



Facultad de Ciencias de la Administración

Carrera de Economía

**ANÁLISIS DE CIERRE DE EMPRESAS DENTRO
DEL SECTOR HOTELERO ECUATORIANO
DURANTE EL PERIODO 2011-2021**

**Trabajo de titulación previo a la obtención del
grado de Economista**

Autor:

Carlos Xavier León Pino

Director:

Iván Felipe Orellana Osorio

Cuenca – Ecuador

2023

DEDICATORIA

Este trabajo de graduación va dedicado a mi madre Karina Pino y a mi padre Carlos León Cherres por su constante e incondicional apoyo a lo largo de estos cuatro años de carrera. Esta tesis también va dedicada a mi hermana María Inés León por siempre estar a mi lado y darme a entender que puedo contar con ella en todo momento.

AGRADECIMIENTO

En primer lugar quiero agradecer a mis padres Carlos León Cherres y Karina Pino Flores por haberme dado la posibilidad de cursar mis estudios dentro de esta gran institución educativa. También quiero agradecer al Ingeniero Iván Orellana Osorio y al Economista Luis Pinos por haberme ofrecido todo su apoyo en la realización del presente trabajo de Titulación, además de dar las más sinceras gracias a todos los docentes que me han acompañado a lo largo de la carrera de economía, brindándome las bases para la realización de esta tesis. Por último quiero agradecer a mis compañeros por haberme apoyado durante todos estos años y de haber hecho de la universidad una experiencia grata.

Índice de Contenidos

DEDICATORIA	i
AGRADECIMIENTO	ii
Índice de Contenidos.....	iii
Índice de Gráficos.....	iv
Índice de Tablas	v
Índice de Ecuaciones.....	vi
RESUMEN y ABSTRACT	vii
1. Introducción	1
1.1. Objetivos	2
1.2. Marco teórico	2
2. Revisión de literatura	3
3. Métodos.....	6
4. Resultados	12
4.1. Análisis descriptivo base de datos	12
4.2. Análisis de correlación.....	14
4.3. Modelo Logístico- Datos de entrenamiento.....	14
4.3.1. Análisis de coeficientes	14
4.3.2. Significancia de las variables	15
4.4. Modelo logístico- Datos de prueba	16
4.4.1. Análisis de coeficientes	16
4.4.2. Significancia de las variables	17
4.5. Redes neuronales	18
5. Discusión.....	19
6. Conclusión	21
7. Referencias.....	23

Índice de Gráficos

Gráfica 1.....	12
Gráfica 2.....	14

Índice de Tablas

Tabla 1	11
Tabla 2	12
Tabla 4	15
Tabla 5	16
Tabla 6	17
Tabla 7	18
Tabla 8	18
Tabla 9	19

Índice de Ecuaciones

Ecuación 1.....	6
Ecuación 2.....	7
Ecuación 3.....	7
Ecuación 4.....	7
Ecuación 5.....	7
Ecuación 6.....	7
Ecuación 7.....	8
Ecuación 8.....	8
Ecuación 9.....	8
Ecuación 10.....	9
Ecuación 11.....	9
Ecuación 12.....	10

Resumen:

La capacidad de financiamiento con la que pueda contar un hotel dentro de territorio ecuatoriano es crucial para su supervivencia. Por lo que el objetivo de este artículo era determinar la probabilidad de cierre de las empresas en el sector hotelero del Ecuador en base a información financiera del periodo 2011-2021. La metodología empleada consistió en la aplicación tanto de un modelo logístico como uno de redes neuronales a una base de datos particionada. Los resultados demostraron que por parte del modelo Logit, la variable endeudamiento fue la más relevante y este modelo contó con una capacidad de predicción del 97,7%. Mientras que para las redes neuronales el nivel de clasificación fue del 63,57%, concluyendo que para este sector el modelo logístico llegó a ser mejor para la predicción de quiebra y que las empresas deberían prestar especial atención a sus respectivas capacidades de endeudamiento.

Palabras clave: Alojamientos turísticos, Cese de actividad, Cierre de empresas, Insolvencia, Quiebra

Abstract:

The financing capacity of a hotel within Ecuadorian territory is crucial for its survival. Therefore, the objective of this article was to determine the probability of the closure of companies in the hotel sector in Ecuador based on financial information for the period 2011-2021. The methodology employed consisted of applying both a logistic model and a neural network model to a partitioned database. The results showed that for the Logit model, the indebtedness variable was the most relevant and this model had a predictive capacity of 97.7%. While for the neural networks, the classification level was 63.57%, concluding that for this sector the logit model was better for bankruptcy prediction and that companies should pay special attention to their respective indebtedness capacities.

Keywords: Necessary Value Tourist accommodations, Business Closure, Business Closure, Insolvency, Bankruptcy

Keywords: Bankruptcy, Cessation of activity, Company closure, Insolvency, Tourist accommodation



Este certificado se encuentra en el repositorio digital de la Universidad del Azuay, para verificar su autenticidad escanee el código QR

Este certificado consta de: 1 página

1. Introducción

El desarrollo de la industria del turismo de un país está sumamente ligada al rendimiento del sector hotelero del mismo, dado que este ofrece uno de los servicios más relevantes para cualquier turista, el alojamiento (Orlandini et al., 2019). En el Ecuador esta industria para el 2011 mostró un crecimiento del 5% y generó ingresos de 1,2 billones de dólares, esta bonanza llegó a darse de una manera similar para el año 2012 (Ministerio de Turismo del Ecuador, n.d.). Igualmente en el 2019 el crecimiento de la industria turística fue del 6,1% llevándola a representar una mayor proporción del PIB ecuatoriano (Ministerio de Turismo del Ecuador, 2019). Sin embargo para el 2020 con la llegada de la pandemia del COVID-19 muchas empresas de alojamiento “no tienen establecidos planes de sostenimiento integrales, los cuáles son necesarios para el desarrollo óptimo de un destino turístico y la satisfacción de los usuarios” (Pinargote y Loor, 2021, p. 2). Debido a esto al nivel de ingresos obtenidos por el sector hotelero en Marzo del 2020 en comparación con el año anterior disminuyó en un 58% (Ministerio de Turismo del Ecuador, 2020).

En base al párrafo anterior se llega a tener una cierta idea de la situación que estaba atravesando el sector hotelero en el país durante la última década. Sin embargo cabe mencionar que dentro de este sector se presentan algunos problemas como la falta de información hacia usuarios externos tales como bancos o inversionistas potenciales. Este problema también aplica a las personas encargadas de administrar los hoteles, puesto que sin la información necesaria es muy difícil para ellos elaborar planes de sostenimiento integrales que sean capaces de contribuir con el desarrollo del sector hotelero y por lo tanto de la industria turística ecuatoriana (Orlandini et al., 2019). Otro problema que enfrenta este sector es la falta de determinación de las variables más relevantes que expliquen la probabilidad de cierre de las empresas dedicadas al alojamiento (Álvarez-Ferrer y Campa-Planas, 2020).

Con todos los problemas planteados anteriormente, el objetivo principal de esta investigación es determinar la probabilidad de cierre de las empresas en el sector hotelero del Ecuador en base a información financiera del periodo 2011-2021. En relación a esto se puede establecer que la importancia de este trabajo radica en primer lugar en saber cuál de los dos modelos empleados tiene una mejor capacidad de predicción de quiebra del sector hotelero, el modelo logístico o el modelo de redes neuronales y mediante los resultados obtenidos a través de estos, llegar a generar información relevante para la toma de decisiones de los usuarios tanto internos como externos (Remache, 2021).

Por último este trabajo está compuesto de 5 apartados. En el primer capítulo se presenta el marco teórico, en el cual se incluyen todos los conceptos más relevantes relacionados con los temas de cierre empresarial, modelo de redes neuronales y modelo logístico. El segundo apartado se centra en la revisión de literatura, aquí se presentan los aspectos más importantes de los artículos que guardan una estrecha relación con el tema de esta investigación, tales como la metodología utilizada, los resultados obtenidos de los modelos empleados y las conclusiones a las que llegaron los autores. En tercer lugar se encuentra la metodología, dentro de este apartado se explica los diversos pasos que tuvieron que llevarse a cabo para la obtención de los resultados, que empieza con el pre procesamiento de los datos para la obtención de una base de datos estandarizada, seguida de una partición de ésta, en datos entrenamiento y de prueba; a esto le siguen la elaboración de los modelos, donde se explica sus respectivas estructuras y las funciones que están inmersas en el proceso de entrenamiento de cada modelo. También se menciona a la matriz de confusión con sus varios elementos de clasificación que ambos modelos deberán cumplir y finalmente se incluyen a las variables empleadas y su relación con la variable dependiente que también es explicada en este apartado.

El cuarto capítulo contiene la parte de resultados. En primer lugar se presenta un análisis descriptivo de las variables, en donde se visualiza la cantidad de datos pertenecientes a la variable dependiente cuando ésta es igual a 0 y 1, además se presentan medidas estadísticas como la media y desviación estándar de cada variable, a esto le sigue una matriz de correlación que incluye a todas las variables explicativas usadas. De ahí se exhiben los resultados obtenidos por los modelos, en el caso del modelo logístico, la variable más relevante es el endeudamiento y este presenta un nivel de precisión del 97,7%. Para las redes neuronales la precisión de identificación de los verdaderos positivos y negativos fue en cambio del 63,58% dando a entender que la capacidad de predicción de quiebra del modelo logístico es mayor que la de redes neuronales. El quinto apartado de la presente investigación está destinado para la discusión, mediante la cual se llega a contrastar en primera instancia las metodologías del pre procesamiento de los datos de este trabajo con otros artículos. En segundo lugar se compara los modelos, las variables empleadas y los resultados obtenidos de esta investigación con la de otros trabajos cuyo enfoque sea el de análisis de quiebra

en el sector hotelero. En el último apartado se realizan las debidas conclusiones de la investigación donde además se mencionan las debilidades y la apertura a futuras ramas de investigación en este tema.

1.1. Objetivos

Objetivo general

Determinar la probabilidad del cierre de las empresas en el sector hotelero del Ecuador en base a información financiera del periodo 2011-2021.

Objetivos específicos

- Revisar literatura actualizada de artículos científicos sobre probabilidad de cierre de empresas.
- Establecer el modelo de redes neuronales en el sector hotelero a través de datos de entrenamiento, permitiendo vincular al mismo con la información de la muestra.
- Elaborar un modelo logístico que sea capaz de determinar el peso que tiene cada variable explicativa sobre la variable dependiente, y realizar una comparación entre ambos modelos.

1.2. Marco teórico

En los siguientes párrafos se presentan los conceptos más relevantes que están relacionadas con el análisis de probabilidad de cierre de empresas dentro del sector hotelero ecuatoriano. En primer lugar se mencionan las definiciones de las variables financieras más importantes que van a ser aplicadas a los modelos posteriormente.

En primera instancia el riesgo de bancarrota según Horváthová y Mokrišová (2018) se refiere a la posibilidad de que una empresa no tenga fondos suficientes para cumplir con sus obligaciones de deuda en el futuro. Además, este riesgo está inversamente relacionado con los flujos de caja de la empresa. En otras palabras, si una empresa tiene menos activos líquidos, el riesgo de bancarrota será mayor. Es importante señalar que esta posibilidad existe durante toda la vida de la empresa. Por último, cabe mencionar que las entidades financieras tienen en cuenta dichos riesgos a la hora de conceder préstamos.

Otro concepto fundamental en el presente trabajo son los riesgos financieros, los cuales Albarracín Muñoz y García Arenas (2017) llegan a definir como todas las variables económicas y financieras que aparecen en el mercado, afectando negativamente la rentabilidad y el nivel de inversión de la empresa. Cabe destacar que el riesgo financiero está compuesto por el riesgo de mercado, el riesgo de crédito y el riesgo de liquidez. Finalmente, este tipo de riesgo se identifica al comprar bienes y servicios ofrecidos en el mercado financiero.

En tercer lugar se encuentra el cierre empresarial, que acorde con Orellana Osorio et al. (2020) se lo puede entender por un lado como una situación en donde una empresa se encuentre imposibilitada de seguir operando dentro de un mercado en específico y por otro como la escasez de recursos financieros, lo cual no permite la continuidad de las operaciones corrientes. Además es importante señalar que el cierre de una empresa está relacionado con varios factores, los cuales pueden estar ligados con el entorno en el que se desenvuelve la industria, la propia entidad y los administradores que desarrollan su actividad. Basándose en lo anterior, se puede decir que el cierre de una empresa puede ser causado por diversos factores no financieros (González Martínez et al., 2018).

