas, en este caso el valor acumulado de estas obligaciones deberá ser igual o superior al diez por ciento del pasivo total de la compañía (Ley 1116, 2006). Esta definición se asocia con el concepto de quiebra, que ha sido definido por diferentes autores entre ellos (Beaver, 1966a), quien define quiebra como la incapacidad de pago de las deudas vencidas posteriormente fue definida como el momento en que la empresa ha dejado de pagar sus obligaciones (Šverko Grdić et al.,

2009). A pesar de que las IPS se entienden que buscan un fin lucrativo basado en relaciones económicas dentro del sector “prestación de servicios para la atención de la salud humana”, y que cuentan con una estructura empresarial que se rige bajo lineamientos de objeto comercial, no son reguladas por la Supersociedades.

Las IPS sean privadas o de naturaleza mixta, siguen la regulación de la superintendencia de salud, y este es el sujeto que compete con la intervención y liquidación de las entidades vigiladas. Ahora bien, el artículo 114 del Estatuto Orgánico del Sistema Financiero, modificado por el artículo 32 de la Ley 795 de 2003, establece las causales de toma de posesión inmediata de una entidad vigilada (Estatuto Orgánico Financiero, 2003), (Resolución 2022320030005874 - 6, 2022). Dentro de las situaciones, causales o características previstas en la anterior disposición se encuentran: cuando la entidad haya suspendido el pago de sus obligaciones, cuando incumpla reiteradamente los lineamientos de la Superintendencia Nacional de Salud, cuando persista en manejar los negocios en forma no autorizada o insegura, cuando su patrimonio se reduzca por debajo del 50% del capital suscrito, cuando existan inconsistencias en la información que suministra a la Superintendencia que a juicio que dicho órgano no permita conocer la situación real de la entidad, entre otras. Por consiguiente, existe una determinación sobre los procesos de intervención y liquidación de IPS teniendo en cuenta la naturaleza comercial de la misma con su objeto de atención según el servicio regulado. Por lo tanto, se extiende el enfoque sobre la causal de disolución en el que se establece la suspensión de la actividad social, finiquitar la operación hasta llegar a la liquidación empresarial de acuerdo con el artículo 218 del código de comercio. En el enfoque se pretende implementar de manera cohesionada sobre otros artículos en los que se establece una condición puntual para ser catalogada en disolución, y es el concepto de las pérdidas acumuladas que reduzcan su patrimonio neto a menos del 50% del capital suscrito (Código de Comercio, 1971).

Esta investigación se centra en un tema de gran relevancia: la predicción de la quiebra en compañías prestadoras de servicio de salud para el año 2023, información relevante para el periodo ya

que para el 2024 se estiman aumentos superiores al 10% en los costos de los servicios de salud (Iñiguez, 2023), con el agravante de una mala situación financiera de algunas entidades de salud que las ha llevado a levantarle la mano al gobierno para solicitar recursos, y donde cuyo gremio ha indicado que de lo contrario no podrán garantizar el acceso a los servicios de salud de calidad (Procuraduría General de la Nación, 2023), y sumado a una posible reforma a la salud que pone en duda el futuro del sector (Orozco, 2024).

En el contexto actual, el sector salud en Colombia enfrenta una crisis financiera sin precedentes, exacerbada por los efectos prolongados de la pandemia de COVID-19, tensiones sociales y fluctuaciones económicas a nivel global. Esta situación ha llevado a entidades significativas como EPS a solicitar su liquidación, evidenciando una tendencia preocupante hacia la insolvencia que podría extenderse a lo largo del sector. El interés en este tema surge de la observación de numerosas empresas del sector que han ingresado a proceso de liquidación, regulado por la ley de Supersalud, indicativo de posibles dificultades financieras que conllevan a la necesidad de aplicar modelos predictivos eficientes para anticipar y, en lo posible, mitigar situaciones de cesación de pagos. Identificar tempranamente las señales de riesgo financiero permitiría a las entidades y reguladores implementar medidas correctivas proactivas, evitando el deterioro de la infraestructura de salud y las consecuencias sociales asociadas, como el desempleo e incluso un deterioro en la calidad de la atención y su alcance. En este sentido, el desarrollo y aplicación de modelos predictivos eficaces constituyen un pilar fundamental para la reforma y fortalecimiento del sistema de salud colombiano, promoviendo un entorno más estable y equitativo en la prestación de servicios médicos y destacando la urgencia y relevancia de este estudio en el actual contexto económico.

En este sentido, es importante resaltar que algunos investigadores han aplicado algunos modelos para predecir la quiebra de empresas del sector en Colombia, como es el caso de un modelo Logit utilizado para modelar el riesgo de insolvencia de las empresas del sector salud, con la información

financiera del periodo 2010-2013 (Támara Ayús et al., 2019). Después, en el año 2020 se aplicó nuevamente un modelo Logit para predecir la insolvencia de entidades prestadoras de servicios de salud privadas (Gómez, 2020). Luego un modelo Altman para predecir la quiebra en los hospitales del país según su clasificación NIIF, con información hasta el 2019 (Mendoza, 2020). Por último, Viana (2021) utilizó un método de análisis discriminante múltiple para predecir la quiebra en los hospitales del país, con información financiera obtenida hasta el 2020. Entre otros que han aplicado modelos de predicción de quiebra para el sector, pero no se ha desarrollado una investigación para determinar el modelo más eficiente aplicado a las empresas del sector, con sus condiciones.

La revisión de estudios anteriores sobre la predicción de quiebra en el sector salud en Colombia muestra limitaciones significativas en los modelos empleados, como los basados en Logit y Altman, que no reflejan adecuadamente las condiciones económicas actuales ni integran variables críticas post-pandemia o el impacto de la migración venezolana. Estos modelos históricos también han omitido factores como cambios en las políticas de financiamiento y regulación gubernamental, lo que puede llevar a predicciones inexactas y limitar su utilidad en proporcionar alertas tempranas eficaces. Este nuevo estudio justifica la necesidad de desarrollar un modelo actualizado que incorpore estas variables emergentes y utilice técnicas analíticas avanzadas para ofrecer una herramienta predictiva más robusta y precisa, mejorando la toma de decisiones estratégicas en el sector salud.

El objetivo general de este estudio es evaluar el modelo predictivo de quiebra más efectivo para las empresas del sector salud en el contexto actual. Para lograr esto, se realizará un análisis concienzudo de diversas metodologías para predicción de quiebra lo que permitirá seleccionar el más adecuado.

Metodológicamente, este estudio se apoya en las bases de datos provenientes de la Supersalud de donde recopilamos la información del registro de las entidades financieramente sanas y en proceso de liquidación, y los estados financieros de las empresas del sector desde el 2018 hasta el 2021, que se

complementa con la descarga en la plataforma EMIS con los estados financieros para el 2022, esto con un tamaño de muestra de 1335 empresas, complementado con análisis de software especializado en analítica de datos realizado en Python, para evaluar y seleccionar el modelo de predicción de quiebra más adecuado para el sector salud, entre redes neuronales, Logit y vectores discriminantes. Además, se integra un enfoque cualitativo que considera el entorno macroeconómico y microeconómico, así como factores específicos del sector. Con el limitante de que solo se encontraron los estados financieros completos de 45 empresas en estado de liquidación.

El marco teórico aborda la problemática de quiebra y la reestructuración empresarial desde una perspectiva histórica y contemporánea, incluyendo teorías sobre el fracaso empresarial y modelos predictivos de quiebra. Este enfoque multidimensional permite no solo entender mejor las dinámicas actuales del sector de la salud en Colombia, sino también contribuir al cuerpo académico con nuevos conocimientos y herramientas para la gestión de riesgos financieros en tiempos de incertidumbre.

Tras abordar la importancia de la solidez financiera empresarial y los desafíos específicos del sector salud en Colombia, esta investigación se sumerge en un análisis profundo para comprender y prever la cesación de pagos en dicho sector. A continuación, se detalla el planteamiento del problema, justificando la relevancia del estudio y formulando preguntas de investigación cruciales. Se establece un marco teórico robusto, seguido por una metodología detallada, objetivos claros, y la presentación de posibles resultados. Posteriormente, se analizan estos resultados, concluyendo con reflexiones críticas y recomendaciones específicas, esenciales para Stakeholders y formuladores de políticas.

#### **4. PLANTEAMIENTO DEL PROBLEMA**

En el contexto económico y empresarial colombiano, el sector salud se ha posicionado como una de las actividades de mayor impacto en la economía colombiana, lo que ha llevado a convertirse en parte fundamental para el bienestar social y económico del país.

Desde la implementación de la Ley 100 de 1993, el sector salud en el país se ha ido desarrollando con fuertes problemas de financiación, en ocasiones limitando el acceso a servicios de calidad, justificados en algunos casos por la necesidad de las entidades privadas de generar utilidades (Echeverri, 2008). Además, de la relación existente que hay entre la salud financiera de las entidades prestadoras de servicio y la calidad del servicio prestado (Dubas-Jakóbczyk et al., 2021). Predominando la lógica del mercado y priorizando los intereses de privados (Yepes Luján, 2010). Como es el caso de un estudio realizado para entidades de cuidados intensivos en Estados Unidos, donde se demuestra que los hospitales con mejor situación financiera brindan servicios de mejor calidad (Dong, 2015), además, en una serie de entrevistas realizadas a personas relacionadas a la prestación de salud, indicaron que uno de los factores importantes para la calidad en la prestación del servicio es la solidez financiera (Mosadeghrad, 2014). Pero también, por las condiciones económicas del país en donde por la cantidad de personas informales y con pocos recursos se requiere mayor participación del estado mediante los aportes al régimen subsidiado, donde según cifras del Dane, para el 2022 el 94,7% de la población residente en el país se encuentra afiliada al sistema de seguridad social en salud, de las cuales el 54,8% están en el régimen subsidiado (DANE, 2023a).

Adicionalmente, para los últimos años estos problemas se han ido agravando con situaciones como el Covid-19 con todas sus implicaciones como los aumentos en los servicios y los costos asociados, el aumento de afiliados extranjeros debido a la migración de venezolanos al país por la crisis en Venezuela que según cifras de Migración Colombia pasaron de 23.573 en el 2014 a 2.875.743 para el 2023 (Migración Colombia, 2023), los aumentos en los costos en el sector a nivel mundial, el envejecimiento de la población por el aumento de la esperanza de vida de los Colombianos que mientras en los años 90 estaba por debajo de los 70 años para el año 2020 estaba alrededor de los 75 años (DANE, 2022); y la reducción de la natalidad colombianos que paso de 669.137 en el 2014 a 573.625 para el 2022 (DANE, 2023b); lo que requiere mayor atención en las mayores edades, entre otros

(Melo-Becerra et al., 2023a). Sin hablar de los problemas de corrupción que afectan las diferentes entidades del sector (Comunicados de prensa - Transparencia por Colombia, 2024).

El problema central de esta investigación se centra en la creciente incidencia de insolvencia financiera entre las Entidades Prestadoras de Salud, que compromete la continuidad y calidad del servicio de atención médica. Este estudio se propone investigar los factores determinantes de la insolvencia financiera en estas entidades, utilizando un enfoque de modelado predictivo para identificar señales tempranas de riesgo y permitir intervenciones proactivas. La clarificación de este problema desde el principio subraya la importancia del estudio para los formuladores de políticas, los gestores de las EPS y los académicos, al proporcionar un marco predictivo que permitirá conocer la realidad de la situación financiera del sistema de salud colombiano. Así mismo, se incluirán las investigaciones pertinentes sobre la insolvencia en Colombia y se indexarán los datos económicos actualizados que integren una cantidad de variables, que generen la mayor aproximación del resultado en el modelo propuesto.

## **5. OBJETIVOS DE LA INVESTIGACIÓN**

### **5.1. Objetivo General**

Evaluar el modelo de predicción de quiebra más efectivo para las empresas del sector salud en el periodo 2022.

### **5.2. Objetivos Específicos**

- Caracterizar los grupos de empresas en quiebra.
- Diagnosticar las compañías del sector salud que quebraron en el periodo del 2018 al 2022.
- Evaluar 3 diferentes modelos de predicción de quiebra para el sector salud.
- Estudiar la condición actual del sector salud en Colombia.

## **6. JUSTIFICACIÓN**

Este estudio es esencial para entender y mitigar los riesgos de insolvencia en el sector salud en Colombia, un sector crítico para la estabilidad económica y el bienestar social, particularmente en un periodo de inestabilidad económica global. Al anticipar la quiebra, se habilitan decisiones proactivas que aseguran la continuidad operativa y evitan repercusiones negativas en la salud pública del país. Además, este análisis proveerá a los empresarios herramientas basadas en datos concretos para anticipar y gestionar riesgos financieros, lo cual es vital para la supervivencia y prosperidad empresarial en tiempos de incertidumbre.

Adicionalmente, el proyecto tiene como objetivo enriquecer la literatura académica con modelos predictivos que son pertinentes al contexto colombiano y sobre todo en el sector salud, proporcionando un marco de referencia útil para futuras investigaciones en otros sectores y países de la región. Esto potencia la comprensión de las dinámicas financieras y fomenta el desarrollo de estrategias más efectivas para enfrentar futuras crisis económicas o pandemias.

En última instancia, al predecir y, por ende, prevenir la quiebra empresarial, no solo se protege el empleo dentro de este sector vital, sino que también se contribuye significativamente al bienestar social, aumentando la resiliencia del sector ante futuras crisis. Este enfoque integral no sólo salvaguarda la infraestructura económica del país, sino que también promueve un entorno más estable y seguro para la salud pública.

## **7. PREGUNTA DE INVESTIGACIÓN**

¿Cuál es el método de predicción de quiebra más idóneo para las empresas del sector salud en Colombia?

## 8. MARCO TEORICO

El estudio de la cesación de pagos y la reestructuración empresarial en el sector salud en Colombia se fundamenta en varias teorías y modelos analíticos que abordan el fracaso empresarial y la predicción de quiebras. Este marco teórico se construye sobre dos pilares fundamentales las teorías del fracaso empresarial y los modelos predictivos de quiebra.

