

TENDENCIAS EN LA DETECCIÓN DE QUIEBRAS CORPORATIVAS: UN ANÁLISIS ENTRE MODELOS

Elsy Lizbeth Gómez Ramos

Héctor Adrián Guerrero Martínez¹

Resumen

El objetivo de la investigación es analizar 30 estudios relacionados con la detección de quiebras corporativas a través de un *mapa de visualización* bajo el criterio: tipo de modelo. Los resultados indican que los modelos más empleados son las *técnicas estadísticas*, seguidas de las *redes neuronales*, mientras que las *fórmulas teóricas* refieren poca frecuencia. Además, se muestra que la aplicación de modelos *híbridos* es la tendencia más reciente, la cual señala la posibilidad de influir bajo una dinámica evolutiva. Adicionalmente, el desempeño entre los modelos indica que frecuentemente las *redes neuronales* superan a las *técnicas estadísticas*; no obstante, los modelos *híbridos* rebasan a su contraparte sin excepción. La limitante de la investigación es que los estudios analizados incluyen diferentes tamaños de empresas y de economías, por lo que los resultados son generalizados. Finalmente, en las conclusiones se señala que las *redes híbridas* no pueden superar algunos argumentos “en contra” (falta de interpretación de los parámetros), lo que explica –al menos en parte– la alta frecuencia de empleo de las *técnicas estadísticas*.

Palabras clave: Quiebras corporativas; Análisis discriminante múltiple, Black-Scholes, Redes neuronales artificiales

Códigos JEL: G33; G17

1 Profesores de la Universidad Autónoma Metropolitana, Unidad Xochimilco y Unidad Iztapalapa, respectivamente.

Abstract

The objective of the research is to analyze 30 researches related with the detection of corporate bankruptcies through a *visualization map* under the criterion: type of model. The results indicate that the most used models are the *statistical techniques* followed by *neural networks*, while the *theoretical formulas* showed a little frequency in the field. On the other hand, it is show that the *hybrid* models are the most recent trend, which show the possibility of permeating under an evolutionary dynamic. Additionally, the performance among the models indicates that *neural networks* often outperform *statistical techniques*, nevertheless the *hybrid* models surpass their counterpart without exception. The limitation is that the studies analyzed include different sizes of firms and of economies, so the results are generalized. Finally, it is concluded that the *hybrid* networks can't overcome some "deficiencies" (lack of interpretation of parameters), which explains –at least in part- the high frequency of using the statistical techniques.

Key words: Corporative bankruptcy, Multiple discriminant analysis, Black-Scholes, Artificial neural networks

JEL classification: G33; G17

1. Introducción

La función de una auditoría es realizar una revisión exhaustiva y ética de los ejercicios contables anteriores o del ejercicio anterior inmediato de una empresa, la cual tiene la facultad y los medios –a través de las notas de auditoría– para señalar irregularidades. Pero si estas irregularidades son imperceptibles o externas, entonces una revisión contable tendrá un sesgo considerable.

Una forma de conocer de manera *aproximada* la condición futura de una empresa es a través del manejo de herramientas capaces de descifrar información relevante. En general, la literatura especializada señala tres tipos de modelos para detectar quiebras corporativas: 1) modelos que se ajustan a funciones lineales o cuadráticas como el análisis discriminante múltiple (ADM), el cual distingue qué variables definen mejor

a las empresas; 2) fórmulas teóricas como el desarrollo Black-Scholes (B-S), que considera a una acción como subyacente de la opción bajo condiciones estables y parámetros conocidos; y 3) modelos no paramétricos como las redes neuronales artificiales (RNA), mismas que se caracterizan por su alta capacidad de adaptación.

Las RNA y, en especial, el perceptrón multicapa (PML) han propiciado otra forma de resolver este tipo de problemas a través de la *clasificación de patrones* (Mehrotra *et al.*, 2000), ya que por su naturaleza no requieren hacer supuestos *a priori* sobre el tipo de información, por lo que su aplicación es ampliamente reconocida en el campo (Pérez, 2006; Zhang *et al.*, 1999), además de reivindicar el uso de información contable como la principal fuente de información.

El objetivo de la investigación es analizar 30 estudios sobre la detección de quiebras corporativas publicados entre el 2000 y el 2015. Para ello se recurre al manejo de un *mapa de visualización* con la idea de ordenar la información bajo el criterio: tipo de modelo empleado. La principal razón de emplear estos *mapas* es porque el número de modelos analizados es amplio y diverso (lineal/no lineal, paramétricos/no paramétricos, además de sus variantes, extensiones y fusiones), por lo que brinda la oportunidad de presentar la información de forma accesible. Adicionalmente, ofrece las siguientes ventajas: 1) permite identificar la mayoría de las opciones, 2) ofrece información sobre la frecuencia de empleo por tipo de recurso, y 3) ofrece indicios sobre las posibles tendencias. Este criterio de análisis junto con el desempeño de los modelos permitirá tener una visión más amplia y objetiva para entender las tendencias actuales en el campo. La limitante de la investigación es que los estudios analizados incluyen diferentes tamaños de empresas y de economías, que abarcan toda una gama de actividades sin excluir aquellos estudios que analizan entidades financieras² (bancos), por lo que los

2 Si bien es cierto que los estudios que analizan empresas productivas excluyen a las entidades financieras, esto no implica que estas no se analicen bajo los modelos aquí expuestos. De hecho, esta exclusión se da por mantener cierta homogeneidad en los criterios de selección e interpretación de los datos, y no porque cambien las motivaciones y consecuencias de una posible quiebra.

resultados son generalizados. El distintivo de la investigación es que, a diferencia de estudios previos (Scott, 1981; Pérez, 2006), se introducen a las RNA junto con las opciones más representativas en el campo.

Los resultados indican que en términos de frecuencia las *técnicas estadísticas* (ADM, modelos logit/probit y otras) son las más empleadas, pero que en términos de desempeño las *redes híbridas* (redes evolutivas, redes difusas y otras) son las más precisas en el campo.

La investigación está organizada en siete secciones. En la segunda se exponen aspectos generales sobre el funcionamiento del PML; en la tercera se exponen los modelos más representativos en el campo y se profundiza sobre sus diferencias; en la cuarta se emplea el *mapa de visualización* para el análisis de las tendencias en la literatura; en la quinta se presenta uno de los modelos *híbridos* más representativos; en la sexta se presenta una discusión, y por último las conclusiones.

2. El Perceptrón multicapa

Aunque pocos estudios se han centrado en comparar el desempeño entre un vasto número de redes, la mayoría concuerda que la *simplicidad* de la red parece ser una constante para un buen desempeño. Lo anterior permite constatar la vigencia de una red como el PML. La sección está dividida en tres partes. En la primera se expone el funcionamiento básico para este tipo de redes, en la segunda se exponen algunas recomendaciones para su diseño, y finalmente se muestra un ejemplo.