También es importante mencionar a la liquidez, dado que se encuentra dentro de un análisis de cierre de empresas, por lo que acorde con Zeballos Puma (2017), este término se refiere a la capacidad de una empresa para cumplir con sus obligaciones de deuda a corto plazo y continuar manteniendo el efectivo después de realizar los pagos. El hecho de que la empresa cuente con suficiente liquidez brinda varias ventajas, como una mejor imagen con las instituciones financieras y también le da a la misma un mayor poder de negociación en relación con varios tipos de endeudamiento.

Dentro de las definiciones de indicadores financieros se encuentra el apalancamiento financiero, que según Sanchez Gil (2018) es cuando una empresa utiliza su deuda a largo plazo para financiar sus operaciones. El objetivo de éste es obtener una mayor ganancia que si el proyecto se financiara únicamente con capital. Cuando el apalancamiento se vuelve positivo, los accionistas se benefician porque el uso de la deuda aumenta la rentabilidad financiera de la empresa. Mientras que un escenario negativo, la deuda dañará los activos de la empresa y por lo tanto, esto perjudicará a los accionistas (Temprano, 2015).

La última definición dentro de los conceptos financieros es la del capital de trabajo, el cual según Palomeque (2008) se refiere a los activos corrientes de la empresa, en otras palabras hace énfasis a los recursos que se pueden convertir en dinero dentro de un año. Por otro lado Arreiza y Gavidia (2019) indican que el capital de trabajo es la diferencia entre el activo y el pasivo a corto plazo, por lo que se convierte en una herramienta muy práctica para determinar la eficiencia financiera de una empresa dentro del periodo de un año.

Por otro lado también se incluyen los conceptos más importantes relacionados con los modelos que se buscan aplicar dentro de esta investigación.

En primer lugar están las redes neuronales, que según Matich (2001) tienen diversas definiciones, entre las cuales resaltan, el ser un nuevo método de computación que está basado en modelos biológicos, y que “son redes interconectadas masivamente en paralelo de elementos simples (usualmente adaptativos) y con organización jerárquica, las cuales intentan interactuar con los objetos del mundo real del mismo modo que lo hace el sistema nervioso biológico” (Matich Damián, 2001, p. 8). En base a esto se puede decir que este método llega a ser de gran ayuda para los análisis de cierre empresarial.

Como complemento al concepto anterior en el trabajo de Matich (2001), se puede estipular que el contar con un aprendizaje adaptivo, llega a ser una de las características más significativas del modelo de redes neuronales, dado que a través de un entrenamiento, estas llegan a aprender cómo realizar determinada tarea, como por ejemplo ser capaces de identificar patrones dentro de un conjunto de datos. Por lo cual se puede establecer que las redes neuronales al ser auto adaptativas llegan a estar en un constante estado de cambio dadas ciertas condiciones, siendo de gran utilidad al momento de determinar la probabilidad de quiebra.

En cuanto a la parte econométrica de esta investigación es importante mencionar que la econometría acorde con Wooldridge (2010), es un conjunto de métodos estadísticos que buscan estimar relaciones económicas, las cuales están sujetas a diversas teorías, además mediante estas herramientas se pretende desarrollar políticas eficientes tanto para sectores públicos como privados.

Otro concepto que se pretende profundizar dentro de la metodología de este trabajo, son los modelos logísticos binarios, los cuales según Sifuentes-Amaya y Ramírez-Valverde (2010) este tipo de modelos se emplean cuando se busca determinar la relación entre una variable cualitativa dicotómica con diversas variables independientes que pueden llegar a ser tanto de índole cuantitativa como cualitativa. En otras palabras estos modelos son de suma utilidad para obtener la probabilidad de que un evento pueda ocurrir dadas ciertas variables explicativas.

2. Revisión de literatura

Los artículos que se presentan a continuación abordan varios temas relacionados con el riesgo de quiebra y bancarrota de empresas pertenecientes a diversos sectores productivos en varios países, especialmente al sector hotelero. Es importante mencionar que a pesar de que se emplean varios métodos de predicción, los artículos más relevantes para este trabajo son aquellos que utilizan modelos logísticos y redes neuronales artificiales o ambos. En primer lugar se presentan artículos en donde solo se ha incorporado un único modelo de análisis de quiebra, comenzando por aquellos relacionados con análisis estadísticos iterativos de discriminación múltiple.

Para el sector hotelero en España se han realizado varias investigaciones en cuanto a su desarrollo. Trabajos como los de Álvarez-Ferrer y Campa-Planas (2020); Vivel-Búa et al. (2018) tienen como finalidad principal analizar la estabilidad financiera y determinar las principales variables que tienen mayor peso dentro del sector ya mencionado en un periodo de tiempo dado. Por un lado en el primer artículo se utiliza dos aplicaciones de la fórmula Z de Altman, mientras que en el segundo se llega a emplear la fórmula ya mencionada pero aplicando datos de panel. En la primera investigación los resultados indican que el nivel de predicción de las fórmulas Z1 y Z2 de Altman para empresas insolventes dentro del sector hotelero se encuentran sobre el 80%. En cambio, con la fórmula ZAMR, se pudo observar que la predicción siempre marca resultados por debajo del 80%, llegando a un 56% en el año -5. En el otro artículo los resultados muestran que la eficiencia es el factor más relevante cuando analizamos el Z-score de Altman, mientras que la liquidez es más relevante cuando se estudia el Z-score. Para concluir en ambos trabajos se incluyeron indicadores financieros como endeudamiento, rentabilidad, y liquidez, también se consideraron variables

exógenas como la localidad de la empresa, la estacionalidad, la concentración del mercado entre otros como las principales variables que afectan a la salud del sector hotelero español.

Otras investigaciones como las de Milašinović et al. (2019); Maharani y Melynda Sari (2021); Horváthová y Mokrišová (2018) pretenden de igual manera analizar la probabilidad de quiebra de empresas que se encuentran en el sector hotelero y turístico a través de modelos Z-Score. Los resultados en el primer trabajo muestran que del 2014 al 2018 no existieron cambios notables en la cantidad de empresas que se encontraban en la zona segura (sin amenaza de bancarrota) y zona problemática (bancarrota), sin embargo los autores determinaron que el método de Data Envelopment Analysis presenta mayor capacidad de predicción que el modelo Z de Altman indicando además que es importante que las empresas se fijen en el ratio de costos. En el segundo artículo los autores plantean que el modelo S-Score mostró un escenario muy pesimista, mientras que el modelo X-Score arrojó resultados bastante positivos en cuanto a empresas en estado de quiebra. A pesar de que cada método presentó diferentes condiciones, algunas empresas fueron clasificadas de igual manera por ambos modelos. Por otro lado, el último artículo indicó que la mayoría de los hoteles obtuvieron un alto riesgo de quiebra. Cabe mencionar que los autores dieron a entender que este modelo presenta una capacidad de predicción de quiebra bastante buena, dado que lo recomendaron para futuras investigaciones.

Además existen ciertos estudios que abordan este problema desde una distinta dirección, como Escribano-Navas y Germar (2021) quienes buscan comprender cómo la variable género de los Directores Ejecutivos puede llegar a afectar la subsistencia de los hoteles. Para esto los autores emplearon un modelo econométrico log-log que cuenta con una estructura de datos de panel. Variables como el logaritmo de la variable de ventas, la experiencia, es decir, años en la industria hotelera antes del período de estudio, el sexo del CEO y la relación entre el capital de trabajo y los activos totales fueron estadísticamente significativas. Este resultado sugiere que las ventas y la experiencia se correlacionan positivamente con la supervivencia del hotel, ya que los hoteles dirigidos por mujeres tienen tasas de supervivencia más altas. En conclusión, se puede argumentar que las empresas hoteleras deberían considerar la elección de mujeres gerentes para ayudar a estas empresas a hacer frente a la crisis provocada por la pandemia de Covid-19.

Por otro lado Gemar et al. (2019); Vivel-Búa et al. (2019); Espino-Rodríguez y Gil-Padilla (2017) buscan mediante sus respectivas investigaciones establecer qué variables son determinantes en la supervivencia de distintos tipos de hoteles. En cuanto a los métodos utilizados en los diferentes artículos, en el primero se hizo uso de una regresión semiparamétrica con riesgos proporcionales de Fox, en el segundo se aplicó modelos de riesgo y para el tercer artículo se empleó un análisis factorial. Los resultados muestran que ciertas variables son consistentes en estos artículos, tales como la ubicación del hotel, el tamaño, la estacionalidad o el ciclo comercial. Como conclusión Gemar et al. (2019) muestra que los hoteles tipo resort necesitan aplicar estrategias de diferenciación que permitan predecir diversos problemas, mientras que los autores del segundo artículo fomentan el desarrollo de modelos de supervivencia empresarial a nivel de industria que tengan en cuenta las especificidades de las industrias estratégicas. Finalmente, el tercer documento concluyó que los hoteles se clasificaron en tres grupos distintos en función de las prioridades en competencia. El primero realiza esto bajo un enfoque de costes y flexibilidad de los mismos, el segundo grupo están enfocados hacia la excelencia y un último grupo que considera relevante las operaciones que guardan relación con la calidad y el servicio.

En cuanto a trabajos relacionados con redes neuronales se encuentran Karahuta et al. (2017) y del Castillo y Fernández (2021), quienes pretenden crear modelos que les permitan realizar una mejor gestión dentro de los sectores hoteleros. Los resultados de ambas investigaciones dan a conocer que el uso de redes neuronales generan una buena bondad de ajuste de los datos empleados, dado que en el primer trabajo mediante un test de Pearson se determinó que las redes neuronales presentan una mayor exactitud predictiva que los modelos convencionales para la toma de decisiones de gestión. En el segundo trabajo también se pudo evidenciar que la capacidad de clasificación de las redes neuronales en relación a si una empresa hotelera se encuentra o no en bancarrota (Patrimonio negativo) son mayores al 90%. De lo mencionado anteriormente, el primer artículo llegó a la conclusión de que los modelos de redes neuronales son más precisos y confiables que otros modelos de predicción convencionales. Por otro lado, la segunda investigación concluyó que la inclusión de variables como el flujo de caja y el margen de beneficio neto en la industria hotelera, de agencias de viajes y de catering ayudaron al modelo global a lograr predicciones más precisas.

Del mismo modo es importante mencionar aquellas investigaciones que llegan a realizar una comparación del poder de predicción entre dos o más modelos. En la mayoría de los artículos que se presentan a continuación, los autores realizan un contraste entre los modelos logísticos y modelos desarrollados en base a redes neuronales artificiales sobre la quiebra de distintos sectores productivos, sin embargo también se llegan a incluir otros modelos como el de Análisis Discriminante Múltiple, modelos estructurales, entre otros. Cabe recalcar que el objetivo principal de estos era por un lado determinar las variables más relevantes del análisis de quiebra y por otro establecer qué modelo es el más apropiado.

En los trabajos donde se concluye, que el uso de redes neuronales brinda una capacidad de mayor predicción, están en primer lugar do Prado et al. (2019), Inam et al. (2019), Kim (2011) y Mansouri et al. (2018), quienes dan a entender que dentro del modelo Logit, cero (0) representa un escenario positivo para la salud del sector en cuestión, mientras que uno (1) indica un escenario pesimista del mismo. Los dos primeros artículos señalan las variables más significativas de sus estudios, siendo para do Prado et al. (2019) retorno sobre patrimonio, margen neto, desglose de deuda entre otros; en cambio los autores del segundo trabajo dieron a entender que lo más significativo son los ratios financieros SALTA y EBITCL. En cuanto al modelo de redes neuronales en ambos casos los datos se llegaron a ajustar de manera adecuada, mostrando resultados similares al modelo logístico. Para el tercer y cuarto artículo en relación al modelo Logit se llegó a determinar que en ambos casos la mayoría de los coeficientes, que representan distintos ratios financieros, mostraban una relación directa con la variable dependiente. Además en ambas investigaciones se emplearon redes neuronales MLP que llegaron a tener un ajuste bastante alto con los datos de la muestra.

Cabe añadir que los artículos de Becerra-Vicario et al. (2020) y Abidin et al. (2020) también indican que las redes neuronales son más efectivas que los modelos logísticos. En relación a los resultados la primera investigación, da a entender que dentro del modelo Logit la mayoría de las variables independientes tienen una relación inversa con la variable explicada, además tanto en este modelo como en el de redes neuronales la variable más significativas fue Beneficio/ Ingreso neto. En la segundo trabajo, el modelo de redes neuronales clasificó a la variable liquidez como la más relevante al realizar pronósticos de quiebra, mientras que el modelo Logit indicó que es importante tener en consideración el tamaño de la junta directiva.