Para entender el concepto de quiebra se toman como referencia a Beaver, quien define quiebra como la incapacidad de pago de las deudas vencidas (Beaver, 1966b), y posteriormente a Šverko quien la definió como el momento en que la empresa ha dejado de pagar sus obligaciones (Šverko Grdić et al., 2009)

En los últimos 3 años el país se ha enfrentado a las llamadas “Crisis Superpuestas” que van desde la emergencia generada por el covid-19 que trajo consigo un cierre de la economía en el 2020, el estallido social con paros y bloqueos en las principales vías del país, y la incertidumbre política (Argáez & Rodríguez, 2023). Además de la guerra de Ucrania y las implicaciones que ha tenido sobre la inflación y las altas tasas de interés. (CEPAL, 2022)

Sin duda, las empresas a través de su historia han atravesado diversas crisis que impactan la economía en el mundo y en cada país, lo que ha llevado a las entidades a realizar una búsqueda exhaustiva de un modelo que les permita predecir el futuro y de esta manera, equiparse con todas las herramientas empresariales para hacer frente a las consecuencias de la contracción de la economía. En ese sentido, existen teorías aplicadas a los procesos de reorganización empresarial que provienen de la quiebra o reestructuración de las compañías, o que han sido objeto de estudio a partir de los procesos de recuperación de la operación, normalización de los pasivos y sus relaciones comerciales, entre otras condiciones. Como primera instancia, en la teoría del fracaso empresarial se explica una lógica que es compatible con el desarrollo de las empresas de diferentes sectores económicos y que dada su

flexibilidad o su rigidez ante la incertidumbre, no es un garante de la subsistencia de la misma (Dupleix, 2021), lo cual indica que las compañías deben tener mayor susceptibilidad ante los cambios de la sociedad, el mercado, la política y todos los factores internos y externos que comprende el entorno de las compañías. Esta apreciación conlleva a analizar el fracaso empresarial a partir del deterioro y reestructuración organizacional en empresas colombianas y específicamente en compañías bogotanas durante los años 1995 al 2010 en el que se analizaron las condiciones y reacciones ante la incertidumbre de la economía en los años 90 en los cuales, mediante un modelo de orden cualitativo se expuso la heterogeneidad no observable de los entes económicos en el periodo de tiempo mencionado, y en el cual se confirman los datos recogidos con la hipótesis que a partir de la reestructuración de las compañías se derivan gran parte de las quiebras empresariales (Espejo, 2014).

Como medición de la crisis económica en donde Colombia ha sido impactada en su historia más reciente, la prospección del riesgo operativo en las MiPymes es de vital importancia dado su gran participación en la producción interna y el gran número de entidades que representan ante el sector empresarial privado. Siendo los patrones detectados a partir de la metodología Rough en los índices financieros y el desempeño operativo de las MiPymes, derivó en una tendencia denominada “Prospección Financiera” la cual es la determinación anticipada del riesgo operativo en las empresas, principalmente de empresas micro, pequeñas y medianas. La investigación determinó que alrededor del 30% de las empresas se consideran plenamente sanas financieramente hablando, en tanto que el 29% de las empresas son propensas a quiebra, el 20% podrían continuar financieramente sanas o quebrar y el 18% no cumplen con las reglas, por lo que es complicado determinar su futuro financiero (González García et al., 2017).

El poder predecir la quiebra le permite tanto a los gerentes de una compañía el tomar acciones para evitarlo, así como a los stakeholders tomar decisiones más efectivas en cuestión de inversión o financiación (Bărbuță-Mișu & Madaleno, 2020) en donde cada vez se buscan tener modelos de

predicción más precisos utilizando diferentes métodos (du Jardin, 2010). Predecir la quiebra empresarial es una prioridad para los diferentes actores del mercado desde la crisis supprime (Lombardo et al., 2022), aunque es importante mencionar que en muchas ocasiones los empresarios no usan los modelos más precisos, sino que recurren a modelos mucho más sencillos, por lo que resulta de gran importancia poder enfocarse en los sectores y educar a los gerentes de su importancia (Lesáková et al., 2020).

Desde el principio del siglo XX se han venido desarrollando modelos de predicción de quiebra, los cuales comenzaron siendo modelos estadísticos que analizaban los indicadores financieros, resaltando el modelo Z-score de Altman, el cual aún sirve de referencia (Altman, 1968), después, a partir del trabajo del modelo de O-score de Ohlson se empezaron a desarrollar modelos más complejos (Kitowski et al., 2022). En los últimos años se tienen modelos cada vez más precisos soportados en modelos de inteligencia artificial y redes neuronales (Valaskova et al., 2023).

En una revisión de 128 publicaciones con respecto a la predicción de quiebra de las empresas publicados entre 1968 y 2005, se encontró que con ese objetivo se habían utilizado casi todas las técnicas inteligentes disponibles a ese momento para hacer las predicciones, además, se determinó que la técnica más usada en los últimos años era la de redes neuronales (Ravi Kumar & Ravi, 2007).

Es importante resaltar que las diferentes condiciones de cada industria hacen necesario un modelo de predicción de quiebra para cada uno de ellos, logrando ser más preciso que un modelo general (Chava & Jarrow, 2004).

Con respecto a la viabilidad de las empresas que se encuentran bajo ley de insolvencia, se abarca una de las diversas investigaciones sobre las variables financieras que influyen en el fracaso empresarial de las empresas de pequeño y mediano tamaño en Colombia. En los resultados, se muestran las principales variables financieras con un mayor apalancamiento a largo plazo y un alto pasivo total sobre la estructura de balance; también indican la baja concentración de activos líquidos

con relación a su activo total, que no alcanzan a cubrir los pagos de la deuda. Por lo tanto, se concluye que los principales problemas de las empresas que fracasan en la industria provienen de la falta de liquidez y un alto endeudamiento, lo cual es consecuente con las causas de insolvencia bajo la que se rige la Superintendencia de Sociedades Colombiana (2012) en las que señala un alto endeudamiento y contracción en ventas, derivando en el ajuste en su liquidez (Espinosa et al., 2015).

Así mismo, es importante traer otra investigación relevante al caso para modelar la situación financiera y predecir la probabilidad de quiebra en empresas bajo régimen de insolvencia. Dicha investigación ha sido pionera en considerar la regulación de insolvencia para explicar los indicadores financieros a considerar en el proceso concursal, y se implementa el algoritmo que impide el sesgo en empresas consultadas. El modelo arroja como resultado los indicadores de rentabilidad como los predictores más precisos en insolvencia de las empresas de la muestra (Mejía & Castaño, 2020).

El fracaso empresarial es analizado también mediante un modelo Fulmer, el cual fue creado en el año 1984 y se enfoca en la capacidad de una empresa para generar liquidez como capital de trabajo e inversión (solventia). Al aplicar este modelo al sector empresarial colombiano con datos del año 2011 a partir de las empresas vigiladas por la Superintendencia de Sociedades se encontró que el 88,25% de las compañías se encuentran en riesgo de insolvencia. La estimación del resultado indica que es efectivo en cuanto a la predicción, sin embargo, se recomienda modificar las variables de la realidad colombiana y su contexto, así como la revisión del segmento del endeudamiento, las utilidades acumuladas y las utilidades operativas o el disponible para generar un mejor diagnóstico al corto plazo. Se encuentra que las condiciones financieras también dependen del tamaño de la empresa por lo cual se sugiere ajustar de manera estructural la clasificación de los datos para un mayor detalle en los resultados que se desean obtener (Galán-Barrera & Torres-García, 2017).

En un modelo desarrollado para evaluar los determinantes financieros para el fracaso de las empresas Pymes en Colombia, se encontró que los determinantes claves fueron: el porcentaje de costos y de gastos administrativos y ventas de la operación (Espinosa et al., 2015).

Después, en uno de los modelos desarrollados para predecir la quiebra en Colombia, se utilizó un modelo Logit para modelar la quiebra en empresas del sector evaluando los periodos entre 2009 y 2019, donde se obtuvo un modelo con una probabilidad de éxito del 98%, soportado en los indicadores del ROA, el pasivo total sobre el total de activos y el patrimonio sobre el total de los activos (Ayús et al., 2023). Años atrás este modelo Logit se comparó contra las máquinas de motores de soporte para predecir el incumplimiento de pago de empresas no financieras de Polonia, donde se validó que el modelo Logit podía pronosticar mejor el incumplimiento de pago (Nehrebecka, 2018).

También, en otro análisis para predecir la quiebra en el sector de alimentos y bebidas en una muestra de empresas listadas en la bolsa de valores de Indonesia, en el periodo entre el año 2016 y 2019, se probaron los modelos Altman, Springate y Grover y se identificó que para este caso el modelo más preciso fue el de Springate (Nasmi & Afriyenti, 2021).

En un entorno donde cada vez se tiene una mayor cantidad de datos, se pueden tener modelos más precisos, como es el caso donde utilizando el algoritmo XGBoost, algoritmo de aprendizaje automático usado en tareas de regresión y clasificación. Que, mediante árboles de decisión mejorados utiliza una técnica llamada "boosting" para producir un modelo predictivo muy eficaz y preciso, con el que se obtuvieron predicciones más precisas (Ben Jabeur et al., 2023).

En la época de COVID las empresas tuvieron un respiro con respecto al costo de la deuda debido a que muchos bancos centrales en el mundo bajaron las tasas de interés a niveles mínimos, impulsando la economía y haciendo frente a los efectos del COVID (Topcu & Gulal, 2020). A pesar de esto muchas empresas se vieron muy afectadas financieramente por lo que se desarrollaron diferentes estudios para

evaluar su impacto. En uno de estos, por medio de la aplicación de un algoritmo de bosque aleatorio a los estados financieros de empresas de 25 países, se determinó que las variables más influyentes en la predicción de quiebra fueron el nivel de endeudamiento, los años de antigüedad de la empresa y el acceso al crédito (Bozkurt & Kaya, 2023). Para el caso de las empresas de Estados Unidos, se encontró que el modelo XGBoost lograba predecir la quiebra en las empresas de este país en el segundo semestre del 2020 (Narvekar & Guha, 2021). Finalmente, en un estudio se revisó la confiabilidad de la predicción de quiebra de las PYMES en la época del COVID y en periodos anteriores, analizando más de 90 mil empresas "Utilizando métodos de predicción basados en árboles de decisión a través de técnicas de 'boosting' como CatBoost, LightGBM y XGBoost, determinaron que en el periodo de crisis por pandemia la predicción de los modelos fue mucho más débil (Papík & Papíková, 2023).

En el caso del sector salud en los últimos años se han desarrollado múltiples análisis tanto nacional como internacionalmente para predecir la quiebra de las entidades prestadoras de servicios de salud. En estudios previos, para el caso de los hospitales de cuidados intensivos en Estados Unidos, se realizó una evaluación de 3 modelos diferentes, Altman Z modificada, Ohlson y Zmijewski, evaluando los estados financieros de 106 entidades, con la mitad de ellas sana y el restante en quiebra, determinando que el modelo más efectivo para este caso era el de Altman e identificando que no había un indicador único que pudiera anticipar un proceso de quiebra, si no que se explicaba por diversos indicadores, todo esto con información financiera entre el 2006 y el 2017 (Puro et al., 2019). En el caso de hospitales rurales del mismo país, se desarrolló un análisis para predecir el índice latente de dificultades financieras, por medio del cual se le asignó un nivel de riesgo de quiebra a las diferentes clínicas evaluadas (Holmes et al., 2017). En un análisis más reciente se aplicaron los modelos máquinas de vectores de soporte lineal, regresión logística y redes neuronales para hallar los factores que más influyen en el estado de quiebra de hospitales en los Estados Unidos, con información financiera del año 2019 correspondientes a 3121 entidades, de las cuales solo 27 quebradas, siendo ingresos netos por

paciente el principal factor seguido por el monto de cuentas por cobrar (Beauvais et al., 2023). En otro estudio realizado en el año 2017, se intentó identificar los indicadores financieros y no financieros más influyentes para predecir la insolvencia de un hospital, con indicadores como Altman Z y Ratio de cobertura de servicio de la deuda dentro de los indicadores financieros y, duración media de la instancia y tasa de ocupación entre los indicadores operativos, encontrando correlaciones entre la solvencia e insolvencia con estos indicadores (Biniek & Gossett, 2017).

## **9. METODOLOGÍA**

En este capítulo se presentan aspectos metodológicos de este estudio como la recopilación y estructuración de la información financiera de las empresas, la selección de las variables independientes, la descripción de los modelos, el análisis descriptivo, la aplicación de los modelos utilizando la herramienta Python, atendiendo a la pregunta de investigación del presente trabajo, en conexión con el objetivo general y objetivos específicos.

### **9.1. Características de la muestra**

Los datos se obtuvieron de las bases de datos de la Supersalud, descargando información financiera de empresas del sector salud entre sanas y en proceso de liquidación, con estados financieros desde el 2018 al 2021, complementando con la información encontrada en EMIS del 2022 para las mismas empresas, de la cuál mediante una meticulosa revisión se seleccionaron las empresas que tuvieran estados financieros. Esta selección permitió obtener un conjunto representativo de datos financieros de 1335 empresas, con el limitante de solo poder tener estados financieros completos de 45 empresas en situación de quiebra ya que el resto están operativas.

Llevar a cabo una investigación y realizar estimaciones con un número relativamente bajo de empresas en quiebra presenta varias limitaciones que afectan la robustez del modelo predictivo. Como

primera medida, la presencia de un menor número de empresas en quiebra puede introducir un sesgo significativo hacia la clase mayoritaria (empresas "sanas"), lo que disminuye la capacidad del modelo para generalizar y detectar correctamente las empresas en riesgo de quiebra. Esto puede resultar en un modelo que es altamente preciso en predecir la clase mayoritaria, pero ineficaz en identificar los casos críticos de quiebra.

Además, la muestra pequeña de empresas en quiebra puede llevar a un modelo menos robusto y más propenso a errores de clasificación. La falta de datos suficientes sobre la clase minoritaria limita la capacidad del modelo para aprender y reconocer los patrones asociados con el riesgo de quiebra.

Para mitigar estas limitaciones, una opción comúnmente utilizada es la implementación de técnicas adicionales de recolección de datos para aumentar el tamaño de la muestra de empresas en quiebra. Otra de las opciones es el uso de técnicas de aprendizaje no supervisado, las cuales pueden identificar patrones en datos no etiquetados, proporcionando información adicional que podría ser indicativa de quiebra. Estas herramientas pueden ayudar a mejorar la robustez del modelo y su capacidad para predecir con mayor precisión las empresas en riesgo de quiebra.