Funcionamiento

El nodo (o unidad de procesamiento) es la unidad básica de una red neuronal. El PML incluye una capa de entrada, otra de salida y al menos una capa oculta. Las capas están formadas por un conjunto de nodos. Cada capa está totalmente interconectada, pero sin establecer conexiones dentro de la misma capa. Las capas de entrada y de salida indican el flujo de información durante la etapa de entrenamiento. En Mehrotra *et al.*, (2000) se describe el algoritmo de retropropagación para una red con conexión hacia adelante (como el PML) que tenga al menos

una capa oculta, es decir, $L = 2$ (la capa de entrada no se contabiliza debido a que no realiza ningún procesamiento). A continuación, se detalla el proceso de aprendizaje utilizando P patrones de entrenamiento, donde la colección de entradas y de salidas constituyen el conjunto de entrenamiento (x_p, d_p) :

1. De la capa de entrada el i -ésimo nodo tiene un valor $x_{p,i}$ para el patrón p .
2. La entrada neta para el j -ésimo nodo de la capa oculta $= net_j^{(1)} = \sum_{i=0}^n w_{j,i}^{(1,0)} x_{p,i}$. El cual incluye el umbral $x_{p,0} = 1$; la conexión del i -ésimo nodo de entrada al j -ésimo nodo de la capa oculta se le asigna un peso de $w_{j,i}^{(1,0)}$.
3. La salida del j -ésimo nodo de la capa oculta $= x_{p,j}^{(1)} = \delta\left(\sum_{i=0}^n w_{j,i}^{(1,0)} x_{p,i}\right)$ donde δ es la función sigmoïdal.
4. La entrada neta del k -ésimo nodo de la capa de salida $= net_k^{(2)} = \sum_j \left(w_{k,j}^{(2,1)} x_{p,j}^{(1)}\right)$, incluye el umbral; la conexión del j -ésimo nodo de la capa oculta al k -ésimo nodo de la capa oculta se le asigna un peso $w_{k,j}^{(2,1)}$.
5. La salida calculada del k -ésimo nodo de la capa salida $o_{p,k} = \delta\left(\sum_j w_{k,j}^{(2,1)} x_{p,j}^{(1)}\right)$.
6. La salida deseada del k -ésimo nodo de la capa de salida es $d_{p,k}$, y su correspondiente error al cuadrado es $l_{p,k}^2 = |d_{p,k} - o_{p,k}|^2$.

Generalizando, con $L > 2$ el peso $w_{k,j}^{(i+1,i)}$ denota el peso asignado a la conexión del nodo j en la i -ésima capa para el nodo k en la $(i+1)$ -ésima capa, y $x_{p,j}^{(i)}$ señala la salida del j -ésimo nodo en la i -ésima capa para el p -ésimo patrón. El error (del patrón p) está dado por $E_p = \sum_k (l_{p,k})^2$, por lo que necesitamos encontrar w , el vector de todos los pesos de la red que minimicen el valor de E_p . Omitiendo el sufijo p por conveniencia la

expresión que se debe minimizar $E = \sum_{k=1}^K (l_k)^2$. Una forma de minimizar E está basado en el método del gradiente descendiente, ya que o_k y E depende de los pesos de la red, de acuerdo al gradiente descendiente, la dirección del cambio del peso de w debiera ser la misma $-\frac{\partial E}{\partial w}$.

Para simplificar el cálculo de $-\frac{\partial E}{\partial w}$ se examina el cambio de un sólo peso, se calcula el valor de para $\frac{\partial E}{\partial w_{k,j}^{(2,1)}}$ cada conexión de la capa oculta a la capa de salida. De forma similar se calcula el valor de $\frac{\partial E}{\partial w_{j,i}^{(1,0)}}$ para cada conexión de la capa de entrada a la capa oculta y los pesos de conexión están, por lo tanto, cambiando; a este método se le conoce como la *regla delta generalizada*. En forma reducida, las ecuaciones 1 y 2 describen el cambio en los pesos.

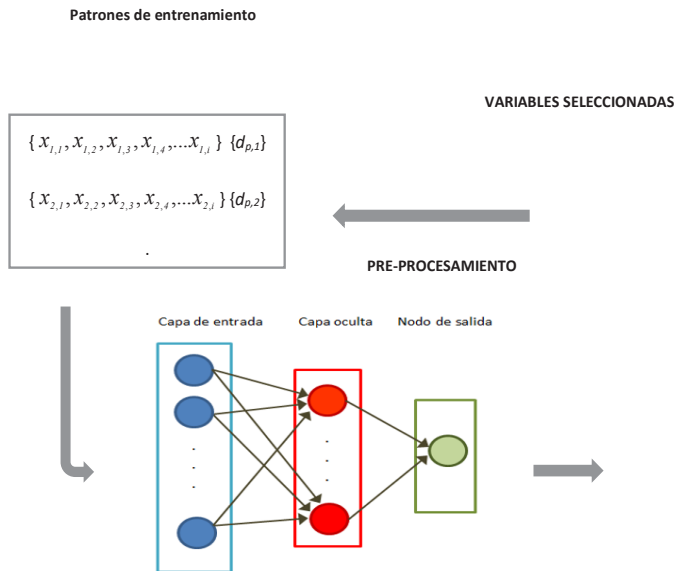
$$\Delta w_{k,j}^{(2,1)} \alpha \left(-\frac{\partial E}{\partial w_{k,j}^{(2,1)}} \right) \quad (1)$$

$$\Delta w_{j,i}^{(1,0)} \alpha \left(-\frac{\partial E}{\partial w_{j,i}^{(1,0)}} \right) \quad (2)$$

Los cálculos para la derivada de E con respecto a los pesos asociados de la capa oculta con la capa de salida, así como las conexiones de la capa de entrada con la capa oculta se realizan con la regla de la cadena, que es aplicada al algoritmo de aprendizaje.

En esencia, el algoritmo de retropropagación es una generalización de mínimos cuadrados que modifica los pesos minimizando el error entre la salida deseada y la actual. Una vez entrenada la red los pesos se fijan y pueden ser utilizados para procesar las salidas ante nuevas entradas (etapa de prueba). En la gráfica 1 se muestra el funcionamiento del PML a través del flujo de información que recibe la capa de entrada, y que se procesa en la capa oculta y en el nodo de salida.

Gráfica 1. Esquema general del PML para la clasificación de patrones



El tamaño de la matriz depende del número de variables seleccionadas y de la muestra de datos.

La selección del vector fila (patrón de entrenamiento) es de forma aleatoria. Cada vez que se selecciona un patrón y se procesa se le llama: *iteración*.

Fuente: Elaboración propia.

En otras palabras, el funcionamiento de la red involucra técnicas estadísticas bien conocidas; sin embargo, las RNA tienen la capacidad de generar o destruir unidades de procesamiento, lo que las hace merecedoras del término: *artificiales*.

2.2 Diseño

El aspecto y funcionalidad de una red depende de su diseño. Existen diversas recomendaciones que ayudan a diseñar al PML cuando se desea clasificar patrones (Mehrotra *et al.*, 2000; Pérez, 2006; Zhang *et*

al., 1999). Aquí se muestran estas recomendaciones agrupadas en dos categorías: arquitectura y parámetros.

Arquitectura

La arquitectura³ de la red incluye el número de nodos y capas (topología), el tipo de función de transferencia (procesamiento) y el tipo de conectividad (hacia adelante). Para el caso del PML, únicamente la topología genera serios problemas, ya que esta depende de la complejidad de cada problema. No obstante, se señalan las siguientes recomendaciones: 1) iniciar con una red pequeña (método constructivo), 2) preferir una forma piramidal (donde la base sea la capa de entrada), y 3) añadir como máximo dos capas ocultas.

Parámetros

La diferencia crucial entre una red tradicional y las RNA es que en estas últimas sus conexiones son *entrenables*, es decir, los valores asignados a las conexiones entre los nodos no son fijas. La definición de “parámetros” en el ámbito de las RNA refiere precisamente a la magnitud de estas conexiones. También, se suelen considerar a las reglas de aprendizaje (término momento y tasa de aprendizaje) y al tamaño del entrenamiento (número de iteraciones) como parámetros, debido a que su magnitud intensifica o deteriora la capacidad de respuesta de la red. Las recomendaciones son: 1) especificar la inicialización de los pesos con rangos moderados, así como de las reglas de aprendizaje, y 2) contar con algún criterio estadístico para definir el tamaño del entrenamiento.