Por otro lado se encuentran los artículos de Fernández-Gámez et al. (2016), da Kammoun y Triki (2016) y Zizi et al. (2021) que llegaron a la conclusión de que los modelos logísticos son más precisos que los modelos de redes neuronales. Los resultados de la primera investigación muestran que la variable más importante a un año antes de la quiebra es el EBITDA a los pasivos circulantes, mientras que si este periodo de tiempo se extiende a dos o tres años anteriores el retorno sobre los activos es el mejor predictor de quiebra. El segundo artículo muestra que dentro del modelo Logit 5 de 7 variables tienen una relación directa con la variable dependiente (riesgo crediticio), además se determinó que el modelo en su conjunto es significativo y tiene un buen ajuste de datos. De igual manera el modelo de redes neuronales también llega a clasificar correctamente los datos en un 72%. Por último los resultados del tercer trabajo muestran que dentro del modelo Logit el valor cero (0) representa a PYMES saludables y el valor de uno (1) si éstas se encuentran afligidas. En cuanto a los coeficientes, la mayoría presentan una relación positiva con la variable dependiente. Por otro lado al hablar de precisión de los modelos, el de redes neuronales presentó una exactitud del 88,33%, mientras que uno de los modelos logísticos llegó a obtener un precisión del 93,33% de los datos.

Siguiendo una dinámica similar, los trabajos de Barreda et al. (2016), Pereira et al. (2017); Wieprow y Gawlik (2021) pretenden predecir y evaluar el riesgo de insolvencia de las empresas de la industria hotelera en los respectivos países. Cabe mencionar que las tres investigaciones en cuestión hicieron uso tanto de modelos logísticos como de modelos de Análisis Discriminante Múltiple Sin embargo los resultados de cada estudio no son los mismos. En Barreda et al. (2016) donde se examina la probabilidad de bancarrota (incapacidad tanto de cubrir obligaciones como de reorganizar las deudas de los activos), la precisión de predicción de ambos modelos fue de 76,7%. En cambio en el segundo artículo donde ambos modelos indicaron que en la primera mitad del 2019, el 55% de las empresas encuestadas estaban en riesgo de bancarrota (declarados insolventes o adquiridas por una tercera parte), mientras que en la primera mitad del 2020 solo dos compañías no presentaban riesgo de quiebra. Por otro lado en el último artículo se indica que en el modelo MDA la variable beneficio operativo/ costos operativos es la más importante, en cuanto al modelo logit la capacidad de predicción sobre bancarrota (Patrimonio negativo) de este sector disminuye conforme se aleja del año en el que fue estimado. En conclusión las investigaciones establecieron que el modelo Logit es más conveniente dado su peso teórico, sin embargo ambos modelos son de suma relevancia. Dado que en el primer artículo la capacidad predictiva fue mayor del 80%, en el segundo trabajo

los autores enfatizan que este modelo no necesita de supuestos de distribuciones normales ni de covarianzas equivalentes entre empresas solventes e insolventes. En cambio los modelos del último artículo presentaron un nivel de predicción del 76,7%.

En cuanto a trabajos donde se realizan pronósticos a futuro de la industria hospitalaria, se encuentra el artículo de Matejić et al. (2022), cuyo propósito fue determinar riesgo de quiebra producido por la crisis sanitaria del COVID-19 en la industria hotelera en la República de Serbia dentro del periodo 2020-2026, en base a una determinada muestra poblacional. Este enfoque consiste en emplear cinco nuevos modelos de series de tiempo estructurales y redes neuronales. Los resultados muestran que la crisis comenzó a tener un impacto negativo en el riesgo de insolvencia en 2020, y se espera que este impacto llegue a incrementarse hasta 2023. No se espera que la situación del sector mejore para el 2024, pero incluso en 2026 el riesgo de insolvencia seguirá siendo alto en contraste con los años anteriores al COVID-19. En base a esto, se concluye que las redes neuronales artificiales y los modelos estructurales presentan resultados bastante similares.

3. Métodos

En primer lugar, el tipo de investigación que se lleva a cabo en este trabajo es de naturaleza cuantitativa, ya que es necesario analizar una base de datos para lograr los objetivos planteados en el apartado anterior. En otras palabras, el trabajo tiene un enfoque positivista, donde se busca sacar conclusiones a partir de resultados obtenidos objetivamente, que a su vez están científicamente fundamentados (Monje, 2017).

La base de datos mencionada anteriormente, que fue otorgada por el Observatorio empresarial de la Universidad del Azuay, se encuentra compuesta de cuentas contables extraídas tanto del estado de situación financiera como del estado de resultados de empresas hoteleras ecuatorianas, pertenecientes al sector de Actividades de Alojamiento y de Servicio de comidas (I-55), durante el periodo 2011 al 2021, la mayoría de estos documentos se encuentran presentes dentro de la Superintendencia de Compañías, Valores y Seguros. También se hallan datos que fueron obtenidos de otras fuentes como el Instituto Nacional de Estadísticas y Censos (INEC).

Es importante dar a conocer que se ha decidido emplear toda la población para el análisis de cierre empresarial del sector hotelero ecuatoriano, el cual está compuesto de 235 empresas. Cabe añadir que las base de datos ya mencionada está compuesta de datos de panel, es decir que la información financiera de los hoteles que son de corte transversal, llegaron a ser recopiladas a lo largo de un periodo, en este caso de diez años, además dichas observaciones son no balanceadas, en otras palabras, algunas unidades de ciertos elementos no llegan a verse reflejados en algunos años dentro del periodo establecido (Wooldridge, 2010).

Sin embargo antes de aplicar la base de datos mencionada en los respectivos modelos, es necesario que estos se encuentren estandarizados, es decir, eliminar los posibles valores atípicos que podrían sesgar los resultados. Para esto se pretende emplear valores z , la cual se define como la distancia que se encuentra un valor en específico sobre o debajo de la media y que es medido en desviaciones estándar (Universidad Técnica del Norte, 2022). La fórmula se expresa en la ecuación 1.

Ecuación 1

$$z = \frac{X - \mu}{\sigma}$$

Al haber obtenido estos valores se busca eliminar aquellos datos cuyo valor z sea mayor o menor a tres desviaciones estándar.

También es necesario mencionar que la base de datos va a ser dividida con el propósito de que con el 80% de los elementos que se encuentran dentro de esta, llegue a generar las estimaciones debidas por parte del modelo logístico y por otro lado que las redes neuronales lleguen a entrenar a dichas conexiones, además con el 20% de los datos restantes se pretende emplearlos para la prueba de estos modelos. Otra cuestión que también es de suma importancia para el análisis de los modelos y que será explicada con mayor profundidad posteriormente es la normalización de los datos numéricos, a excepción de aquellos pertenecientes a la variable dependiente, esto se realiza con el objetivo de interpretar los coeficientes del modelo. En cuanto a las variables nominales, estas serán transformadas en variables cualitativas de respuesta cero o uno. Por último se pretende eliminar a una de dos variables en el caos que estas presenten

una correlación mayor al 90%, dado que esto puede llegar a indicar indicios de multicolinealidad entre variables explicativas.

Otro aspecto a incluir dentro de esta sección es el llegar a establecer las distintas metodologías que se esperan aplicar a cada objetivo específico. Dado que el primero de ellos incluye una revisión de la literatura relevante sobre el fracaso empresarial en varios sectores productivos de la economía, es importante señalar que el conocimiento bibliográfico existente sobre el tema permite tener una imagen más clara a cerca de qué están tratando de resolver otros autores, qué métodos utilizan y qué limitaciones enfrentan; dicho de otra manera, si se cuenta con este tipo de información, el investigador puede adquirir mayor conocimiento teórico sobre el problema que pretende resolver (Monje, 2017).

Para obtener solo información relevante sobre los diversos artículos considerados, se planea crear un fichaje donde se llegue a registrar la fecha de publicación del artículo, el problema planteado por el autor, los métodos utilizados, los resultados obtenidos y las conclusiones más importantes.

Como ya se mencionó anteriormente, dentro de esta investigación se busca aplicar un modelo logístico, cuya definición fue presentada dentro del marco teórico de este trabajo, sin embargo antes de profundizar con respecto a la metodología del mismo, es importante conocer que éste, al clasificarse como un modelo de respuesta binaria, tiene como principal enfoque la obtención de probabilidad de un suceso en específico (Wooldridge, 2010), esto se ve reflejado en la Ecuación 2.

Ecuación 2

$$P(y = 1|x) = P(y = 1|x_1, x_2, \dots, x_k)$$

En primer lugar la formula indica que el valor de “Y”, que puede llegar a adoptar los valores de cero y uno, va a depender de ciertas variables independientes (x), en otras palabras la probabilidad de que suceda un suceso Y va a estar determinado de ciertas variables explicativas (Wooldridge, 2010). Bajo esta lógica se pueden formular diversos modelos de respuesta binaria que llegan a estructurarse de la siguiente manera (Ecuación 3).

Ecuación 3

$$P(y = 1|x) = G(\beta_0 + \beta_1 x_1 + \dots + \beta_k x_k) = G(\beta_0 + x\beta)$$

Como se planteó anteriormente “Xn” representa cada una de las variables explicativas, mientras que los betas son los coeficientes respectivos obtenidos en base a un conjunto de datos en particular. Por otro lado “G” representa una función donde solo se pueden obtener valores reales entre cero y uno. Existen diversas funciones para “G”, en el caso del modelo Logit, “G” se presenta en la Ecuación 4 (Wooldridge, 2010).

Ecuación 4

$$G(z) = \frac{1}{1 + e^{-z}} = \frac{e^z}{1 + e^z}$$

Aquí “zi” representa todos los números reales entre cero y uno que llega a obtenerse mediante la función expresada anteriormente, es decir, “z” se puede expresar de la siguiente manera (Ecuación 5).

Ecuación 5

$$z = (\beta_0 + \beta_1 x_1 + \dots + \beta_k x_k)$$

Cabe recalcar que este modelo al no operar con datos de origen lineal, es decir, que al no emplear una metodología de MCO, no es posible estimar este modelo basándose directamente en los Betas, por lo que es necesario obtener los Odd ratios, los cuales se pueden definir como el logaritmo natural de la razón entre la probabilidad de que se cumpla cierta condición y la probabilidad de que la misma no se dé (Gujarati y Porter, 2010). Esto se puede observar en la Ecuación 6.

Ecuación 6

$$Li = \ln\left(\frac{Pi}{1 - Pi}\right) = zi = \beta_0 + \beta_1 X_{1i} + \dots + \beta_k X_{ki}$$

Mediante este proceso se puede ver que tanto las variables explicativas como los parámetros llegan a linealizarse.

Una vez ya obtenidos los Odd ratios respectivos es de suma relevancia tener en cuenta que si el valor es cercano o igual a 1 la variable no genera una clasificación correcta del modelo, por lo que no es relevante. Por otro lado si el valor es menor a 1, este se debe ser dividido para 1, mientras que si el valor es mayor a 1 se llega a interpretar el coeficiente de manera directa.

Otro aspecto importante dentro del análisis de cierre de empresas del sector hotelero ecuatoriano es la elaboración de un modelo de predicción de cierre mediante el uso de redes neuronales, cuya arquitectura será de tipo secuencial, es decir, que los nodos de una capa van estar conectados con aquellos que se encuentran en capas anteriores y en próximas. En cuanto a la estructura interna, la red se divide en tres partes principales, la primera es la función de entrada, la cual mediante un vector entrada (conjunto de todas las entradas) busca combinar todos los valores que se encuentran dentro de la entrada de la red, esto se expresa de forma matemática en la Ecuación 7. (Matich, 2001).

Ecuación 7

$$\text{inputi} = (in_{i1} \times w_{i1}) * (in_{i2} \times w_{i2}) * \dots * (in_{in} \times w_{in})$$

Dentro de esta función “*” representa un cierto operador como por ejemplo una sumatoria o una productoria, “n” es la cantidad de entradas que llegan a la neurona, la cual en este caso será igual al número de variables empleadas. “Ni” y “wi” hace referencia al peso que tiene cada valor de entrada, la función de este elemento al multiplicarse con cada entrada es determinar la influencia que cada valor dentro de la neurona va a tener (Matich, 2001).