## **9.2. Análisis descriptivo**

De acuerdo con la bibliografía consultada, se seleccionan los indicadores para predecir la quiebra: días de cartera, días de proveedores, ciclo de conversión de efectivo, endeudamiento total, apalancamiento, deuda corriente sobre operación, relación corriente, cobertura del gasto financiero, razón de tesorería, margen operacional, margen neto, ROA, ROE, cobertura de servicio de la deuda y el indicador de solvencia. Con estos indicadores se procedió con un análisis descriptivo, analizando las correlaciones, promedios, desviación estándar y los cuartiles y un análisis detallado de estos indicadores y su comportamiento en el sector.

## **9.3. Definición de los modelos**

### 9.3.1. Regresión logística

La regresión logística es un método estadístico que se emplea para predecir el valor de una variable dependiente binaria, es decir, una variable con dos posibles resultados. Este modelo es especialmente útil para determinar la probabilidad de ocurrencia o no ocurrencia de un evento, basándose en uno o más predictores (variables independientes).

$$P(y=1) = \frac{1}{1+e^{-(\beta_0+\beta_1x_1+\dots+\beta_nx_n)}}$$

Donde  $P(y=1)$  es la probabilidad de que la variable dependiente sea 1 (en riesgo de quiebra),  $e$  es la base del logaritmo natural,  $\beta_0, \beta_1, \dots, \beta_n$  son los coeficientes, y  $x_1, \dots, x_n$  son las variables independientes (Hosmer et al., 2013).

### 7.3.2. Clasificador SVM Lineal

El clasificador de Máquina de Vectores de Soporte (SVM) lineal es una técnica robusta de aprendizaje supervisado, empleada tanto en clasificación como en regresión. El principio central de un SVM lineal es determinar el mejor hiperplano que separa los conjuntos de datos en un espacio de características N-dimensional, donde N representa el número de características. Este hiperplano actúa como una frontera de decisión entre las diferentes clases.

$$w \cdot x + b = 0$$

Donde  $w$  es el vector de pesos,  $x$  representa los vectores de características, y  $b$  es el sesgo. El objetivo es maximizar el margen entre este hiperplano y los puntos de datos más cercanos de cada clase, conocidos como vectores de soporte (Ketkar & Moolayil, 2021).

### 7.2.3. Red Neuronal

Una Red Neuronal Artificial es un modelo computacional inspirado en la estructura y funcionamiento de las redes neuronales biológicas del cerebro humano. Está compuesta por unidades de procesamiento llamadas nodos o neuronas, que están interconectadas por enlaces que tienen pesos asignados. Estas redes son excepcionalmente buenas para detectar patrones complejos y no lineales en los datos gracias a su capacidad para aprender de manera adaptativa.

Fórmula Matemática:

$$f(x) = \phi(\sum(w_i x_i) + b)$$

Donde  $f(x)$  es la salida de la neurona,  $\phi$  es una función de activación (por ejemplo, ReLU o Sigmoides),  $w_i$  son los pesos,  $x_i$  son las entradas, y  $b$  es el sesgo. La red se compone de múltiples capas de estas neuronas, y el entrenamiento implica ajustar los pesos y sesgos para minimizar un error, generalmente mediante un proceso llamado retro propagación y un algoritmo de optimización como el Descenso del Gradiente (Ketkar & Moolayil, 2021).

#### **9.4. Proceso de Modelado**

En este estudio se implementó un conjunto de técnicas avanzadas de modelado predictivo con el fin de cuantificar y pronosticar el riesgo de insolvencia dentro del ámbito sanitario en Colombia. Se han integrado tres metodologías de modelado estadístico y computacional distintas: Regresión Logística, Support Vector Machine (SVM) con un kernel lineal y Redes Neuronales Artificiales. Cada técnica se seleccionó por su capacidad para manejar complejidades inherentes a datos financieros y por su uso establecido en la predicción de eventos binarios.

Para validar la precisión y la generalización de estos modelos, se empleó una técnica de validación cruzada k-fold, lo que permite una evaluación integral de cada modelo al utilizar diferentes subconjuntos de datos como prueba y entrenamiento. Este procedimiento es crucial para mitigar

cualquier sesgo de sobreajuste y para proporcionar una medida confiable de la capacidad predictiva del modelo.

Además, se efectuó un análisis comparativo riguroso utilizando una batería de métricas de rendimiento, incluyendo la precisión, la sensibilidad, la especificidad y el área bajo la curva ROC, entre otras. Este análisis multifacético asegura la selección de un modelo predictivo no solo con la mayor capacidad de discriminación sino también con la mayor relevancia práctica para los interesados en el ámbito financiero de la salud en Colombia. La elección del modelo final se basó en su rendimiento equilibrado a través de estas métricas, garantizando así la aplicabilidad y la relevancia de los pronósticos de riesgo de quiebra.

#### **9.4.1. División del conjunto de datos**

El estudio procedió con una cuidadosa estratificación del conjunto de datos completo, asignando un 80% a la fase de entrenamiento del modelo, lo cual se alinea con las prácticas estándar de ciencia de datos para asegurar una base sólida y representativa sobre la cual los modelos pueden aprender. El 20% restante se destinó a la fase de evaluación, constituyendo una muestra de prueba robusta que posibilita un examen exhaustivo de la capacidad predictiva del modelo y su aplicabilidad práctica en escenarios no expuestos previamente. Este esquema de partición es crucial para mitigar el sobreajuste y garantizar que el desempeño del modelo es evaluado sobre una base de datos independiente, proporcionando así una estimación más fidedigna de su eficacia predictiva en entornos operativos reales.

Se dividió la base de datos compuesta por 1,293 empresas en dos segmentos: un 80% (1,034 empresas) se destinó al entrenamiento del modelo, mientras que el restante 20% (259 empresas) se reservó para la validación final (Tabla 4). Este proceso se llevó a cabo utilizando la función “train\_test\_split”, garantizando así una partición equitativa y sistemática.

**Tabla 1***División de datos*

Total empresas	1293
Train	1034
Test	259

**Nota:** Elaboración propia, en base a la cantidad de datos designada para el proceso de entrenamiento y el de evaluación.

#### 9.4.2. Estandarización de Datos

Tras la distribución del conjunto de datos, se efectuó una normalización de las características de los datos de entrenamiento. Esta técnica de preprocesamiento es esencial, ya que transforma las variables para que operen en una escala estándar con una media aritmética de cero y una desviación estándar unitaria. Este paso es de importancia crítica ya que alinea las escalas de los distintos atributos, optimizando así la eficiencia y la precisión de los algoritmos predictivos al minimizar sesgos en la contribución de cada característica debido a su magnitud. Es crucial recalcar que la normalización se basó en los parámetros derivados exclusivamente del conjunto de entrenamiento para evitar la contaminación de los datos de prueba, manteniendo la integridad del proceso de validación. La misma transformación para estandarizar se aplicó después al conjunto de prueba para asegurar coherencia y precisión en la evaluación del modelo. En donde utilizando la librería StandardScaler, se estandarizan las características numéricas de nuestro conjunto de datos.

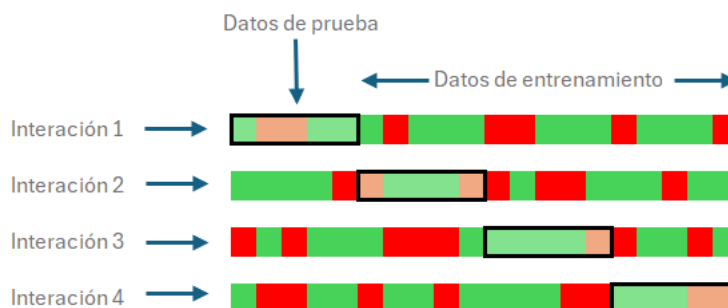
#### 9.4.3. Validación Cruzada de K iteraciones

Para garantizar una evaluación rigurosa de los modelos predictivos, se implementó la estrategia de validación cruzada de K iteraciones, un método estadístico robusto para estimar la eficacia de los modelos en conjuntos de datos independientes. Esta técnica consiste en subdividir aleatoriamente el conjunto de datos en K subconjuntos o 'folds'. En cada una de las K iteraciones del procedimiento, se elige un fold diferente como conjunto de validación, mientras que los K-1 folds restantes constituyen el

conjunto de entrenamiento. Esto se lleva a cabo mediante la función “cross\_validate”, que permite evaluar la efectividad y la estabilidad del modelo bajo diferentes subconjuntos de datos.

**Figura 1**

*Validación cruzada*



**Nota:** Elaboración propia, ilustra el proceso de las interacciones realizadas para la validación.

#### 9.4.4. Métricas de Evaluación

El análisis de la eficiencia predictiva de los modelos desarrollados se llevó a cabo mediante una serie de métricas de evaluación diferenciadas, que son críticas para desglosar y entender en profundidad el comportamiento y la efectividad de cada algoritmo. Entre estas métricas se incluyen, pero no se limitan a, la precisión global, la sensibilidad (tasa de verdaderos positivos), la especificidad (tasa de verdaderos negativos), la precisión de clase, el valor predictivo positivo, la tasa de falsos positivos, y el área bajo la curva de la función operativa del receptor (ROC AUC).

Es relevante anotar que, en conjuntos de datos desbalanceados, la sensibilidad de las métricas se ve notablemente afectada. Aunque la precisión (accuracy) puede ser alta al predecir predominantemente la clase mayoritaria, esta métrica no refleja la capacidad real del modelo para identificar empresas en riesgo de quiebra. En estos casos, la sensibilidad es fundamental, ya que mide la proporción de verdaderos positivos correctamente identificados. Un bajo nivel de sensibilidad indica que el modelo no está detectando adecuadamente las empresas en riesgo de quiebra, lo cual es crítico en este contexto. Por lo tanto, cuando se observan inconvenientes en una clasificación desbalanceada, es esencial priorizar la

sensibilidad para asegurar que el modelo identifique la mayor cantidad posible de empresas en quiebra, minimizando así los falsos negativos y mejorando la capacidad predictiva del modelo en situaciones de alta importancia.

#### 9.4.5. Matriz de Confusión

La matriz de confusión es una herramienta crítica para el análisis detallado del rendimiento de los modelos predictivos aplicados en este estudio. Al categorizar las predicciones de los modelos en cuatro cuadrantes específicos—Verdaderos Positivos (VP), Falsos Positivos (FP), Verdaderos Negativos (VN), y Falsos Negativos (FN)—, la matriz de confusión permite una evaluación precisa y detallada de la capacidad de clasificación del modelo.

Los Verdaderos Positivos y Verdaderos Negativos representan las predicciones donde el modelo ha acertado en identificar correctamente las categorías positivas y negativas, respectivamente. Por otro lado, los Falsos Positivos y Falsos Negativos reflejan los errores donde el modelo ha predicho incorrectamente las categorías negativas como positivas y las positivas como negativas, respectivamente.

A partir de la matriz de confusión, es posible extraer métricas esenciales que proporcionan una evaluación detallada y matizada del rendimiento de los modelos predictivos. Estas métricas, que incluyen la Exactitud (Accuracy), la Precisión (Precisión) y la Sensibilidad (Recall), son fundamentales para analizar cómo el modelo maneja las clasificaciones correctas e incorrectas en diferentes contextos.

- Exactitud (Accuracy): Esta métrica es un indicador general de la eficacia del modelo, reflejando la proporción de predicciones totales que el modelo clasificó correctamente.

$$\text{Exactitud} = \frac{VP+VN}{VP+VN+FP+FN}$$

- Precisión (Precision): Esta métrica evalúa la exactitud de las predicciones positivas del modelo, es decir, la proporción de predicciones de quiebra que resultaron ser correctas.

$$\text{Precisión} = \frac{VP}{VP+FP}$$

- Sensibilidad (Recall): También conocida como la tasa de verdaderos positivos, esta métrica mide la habilidad del modelo para identificar correctamente todos los casos reales de quiebra. Se calcula como:

$$\text{Sensibilidad} = \frac{VP}{VP+FN}$$

Estas métricas permiten no solo evaluar la precisión global del modelo, sino también su eficiencia en prevenir errores críticos como falsos negativos y falsos positivos. En el ámbito de la gestión de riesgo financiero en el sector salud, el balance entre estas métricas es vital para desarrollar un modelo que no solo prediga con precisión, sino que también minimice el riesgo de decisiones erróneas basadas en la predicción.

#### 9.4.6. Curva ROC

La Curva de Característica Operativa del Receptor (ROC) junto con el Área bajo la Curva (AUC) son herramientas estadísticas críticas para evaluar la capacidad de discriminación de un modelo en la clasificación binaria, es decir, su eficiencia para distinguir entre dos clases, como podría ser la clasificación de empresas según su riesgo de quiebra.

La curva ROC es una representación gráfica que ilustra la relación entre la sensibilidad (tasa de verdaderos positivos) y la especificidad (1 - tasa de falsos positivos) a lo largo de varios umbrales de decisión. Esta curva proporciona un resumen visual de la trade-off entre la capacidad del modelo para identificar correctamente las clases positivas y su tasa de error al clasificar las clases negativas como positivas.

El AUC, por su parte, es una métrica integral que mide el grado de separabilidad entre las clases. Cuantifica el área total bajo la curva ROC; un AUC de 1.0 representa un modelo perfecto que logra una

separación completa entre las clases en todos los umbrales, mientras que un AUC de 0.5 sugiere un rendimiento no mejor que el azar. En términos prácticos, cuanto mayor sea el AUC, mejor será el modelo en discriminar entre las entidades en riesgo de quiebra y aquellas estables.

#### 9.4.7. SMOTE

SMOTE, acrónimo de “Synthetic Minority Over-sampling Technique”, es una metodología avanzada diseñada para combatir el desequilibrio de clases en los conjuntos de datos utilizados para análisis predictivos. A diferencia de las técnicas sobre muestreo tradicionales que replican simplemente los ejemplos existentes, SMOTE incrementa la representatividad de la clase minoritaria mediante la creación de ejemplos sintéticos. Esto se logra seleccionando muestras aleatorias de la clase minoritaria y generando nuevos ejemplos que son interpolaciones de estas muestras con sus vecinos más cercanos en el espacio de características.

La adopción de SMOTE puede mejorar significativamente la capacidad de generalización de los modelos predictivos, al proporcionar un contexto de aprendizaje más balanceado y diverso. Al permitir que los algoritmos de aprendizaje automático exploren un espacio más amplio de características representativas de todas las clases, SMOTE ayuda a reducir el sesgo hacia la clase mayoritaria, que puede prevalecer en conjuntos de datos desequilibrados.

