

LA PREDICCIÓN DE SERIOS PROBLEMAS FINANCIEROS CON REDES NEURONALES ARTIFICIALES

EL CASO ARGENTINO¹

RICARDO PASCALE

Profesor Titular de Finanzas de Empresas
Universidad de la República
Montevideo, Uruguay

Mayo, 2000

¹ Remarco en este trabajo una renovada deuda de gratitud con el Prof. Edward I. Altman, por su impulso y guía. Se agradecen los aportes efectuados al mismo por los Profs. Miguel Galmés, Laura Nalbarte, Oscar Perelló y Raúl Ramírez así como de la Ec. Elizabeth Bucacos y del del Cr. Juan P. Rodríguez. La responsabilidad sobre el trabajo permanece con el autor.

INDICE

	Pág.
1. INTRODUCCION	1
2. BREVE REVISION DE LA LITERATURA	2
3. PROBLEMA, DEFINICIONES Y METODO DE INVESTIGACION	4
4. SELECCIÓN DE VARIABLES, RECOLECCIÓN DE DATOS E INFORMACION ESTADISTICA DE LOS MISMOS	7
5. ANALISIS DISCRIMINANTE MULTIVARIANTE. BREVE INTRODUCCION	11
6. REDES NEURONALES ARTIFICIALES. BREVE INTRODUCCION	16
7. EXPERIMENTO	22
8. DISCUSION	43
9. CONCLUSIONES	47

REFERENCIAS BIBLIOGRAFICAS

E. L. Altman, *Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy*, The Journal of Finance 23 (September 1968), pp. 589-609.

E. Altman, M. Margaine, M. Schlosser and P. Vernimmen, 1974, *Statistical Credit Analysis in the Textile Industry: A French Experience*, Journal of Financial and Quantitative Analysis, March.

E. Altman, 1977, *Predicting Performance in the Savings and Loan Industry*, Journal of Monetary Economics, October.

E. Altman, R. Haldeman and P. Narayanan, 1977, *ZETA Analysis; a New Model to Identify Bankruptcy Risk of Corporations*, Journal of Banking and Finance, June.

T.W. Anderson, 1949, *Classification by Multivariate Analysis*, Psychometrika 16.

R.A. Collins and R.D. Green, *Statistical Method for Bankruptcy Forecasting*, Journal of Economics and Business 32 (1982), pp. 349-54.

E. Deakin, 1972, *A discriminant Analysis of Predictors of Business Failure*, Journal of Accounting Research, Spring.

R. Edmister, 1972, *An Empirical Test of Financial Ratio Analysis for Small Business Failure Prediction*, Journal of Financial and Quantitative Analysis, March.

R. Eisenbeis, 1977, *Pitfalls in the Applications of Discriminant Analysis of Business*, Finance and Economics, The Journal of Finance, June.

R. A. Fisher, 1936, *The Use of Multiple Measurements in Taxonomic Problems*, Annals of Eugenics, September.

C. Harris, *An Expert Decision Support System for Auditor Going Concern Evaluation* Ph.D. Diss., University of Texas at Arlington, 1989.

S. Haykin, *Neural Networks: A Comprehensive Foundation*, MacMillan College Publishing Co., New York, 1994.

J.O. Horrigan, *A Short History of Financial Ratio Analysis*, The Accounting Review 43 (April 1968), pp.284-94.

- S. Kanitz, 1974**, *Como prever la Falencia de Empresas*, Examen, Dec.
- G.V. Karels and A. J. Prakash**, "Multivariate Normality and Forecasting of Business Bankruptcy, *Journal of Business Finance & Accounting* 14 (Winter 1987), pp. 573-93.
- P. Lachenbruch, 1967**, *An almost unbiased method of obtaining confidence. Intervals for the Probability of misclassification in discriminant Analysis*, *Biometrics*, Dec.
- P. Meyer and H. Pifer**, 1970, *Prediction of Bank Failures*, *Journal of Finance*. September.
- D.F. Morrison**, 1976, *Multivariate Statistical Methods*, 2nd ed. (McGraw-Hill, New York).
- R. Pascale**, *Decisiones Financieras*, 3^a. Edición, Ediciones Macchi, Buenos Aires, 1998.
- R. Pascale y L. Preve** *Modelo Discriminante Multivariante para serios problemas financieros en empresas concesionarias de automóviles*, Buenos Aires, Diciembre 1999.
- R. Pascale** 1988 *Un modelo Multivariante para predecir serios problemas financieros en empresas: el caso uruguayo*, *Studies in Banking and Finance* 7 /1988) 171-182 North Holland.
- J. Sinkey**, 1975, *A Multivariate Analisis of the Characteristics of Problem Banks*, *Journal of Finance*, March.
- R. Taffler and H. Tisshaw**, 1977, *Going, going, gone - Four Factors Which Predict, Accountancy*, March.
- Robert Trippi, y Efrain Turban** (1996), *Neural Networks in Finance and Investing*, Irwin Professional Publishing.
- A. Wald**, 1944, *On a Statistical Problem Arising in the Classification of an Individual into one of Two Groups*, *Annals of Mathematical Statistics* 15.

1. INTRODUCCION

Esta investigación exploratoria se encuadra dentro del tema general del análisis financiero. Dentro de él, uno de los tópicos que ha despertado más interés ha sido la predicción de bancarrota de empresas, o en términos más amplios y frecuentemente más ajustados a varios países de América Latina, la predicción de serios problemas financieros en empresas. El objetivo general del trabajo es explorar las capacidades de dos de las técnicas utilizadas para estos efectos, como son el análisis discriminante multivariante (ADM) y las redes neuronales artificiales (RNA).

Como objetivo específico, el trabajo se propone determinar y analizar la calidad de la predicción de ambas técnicas aplicadas a un caso concreto (el de las empresas concesionarias de automotores en la República Argentina) y explorar las posibilidades de su uso conjunto.

La Sección 2 del trabajo repasa la reciente literatura sobre el tópico, en tanto que la Sección 3 considera el *problema* de la investigación así como las definiciones y los métodos utilizados.

La selección de las variables a utilizar, la recolección de datos así como cierta información estadística de los mismos se presentan en la Sección 4.

Las Secciones 5 y 6 se destinan a repasar una breve introducción del ADM y de las RNA. La Sección 7 está dedicada a desarrollar el experimento efectuado que consta de cuatro *etapas*, que incluyen ADM y RNA, y exponer los resultados del mismo.

La discusión de los resultados del modelo, y la presentación de los principales hallazgos y conclusiones se expone en la Sección 8. Finalmente, se presentan las conclusiones finales y campos de futura investigación.

2. BREVE REVISION DE LA LITERATURA

La predicción de bancarrota o de serios problemas financieros en empresas ha sido objeto de amplia investigación, en particular en los últimos 40 años.

La evolución de modelos univariados, ha ido desde la utilización de ratios para esta predicción (**Horrigan**, 1969), a modelos no cuantitativos y, a modelos de regresión de una variable con otra para predecir tendencias, se ha pasado a integrar los ratios en técnicas más avanzadas como modelos econométricos o estadísticos. En esta línea, tres han sido los instrumentos básicos utilizados, el ADM, el análisis logístico y el RPA (Recursive Partitioning Algorithm). A los efectos de este trabajo se repasan algunos ajustes a la literatura afin.

El trabajo pionero en el tema de la aplicación de ADM a la predicción de bancarrota se debe a **Edward Altman** (1968) que desarrolló el primer modelo en Estados Unidos para predecir bancarrota de firmas manufactureras. El ADM como se verá más adelante con mayor detalle, es una técnica estadística que permite clasificar una nueva observación en uno de los dos o más grupos definidos a priori. La clasificación frecuentemente se obtiene por una función discriminante que es una combinación lineal de las variables. Esta función que se asienta en varios supuestos y es derivada de forma tal de minimizar los errores de clasificación. En el mismo país pueden encontrarse entre otros, los trabajos de **Edmister** (1972) para pequeñas empresas, **Deakin** (1972) para manufactureras, **Meyer y Pifer** (1970) para bancos, **Sinkey** (1975) bancos, **Altman y Lorriss** (1976) para operadores de bolsas, **Altman, Haldeman y Narayanan** (1977) manufactura y minoristas y **Altman** (1977) para asociaciones de ahorro y préstamo.

Asimismo, en Gran Bretaña, se encuentran los trabajos de **Taffler** (1977), en Francia, **Altman, Margaine, Scholsser y Vernimmen** (1973) así como en otros países de Europa (Italia, Noruega, Finlandia, Alemania) se encuentran trabajos de análisis discriminante para predecir bancarrota de firmas. Para América Latina puede verse el trabajo de **Pascale** (1988).

Collins y Green (1982) han repasado y comparado las distintas técnicas estadísticas para la predicción citada.

El análisis logístico es un modelo logístico asentado en la función acumulativa logística, de especial utilidad cuando se trata de variables binarias dependientes, así como también cuando existen variables explicativas tanto cuantitativas como cualitativas en este caso el discriminante es un discriminante logístico. Conforme a los autores, si bien el análisis logístico parece presentar menos errores de clasificar empresas como "buenas" que son "malas", en su utilización no llega a ser superior en la precisión de la predicción del análisis discriminante.

El RPA es una técnica no parametrizada de clasificación que se basa en patrones de reconocimiento. El modelo va asignando los comportamientos en base a un árbol binario de clasificación derivado en forma de minimizar el costo de los errores. El modelo ha mostrado en general menor contundencia de resultados, según los autores, que el ADM.

El análisis llevado a cabo por **Harris (1989)** con empresas que habían ingresado en bancarrota y otras que no, arrojó para la muestra con la que se elabora el modelo y para una muestra separada para testing los siguientes resultados: el ADM, para la muestra del modelo reportó una precisión del 84.5% y en la de testing el 85%, para el análisis *logit*, 87 y 78% respectivamente y para RPA, 84% (para bancarrota) y 90% (para no bancarrota) en la muestra de análisis y 76% (para bancarrota) y 90% para bancarrota en la muestra de prueba.

Más recientemente han aparecido investigaciones que aplican redes neuronales artificiales a la predicción de serios problemas financieros en empresas. Entre las más conocidas están trabajos de **Odom y Sarda** (1990), **Coats et al** (1993) y, **Coleman et al** (1992).

Luego del desarrollo de los primeros modelos de RNA, han aparecido investigaciones que buscan la comparación de las RNA y el ADM para predicción de bancarrotas, entre ellas **Altman et al** (1994). En este último trabajo si bien se reconoce la importancia y el futuro potencial de desarrollo de las RNA, concluyen

que no es claro el dominio de esta técnica con respecto al ADM. Otros autores, probablemente la mayoría, como **Raghupati**, et al (1991) **Rahimian** et al (1992) sostienen definitivamente la importancia de las RNA en la predicción de serios problemas financieros en empresas.

3. PROBLEMA, DEFINICIONES Y METODOS DE INVESTIGACION

Problema

Diversos métodos estadísticos se han desarrollado para dar soluciones al problema de la clasificación. En el caso de esta investigación, el centro es la clasificación entre empresas que presentan serios problemas financieros (PF) y las que no presentan serios problemas financieros (NPF). Son modelos que buscan establecer a cuál de las categorías definidas a priori se clasifica un nuevo caso.

