

PREDICCIÓN DE INSOLVENCIA EN EL SECTOR RESTAURACIÓN

Eva Aranda Llamas

Departamento de Finanzas y Contabilidad. Universidad de Málaga.

Ángela Callejón Gil

Departamento de Finanzas y Contabilidad. Universidad de Málaga.

José R. Sánchez Serrano

Departamento de Finanzas y Contabilidad. Universidad de Málaga.

David Alaminos Aguilera

Programa de Doctorado en Ingeniería Mecánica y Eficiencia Energética. Universidad de Málaga.

Resumen

Usando técnicas de regresión logística y de redes neuronales artificiales, este estudio busca mejorar la capacidad de los modelos de predicción de insolvencia existentes para la industria de la restauración. Nuestro objetivo es construir un modelo de predicción de insolvencia de restaurantes que proporcione una alta precisión, utilizando información distante de la situación de insolvencia. Con una muestra de restaurantes españoles correspondiente al periodo 2008-2017, los resultados indican que los mejores predictores de insolvencia son variables financieras relacionadas con la rentabilidad y el endeudamiento, y que las redes neuronales artificiales superan a la regresión logística en capacidad predictiva.

Área temática: Turismo

Palabras clave: predicción de insolvencia, restaurantes, redes neuronales artificiales, regresión logística

1. INTRODUCCIÓN

El objetivo de este estudio es estimar modelos de predicción de insolvencia para las empresas pertenecientes a la industria de la restauración. El hecho de centrarnos en esta industria deriva de dos motivos esenciales: Primero, por la importancia de dicha rama de actividad, tanto en el ámbito de la economía española, en la cual tiene un peso especial el sector servicios, como en la participación de la misma en el ámbito internacional. Y segundo, por el notable incremento de las situaciones concursales de las empresas pertenecientes a este sector, con importante incidencia, incluso, en el primer año de actividad. Estas circunstancias han motivado el interés en analizar las causas que propician la insolvencia en la industria de la restauración e intentar facilitar herramientas o estrategias a los gestores de las mismas, con vistas a evitarla y asegurar la permanencia de sus empresas.

Los trabajos de investigación realizados hasta el momento se han centrado en empresas americanas, analizándose la insolvencia 1 o 2 años antes de la insolvencia (Li et al. 2019). En consecuencia, la inexistencia de trabajos empíricos con muestras de empresas españolas de este sector de actividad hace especialmente interesante que analicemos las causas que la propician, e incluso que comparemos si los modelos existentes para otros ámbitos internacionales podrían utilizarse en el caso español.

Por otra parte, y con objeto de contar con un mayor margen para realizar estrategias que eviten la insolvencia de estas empresas, pretendemos obtener modelos que tengan con objeto predecirla 1, 2 y 3 años antes de que ésta se produzca, comparándose las similitudes y diferencias de dichos modelos a medida que nos alejamos del momento de la insolvencia. Así mismo, y con objeto de obtener unos modelos robustos, se ha considerado un período total de análisis de diez años, comprendidos entre el ejercicio 2008 y el 2017, periodo que abarcaría varios ciclos económicos y, en consecuencia, evita el riesgo de obtener modelos sólo válidos para épocas de crecimiento o en su caso de decrecimiento económico.

2. REVISIÓN DE LA LITERATURA

Si bien son numerosos los estudios sobre predicción de insolvencia (Piñeiro, Del Llano y Rodríguez, 2018), resultan escasos los realizados específicamente para el sector de la hostelería, a pesar del reconocimiento de la alta vulnerabilidad a la quiebra de las empresas pertenecientes a dicho sector. Gu (2002) llevó a cabo el primer estudio empírico centrado exclusivamente en la predicción de insolvencia en el sector de restauración utilizando en su muestra empresas americanas. En los siguientes trabajos se comenzó a utilizar métodos estadísticos más avanzados. Tal ha sido el caso de regresión logística (LOGIT) y de redes neuronales artificiales (NN). Los primeros en utilizar un modelo LOGIT fueron Kim y Gu (2006a), los cuales utilizaron LOGIT aplicado a 32 empresas del sector hostelería (16 en quiebra y otras 16 no quebradas). Asimismo Kim y Gu (2006b) estimaron un modelo LOGIT utilizando la misma muestra de restaurantes utilizada por Gu (2002), con objeto de comparar la capacidad predictiva de un modelo LOGIT respecto al modelo estimado por Gu (2002). Sus resultados demostraron mayor precisión, aunque no con elevada diferencia, del modelo LOGIT. Park y Hancer (2012) realizaron una predicción estadística de la quiebra utilizando una muestra de 80 empresas (40 en quiebra y 40 solventes), que combinaba empresas hoteleras, del sector de la restauración y empresas de servicios de entretenimiento. Compararon la capacidad predictiva de un modelo LOGIT y un NN, concluyendo que las NN predicen mejor dentro de la muestra, pero que la capacidad predictiva de ambos modelos es similar en la muestra externa utilizada para comprobar la precisión de los modelos.