2.3 Un ejemplo

Con la idea de mostrar cómo se construyen los patrones de entrenamiento y el diseño de la red para detectar una quiebra corporativa, se expone el siguiente ejemplo. Suponga que se cuenta con la información disponible de una empresa antes y después de una quiebra. Imagine también que se desea conocer, si el PML es capaz de identificar la si-

3 También es común referirse a la topología como la arquitectura.

tuación futura de la empresa justo antes de la quiebra. De acuerdo con los estados financieros de los últimos cuatro años de la “*Empresa X*” (cuadro 1 y 2), se analiza su desempeño a través de las razones financieras. De acuerdo con las Normas de Información Financiera (NIF), estas se pueden clasificar en: solvencia, liquidez, eficiencia operativa y rentabilidad. Las primeras definen el exceso de activos sobre pasivos y la suficiencia del capital contable, las segundas refieren a la disponibilidad de fondos suficientes para satisfacer los compromisos financieros a su vencimiento, las terceras definen el grado de actividad con que se mantienen niveles de operación adecuados, y las últimas refieren a la capacidad para generar utilidades o incremento en los activos netos.

Las razones financieras seleccionadas con base en Zhang (*et. al.*, 1999) y en las NIF son las siguientes: cobertura de deuda (CD), prueba de liquidez (PL), margen de seguridad (MS), rotación de activos totales (RAT), margen de utilidad bruta (MUB) y retorno de activos (RdA). En el cuadro 3 se muestran los resultados de las razones financieras para la “*Empresa X*”. A continuación, se detallan seis pasos para la etapa de entrenamiento, y uno para la etapa de prueba⁴.

4 Este hecho no afecta, ya que los datos para la etapa de prueba deberán seguir el mismo tratamiento (paso 1) y omitir del paso 2 al paso 6.

Cuadro 1. Balance general

	2011				2010				2009				2008			
	1Q	2Q	3Q	4Q	1Q	2Q	3Q	4Q	1Q	2Q	3Q	4Q	1Q	2Q	3Q	4Q
Activo total	16 626	16 138	16 289	15 933	15 822	16 534	16 145	16 229	12 090	11 782	11 927	15 758	11 021	11 152	11 162	11 863
Activo circulante	2 384	2 412	2 260	1 952	2 164	2 584	2 307	2 319	1 776	1 805	1 891	2 004	1 624	1 623	1 627	1 708
Activo fijo	8 943	8 710	8 817	8 963	8 749	9 094	9 128	8 940	6 936	6 647	6 587	9 034	6 129	6 536	6 523	6 921
Otros activos	5 299	5 016	5 212	5 018	4 909	4 856	4 710	4 970	3 378	3 330	3 449	4 720	3 268	2 993	3 012	3 234
Pasivo total	8 250	8 611	9 148	8 982	8 166	8 602	8 478	8 076	5 768	5 690	6 036	8 110	5 587	5 595	5 477	5 428
Pasivo circulante	2 328	2 472	3 558	2 954	3 040	3 380	3 303	2 346	1 785	1 792	2 338	3 920	1 879	1 847	2 214	1 826
Pasivo largo plazo	4 394	4 685	4 031	4 373	3 443	3 493	3 568	4 344	3 320	3 116	2 871	2 709	3 544	3 435	2 953	3 340
Otros pasivos	1 528	1 454	1 559	1 655	1 683	1 729	1 607	1 386	663	782	827	1 481	164	313	310	262
Capital contable	8 376	7 527	7 141	6 951	7 656	7 932	7 667	8 153	6 322	6 092	5 891	7 648	5 434	5 557	5 685	6 435

Fuente: Elaboración propia.

Cuadro 2. Estado de resultados

	2011				2010				2009				2008			
	1Q	2Q	3Q	4Q	1Q	2Q	3Q	4Q	1Q	2Q	3Q	4Q	1Q	2Q	3Q	4Q
Ventas netas	1 571	1 741	1 720	1 620	1 581	1 815	1 757	1 676	1 325	1 360	1 393	1 510	1 119	1 198	1 239	1 236
Costo de ventas	-865	-946	-959	-945	-887	-982	-952	-897	-689	-685	-700	-813	-638	-646	-679	-702
Utilidad bruta	706	795	761	675	694	833	805	779	636	675	693	697	481	552	560	534
Utilidad de oper.	320	403	353	256	-304	-373	409	371	-238	-259	-258	-300	316	384	-171	-194
Utilidad neta	299	71	33	168	329	466	122	395	274	236	273	287	317	260	261	182
Flujo operación	473	555	505	417	532	611	561	519	486	502	516	513	398	464	476	440

Fuente: Elaboración propia.

Cuadro 3. Resultados

Fecha	CD	PL	MS	RAT	MUB	RdA
2011-1	0.070365	1.024054	0.024054	0.094490	0.449395	0.017983
2	0.077546	0.975728	-0.024278	0.107882	0.456634	0.004399
3	0.066543	0.635188	-0.364811	0.105592	0.442441	0.002025
4	0.056912	0.659682	-0.339201	0.101675	0.416666	0.010544
2010-1	0.082060	0.711842	-0.288157	0.099924	0.438962	0.020793
2	0.088898	0.764449	-0.235502	0.109773	0.458953	0.028184
3	0.081647	0.698455	-0.301544	0.108826	0.458167	0.007556
4	0.077578	0.988491	-0.011508	0.103271	0.464797	0.024339
2009-1	0.095200	0.994957	-0.005042	0.109594	0.480000	0.022663
2	0.102281	1.007250	0.007254	0.115430	0.496323	0.020030
3	0.099059	0.808810	-0.018819	0.116793	0.497487	0.022889
4	0.077380	0.511224	-0.488775	0.095824	0.461589	0.018212
2008-1	0.073391	0.864289	-0.135710	0.101583	0.429848	0.028763
2	0.087845	0.878722	-0.121277	0.107424	0.460767	0.023314
3	0.092123	0.734869	-0.265130	0.111000	0.453074	0.023382
4	0.085172	0.935377	-0.064622	0.104189	0.432038	0.015341

Cobertura de deuda $CD = FO/DT$, donde FO es el flujo operativo y DT es la deuda total.

Prueba de liquidez $PL = AC/PC$, donde AC es el activo circulante y PC es el pasivo circulante. Margen de seguridad $MS = CT/PC$, donde CT es el capital de trabajo. Rotación de activos totales $RAT = VN/AT$, donde VN son las ventas netas y AT son los activos totales. Margen de utilidad bruta $MUB = UB/VN$, donde UB es la utilidad bruta. Retorno de activos $RdA = UN/AT$, donde UN es la utilidad neta.

Fuente: Elaboración propia con base en los cuadros 1 y 2.

PASO 1. Pre-procesamiento de los datos

El pre-procesamiento se hace necesario debido a que la función sigmoide requiere de un rango por debajo de sus puntos de saturación [0-1] para el correcto procesamiento de los datos. Una forma podría ser con la siguiente fórmula:

$$V' = (V - \min) * (\text{newmax} - \text{newmin}) / (\max - \min) + \text{newmin} \quad (3)$$

Aplicando la fórmula (3) al primer dato del extremo izquierdo del cuadro 3, se tiene que: V es el dato a normalizar (0.070365), \min es el valor mínimo del grupo CD (0.056912), newmin es el nuevo mínimo que se desee (0.1), \max es el valor máximo del grupo CD (0.10228), newmax es el nuevo máximo que se desee (0.2) y V' es el dato normalizado (0.1). Una vez normalizados los datos es necesario identificar patrones, o bien establecer algún criterio. En el cuadro 4 se muestran los datos normalizados, y se detecta un punto de inflexión⁵, es decir, durante los periodos 2011-3 y 2011-4 todos los datos entran en los rangos inferiores. De este modo, se etiqueta a este patrón con (0.1) peligro; en caso contrario, se etiqueta con (0.7) bajo riesgo, y para los casos en donde existe una combinación se etiqueta con (0.5) riesgo medio y (0.3) riesgo alto.