Uno de los métodos más extendidos, a estos efectos, es el **análisis discriminante multivariante (ADM)**. En este caso, la existencia de supuestos en el relacionamiento de las variables involucradas en el análisis lleva a que cuando la forma funcional no es la adecuada, esta situación es confirmada y los modelos no prevén en la forma correcta. Se supone, por ej. que la distribución conjunta de las variables discriminantes es una función normal multivariada. Se ha estudiado, **Karels y Prakash** (1987), que los resultados, cuando no se da ese supuesto son erróneos. Se trata, en este caso de modelos poco "robustos".

Las **redes neuronales artificiales, (RNA)**, no operan bajo los supuestos de normalidad y, en general son modelos más "robustos" en el sentido que sus resultados no dependen de supuestos de partida. Las redes neuronales son un mecanismo dirigido para la inferencia estadística no parametrizada. En términos estadísticos, puede expresarse que

dada una muestra $D_n = \{(x_i, y_i)\} \quad i= 1, 2, \dots, n$

generada por la función desconocida $\varphi(\mathbf{x})$ con la adición de un componente estocástico ε , se tiene que

$$y_i = \varphi(\mathbf{x}_i) + \varepsilon_i$$

y a través del aprendizaje busca construir un estimador $\hat{g}(\mathbf{x}; \mathbf{w}; D_n) \equiv \hat{\varphi}(\mathbf{x})$ de $\varphi(\mathbf{x})$, donde $\mathbf{W} = (w_1, w_2, w_3, \dots, w_p)$ es un set de parámetros libres que se conocen como pesos de conexión (se verá más adelante como vector de pesos) y D_n el set finito de observaciones.

Por lo tanto, en RNA no existen supuestos *a priori* de la normalidad de $\varphi(\mathbf{x})$ y el modelo neural $\hat{g}(\mathbf{x}; \mathbf{w}; D_n)$ es un estimador *no parametrizado* en oposición a estimadores *parametrizados*, donde la forma funcional es supuesta a priori. La Sección 6 de este trabajo se refiere con mayor detalle a las RNA.

El ADM, en su comparación analítica con las redes neuronales artificiales muestra más claramente los conceptos financieros subyacentes y el modelo lleva a interpretaciones que se entroncan con más nitidez con los fundamentos económicos.

En consonancia con los elementos desarrollados, las RNA han mostrado en su utilización un amplio poder y flexibilidad. Y han mostrado prever un marco de referencia más general para establecer relaciones entre los datos y que no requieren de la especificación de ninguna forma funcional.

Definiciones

La investigación trata acerca de la clasificación a priori de empresas con serios problemas financieros o sin ellos. Se trata, como se ha señalado de empresas concesionarias de automotores en la República Argentina.

Por tanto, para efectos de esta investigación se entiende por empresas con serios problemas financieros a aquéllas que no solamente han ingresado en figuras jurídicas como la quiebra o la bancarrota. Por el contrario, incluye toda una amplia gama de

situaciones en las que pueden aparecer figuras jurídicas como el concordato o similares e incluso situaciones no estipuladas en forma especial por el derecho, como es el caso de severas restricciones financieras que dificultan gravemente el servicio de deudas así como el financiamiento de la actividad corriente. Esta área menos precisa jurídicamente aparece, empero, muy extendidamente en la evidencia empírica.

Metodología

Dos técnicas se han utilizado en la presente investigación para alcanzar el objetivo establecido: el ADM y las RNA.

En primer término se exponen las variables a utilizar, esto es, las variables que se han seleccionado para explicar la clasificación, así como la recolección de los datos que representan a las mismas a lo que se agregan algunas de sus principales características estadísticas.

El proceso metodológico escogido se beneficia de un modelo discriminante multivariante, (**Pascale** et al, 1999), desarrollado para la clasificación que importa a este trabajo (esto es la clasificación en base a una definición a priori de firmas "buenas" y "malas" financieramente entre empresas concesionarias de automotores de las distintas marcas que operan en la República Argentina. Los datos que luego se utilizan para redes neuronales son extraídos de esa investigación.

En esta investigación, el modelo de ADM, se expone como *Etapa Experimental 1* y, representa la aproximación que arrojó mejores resultados de clasificación, que fue la función discriminante lineal obtenida por el método *stepwise*.

Se ingresa a las RNA, en la *Etapa experimental 2*, en donde utilizando como se estableció los mismos datos con los cuales se obtuvo el modelo de ADM. En este experimento, se utilizan como *inputs* todas las variables seleccionadas, aplicando el algoritmo de aprendizaje, *Back-propagation*.

La *Etapa experimental 3*, aplica también RNA, con *inputs* en las dos variables que seleccionó el modelo de ADM. En este experimento el algoritmo utilizado es *Perceptron*.

La *Etapa Experimental 4*, es la aplicación de RNA, en el algoritmo *Back-Propagation*, utilizando como *inputs* las variables explicativas del modelo de ADM, que fueron dos.

Con los resultados obtenidos en los cuatro experimentos se pasa a la discusión de los resultados y a sintetizar los principales hallazgos.

4. SELECCIÓN DE LAS VARIABLES, RECOLECCION DE DATOS E INFORMACION ESTADÍSTICA DE LOS MISMOS

Selección de variables

1. Variable cualitativa

La variable cualitativa a investigar en el problema del presente trabajo es la presencia de serios problemas financieros de empresas concesionarias de automotores correspondientes a las fábricas de las nueve marcas principales. La existencia o no de esta variable cualitativa en presencia de un nuevo caso a investigar, se efectuará a través de variables cuantitativas que se exponen en la sección siguiente.

2. Variables cuantitativas

En la selección de las variables cuantitativas se procedió por etapas.

La *primera*, consistió en revisar la etiología de los grandes problemas financieros de empresas.

La *segunda* etapa, fue a partir de esa etiología cuales pueden ser las variables cuantitativas que podrían explicar el fenómeno.

Se consideraron, luego de varias depuraciones en esta etapa primariamente 9 ratios, algunos de ellos manejando sólo variables de flujo, en otros sólo variables de stock y en una tercera categoría variables de stock y flujos.

CUADRO N° 1

<i>Variables cuantitativas seleccionadas</i>	
Ratio 1.	Patrimonio Neto / Deudas Totales
Ratio 2.	Ventas / Cuentas a Cobrar
Ratio 3.	Margen de Ventas Netas
Ratio 4.	Ganancia Neta / Total de Deudas
Ratio 5.	Cuentas a Cobrar + Inventarios / Cuentas a Pagar
Ratio 6.	Capital de Trabajo / Activos Totales
Ratio 7.	GAI / Activos Totales
Ratio 8.	Ventas / Deudas Totales
Ratio 9.	Ganancia Neta / Activo Total

Con ellos fue que se comenzó el trabajo de extracción de las funciones discriminantes.

Estos ratios fueron calculados para las 79 empresas de elaboración del modelo y las 64 de testing, que finalmente se utilizaron.

Recolección de datos

La base de datos para la confección del modelo fueron informaciones contables de empresas concesionarias de automotores que fueron suministradas por 9 empresas

fabricantes de automotores asociadas a ADEFA ¹, que se verificó en el segundo semestre de 1999.

En conjunto se enviaron los **estados contables de 94 empresas** concesionarias de automotores, las cuales incluían variada información en términos de cantidad y calidad de la misma.

Las 94 empresas originales se procedieron a depurar en una aproximación más refinada para llegar al grupo final de empresas a tomar como base para elaborar el modelo. De esta forma, algunas empresas fueron eliminadas por tener información incompleta, o tenerla defectuosa en el sentido de lógicas económicas en otros casos por ser estadísticamente *outliers*, así como por encontrarse alguna inconsistencia contable.

Luego de esta depuración, quedó una base de datos de 79 empresas para la extracción de funciones discriminantes multivariantes, distinguiéndose en 41 "buenas" y 38 "malas", las que se denotan como:

$$\begin{aligned}n_1 &= 41 \\n_2 &= \frac{38}{79}\end{aligned}$$

La muestra es bien significativa en tamaño y distribución.

Para el testing

A estos efectos se tomaron los datos de empresas no consideradas en la elaboración del modelo. El número inicial fue de 83 empresas.

Efectuada la depuración de los 83 datos para efectuar testing, ya sea también por información incompleta, por ser estadísticamente *outliers* o por encontrarse inconsistencias marcadas en las cifras, se obtuvo un total de 64 empresas, 30 "buenas" y 34 "malas", que serían:

$$N_1 = 30$$

¹ Asociación de Fabricantes de Automotores de la República Argentina.

$$N_2 = \frac{34}{64}$$

A las 79 empresas para la extracción de la función discriminante se le efectuaron los cálculos de las variables cuantitativas a utilizarse.

Otro tanto se efectuó para las 64 para la fase de uno de los tests de validación.

Información estadística de las variables

Estadísticas descriptivas

La siguiente tabla presenta la media y el desvío estándar de cada uno de los ratios en cada grupo de empresa:

CUADRO N° 2

Grupo al que pertenece	RATIO	Media	Desviación Estándar
BUENA (1)	Patrimonio Neto / Deudas Totales	0,85	0,79
	Ventas / Ctas. a cobrar	27,53	28,77
	Margen de ventas netas	0,03	0,02
	Ganancia neta / total de deudas	0,14	0,13
	Ctas. a cobrar + Inventarios / Activos Totales	2,30	2,94
	Capital de trabajo / activos totales	0,30	0,14
	GAI / activos totales	0,06	0,08
	Ventas / deudas	4,68	2,04
Ganancia neta / activo total	0,07	0,05	
MALA (2)	Patrimonio Neto / Deudas Totales	0,39	0,33
	Ventas / Ctas. a cobrar	12,33	12,99
	Margen de ventas netas	-0,03	0,09
	Ganancia neta / total de deudas	-0,02	0,09
	Ctas. a cobrar + Inventarios / Activos Totales	1,80	1,11
	Capital de trabajo / activos totales	0,11	0,25
	GAI / activos totales	-0,03	0,09
	Ventas / deudas	2,03	0,93
Ganancia neta / activo total	-0,02	0,08	

En el Anexo 1 se presentan los diagramas de cajas para los valores de cada ratio. El diagrama de caja es un gráfico basado en los cuartiles en la mediana y los valores extremos que contiene información sobre la simetría y la dispersión de la distribución

de los ratios. En estos diagramas se puede observar como se distribuyen los datos de las empresas para cada ratio según el grupo.

5. ANALISIS DISCRIMINANTE MULTIVARIANTE. BREVE INTRODUCCION ²

1. Consideraciones generales

En el mundo académico, la utilidad del uso del tradicional análisis de ratios, en la evaluación de la situación de las empresas, ha venido siendo motivo de críticas.

El análisis tradicional de ratios distinguiendo por ej. entre ratios de liquidez, endeudamiento, rentabilidad, actividad etc., se ha efectuado al margen de un contexto bien definido de teoría de toma de decisiones. La crítica más importante deriva del hecho que el mismo, no se ha acompasado a los desarrollos que en algunos campos iban teniendo la economía y las finanzas.

En el moderno análisis financiero, se insiste en que las técnicas de análisis financieros, como el caso de los ratios, se inserten en un marco de modelos decisorios.

Por tanto, una diferencia sustantiva entre el enfoque tradicional y moderno es que, en el primero el análisis se efectúa sin un contexto bien definido de teoría decisoria, lo que sí busca el segundo.

En esta línea de pensamiento, se vienen desarrollando esfuerzos por introducir el análisis de ratios, entre otras aplicaciones, dentro del modelo de predicción de las ganancias de las empresas, el crecimiento de éstas y la constitución de portafolios eficientes.