El segundo componente dentro de la neurona es la función de activación, que se encuentra en las capas intermedias de la red y para este análisis solo se contará con una capa que estará conformada de cinco nodos. La tarea de esta función consiste en que mediante una modificación de la entrada global (se emplea este término cuando ciertos valores de entrada son tratados como uno solo) se llegue a obtener un valor de activación que se encuentra dentro de un rango, el cual dependerá de la función aplicada (Matich, 2001). Dado que se busca comparar la efectividad de las redes neuronales con el modelo logístico se planteó implementar la función de activación Sigmoide, cuyo rango de valores es de 0 a 1 y que se define de la siguiente manera (Ecuación 8).

Ecuación 8

$$\sigma(z) = \frac{1}{1 + e^{-z}}$$

Dentro de la Ecuación 8 e representa una constante exponencial, mientras que “z” es equivalente a la suma ponderada de entradas dentro de la red neuronal. Esta se expresa de la siguiente manera.

$z = b + \sum_i w_i x_i$ (“b” es denominado como un sesgo, el cual determina la predisposición de la neurona en dar un uno o cero independientemente de los pesos de las entradas) (Github, n.d.)

La última parte dentro de la estructura de una red neuronal es la función de salida, la cual determina el valor que va a ser expulsado de la neurona en cuestión hacia aquellas que estén vinculadas, cabe recalcar que las neuronas no llegan a aceptar cualquier tipo de valores de entrada, por lo que los valores de salida deben estar dentro de un rango como de 0 a 1 (Matich, 2001). Cabe mencionar que dicha función se ubica en la capa de salida de la red y para este artículo solo está compuesta de un nodo.

El siguiente paso para la elaboración de la red neuronal es la compilación del modelo, en donde uno de los aspectos importantes a tomar en consideración es la función de costo. Para este caso se va a emplear la función de costo de entropía binaria cruzada, la cual se expresa mediante la siguiente ecuación.

Ecuación 9

$$C = -\frac{1}{n} \sum_x [y \ln a + (1 - y) \ln(1 - a)]$$

Dentro de la Ecuación 9 “n” representa el número de elementos de datos de entrenamiento, “x” representa las entradas de entrenamiento, que la función mediante la sumatoria busca cubrir, “y” por otro lado se define como la salida deseada o esperada, mientras que “a” es la verdadera salida de la neurona (Nielson, n.d.).

Es importante añadir que existen ciertas propiedades que hacen posible que la función de entropía binaria cruzada sea una función de costo. En primer lugar el resultado obtenido va a ser un número no negativo dado que por un lado todos los logaritmos pertenecen a números que se encuentran en el rango de 0 a 1,

por lo que los términos de forma individual dentro de la suma son negativos, y estos son compensados por el signo negativo ubicado al inicio de la parte derecha de la función. Otra propiedad importante a mencionar es que si la salida obtenida es cercana a la salida deseada para todos las entradas de entrenamiento, entonces la función de entropía cruzada será cercana a 0. Dado que el presente trabajo es un análisis de clasificación, los resultados van a ser cero o uno (Nielson, n.d.).

Otro aspecto a tener cuenta dentro de este paso es la determinación del optimizador a emplear, para este trabajo se pretende emplear el algoritmo de optimización ADAM, que se define como un método de dos momentos adaptativos en el tema de optimización estocástica, el cual a través de estrategias heurísticas realiza una serie de pasos para la resolución de un problema. Comenzando por entender dicho problema, seguido de la elaboración de un plan de solución, se llega a ejecutar el plan, continuado de un análisis de resultados. Las ventajas de este algoritmo llegan a ser desde su simplicidad, eficiencia computacional, invariancia y parámetros intuitivos de fácil calibración (Rojano et al., 2021).

El tercer paso consiste en ajustar el modelo de redes neuronales mediante la aplicación de backpropagation a la data para obtener los pesos óptimos. Para que este tipo de red de aprendizaje supervisado tome lugar en el proceso es necesario que se perciba un patrón en la entrada de la red, el cual se propagará en el resto de capas internas hasta obtener una salida, esta llega a ser comparada con una salida deseada y la diferencia llega a ser denominada como señal de error que es calculada en cada salida (Cornejo y Quispe, 2011).

Estas salidas de error llegan a propagarse desde la capa de salida hacia las capas internas de la red, sin embargo cada neurona ubicada en las capas intermedias solo perciben un valor relativo del error el cual va a depender de la contribución de dicha neurona para la obtención del resultado final, este proceso se repite hacia las otras capas hasta que todas las neuronas involucradas perciban aquella proporción de la señal de error total. Mediante este proceso cada nodo es capaz de ajustar los pesos de conexión permitiendo a la red generar una mejor clasificación de los datos empleados (Cornejo y Quispe, 2011).

Una vez culminado el entrenamiento de la neurona mediante los ajustes mencionados en el párrafo anterior, los nodos que se encuentran en las capas internas de la red serán capaces de responder con una salida activa si el patrón empleado contiene aquella característica que ha estado presente con los datos de entrenamiento (Cornejo y Quispe, 2011).

En cuanto al proceso de entrenamiento de la red neuronal para esta investigación, es importante indicar que el proceso que se mencionó en los párrafos anteriores llega a denominarse como un batch, cuya información obtenida va a servir como base para un segundo batch y así sucesivamente hasta llegar a quinientos batches, toda esta cantidad de elementos se denomina como epoch, Al haber culminado dentro del primer epoch la información del último batch irá al primero del segundo epoch, en donde se seguirá el mismo procedimiento hasta que se generen 100 epochs, cada uno de 500 batches.

Tanto para el modelo logístico como para el que se basa en redes neuronales, se va a emplear tablas de métricas de predictibilidad, donde se analizará en primer lugar la exactitud de clasificación, que llega a representar aquella proporción que tuvo éxito en la clasificación de datos tanto positivos como negativos (Souza et al., 2021). Esta se obtiene mediante la siguiente fórmula que se expresa en la Ecuación 10 .

$$\frac{TP + TN}{(TN + FP + FN + TP)}$$

TP=Verdaderos positivos TN= Verdaderos negativos FP = Falsos positivos FN = Falsos negativos

Otro elemento importante que se encuentra dentro de esta matriz es el análisis de sensibilidad, la cual se define como la proporción de verdaderos positivos correctamente identificados sobre el total de de los mismos más los falsos negativos (Souza et al., 2021) . Este es representado en la Ecuación 11

$$\frac{TP}{(FN + TP)}$$

También es importante tener en cuenta la especificidad del modelo, es decir, la proporción de los verdaderos negativos que se llegaron a clasificar de manera correcta en relación a la suma de los mismos con los falsos

positivos (Souza et al., 2021), esto se expresa matemáticamente de la siguiente manera dentro de la Ecuación 12.

Ecuación 12

$$\frac{TN}{(FP + TN)}$$

En cuanto a las variables que se pretenden emplear dentro de este análisis son pertenecientes al modelo de Ohlson (1980), el cual presenta un modelo Logit cuyo propósito es determinar la probabilidad de que una empresa se encuentre en riesgo de insolvencia, generando un valor de uno, cuando el patrimonio de esta es negativa y obtiene pérdida al final del periodo (Orellana et al., 2020). En la Tabla 1 se puede observar un breve análisis de las variables independientes que conforman este modelo, es decir, se puede llegar a ver qué fórmulas fueron aplicadas a cada variable, de qué fuente se llegan a obtener los datos, la relación que cada una de estas tiene con la variable dependiente y una descripción de cómo se llegan a interpretar. Es importante añadir que la variable dependiente está estructurada bajo el criterio de que una empresa va a considerarse en situación de quiebra en un determinada año cuando esta presenta tanto un patrimonio negativo como pérdida neta.

Tabla 1

Variables independientes

Variables	Var. desagregadas	Fuente	Relacion	Descripción
Tamaño				Dentro de la base de datos se presentan empresas de distintos tamaños: Microempresas, pequeñas, medianas y grandes empresas de alojamiento
pasivos totales/A. totales	Pasivos totales Activos totales	Superintendencia de Compañías	Directa	Al haber un incremento en el endeudamiento de la empresa, la probabilidad de que esta quiebre será mayor
Capital de trabajo/ A. totales	Cap. De trabajo	Superintendencia de Compañías	Inversa	Llega a ser un indicador de liquidez que al aumentar, la probabilidad de quiebra de una empresa disminuirá
Pasivo corriente/ Activo Corriente	Pasivo Corriente Activo corriente	Superintendencia de Compañías Superintendencia de Compañías	Directa	Mientras menor sea el endeudamiento a corto plazo, la probabilidad de quiebra va a reducirse
Dummy de solvencia		Superintendencia de Compañías	Inversa	Variable dicotómica, es igual a 1 cuando en los dos últimos años los pasivos hayan sido mayores a los activos en su totalidad
Utilidad neta/ Activos totales	Utilidad neta	Superintendencia de Compañías	Inversa	A medida que se incrementa el ROA, la probabilidad de quiebra disminuye.
Resultado op./ Total obligac	Resultado operativo Total obligaciones	Superintendencia de Compañías Superintendencia de Compañías	Inversa	Si la utilidad operacional es mayor al pasivo total, la probabilidad de quiebra se reduce
ROE	Ingresos netos Patrimonio total medio	Superintendencia de Compañías	Inversa	A medida que se incrementa el ROE, la probabilidad de quiebra disminuye.

Nota: Descripción y estructura de las variables independientes y su relación con la variable dependiente
Fuente: (Orellana et al., 2020)

4. Resultados

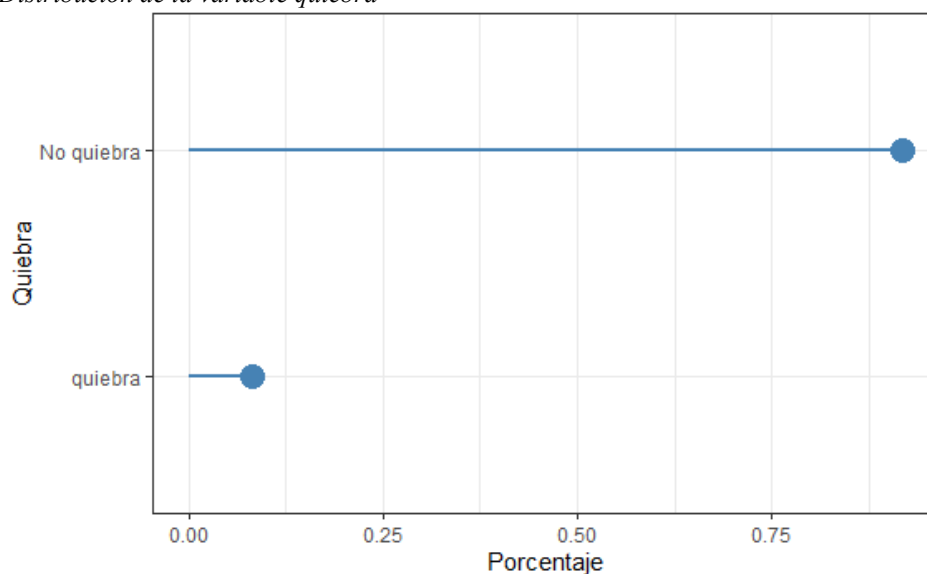
4.1. Análisis descriptivo base de datos

Variable dependiente

En la siguiente gráfica se puede observar la distribución de la variable quiebra.

Gráfica 1

Distribución de la variable quiebra



Nota: Se muestra el porcentaje de empresas de alojamiento que llegaron a clasificarse en quiebra

En la Gráfica 1 claramente se ve que alrededor del 87% de las observaciones que se encuentran dentro de la variable dependiente indican que las empresas hoteleras a través de los años dentro del periodo establecido no presentan quiebra, es decir, no llegan a constar de patrimonio negativo ni presentan pérdida neta alguna durante un año en específico. Por otra parte cerca del 13% de los datos indican que una empresa dentro del sector hotelero presentó pérdida neta y patrimonio negativo.

La Tabla 2 muestra la relación entre la variable dependiente con las distintas variables significativas mediante diversos estadísticos descriptivos.