PASO 2. *Arquitectura*

Para el caso de la capa de entrada, el número de nodos es igual al número de razones financieras (seis) seleccionadas: CD, PL, MS, RAT, MUB y RdA. En cuanto al número de nodos de la capa oculta, se proponen tres (aunque también se pueden proponer 5, 4, o 2 nodos; en este caso, lo mejor es realizar los cálculos para cada propuesta) y uno de la capa de salida.

PASO 3. *Selección de parámetros*

Generalmente, la tasa de aprendizaje oscilar entre 0.1 (cambio mínimo en los pesos) y 1 (cambio importante en los pesos), y el término momento entre 0 (el pasado no tiene efecto sobre el cambio en los pesos) y 1 (el pasado tiene un gran efecto sobre el cambio en los pesos). Los valores propuestos para estos parámetros son: 0.5 y 0.4, respectivamente. En cuanto a la inicialización de los pesos se selecciona un rango entre [-0.1 y 0.1].

5 Este criterio sólo debe entenderse como un ejercicio.

Cuadro 4. Datos normalizados

Fecha	CD	PL	MS	RAT	MUB	RdA	Clase	Etiqueta
2011-1	{0.1}	0.2	0.2	0.1	0.1	0.2}	{0.5}	MEDIO
2	{0.1}	0.2	0.2	0.2	0.1	0.1}	{0.5}	
3	{0.1}	0.1	0.1	0.1	0.1	0.1}	{0.1}	PELIGRO
4	{0.1}	0.1	0.1	0.1	0.1	0.1}	{0.1}	ALTO
2010-1	{0.2}	0.1	0.1	0.1	0.1	0.2}	{0.3}	MEDIO
2	{0.2}	0.1	0.1	0.2	0.2	0.2}	{0.5}	
3	{0.2}	0.1	0.1	0.2	0.2	0.1}	{0.5}	
4	{0.1}	0.2	0.2	0.1	0.2	0.2}	{0.5}	BAJO
2009-1	{0.2}	0.2	0.2	0.2	0.2	0.2}	{0.7}	MEDIO
2	{0.2}	0.2	0.2	0.2	0.2	0.2}	{0.7}	
3	{0.2}	0.2	0.2	0.2	0.2	0.2}	{0.7}	BAJO
4	{0.1}	0.1	0.1	0.2	0.2	0.2}	{0.5}	
2008-1	{0.1}	0.2	0.2	0.1	0.1	0.2}	{0.5}	MEDIO
2	{0.2}	0.2	0.2	0.2	0.2	0.2}	{0.7}	
3	{0.2}	0.1	0.1	0.2	0.1	0.2}	{0.5}	
4	{0.2}	0.2	0.2	0.1	0.1	0.1}	{0.5}	

Fuente: Elaboración propia.

PASO 4. *Tamaño del entrenamiento*

El tamaño del entrenamiento se puede basar en el nivel de confianza deseado, entonces:

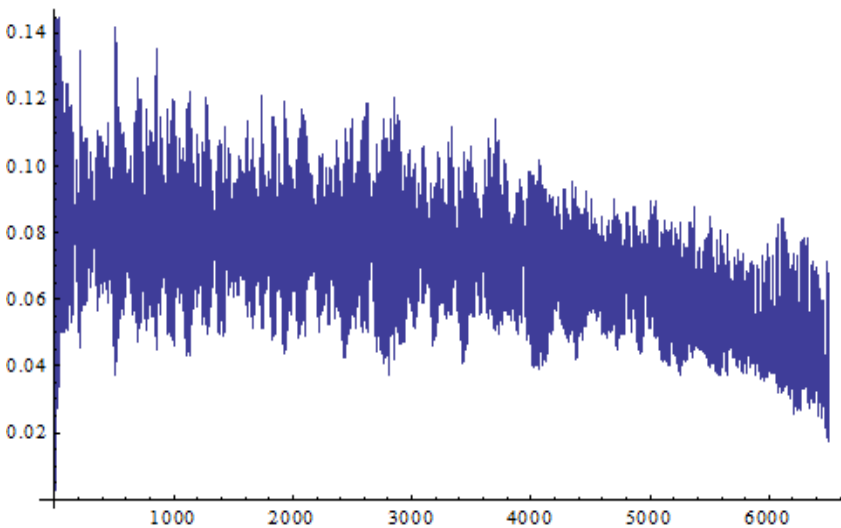
$$P > \frac{|W|}{(1-a)} \log \frac{n}{1-a} \quad (4)$$

Donde P es el número de patrones, $|W|$ es el número de pesos a entrenar ($6 \times 3 = 18 + 3 = 21$), a es el nivel de confianza esperado y n es el número de nodos ($6 + 3 + 1 = 10$). Si se espera que el nivel de confianza sea del 98%, entonces el tamaño del entrenamiento está definido por:

$$P > \frac{|21|}{(1-.98)} \log \frac{10}{1-.98}$$

El resultado indica que el entrenamiento debe ser mayor a 2,833 *iteraciones* o patrones presentados. En la gráfica 2 se muestra el error cuadrático medio (ECM) durante la etapa de entrenamiento⁶. En consecuencia, si estos errores tienden a disminuir, entonces la etapa de entrenamiento concluye de forma satisfactoria y, por tanto, la red ha podido detectar características propias de cada clase.

Gráfica 2. Errores durante la fase de entrenamiento



Fuente: Elaboración propia con base en la programación propuesta en Freeman (1993) con *Mathematica*.

PASO 6. *Pesos óptimos*

Como resultado del paso anterior se obtienen los pesos “óptimos” (21 pesos), es decir, el valor de cada una de las conexiones de la red después de su entrenamiento. Estos valores se mantendrán fijos para la siguiente etapa.

6 Véase, apéndice 1, cuadro A.1, para detalles sobre la programación empleada.

PASO 7. Resultados

Al insertar los vectores de prueba, la red genera sus salidas. Estas salidas se cotejan con las clases de la fase de entrenamiento para decidir a cuál clase es más próxima, pero si se cuentan con los valores de las clases (salidas deseadas), entonces se podrá calcular el error cometido en esta etapa.

3. Modelos

En esta sección se exponen brevemente los tres modelos pioneros por su naturaleza empleados en la detección de quiebras corporativas. Adicionalmente, se realiza un comparativo entre ellos que permite identificar mejor su naturaleza.

3.1 Análisis Discriminante Múltiple

El estudio de Altman (1968) unificó la evidencia sobre la capacidad predictiva proveniente de los estados financieros a través de las razones financieras con una técnica estadística rigurosa capaz de modelar diversas variables durante diferentes periodos para todo un conjunto de empresas. Al introducir esta información el modelo genera una *combinación lineal* entre las características que mejor discriminan a los grupos (quiebra-no quiebra). Los resultados señalan que, si se considera una muestra amplia previa a la quiebra, el modelo logra identificar con precisión a las empresas en peligro, pero si se considera una muestra reducida el modelo se deteriora.

3.2 Black - Scholes

La derivación de la fórmula para valorar una opción (europea) teniendo como subyacente a una acción fue propuesta por Black y Scholes (1973). El valor de mercado de las obligaciones corporativas (acciones comunes y bonos) es visto como una opción, que será valuada al tiempo que su deuda expire. Sin embargo, los supuestos de la formulación son fuertes, en especial cuando se tratan de omitir los pagos periódicos de

las obligaciones o al mitigar el impacto de intereses e impuestos ante una posible restructuración del capital.