El uso de la técnica SMOTE en el análisis predictivo puede considerar ciertos efectos secundarios que deben ser expuestos y mitigados para asegurar la precisión y la robustez del modelo. Uno de los principales riesgos asociados a esta técnica es la generación de datos “ruidosos” al crear ejemplos sintéticos. Estos datos emitidos artificialmente pueden no representar perfectamente las instancias reales, lo que puede conllevar a un sobreajuste del modelo. Esto significa que el modelo podría aprender patrones específicos de los datos sintéticos en lugar de los patrones verdaderamente representativos de las instancias reales de quiebra.

Adicional a esto, el balanceo artificial de clases mediante esta técnica puede hacer que el modelo aprenda patrones que no son generalizables. Este fenómeno puede resultar en una disminución de la capacidad del modelo para predecir correctamente nuevos datos no vistos, especialmente en situaciones donde la distribución de clases es naturalmente desbalanceada. En efecto para mitigar estas dificultades, se implementaron algunas estrategias:

Como primera estrategia se utilizó una validación cruzada estricta para asegurar que los modelos no estuvieran sobre ajustados a los datos sintéticos. Esta técnica permite evaluar el rendimiento del modelo en múltiples subconjuntos de datos, proporcionando una medida más confiable de su capacidad de generalización. Luego se aplicaron métodos de regularización para penalizar la complejidad del modelo, reduciendo así la probabilidad de que el modelo se ajuste demasiado a los datos de entrenamiento. Posteriormente, se evaluó el modelo utilizando múltiples métricas, como la precisión, sensibilidad, especificidad y AUC-ROC, para asegurar una evaluación comprehensiva de su desempeño y robustez en diferentes escenarios. Estas medidas combinadas ayudan a garantizar que el uso de SMOTE mejora la capacidad del modelo para predecir la quiebra sin comprometer su aplicabilidad práctica y generalización.

### **9.5. Optimización de Hiperparámetros para Modelos Avanzados**

En el caso de modelos más complejos, como las Máquinas de Soporte Vectorial (SVM) y las Redes Neuronales, un paso crucial es la selección de los hiperparámetros óptimos. Este proceso se realiza mediante una búsqueda exhaustiva en la que se evalúan diferentes combinaciones de hiperparámetros para determinar cuáles maximizan el rendimiento del modelo.

### **9.6. Estrategias para la comparación de Modelos**

Para asegurar una comparación meticulosa y objetiva de los modelos predictivos, se implementa un enfoque estructurado en tres etapas clave:

- **Aplicación de la Validación Cruzada:** Inicialmente, se evalúa el desempeño de los modelos mediante la técnica de validación cruzada de K iteraciones (con  $k = 10$ ). Este método permite comparar la eficacia de cada modelo no solo en términos de precisión media, sino también respecto a su consistencia y robustez a lo largo de diferentes particiones del conjunto de datos. La validación cruzada ayuda a mitigar el riesgo de sobreajuste, proporcionando una visión más fidedigna del rendimiento del modelo bajo variadas condiciones de muestra.
- **Selección del Modelo Óptimo:** Posterior a la validación cruzada, se seleccionan uno o varios modelos que no solo demostraron superioridad en términos de rendimiento global, sino que también mostraron la mayor consistencia en los resultados a lo largo de los múltiples pliegues. Este criterio asegura que el modelo elegido no solo es potente en promedio, sino que opera de manera confiable independientemente de las variaciones específicas del conjunto de datos.
- **Evaluación Final en el Conjunto de Prueba:** Finalmente, el modelo o modelos seleccionados son sometidos a una evaluación definitiva utilizando el conjunto de prueba, que consiste en datos que no fueron utilizados en ninguna fase de entrenamiento. Este paso es crucial para validar la generalización y la aplicabilidad práctica del modelo. La evaluación en el conjunto de prueba proporciona un indicador crítico de cómo se espera que el modelo se comporte en un entorno real, ofreciendo así una validación adicional de su rendimiento.

Este enfoque tridimensional garantiza una comparación rigurosa y justa de los modelos, enfocándose no solo en la precisión sino también en la estabilidad y la capacidad de generalización, elementos esenciales para la implementación efectiva de soluciones de modelado predictivo en contextos prácticos.

### 9.7. Aplicación de los modelos

### 9.7.1. Modelo logístico

El modelo logístico se implementó en Python utilizando la librería “scikit-learn”. Este modelo incluyó varios pasos: inicialmente, se realizó la partición de los datos en conjuntos de entrenamiento y prueba, seguido por la estandarización de los datos. Posteriormente, los datos fueron balanceados utilizando el método SMOTE para abordar el problema del desequilibrio de clases. Por último, se llevó a cabo la validación cruzada para evaluar la robustez y la generalidad del modelo.

Según lo anterior y bajo los coeficientes proporcionados para el modelo logístico, la fórmula para la predicción de la probabilidad de quiebra de una empresa en el sector salud es:

$$\log\left(\frac{P(\text{Quiebra})}{1 - P(\text{Quiebra})}\right) = -2.77859 + 0.02160 * X1 - 0.01731 * X2 - 0.01915X3 + 1.30289 * X4 - 0.00221 * X5 - 0.01251 * X6 - 2.31787 * X7 - 0.00421 * X8 - 1.68736 * X9 + 0.00030 * X10 - 0.45225 * X11 - 0.26726 * X12 - 0.01860 * X13 - 0.16684 * X14 - 0.68457 * X15$$

Donde:

- X1 = Días de Cartera
- X2 = Días Proveedores
- X3 = Ciclo de Conversión de Efectivo
- X4 = Endeudamiento Total
- X5 = Apalancamiento
- X6 = Deuda Corriente sobre Operación
- X7 = Relación Corriente
- X8 = Cobertura del Gasto Financiero
- X9 = Razón de Tesorería
- X10 = Margen Operacional
- X11 = Margen Neto
- X12 = ROA
- X13 = ROE
- X14 = Debt Service Coverage Ratio
- X15 = Indicador de Solvencia

En el desarrollo del modelo logístico es necesario abordar el desbalance de clases en los datos ya que este proceso inició separando los datos en conjuntos de entrenamiento y prueba.

Posteriormente, se implementó un pipeline en Python utilizando scikit-learn y imblearn, que incluyó la estandarización de las características para generar ejemplos sintéticos de la clase minoritaria y el ajuste del modelo de regresión logística.

La implementación de la técnica de SMOTE fue indispensable para equilibrar la proporción de empresas "sanas" y "quebradas", optimizando la capacidad del modelo para identificar de manera correcta las empresas en riesgo de quiebra. Además, se empleó una validación cruzada de  $k=10$  para asegurar la robustez y la capacidad de generalización del modelo, evaluando su rendimiento en diferentes subconjuntos de datos.

### **9.7.2. Modelo SVM (Máquina de Soporte Vectorial)**

Para el desarrollo y la evaluación del modelo de máquina de soporte vectorial (SVM), se empleó la biblioteca scikit-learn. Para la selección de los hiperparámetros más adecuados, implementamos un proceso de búsqueda de cuadrícula exhaustiva, evaluando las combinaciones posibles:

1. **kernel**: Este es un hiperparámetro que especifica el tipo de kernel a utilizar en el algoritmo SVM. Se añaden tres tipos de kernels al diccionario de parámetros:
  - 'linear': Kernel lineal.
  - 'rbf': Kernel de base radial (Gaussian).
  - 'poly': Kernel polinomial. Aquí también incluimos el parámetro degree, que define el grado del polinomio, se tomaron los alores [2, 3, 4].
2. **C**: Este es un hiperparámetro que controla la penalización de los términos de error en el clasificador SVM. Específicamente, regula la compensación entre lograr un margen más grande y asegurar que cada ejemplo de entrenamiento sea clasificado correctamente. C

puede tomar varios valores, como se define en `params_svm = {'C': [0.1, 1, 10, 100]}`. Estos valores indican que se probarán diferentes niveles de regularización:

- 0.1: Alta regularización, mayor tolerancia a errores en la clasificación, pero con un margen más grande.
- 1: Regularización moderada.
- 10: Regularización baja, menos tolerancia a los errores de clasificación, buscando un margen más ajustado.
- 100: Mínima regularización, se prioriza clasificar correctamente todos los puntos de entrenamiento.

Al final del proceso, “GridSearchCV” proporciona el mejor clasificador y su rendimiento asociado. La propiedad “best\_estimator\_” brinda el mejor modelo encontrado y “best\_score\_” muestra el mejor puntaje obtenido en la validación cruzada, basado en la métrica por defecto para clasificación, que es la precisión. Esto permitirá comparar directamente qué configuración de kernel y parámetros da los mejores resultados en tu conjunto de datos.

El resultado de la búsqueda de cuadrícula reveló que el mejor modelo se obtuvo utilizando un kernel, a continuación, los mejores hiperparámetros:

- **Kernel:** rbf
- **C:** valor 1 y un grado de polinomio de 2

Posteriormente, este conjunto óptimo de hiperparámetros se utilizó para realizar una validación cruzada detallada.

### 9.7.3. Modelo Red neuronal

Utilizando la herramienta “MLPClassifier” de “scikit-learn”, se puede experimentar una amplia variedad de configuraciones. Esto permite afinar el modelo de acuerdo con las características específicas de nuestros datos. Esta sección, se centra el análisis en tres hiperparámetros esenciales:

1. `max_iter`: Este hiperparámetro en `MLPClassifier(max_iter=1000)` especifica el número máximo de iteraciones o épocas que el optimizador pasa sobre el conjunto de datos de entrenamiento. Un valor de 1000 indica que el algoritmo detendrá el entrenamiento si no ha convergido antes de completar 1000 épocas. Este parámetro es importante para evitar que el entrenamiento continúe por demasiado tiempo, especialmente en casos donde el modelo no converge rápidamente.
2. `hidden_layer_sizes`: Este hiperparámetro define la arquitectura de las capas ocultas del perceptrón multicapa. Cada elemento en la tupla representa una capa oculta y el número indica la cantidad de neuronas en esa capa. En `params_mlp = {'hidden_layer_sizes': [(100, 50, 50), (100, 50, 25)]}`, se especifican tres configuraciones diferentes para probar:
  - `(100,)`: Una capa oculta con 100 neuronas.
  - `(50, 50)`: Dos capas ocultas, cada una con 50 neuronas.
  - `(100, 50, 25)`: Tres capas ocultas con 100, 50 y 25 neuronas respectivamente.
3. `alpha`: Este es un parámetro de regularización también conocido como el término de penalización L2. Ayuda a evitar el sobreajuste penalizando pesos más grandes en el modelo. En `params_mlp = {'alpha': [0.0001, 0.001, 0.01]}`, se evalúan tres valores diferentes:
  - `0.0001`: Regularización muy ligera.
  - `0.001`: Regularización moderada.

- 0.01: Regularización más fuerte.

La búsqueda exhaustiva de hiperparámetros mediante GridSearchCV ha culminado con la identificación de una configuración óptima para nuestro modelo de red neuronal. El mejor clasificador resultante se configura con un total de 1000 iteraciones y la función de activación 'relu'. Entre las combinaciones de parámetros evaluadas, la configuración que ha demostrado ser más efectiva incluye:

- hidden\_layer\_sizes: Una sola capa oculta con 100 neuronas
- alpha: Un valor de 0.0001, que implica una regularización muy ligera.

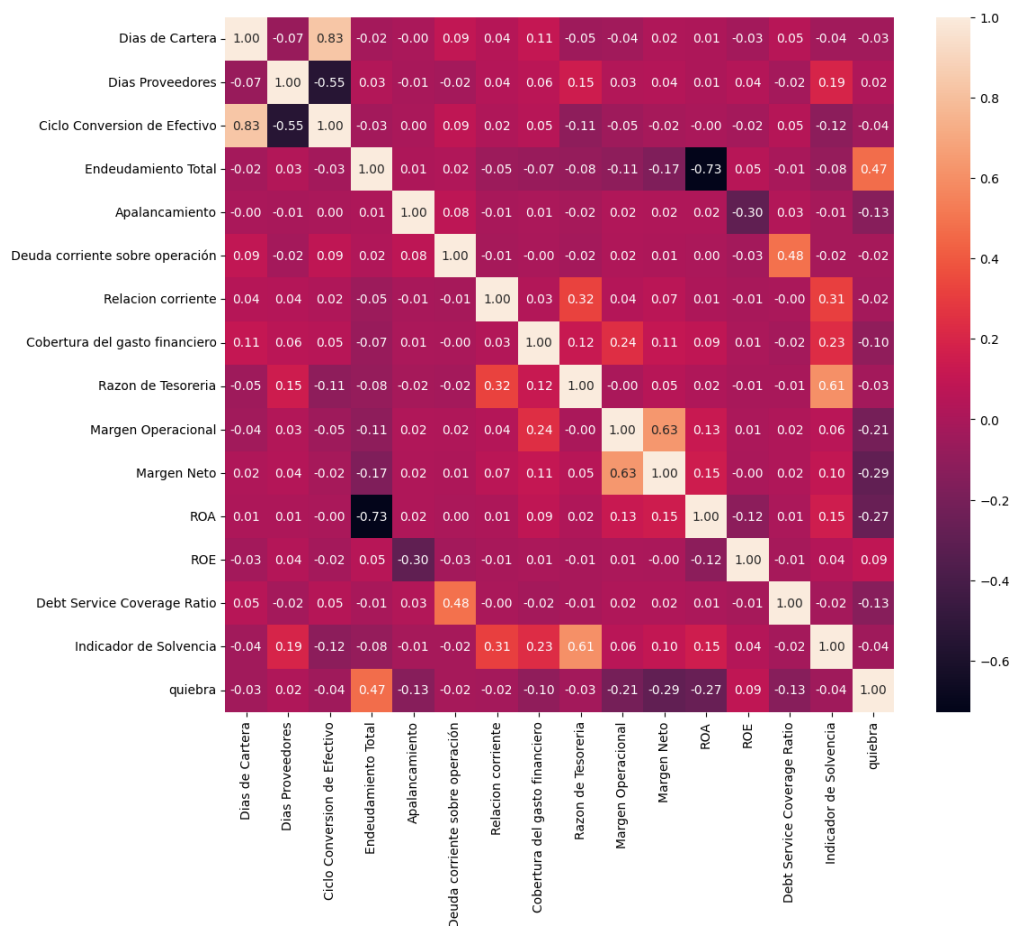
Con estos mejores valores de hiperparámetros se estimó el modelo y se realizó la validación cruzada

## **10. RESULTADOS**

### **10.1. Análisis Exploratorio**

#### **Figura 2**

*Matriz de correlación*



**Nota:** Elaboración propia, ilustra la correlación entre las diferentes variables independientes.

La matriz de correlación usada como una herramienta estadística para medir la fuerza y la dirección de la relación lineal entre dos variables que, para el caso de la muestra aplicada, ayuda a entender como los diferentes indicadores financieros se relacionan entre si dentro de un conjunto de empresas.