De los estudios empíricos de predicción de insolvencia analizados anteriormente, sólo tres de ellos habían utilizado datos de empresas de restaurantes de forma exclusiva, y otros tres incluyeron empresas de restaurantes de una forma mixta (incluyendo a su vez empresas de otras industrias). El primero de los trabajos realizados en este ámbito fue el de Gu y Gao (2000), quienes estimaron un modelo de análisis multidiscriminante (MDA) de predicción de insolvencia basado en una muestra de 14 empresas en quiebra, incluyendo 4 hoteles y 10 restaurantes, y un número similar de empresas no quebradas. El modelo pudo clasificar las empresas dentro de la muestra con una precisión del 93%. Sin embargo, como el modelo se derivó de una muestra mixta de hoteles y restaurantes y

no fue probado con restaurantes fuera de la muestra, su aplicabilidad a la predicción de quiebra de restaurantes es limitada. Gu (2002) utilizó una muestra de tamaño algo superior, considerando 18 restaurantes en quiebra, emparejados con otros tantos restaurantes no quebrados, con datos del periodo 1986-1998. El modelo estimado obtuvo un nivel de predicción fuera de la muestra del 80%. Posteriormente Kim y Gu (2006b) estimaron un modelo LOGIT utilizando los mismos datos del modelo MDA de Gu (2002). El modelo LOGIT, resultante de una más avanzada sofisticación estadística, predecía correctamente un año antes de la quiebra el 94% de las empresas que habían entrado en bancarrota, así como el 93% de las empresas utilizadas en la muestra externa para verificar la capacidad de predicción. Con estos resultados se demostraba que el modelo LOGIT aportaba una capacidad predictiva algo superior (94% frente al 92% del MDA de Gu, 2002), sin considerar por otro lado que el modelo LOGIT cuenta con una mayor solidez teórica.

Por su parte, Kim y Gu (2006) estimaron dos modelos LOGIT para predecir la quiebra 1 y 2 años antes de empresas pertenecientes al sector hotelero, restaurantes y empresas que prestan servicios de entretenimiento. La capacidad predictiva del modelo ascendió al 91% para 1 año antes de la quiebra y del 84% para 2 años antes. Los autores concluyeron, en base a los modelos estimados, que las empresas de este sector están más próximas a la quiebra si tienen un menor flujo de efectivo y un mayor endeudamiento. Youn y Gu (2010a) fue el estudio empírico posterior que se centró en la industria de restaurantes utilizando información de empresas que cotizaban o habían cotizado en bolsa. Estos autores fueron los pioneros en aplicar NN en la predicción para esta industria. Conjuntamente estimaron un modelo LOGIT con objeto de comparar la capacidad predictiva de ambas técnicas estadísticas, estimando modelos para predecir la quiebra 1 y 2 años antes de la misma. Determinaron que, si bien las NN predicen bien, no aportan mejor predicción que el LOGIT, especialmente en la muestra externa. Finalmente, Park y Hancer (2012) estimaron modelos LOGIT y NN como predictivos de insolvencia de este sector, realizándose estimaciones de ambos tipos de modelos para 1 año antes de la quiebra. En función de los resultados empíricos de dos metodologías, NN obtuvo un mayor nivel de predicción que el modelo LOGIT dentro de la muestra (97.5% frente al 90% del modelo LOGIT). Sin embargo, cuando se realiza el test para su verificación en la muestra externa (formada por 8 empresas solventes y 8 insolventes que no se habían incluido en el modelo inicial), ambos modelos predicen en un 100%.

3. MÉTODOS

En el presente estudio utilizamos dos metodologías diferenciadas para predecir la insolvencia: modelo de regresión logística (LOGIT) y modelo de red neuronal (NN). Si bien los modelos LOGIT han tenido y siguen manteniendo una especial relevancia en los estudios realizados en esta materia en los últimos treinta años, los modelos NN se corresponden con metodologías más avanzadas, que han mostrado tener un importante potencial en el ámbito de la predicción. La principal ventaja de los modelos LOGIT reside, no sólo en la capacidad de predecir previamente si una empresa se espera resulte solvente e insolvente, sino en facilitar información respecto a cuáles son las variables que resultan significativamente explicativas de la insolvencia, y en consecuencia, permiten deducir estrategias adecuadas en la gestión de la empresa con objeto de asegurar la solvencia de la misma. Por su parte, los modelos NN presentan un gran potencial de clasificación, superando en muchos casos a LOGIT, si bien no cuentan con la utilidad explicativa de éste. En el presente trabajo comparamos la utilidad de ambas metodologías en la predicción de la insolvencia en la industria de la restauración en España, contrastándose sus resultados con los obtenidos en los trabajos previos realizados.