3.3 Perceptrón multicapa

El estudio de Odom y Sharda (1990) respalda la importancia de considerar razones financieras como fuente de información confiable sobre la salud de la empresa. Por ello compara la capacidad predictiva del ADM y del PML utilizando las mismas razones financieras seleccionadas por el estudio de Altman (1968). Los resultados indican que la red tuvo el mejor desempeño tanto en la muestra original como en la ajustada.

En el cuadro 5 se resumen las diferencias más sobresalientes sobre los modelos expuestos. En general, se detectaron tres diferencias: 1) *supuestos*, 2) *interpretación de los parámetros*, y 3) *fuentes de información*. En este sentido, los supuestos en el ADM se derivan de la rigidez en su planteamiento, mientras que para el B-S son las que justifican su formulación estilizada. Para el caso de las RNA, el análisis es diferente, ya que como “directamente” no se requiere información *a priori* sobre la naturaleza de la información, entonces la discusión no está en definir el problema, sino en el diseño de la red.

Lo anterior, no significa que no existan *supuestos*, sólo que no es posible definirlos de manera explícita. Para el caso de la *interpretación de los parámetros*, nuevamente las RNA difieren considerablemente con los otros dos modelos. La razón principal es que las RNA son modelos *no paramétricos* (desde el punto de vista estadístico) que tienen como finalidad minimizar el error en cada iteración, por lo que su diseño no obedece a interpretaciones sino a resultados. Sin embargo, esto no les permite ofrecer un conocimiento claro sobre cómo solucionan el problema. Mientras que los *parámetros* en el modelo ADM y la formulación B-S ofrecen una interpretación de sus resultados y de sus implicaciones (por lo menos de manera aproximada). Finalmente, las *fuentes de información* pueden ser contables, macroeconómicas, financieras u otras. La integración de las RNA ha permitido considerar nuevamente el uso de este tipo de información sin excluir la posibilidad de integrar información bursátil o macroeconómica.

Cuadro 5. Comparativo entre modelos pioneros

Modelo	Altman (1968) ADM	Black-Scholes (1973) B-S	Odom y Sharda (1990) PML
Naturaleza	Estadístico	Físico - Matemático	Biológico
Tipo	Empírico	Teórico	Empírico
Modalidad	Clasificador de patrones	Valuación de opciones	Clasificador de patrones
Formulación ^{1/}	$z = v_1x_1 + v_2v_2 + \dots + v_nx_n$	$w(x,t) = xN(d_1) - ce^{r(t-t^*)}N(d_2)$	
Especificaciones		-El valor de mercado del capital de una empresa es visto como una opción (europea)	Selección apropiada de la arquitectura, <i>parámetros</i> y codificación
Supuestos	-Existe una relación <i>lineal</i> entre las variables -Las variables seleccionadas son apropiadas	-Condiciones ideales para el mercado y la opción -Parámetros conocidos -El subyacente es una acción	n.a
Parámetros	Coefficientes discriminantes	Tasa de interés y varianza	Pesos
Información requerida	Razones financieras y otros	Precio de la acción y de ejercicio, fecha de inicio y de vencimiento, y parámetros	Razones financieras, variables macroeconómicas, etc.
Ventajas	-Considera diversas características -El modelo es parsimonioso -Existen técnicas auxiliares para mejorar el desempeño	-Existe un fundamento teórico -La formulación no depende de variables contables	-No se requieren supuestos <i>a priori</i> sobre la naturaleza de la información -Se puede utilizar información escasa o faltante -Es posible modelar relaciones <i>no lineales</i>

Criticas	Utiliza información contable Existe sobre ajuste en el modelo Existe rigidez en el planteamiento	Existe un número importante de <i>supuestos</i> La formulación está acotada Solo se pueden analizar a las empresas que cotizan en bolsa	-No existe ninguna interpretación acerca de los <i>parámetros</i> -No existen reglas para seleccionar la topología -No es consistente la manera de seleccionar a las variables -Utiliza información contable
-----------------	--	---	---

Nota: ^{1/}Dónde: v_1, v_2, \dots, v_v son los coeficientes discriminantes, x_1, x_2, \dots, x_v son las variables independientes, $w(x, t)$ es el precio de la opción en función de la acción y del tiempo, c es el precio de ejercicio, $N(d)$ es la función de densidad normal acumulada y r es la tasa de interés.

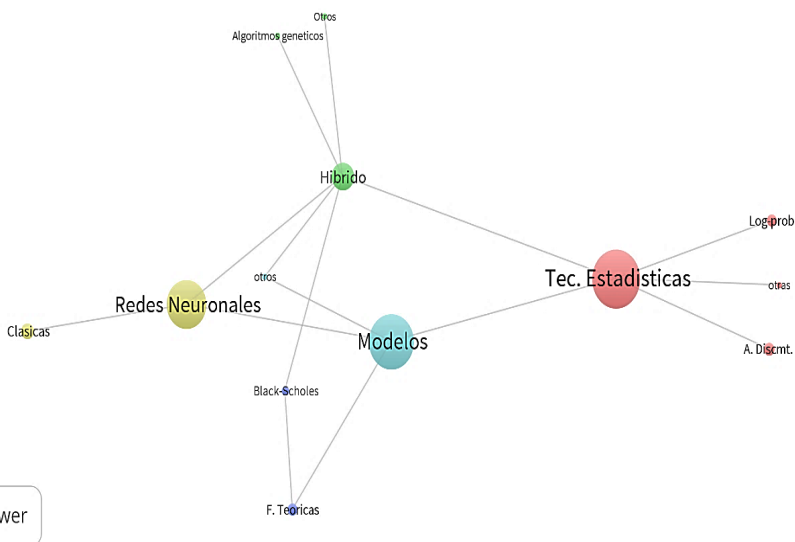
Fuente: Elaboración propia.

4. Tendencias en la literatura

El análisis comprende 30 estudios⁷ publicados en revistas internacionales de alto prestigio entre el 2000 y el 2015. Con la idea de analizar las tendencias en la literatura se recurre al empleo de un *mapa de visualización*, el cual permite organizar un cúmulo de referencias a través de concentraciones (*cluster*) y conexiones entre grupos. El tamaño de las concentraciones representa la frecuencia de empleo de cada grupo o subgrupo, mientras que las conexiones expresan la relación existente entre ellos.

7 Véase, apéndice 2, cuadro A.2, para mayor detalle sobre los estudios analizados.

Gráfica 3. Frecuencia por tipo de modelo empleado (2000-2015)



Fuente: Elaboración propia.

Esta frecuencia tiene las siguientes premisas: 1) las investigaciones que emplean diversos modelos de la misma naturaleza sólo se contabilizan una vez, 2) las investigaciones que emplean variantes y/o extensiones se consideran de la misma clase, y 3) las investigaciones que fusionan modelos de la misma naturaleza permanecen bajo la misma clasificación. Para el caso de los grupos, estos son los modelos expuestos en la sección anterior; sin embargo, si una referencia no se apega a alguno de ellos, se incluye en *Otros*.

Para generalizar la información, es necesario redefinir a cada grupo: el grupo 1 (ADM) por *técnicas estadísticas*, el grupo 2 (B-S) por *fórmulas teóricas* y el grupo 3 (PML) por *redes neuronales*. Así, los cuatro grandes grupos son: *técnicas estadísticas* (ADM, modelos logit/probit y otras), *fórmulas teóricas* (B-S y otras), *redes neuronales*

(PML, función de base radial, mapas auto-organizados⁸, máquina de vector soporte⁹ y otras) y *Otros* (redes bayesianas y otras).

Existe un subgrupo denominado *híbrido*, el cual concentra las herramientas complementarias como la lógica difusa, minería de datos y, en especial, los algoritmos genéticos (AG), así como también incluye la fusión de modelos de diferente naturaleza. En términos generales, el grupo de *redes neuronales* es el más recurrente a emplear estas herramientas, por lo que se realiza una división entre redes clásicas e *híbridas*. El software que se utiliza es *VOSviewer*.