Un área donde se ha desarrollado el enfoque moderno, es en cuanto a la predicción de serios problemas financieros en las firmas, caso en el cual se han utilizado los ratios insertados en técnicas estadísticas multivariantes, como es el caso del análisis discriminante.

² Esta sección se basa en lo establecido en el trabajo *Un Modelo Multivariante para Predecir Serios Problemas Financieros en Empresas: El Caso Uruguayo*, de Ricardo Pascale, 1988

El análisis estadístico multivariante tiene relación con datos que se obtienen de diversas dimensiones de una misma empresa, o en otros temas (por ej. ciencias biológicas, o médicas) de un mismo individuo.

En la aplicación de predecir serios problemas financieros, el análisis multivariante, del tipo discriminante, busca de resolver, cuando se está en presencia de una nueva observación, a cuál de las poblaciones definidas a priori, ésta debe ser asignada en forma óptima.

El rasgo más significativo de este análisis discriminante multivariante es la consideración simultánea de diversos indicadores en el proceso de predicción. Estos indicadores aparecen adecuadamente ponderados, conforme a la técnica, de forma de obtener un índice general. Según el resultado que arroje el índice, aplicado a un caso concreto, nos brinda un elemento para clasificar, el caso estudiado, dentro de uno de los grupos definidos a priori.

El tema de la predicción de problemas financieros, a través del análisis discriminante multivariante, ha venido siendo objeto de amplio tratamiento en diversos países.

El objeto de este trabajo se encuadra en la tipología de las anteriores investigaciones.

2. Elementos del análisis discriminante multivariante

El análisis discriminante es una técnica estadística multivariante a través de la cual, dependiendo de sus características individuales, se clasifica una observación en uno de los varios grupos definidos a priori.

Esta técnica, ha sido utilizada en muchos campos del conocimiento a partir de su primer uso por **Ronald A. Fischer** (1936) y ha recibido, luego de un comienzo en la

biología y ciencias del comportamiento, una creciente aplicación recientemente, en finanzas y economía.

En el caso de este trabajo, se opera con dos grupos: empresas concesionarias de automóviles que han tenido serios problemas financieros como insolvencia (CPF) y empresas que no presentan problemas de insolvencia (SPF). Se supone que el comportamiento financiero de una firma está asociado a ciertos atributos observados. De esta forma, los resultados obtenidos pueden utilizarse a efectos de predecir.

Trabajando con las características de cada uno de los grupos o poblaciones, se llega a determinar una combinación de las mismas, en nuestro caso de tipo lineal o cuadrática que es la que "mejor discrimina" entre aquellas poblaciones. La idea de "mejor" debe interpretarse que bajo el método se maximiza la separación entre las poblaciones.

A estas características, a través del método, se le asignan adecuadas ponderaciones a efectos de establecer una única medida.

Ante un caso concreto, tomando las ponderaciones obtenidas y utilizando las características del caso tratado se llega a determinar un valor de su índice. Según sea este valor, la empresa se clasifica en un grupo o en el otro.

El análisis discriminante supone que las poblaciones siguen una distribución multinormal con diferentes vectores de medias aritméticas μ_1 y μ_2 pero con igual matriz de covarianzas Σ .

Dado que se trabaja con una muestra de esas poblaciones, a partir de ellas se puede determinar los vectores de medias de la muestra \bar{x}_1 y \bar{x}_2 y la matriz de covarianza estimada, S.

Considerando dos supuestos, que la probabilidad a priori de una observación de pertenecer a una u otra población es la misma y, que los costos de incurrir en un error de asignar un caso a la población 1 cuando pertenece a la población 2, es igual al de asignar un caso de la población 2 cuando pertenece a la 1, se minimiza la probabilidad de equivocarse en asignar un caso a la población CPF si es mayor la

probabilidad (la densidad de probabilidad) de que esa observación pertenezca a esa población que a la otra.

Si $f_1(x)$ y $f_2(x)$ son las correspondientes funciones de densidad para las poblaciones 1 y 2 respectivamente en un espacio k dimensional, el criterio expuesto se puede establecer como:

a: Asignar un caso observado a la población 1, si $\frac{f_1(x)}{f_2(x)} > 1$, y

b: Asignar un caso observado a la población 2, si $\frac{f_1(x)}{f_2(x)} < 1$.

A partir de las muestras, y siguiendo una aproximación lineal sabiendo que son funciones de densidad multinormales de k dimensiones se llega a la siguiente función discriminante:

$$W = \bar{x}' S^{-1} (\bar{x}_1 - \bar{x}_2) - 1/2 (\bar{x}_1 + \bar{x}_2)' S^{-1} (\bar{x}_1 - \bar{x}_2)$$

A partir del cual una nueva observación se asigna a la población 1 toda vez que $W > 0$ y a la población 2 cuando $W < 0$ ³

Puede observarse que W está formado por dos términos. El segundo de ellos que podemos denotar como a_0 , no depende del valor determinado para un caso particular de x observado, el mismo está compuesto por la varianza común S y los vectores de medias \bar{x}_1 y \bar{x}_2 .

El primer término constituye una combinación de tipo lineal del vector x , la que puede exponerse como:

$$a_1x_1 + a_2x_2 + a_3x_3 + \dots + a_kx_k$$

³ Una visión amplia sobre análisis discriminante multivariante puede encontrarse en **Morrison** (1976) y comentarios sobre sus limitaciones en **Eisenbeis** (1977). El clásico en este tema es el libro de **Anderson T.W.** *An Introduction to Multivariate Statistical Analysis*, (1958), J. Wiley & Sons.

La función discriminante incluyendo el término constante puede establecerse como:

$$W = a_0 + a_1x_1 + a_2x_2 + a_3x_3 + \dots + a_Kx_K$$

donde:

$W =$ es el valor que se obtiene para un caso determinado de la función lineal discriminante multivariante.

$a_0 =$ término constante.

$a_1 \dots a_K =$ coeficientes discriminantes o ponderaciones.

$x_1 \dots x_K =$ variables explicativas.

El modelo que se presenta en este trabajo determina el valor del término constante a_0 y de los coeficientes discriminantes $a_1 \dots a_K$, los que son utilizados ante un caso específico, multiplicando estos últimos por cada una de las variables explicativas observadas en el caso. Sumados estos productos junto con a_0 , se obtiene un valor W , que es utilizado para clasificar ese caso en uno de los grupos establecidos a priori. La regla de clasificación, es asignar a la población 1 si $W > 0$ y a la población 2 si $W < 0$.

6. REDES NEURONALES ARTIFICIALES.

BREVE INTRODUCCION

Las Redes Neuronales Artificiales (RNA) que son parte de la Inteligencia Artificial (IA), comienzan su desarrollo a partir de los '40, en un intento de replicar aspectos del funcionamiento del cerebro humano, en particular su lado izquierdo. Las investigaciones sostienen como la mitad izquierda del cerebro humano dedicada a la lógica, las funciones matemáticas. La otra mitad, la derecha se ocupa de las sensaciones instintivas y emociones. Si bien las personas en general utilizan las dos mitades del cerebro, sin embargo en la mayoría de los casos hay una mitad que prevalece. Los científicos se señalan como gente que usa el lado izquierdo en tanto que los artistas y escritores como personas en las cuales el lado derecho del cerebro como dominante.

El cerebro humano, y los cerebros biológicos en términos más generales, están compuestos por células llamadas *neuronas*. Un cerebro humano puede tener 100.000 millones de neuronas, las que están densamente interconectadas.

En la fig. 1 se aprecia la estructura de una neurona cerebral, que consta de un *campo celular* o soma, del cual parten *las dendritas* que son ramificaciones del soma y a través de las cuales ingresan a aquél la mayor parte de las señales que se reciben. Asimismo, cada neurona tiene un *axon*, que suele conocerse como la fibra nerviosa de la neurona. Estos son responsables de la transmisión de las señales hacia otra neurona. La terminación del axon se efectúa a su vez en muchas ramificaciones en cuyo final se encuentra una terminal nerviosa especializada llamada *botón sináptico*, el que se encuentra contra la membrana de la *dendrita* o a veces contra el propio soma de otra neurona. La *sinapsis*, es donde se contactan ambas neuronas y a través de la cual son transmitidos los impulsos nerviosos de una neurona a otra. Estas estructuras neuronales operan electroquímicamente, esto es que el cerebro opera con impulsos eléctricos que son originados en acciones químicas.

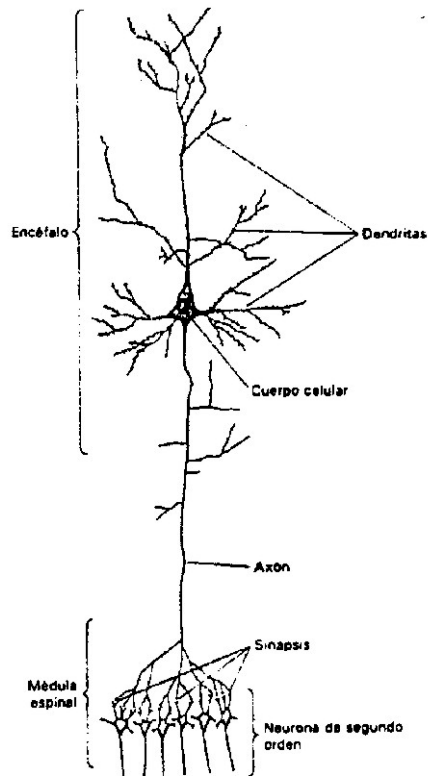


Fig. 1. Estructura de una neurona

Fuente: Guyton, A.C-. "Human Physiology", 1987

Mc Culloch y **Pitts** (1943) desarrollan la primera modelización de una neurona artificial que busca explicar rudimentariamente la neurona humana. El modelo de **Mc Culloch** y **Pitts** se puede representar como:

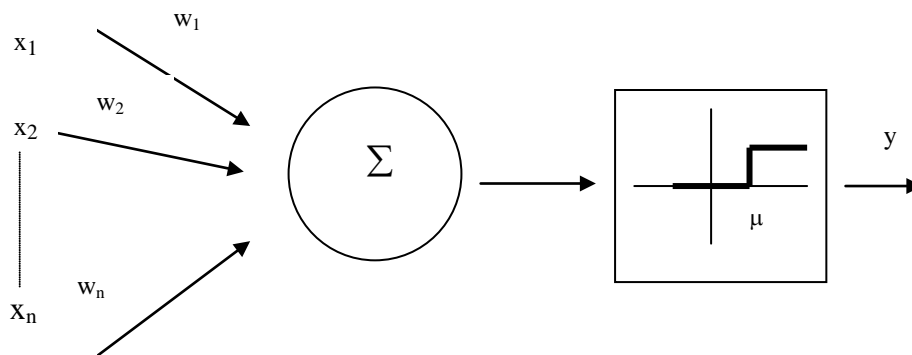


Fig. 2. Modelo de neurona artificial de McCulloch y Pitts

La neurona artificial tiene un *vector de entrada* $\mathbf{x} = [x_1, x_2, \dots, x_n]$ que son los impulsos remitidos desde la *i*-ma dendrita.

El *vector de pesos* $\mathbf{w} = [w_1, w_2, w_3, \dots, w_n]$ representa en cada aproximación a la neurona biológica la sinapsis, y el cuerpo celular opera a través de una función de transferencia o de activación que representa la actividad del soma.