3.1. Regresión logística

El modelo LOGIT es un modelo no lineal, a pesar de que contiene una combinación lineal de parámetros y observaciones de las variables explicativas. La función logística se encuentra acotada entre 0 y 1, proporcionando, por tanto, la probabilidad de que un elemento se encuentre en uno de los dos grupos establecidos. A partir de un suceso dicotómico, el modelo LOGIT predice la probabilidad de que el suceso tenga o no lugar. Si la estimación de la probabilidad es superior a 0,5 entonces la predicción es que sí pertenece a ese grupo, y en caso contrario, supondría que pertenece al otro grupo considerado.

Para estimar el modelo se parte del cociente entre la probabilidad de que un suceso ocurra y la probabilidad de que éste no ocurra. La probabilidad de que un suceso ocurra vendrá determinada por la expresión (1).

$$P(Y_i = 1/x_i) = \frac{e^{(\beta_0 + \beta_1 X_1 + \dots + \beta_k X_k)}}{1 + e^{(\beta_0 + \beta_1 X_1 + \dots + \beta_k X_k)}} = \frac{1}{1 + e^{-(\beta_0 + \beta_1 X_1 + \dots + \beta_k X_k)}} \quad (1)$$

donde β_0 es el término constante del modelo y los β_1, \dots, β_k los coeficientes de las variables. Si finalmente se aplican logaritmos en (1), se obtiene una expresión lineal del modelo Y_i^* , en la forma:

$$Y_i^* = \ln \frac{P(Y_i = 1)}{1 - P(Y_i = 1)} = \ln e^{(\beta_0 + \beta_1 X_1 + \dots + \beta_k X_k)} = \beta_0 + \beta_1 X_1 + \dots + \beta_k X_k \quad (2)$$

Los coeficientes del modelo ($\beta_0, \beta_1, \dots, \beta_k$) se estiman aplicando el método de máxima verosimilitud, lo cual conllevaría una serie de pasos: en primer lugar especificar la función de máximo verosimilitud el modelo que recoge la probabilidad conjunta para las observaciones independientes consideradas, en segundo lugar calcular el logaritmo neperiano de la función de verosimilitud, en tercer lugar el cálculo de la derivada del logaritmo neperiano de dicha función respecto a los parámetros que se desean estimar, y en último lugar, obtener las soluciones al sistema de k ecuaciones planteado o estimadores verosímiles (Levy y Varela 2003).

3.2. Redes neuronales artificiales

El Perceptrón Multicapa (MLP) es el modelo de NN utilizado en el presente estudio. Es una red de alimentación hacia delante que estaría compuesta por una capa de unidades de entrada (sensores), otra capa de salida y un número determinado de capas intermedias, denominadas capas ocultas en tanto que no tienen conexiones con el exterior. Cada sensor de entrada estaría conectado con las unidades de la segunda capa, y éstas a su vez con las de la tercera capa, y así sucesivamente. La red tendrá como objetivo establecer una correspondencia entre un conjunto de entrada y un conjunto de salidas deseadas.

Mediante la aplicación del “Teorema de Kolmogorov” (Kolmogorov 1957), Hetch-Nielsen (1987, 1990) demostró que una arquitectura de características similares al MLP y con una única capa oculta resultaba ser un aproximador universal de funciones, no

siendo necesarias arquitecturas de redes más complejas. Además, Nuñez de Castro y Von Zuben (1998) confirmaron que el aprendizaje en MLP constituían un caso especial de aproximación funcional, donde no existe ninguna asunción acerca del modelo subyacente a los datos analizados. El proceso de aprendizaje supone encontrar una función que represente correctamente los patrones de aprendizaje además de llevar a cabo un proceso de generalización que permita tratar de forma eficiente a individuos no analizados durante el aprendizaje (Flórez y Fernández 2008). Para ello se procede al ajuste de pesos W a partir de la información procedente del conjunto muestral, considerando que tanto la arquitectura como las conexiones de la red son conocidas, siendo el objetivo obtener aquellos pesos que minimicen el error de aprendizaje.

Dado, pues, un conjunto de pares de patrones de aprendizaje $\{(x_1, y_1), (x_2, y_2) \dots (x_p, y_p)\}$ y una función de error $\varepsilon(W, X, Y)$, el proceso de entrenamiento implica la búsqueda del conjunto de pesos que minimiza el error de aprendizaje $E(W)$ (Shang y Benjamin 1996), tal y como se expresa en (3).

$$\min_W E(W) = \min_W \sum_{i=1}^p \varepsilon(W, x_i, y_i) \quad (3)$$

La mayor parte de los modelos analíticos utilizados para minimizar la función de error emplean métodos que requieren la evaluación del gradiente local de la función $E(W)$, pudiendo considerarse también técnicas basadas en derivadas de segundo orden (Flórez y Fernández 2008).