En la gráfica 3 se indica que los modelos más empleados son las *técnicas estadísticas*. Por su parte, las *redes neuronales* han ganado terreno de manera significativa, mientras que las *fórmulas teóricas* muestran poca frecuencia en el campo. El subgrupo *híbrido* se define como la tendencia más reciente en la literatura, siendo los AG la técnica más empleada en este ámbito.

En el cuadro 6 se muestran los porcentajes del mejor desempeño (de acuerdo con los criterios de cada autor) por grupo. Con la idea de mantener cierta homogeneidad con los resultados anteriores se omiten aquellos estudios donde: 1) el comparativo sea entre modelos de la misma naturaleza, y 2) el comparativo resulte imparcial o contradictorio. Los resultados indican que frecuentemente las *redes neuronales* superan a las *técnicas estadísticas*; no obstante, los modelos *híbridos* superan a su contraparte sin excepción.

8 Los mapas auto-organizados son un tipo de red que tiene un aprendizaje no supervisado.

9 La máquina de vector soporte se incluye dentro de este grupo, debido a que cumple con el requisito básico (conexiones entrenables) para considerarla como una red neuronal.

Cuadro 6. Comparativo del mejor desempeño por grupo

	Grupos	Porcentaje
1	Técnicas estadísticas	-
2	Redes neuronales	40
3	B-S	10
4	Otros	10
	Híbridos	40

Fuente: Elaboración propia

5. Redes híbridas

Dentro de todas las posibilidades de *redes híbridas* las redes evolutivas¹⁰ ocupan un papel esencial en el campo. Este tipo de redes han logrado superar aspectos como: la búsqueda de soluciones cercanas a las “óptimas” entre una cantidad importante de posibles soluciones (Castillo *et al.*, 2001).

5.1 Redes evolutivas

La computación evolutiva (CE) fusiona los avances en la evolución humana junto con los avances computacionales. Dentro de la CE existen diversas metodologías como las estrategias evolutivas, la programación evolutiva y los AG. Los primeros difieren del último, principalmente en los operadores genéticos, en la reproducción y en la selección de estrategias (Mehrotra *et al.*, 2000). El objetivo de los AG es diseñar e implementar *sistemas adaptativos robustos*, siguiendo los paradigmas de la naturaleza a partir de las estructuras genéticas de la evolución.

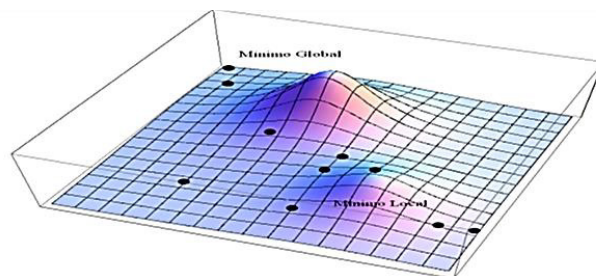
En 1975 con la publicación del libro *Adaptation in Natural and Artificial Systems*, Holland presenta a los AG como una abstracción de la evolución biológica y los define como un método de cambio de una población de cromosomas a otra usando un tipo de *selección natural* junto

10 Los AG pueden fusionarse con cualquier tipo de red; sin embargo, el PML continúa siendo una buena guía.

con los operadores genéticos (recombinación, mutación e inversión). Teniendo como principal atractivo utilizarlos como métodos de búsqueda ante problemas altamente complejos o incluso como diseñadores de soluciones innovadoras, entre otros (Mitchell, 1999). Por ejemplo, una solución candidata en un óptimo global tiene la *mejor o igual* calidad que cualquier otra solución, mientras que una solución candidata en un óptimo local tiene una calidad que *no puede ser mejorada* por un movimiento en la superficie, ya que la vecindad en los estados del espacio es de menor calidad, además en superficies complejas donde existen múltiples óptimos locales y ruido, es posible no acceder de forma sencilla al óptimo global.

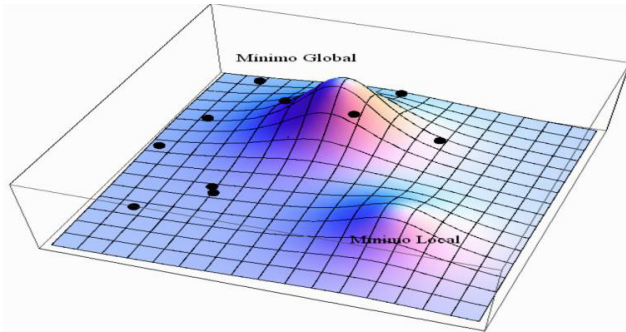
Por ello, el proceso evolutivo de búsqueda genera una población inicial¹¹ de posibles soluciones candidatas, las cuales realizan una búsqueda generalizada en toda la superficie. Los operadores genéticos toman relevancia al generar nuevas poblaciones que después de n generaciones garanticen soluciones *cercanas al óptimo global* (Jacob, 2001). En la parte superior de la gráfica 4 se muestra cómo el proceso de búsqueda se inicia generando una población de soluciones candidatas, mientras que en la parte inferior se muestra como después de n generaciones se obtienen mejores soluciones.

Gráfica 4. Proceso evolutivo de búsqueda
Población inicial de soluciones candidatas



11 La población inicial supone que contiene algunos elementos de calidad para solucionar el problema.

Población de soluciones después de n generaciones



Fuente: Elaboración propia con base en Jacob (2001) con *Mathematica*.

La idea central de utilizar a los AG en el diseño de las RNA es porque establecen una *guía formal* que ayuda a encontrar buenas soluciones de una forma rápida y eficaz. Esta combinación se visualiza bajo tres vertientes: 1) encontrar el valor óptimo en la inicialización de los pesos de la red, 2) encontrar de manera eficiente la *arquitectura*¹² (o topología) de la red, y 3) definir la magnitud de las reglas de aprendizaje.

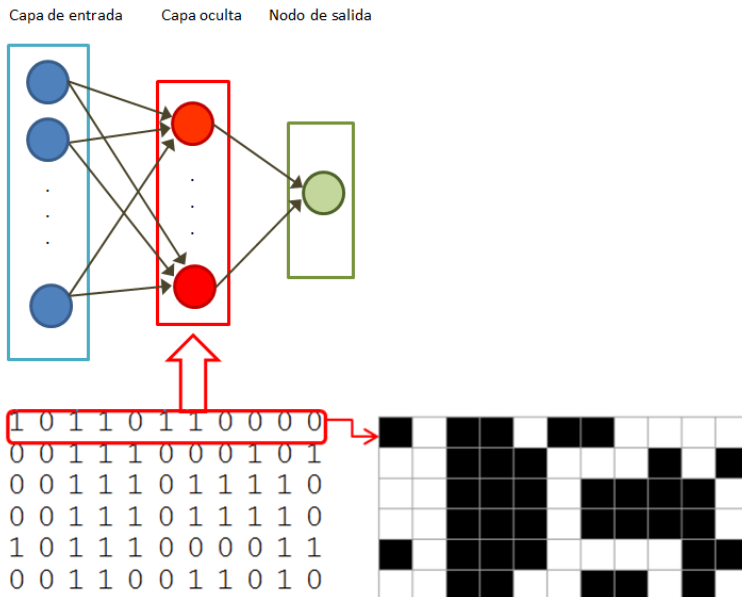
Los criterios para que las RNA tengan un buen desempeño son diversos; sin embargo, la mayoría considera que hacer *evolucionar a la arquitectura* a través de los AG es un elemento esencial. La arquitectura del PML está definida por una sola capa oculta, por lo que el problema se reduce a sólo encontrar *el número óptimo de nodos de la capa oculta* y, en su caso, de *la capa de entrada*.