Las neuronas artificiales computan la suma ponderada de n *inputs*, que en las aplicaciones pueden ser los valores de ratios financieros, datos de comportamiento de los mercados, etc.. El cuerpo celular opera a través de dos funciones, una la suma ponderada de los *inputs* y la otra la función de transferencia o activación.

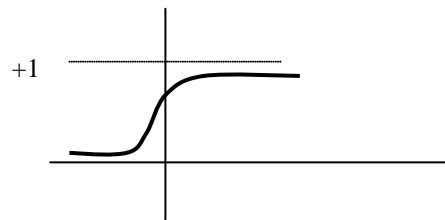
Matemáticamente,

$$y_i = f \left[\sum_{j=1}^n w_j x_j \right]$$

Donde f es la función de activación.

Cada neurona recibe señales a partir de sus vínculos de entrada y calcula un nuevo nivel de activación. El cálculo de este nivel está dividido en dos componentes. Uno de ellos es lineal, la función de entrada, que calcula la suma ponderada de los valores de entrada y un segundo componente no lineal, función de activación o de transferencia, que transforma la suma ponderada en el valor final que sirve como valor de activación de la neurona en cuestión.

El modelo básico de Mc Culloch y Pitts fue extendido en numerosas direcciones, siendo una de ellas la aparición de otras funciones de transferencia o de activación que la *hard limit* (función escalón) o que toma valor que saltan de 0 a 1. Una de estas funciones, de mucha utilización es la *sigmoide*, con sus características de hacer *smooth* en los valores y con propiedades asintóticas (Fig. 3)



FUNCION SIGMOIDE

Fig. 3

La función logística o sigmoide, es una forma de transformar el resultado potencial de una neurona en su output Y que se presentó como y en la fig. 2):

$$Y = \frac{1}{1 + e^{-y}}$$

Obsérvese que $Y(y) > 0$ y que el $\lim_{y \rightarrow -\infty} Y = 0$ y que el $\lim_{y \rightarrow +\infty} Y = 1$

$$y \rightarrow -\infty \qquad y \rightarrow +\infty$$

por lo que representa una función de distribución.

El valor de Y puede ser el valor final del output o un input de otras neuronas. Se advierte que la función de transferencia de este tipo tiene condiciones de responder en forma no lineal a un impulso recibido.

Las RNA, pueden en cierta forma verse como redes gráficas dirigidas en las cuales, las neuronas son interconectadas. Los patrones de conexión de las redes conforman lo que se conoce como *arquitectura* de la red.

Las redes neuronales y su arquitectura varían en complejidad según los casos. Desde algunas muy simples que tiene una sola capa neuronal (la que recibe un *input* y produce un *output*), hasta otras que están estructuradas en varias capas, las que pueden tener variado número de neuronas, desde una a cientos. La primer capa, en estos casos, se conoce como capa de entrada (*input layer*), luego vienen las capas ocultas (*hidden layer*) que pueden ser una o varias con diversas combinaciones de neuronas y, finalmente, aparece la capa de salida (*output layer*). En la figura 4, se observa una red con una capa de entrada con 5 neuronas, una capa oculta con igual número de neuronas y una capa de salida con una neurona.

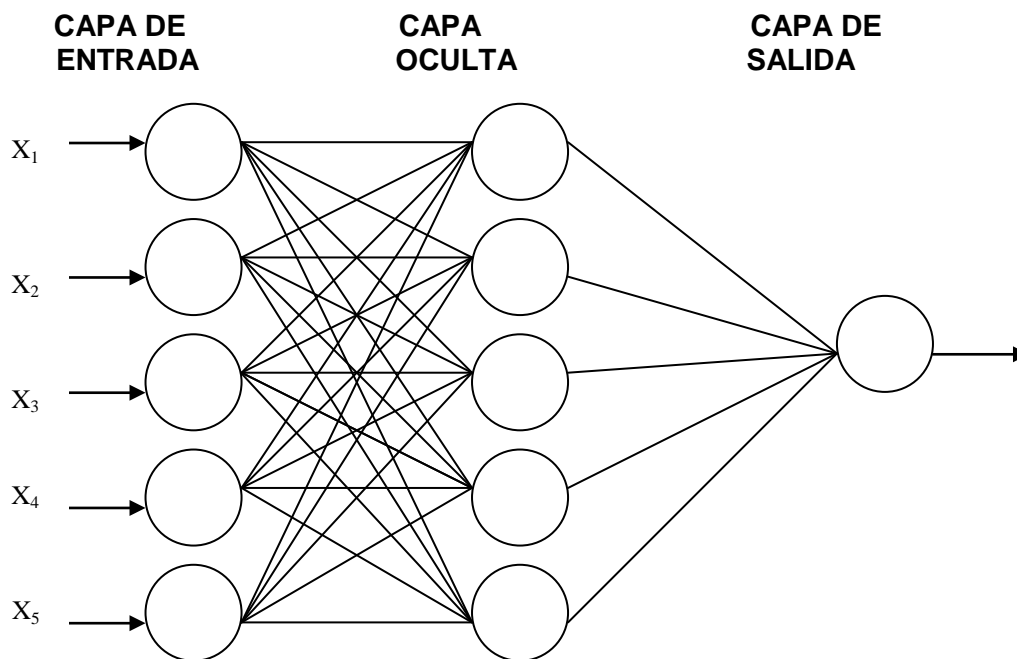


Fig. 4 Red neuronal

La arquitectura de las redes se agrupan en dos grandes tipos: a) *feed forward*, en donde las señales son transmitidas desde una neurona de una capa a la otra, no hay retroalimentación, y b) *feedback* o recurrentes, en las cuales las señales de una capa pueden ir a capas anteriores, esto es pueden tener retrocesos y que son más dinámicos.

Las redes se someten a un proceso de aprendizaje, en el cual varían los pesos y también la arquitectura a efectos de lograr un resultado más eficiente.

Este aprendizaje tiene tres *paradigmas*: el *supervisado*, *no supervisado* e *híbrido*. En el primero, los pesos se van adecuando a la correcta respuesta del *output* a los efectos de llegar a la solución, para lo cual cada input se provee de una respuesta correcta. En el segundo paradigma, no tiene la respuesta correcta asociada a cada input, y va operando tomando y procesando información de los datos adyacentes al problema. El paradigma híbrido, ajusta sus pesos en parte con un aprendizaje supervisado y parte con el no supervisado.

Establecido el paradigma de aprendizaje, las RNA manejan otro aspecto muy importante que son las *reglas de aprendizaje*, que especifican como los pesos ni se modificarán.⁴

Las cuatro reglas de aprendizaje más importantes son: error-corrección, **Hebbian**, **Boltzmann** y aprendizaje competitivo.

Asimismo, la fase de aprendizaje debe seguir un algoritmo, conocido como *algoritmos de aprendizaje* que son los procedimientos que se utilizan para ajustar los pesos a través de la red. Existen numerosos algoritmos de aprendizaje. Uno de los más utilizados, en particular para problemas a que se refiere este trabajo es el algoritmo Back-propagation.

⁴ Un tratamiento en profundidad de RNA está fuera de los objetivos del trabajo. A aquellos efectos puede verse en **Haykin** (1994) y, una amplia visión de la evidencia empírica se encuentra en **Trippi, R. y E. Turban** (1993).

7. EXPERIMENTO

El experimento consta de cuatro fases, que son:

- **Etapa Experimental 1** - Análisis discriminante en el método *stepwise*.
- **Etapa Experimental 2** - Redes neuronales artificiales con todas las variables como inputs en el algoritmo *Back-propagation*.
- **Etapa Experimental 3** -Redes neuronales artificiales con las variables seleccionadas por el ADM en el algoritmo de *Perceptron*.
- **Etapa Experimental 4** -Redes neuronales artificiales con las variables seleccionadas por el ADM en el algoritmo de *Back-propagation*.

En cada una de estas fases se exponen dos aspectos: a) desarrollo y características del modelo aplicado en ella y b) resultados obtenidos.

Etapa Experimental 1 - Análisis Discriminante Multivariante

Desarrollo del modelo

*Función discriminante obtenida*⁵

La selección paso a paso - **Stepwise** - consiste en ir seleccionando los ratios con el propósito de que solo aquellos que satisfagan ciertas condiciones sean incorporados al modelo. Los ratios van entrando o saliendo del modelo según su nivel de aproximación al criterio de selección que se establece.

Para la determinación del ingreso de una variable o la remoción se deberá tener en cuenta la conjunción de los valores de un estadístico así como criterio de entrada o salida. En este caso el estadístico que se tiene en consideración es el lambda de wilks, el que mide las desviaciones de cada grupo respecto a las desviaciones totales sin distinguir grupos, en el espacio generado por las p variables.

La aplicación del **Lambda de Wilks** y del estadístico **F** permiten establecer si las medias de las variables son diferentes entre los dos grupos. El **Lambda de Wilks** (cuando se considera cada variable por separado) se obtiene como cociente de la dispersión dentro de los grupos respecto a la dispersión total.

$$\lambda_{Wilks} = \frac{\text{Suma de Cuadrados Intragrupos}}{\text{Suma de Cuadrados Totales}} = \frac{\sum_{i=1}^{i=2} \sum_{j=1}^{n_i} (x_j - \bar{x}_i)^2}{\sum_{j=1}^{j=79} (x_j - \bar{x}_T)^2}$$

donde:

$n_1 = 41$, $n_2 = 38$, x_j = puntaje de la empresa j en el ratio R , \bar{x}_1 = media del ratio R para las empresas Buenas , \bar{x}_2 = media del ratio R para las empresas Malas y \bar{x}_T = media total del ratio R .

⁵ Trabajo de **Pascale** et al (1999)

$\sum_{i=1}^2 \sum_{j=1}^{n_i} (x_j - \bar{x}_i)^2$. *Suma de cuadrados intragrupo*. Indica la suma de los cuadrados de la diferencia entre cada valor y la media de su grupo y expresa la variación de las empresas *dentro* de cada grupo, en torno a la media de su grupo.

$\sum_{j=1}^{79} (x_j - \bar{x}_T)^2$. *Suma de cuadrados totales*. Indica la suma de los cuadrados de las diferencias entre cada valor y la media total y expresa la variación *total* de las 79 empresas en torno a dicha media total.

La aplicación del estadístico **F** es el cociente entre la media cuadrática intergrupo y la media cuadrática intragrupo.

$$F = \frac{\text{Media Cuadrática Intergrupo}}{\text{Media Cuadrática Intragrupo}} = \frac{79 \sum_{i=1}^2 (\bar{x}_i - \bar{x}_T)^2}{\frac{\sum_{i=1}^2 \sum_{j=1}^{n_i} (x_j - \bar{x}_i)^2}{2*(79-1)}}$$

donde:

$79 \sum_{i=1}^2 (\bar{x}_i - \bar{x}_T)^2$. *Suma de cuadrados intergrupo*. Indica la suma de los

cuadrados de las diferencias entre la media de cada grupo y la media total (multiplicada esta suma por 79) y expresa la variación de las empresas de *grupo a grupo* .

Como existen únicamente dos grupos de empresas (Buenas y Malas) la *Suma de cuadrados intergrupo* coincide con la *Media Cuadrática Intergrupo*.

Las restricciones impuestas tienen relación con superar una prueba de tolerancia y una aplicación del estadístico **F** de **Fisher - Snedecor**. De este modo, se garantiza que los ratios incluidos en el modelo aportan al menos un nivel de información mínimo.