4. DATOS Y VARIABLES

La base de datos utilizada en el presente trabajo está formada por 230 empresas tanto solventes como insolventes pertenecientes a la industria de restauración y cuya actividad se desarrolla o se ha desarrollado en territorio español. Dichos datos se han obtenido de la base de datos SABI (Sistema de Análisis de Balances Ibéricos), la cual se define como una base de datos económica-financiera que incluye más de 1.250.000 empresas españolas y más de 400.000 empresas portuguesas. Con objeto de validar los modelos a

estimar y comprobar su capacidad predictiva, se utilizan así mismo muestras de testeo, diferentes y ajenas a las utilizadas en la estimación de los modelos.

Concretamente, el presente estudio ha considerado tres muestras distintas con objeto de analizar la predicción de insolvencia de las empresas de la industria de la restauración en España, para 1, 2 y 3 años antes de la quiebra. En las 3 muestras realizadas se ha considerado el mismo número de empresas solventes que de insolventes, norma general realizada en todos los estudios de predicción de insolvencia, así como la mantenida en el caso concreto de la industria de la restauración. Igualmente, y dado que NN requiere propiamente de otra sub-muestra más, se realizará una partición, utilizando una parte de los datos como muestra de validación. Esta muestra difiere totalmente en concepto y utilidad con respecto a la muestra de testeo, pues la muestra de validación es requerida sólo para la correcta estimación de la NN, evitando su sobre-entrenamiento. La muestra de testeo, por su parte, es la utilizada para constatar la capacidad de generalización de los modelos con datos ajenos a los utilizados propiamente para la obtención de los mismos. La tabla 1 ofrece detalle del número de empresas de la muestra.

Tabla 1. Tamaño de la muestra

Muestra I (MI): 230 empresas solventes y 230 insolventes, siendo los datos contables de las empresas insolventes un año antes de la entrada en quiebra.

Muestra II (MII): 224 empresas solventes y 224 insolventes, siendo los datos contables de las empresas insolventes dos años antes de la entrada en quiebra.

Muestra III (MIII): 192 empresas solventes y 192 insolventes, siendo los datos contables de las empresas insolventes tres años antes de la entrada en quiebra.

En el estudio se ha considerado un total de 24 variables o ratios financieros, obtenidos de la revisión de la literatura de los trabajos de investigación previos sobre la predicción de insolvencia en la industria de la restauración (Li et al. 2019). Todas las variables que se han considerado son de tipo cuantitativo, correspondiéndose con distintos ratios económicos obtenidos a partir de la información contable de las empresas utilizadas en las distintas muestras. Estos ratios se han clasificado, a su vez, en las categorías de eficiencia, de liquidez, de cash-flow y de rentabilidad (Tabla 2). Además de los ratios

financieros objeto de análisis, se ha utilizado una variable dummy, de tipo binomial, que es la variable dependiente para identificar a la empresa como solvente o como insolvente.

Tabla 2. Descripción de las variables independientes

EFICIENCIA

Ingresos/Activo No Corriente	VE1
Ingresos/Total Activo	VE2
Coste de las ventas/Inventario Medio	VE3

LIQUIDEZ Y CASH-FLOW

(Tesorería+Inv. financieras a corto+deudores)/ Pasivo Corriente	VL1
Activos Corrientes/Pasivos Corrientes	VL2
EBIT/Pasivo Corriente	VL3
EBITDA/Pasivo Corriente	VL4
Fondo Maniobra/Total Activo	VL5
(Cash-flow libre + gastos financieros)/Total Pasivo	VC1
(Cash-flow libre + dividendos)/Total Pasivo	VC2
Cash -flow operativo/Total Pasivo	VC3

RENTABILIDAD

Beneficio neto/Ingresos	VR1
Beneficio neto/Patrimonio Neto (ROE)	VR2
Beneficio neto/Total Activo (ROA)	VR3
EBIT/Total Activo	VR4
Margen de beneficio bruto	VR5

SOLVENCIA

Deudas a largo plazo/Total Activo	VS1
EBIT/Gastos financieros	VS2
EBIT/Total Pasivo	VS3
EBITDA/Total Pasivo	VS4
Patrimonio Neto/Pasivo No Corriente	VS5
Total Activo/Patrimonio Neto	VS6
Total Pasivo/Patrimonio neto	VS7
Total Pasivo/Total Activo	VS8