Tabla 2

Análisis descriptivo variables independientes

Variable	Observaciones		Media		Des. Estandar		Mínimo		Máximo	
	No quiebra	Quiebra	No quiebra	Quiebra	No quiebra	Quiebra	No quiebra	Quiebra	No quiebra	Quiebra
Tamaño	1213	109	2,015664	2,055046	0,8456364	0,9111944	1	1	4	4
Endeudamiento	1213	109	0,5116623	2,80417	0,4807285	3,7209	0,0002294	1,001358	9,533525	26,15557
Indicador de liquidez	1213	109	0,0305402	-1,51421	0,4311882	3,452189	-8,533525	-25,15557	0,9978629	0,9877204
Endeudamiento corto plazo	1213	109	5,002613	11,1694	34,23417	22,83741	0,0021371	0,0119381	753,6465	124,6556
Solvencia	1213	109	0,0568838	0,0458716	0,2317159	0,2101728	0	0	1	1
ROA	1213	109	0,0119982	-0,8417874	0,2177727	1,494018	-3,144163	-6,985287	2,236139	-0,0018506
U. operac./ Pasivos totales	1213	109	-0,0043121	-0,2279818	2,049226	0,3163562	-21,47035	-1,49404	21,24271	0,6788458
ROE	1213	109	-0,0341583	0,0264383	1,711821	0,6965267	-19,55577	-2,917124	18,0563	2,803948

Nota: Esta tabla muestra ciertos estadísticos descriptivos para cada variable independiente, cuando la variable dependiente es igual tanto a Quiebra como No Quiebra.

Como se puede observar gran parte de las observaciones de la variables explicativas llegan a estar asociadas con empresas hoteleras que dentro del periodo establecido no han mostrado patrimonio negativo, ni pérdidas netas, es decir que no se encuentran en quiebra. En primer lugar se encuentra la variable Tamaño, cuyo promedio en ambos casos no varía de manera notable, dando a entender que los valores en los que se encuentran clasificados los distintos tamaños de empresas de alojamiento para cuando la variable dependiente es “No quiebra” y “Quiebra” son bastante similares. Algo parecido se evidencia con la desviación estándar, la cual también consta de valores muy parecidos, indicando que la dispersión de los datos en relación a la media es muy semejante cuando la variable dependiente llega a ser clasificada tanto

como “Quiebra” o “No quiebra”. Dado que esta es una variable cualitativa, el rango de valores es de tres en los dos casos.

En relación con la variable endeudamiento, se puede ver que los datos dentro de la categoría “No quiebra” presentan una menor dispersión, también se puede observar que la media, cuando la variable dependiente es igual a “Quiebra”, es notablemente mayor, dando a entender que el nivel de endeudamiento de las empresas pertenecientes a este grupo es mucho mayor. El indicador de liquidez por su parte muestra que en promedio las empresas que presentan un patrimonio negativo y pérdidas netas llegan a contar con una menor liquidez. Además mediante esta tabla se puede determinar que cuando la variable dependiente es igual a “Quiebra”, las observaciones relacionadas con el índice de liquidez se encuentran mucho más dispersos en relación a la media. Por último la diferencia entre los valores mínimos y máximos en ambas categorías es notorio.

Como cuarta variable está el Endeudamiento a corto plazo, la cual indica que en promedio las observaciones que se encuentran dentro de la categoría “Quiebra” de la variable dependiente muestran un mayor endeudamiento a corto plazo, pero un aspecto importante a tomar en consideración es que la desviación estándar del endeudamiento a corto plazo cuando la variable explicada presenta la etiqueta de “No quiebra” es notablemente mayor, esto se ve reflejado en los rangos de la variable, especialmente al contrastar los valores máximos. Por otro lado tenemos a la variable solvencia, que en cuanto a su media poblacional, se puede ver que los valores de ambas categorías no difieren de manera notable, la diferencia puede llegar a darse por la cantidad de datos. Asimismo la desviación estándar tampoco varía notablemente, dando a entender que los datos se encuentran igualmente dispersos en relación a su media respectiva.

En sexto lugar se encuentra la variable ROA, cuyos resultados muestran que la dispersión de los datos es mayor dentro de la categoría “Quiebra”, ya que cuenta con una mayor desviación estándar en relación a la media. En cuanto a esta se puede ver que el ROA promedio cuando la variable dependiente da como resultado “No quiebra” es positivo, mientras que cuando esta es igual a “Quiebra” el Retorno medio sobre los Activos es negativo, indicando que en base a los datos empleados, en promedio si no se llega a contar con un patrimonio negativo y pérdida neta a su vez el Retorno sobre Activos no genera beneficio alguno.

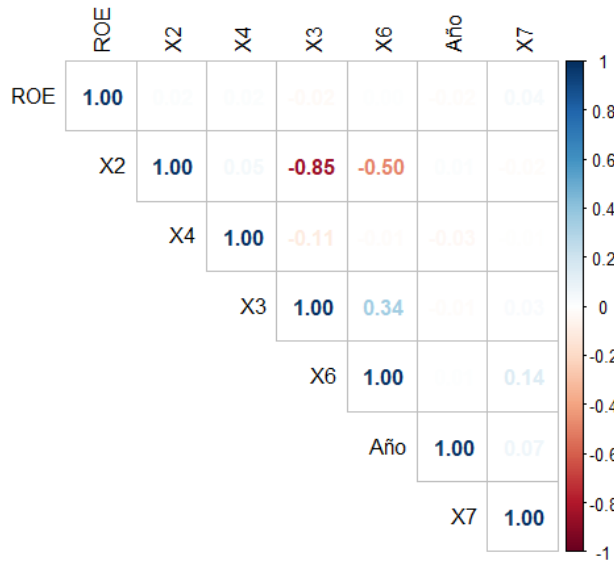
También se encuentra presente el ratio Utilidad operacional/ Pasivos totales, el cual, en base a los resultados de la tabla en cuestión, indica que la desviación estándar para el conjunto de datos en donde la variable dependiente es “No quiebra” es mayor. Sin embargo es importante mencionar que dentro de este grupo el valor máximo muestra que la utilidad operacional es notablemente mayor a los pasivos totales de dicha empresa hotelera, mientras que en la otra categoría el valor máximo indica que la utilidad operacional del hotel analizado no supera sus obligaciones totales, llegando a incrementar su probabilidad de riesgo de quiebra. Las medias en ambos casos son negativas, indicando que en ambos grupos un gran número de empresas dentro de la población presentan pérdidas operativas.

Por último se encuentra la variable ROE, donde se puede ver en primera instancia que la media de la misma, cuando la variable dependiente está clasificada bajo la etiqueta de “Quiebra”, es positiva Sin embargo menor a 1, dando a entender que la utilidad operativa promedio es menor al patrimonio medio. Por otro lado se puede observar que de igual manera el valor máximo de la categoría “No quiebra” es mayor al valor máximo del otro grupo, pero en ambos casos se muestra que la utilidad operativa de las empresas es mayor que el patrimonio de las mismas.

4.2. Análisis de correlación

Gráfica 2

Correlación variables independientes



Nota: Matriz de correlación entre variables explicativas.

En la Gráfica 2 se puede observar que la correlación que se presenta entre las diversas variables independientes es baja, las correlaciones más notables son en primer lugar entre las variables X3(Capital de trabajo/Activos totales) y X6 (Utilidad neta) donde se presenta una relación directa débil entre las variables, en otras palabras a medida que el indicador de liquidez aumente, el ROA también lo hará. Otra correlación que destaca dentro de la presente matriz, es la de la variable X2(Endeudamiento) con la variable X5(Dummy de solvencia), la cual a diferencia de la anterior presenta una relación inversa y un tanto más moderada. Finalmente se encuentra la correlación entre las variables X2 y X3 que también presentan una relación inversa, dicho de otra manera, mientras el endeudamiento de la empresa llega a incrementarse, el indicador de liquidez será cada vez menor. Cabe mencionar que a pesar de ser considerada una correlación alta, ésta no supera el límite de 0,9 por lo que se pretende quedar con todas las variables analizadas anteriormente.

4.3. Modelo Logístico- Datos de entrenamiento

4.3.1. Análisis de coeficientes

En la Tabla 4 se puede apreciar los coeficientes ajustados de cada una de las variables en cuestión con sus respectivos indicadores de significancia.

Tabla 3*Modelo logístico con datos de entrenamiento*

Variable	Coefficiente Beta	Desviación estándar	Estadístico	Valor p
Intercepto	-2,62	0,227	11,5	7,94E-31
ROE	-0,0218	0,163	0,134	8,94E-01
X2	3,37	0,539	-6,25	4,07E-10
X3	-0,125	0,435	0,289	7,73E-01
X4	0,174	0,0797	-2,18	2,94E-02
X6	-1,42	0,36	3,96	7,62E-05
X7	0,329	0,198	-1,66	9,75E-02
Tamaño Microempresa	-0,361	0,368	0,98	3,27E-01
Tamaño Mediana	0,0449	0,378	-0,119	9,05E-01
Tamaño Grande	0,286	0,604	-0,474	6,35E-01
Tamaño (otra)	-0,319	0,774	0,412	6,80E-01

Nota: Esta tabla muestra el signo, los coeficientes, desviación estándar, estadístico y valor p de cada variable independiente.

En primer lugar es importante reconocer qué variable es la más relevante para el presente modelo y es así como a través de los coeficientes ajustados se pudo llegar determinar que la variable endeudamiento (X2) es la variable que más peso tiene dentro del modelo logístico en base a los datos empleados, dando entender que las empresas hoteleras deben tener especial atención a los indicadores de endeudamiento al momento de analizar la salud de la misma. La variable ROA también llega a ser de gran importancia para determinar la probabilidad de no quiebra de una empresa, por lo que los gerentes también deberían tener especial consideración con este indicador financiero.

En cuanto a la relación que guardan las variables explicativas con la variable quiebra, se puede establecer en primer lugar que el Retorno sobre el Patrimonio (ROE) presenta una relación negativa con la variable no quiebra, es decir, a medida que aumente el ROE la probabilidad de quiebra de una empresa dentro del sector hotelero ecuatoriano será menor. En segundo lugar está la variable endeudamiento, la cual indica que mientras mayor sea el endeudamiento la probabilidad de quiebra dentro del sector hotelero ecuatoriano será mayor. Por otro lado la variable X3 da a conocer que a medida que el cociente entre el Capital de trabajo sobre los activos totales aumente, la probabilidad de quiebra disminuirá.

La variable X4 muestra una relación directa con la variable dependiente, en otras palabras a medida que se incrementa el endeudamiento a corto plazo la probabilidad de quiebra dentro del sector hotelero será mayor. La variable X6 en cambio indica que a medida que el ROA de una empresa aumenta, la probabilidad de quiebra se verá reducida. En cuanto a la variable X7 se puede observar que a medida que la utilidad operacional es mayor a los pasivos de una empresa dentro del sector hotelero, la probabilidad de quiebra aumentará. Por último la variables relacionadas con el tamaño de la empresa muestran que cuando un hotel se clasifica como Microempresa, entonces guarda una relación inversa con la probabilidad de quiebra, caso contrario la variable tamaño muestra una relación directa con la variable dependiente.

4.3.2. Significancia de las variables

En primer lugar es necesario plantear las debidas hipótesis para determinar la significancia de una variable independiente dentro del modelo en cuestión. La hipótesis nula establece que los coeficientes de las variables explicativas son iguales a 0, es decir que no son significativas para el modelo, mientras que la hipótesis alternativa da a entender que los coeficientes difieren de 0, indicando de esta forma que la variable analizada sí llega a aportar al modelo. En segundo lugar se decidió implementar un nivel de significancia de 0,05; dado que este nivel es el más común para proyectos de investigación. A través de este nivel es posible comparar los valores p de cada una de las variables independientes y se llega a rechazar la Hipótesis nula en caso de que el valor p sea menor al 5%.

Como se puede observar en la tabla 4, en base al conjunto de datos aplicado y con un nivel de significancia de 0,05, las variables que llegan a ser consideradas como significativas para el modelo son: el endeudamiento, el endeudamiento a corto plazo y el ROA.

Tabla 4

Métricas de predictibilidad (solo con datos de entrenamiento)

Métrica	Estimador	Resultado
accuracy	binario	0,958
kap	binario	0,457
sens	binario	0,357
spec	binario	0,992
ppv	binario	0,714
npv	binario	0,965
mcc	binario	0,487
j_index	binario	0,349
bal_accuracy	binario	0,675
detection_prevalence	binario	0,0264
precision	binario	0,714
recall	binario	0,357
f_meas	binario	0,476

Nota: Esta tabla muestra el porcentaje de precisión, sensibilidad y especificidad del modelo.

En primer lugar se encuentra el porcentaje de precisión de clasificación del presente modelo, el cual al ser mayor al 80% da a entender que el modelo realizó una buena clasificación tanto de los verdaderos positivos como de los verdaderos negativos. Sin embargo en esta tabla se puede observar que la sensibilidad del modelo logístico se encuentra muy por debajo de los niveles esperados, lo cual indica una mala identificación de los verdaderos positivos de la base de datos empleada. Por último la especificidad de dicho modelo llega a estar por encima del 0,8; mostrando de esta manera que el modelo Logit sí llega a clasificar correctamente los verdaderos negativos.