La prueba de tolerancia se realiza para detectar ratios que son combinación lineal de ratios previamente incorporados al modelo.

La **F de ingreso** se usa como prueba de significación para evaluar la relevancia de la información aportada por el nuevo ratio considerando los ratios incluidos en el modelo. Es un estadístico que evalúa la capacidad discriminatoria introducida por el nuevo ratio.

La **F de salida** es un estadístico multivariable que evalúa la significación del decremento en la discriminación si una variable se retira del modelo. Se realiza en cada paso para comprobar la calidad de los ratios incorporados.

En este estudio se utilizó un valor **F** mínimo como criterio para que los ratios entren o salgan del modelo de 3,84 y 2,71 respectivamente.

Se obtienen valores altos de **F** cuando la diferencia de medias es significativa en los grupos. El nivel de significación asociado a **F** permite concluir que la media de las empresas Buenas y la media de las empresas Malas en el ratio "Ctas. a cobrar + Inventarios / Activos Totales" no son significativamente diferentes. El valor de significación que figura en la columna con el encabezado Sig. es superior al valor 0,05. La tabla siguiente presenta las pruebas de significación estadística mencionadas.

CUADRO N° 3

Test de igualdad de las medias grupales

Ratio	Wilks' Lambda	F	Sig.
Patrimonio Neto / Deudas Totales	0,874	11,052	0,001
Ventas / Ctas. a cobrar	0,896	8,917	0,004
Margen de ventas netas	0,843	14,387	0
Ganancia neta / total de deudas	0,669	38,051	0
Ctas. a cobrar + Inventarios / Activos Totales	0,988	0,975	0,327
Capital de trabajo / activos totales	0,813	17,737	0
GAll / activos totales	0,801	19,116	0
Ventas / deudas	0,589	53,702	0
Ganancia neta / activo total	0,676	36,878	0

El estadístico **Lambda de Wilks**, de utilización en el Modelo de Análisis Discriminante Multivariante, permite *seleccionar* variables en base a su poder discriminante entre los grupos analizados. Es claro entonces que tanto más pequeña sea el Lambda de Wilks, la variabilidad total reside entre los grupos. En estos casos es advertible una mejor discriminación de los grupos.

Las variables candidatas, luego de la selección efectuada por el estadístico Lambda de Wilks, deben someterse a un criterio de *eliminación*. Este se efectúa a través del estadístico **F**. El mismo permitiría llegar a eliminar aquella variable que su eliminación provoque el incremento mínimo del Lambda de Wilks.

El modelo que se obtuvo después de aplicar este método es el siguiente:

$$Z_{pp} = -1,76 + 0,45 R8 + 7,73 R9$$

El valor de corte es $Z_{pp} = 0$ y la zona de ignorancia opera entre -0.4097 y $+0.4123$.

Los coeficientes de la función discriminante no están estandarizados. Debido a esto, el valor de cada coeficiente depende de las unidades de medida y no son comparables para determinar el ratio que más influye en la discriminación. En cambio, los coeficientes estandarizados permiten determinar el ratio que más influye en la

discriminación. Los coeficientes estandarizados de los ratios "Ventas / Deudas" (R8) y "Ganancia neta / Activo total" (R9) son 0,722 y 0,507 respectivamente. El mayor valor corresponde a "Ventas / Deudas" y por lo tanto es el ratio de mayor poder discriminante.

Resultados

*Prueba en Validación cruzada*⁶

El resumen de la clasificación de las empresas realizada mediante el método Stepwise se presenta en la tabla siguiente:

⁶ La misma consiste en realizar la regla de clasificación con n-1 observaciones y clasificar la enésima observación, repitiendo el proceso hasta que clasificó las n observaciones.

CUADRO N° 4

Resultados de la clasificación

		Grupo al que pertenece	Grupo asignado por la Función Discriminante		Total
			1	2	
Original	Número de casos	1	36	5	41
		2	3	35	38
	Porcentaje	1	87,80%	12,20%	100%
		2	7,89%	92,11%	100%
Validación cruzada	Número de casos	1	36	5	41
		2	3	35	38
	Porcentaje	1	87,80%	12,20%	100%
		2	7,89%	92,11%	100%

Puede observarse en la tabla anterior que la tasa de error por este método es de 10,1% ($(3+5) / 79$).

En el gráfico siguiente (Fig. 5) se representan las 79 empresas en el plano formado por los ratios "Ventas / Deudas" y "Ganancia neta / Activo total" y la función discriminante determinada.

FUNCION DISCRIMINANTE

EMPRESAS USADAS EN EL MODELO

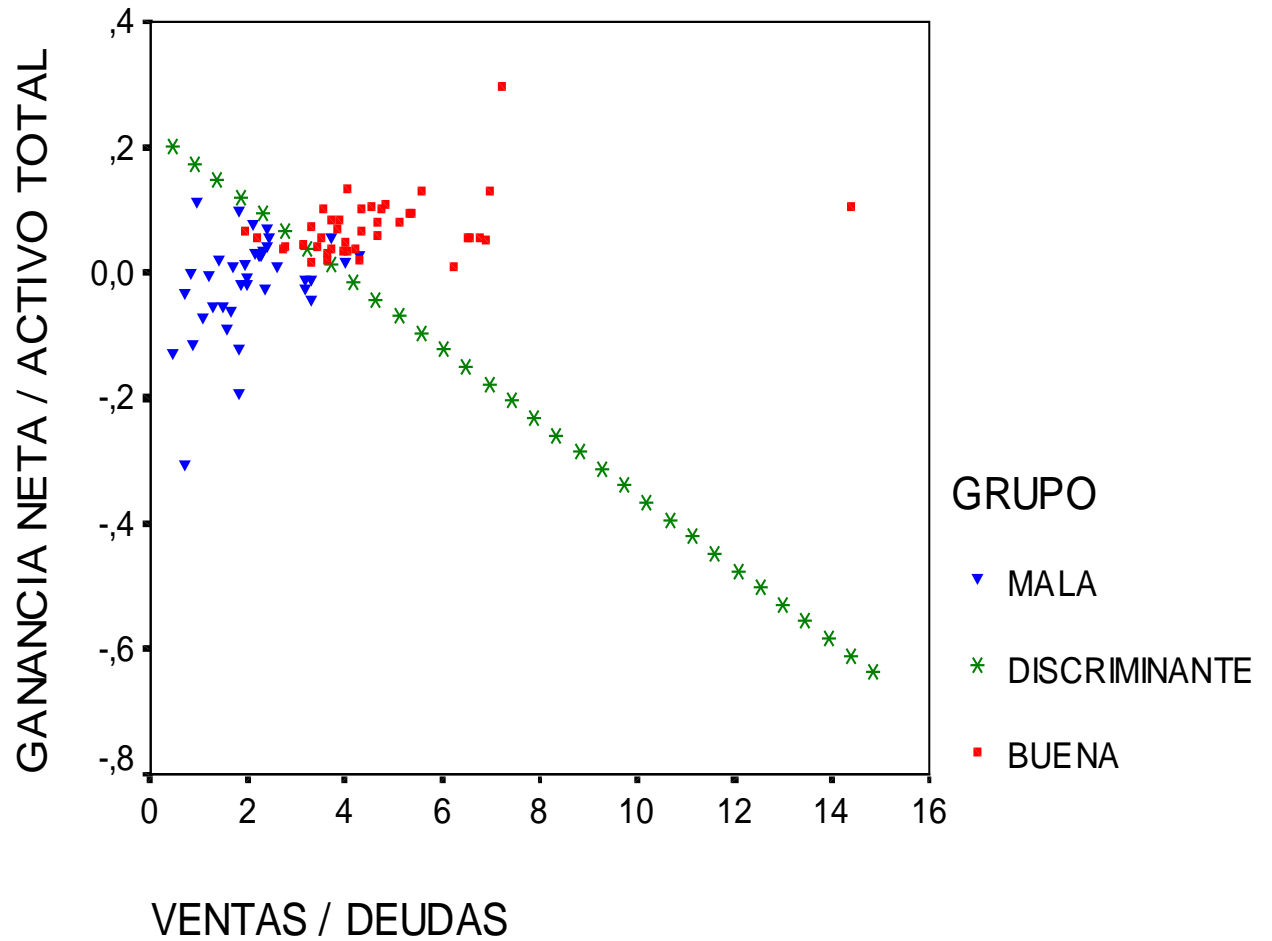


Fig. 5

En la figura anterior se observa que hay zonas confiables para clasificar y un área menos confiable (confusa).

El área confusa es la zona donde empresas "Buenas" se ubican en la zona de las empresas "Malas" y viceversa. Si proyectamos cada una de las empresas en un eje ortogonal a la línea discriminante, se observa con más detalle el área donde la discriminación no es correcta. La zona que corresponde a la superposición de empresas "Buenas" y "Malas" da lugar al área confusa, delimitada por los puntajes discriminantes $-0,3875$ y $0,3831$. El gráfico siguiente (Fig. 6) muestra los puntajes discriminantes de las empresas, donde la zona confusa se aprecia claramente.

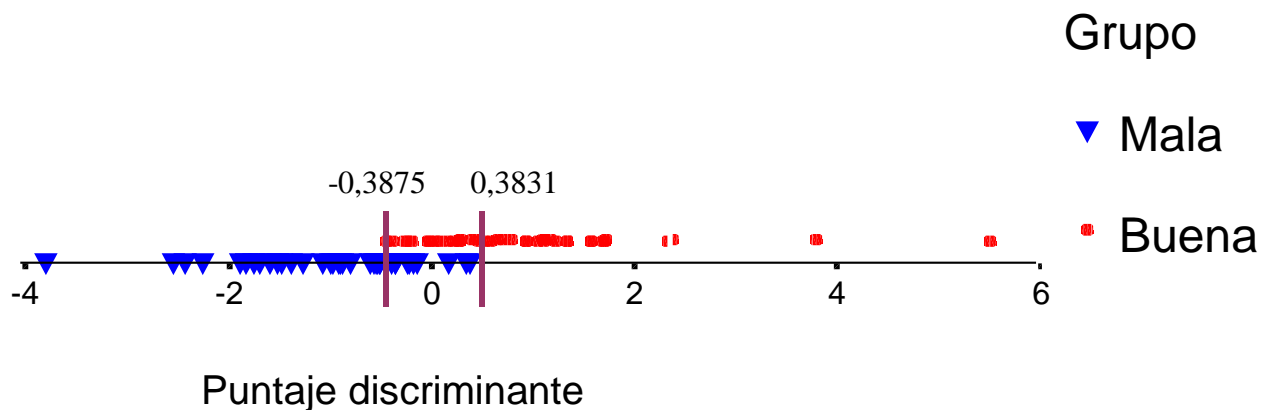


Fig. 6

Testing en Tabulación cruzada

El método Stepwise es el que posee menor tasa de error dentro de las funciones discriminantes lineales. Por este motivo, y debido a su simplicidad, se utilizó la función discriminante resultante de la aplicación del mismo para clasificar a un grupo de empresas que no intervinieron en la construcción de la función discriminante. Como se conocía el grupo al que pertenece cada empresa, se utilizó a éstas como grupo de testeo.

El procedimiento de testeo consiste en comparar el grupo asignado por la función discriminante con el grupo al que realmente pertenece la empresa.

La tabla siguiente resume la información de esta clasificación.

CUADRO N° 5

Empresas de Testeo

Grupo al que pertenece	Grupo asignado por la Función Discriminante		Total
	Buena	Mala	
Buena	28	2	30
Mala	4	30	34
Total	27	37	64

La tasa de error es de tan sólo 9,4 % ($[4+2] / 64$).