5. RESULTADOS

5.1. Análisis exploratorio

El análisis exploratorio propuesto en este estudio incluye un análisis descriptivo de las variables, y que se presenta de forma diferenciada para las empresas solventes e insolventes de cada una de las tres muestras, con objeto de comparar los parámetros obtenidos, según las empresas sean solventes o insolventes (Tablas 3, 4. y 5). Observando las medias obtenidas para cada una de las variables, diferenciando las empresas solventes de las empresas insolventes, se deduce que en la muestra M.I el 62,5% de las variables seleccionadas tienen una media con signo distinto, según se trate de las empresas solventes o de las empresas insolventes, y que entre ellas estarían todas las variables consideradas de rentabilidad y la mayor parte de las variables clasificadas como de liquidez y cash-flow, así como la mitad de las variables de solvencia. Si bien es necesario resaltar que, como norma general, hay variables que no presentan nunca signo negativo (tal es el caso de la variable VE2 representativa de los Ingresos/Total Activo). Al analizar las medias de la muestra M.II se constata que tan solo el 46% de las variables cuentan con distinto signo según se trata de empresas solventes e insolventes, siendo esta proporción incluso menor en la muestra M.III, en la que tan solo se produce en el 25% de las variables. Estos resultados nos hacen deducir que a medida que nos alejamos del momento de la quiebra, las diferencias entre las variables de las empresas solventes e insolventes se atenúan, y en consecuencia, es previsible que sea más complicado poder realizar una predicción de la solvencia con mayor exactitud.

Tabla 3. Estadísticos descriptivos. M.I.

		MEDIA		MEDIANA		DESV. TÍPICA		MÍNIMO		MÁXIMO	
		Solv	Insolv	Solv	Insolv	Solv	Insolv	Solv	Insolv	Solv	Insolv
EFICIENCIA	VE1	5,64	6,75	3,89	3,78	5,76	12,94	0,31	0,01	41,15	98,81
	VE2	2,06	1,20	1,67	0,87	1,55	1,16	0,15	0,00	6,96	7,65
	VE3	17,70	26,21	5,79	4,24	35,51	112,37	0,00	0,00	289,82	1.192,94
LIQUIDEZ Y CASH-FLOW	VC1	0,16	-0,03	0,06	-0,02	1,26	0,35	-1,55	-1,35	15,29	1,87
	VC2	0,13	-0,05	0,03	-0,03	1,18	0,34	-1,55	-1,35	14,18	1,87
	VC3	0,14	-0,02	0,06	0,02	1,32	0,28	-1,66	-1,12	16,17	0,80
	VL1	2,42	0,67	1,12	0,35	3,73	0,90	0,00	0,00	25,97	6,10
	VL2	2,29	0,75	1,21	0,44	2,99	0,88	0,00	0,01	24,15	5,21
	VL3	0,26	-0,23	0,13	-0,16	0,48	0,49	-1,56	-2,82	3,16	1,66
	VL4	0,52	-0,13	0,32	-0,09	0,76	0,51	-0,08	-2,82	5,95	2,00
	VL5	0,13	-0,48	0,09	-0,29	0,35	0,91	-0,68	-7,36	0,89	0,69
RENTABILIDAD	VR1	0,02	-0,56	0,01	-0,15	0,04	1,78	-0,03	-12,17	0,23	1,02
	VR2	0,13	-0,41	0,07	0,24	0,26	3,86	-1,35	-28,69	1,72	4,01
	VR3	0,03	-0,21	0,02	-0,11	0,04	0,33	-0,07	-2,02	0,23	0,36
	VR4	0,05	-0,20	0,04	-0,10	0,06	0,36	-0,09	-2,69	0,30	0,51
	VR5	0,04	-0,47	0,03	-0,12	0,05	1,55	-0,04	-11,92	0,32	1,65
SOLVENCIA	VS1	0,25	0,46	0,17	0,39	0,26	0,39	0,00	0,00	0,85	2,11
	VS2	14,34	-11,36	2,27	-3,02	51,84	33,21	-5,23	-316,36	554,97	51,14
	VS3	0,15	-0,12	0,07	-0,09	0,30	0,20	-0,10	-0,91	2,71	0,82
	VS4	0,26	-0,07	0,15	-0,05	0,35	0,21	-0,08	-0,91	2,78	0,99
	VS5	2,62	-0,07	0,36	0,00	7,12	3,26	0,00	-19,73	63,83	15,23
	VS6	6,26	2,93	2,88	-0,56	7,95	28,38	1,07	-143,01	45,32	187,76
	VS7	5,26	1,93	1,88	-1,56	7,95	28,38	0,07	-144,01	44,32	186,76
	VS8	0,62	1,29	0,65	1,05	0,27	0,91	0,06	0,30	0,98	8,36

Tabla 4. Estadísticos descriptivos. M.II.