4.4. Modelo logístico- Datos de prueba

4.4.1. Análisis de coeficientes

En base al 20% de la base de datos, en la Tabla 6 se puede observar los coeficientes ajustados de cada una de las variables pertenecientes al presente modelo

Tabla 5*Modelo logístico con datos de prueba*

Variable	Coefficiente Beta	Desviación estándar	Estadístico	Valor p
Intercepto	-6,32	1,52	4,16	0,0000314
ROE	1,01	1,17	-0,861	0,389
X2	4,78	1,35	-3,54	0,000408
X3	-1,22	0,688	1,78	0,0752
X4	-0,667	0,397	1,68	0,0934
X6	-1,09	1,29	0,846	0,397
X7	1,72	6,81	-0,253	0,801
Tamaño Microempresa	-1,7	2,03	0,835	0,404
Tamaño Mediana	0,752	1,44	-0,523	0,601
Tamaño Grande	1,45	1,68	-0,862	0,388
Tamaño (otra)	-13,4	2429	0,00553	0,996

Nota: Esta tabla muestra el signo, los coeficientes, desviación estándar, estadístico y valor p de cada variable independiente.

Al igual que en la sección de datos de entrenamiento, aquí la variable Endeudamiento (X2) llega a ser la variable más relevante del modelo logístico, dado que su coeficiente ajustado es el mayor entre todas las variables. Sin embargo un aspecto de este modelo que llega a diferir del anterior es que la variable ROA llega a perder influencia dentro del modelo para determinar la probabilidad de no quiebra, siendo ésta remplazada por el ratio financiero Utilidad operativa/Pasivo total.

Por otro lado algunos de los signos que presentan los coeficientes de cada variable llegan a diferir del modelo anterior. En primer lugar está la variable ROE, la cual indica que a medida que aumente el retorno sobre el patrimonio de las empresas del sector hotelero, su probabilidad de quiebra va a ser mayor. Con la variable Endeudamiento a corto plazo (X4) se llega una situación similar, dado que el signo en este modelo indica que a medida que el endeudamiento a corto plazo aumenta, la probabilidad de quiebra disminuirá.

Existen variables que no fueron mencionadas y esto se debe a que los signos que presentan en el modelo son los mismos que el anterior. Dando a entender que cuando las variables Endeudamiento de las empresas(X2), Utilidad operativa/Pasivo total de los hoteles(X7), Tamaño de empresa mediana y Tamaño de empresa grande aumentan, la probabilidad de quiebra del sector hotelero ecuatoriano aumenta. En cambio cuando las variables Capital de trabajo/Activos totales (X3), Retorno sobre Activos (X6) y Tamaño Microempresa aumentan, la probabilidad de quiebra del sector productivo en cuestión será menor.

4.4.2. Significancia de las variables

Al igual que en la sección anterior la hipótesis nula hace referencia a que la presente variable no es significativa para el modelo en base a los datos empleados, mientras que la hipótesis alternativa indica lo contrario. Del mismo modo se aplica un nivel de significancia del 5%, es decir que se va a rechazar la H0 de cada variable independiente en el caso que el p valor sea menor a dicho nivel.

Como se puede observar en la Tabla 6, acorde a la proporción de base de datos empleada y con un nivel de significancia de 0,05; la variable endeudamiento es la única variable que llega a considerarse como significativa para el modelo.

Tabla 6*Métricas de predictibilidad con datos de prueba*

Métrica	Estimador	Resultado
accuracy	binaria	0,977
kap	binaria	0,757
sens	binaria	0,714
spec	binaria	0,992
ppv	binaria	0,833
npv	binaria	0,984
mcc	binaria	0,760
j_index	binaria	0,706
bal_accuracy	binaria	0,853
detection_prevalence	binaria	0,045
precision	binaria	0,833
recall	binaria	0,714
f_meas	binaria	0,769

Nota: Esta tabla muestra el porcentaje de precisión, sensibilidad y especificidad del modelo.

Al igual que en la matriz de confusión del modelo logístico con datos de entrenamiento, la precisión de este para clasificar tanto los verdaderos positivos como negativos es bastante alta (98%). En cuanto a la especificidad se puede observar el mismo resultado para ambos casos, es decir, que tanto al emplear datos de entrenamiento como datos de prueba el modelo Logit logra una efectividad de clasificación de los verdaderos negativos del 99%. Sin embargo dentro de la sensibilidad del modelo se pudo determinar que al aplicar los datos de prueba la capacidad de identificación de verdaderos positivos por parte del modelo es mucho mayor (71%), aunque aún se encuentra debajo de los niveles aceptables (0,8).

4.5. Redes neuronales

Al haber culminado con el análisis del modelo logístico, a continuación se presenta la Tabla 8 donde se muestra la estructura de la red neuronal empleada para el conjunto de datos.

Tabla 7*Estructura de la red neuronal*

Capa (tipo)	Forma de salida	Núm. Parámetros
Intermedia	5	50
Salida	1	6
Parámetros totales		56
Parámetros entrenables		56
Parámetros no entrenables		0

Nota: Esta tabla muestra la estructura funcionamiento de la capa intermedia y de salida de la red neuronal aplicada.

Como se puede observar se presentan dos tipos de capas. En primer lugar está la capa intermedia del modelo, que como ya se mencionó en la metodología de este trabajo está conformada de cinco neuronas y tiene 50 parámetros. Por otro lado la segunda capa, hace referencia a la capa de salida del modelo de la red, la cual como ya se explicó en las secciones pasadas está conformada de solo un nodo de salida, además esta capa contiene seis parámetros. En base a lo mencionado se puede decir que el modelo de redes neuronales presente tiene un total de 56 parámetros.

Tabla 8*Métricas de predictibilidad -Redes neuronales*

Métrica	Estimador	Resultado
accuracy	binario	0,636
kap	binario	0,178
sens	binario	0,778
spec	binario	0,618
ppv	binario	0,205
npv	binario	0,958
mcc	binario	0,253
j_index	binario	0,395
bal_accuracy	binario	0,697
detection_prevalence	binario	0,427
precision	binario	0,205
recall	binario	0,778

Nota: Esta tabla muestra la precisión, sensibilidad y especificidad del modelo de redes neuronales.

Mediante esta matriz de confusión es posible observar, al igual que en el otro modelo, la capacidad de clasificación del modelo de redes neuronales, donde en primer lugar muestra una precisión de del 64% aproximadamente. Si se llega a evaluar este resultado con los del modelo logístico, se puede ver que la precisión de clasificación de verdaderos positivos y verdaderos negativos en relación con el total de datos en este modelo es menor. Algo similar se evidencia en la especificidad del modelo, dado que la capacidad de clasificación correcta de verdaderos negativos de la base de datos empleada es menor para las redes neuronales (62%) que para el modelo logístico. Sin embargo en relación con la sensibilidad, se puede observar que este modelo presenta una capacidad de identificación de verdaderos positivos del 77%, que llega a ser mayor a la del modelo logístico.

5. Discusión

Con todo lo planteado en las secciones anteriores se busca generar un contraste tanto en el aspecto metodológico como en el de resultados, entre esta investigación y otros artículos científicos similares. En relación con la transformación de la base de datos, se llevó a cabo un re muestreo en donde se llegó a dividir la base de datos en dos conjuntos, el primero estuvo compuesto del 80% de los datos que fueron empleados para el entrenamiento de los modelos y el otro grupo con el 20% de la base total, fue utilizado para la prueba de los mismos; todo esto con el propósito de tener un conjunto de datos balanceados. Cabe mencionar que varios artículos científicos llegan a aplicar metodologías similares, las cuales se presentan en los siguientes párrafos.

Dentro del artículo de Abidin et al. (2020); se realiza una comparación en la capacidad de predicción de quiebra en el sector hospitalario en Malasia y para realizar dicho análisis el 70% de la base de datos total empleada fue destinada para la estimación del modelo logístico y el de redes neuronales siendo el 30% restante empleado para probar el poder predictivo de cada uno de los modelos en cuestión, los autores no llegan a indicar ningún tipo de tratamiento en cuanto a la estandarización de los datos.

Otro trabajo en donde se aplicó una metodología análoga es el de Zizi et al. (2021), cuyo objetivo principal era determinar los ratios más relevantes para la predicción de quiebra empresarial en Marruecos. En cuanto al balanceo de datos los autores decidieron aplicar el método de Sobre-muestreo que busca generar un ratio aceptable entre clase minoritaria con clase mayoritaria, para esto se aplicó el algoritmo SMOTE que les permite generar muestras sintéticas para las clases minoristas. Además también se llegó a dividir la muestra en dos submuestras, una que contiene el 75% de los datos que fue utilizada para el entrenamiento de los modelos y el 25% restante de los datos de la muestra fueron empleados en la validación de estos.

Kroeze et al. (2019), que tienen como objetivo principal contrastar el modelo de predicción de quiebra de Altman con un modelo alternativo de quiebra de tres variables para la industria de aerolíneas, generaron su base de datos mediante un enfoque censal en vez de muestreo aleatorio. Los autores además decidieron aplicar una muestra reservada con el propósito de no superponer la data que buscan emplear para el entrenamiento de los modelos con el conjunto destinado para la predicción de los mismos. El primer grupo mencionado está compuesto de 62 puntos de datos, mientras que el segundo grupo está conformado de 22 puntos de datos.

En el trabajo de Mansouri et al. (2018), que buscan comparar la capacidad de predicción de quiebra entre el modelo logístico y de redes neuronales para empresas que cotizan en bolsa en Tehran, llegaron a emplear un muestreo longitudinal para la obtención de la base de datos, la cual al igual que en los otros artículos llegaron a dividirla en datos de entrenamiento y de prueba para ambos modelos, sin embargo para el primer conjunto emplearon únicamente los datos del año 2013, mientras que para el conjunto del test usaron las observaciones del periodo 2010-2012.

Por otro lado también se encuentran artículos que emplean una metodología diferente en cuanto al procesamiento de las bases de datos, como es el caso de Kim (2011), quien mediante este análisis busca generar una predicción acertada de la probabilidad de bancarrota en el sector hotelero de Estados Unidos. Para la construcción de la base de datos se empleó la metodología de bases emparejadas, cuyos emparejamientos se realizaron máximo hasta el tercer años antes de que los hoteles se consideraran en quiebra. La principal premisa utilizada para resaltar la diferencia entre empresas hoteleras quebradas y no quebradas es la cantidad de activos totales que poseen.

En cuanto a la parte de resultados, en esta sección se busca contrastar los hallazgos obtenidos en el presente trabajo, con aquellos de otros artículos científicos cuyo objetivo principal es determinar la capacidad de predicción de dos o más modelos en la probabilidad de quiebra de empresas dentro del sector hotelero de un país en específico. En relación con los modelos aplicados en primer lugar están los artículos de Park y Hancer (2012); y de Youn y Gu (2010), que al igual que en este trabajo aplicaron modelos logísticos y de redes neuronales para determinar la quiebra dentro de los sectores hoteleros de Estados Unidos y Corea del Sur respectivamente. En cambio el trabajo de Vivel-Búa et al. (2016) solo emplea modelos logísticos Logit y Probit para el sector hotelero español, no obstante Castillo y Fernández (2021) realizan un análisis de probabilidad de quiebra en los distintos sectores de la industria turística de España utilizando únicamente un modelo de redes neuronales Perceptrón Multicapa.

Antes de comparar los modelos se pretende mencionar las variables implementadas en cada uno de los artículos. Park y Hancer (2012) emplean 16 ratios financieros ligados a la liquidez, solvencia, rentabilidad, eficiencia y flujo de caja. En cuanto a la variable dependiente, 0 representa no bancarrota y 1 representa aquellas empresas que legalmente se han declarado en bancarrota. Youn y Gu (2010) implementan 10 ratios financieros similares al anterior artículo, sin embargo ellos no consideran la rentabilidad dado que las variables dentro de esta categoría están sumamente ligadas al ingreso neto, el cual se encuentra dentro de la variable dependiente, que adopta el valor de 1 si la empresa ha presentado pérdidas netas durante tres años consecutivos. Vivel-Búa et al. (2016) también llega a implementar ratios financieros que se encuentran dentro de las mismas categorías ya mencionadas, sin embargo ellos además incluyen variables externas relacionadas con la industria hospitalaria. En este caso la variable dependiente adoptará el valor de 1 para aquellas empresas que dentro del periodo 2008-2011 hayan estado enfrentando problemas de solvencia o estaban a punto de hacerlo. En último lugar del Castillo y Fernández (2021) emplean ratios financieros pertenecientes a las mismas categorías de los anteriores artículos y la variable dependiente está relacionada con la situación legal de bancarrota. En base a esto las variables explicativas utilizadas para esta investigación son similares a las de estos artículos, sin embargo para la variable dependiente ninguno de los autores emplea al mismo tiempo los dos criterios que fueron empleados en este trabajo (Pérdida neta y Patrimonio negativo).