La matriz de correlación en la Figura 2 detalla las correlaciones entre diversos indicadores financieros junto con una variable binaria que indica quiebra (1) o estabilidad (0). Así mismo, los valores de correlación varían entre -1 y 1, donde -1 indica una correlación negativa perfecta, 0 ninguna correlación y 1 una correlación positiva perfecta. Posteriormente, se halla lo siguiente:

Existe una correlación alta y positiva de 0.83 en "Días de Cartera" y 0.55 en "Días Proveedores" con respecto al "Ciclo de Conversión de Efectivo"; lo cual tiene sentido ya que el ciclo de conversión de efectivo depende de cuánto tiempo lleva convertir las cuentas por cobrar comerciales y los inventarios en efectivo, y están directamente relacionados con el tiempo en que la empresa tarda en cobrar y pagar sus cuentas.

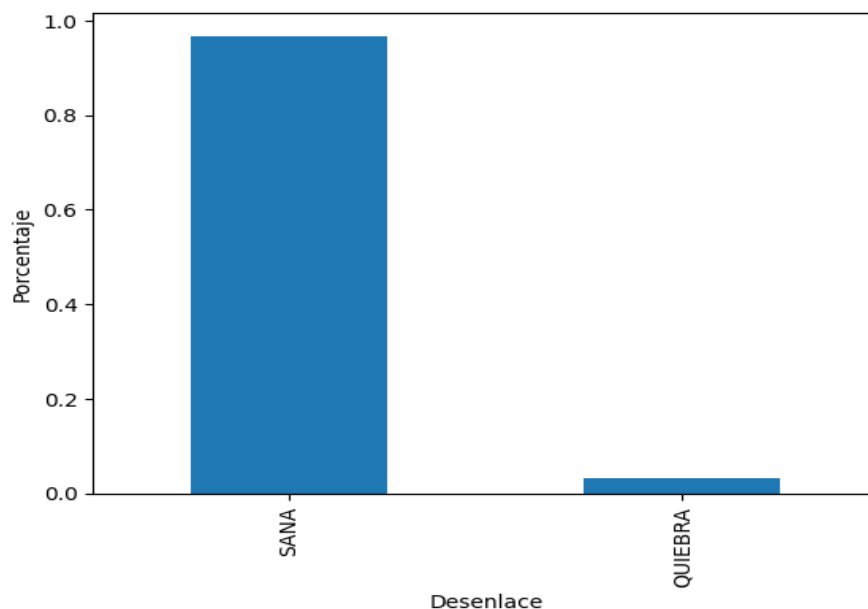
Por el contrario, existe una alta correlación negativa de -0.73 entre el "Endeudamiento Total" y "ROA", lo que puede indicar que a medida que el endeudamiento de una empresa aumenta, su rendimiento sobre los activos tiende a disminuir. Es decir, el apalancamiento excesivo puede estar afectando negativamente la eficiencia con la que las empresas están utilizando sus activos.

La relación con la variable "Quiebra" muestra correlaciones con otros indicadores que podrían interpretarse como factores de riesgo como potencial insolvencia o protección contra la quiebra. Es así como una correlación de 0.47 entre "Quiebra" y "Endeudamiento Total" indicaría que mayores niveles de endeudamiento pueden estar asociados con un mayor riesgo de insolvencia.

La correlación de 0.61 entre "Indicador de Solvencia" y "Razón de Tesorería" es significativamente positiva, lo cual indica que gran parte de empresas muestran una mayor capacidad de pagar deudas a largo plazo está, lo cual se asocia con la generación de más recursos líquidos disponibles para cubrir obligaciones a corto plazo. A pesar de que uno mide la estabilidad financiera a largo plazo y el otro la liquidez a corto plazo, esta correlación podría indicar que las empresas con sólidas posiciones financieras a largo plazo suelen manejar oportunamente su liquidez inmediata.

### **Figura 3**

*Equilibrio de clases*



**Nota:** Elaboración propia, aquí se compara la cantidad de información de las empresas sanas frente a las empresas en quiebra.

La grafica ilustra la distribución de las empresas del sector salud en términos de su solidez financiera. Se evidencia la posición de empresas “sanas” que corresponde al 95.75% del total de la muestra recolectada y que esta muestra un indicativo de robustez o estabilidad financiera, que les permite operar sin signos aparentes de dificultad económica.

Por el contrario, se evidencia un marginal que representa el 3.25% de empresas que se encuentran en la clasificación de “quebradas” o “en quiebra”, las cuales reflejan una aparente dificultad en términos de condiciones financieras que los llevan a una condición de insolvencia o en causal de disolución.

El desbalance en la distribución de categorías “sanas” y “quebradas” sugieren una asimetría en las condiciones financieras de las entidades que prestan servicios de salud en Colombia, lo cual podría

tener implicaciones significativas para la estabilidad del sector salud, así como la calidad en la atención a la salud humana y los riesgos que el contexto pueda comprender.

## 10.2. Análisis Descriptivo

**Tabla 2**

*Resumen de indicadores financieros en las empresas en quiebra*

	Promedio	Desviación estándar	Mínimo	Cuartil 1	Cuartil 2	Cuartil 3	Máximo
<b>Días de Cartera</b>	62.20	133.89	0.00	0.00	7.22	63.67	765.96
<b>Días Proveedores</b>	42.29	49.87	0.00	6.26	23.98	66.43	229.82
<b>Ciclo Conversión de Efectivo</b>	34.90	155.71	-158.95	-26.11	-4.18	79.08	735.23
<b>Endeudamiento Total</b>	2.55	3.72	1.03	1.23	1.49	2.38	24.27
<b>Apalancamiento</b>	-6.10	9.01	-40.80	-5.29	-3.03	-1.73	-1.04
<b>Deuda corriente sobre operación</b>	1.17	12.39	-37.72	-3.40	-1.11	4.12	34.12
<b>Relación corriente</b>	0.43	0.71	0.00	0.01	0.23	0.44	3.36
<b>Cobertura del gasto financiero</b>	-22.10	73.83	-387.62	-13.01	-0.11	2.24	57.75
<b>Razón de Tesorería</b>	0.11	0.38	0.00	0.00	0.01	0.06	2.46
<b>Margen Operacional</b>	-72.75	218.96	-1350.02	-62.34	-22.38	4.33	50.38
<b>Margen Neto</b>	-0.88	2.83	-18.02	-0.64	-0.22	0.00	0.27
<b>ROA</b>	-1.11	3.44	-22.18	-1.05	-0.22	0.00	0.38
<b>ROE</b>	1.07	1.75	-1.63	0.01	0.65	1.28	7.35
<b>Debt Service Coverage Ratio</b>	-2.07	8.87	-55.33	-0.02	0.00	0.00	2.05
<b>Indicador de Solvencia</b>	-0.29	0.40	-1.32	-0.41	-0.16	0.00	0.31

**Nota:** Elaboración propia, evaluación estadística de los indicadores financieros independientes de las empresas en quiebra.

La tabla 2 presenta las agrupaciones estadísticas de los indicadores financieros sobre las empresas que pertenecen al marginal “quiebra” con un amplio rango de desviación estándar en todos los indicadores financieros, lo que supone un alto grado de heterogeneidad en la estructura financiera de estas compañías.

Los días de cartera promedio, con un valor de 62.20 y una desviación estándar de 133.89, junto con un máximo de 765.96, señalan una gestión de cobros posiblemente ineficiente o condiciones de crédito muy laxas, a su vez que las condiciones de transferencia de recursos por parte de EPS o entidades públicas, contribuyan a impactar negativamente a las entidades privadas que se encuentran en el frente de la atención al paciente, de modo que no realizan las operaciones eficientemente oportunas en cuanto al tiempo en el que se está condonando la cartera por servicios prestados. Posteriormente, se podría observar un significativo deterioro en el estado de flujo de efectivo.

Los días de proveedores muestran una media de 42.29 días y un máximo de 229.82, con variaciones que podrían indicar desde condiciones de pago negociables hasta problemas de liquidez que obligan a extender los plazos de pago, retrasando el desembolso a proveedores.

El ciclo de conversión de efectivo promedio es de 34.90 días, pero el rango negativo a máximo sugiere que hay empresas que están convirtiendo su inventario y cuentas por cobrar en efectivo de manera mucho más lenta que otras, lo que puede llevar a insuficiencias en flujo de caja y una posible condición de liquidez desfavorable.

El apalancamiento, medido por el endeudamiento total, tiene un promedio de 2.55, lo que podría indicar un uso significativo e incluso desbordado de la deuda para financiar operaciones. Sin embargo, la presencia de cifras negativas en la deuda corriente sobre operación apunta a la posibilidad de que algunas empresas estén operando con capital de trabajo negativo, lo cual conlleva de manera secuencial a que las dificultades de liquidez y caja deriven en mayor profundidad de insolvencia.

La relación corriente promedio está por encima de 1, lo cual es típicamente positivo, ya que esta se puede ver influida por la acumulación de cartera; contrariamente la variabilidad sugiere que algunas empresas no podrían cubrir sus obligaciones a corto plazo.

La cobertura del gasto financiero presenta un promedio negativo muy alarmante del -22.10, lo que indica que las empresas no estarían generando los suficientes ingresos, o registran altos costos de operación que implican pérdidas operacionales o un margen Ebitda negativo; dificultando el cubrimiento del gasto financiero, el cual podría tener serias implicaciones en cuanto a la participación de la deuda bancaria sobre el endeudamiento total, además de otras consecuencias de carácter jurídico que aumentarían el riesgo de crédito de las mismas.

Los márgenes operacionales y netos negativos sugieren que las empresas están incurriendo en pérdidas que erosionan el patrimonio de los accionistas y aumenta el riesgo de insolvencia. Igualmente se podría indicar que la eficiencia es negativa en cuanto a la fuerza que incrementa las ventas o el estamento que pueda controlar oportunamente los costos de operación, gastos de ventas y administración, entre otros rubros intrínsecos.

Los indicadores de rentabilidad, como el ROA y el ROE, están en territorio negativo en promedio, lo que indica que las empresas no están generando un retorno adecuado sobre sus activos ni sobre el capital invertido.

El Debt Service Coverage Ratio (DSCR) negativo implica que las empresas no están generando suficiente EBITDA para cubrir sus obligaciones de servicio de deuda, lo que aumenta el riesgo de default.

El indicador de solvencia en promedio es negativo, y un máximo de apenas 0.31, lo que sugiere que las empresas tienen más pasivos que activos, una señal clara de insolvencia potencial. De manera general, cuando es mayor a 1 (o 100%) indica que la empresa tiene amplia capacidad para cubrir sus deudas y es una señal de buena salud financiera. Cuando es igual a 1 se aprecia una ajustada capacidad de cubrimiento o un equilibrio precario, por lo que cualquier disminución en ingresos o un aumento potencial de pasivos, podría incurrir en un potencial riesgo. Así mismo, si el indicador es menor a 1,

mientras insuficiencia en el cubrimiento del pasivo total, generando una advertencia sobre la solvencia de la compañía en el futuro.

**Tabla 3**

*Resumen indicadores financieros de empresas sanas.*

	Promedio	Desviación estándar	Mínimo	Cuartil 1	Cuartil 2	Cuartil 3	Máximo
<b>Días de Cartera</b>	86.87	130.98	0.00	0.00	32.37	127.66	947.37
<b>Días Proveedores</b>	33.05	83.92	0.00	4.57	14.12	32.40	1613.06
<b>Ciclo Conversión de Efectivo</b>	72.80	167.58	-1559.33	-6.88	33.40	128.24	947.37
<b>Endeudamiento Total</b>	0.41	0.26	0.00	0.21	0.39	0.59	1.36
<b>Apalancamiento</b>	1.37	9.81	-264.86	0.26	0.61	1.39	137.25
<b>Deuda corriente sobre operación</b>	3.60	25.71	-121.91	0.36	1.14	3.12	602.63
<b>Relación corriente</b>	4.01	30.98	0.00	0.41	1.18	2.29	1008.14
<b>Cobertura del gasto financiero</b>	16.66	67.21	-876.00	0.00	4.81	16.55	932.49
<b>Razón de Tesorería</b>	1.62	9.04	-0.05	0.04	0.22	0.84	137.74
<b>Margen Operacional</b>	8.03	54.35	-1135.14	4.78	10.66	18.56	100.00
<b>Margen Neto</b>	0.07	0.23	-2.84	0.02	0.05	0.11	3.66
<b>ROA</b>	0.10	0.49	-11.99	0.02	0.08	0.16	11.01
<b>ROE</b>	0.23	1.71	-26.88	0.04	0.14	0.31	33.74
<b>Debt Service Coverage Ratio</b>	0.60	3.18	-21.83	0.00	0.00	0.32	35.92
<b>Indicador de Solvencia</b>	0.95	5.73	-13.83	0.04	0.20	0.55	138.15

**Nota:** Elaboración propia, evaluación estadística de los indicadores financieros independientes de las empresas en quiebra.

La tabla 3 presenta los indicadores financieros del equilibrio de la muestra sobre las empresas “sanas” o “estables” en donde resaltan las siguientes observaciones de la agrupación de resultados.

El promedio de los días de cartera es de 86.87 días, con una desviación estándar de 130.98 días, lo que indica una considerable dispersión en la eficiencia en la gestión de recuperación de la cartera comercial de estas empresas. A pesar de la alta media, la presencia de empresas con 0 días de cartera

como valor mínimo, sugiere la existencia de compañías con un ciclo de cobranza excepcionalmente corto. Por el contrario, la presencia de empresas con 947.37 días como indicador máximo, indicaría la desaceleración de la dinámica comercial que conlleva a una baja en los ingresos, que podría combinar las cuentas por cobrar rezagadas, con alto vencimiento que incrementaría los días de cartera.