		MEDIA		MEDIANA		DESV. TÍPICA		MÍNIMO		MÁXIMO	
		Solv	Insolv	Solv	Insolv	Solv	Insolv	Solv	Insolv	Solv	Insolv
EFICIENCIA	VE1	5,10	4,95	3,86	3,18	4,31	5,71	0,31	0,00	24,19	33,88
	VE2	2,02	1,21	1,64	0,88	1,53	1,11	0,15	0,00	6,96	6,39
	VE3	17,04	14,10	4,60	4,56	36,32	24,67	0,00	0,00	289,82	140,23
LIQUIDEZ Y CASH-FLOW	VC1	0,17	-0,04	0,06	0,00	1,32	0,34	-1,55	-1,12	15,29	1,66
	VC2	0,14	-0,07	0,03	-0,02	1,23	0,34	-1,55	-1,16	14,18	1,61
	VC3	0,15	0,06	0,06	0,05	1,39	0,27	-1,66	-0,83	16,17	1,62
	VL1	2,07	0,72	1,19	0,47	2,52	0,71	0,00	0,01	14,18	3,43
	VL2	2,02	0,85	1,24	0,58	2,01	0,81	0,00	0,03	10,88	4,20
	VL3	0,22	-0,09	0,13	-0,05	0,31	0,27	-0,34	-1,02	1,84	0,66
	VL4	0,44	0,02	0,31	0,02	0,44	0,29	-0,03	-0,82	2,28	0,81
RENTABILIDAD	VR1	0,02	-0,18	0,01	-0,05	0,04	0,34	-0,02	-1,76	0,23	0,37
	VR2	0,13	0,22	0,07	0,12	0,20	2,20	-0,22	-17,68	1,11	11,22
	VR3	0,03	-0,10	0,02	-0,04	0,04	0,19	-0,06	-0,98	0,18	0,27
	VR4	0,05	-0,08	0,04	-0,03	0,05	0,21	-0,06	-1,32	0,26	0,41
	VR5	0,04	-0,13	0,03	-0,02	0,05	0,34	-0,02	-1,66	0,32	0,83
SOLVENCIA	VS1	0,25	0,44	0,17	0,44	0,26	0,33	0,00	0,00	0,85	1,53
	VS2	10,12	-2,40	2,25	-1,06	26,71	14,45	-5,23	-84,70	250,60	91,70
	VS3	0,13	-0,05	0,06	-0,02	0,22	0,15	-0,08	-0,71	1,84	0,42
	VS4	0,25	0,00	0,15	0,01	0,30	0,15	-0,03	-0,69	2,28	0,45
	VS5	2,83	0,12	0,35	0,00	7,45	5,91	0,00	-57,23	63,83	32,05
	VS6	6,02	8,13	2,83	2,13	7,38	36,00	1,07	-144,24	45,32	176,21
	VS7	5,02	7,13	1,83	1,13	7,38	36,00	0,07	-145,24	44,32	175,21
	VS8	0,62	1,07	0,65	0,98	0,27	0,51	0,06	0,29	0,98	4,97

Tabla 5. Estadísticos descriptivos. M.III.

		MEDIA		MEDIANA		DESV. TÍPICA		MÍNIMO		MÁXIMO	
		Solv	Insolv	Solv	Insolv	Solv	Insolv	Solv	Insolv	Solv	Insolv
EFICIENCIA	VE1	5,58	5,03	3,92	3,56	5,31	4,97	0,58	0,00	35,92	25,15
	VE2	1,86	1,42	1,58	1,14	1,37	1,13	0,15	0,00	6,47	5,43
	VE3	17,64	14,78	5,70	4,76	37,31	23,14	0,00	0,00	289,82	120,84
LIQUIDEZ Y CASH-FLOW	VC1	0,21	0,00	0,07	0,01	1,46	0,58	-1,55	-1,57	15,29	4,74
	VC2	0,18	-0,03	0,06	-0,03	1,36	0,58	-1,55	-1,60	14,18	4,74
	VC3	0,16	0,08	0,05	0,07	1,53	0,31	-1,66	-1,41	16,17	1,23
	VL1	1,71	0,80	1,02	0,63	1,82	0,72	0,00	0,01	9,66	4,11
	VL2	1,76	0,88	1,10	0,66	1,70	0,75	0,00	0,06	7,89	3,97
	VL3	0,20	0,00	0,13	0,01	0,24	0,27	-0,05	-0,95	1,39	0,86
	VL4	0,44	0,13	0,32	0,06	0,41	0,30	-0,03	-0,55	2,23	1,22
	VL5	0,08	-0,20	0,06	-0,12	0,30	0,43	-0,63	-1,53	0,81	0,69
RENTABILIDAD	VR1	0,02	-0,09	0,01	0,00	0,04	0,30	-0,02	-2,29	0,23	0,20
	VR2	0,13	0,16	0,06	0,17	0,18	1,48	-0,03	-9,54	1,04	6,20
	VR3	0,03	-0,05	0,02	-0,01	0,04	0,14	-0,02	-0,54	0,18	0,31
	VR4	0,05	-0,02	0,04	0,00	0,05	0,16	-0,02	-0,50	0,26	0,47
	VR5	0,04	-0,06	0,03	0,00	0,05	0,31	-0,02	-2,09	0,32	1,00
SOLVENCIA	VS1	0,25	0,39	0,18	0,33	0,25	0,30	0,00	0,00	0,85	1,21
	VS2	11,42	1,24	2,30	0,00	29,29	16,14	-5,23	-41,39	250,60	92,06
	VS3	0,12	0,01	0,06	0,00	0,17	0,18	-0,05	-0,43	0,99	0,68
	VS4	0,24	0,07	0,15	0,05	0,25	0,19	-0,03	-0,34	1,11	0,79
	VS5	3,37	0,79	0,60	0,08	8,18	4,50	0,00	-18,70	63,83	30,97
	VS6	5,39	9,24	2,69	3,52	5,50	25,08	1,08	-77,81	30,27	107,85
	VS7	4,39	8,24	1,69	2,52	5,50	25,08	0,08	-78,81	29,27	106,85
	VS8	0,62	0,98	0,63	0,94	0,26	0,37	0,07	0,17	0,97	2,42