Para los modelos de redes neuronales Park y Hancer (2012) emplearon una capa de entrada de seis neuronas, esto debido a que de las 16 variables solo se quedaron con las significativas para el modelo; una capa intermedia de dos neuronas y una capa de salida solo con una neurona. El proceso de aprendizaje fue el mismo al de la presente investigación, es decir, se llegó a aplicar el algoritmo de Backpropagation, en donde se empleó la función de entrada, la función de activación Sigmoidal y la función de salida que fueron ya explicadas en la sección de la metodología. Youn y Gu (2010) por su parte decidieron implementar de igual manera tres capas, una capa de entrada, cuya cantidad de neuronas es equivalente al número de variables presentes que son diez; una capa intermedia oculta, en donde se emplea también la función de activación

Sigmoide y una capa de salida que contiene una sola neurona y genera resultados entre los valores de 0 y 1. Por otra parte la red MLP empleada por del Castillo y Fernández (2021) está conformada igualmente de una capa de entrada, una capa intermedia y una capa de salida. El procesamiento de información se realiza mediante un algoritmo de corrección de errores, dado que llega a determinar la desviación entre los valores de entrada y salida y en base a esto la red modifica sus parámetros para minimizar dicha diferencia. A esto los autores adicionan un análisis de sensibilidad para determinar el peso que tiene cada variable para explicar el problema en cuestión.

Al contrastar los resultados obtenidos con el modelo de redes neuronales, entre esta investigación con los artículos mencionados en el párrafo anterior, se obtiene que la capacidad de precisión de clasificación de sus modelos es mayor al del presente trabajo. En Park y Hancer (2012) la precisión total de las redes neuronales para identificar los verdaderos positivos y negativos de la base de datos es del 97,5%. Youn y Gu (2010) por su parte muestran que con los datos de entrenamiento la capacidad de precisión del modelo tanto para los verdaderos positivos como para los negativos es de 90,20% y 85,29% respectivamente. Sin embargo con los datos de prueba la precisión de identificación de verdaderos positivos disminuye a 72,73%, mientras que para los verdaderos negativos se incrementa a un 90,91%. En el caso de del Castillo y Fernández (2021) los resultados indican que en base a los datos de entrenamiento (80%) el modelo aplicado a este sector cuenta con una capacidad de precisión de clasificación del 91,73%, mientras que con los datos de prueba (20%) la precisión de la red neuronal es de 90,21%.

En el caso del modelo logístico Park y Hancer (2012) muestran que la variable que llega a ser la más representativa para la predicción de quiebra del sector hotelero de Estados Unidos es el endeudamiento (Pasivos totales/Activos totales), además este modelo presenta un nivel de precisión de clasificación de datos del 90% de un total de 80 empresas del sector. En cambio Youn y Gu (2010) indican que la única variable que llegó a ser significativa y por lo tanto es considerada como la más relevante para la determinación de quiebra es la cobertura de interés, que guarda una relación inversa con la variable dependiente. En cuanto al nivel de clasificación del modelo, empleando datos de entrenamiento, este muestra una capacidad de precisión de verdaderos positivos y verdaderos negativos de 80,39% y 86,27% respectivamente. Sin embargo al haber empleado los datos de test, la precisión del modelo al identificar los verdaderos positivos disminuye a 63,64%, mientras que ésta se incrementa para los verdaderos negativos (90,91%). Por último en el trabajo de Vivel-Búa et al. (2016) se llegó a generar cuatro diferentes modelos, cada uno con un modelo Logit y Probit. En base a estos, los autores llegaron a determinar que dentro de los ratios financieros la variable Rentabilidad (Ingreso operativo/Total de activos) es la más relevante dado que presenta el coeficiente más alto. Para las variables del sector, la concentración del mercado hospitalario en cada destino turístico llega a ser la más relevante. En cuanto a la capacidad de clasificación, el modelo tres, en donde no se toma en cuenta la variable Delegaciones, tiene la mayor capacidad de identificación de los verdaderos negativos, mientras que el modelo cuatro, que toma en cuenta todas las variables tiene una mejor capacidad de identificar los verdaderos positivos. Al comparar los resultados de estos artículos con los de la presente investigación, se puede apreciar en primer lugar, que en ambos casos las variables más significativas están ligadas al endeudamiento de las empresas y por otro lado la capacidad de predicción del modelo logístico en este trabajo es mayor que en todos los artículos mencionados.

6. Conclusión

El principal propósito de esta investigación era determinar la probabilidad de quiebra del sector hotelero ecuatoriano durante un periodo de 10 años. Para lograr dicho objetivo en primer lugar fue necesario la construcción de la base de datos, la cual estuvo conformada de cuentas contables de los Estados de Situación Financiera y de Resultados de 235 empresas pertenecientes al sector anteriormente mencionado. En base a esta información fue posible llegar a estructurar una nueva base de datos, conteniendo tanto la variable dependiente como las variables explicativas.

Para responder a los objetivos de la investigación, metodológicamente se desarrollaron los siguientes pasos:

- A. Preprocesamiento de los datos, que incluyó tratamiento de valores atípicos, nivelación, estandarización y partición de la base de datos en set de entrenamiento y set de prueba.
- B. La estimación del modelo para lo cual fue indispensable determinar las variables más relevantes que se explican en el siguiente párrafo.

En relación al horizonte temporal que se tomó para el estudio, este fue del 2011 al 2021, es decir que se consideró para la presente investigación, toda empresa que provee servicios de alojamiento en territorio Ecuatoriano que haya estado activa de manera completa o parcial dentro de este periodo. En cuanto a las variables empleadas como ya se mencionó en los párrafos anteriores la variable dependiente representa quiebra, es decir, si la empresa para un determinado año de periodo presenta quiebra, presenta patrimonio negativo y a su vez pérdida neta. Las variables explicativas Tamaño y Solvencia también son de índole cualitativa, por lo que también solo adoptaron valores de 0 y 1. Por otro lado las variables independientes ROE, Endeudamiento, Indicador de Liquidez, Endeudamiento a corto plazo, ROA y Utilidad Operacional sobre Pasivos totales son ratios financieros de naturaleza numérica.

Con respecto a los modelos empleados en esta investigación, el primero fue un modelo logístico, mediante el cual se pretendía en primer lugar obtener los coeficientes ajustados de cada variable independiente para determinar cuál es la más relevante al momento de determinar la probabilidad de quiebra del sector, además de analizar los signos que se encuentran antes de los coeficientes de cada variable explicativa y en segundo lugar se buscaba ver la capacidad de clasificación del modelo. Por otro lado está el modelo de redes neuronales, que se compuso de una capa de entrada con nueve neuronas, una capa intermedia oculta con cinco neuronas en donde se llegó a aplicar una función de activación Sigmoide y una capa de salida de una neurona. Para la optimización de la red se usó el algoritmo ADAM y el entrenamiento de esta se realizó mediante Backpropagation. De igual manera se buscaba determinar la capacidad de predicción de este modelo.

Los resultados obtenidos por parte del modelo logístico mostraron que la variable independiente más significativa del mismo, es el Endeudamiento (Pasivos totales sobre Activos totales) tanto al emplear los datos de entrenamiento como los datos de prueba. Por otro lado a pesar de que el modelo no llegó a contar con una muy buena sensibilidad (71,4%), este sí logró obtener muy buenos porcentajes en la especificidad (99,2%) y en la precisión de clasificación de los verdaderos positivos y verdaderos negativos (97,7%). En el caso del modelo de redes neuronales, este cuenta con una sensibilidad del 77,8%, mientras que la especificidad es del 61,7% y la capacidad de precisión de clasificación del modelo es del 63,5%. En base a estos resultados se pudo determinar que con este conjunto de datos el modelo logístico llega a tener una mejor predicción de quiebra que el modelo de redes neuronales.

En cuanto a las debilidades que se presentaron dentro de investigación, una de ella es que los datos se encontraban inicialmente desbalanceados, por lo que si no se realizaba una partición a la base de datos, los resultados de los modelos podían haberse encontrado sesgados y no hubiera sido posible realizar una comparación adecuada en la capacidad de predicción de quiebra. Otra debilidad que estuvo presente dentro de este trabajo es el hecho de que las Redes Neuronales no llegan a generar un modelo explicable, es decir, que a diferencia de los modelos tradicionales no es posible analizar las variables de manera separada, ni determinar su significancia o relevancia en relación con la variable dependiente.

En último lugar, para futuras investigaciones sería interesante que se realice un análisis sobre este sector pero desde un enfoque demográfico empresarial. También sería atractivo que dentro de la industria hotelera ecuatoriana se apliquen otros modelos de quiebra empresarial como aquellos que emplean Análisis Discriminante Multivariado, o modelos que operen bajo la metodología de Mínimos Cuadrados Ordinarios. Asimismo sería bastante novedoso aplicar métodos como árbol de decisiones o Vector Machine Learning dentro de la probabilidad de quiebra del sector hospitalario en Ecuador. En cuanto al uso de los modelos logísticos y redes neuronales, muchos otros sectores productivos en el Ecuador podrían beneficiarse de forma notable al emplear estos para la determinación de riesgo de quiebra, permitiendo a las diversas empresas tomar las medidas necesarias para evitarlo.