En cuanto a los días proveedores, la media de 33.05 días con una desviación estándar de 83.92, reflejando prácticas de pago variadas. Algunas empresas pueden estar aprovechando términos de pago extendidos para optimizar su flujo de efectivo. La presencia de 0 días en el pago de proveedores podría reflejar que algunas compañías cuentan con el flujo de caja suficiente para atender los pagos de manera inmediata o de contado, aprovechando descuentos comerciales que optimizarían el costo de oportunidad y mayor utilidad del ejercicio, robusteciendo el valor patrimonial y accionario. Por el contrario, existe un indicador máximo de 1613.06 días, que podría suceder en casos excepcionales o situaciones extraordinarias a revisar detalladamente según sea la condición puntual de la compañía.

El ciclo de conversión de efectivo promedio es de 72.80 días, pero la amplia desviación estándar y los valores negativos mínimos indican que algunas empresas pueden estar manejando de manera muy eficiente su capital de trabajo, mientras que otras podrían tener exceso de inventario o cartera rezagada.

El bajo endeudamiento total promedio de 0.41 es indicativo de una estructura de capital conservadora, posiblemente reflejando un menor riesgo financiero. El apalancamiento promedio de 1.37 con un amplio rango sugiere que hay una variedad de enfoques hacia el uso de deuda.

Una deuda corriente sobre operación de 3.60 sugiere que, en promedio, las empresas tienen la capacidad de cubrir sus deudas a corto plazo con la generación de efectivo de sus operaciones, aunque la alta desviación indica diferencias significativas entre las clases de empresas.

La relación corriente muy superior a 1, con un promedio de 4.01, implica que las empresas tienen más que suficiente capacidad para cubrir sus obligaciones a corto plazo con activos líquidos, lo cual es positivo para la liquidez y solvencia. Sin embargo, es importante generar un detalle de la concentración de su capital de trabajo, ya que las compañías del sector salud presentan variaciones significativas en cuanto al valor y recuperación de la cartera comercial, por lo que dicha liquidez podría verse influida y diferenciada ante el flujo de efectivo que esta relación corriente pueda generar.

La cobertura del gasto financiero, con un promedio de 16.66, indica que las empresas generan ingresos suficientes junto con una adecuada estructura de costos, que deriven en un resultado Ebitda que permita el pago de los intereses financieros, lo cual es una señal de fortaleza financiera y un bajo riesgo de insolvencia. La desviación no representa una medida de alerta, y los mínimos y máximos podrían ser consecuentes a una condición puntual ya que no tiene una relación cercana con su promedio.

La razón de tesorería promedio es saludable a 1.62, la desviación es de 9.04 y el valor mínimo es de -0.05, apoyando la visión de que estas empresas mantienen una liquidez inmediata.

Los márgenes operacionales y netos positivos sugieren eficiencia en la operación y la capacidad de retener ganancias después de todos los gastos.

El ROA y el ROE son bajos, pero positivos, indicando un retorno modesto sobre los activos y el capital invertido respectivamente. Los valores mínimos con un resultado negativo provienen de las posibles pérdidas netas de algún grupo de compañías, que podrían tener condiciones específicas; sin embargo, aun cuentan con el respaldo suficiente que no supera el causal de disolución o que obedece a términos expresamente contables.

El Debt Service Coverage Ratio (DSCR) con un promedio de 0.60 y representa una muestra representativamente baja, teniendo en cuenta que un resultado igual a 1 es en el que el Ebitda cubre

exactamente los pagos programados con terceros financieros, por lo que esto sería una alerta en la que las compañías no están generando el suficiente Ebitda para cubrir el servicio de la deuda, posiblemente por ajustes operativos relacionados a la baja en ventas o una alta estructura de costos. Otra razón de esta condición podría hacer alusión a un ajustado flujo de caja operativo o términos de deuda desfavorables.

El indicador de solvencia de 0.95 indica un balance levemente inclinado entre activos y pasivos, señalando una situación patrimonial generalmente saludable. Aunque lo ideal es que el indicador sea mayor a 1 en cuanto a la capacidad de cubrimiento de las deudas, se observa que generalmente existe un apalancamiento y endeudamiento considerable que podría afectar la solvencia, asumiendo que estas pueden cumplir con las mismas con un mayor esfuerzo operativo.

### 10.3. Análisis de modelos

#### 10.3.1. Modelo Logístico

Los resultados de la validación cruzada se han resumido calculando el promedio de las diez estimaciones:

**Tabla 4**

*Validación cruzada Modelo Logístico – Métricas*

	Promedio
<b>Exactitud (Accuracy)</b>	0.9894
<b>Precisión (Precision)</b>	0.7988
<b>Sensibilidad (Recall)</b>	1
<b>AUC ROC</b>	0.9963

**Nota:** Elaboración propia

**Tabla 5***Matriz de confusión Modelo Logístico*

<b>Real</b>	<b>Predicción</b>	
	No	Si
No	253	1
Si	1	4

**Nota:** Elaboración propia.

Se calcularon las métricas bajo esta matriz de confusión anterior:

**Tabla 6***Métricas datos validación cruzada Modelo Logístico*

	<b>Score</b>
<b>Exactitud (Accuracy)</b>	0.9923
<b>Precisión (Precision)</b>	0.8
<b>Sensibilidad (Recall)</b>	0.8
<b>AUC ROC</b>	0.9984

**Nota:** Elaboración propia.

La exactitud del modelo sigue aproximadamente igual (0.99). La precisión se mantuvo más o menos igual, lo cual es bueno porque significa que cuando el modelo predice que algo sucederá, generalmente tiene razón. Sin embargo, la sensibilidad bajó del 100% al 80%. Esto quiere decir que antes, el modelo era perfecto para detectar todos los casos importantes, pero ahora se está perdiendo algunos en situaciones nuevas. A pesar de eso, la habilidad del modelo para distinguir entre situaciones diferentes mejoró, como lo muestra el aumento en la puntuación AUC ROC a casi 100%. Ahora se estima los coeficientes del modelo:

**Tabla 7***Coefficientes Modelo Logístico*

	<b>Coefficiente</b>	<b>Std. Error</b>	<b>Z</b>	<b>Valor p</b>
<b>Intercepto</b>	-2.77859	0.439768	-6.31830	0
<b>Días de Cartera</b>	0.02160	0.011657	1.85299	0.032
<b>Días Proveedores</b>	-0.01731	0.011747	-1.47338	0.07
<b>Ciclo Conversión de Efectivo</b>	-0.01915	0.011624	-1.64777	0.05
<b>Endeudamiento Total</b>	1.30289	0.456172	2.85613	0.002
<b>Apalancamiento</b>	-0.00221	0.010749	-0.20583	0.418
<b>Deuda corriente sobre operación</b>	-0.01251	0.010067	-1.24267	0.107
<b>Relación corriente</b>	-2.31787	0.835566	-2.77401	0.003
<b>Cobertura del gasto financiero</b>	-0.00421	0.005083	-0.82863	0.204
<b>Razón de Tesorería</b>	-1.68736	2.068869	-0.81560	0.207
<b>Margen Operacional</b>	0.00030	0.003427	0.08715	0.465
<b>Margen Neto</b>	-0.45225	0.690961	-0.65452	0.256
<b>ROA</b>	-0.26726	0.787812	-0.33924	0.367
<b>ROE</b>	-0.01860	0.081664	-0.22777	0.41
<b>Debt Service Coverage Ratio</b>	-0.16684	0.059191	-2.81872	0.002
<b>Indicador de Solvencia</b>	-0.68457	0.879594	-0.77828	0.218

**Nota:** Elaboración propia.

La tabla 7 derivada del análisis logístico, proporciona un entendimiento cuantitativo de cómo varios indicadores financieros afectan la probabilidad de que empresas del sector salud en Colombia sean clasificadas como financieramente estables o quebradas.

Analizando los coeficientes estimados, se observa que el coeficiente para el endeudamiento total es positivo y significativamente distinto de cero, con un valor de 1.30289 y un valor p de 0.002, indicando que a un nivel de significancia del 0.1%, un incremento en el endeudamiento total está asociado positivamente con la estabilidad financiera de las empresas. Esto podría reflejar una contrariedad del endeudamiento, en donde las empresas más endeudadas son aquellas con acceso a financiamiento debido a su solidez y reputación en el mercado, lo que les permite apalancarse estratégicamente para financiar operaciones y expandirse.

En contraste, la relación corriente con un coeficiente de  $-2.31787$  y un valor  $p$  de  $0.003$ , tiene una asociación negativa con la estabilidad financiera de la muestra. Este indicador, tradicionalmente interpretado como una medida de liquidez, podría en este contexto sugerir que un exceso de activos corrientes o una gestión ineficiente de los mismos, no se está utilizando de manera productiva y, por tanto, podría estar correlacionado con un menor rendimiento operativo.

Otro coeficiente de interés es el "Debt Service Coverage Ratio" (DSCR), con un valor de  $-0.16684$  y un valor  $p$  significativo de  $0.002$ , estableciendo una relación negativa con la probabilidad de ser clasificado como estable. Un DSCR bajo puede indicar que la empresa está generando insuficientes flujos de caja operativos con relación a las obligaciones de servicio de la deuda, lo que a generalmente es interpretado como una señal de vulnerabilidad financiera.

El coeficiente para los "Días de Cartera" es bajo, aunque significativo de  $0.02160$  con un valor  $p$  de  $0.032$ . Su significancia estadística sugiere que un aumento en el número de días en los que la empresa tarda en cobrar su cartera comercial puede ser indicativo de una política de crédito expansiva que, si bien podría reflejar confianza en el mercado, también podría señalar una potencial acumulación de cuentas incobrables y una condición en la que incurra en dificultades de liquidez.

Los demás indicadores financieros como "ROA" y "ROE" no resultaron estadísticamente significativos, puede indicar que la rentabilidad operativa y la rentabilidad del patrimonio no son predictores distintivos de la estabilidad financiera dentro de la muestra. Esto puede sugerir que otros factores, posiblemente externos o no financieros, podrían estar desempeñando roles más críticos en la determinación de la estabilidad financiera de estas empresas. Así mismo, en cuanto a la significancia estadística de los indicadores "Apalancamiento" y "Cobertura del Gasto Financiero" sugieren que no hay evidencia suficiente para afirmar que estos factores influyen en la probabilidad de que una empresa sea considerada estable o en riesgo de quiebra, dada la muestra y el periodo estudiado.

Con base a los coeficientes anteriores, en conjunto óptimo de hiperparámetros se utilizó para realizar una validación cruzada detallada, cuyos resultados se resumen a continuación:

### 10.3.2. Modelo SVM

Los resultados de la validación cruzada se han resumido calculando el promedio de las diez estimaciones:

**Tabla 8**

*Validación cruzada Modelo SVM - Métricas*

	Promedio
<b>Exactitud (Accuracy)</b>	0.9975
<b>Precisión (Precision)</b>	0.995
<b>Sensibilidad (Recall)</b>	1
<b>AUC ROC</b>	0.9987

**Nota:** Elaboración propia.

El modelo SVM demostró un rendimiento sobresaliente en la predicción de quiebra para las empresas del sector salud, como lo indican sus elevadas métricas de validación cruzada. Alcanzó una exactitud de 0.9975, evidenciando una excelente capacidad de generalización en distintos conjuntos de datos. La precisión alcanzada fue de 0.995, lo que refleja la alta confiabilidad del modelo al clasificar correctamente las instancias de empresas en riesgo de quiebra. De manera destacada, el modelo logró una sensibilidad perfecta de 1, garantizando la identificación de todos los verdaderos positivos, lo que es crucial para la implementación de estrategias de intervención temprana y precisa. Además, el área bajo la curva ROC de 0.9987 subraya la capacidad excepcional del modelo para discriminar eficazmente entre las empresas en riesgo y las que no lo están, con una mínima tasa de falsos positivos.

Ahora, utilizando el conjunto de datos de prueba, se ha recalibrado el modelo para optimizar su rendimiento.

**Tabla 9***Matriz de confusión Modelo SVM*

	Predicción	
Real	No	Si
No	252	2
Si	0	5

**Nota:** Elaboración propia.**Tabla 10***Métricas datos de validación*

	Modelo SVM
Exactitud (Accuracy)	0.9923
Precisión (Precision)	0.7143
Sensibilidad (Recall)	1
AUC ROC	0.9969

**Nota:** Elaboración propia.

En la evaluación final utilizando el conjunto de prueba (Tabla 10), el modelo SVM demostró una sólida exactitud de 0.9923, reflejando su capacidad para clasificar correctamente la mayoría de las instancias. No obstante, la precisión observada fue significativamente menor, registrando un valor de 0.7143. A pesar de esta caída en la precisión, el modelo mantuvo una sensibilidad perfecta de 1, asegurando que todos los verdaderos positivos fueran capturados sin excepciones. Además, el AUC ROC se mantuvo alto en 0.9969, subrayando la habilidad del modelo para diferenciar entre clases con alta efectividad. Estos resultados indican que, aunque el modelo es extremadamente efectivo en identificar y clasificar correctamente los casos positivos, podría beneficiarse de ajustes que mejoren la precisión sin comprometer otras métricas de rendimiento.

### 10.3.3. Modelo Red neuronal

**Tabla 11***Validación cruzada Modelo Red Neuronal – Métricas*

	Promedio
<b>Exactitud (Accuracy)</b>	0.9959824
<b>Precisión (Precision)</b>	0.9960394
<b>Sensibilidad (Recall)</b>	0.9959848
<b>AUC ROC</b>	0.9978803

**Nota:** Elaboración propia.

Los resultados obtenidos a través de la validación cruzada demuestran un desempeño excepcional, con valores de exactitud, precisión y sensibilidad cercanos al 99.6%, y un área bajo la curva ROC de aproximadamente 99.8%. Esta consistencia en las métricas de rendimiento sugiere que el modelo posee una capacidad notable para identificar correctamente tanto las empresas que están en riesgo de quiebra como aquellas que no lo están. La alta precisión y sensibilidad indican que el modelo es altamente confiable para señalar las empresas en peligro, minimizando el riesgo de falsas alarmas, lo cual es crítico en el sector salud dado su impacto directo en la estabilidad económica y social. El área bajo la curva ROC reafirma la robustez del modelo, destacando su eficacia en clasificar adecuadamente las entidades bajo estudio.

En conjunto, estos resultados resaltan la potencial aplicabilidad del modelo en contextos donde las decisiones precisas son cruciales para prevenir consecuencias financieras adversas.