El gráfico siguiente muestra las 64 empresas ubicadas en el plano que conforman las variables discriminantes ("Ventas / Deudas" y "Ganancia neta / Activo total").

FUNCIÓN DISCRIMINANTE

EMPRESAS USADAS EN EL TESTING

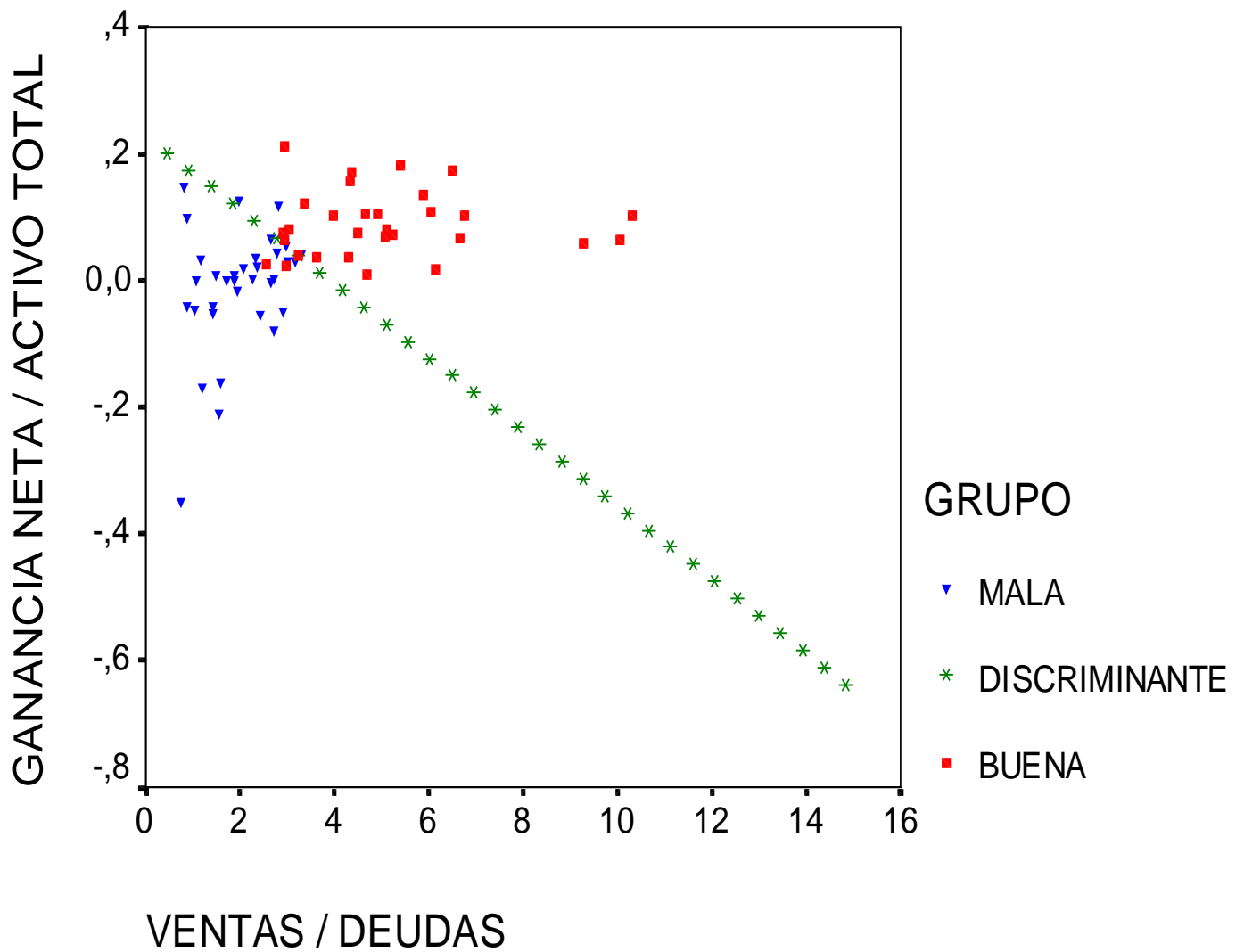


Fig. 7

Al igual que para las empresas usadas para construir el modelo, se observa en el diagrama anterior un área confusa. Esta área corresponde a empresas Buenas que están ubicadas en la zona de empresas Malas (área inferior con respecto a la línea discriminante) y viceversa. Los extremos inferior y superior de la zona confusa son - 0,4097 y 0,4123 respectivamente.

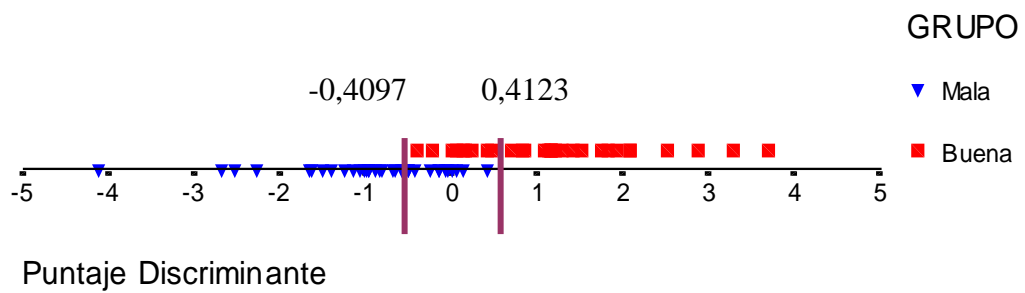


Fig. 8

Etapa Experimental 2 - Redes Neuronales artificiales utilizando 9 variables y algoritmo de aprendizaje Back-propagation

Desarrollo

Aprendizaje

La arquitectura de la red fue, en esta etapa del experimento, de multicapas, del tipo *feed-forward, full interconnected*, contando con una capa de entrada de 9 neuronas, dos capas interiores en las cuales el número de neuronas varió entre 5 y 15 y una capa de salida con una sola neurona que representó una clasificación binaria con 1 para las empresas "buenas" y 0 para las empresas "malas".

La etapa contó con una fase de *entrenamiento* y otra de *testing*. Para la fase de entrenamiento se utilizó como *paradigma de aprendizaje*, el supervisado como *regla de aprendizaje*, (aquellas que gobiernan la actualización de los pesos), el error-corrección en su variante Levenberg Marquandt. Como *algoritmo de aprendizaje*, (procedimientos con los cuales las reglas de aprendizaje son utilizadas para ajustar los pesos) se aplicó Back-propagation. La función de transferencia o de activación utilizada fue la Sigmoide en la cual los valores de la función oscilan entre 0 y 1, para lo cual los datos de entrada (los ratios seleccionados), han debido ser normalizados. A estos efectos se utilizó la siguiente fórmula:

$$Y = \frac{(X - X_{\min})}{(X_{\max} - X_{\min})}$$

Se normalizaron, para el caso de la etapa 2, 9 ratios en las 77 empresas utilizadas en la fase de aprendizaje.

El cuadro que sigue sintetiza algunos de los parámetros utilizados (en Anexo N° 2 se exponen en mayor detalle).

CUADRO N° 6

Parámetros del Modelo de Redes Neuronales	
Umbral Superior	0-7
Umbral Inferior	0.3
Error SSE	0.01
Tasa de aprendizaje	0.01 - 0.001

Cada configuración se entrenó en la fase de aprendizaje con un máximo de 5000 *epochs*.

Testing

Luego de entrenada la red, se efectuó el testing en las 63 empresas con que se contó para esta fase, se analizaron 66 casos con diferentes configuraciones, con pesos que permanecieron constantes de los obtenidos en la fase de aprendizaje.

Cada caso, implicó ingresarle las 9 variables correspondiente a una empresa y luego del paso por la red, se obtuvo un valor de salida entre 0 y 1, utilizándose para considerarla buena o mala los umbrales que se usaron en la fase de aprendizaje.

Resultados

La figura 9, muestra la curva de error típica en las configuraciones obtenidas en la fase de aprendizaje, con descensos en el error promedio a medida que se aumentan las iteraciones. Las distintas curvas que se obtienen a partir de cada configuración no tienen variaciones sustanciales, debiéndose por tanto, dilatar la presentación de una configuración con atributos de generalización, a la fase de testing.

Curva de Error Tipica con 9 variables

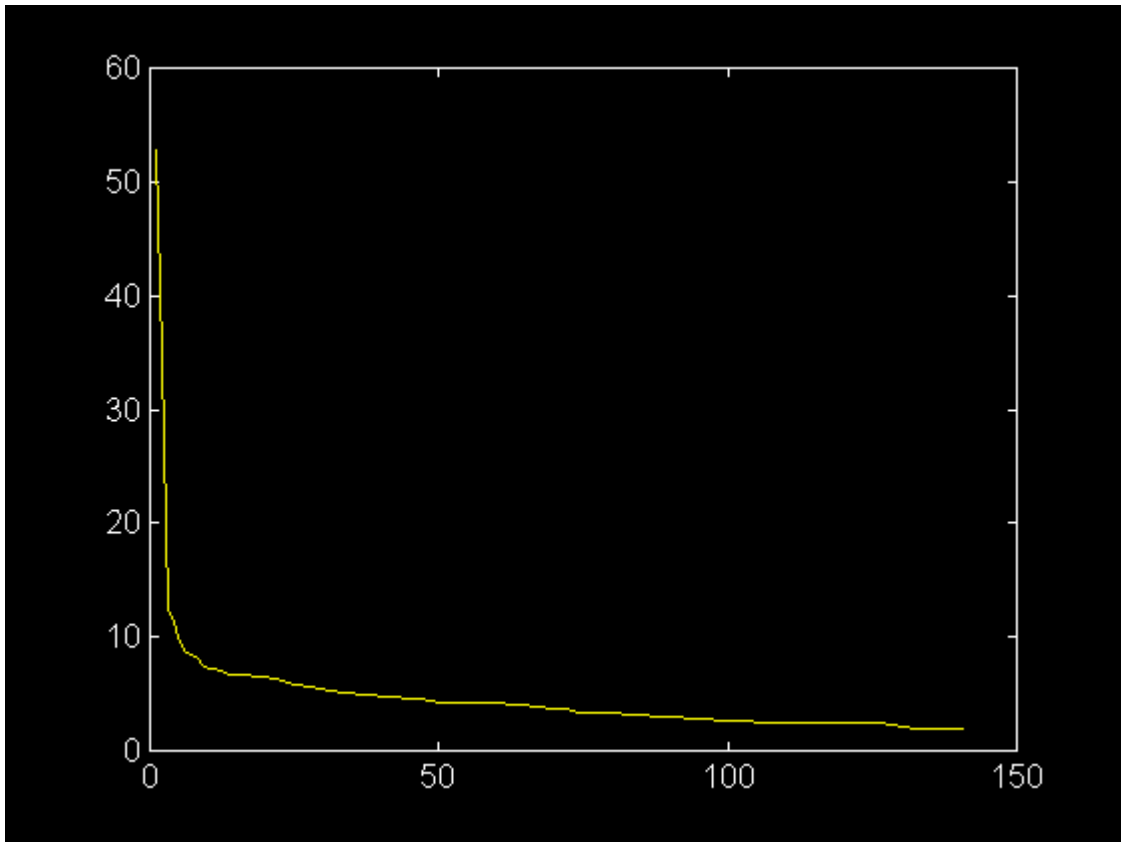


Fig. 9

Los volúmenes obtenidos en la fase de testing, para distintas neuronas en la primer y segunda capa oculta, y el porcentaje de acierto de cada configuración, se muestra en las Fig. 10.