Aunque en este análisis descriptivo se ha podido apreciar indicios de qué variables podrán resultar relevantes en el estudio de la insolvencia, no es posible llegar aún a concluir si estas variables son realmente significativas, en tanto este primer análisis resultaría insuficiente para valorar si las diferencias presentadas son precisamente por su significación o por la variabilidad propia que presentan las variables económicas. Esto lleva a requerir la realización de un análisis confirmatorio que permitan valorar la significación de tales variables en el análisis de la insolvencia.

5.2. Análisis confirmatorio

En este apartado se realizará un análisis confirmatorio de los resultados obtenidos en el análisis exploratorio. En este sentido se estimarán dos modelos para cada una de las muestras seleccionadas con objeto de predecir la insolvencia 1, 2 y 3 años de producirse la quiebra. Una vez obtenidos los distintos modelos para cada una de las tres muestras, y a efectos de su comparación, a continuación se presenta un resumen de sus resultados (tablas 6, 7 y 8).

Tabla 6. Resultados de los modelos estimados. M.I

VARIABLES	LOGIT			MLP
	β	Odds ratio	Sig. (Wald)	
VE2	-0,471	0,624	(0,001)	24 variables de entrada
VR3	-13,124	0,000	(0,000)	
VS8	4,121	61,645	(0,000)	
Constante	-2,961	0,052	(0,000)	
MATRIZ DE CLASIFICACIÓN				
Dentro de la muestra		83,80%		85,90%
Testeo		80,00%		89,60%
AJUSTE MODELO				
Prueba RV coefic.		(0,000)		
Hosmer-Lemeshow		0,200		
-2 log verosimilitud		221,650		
R2 Cox-Snell		0,495		
R2 Nagelkerke		0,660		
Curva COR		0,919		0,950

Tabla 7. Resultados de los modelos estimados. M.II

VARIABLES	LOGIT			MLP
	β	Odds ratio	Sig. (Wald)	
VE2	-0,374	0,688	(0,009)	24 variables de entrada
VR3	-17,077	0,000	(0,000)	
VS1	-1,731	0,177	(0,023)	
VS8	4,786	119,801	(0,000)	
Constante	-2,694	0,068	(0,000)	
MATRIZ DE CLASIFICACIÓN				
	Dentro de la muestra	80,80%		87,30%
	Testeo	78,00%		85,40%
AJUSTE MODELO				
	Prueba RV coefic.	(0,000)		
	Hosmer-Lemeshow	(0,066)		
	-2 log verosimilitud	228,943		
	R2 Cox-Snell	0,435		
	R2 Nagelkerke	0,580		
	Curva COR	0,888		0,954

Tabla 8. Resultados de los modelos estimados. M.III

VARIABLES	LOGIT			MLP
	β	Odds ratio	Sig. (Wald)	
VE2	-0,242	0,785	(0,070)	24 variables de entrada
VR1	-6,498	0,002	(0,028)	
VS8	3,836	46,352	(0,000)	
Constante	-2,666	0,070	(0,000)	
MATRIZ DE CLASIFICACIÓN				
	Dentro de la muestra	70,90%		81,30%
	Testeo	74,00%		75,50%
AJUSTE MODELO				
Prueba RV coefic.		(0,000)		
Hosmer-Lemeshow		(0,244)		
-2 log verosimilitud		225,580		
R2 Cox-Snell		0,312		
R2 Nagelkerke		0,416		
Curva COR		0,820		0,774

Como conclusión es posible resaltar la relevancia de las variables VR3, VR1 y VS8 como variables explicativas en los modelos de predicción de insolvencia. Cuanto más alto son los valores de las variables VR3 y VR1 más solvente se espera sea la empresa, mientras que cuanto más alto sea el valor de la variable VS8, más probabilidad existirá de que la empresa sea insolvente. Así mismo otras variables, tales como VE2 y VS1, si bien no han resultado tan relevantes, han resultado necesarias para obtener modelos robustos. Las variables VE2 y VS8 han resultado significativas en todos los modelos LOGIT estimados. Igualmente la variable VR3 también resultó significativa en más de un modelo (M.I y M.II), siendo las variables VS1 y VR1 significativas tan solo en los modelos M.II y M.III respectivamente.