Ahora con los datos del test, se estima el modelo de tal manera que podamos tener una capacidad del mejor modelo con el fin de asegurar la eficacia del modelo en situaciones reales, se utilizarán los datos de prueba para ajustar y perfeccionar el modelo.

### **Matriz de confusión**

**Tabla 12**

*Matriz de confusión Modelo Red Neuronal*

Real	Predicción	
	No	Si
No	253	1
Si	2	3

**Nota:** Elaboración propia.

**Tabla 13**

*Métricas datos de validación Modelo Red Neuronal*

	Modelo Red Neuronal
<b>Exactitud (Accuracy)</b>	0.996139
<b>Precisión (Precision)</b>	0.833333
<b>Sensibilidad (Recall)</b>	1
<b>AUC ROC</b>	0.99685

**Nota:** Elaboración propia.

Lo anterior se traduce a una exactitud una exactitud (Accuracy) de 99.61%, indicando una alta eficacia en identificar correctamente las empresas en riesgo y aquellas que no lo están. Con una precisión (Precision) de 83.33% y una sensibilidad (Recall) perfecta de 100%, el modelo demuestra gran confiabilidad al clasificar empresas en riesgo. Además, un AUC ROC de 99.685% subraya la sobresaliente capacidad del modelo para discriminar eficazmente entre las diferentes clasificaciones de riesgo.

#### 10.4. Comparación de los modelos

Tomando como base los modelos y las métricas anteriores, se presenta a continuación un resumen:

**Tabla 14.**

*Resumen niveles de precisión de los tres modelos*

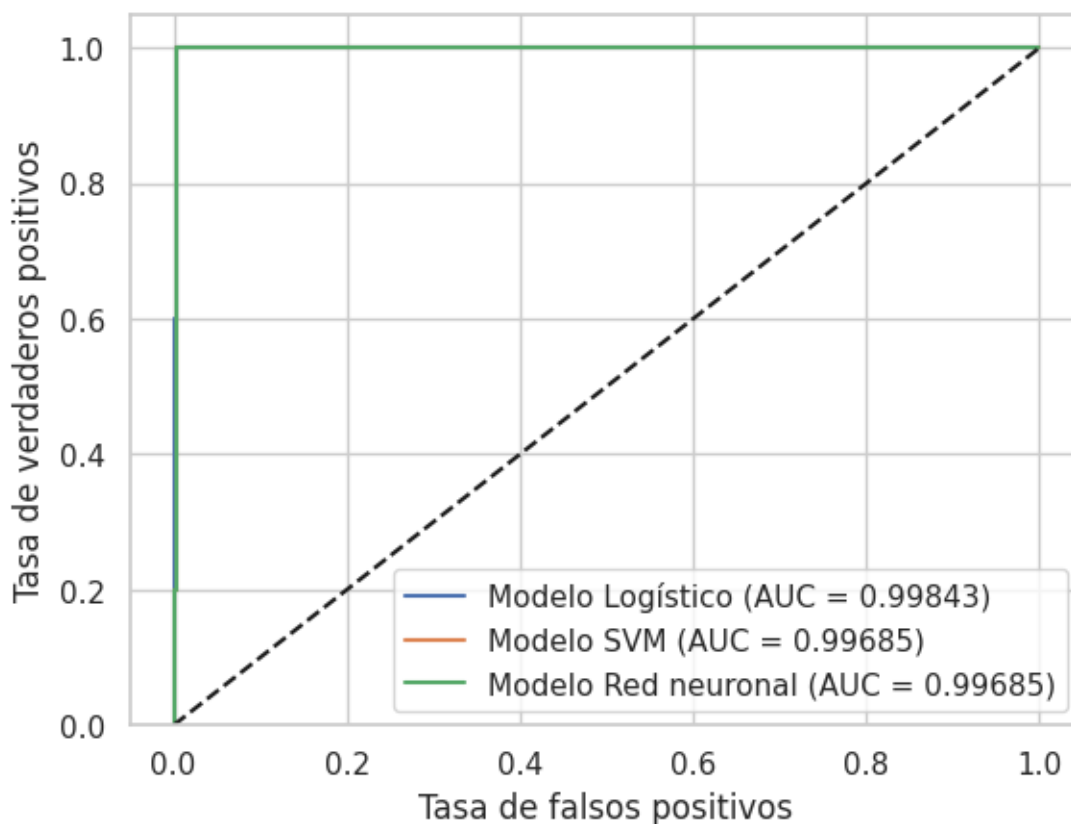
Métrica	Modelo Logístico	SVM	Red Neuronal
<b>Exactitud (Accuracy)</b>	0.992278	0.992278	0.996139
<b>Precisión (Precision)</b>	0.8	0.714286	0.833333
<b>Sensibilidad (Recall)</b>	0.8	1	1
<b>AUC ROC</b>	0.998425	0.99685	0.99685

**Nota:** Elaboración propia.

La curva ROC de cada modelo se generó de la siguiente manera:

**Figura 4**

Curva ROC



**Nota:** Elaboración propia.

Considerando las métricas de desempeño presentadas para los modelos de Regresión Logística, SVM lineal y Red Neuronal, la Red Neuronal demuestra ser el modelo más adecuado para predecir la quiebra en el sector salud en Colombia para el periodo 2022. Este modelo tiene la mayor exactitud con un 99.61%, también la precisión es alta con valor de 0.833333, esto indica una mejor capacidad para identificar correctamente las empresas en riesgo de quiebra.

Claro está que el modelo SVM y la Red Neuronal tienen una sensibilidad y AUC ROC iguales, señalando una excelente capacidad para identificar positivos verdaderos y un excelente desempeño

bajo la curva ROC, por tanto, la red neuronal es la mejor opción dado que tiene la exactitud y precisión alta.

### Probabilidades

Mediante el modelo de regresión logística, se evalúa la “Probabilidad de Quiebra” en el conjunto de empresas pertenecientes al sector salud, clasificadas inicialmente como “Financieramente Sanas”.

**Tabla 15**

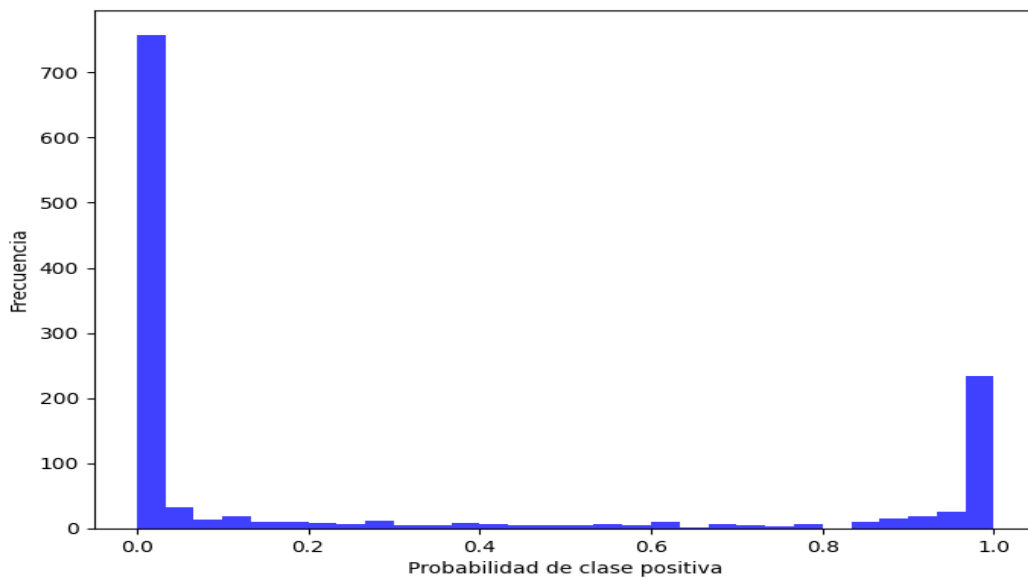
*Métricas Modelo Logístico – Probabilidad muestra de empresas sanas*

Total empresas	1251
Promedio	0.283837343
Desv. Estándar	0.415234648
Mínimo	3.254914E-14
Cuartil 1	4.53E-06
Cuartil 2 – Mediana	0.0023661747
Cuartil 3	0.7189773833
Máximo	1

**Nota:** Elaboración propia.

**Figura 5**

*Distribución de las probabilidades quiebra de empresas sanas*



**Nota:** **Nota:** Elaboración propia.

Se encontró que la probabilidad media de ser consideradas “Financieramente Sanas” es del 28%, con una considerable variabilidad. Este alto grado de dispersión en las probabilidades refleja una notable disparidad en sus indicadores financieros. Por lo tanto, el 72% de la probabilidad media de la muestra, indica que esta población de empresas se clasificaría como “Probabilidad de Quiebra” las cuales reflejan un riesgo financiero significativo.

El resultado de la alta probabilidad de quiebra podría ser consecuencia de los factores inherentes al sector salud, como las fluctuaciones en las tasas de interés para obtener financiación, cambios en la regulación, aumento de los costos en insumos y gasto de personal altamente calificado, y las condiciones del mercado que pueden erosionar los márgenes de rentabilidad y la liquidez como capital de trabajo. Además, el impacto de las condiciones políticas como las reformas estructurales en la normativa que la rige, crisis económicas o ausencia del estado en apoyo a la coyuntura el sector, etc.

## **11. CONCLUSIONES**

El sector salud en Colombia es un pilar fundamental en el bienestar social y económico del país que, a pesar de su importancia, enfrenta desafíos significativos como el retraso en la transferencia de recursos desde las EPS a las entidades privadas, lo cual ha sido producto de las regulaciones del ente administrador de los recursos públicos y la inadecuada legislación que rige a toda la cadena productiva de la misma. El sector también ha manifestado la preocupación por el envejecimiento de la población y la baja implementación en los hábitos de vida saludable, lo que conlleva a una mayor demanda de atención y mayor inversión en aquellos procesos que mitiguen el impacto de la morbilidad en la población, así como evitar propagación de enfermedades que agraven la salud pública. Esto por supuesto, sumado al aumento de la afiliación de población migrante, la cual se encuentra acobijada bajo el régimen subsidiado, ocupando mayor capacidad instalada de las instituciones, entre otras implicaciones que podrían deteriorar la estabilidad financiera de estas compañías en un futuro próximo.

Por lo anterior, la elección de abarcar el tema propuesto y la investigación realizada es una clara respuesta a la urgencia de aplicar modelos predictivos eficaces que anticipen situaciones de quiebra a partir de los múltiples desafíos que se enfrenta este sector productivo, y que promuevan la toma de medidas preventivas para poder asegurar la continuidad de los servicios de salud en la población.

La legislación sobre insolvencia o procesos de reorganización empresarial no está completamente adaptada a las particularidades del sector salud ya que las instituciones de salud privadas no están incluidas en la ley de insolvencia empresarial que protege el crédito y la recuperación empresarial (ley 1116 del 2006); por el contrario, el sector se encuentra regulado por la Superintendencia Nacional de Salud, la cual no es especialista en cuanto a la normatividad aplicable en el código de comercio que rige a compañías privadas, con fines de lucro, en contraprestación de servicios o canje de productos con remuneración directa y sin manejo de recursos girados por el estado. Esto crea un entorno en el que es difícil para estas entidades reestructurarse o recuperarse de manera efectiva una vez que enfrentan problemas financieros estructurales. En ese sentido, la Superintendencia Nacional de Salud genera acciones de intervención administrativa en la cual diagnostica y controla en un tiempo estimado la entidad, para llevar a cabo la disolución de las falencias de carácter financiero y operativo. Y, aunque el gobierno ha hecho esfuerzos para mejorar el sistema a partir de reformas y resoluciones de tipo administrativo, que finalmente busca saldar las deudas históricas del sistema con las EPS e IPS, así como el aseguramiento de calidad en la prestación de servicios a la salud humana en toda la población con cobertura nacional, aún le quedan retos significativos para asegurar la sostenibilidad financiera y operativa del sistema. Por lo cual, la comprensión de estos desafíos y la implementación de políticas estructurales son fundamentales para mejorar la estabilidad financiera del sector y garantizar un acceso equitativo.

Los modelos implementados en la investigación reflejan capacidades variadas para predecir la quiebra en el sector salud. En el caso del modelo de Regresión Logística, por ser un modelo

relativamente simple y transparente, permitió la interpretación y diagnóstico de los resultados, aunque pudo ser limitado en manejar la complejidad y no linealidad de algunos de los indicadores financieros. Por su parte, los modelos SVM y Redes Neuronales son más sofisticados, pudieron capturar relaciones no lineales más complejas entre las variables. Sin embargo, esto también los hizo más susceptibles al sobreajuste, especialmente en conjuntos de datos con variabilidad limitada y desbalance de clases. La validación cruzada ayudó a mitigar este riesgo, proporcionando una evaluación más robusta de su desempeño.

Posterior a la modelación de datos, se evidencian los resultados más sobresalientes en el análisis exploratorio y descriptivo, con respecto a la alta correlación entre el nivel de endeudamiento total y el ROA, en donde el excesivo apalancamiento suele estar asociado con mayores cargas financieras, reduciendo la capacidad operativa de la empresa, afectando negativamente la eficiencia con la que las empresas están utilizando sus activos, aumentando el nivel de riesgo de insolvencia por un ciclo de gestión inadecuado. Otro aspecto crítico es la liquidez, evidenciando correlaciones significativas en indicadores como la razón corriente y la razón de tesorería, sugiriendo dificultades para cubrir las obligaciones a corto plazo sin obtener financiación adicional; correlacionado con el indicador de rotación de la cartera, expuesto por condiciones de los entes reguladores y administradores de los recursos públicos, los cuales han sido históricamente condicionados a los retrasos en la transferencia de recursos a las entidades de la muestra. Por lo tanto, esto demuestra insuficiencia de liquidez en las entidades, derivando en mayor riesgo de liquidez y posible default.

Gran parte de las limitaciones principales en la realización de la presente investigación fue el acceso a los estados financieros completos de las entidades que componen todo el sector salud bajo la entidad reguladora, sobre todo los de las empresas que se encuentran catalogadas en quiebra por la Supersalud. Estos presentaban diferencias en tipo de formato, periodos contables incompletos, partidas contables contrarias a su naturaleza, entre otras observaciones que fueron minuciosamente ajustadas en

la base de datos total. Estas condiciones podrían afectar la construcción de los modelos predictivos desarrollados, comprendiendo la necesidad de mejorar las bases de datos públicas y su transparencia.