Superficies obtenidas en la fase de Testing con 9 variables

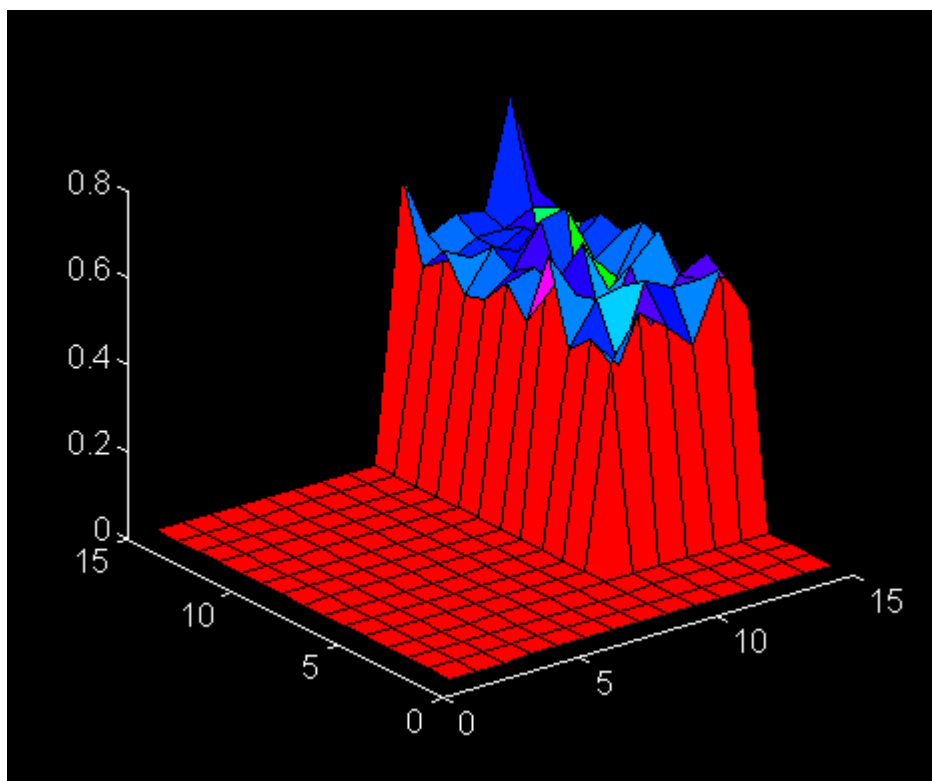


Fig. 10

Los porcentajes de configuraciones con aciertos se sintetizan seguidamente:

CUADRO N° 7

	100%	98 a 99%	< 98%
Aprendizaje	0(0%)	7 (11.6)	59 (88.4)
Testing	0(0%)	0 (-)	66 (100%)

Por otra parte, en la fase de testing el 100% de las configuraciones muestran un porcentaje de acierto menor al 98%, de las cuales a su vez en su totalidad (100%) están por debajo del 60% de acierto.

Etapa Experimental 3 - Redes neuronales artificiales utilizando 2 variables y algoritmo de aprendizaje Perceptron

Desarrollo

Aprendizaje

Los *inputs* en la etapa 2 del experimento fueron las 9 variables seleccionadas utilizando el algoritmo de aprendizaje Back-propagation. Las etapas 3 y 4 del experimento solamente van a tomar como variables de entrada las dos seleccionadas por el modelo de ADM en su versión *stepwise*, estos son los ratios Ventas/Deudas y Ganancia Neta/Activos Totales.

Esta etapa 3 utilizará como algoritmo de aprendizaje Perceptron, en tanto que la etapa 4 aplicará Back-propagation.

El modelo de Perceptron utilizado en esta fase 3, tiene una caja de entrada de 2 neuronas, una capa oculta, con neuronas cuyo número varían entre 5 y 60 y una capa de salida, en una arquitectura de red del tipo *feed-forward, full-interconnected*.

Para la fase de aprendizaje se utilizó como *paradigma de aprendizaje*, el supervisado y como *regla de aprendizaje*, el error-corrección y como *algoritmo de aprendizaje*, Perceptron.

La función *de transferencia* o *de activación* tiene un umbral superior de 0,7 e inferior de 0.3 respectivamente para las empresas "buenas" y "malas". El número de *epochs* en la que se entrenó la red fue de 500. Los demás parámetros utilizados en Perceptron se presentan en Anexo N° 3.

Testing

El testing efectuado luego de entrenada la red se aplicó a 64 empresas de aquella fase. Se le ingresaron con este propósito los valores de las dos variables que se habían seleccionado y luego de transitar por la red se clasificó como "buena" o "mala" teniendo en consideración los umbrales definidos.

Resultados

No se logró con el número de *epochs* utilizado una separación lineal que demarcan las dos poblaciones, por lo que no se alcanzó la convergencia deseada.

Los resultados más significativos obtenidos utilizando el modelo Perceptron fueron:

CUADRO N° 8

Configuración	Aciertos Training (79 empresas)	Aciertos Testing (64 empresas)
Perceptron (2,45)	96,2%	87,5%
Perceptron (2,50)	94,94%	85,93%
Perceptron (2,60)	94,94%	81%

Etapa Experimental 4 - Redes neuronales artificiales utilizando 2 variables y algoritmo de aprendizaje Back-propagation

Desarrollo

Aprendizaje

En esta fase 4, el experimento se centra en utilizar como variables de ingreso, las dos que se obtuvieron con mayor poder discriminante en el modelo de ADM y que se usaron en la fase 3 en el modelo Perception. En esta fase 4 se utilizó Back-Propagation, con los mismos parámetros definidos en la fase 2. El resultado obtenido en la etapa de testing se utilizó para un reentrenamiento de la red, incluyendo aquellas empresas que en la fase de testing, en general no eran correctamente clasificadas, provocando modificaciones en los pesos y en los resultados finales.

Testing

Las mismas consideraciones efectuadas para la etapa 2, aplican a este caso aplicándose a una experiencia con análisis de 66 casos, considerando 64 empresas para esta fase.

La curva de error típica obtenida en las distintas configuraciones (con las variaciones correspondientes) se presenta en la Fig. 11 y en la Fig. 12 aparecen los volúmenes, de los aciertos ante variaciones del número de neuronas en la capa oculta.

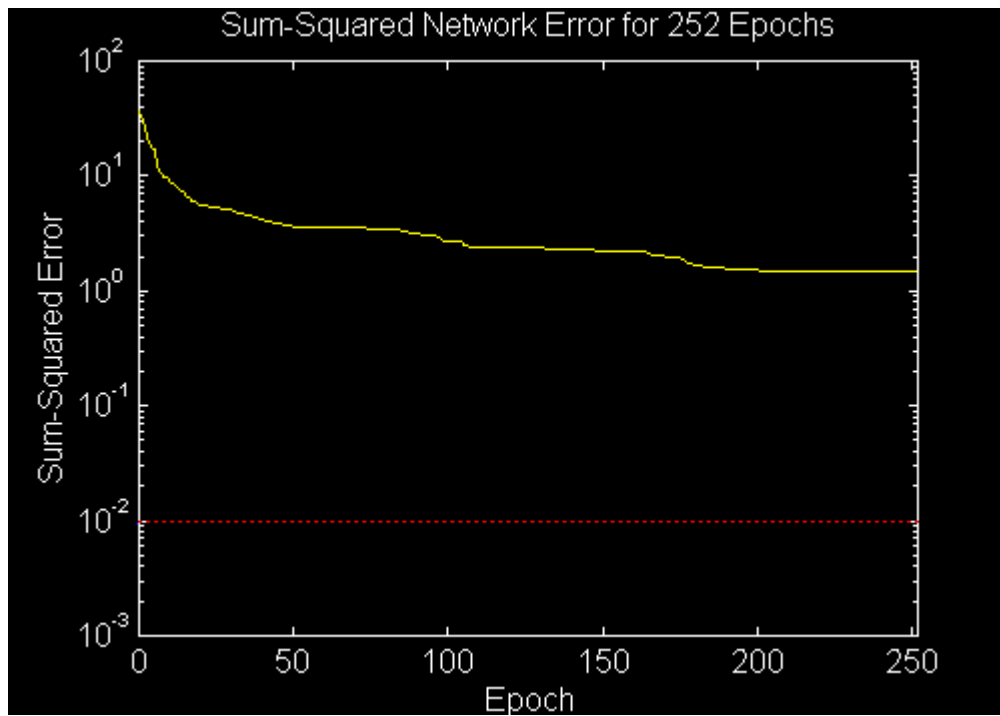


Fig. 11

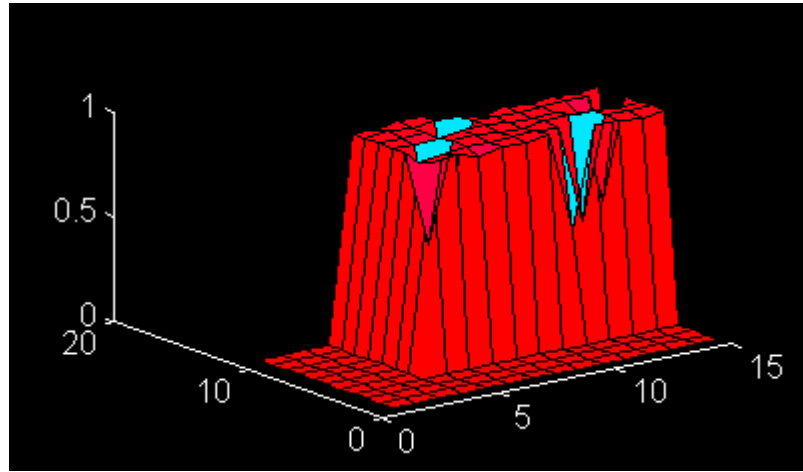


Fig. 12

Para los 66 casos analizados, se obtuvieron los siguientes resultados:

Variación de neuronas en la primer capa oculta	5 a 10
Variación de neuronas en la segunda capa oculta	5 a 15

CUADRO N° 9

	100%	98 a 99%	< 98%
Aprendizaje	28(42.4%)	26 (39.4%)	12(18,2%)
Testing	45(68.2%)	15 (22.7%)	6(9.01%)

Este cuadro representa el porcentaje de aciertos considerando las 66 configuraciones estudiadas. Analizada la fase de testing, parece sugerir que sería posible generalizar mejor con R (2, 6 y 5) (Ver Anexo N° 4)

8. DISCUSION

Algunos hallazgos y conclusiones preliminares que surgen de esta investigación exploratoria son:

- a. En su utilización considerando todas las variables como *inputs*, y con el número de ejemplos del caso analizado las redes neuronales artificiales, en el caso tratado, muestran un comportamiento que refleja una escasa mejoría relativa de acierto en la clasificación entre empresas financieramente "buenas" y "malas", con respecto al análisis discriminante multivariante.
- b. La existencia de dos capas en el modelo mejora la acuciosidad de la clasificación. Estas son conclusiones que surgen de los resultados en el uso de los algoritmos de aprendizaje Back-propagation con dos capas ocultas representa superiores resultados que Perceptron con una sola capa oculta. Una apreciación más contundente a este respecto debería surgir de la utilización de Perceptron con dos capas ocultas, lo que no fue posible en esta investigación.
- c. Cuando se introducen como inputs a las RNA, en el algoritmo Back-Propagation, los hallazgos obtenidos en el modelo de ADM siguiendo el método Stepwise, se potencian sensiblemente los resultados de acierto de las primeras. En este caso se habían hallado en ADM dos variables explicativas con el mayor poder discriminante. Son los ratios Ventas/Deudas Totales y Ganancia Neta/Activo Total. De esta forma, introducidos ellos como variables de entrada, se alcanzan valores muy importantes de acierto en la clasificación con RNA. Esto se aprecia en los resultados obtenidos en la etapa 4 del experimento.
En el caso estudiado la mayor potencia de clasificación se obtiene cuando se utilizan las RNA, tomado como inputs las variables halladas en el modelo de ADM.
- d. La clara potencia de las Redes Neuronales Artificiales en los problemas de clasificación en el caso de empresas ha sido, en general, confirmada por numerosas investigaciones. En todo caso, el arrastre que deviene de la trampa de "*overfitting*",

el tiempo de proceso de las redes en la fase de aprendizaje y la amplia cantidad de casos para utilizar en el testing, así como la observancia de encontrar casos *ilógicos* desde el punto de vista del análisis financiero, que derivan de la propia naturaleza de los modelos matemáticos que están en la base de las redes, operando importantes números de variables en muy numerosos casos de una manera compleja y, la dificultad de una lectura clara de la causalidad de la clasificación, que la ubican como un método "black-box", sugieren la necesidad de recordar estos aspectos con miras a una más balanceada interpretación de los resultados.