En la gráfica 5 los *bits* de la cadena binaria representan los nodos de las capas del PML, así cada vector fila es una posible solución, y el conjunto de las posibles soluciones es la población. Para que esta población evolucione es necesario asociarle una “función de aptitud” (la cual

12 En esta investigación se ha dado énfasis al PML, sin embargo, es importante considerar que la búsqueda de la arquitectura no es propia de esta red. De hecho, es un problema que atañe a diversos tipos de redes, y sólo algunas escapan de este inconveniente como son: las redes modulares, las redes dinámicas y otras.

emula el proceso de *selección natural* de Darwin). Para el caso de hacer evolucionar la arquitectura del PML, basta con asociarle a cada posible solución el error cometido (etapa de entrenamiento). Mientras los operadores genéticos generan diversidad y dirección con cada generación.

Gráfica 5. Soluciones candidatas en el diseño evolutivo del PML



El número de nodos de entrada y ocultos del PML está representado por el sistema binario, donde 1 indica nodo activo y 0 nodo no activo. La matriz de 10×6 del lado izquierdo es la misma que se presenta del lado derecho, sólo que el 1 está representado por el color negro y el 0 por el color blanco, lo cual ayuda a representar la evolución de las soluciones candidatas. El tamaño de la matriz va a depender de la longitud de la cadena binaria y del tamaño de la población.

Fuente: Elaboración propia con base en Jacob (2001).

6. Discusión

En esta sección se discuten tres temas esenciales: las *fuentes de información*, las *redes híbridas* y las *técnicas estadísticas*. El primero por-

que ha generado una gran controversia entre los modelos, el segundo porque es un tema muy atractivo en el campo, y el último porque ayuda a comprender todos los argumentos anteriores.

Primero ¿por qué genera tantas críticas el uso de las razones financieras? Porque éstas se generan con información contable, la cual es estática y está correlacionada. Mientras la información bursátil es altamente dinámica y se vincula con las fuerzas del mercado. Entonces, si es tan efectiva la información bursátil, ¿por qué no usar únicamente este tipo de información? O más aun ¿por qué las *fórmulas teóricas* son tan poco atractivas para este tipo de aplicaciones? La información bursátil refleja la sensibilidad del mercado e incluso de otros factores de diversa naturaleza (psicológica, política, social, etc.) y, por tanto, el precio de una acción no mantiene una relación estrecha con el funcionamiento interno de la empresa, es decir, el precio es reflejo de efectos exacerbados que, sin duda, tendrá una repercusión en la empresa, pero con una influencia distinta a la reflejada con la información contable. Por otro lado, si se considera que para cotizar en bolsa se requiere presentar “frecuentemente” los estados financieros de la empresa, entonces la relación entre la información contable y bursátil puede coexistir (García y Morales, 2016); de hecho, la primera sustenta a la segunda. Lo anterior, no debe entenderse en términos de superioridad, sino de una posible *complementariedad*, lo que pone en relieve las connotaciones peyorativas hacia la información contable.

Segundo ¿cuál es el principal atractivo de las RNA en la detección de quiebras corporativas? y ¿las *redes híbridas* son los modelos “perfectos”? El principal atractivo de estas redes es que ofrecen la oportunidad de modelar agentes con “mente” bajo un entorno estructurado (la empresa y/o la bolsa). Aquí las quiebras corporativas se analizan con agentes heterogéneos, con comportamientos no lineales y con *racionalidad limitada* (Simon, 2006). Esto permite acercamientos más reales con el entorno, pero con la implicación de que los agentes cometen “errores”. En cuanto a la segunda pregunta, la respuesta es no. Los límites (o deficiencias) de las RNA son diversos; sin embargo, la mayoría se suelen superar con herramientas complementarias, pero hay otros

como *la falta de interpretación de los parámetros* que simplemente no es posible superar.

Tercero ¿por qué ha perdurado el empleo de las *técnicas estadísticas*?, o más bien ¿los resultados obtenidos son tan solo parte de un proceso de adaptación? Desde la incorporación de las RNA a las finanzas, se ha mostrado amplia evidencia de la superioridad de éstas sobre las *técnicas estadísticas*, es decir, es un proceso que lleva más de 20 años. Por lo tanto, la complejidad del tema no solo abarca un simple proceso de adaptación, sino también una cierta “preferencia” por las técnicas estadísticas, ya que brindan una metodología bien establecida, una interpretación clara de los parámetros y una evidente transparencia en la resolución del problema, aspectos que simplemente las *redes neuronales* no pueden ofrecer. Si se entiende que por la naturaleza de éstas no es posible ofrecer el mismo mecanismo de análisis e interpretación, entonces se habrá entendido también que su sustento ideológico es diferente a la visión ortodoxa del pensamiento económico y, en consecuencia, de las teorías financieras tradicionales.

7. Conclusiones

La importancia de analizar simultáneamente la frecuencia y el desempeño de los modelos en la detección de quiebras corporativas ayuda a tener una idea más amplia y objetiva sobre las tendencias en la literatura, sobre todo cuando las posibilidades son tan diversas.

En esencia, los modelos más empleados son las *técnicas estadísticas* seguidas de las *redes neuronales*, mientras que las *fórmulas teóricas* muestran poca frecuencia en el campo. La tendencia más reciente es la aplicación de modelos *híbridos*, la cual muestran la posibilidad de influir bajo una dinámica evolutiva. Lo anterior sugiere que: 1) *las técnicas estadísticas* no solo se sustentan como un punto de referencia, sino que en algunos casos es la única opción; y 2) *las redes evolutivas* (modelo *híbrido* más empleado) no pueden superar algunas deficiencias (falta de interpretación de los parámetros), lo que podría explicar –al menos en parte– el argumento anterior.

Por otro lado, el desempeño de los modelos señala que frecuentemente las *redes neuronales* superan a las *técnicas estadísticas*; no obstante, los modelos *híbridos* rebasan a su contraparte sin excepción. Los resultados anteriores sugieren que: 1) las *redes neuronales* efectivamente son capaces de resolver problemas complejos debido a su alto grado de flexibilidad; 2) las *redes híbridas* mejoran notablemente su desempeño al emplear herramientas complementarias (optimizar parámetros, selección de información, etc.), ya que se liberan de procesos de error y prueba; y 3) las *redes evolutivas* (red *híbrida* más empleada) sustentan que hacer evolucionar la arquitectura de la red es la “clave” para garantizar un buen desempeño.

Apéndice 1

Cuadro A.1. Programación básica para el PML (primera iteración)

Primero, se establece el número de nodos para cada capa. En este caso, tenemos seis para la capa de entrada, tres para la capa oculta y uno para la capa de salida.
<pre>inNumber = 6 hidNumber = 3 outNumber = 1</pre>
Posteriormente, se establece el rango para la inicialización de los pesos. Y como todos los nodos de la capa de entrada se encuentran conectados con la siguiente, tenemos en total 21 pesos o valores iniciales para cada conexión.
<pre>hidWts=Table[Table[Random[Real,{-0.1,0.1}],{inNumber}],{hidNumber}] {{ { 0 . 0 7 5 1 7 5 6 , 0 . 0 9 2 4 7 3 5 , - 0 . 0 8 4 3 4 9 7 , - 0.0630787,0.0629712,0.0706455 }, {0.0653188,-0.00464515,- 0.0756743,0.0961266,-0.0926137,-0.059189}, {0.0733084,0.0254089,- 0.0143054,0.0946497,0.0603946,-0.0646163}}</pre>
<pre>outWts=Table[Table[Random[Real,{-0.1,0.1}],{hidNumber}],{outNumber}] {{{-0.0197106,-0.0875833,-0.0932606}}</pre>

Después se realiza la selección del vector fila (patrón de entrenamiento) de forma aleatoria. Su procesamiento inicia con la función sigmoïdal tanto para la capa oculta como para la capa de salida.
<code>ioP=ioPairs[[Random[Integer,{1,Length[ioPairs]}]]]</code> <code>{{0.2,0.2,0.2,0.1,0.1,0.1},{0.5}}</code>
<code>inputs=ioP[[1]]</code> <code>{0.2,0.2,0.2,0.1,0.1,0.1}</code>
<code>outDesired=ioP[[2]]</code> <code>{0.5}</code>
<code>sigmoid[x_]=1/(1+e-x)</code>
<code>hidOuts=sigmoid[hidWts.inputs]</code> <code>{0.505928,0.497858,0.506481}</code>
<code>outputs=sigmoid[outWts.hidOuts]</code> <code>{0.474819}</code>
Finalmente, se obtiene el error:
<code>outErrors=outDesired-outputs</code> <code>{0.0251814}</code>

Fuente: Elaboración propia con base en la programación propuesta en Freeman (1993) con *Mathematica*.