7. Referencias

- Abidin, Z. J., Abdullah, N. A. H., & Khaw, K. L.-H. (2020). Predicting SMEs Failure: Logistic Regression vs Artificial Neural Network Models. *Capital Markets Review*, 28(2), 29–41. https://www.mfa.com.my/cmrv28_i2_a3/%0Ahttp://creativecommons.org/licenses/by/4.0/
- Albarracín Muñoz, M., & García Arenas, L. F. (2017). Riesgo financiero: una aproximación cualitativa al interior de las mipymes en Colombia. *Aglala*, 8(1), 139–160. <https://doi.org/10.22519/22157360.1029>
- Álvarez-Ferrer, A., & Campa-Planas, F. (2020). The prediction of business failure in the hospitality business. *Cuadernos de Turismo*, 45, 33–59. <https://doi.org/10.6018/turismo.426031>
- Arreiza, E., & Gavidia, J. (2019). Gestión del capital de trabajo como estrategia financiera para el desarrollo empresarial. *Revista Valor Contable*, 6(1), 65–77.
- Barreda, A. A., Kageyama, Y., Singh, D., & Zubieta, S. (2016). Hospitality Bankruptcy in United States of America: A Multiple Discriminant Analysis-Logit Model Comparison. *Journal of Quality Assurance in Hospitality and Tourism*, 18(1), 86–106. <https://doi.org/10.1080/1528008X.2016.1169471>
- Becerra-Vicario, R., Alaminos, D., Aranda, E., & Fernández-Gámez, M. A. (2020). Deep recurrent convolutional neural network for bankruptcy prediction: A case of the restaurant industry. *Sustainability (Switzerland)*, 12(12), 1–15. <https://doi.org/10.3390/su12125180>
- Cornejo, D., & Quispe, G. (2011). *Aplicación del algoritmo Backpropagation de redes neuronales para determinar los niveles de morosidad en los alumnos de la Universidad Peruana Unión*. 21–31. https://revistas.upeu.edu.pe/index.php/ri_bi/article/download/908/876
- da Kammoun, A., & Triki, I. (2016). Credit Scoring Models for a Tunisian Microfinance Institution: Comparison between Artificial Neural Network and Logistic Regression. *Review of Economics & Finance*, 6(1), 61–78.
- del Castillo, A., & Fernández, S. (2021). *Predictive potential of the global bankruptcy models in the tourism industry*. 17(4), 23–31.
- do Prado, J. W., de Melo Carvalho, F., de Benedicto, G. C., & Lima, A. L. R. (2019). Analysis of credit risk faced by public companies in Brazil: An approach based on discriminant analysis, logistic regression and artificial neural networks. *Estudios Gerenciales*, 35(153), 347–360. <https://doi.org/10.18046/j.estger.2019.153.3151>
- Escribano-Navas, M., & Germar, G. (2021). Gender and Bankruptcy : A Hotel Survival Econometric Analysis. *MDPI*, 13(12), 1–13.
- Espino -Rodríguez, T. F., & Gil -Padilla, A. M. (2017). Configuraciones en el sector hotelero basadas en las prioridades competitivas y su relación con el tamaño, categoría y el resultado organizativo. *PASOS - Revista de Turismo y Patrimonio Cultural*, 15(1), 211–228.
- Fernández-Gámez, M. Á., Cisneros-Ruiz, A. J., & Callejón-Gil, Á. (2016). Applying a probabilistic neural network to hotel bankruptcy prediction. *Tourism & Management Studies*, 12(1), 40–52. <https://doi.org/10.18089/tms.2016.12104>
- Gemar, G., Soler, I. P., & Guzman-Parra, V. F. (2019). Predicting bankruptcy in resort hotels: a survival analysis. *International Journal of Contemporary Hospitality Management*, 31(4), 1546–1566. <https://doi.org/10.1108/IJCHM-10-2017-0640>
- Github. (n.d.). *Neural networks*. Github. https://ml4a.github.io/ml4a/es/neural_networks/
- González Martínez, R., Arteaga Cervantes, A. L., & Ruíz Torres, M. del R. (2018). Cierre empresarial en la región Laja- Bajío. *Management Review*, 3(2). <https://doi.org/10.18583/umr.v3i2.119>
- Gujarati, D. N., & Porter, D. C. (2010). *Econometría*. <http://journal.um-surabaya.ac.id/index.php/JKM/article/view/2203>
- Horváthová, J., & Mokrišová, M. (2018). Risk of bankruptcy, its determinants and models. *Risks*, 6(4). <https://doi.org/10.3390/risks6040117>
- Inam, F., Inam, A., Mian, M. A., Sheikh, A. A., & Awan, H. M. (2019). Forecasting Bankruptcy for organizational sustainability in Pakistan. *Journal of Economic and Administrative Sciences*, 35(3), 183–201. <https://doi.org/10.1108/jeas-05-2018-0063>
- Karahuta, M., Gallo, P., Matuľáková, D., Ľenková, A., & Ľambronská, K. (2017). Forecast of Using Neural Networks in the Tourism Sector. *CBU International Conference Proceedings*, 5, 218–223. <https://doi.org/10.12955/cbup.v5.928>
- Kim, S. Y. (2011). Prediction of hotel bankruptcy using support vector machine, artificial neural network, logistic regression, and multivariate discriminant analysis. *Service Industries Journal*, 31(3), 441–468. <https://doi.org/10.1080/02642060802712848>
- Kroeze, C., Zemke, D., Marie, V., & Raab, C. (2019). Improving Airline Bankruptcy Prediction. *The Journal of Hospitality Financial Management*, 26(2), 104–113.
- Maharani, R., & Melynda Sari, T. A. (2021). TOURISM INDUSTRY BANKRUPTCY ANALYSIS OF

- IMPACT COVID-19 PANDEMIC. *International Journal of Economics, Business and Accounting Research (IJEBAR)*, 5(4), 339–350.
- Mansouri, A., Nazari, A., & Ramazani, M. (2018). A Comparison of Artificial Neural Network Model and Logistics Regression in Prediction of Companies Bankruptcy (A Case Study of Tehran Stock Exchange). *SSRN Electronic Journal*, 6(24). <https://doi.org/10.2139/ssrn.2787308>
- Matejić, T., Knežević, S., Arsić, V. B., Obradović, T., Milojević, S., Adamović, M., Mitrović, A., Milašinović, M., Simonović, D., Milošević, G., & Špiler, M. (2022). Assessing the Impact of the COVID-19 Crisis on Hotel Industry Bankruptcy Risk through Novel Forecasting Models. *Sustainability (Switzerland)*, 14(8). <https://doi.org/10.3390/su14084680>
- Matich, D. (2001). Redes Neuronales: Conceptos Básicos y Aplicaciones. *Historia*, 55. <ftp://decsai.ugr.es/pub/usuarios/castro/Material-Redes-Neuronales/Libros/matich-redesneuronales.pdf>
- Milašinović, M., Knežević, S., & Mitrović, A. (2019). Bankruptcy forecasting of hotel companies in the Republic of Serbia using Altman's Z-score model. 7(2), 87–95. <https://doi.org/10.5937/menhottur1902087M>
- Ministerio de Turismo del Ecuador. (n.d.). *Ecuador muestra un crecimiento del 15% en la industria del turismo*. Ministerio de Turismo Del Ecuador.
- Ministerio de Turismo del Ecuador. (2019). *Promedio de llegada de visitantes extranjeros a Ecuador creció 4% en 2019*. Ministerio de Turismo Del Ecuador. <https://www.turismo.gob.ec/promedio-de-llegada-de-visitantes-extranjeros-a-ecuador-crecio-4-en-2019/>
- Ministerio de Turismo del Ecuador. (2020). *Plan de reactivación turística 2020*. 73. https://www.turismo.gob.ec/wp-content/uploads/2021/05/Plan-Reactivacion-Turistica-Red_compressed.pdf
- Monje, C. A. (2017). *Metodología de la investigación cuantitativa y cualitativa*. <https://www.uv.mx/rmipe/files/2017/02/Guia-didactica-metodologia-de-la-investigacion.pdf>
- Nielson, M. (n.d.). 3.1: The cross-entropy cost function. 1–12. [https://batch.libretexts.org/print/url=https://eng.libretexts.org/Bookshelves/Computer_Science/Applied_Programming/Book%3A_Neural_Networks_and_Deep_Learning_\(Nielsen\)/03%3A_Improving_the_way_neural_networks_learn/3.01%3A_The_cross-entropy_cost_function.pdf](https://batch.libretexts.org/print/url=https://eng.libretexts.org/Bookshelves/Computer_Science/Applied_Programming/Book%3A_Neural_Networks_and_Deep_Learning_(Nielsen)/03%3A_Improving_the_way_neural_networks_learn/3.01%3A_The_cross-entropy_cost_function.pdf)
- Ohlson, J. A. (1980). Ratios financieros y la predicción probabilística de la bancarrota. *Journal of Accounting Research*, 18(1), 109. <https://www.jstor.org/stable/10.2307/2490395?origin=crossref>
- Orellana, I., Tonon, L., Reyes, M., Pinos, L., & Cevallos, E. (2020). Riesgos financieros en el sector manufacturero del Ecuador. *Universidad Del Azuay, November*, 31–36.
- Orellana Osorio, I., Pinos Luzuriaga, L., Tonon Ordóñez, L., Reyes Clavijo, M., & Cevallos Rodríguez, E. (2020). Análisis de cierre empresarial en el sector manufacturero de Ecuador, periodo 1901 - 2018. *Ecos de Economía*, 24(50), 45–79. <https://doi.org/10.17230/ecos.2020.50.3>
- Orlandini, I., Paco, P., & Torricos, P. (2019). Crecimiento económico y la industria hotelera un análisis en dos ciudades patrimoniales del sur de Bolivia. *Revista Digital Investigación y Negocios*, 12(19), 36–45. http://www.scielo.org.bo/scielo.php?script=sci_arttext&pid=S2521-27372019000100005
- Palomeque, M. P. (2008). Administración del capital de trabajo. *Perspectivas*, 21, 161–172. <http://www.redalyc.org/articulo.oa?id=425942157009>
- Park, S. S., & Hancer, M. (2012). A comparative study of logit and artificial neural networks in predicting bankruptcy in the hospitality industry. *Tourism Economics*, 18(2), 311–338. <https://doi.org/10.5367/te.2012.0113>
- Pereira, J. M., Basto, M., & Ferreira das Silva, A. (2017). Comparing logit model with discriminant analysis for predicting bankruptcy in Portuguese hospitality sector. *European Journal of Tourism Research*, 16, 276–280.
- Pinargote, G., & Loor, T. (2021). *El Covid-19 Y Su Impacto Económico En Las Empresas Hoteleras De Manta-Ecuador Covid-19 and Its Economic Impact on the Hotel Companies of Manta-Ecuador*. 12, 152–168. <https://revistas.utm.edu.ec/index.php/ECASinergia>
- Remache, I. (2021). Análisis Discriminante Múltiple para predicción de quiebra, un estudio aplicado a empresas hoteleras de Ecuador. *Universidad Técnica de Ambato*, 1–129. <https://repositorio.uta.edu.ec/bitstream/123456789/27090/1/T4158e.pdf>
- Rojano, A., Salazar, R., Miranda, L., & Ojeda, W. (2021). Algoritmo ADAN en la inteligencia artificial. *Sexto Congreso Nacional de Riego, Drenaje y Biosistemas*, 1–8. <https://www.riego.mx/congresos/comeii2021/files/ponencias/extenso/COMEII-21005.pdf>
- Sanchez Gil, M. (2018). EL APALANCAMIENTO FINANCIERO Y SU INFLUENCIA EN LA RENTABILIDAD DE LAS MYPES DEL SECTOR COMERCIO RUBRO LIBRERIAS DEL DISTRITO DE UCHIZA, 2017. *Universidad Catolica Los Angeles de Chimbote*, 108. http://repositorio.uladech.edu.pe/bitstream/handle/20.500.13032/4116/APALANCAMIENTO_FIN

- ANCIERO_SANCHEZ_GIL_MILKA_ELAMI.pdf?sequence=5&isAllowed=y
- Sifuentes-Amaya, R., & Ramírez-Valverde, G. (2010). Efectos de especificar un modelo incorrecto para regresión logística, con dos variables independientes correlacionadas. *Agrociencia*, 44(2), 197–207.
- Souza, C. D. De, Sanches, A. G., Moisés, G. C., Quispechire, A., Luís, J., Francisco, E., André, J. C., & Rondina, J. M. (2021). *Open Access Machine Learning for Inconsistency Detection in Hospital Cesarean Bills*. 11(2006), 44785–44788.
- Temprano, A. (2015). Análisis del efecto del apalancamiento financiero [Universidad de Oviedo]. In *Repositorio Institucional de la Universidad de Oviedo*. https://digibuo.uniovi.es/dspace/bitstream/handle/10651/32509/TFM_AinhoaVallina.pdf?sequence=6&isAllowed=y
- Universidad Técnica del Norte. (2022). *Cap. 1 El valor Z*. 8. <http://www1.frm.utn.edu.ar/estadistica/documentos/02NotaClase.pdf>
- Vivel-Búa, M., Lado-Sestayo, R., & Otero-González, L. (2016). Impact of location on the probability of default in the Spanish lodging industry: A study of MSMEs. *Tourism Economics*, 22(3), 593–607. <https://doi.org/10.5367/te.2015.0461>
- Vivel-Búa, M., Lado-Sestayo, R., & Otero-González, L. (2018). Risk determinants in the hotel sector: Risk credit in MSMEs. *International Journal of Hospitality Management*, 70(January 2017), 110–119. <https://doi.org/10.1016/j.ijhm.2017.11.004>
- Vivel-Búa, M., Lado-Sestayo, R., & Otero-González, L. (2019). Influence of firm characteristics and the environment on hotel survival across MSMEs segments during the 2007–2015 period. *Tourism Management*, 75(July), 477–490. <https://doi.org/10.1016/j.tourman.2019.06.015>
- Wieprow, J., & Gawlik, A. (2021). The use of discriminant analysis to assess the risk of bankruptcy of enterprises in crisis conditions using the example of the tourism sector in poland. *Risks*, 9(4). <https://doi.org/10.3390/risks9040078>
- Wooldridge, J. (2010). Introduccion a la econometria un nefoque moderno. In *Angewandte Chemie International Edition*, 6(11), 951–952. <https://herioscarlanda.files.wordpress.com/2018/10/wooldridge-2009-introducccic3b3n-a-la-econometrc3ada-un-enfoque-moderno.pdf>
- Youn, H., & Gu, Z. (2010). Predicting Korean lodging firm failures: An artificial neural network model along with a logistic regression model. *International Journal of Hospitality Management*, 29(1), 120–127. <https://doi.org/10.1016/j.ijhm.2009.06.007>
- Zeballos Puma, D. M. (2017). Políticas de Crédito y su incidencia en la liquidez de las empresas comercializadoras de productos químicos en el distrito de Lima-Cercado 2014. *Universidad Cesar Vallejo*, 92.
- Zizi, Y., Jamali-Alaoui, A., El Goumi, B., Oudgou, M., & El Moudden, A. (2021). An optimal model of financial distress prediction: A comparative study between neural networks and logistic regression. *Risks*, 9(11). <https://doi.org/10.3390/risks9110200>