- e. Los mejores resultados obtenidos tanto en el Análisis Discriminante Multivariante como en las Redes Neuronales Artificiales, remarcan la importancia de algunas variables en la predicción temprana de serios problemas financieros de las empresas. Estas variables son los ratios: Ventas/Deudas Totales y Ganancia Neta/Activos Totales.

Los hallazgos en esta investigación aplicada a una rama del sector comercial en la República Argentina, (empresas concesionarias de automotores), son conceptualmente coincidentes con las obtenidas en anteriores investigaciones (**Pascale**, 1988), aplicando Análisis Discriminante Multivariante a la clasificación de empresas, con y sin serios problemas financieros en empresas de la industria manufacturera de Uruguay. El trabajo de 1988, reportó tres ratios con mayor poder discriminante ⁷ en la escala que se expone: Ventas/Deudas Totales, Ganancia Neta/Activos Totales y Deuda de Largo Plazo/Deuda Total.

En la presente investigación no aparece el último de los ratios mencionados, Deudas a Largo Plazo/Deudas totales. El mismo no fue considerado para el análisis, debido a que ninguna de las 79 empresas de la fase de aprendizaje ni las 64 de la fase de testing, por la propia naturaleza de la rama en estudio, mantenía deudas de largo plazo.

⁷ El modelo de 1988 es $Z_p = -3,70992 + 0.99418x_1 + 6.5534x_2 + 5.51253x_3$

donde:

x_1 = Ventas/Deudas Totales; x_2 = Ganancia Neta/Total de Activos; x_3 = Deuda Largo plazo/Deuda Total. El modelo tiene un valor de corte de $Z = 0$ y una zona de ignorancia entre -1.05 y 0.4.

Los otros dos, Ventas/Deudas Totales y Ganancia Neta/Activos Totales, con años de distancia en los modelos, con ramas de actividad diferentes y países distintos, siguen mostrando ampliamente su superioridad discriminante frente a las otras variables también incorporadas originalmente como potencialmente explicativas del problema, claro está que con ponderaciones diferentes. El ratio Ventas/Deudas Totales, que surge como el de mayor poder discriminante en el modelo de 1988, y su reaparición en el de 1999 y en las RNA de este trabajo es prácticamente inexistente en los modelos discriminantes de RNA desarrollados en países industrializados para predecir serios problemas financieros en empresas. Sin embargo lo he considerado dentro de la lógica del análisis financiero de los países de América Latina. En América Latina la implementación de distintas políticas encaminadas a obtener una mayor estabilidad monetaria, padecen con amplia generalidad del *síndrome de altas tasas de interés reales positivas*. Es frecuente encontrar tasas bancarias para operaciones de préstamos que oscilan en términos anuales en el 20/30% y muchas veces más altas aún, en dólares, en términos reales. La causalidad de este síndrome, todavía en proceso de clara definición, pivotea sobre varios factores, entre los cuales aparecen: la tasa de riesgo país que ya aporta un "piso" a la tasa de interés doméstica, el poder oligopólico de los bancos que le permite mantener altos sus precios por el crédito, las expectativas de una depreciación significativa de la moneda nacional y los altos costos operativos de los sistemas financieros. La intensidad de cada uno es variable según la situación analizada, pero en todo caso las altas tasas de interés activas en términos reales se aprecian extendidamente en la mayor parte de los países de América Latina, a los que incuestionablemente no escapan Argentina y Uruguay.

La potencialidad de este ratio (Ventas/Deudas Totales) en la explicación de la causalidad, deriva en gran medida que es un subrogante cuantitativo primario bruto de la capacidad de la empresa de generar fondos para atender los servicios de deudas. La clasificación se ve reforzada cuando se agrega al mismo el ratio Ganancia Neta/Activos Totales, que daría un subrogante cuantitativo primario bruto de la capacidad de la empresa para sostener el costo del financiamiento en condiciones de incertidumbre, ya sea a través de fondos ajenos como de fondos propios.

- f. La utilidad en el diagnóstico económico financiero de técnicas como las RNA y el ADM, a esta altura de los desarrollos teóricos y de la evidencia empírica, es incuestionable. En todo caso es importante recordar que se trata de técnicas de *paraclínica financiera*.
- g. El algoritmo de diagnóstico financiero (**Pascale**, 1998) incluyendo el de bancarrota, se establece a partir de numerosos indicadores cuantitativos, pero aparece como deseable la incorporación de la *clínica financiera* para el diagnóstico tanto etiológico como nosológico. La clínica financiera frecuentemente no está asociada a cifras o a la historia de una firma, puede tener que ver con aspectos más cualitativos, que frecuentemente definen el diagnóstico, entre los cuales se pueden señalar para sólo ejemplificar su importancia, la competitividad del sector en un proceso creciente de integración, aparición de firmas de gran potencia competitiva en el mercado, la existencia de litigios, o de una administración basada en una aproximación "one man show".

A la hora del diagnóstico financiero deben balancearse tanto la clínica como la paraclínica financiera. A esta última, el aporte de técnicas como las RNA y el ADM, es indudable. Empero no son el único aporte necesario para un adecuado diagnóstico.

9. CONCLUSIONES

Esta investigación exploratoria se centró en la aplicación de RNA a la predicción de serios problemas financieros en empresas, buscando integrarlo con técnicas estadísticas más tradicionales utilizadas para resolver el mismo tipo de problemas de clasificación como es el caso del ADM. Las RNA como modelo no parametrizado muestra una gran flexibilidad en la solución de la clasificación.

El caso estudiado a la vez que confirma la potencia de las RNA para resolver problemas de clasificación, ha puesto de manifiesto que la potencia de aquéllas es aún mayor cuando los inputs son las variables relevantes para el ADM.

Conclusiones más terminantes sobre la eficiencia agregada del uso concomitante de técnicas estadísticas y RNA, requiere de un mayor esfuerzo investigativo, que incluyen otras técnicas como clusters y regresión logística para luego ADM y RNA, así como en estas últimas utilizar Perceptron con dos capas ocultas o algoritmos de "daño cerebral óptimo".

En todo caso, más allá del interés que puedan derivarse de los hallazgos, necesariamente preliminares de esta investigación exploratoria, la misma debe interpretarse como un paso en el camino de cubrir los objetivos y el problema planteado.

ANEXO N° 1

ANEXO N° 2

Todas las configuración fueron evaluadas utilizando los siguientes parámetros generales

Rango de Neuronas para Primer capa oculta	5 a 15
Rango de Neuronas para Segunda capa oculta	5 a 15
Total de casos probados	90 a 121 para cada configuración
Tasa de Aprendizaje	entre 0.01 y 0.001
Numero máximo de epochs	5000
Convergencia promedio	entre 500 y 3500 epochs
Frecuencia de display para Curva de error	20
SSE (sum squared error)	<0.01
	$SSE = \sum_{1,n} (T_k - O_k)^2$ n = numero de datos de entrada en la fase de training T _k es el valor deseado de salida O _k es el valo generado por la red
Intervalo de Aceptación buenas	0.7 y 1
Intervalo de Aceptación malas	0 y 0.3
Paradigma de Aprendizaje	Supervisado
Arquitectura de Red	feed-forward, full interconnected
Regla de Aprendizaje	variante Levenberg Marquadt
	Reglas que gobiernan la actualización de pesos
Algoritmo de aprendizaje	Back-propagation
	Como las reglas de aprendizaje son utilizadas para ajustar los pesos
Tiempo de procesamiento de cada caso	entre 2 y 6 minutos

Minimum gradient	0.0001
Valor inicial para MU	0.000001
Multiplier for increasing MU	0.1
Multiplier for decreasing MU	0.1
Maximum value for MU	10

ANEXO N° 3

Todas las configuración de Perceptron fueron evaluadas utilizando los siguientes parámetros generales

Variables de Entrada	2 a 3
Rango de Neuronas para capa oculta	5 a 60
Total de casos probados	50 para cada configuración
Numero máximo de epochs	500
Convergencia promedio	500 (no se llega a la separación lineal)
Frecuencia de display para Curva de error	10
Intervalo de Aceptación buenas	0.7 y 1
Intervalo de Aceptación malas	0 y 0.3
Paradigma de Aprendizaje	Supervisado
Arquitectura de Red	feed-forward, full interconnected
Tiempo de procesamiento	entre 1 y 2 minutos cada caso
Tiempo de procesamiento de configuraciones	entre 40 y 60 minutos
Error exigido en la fase de Training	menor a 8
Error exigido en la fase de Testing	menor a 8

ANEXO N° 4

Datos de la Red Neuronal Seleccionada RNA(2,6,5)

Numero de Capas
1 de entrada
2 ocultas
1 de salida

Dimensiones

Dim (2,6) entre capa de entrada y primer capa oculta
Dim (6,5) entre primer capa oculta y segunda capa oculta
Dim (5,1) entre segunda capa oculta y capa de salida

Ciclos de entrenamiento

141

Tasa de Aprendizaje

0.001

Detalles de Pesos calculados

W1 =

-46.9528 13.5180
-13.6325 26.1025
6.1615 42.5317
12.6374 -6.8052
-45.3819 45.7800
-25.6859 -19.7166

W2 =

-6.9606 -19.9533 -3.6825 -16.1883 47.1477 -1.7717
1.0593 -0.1725 -6.9912 3.0812 -3.2502 -11.1860
-39.9236 -10.2666 -15.4125 13.7736 -3.8823 5.1667
40.2988 -12.4551 -33.2695 1.7875 -38.2800 20.1644
24.2567 1.4516 5.7861 -5.0809 28.0664 -15.6172

W3 =

-48.8932 -0.7109 26.0559 -48.1993 37.4994

Bias

B1 =

22.5745
0.7001
-14.8076
1.6404
19.7054
25.0822

B2 =

-9.3557
1.8010

13.0689
-3.6718
-1.2806

B3 =

-0.6887