Nuestra motivación al realizar la investigación actual con el objetivo de profundizar en el sector salud y en aras de poner en marcha las herramientas suministradas por la academia, el proceso de investigación y modelación de información financiera junto con sus resultados nos ha permitido visualizar futuras investigaciones que podrían enfocarse, por ejemplo, en la exploración de nuevas metodologías predictivas con un enfoque de automático y secuencial, que permita capturar estacionalidades y patrones en la información financiera. Adicionalmente, incorporar modelos estadísticos que permitan incluir variables macroeconómicas y socioeconómicas, que permitan evaluar factores externos y su impacto en la probabilidad de quiebra en empresas del sector.

## 12. REFERENCIAS

- Altman, E. I. (1968). Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy. *The Journal of Finance*, 23(4). <https://doi.org/10.2307/2978933>
- Argáez, C. C., & Rodríguez, G. D. M. (2023). *Los desafíos de la economía colombiana en la tercera década del siglo XXI* (Universidad de los Andes, Ed.; 1st ed.).
- Ayús, A., Valencia, R., & Bedoya, J. (2023). Modelación de la quiebra empresarial en el sector de la construcción en Colombia. *Cuadernos Del Cendes*, 40(113), 125–147.
- Banco Mundial. (2024, March 13). *Salud*. <https://www.bancomundial.org/es/topic/health/overview#1>.
- Bărbuță-Mișu, N., & Madaleno, M. (2020). Assessment of Bankruptcy Risk of Large Companies: European Countries Evolution Analysis. *Journal of Risk and Financial Management*, 13(3). <https://doi.org/10.3390/jrfm13030058>
- Beauvais, B., Ramamonjiarivelo, Z., Betancourt, J., Cruz, J., & Fulton, L. (2023). The Predictive Factors of Hospital Bankruptcy—An Exploratory Study. *Healthcare (Switzerland)*, 11(2). <https://doi.org/10.3390/healthcare11020165>
- Beaver. (1966a). Financial Ratios as Predictors of Failure. *Empirical Research in Accounting: Selected Studies 1966*. *Journal of Accounting Research*, 4(1966).
- Beaver, W. H. (1966b). Financial Ratios As Predictors of Failure. *Journal of Accounting Research*, 4. <https://doi.org/10.2307/2490171>
- Ben Jabeur, S., Stef, N., & Carmona, P. (2023). Bankruptcy Prediction using the XGBoost Algorithm and Variable Importance Feature Engineering. *Computational Economics*, 61(2). <https://doi.org/10.1007/s10614-021-10227-1>

- Biniek, R., & Gossett, K. (2017). Sample indicators for predicting U.S. publicly-traded for-profit hospital financial solvency. *Global Journal of Accounting and Finance*, 1(1).  
<https://doi.org/10.47177/gjaf.01.01.2017.049>
- Bozkurt, İ., & Kaya, M. V. (2023). Foremost features affecting financial distress and Bankruptcy in the acute stage of COVID-19 crisis. *Applied Economics Letters*, 30(8).  
<https://doi.org/10.1080/13504851.2022.2036681>
- CEPAL. (2022). Repercusiones en América Latina y el Caribe de la guerra en Ucrania: ¿cómo enfrentar esta nueva crisis? In *Revista Comisión Económica para América y el Caribe (CEPAL)*.
- Chava, S., & Jarrow, R. A. (2004). Bankruptcy Prediction with Industry Effects\*. *Review of Finance*, 8(4).  
<https://doi.org/10.1093/rof/8.4.537>
- Código de Comercio, Pub. L. No. Artículo 218 (1971).
- Comunicados de prensa - Transparencia por Colombia. (2024). *Más de \$21 billones de pesos perdidos y 15 millones de personas afectadas por la corrupción en Colombia revela la Radiografía de la Corrupción 2016 - 2022*. <https://transparenciacolombia.org.co/comunicados-de-prensa/>
- DANE. (2022). *Años de esperanza de vida perdidos 2017 - 2019: Un análisis regional y por grupos etarios*.
- DANE. (2023a). *Encuesta nacional de calidad de vida*.
- DANE. (2023b). *Estadísticas Vitales*.
- Dong, G. N. (2015). Performing well in financial management and quality of care: Evidence from hospital process measures for treatment of cardiovascular disease. *BMC Health Services Research*, 15(1).  
<https://doi.org/10.1186/s12913-015-0690-x>

du Jardin, P. (2010). Predicting bankruptcy using neural networks and other classification methods: The influence of variable selection techniques on model accuracy. *Neurocomputing*, 73(10–12).

<https://doi.org/10.1016/j.neucom.2009.11.034>

Dubas-Jakóbczyk, K., Kocot, E., Tambor, M., & Quentin, W. (2021). The association between hospital financial performance and the quality of care—a scoping review protocol. *Systematic Reviews*, 10(1). <https://doi.org/10.1186/s13643-021-01778-3>

Dupleix, M. D. (2021). La teoría efectual y el fracaso empresarial. *Innovar*, 31(81).

<https://doi.org/10.15446/innovar.v31n81.95581>

Echeverri, O. (2008). Mercantilización de los servicios de salud para el desarrollo: El caso de Colombia. *Revista Panamericana de Salud Publica/Pan American Journal of Public Health*, 24(3).

<https://doi.org/10.1590/s1020-49892008000900008>

Espejo, M. C. M. (2014). Declive organizativo, fracaso y reestructuración organizacional en empresas colombianas. *Contaduría y Administración*, 59(3). [https://doi.org/10.1016/s0186-1042\(14\)71271-9](https://doi.org/10.1016/s0186-1042(14)71271-9)

Espinosa, F. R., Molina, Z. A. M., & Vera-Colina, M. A. (2015). Fracaso empresarial de las pequeñas y medianas empresas (pymes) en Colombia. *Suma de Negocios*, 6(13).

<https://doi.org/10.1016/j.sumneg.2015.08.003>

Estatuto Organico Financiero, Pub. L. No. Ley 795 de 2003 (2003).

Galán-Barrera, J. A., & Torres-García, L. F. (2017). El fracaso empresarial en Colombia: Aproximación a través del modelo de Fulmer. *Civilizar de Empresa y Economía*, 13(1).

Gomez, V. (2020). *Modelo de evaluación para el análisis de insolvencia de instituciones prestadoras de servicios de salud privadas de Colombia con base en el análisis de las mejores clínicas y hospitales del ranking de América economía*. CESA.

González García, L. M., Viga Juárez, C. A., & Fierro Martínez, S. D. (2017). Prospección del riesgo operativo de las Mipymes en Colombia. *Suma de Negocios*, 8(18).

<https://doi.org/10.1016/j.sumneg.2017.11.004>

*Health at a Glance 2023*. (2023). OECD. <https://doi.org/10.1787/7a7afb35-en>

Holmes, G. M., Kaufman, B. G., & Pink, G. H. (2017). Predicting Financial Distress and Closure in Rural Hospitals. *Journal of Rural Health*, 33(3). <https://doi.org/10.1111/jrh.12187>

Hosmer, D. W., Lemeshow, S., & Sturdivant, R. X. (2013). Applied Logistic Regression: Third Edition. In *Applied Logistic Regression: Third Edition*. <https://doi.org/10.1002/9781118548387>

Iñiguez, P. (2023, November 15). *Informe de tasas de tendencias médicas globales 2024*.

<https://noa.aon.es/2024-global-medical-trend-rates-report/>

Ketkar, N., & Moolayil, J. (2021). Deep Learning with Python. In *Deep Learning with Python*.

<https://doi.org/10.1007/978-1-4842-5364-9>

Kitowski, J., Kowal-Pawul, A., & Lichota, W. (2022). Identifying Symptoms of Bankruptcy Risk Based on Bankruptcy Prediction Models—A Case Study of Poland. *Sustainability (Switzerland)*, 14(3).

<https://doi.org/10.3390/su14031416>

Lesáková, L., Gundová, P., & Vinczeová, I. (2020). The practice of use of models predicting financial distress in Slovak companies. *Journal of Eastern European and Central Asian Research*, 7(1).

<https://doi.org/10.15549/jeecar.v7i1.369>

Ley 1116, Diario Oficial No. 46.494 (2006).

- Lombardo, G., Pellegrino, M., Adosoglou, G., Cagnoni, S., Pardalos, P. M., & Poggi, A. (2022). Machine Learning for Bankruptcy Prediction in the American Stock Market: Dataset and Benchmarks. *Future Internet*, 14(8). <https://doi.org/10.3390/fi14080244>
- López Rodríguez, A. L., López Rodríguez, S. A., Pinzón Hoyos, B., & Vásquez Bernal, O. A. (2023). Turismo médico en Colombia: dinámica y ventaja competitiva. *Revista CEA*, 9(20). <https://doi.org/10.22430/24223182.2407>
- Mejía, D., & Castaño, M. (2020). Indicadores financieros como instrumento poderoso para predecir la insolvencia; un estudio usando el algoritmo boosting en empresas colombianas. *Journal of Management and Economics for Iberoamerica*, 36(155), 229–238.
- Melo-Becerra, L. A., Arango-Thomas, L. E., Ávila-Montealegre, Ó., Ayala-García, J., Bonilla-Mejía, L., Botero-García, J. A., Cardona-Badillo, M., Crispin-Fory, C., Gallo-Montaño, D. del P., Granger-Castaño, C. J., Guzmán-Finol, K., Iregui-Bohórquez, A. M., Ospina-Tejeiro, J. J., Pinilla-Alarcón, D., Posso-Suárez, C., Ramírez-Giraldo, M. T., Ramos-Forero, J. E., Ramos-Velozza, M., Restrepo-Tobón, D. A., ... Vásquez-Escobar, D. M. (2023a). Aspectos financieros y fiscales del sistema de salud en Colombia. *Ensayos Sobre Política Económica*, 106, 1–92. <https://doi.org/10.32468/espe106>
- Melo-Becerra, L. A., Arango-Thomas, L. E., Ávila-Montealegre, Ó., Ayala-García, J., Bonilla-Mejía, L., Botero-García, J. A., Cardona-Badillo, M., Crispin-Fory, C., Gallo-Montaño, D. del P., Granger-Castaño, C. J., Guzmán-Finol, K., Iregui-Bohórquez, A. M., Ospina-Tejeiro, J. J., Pinilla-Alarcón, D., Posso-Suárez, C., Ramírez-Giraldo, M. T., Ramos-Forero, J. E., Ramos-Velozza, M., Restrepo-Tobón, D. A., ... Vásquez-Escobar, D. M. (2023b). Aspectos financieros y fiscales del sistema de salud en Colombia. *Ensayos Sobre Política Económica*, 106, 1–92. <https://doi.org/10.32468/espe106>
- Mendoza, L. (2020). *Sistema de alerta temprana de riesgo financiero para los hospitales públicos de Colombia*. Universidad Militar Nueva Granada.

- Migración Colombia. (2023). *Distribución de Migrantes Venezolanos*.
- Mosadeghrad, A. M. (2014). Factors influencing healthcare service quality. *International Journal of Health Policy and Management*, 3(2). <https://doi.org/10.15171/ijhpm.2014.65>
- Narvekar, A., & Guha, D. (2021). Bankruptcy prediction using machine learning and an application to the case of the COVID-19 recession. *Data Science in Finance and Economics*, 1(2), 180–195.  
<https://doi.org/10.3934/DSFE.2021010>
- Nasmi, W., & Afriyenti, M. (2021). Analisis Prediksi Kebangkrutan pada Perusahaan Manufaktur Sub Sektor Food & Beverages Yang Terdaftar di Bursa Efek Indonesia: Menggunakan Model Altman, Springate, dan Grover. *JURNAL EKSPLORASI AKUNTANSI*, 3(4), 749–763.  
<https://doi.org/10.24036/jea.v3i4.421>
- Nehrebecka, N. (2018). Predicting the default risk of companies. Comparison of credit scoring models: logit vs support vector machines. *ECONOMETRICS*, 22(2). <https://doi.org/10.15611/eada.2018.2.05>
- Orozco, B. (2024, February). Equipos médicos podrían aumentar más de lo previsto a causa de Reforma a la Salud. *La Republica*.
- Papík, M., & Papíková, L. (2023). Impacts of crisis on SME bankruptcy prediction models' performance. *Expert Systems with Applications*, 214. <https://doi.org/10.1016/j.eswa.2022.119072>
- Procuraduría General de la Nación. (2023). *Procuraduría advierte graves consecuencias en el país ante la crisis financiera de las EPS*.
- Puro, N., Borkowski, N., Hearld, L., Carroll, N., Byrd, J., Smith, D., & Ghiasi, A. (2019). Financial Distress and Bankruptcy Prediction: A Comparison of Three Financial Distress Prediction Model in Acute Care Hospitals. *Journal of Health Care Finance*.

Ravi Kumar, P., & Ravi, V. (2007). Bankruptcy prediction in banks and firms via statistical and intelligent techniques - A review. *European Journal of Operational Research*, 180(1).

<https://doi.org/10.1016/j.ejor.2006.08.043>

Resolución 2022320030005874 - 6, Pub. L. No. 2022320030005874-6 (2022).

Šverko Grdić, Z., Radolović, J., & Bagarić, L. (2009). Solventnost poduzeća u Republici Hrvatskoj iu Europskoj unij. *Ekonomski Pregled*, 60(5-6), 250-266.

Támara Ayús, A. L., Villegas, G. C., Leones Castro, M. C., & Salazar Bocanegra, J. A. (2019). Modelación del riesgo de insolvencia en empresas del sector salud empleando modelos logit. *Revista de Métodos Cuantitativos Para La Economía y La Empresa*, 26.

<https://doi.org/10.46661/revmetodoscuanteconempresa.2757>

Topcu, M., & Gulal, O. S. (2020). The impact of COVID-19 on emerging stock markets. *Finance Research Letters*, 36. <https://doi.org/10.1016/j.frl.2020.101691>

Valaskova, K., Gajdosikova, D., & Belas, J. (2023). Bankruptcy prediction in the post-pandemic period: A case study of Visegrad Group countries. *Oeconomia Copernicana*, 14(1).

<https://doi.org/10.24136/oc.2023.007>

Yepes Luján, F. J. (2010). Luces y sombras de la reforma de la salud en Colombia. Ley 100 de 1993.

*Revista Gerencia y Políticas de Salud*, 9(18).