Cuadro A.2 Aplicaciones de quiebras corporativas

	Autor (es)	Año	Tipos de Modelo ^{1/}					País ^{2/}	Tamaño de las empresas/ Resultados ^{3/} (mejor desempeño)	Fuente
			1	2	3	4	5			
1	Blanco, et al.	2015	•					RU	Pequeñas/red neuronal	Journal of Economics and Finance
2	López y Pastor	2015	•			•		EU	Grandes/red híbrida	Expert Systems with Applications
3	Heo e Yang,	2014	•			•		COR	Diversos/ Otros	Applied Soft Computing
4	Kulsár	2014	•					RUM	Med. y peq. / n.a	University of Oradea
5	Briñdescu y Goleț	2013	•					RUM	Diversos/ n.a4/	Annals of the University of Oradea.
6	Kasgari, et al.	2013	•			•		IRN	Grandes/ red neuronal	International Journal of Academic Research in Accounting, Finance and Management Sciences
7	Ahmadpour, et al.	2013	•			•		IRN	Grandes/ red neuronal	
8	Deligianni y Kotsiantis	2012				•		GRE	Grandes/ Otros	Artificial Intelligence: Theories and Applications
9	Kiyak y Labanauskaitė	2012	•					LIT	Med. y peq./ n.a	Economics and Management
10	Peat y Jones	2012	•			•		AUS	Diversos/red neuronal	Intelligent Systems in Accounting, Finance and Management
11	Telmoudi, et al.	2011				•		TUN	red híbrida	Intelligent Systems in Accounting, Finance and Management

12	Mandru, <i>et al.</i>	2010	•					RUM	Med./n.a	Recent Advances in Artificial Intelligence, Knowledge Engineering and Data Bases
13	Tzong-Huei	2009	•	•				TAI	n.a	Neurocomputing
14	Xu y Zhang	2009	•	•		•		JAP	Grandes /modelo híbrido	Review of Accounting Studies
15	Hsueh-Ju, <i>et al.</i>	2009	•	•		•		TAI	Grandes/ red híbrida	Expert Systems with Applications
16	Patel y Pereira	2007	•	•				RU	Grandes/fórmulas teóricas	The Journal of Real Estate Finance and Economics
17	Sun y Shenoy	2007				•			n.a	European Journal of Operational Research
18	Bo y Zheng	2006	•	•				CHN	Grandes/red neuronal	International Conference on Natural Computation
19	Merkevicius, <i>et al.</i>	2006	•	•		•		EU	Grandes/ red híbrida	International Conference on Computational Science
20	Min e Young-Chan	2005	•	•		•		COR	Med. y peq./ red neuronal	Expert Systems with Applications
21	Härde, <i>et al.</i>	2005		•				EU	Grandes/ n.a	Statistical Tools in Finance & Insurance
22	Prasad y Puri	2005	•					EU	Grandes/ n.a	The International Journal of Finance
23	Brabazon y Keenan	2004	•	•		•		EU	Grandes/ red híbrida	Computational Management Science
24	Charitou, <i>et al.</i>	2004	•	•		•		RU	Grandes/ red neuronal	European Accounting Review

25	Hillegest, <i>et al.</i>	2004	•			EU	Grandes/fórmulas teóricas	Review of Accounting Studies
26	Kyung-Shik, <i>et al.</i>	2004		•		COR	/ n.a	Neural Information Processing
27	Anandarajan, <i>et al.</i>	2001	•		•	EU	Grandes/red híbrida	International Journal of Intelligent Systems in Accounting, Finance and Management
28	Shirata y Terano	2000	•		•	JAP	Grandes/técnica estadística híbrida	Knowledge Discovery and Data Mining
29	Charalambous, <i>et al.</i>	2000	•		•	EU	Diversos/red neuronal	Annals of Operations Research
30	Refait, C.	2000		•		FR	n.a	Finance

^{1/} Los tipos de modelos son: 1) técnicas estadísticas, 2) fórmulas teóricas, 3) redes neuronales, 4) otros, y 5) modelos híbridos. ^{2/} Australia (AUS), Estados Unidos (EU), Francia (FR), Irán (IRN), Japón (JAP), Corea (COR), Reino Unido (RU), Rumania (RUM). ^{3/} De acuerdo con la clasificación propuesta (tipos de modelos). ^{4/} Refiere a uno de los siguientes casos: a) el comparativo es entre el mismo tipo de modelos, b) los resultados son contradictorios o no decisivos, o c) el estudio analiza sola una propuesta.

Fuente: Elaboración propia.

Bibliografía

- Altman, E. (1968). "Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy". *The Journal of Finance*, 23 (4), pp. 589-609.
- Black, F. y M. Scholes (1973). "The Pricing of Options and Corporate Liabilities", *Journal of Political Economy*, 81 (3), pp. 637-659.
- Castillo, P., J. Castellano, J. Merelo y A. Prieto (2001). "Diseño de Redes Neuronales Artificiales mediante algoritmos evolutivos". *Revista Iberoamericana de Inteligencia Artificial*, 14, pp. 2-32.
- Freeman, J. (1993). *Simulating Neural Networks with Mathematica*. Addison-Wesley Professional.
- García, O. y A. Morales (2016). "Desempeño Financiero de las Empresas: una Propuesta de Clasificación por RNA". *Dimensión Empresarial*, 14, pp. 11-22.
- Jacob, C. (2001). *Illustrating Evolutionary Computation with Mathematica*. Morgan Kaufmann Publishers.
- Mehrotra, K., C. K. Mohan, y S. Ranka (2000). *Elements of Artificial Neural Networks*. MIT Press. Cambridge, Massachusetts.
- Mitchell, M. (1999). *An Introduction to Genetic Algorithms*. MIT press. Cambridge, Massachusetts.
- Odom, M. y R. Sharda (1990). "A Neural Network for Bankruptcy Prediction". 1990 *International Joint Conference on Neural Networks (IJCNN'90)*.
- Pérez, M. (2006). "Artificial Neural Networks and Bankruptcy Forecasting: A State of the Art". *Neural Computation & Application*, 15 (2), pp.154-163.
- Scott, J. (1981). "The Probability of Bankruptcy: A Comparison of Empirical Predictions and Theoretical Models". *Journal of Banking and Finance*, 5 (1), pp. 317-344.

Simon, H. (2006). *Las Ciencias de lo Artificial*, Comares.

Zhang, G., M. Y. Hu, B. E. Patuwo y D. C. Indro (1999). “Artificial Neural Networks in Bankruptcy Prediction: General Framework and Cross-validation Analysis”. *European Journal of Operational Research*, 116 (1), pp.16-32.