Comparando los niveles de predicción de los modelos elegidos (LOGIT y MLP) se puede comprobar que en todos los casos el nivel de acierto MLP es superior que al de LOGIT, tanto dentro como fuera de la muestra (muestra de testeo). Asimismo, si comparamos el nivel de ajuste medido por el área bajo la curva COR de cada modelo, también se constata

que el ajuste en el caso del modelo MLP superaría al de LOGIT para M.I y M.II, si bien es inferior en el caso de M.III. La capacidad de predicción, a medida que nos alejamos de la quiebra, ha ido disminuyendo, siendo en consecuencia dicha capacidad del modelo obtenido para M.III inferior al del resto de modelos. Este hecho mostraría que la predicción de la insolvencia es más precisa a medida que nos acercamos al momento de la quiebra, en tanto las variables económicas utilizadas mostrarían más diferencias entre las empresas solventes e insolventes.

6. CONCLUSIONES

Los resultados del presente estudio indican, en primer lugar, que la utilización de variables financieras permite estimar modelos con buen ajuste y nivel de clasificación para predecir la insolvencia en el sector de la restauración, inclusive abarcado un periodo suficientemente grande donde han tenido lugar distintos ciclos económicos. En segundo lugar, que la utilización de técnicas estadísticas y computacionales proporcionan modelos robustos de predicción, si bien los construidos con NN arrojan un mayor poder de clasificación.

En tanto que la industria de la restauración ampara tres tipos de actividades relacionadas pero diferenciadas: restaurantes, servicios de catering y establecimientos de bebidas, como propuesta de futura línea de investigación, sería interesante analizar si las conclusiones obtenidas en el presente trabajo para la totalidad de la industria sería similar a los resultados que se obtendrían si analizáramos de forma pormenorizada cada una de estas actividades. Asimismo también se propone como futura línea de investigación analizar este sector económico en los distintos países europeos y comparar los resultados obtenidos entre los distintos países, con objeto de contrastar si las conclusiones obtenidas serían generalizables o no al resto de países europeos. De esta forma también podría verificarse si las conclusiones obtenidas por los estudios previos y que se han centrado en muestras de empresas americanas, podrían o no extrapolarse a estos países, o al igual que han ocurrido en el presente trabajo con la muestra de empresas españolas, se verifica que las conclusiones obtenidas en empresas americanas tampoco serían extrapolables para el resto de países europeos.

REFERENCIAS

- Flórez, R. y Fernández, J.M. (2008). *Las Redes Neuronales Artificiales. Fundamentos teóricos y aplicaciones prácticas*. Ed. Netbiblo. Coruña.
- Gu, Z. (2002). Analyzing bankruptcy in the restaurant industry. A multiple discriminant model. *International Journal of Hospitality Management*, 21(1), 25-42.
- Gu, Z. y L. Gao (2000). A Multivariate Model for Predicting Business Failures of Hospitality Firms. *Tourism and Hospitality Research: A Survey Quaterly Review*, 2(1), 37-49.
- Hecht-Nielsen, R. (1988). Neurocomputing: Pricking the Human Brain. *IEEE Spectrum*, 25, 36-41.
- Hecht-Nielsen, R. (1990). *Neurocomputing*. Addison Wesley. Nueva York.
- Kim, H y Z. Gu (2006a). A logistic regression analysis for predicting bankruptcy in the Hospitality Industry. *The Journal of Hospitality Financial Management*, 14(1), 17-34.
- Kim, H y Z. Gu (2006b). Predicting Restaurant Bankruptcy. A Logit Model in Comparison with a Discriminant Model. *Journal o Hospitality and Tourism Research*, 30(4), 474-493.
- Levy, J. P. y Varela, J. (2003). *Análisis Multivariable para las Ciencias Sociales*. Ed. Prentice Hall. Madrid.
- Li, H., Xu, Y. H., Xiang, R. y Xu, L. H. (2019). Failure analysis of corporations with multiple hospitality businesses. *Tourism Management*, 73, 21-34.
- Núñez de Castro, L. y von Zuben, F. J. (1998). *Optimised Training Techniques for Feedforward Neural Networks*. Technical Report DCA-RT 03/98. Department of Computer Engineering and Industrial Automation. FEE/UNICAMP, Brasil.

Park, S. M. Hancer (2012). A comparative study of logit and artificial neural networks in predicting bankruptcy in the hospitality industry. *Tourism Economics*, 18(2), 311-338.

Piñeiro, C., Del Llano, P. y Rodríguez, M. (2017). Fracaso e insolvencia empresarial: Una reinterpretación en términos de recursos y capacidades. *AECA*, 119, 69-71.

Shang, Y. W. W. Benjamin (1996). Global Optimization for Neural Network Training. *IEEE Computer*, 29(3), 45-54.

Youn, H, y Gu, Z. (2010). Predict US restaurant firm failures: the artificial neural network mode versus logistic regression model. *Tourism and Hospitality Research*, 10(3), 171-187.