

Herramientas para el Pronóstico de la Calificación Crediticia
de las
Finanzas Públicas Estatales en México:
Redes Neuronales Artificiales, Modelo Probit Ordenado y Análisis
Discriminante

Categoría de Investigación
Premio Nacional Bolsa Mexicana de Valores

Herramientas para el Pronóstico de la Calificación Crediticia de las Finanzas Públicas Estatales en México: Redes Neuronales Artificiales, Modelo Probit Ordenado y Análisis Discriminante

Resumen Ejecutivo

Una calificación de crédito es una opinión sobre la solidez de las finanzas de una empresa, una emisión bursátil o de las finanzas públicas de un estado o un municipio, que indica a su vez la capacidad y voluntad de pago de las obligaciones financieras contraídas con acreedores. Algunos autores consideran incluso que la calificación de riesgo es una aproximación de la probabilidad de incumplimiento y, en ese sentido, las calificaciones de riesgo son un buen indicador de la salud o deterioro financiero de las entidades federativas y de las finanzas públicas locales en su conjunto.

El proceso de asignación de calificaciones de riesgo de las Finanzas Públicas de entidades federativas en México es relativamente nuevo. La introducción de calificadoras de riesgo desde el 2001 ha permitido, entre otras cosas, mejorar la claridad en el otorgamiento de créditos a Estados de acuerdo con las condiciones financieras particulares de estas entidades. Esto ha evitado, entre otros fenómenos, los problemas de información asimétrica que se hicieron evidentes en los gobiernos locales durante la crisis de 1995 y hasta el año 2000. Ahora, en las condiciones actuales, con una crisis económica y financiera que ya se ha transmitido de forma virulenta a los gobiernos locales, es importantísimo monitorear el desempeño de las finanzas públicas estatales, debido a la posibilidad real de que alguno de estos fenómenos se manifieste poniendo en riesgo el actual sistema financiero y de coordinación fiscal.

También, con las implicaciones evidentes para la regulación, financiamiento y monitoreo de la calidad crediticia de las finanzas públicas sub-soberanas, resulta de utilidad para diversos actores en

este esquema fiscal y financiero pronosticar y clasificar adecuadamente la calificación de crédito de los estados de acuerdo a sus principales determinantes. En esta investigación se comparan de manera concreta tres métodos comúnmente usados en la literatura para el pronóstico de calificaciones crediticias: un algoritmo de Redes Neuronales Artificiales, el Análisis de Discriminantes y un modelo de variable dependiente limitada conocido como Probit Ordenado.

Entre los resultados se encuentra que el algoritmo de Redes Neuronales Artificiales empleado tiene una mejor capacidad de pronóstico puntual para datos dentro y fuera de la muestra. Dentro de la muestra se alcanza un pronóstico exacto de las calificaciones crediticias asignadas a los estados, mientras que la capacidad de pronóstico disminuye sustancialmente cuando se consideran calificaciones fuera de la muestra.

También se observa que el Análisis Discriminante tiene un mejor pronóstico por intervalos que las Redes Neuronales o los modelos Probit. En términos generales sin embargo se muestra que las redes neuronales proveen una mejor estimación puntual, mientras que el Análisis Discriminante genera una mejor estimación por intervalos y para datos dentro de la muestra. El modelo Probit Ordenado no logra clasificar de manera satisfactoria la calidad crediticia de los estados en México. En otras palabras, si el analista busca una estimación de la calificación crediticia en intervalos de dos o tres *notches* el Análisis Discriminante supera sustancialmente al algoritmo de Redes neuronales aquí propuesto. Si el analista busca estimar la calificación crediticia de manera puntual, la Red Neuronal Artificial es sin duda la mejor herramienta para hacerlo.

En el entorno actual de crisis financiera de los gobiernos locales, este trabajo de investigación arroja información muy importante acerca de cómo el regulador, el secretario de finanzas estatal, el acreedor o los órganos fiscales, pueden pronosticar la calidad crediticia de los gobiernos locales con métodos eficientes. Aún más,

en este estudio se sugiere cuál método de estimación es el más apropiado dependiendo del objetivo del analista. Entre otros, este pronóstico puede usarse por el funcionario local para implementar acciones que ayuden a mejorar la calificación crediticia de sus finanzas públicas; por el supervisor/regulador para detectar casos de deterioro financiero significativo que requieran acción inmediata; y por el acreedor para estimar las tasas de interés de los financiamientos que otorga a los gobiernos estatales y para estimar las reservas de capital que requiere. Son estas razones de índole práctico apremiante las que justifican que esta investigación sea la ganadora del premio de la Bolsa Mexicana de Valores.

Este trabajo se divide en cinco secciones. La primera sección introduce el trabajo de investigación en términos concretos mientras que en la sección II se hace una revisión breve de la literatura sobre los métodos empleados comúnmente para el pronóstico de calificaciones crediticias, así como sus aplicaciones en distintos campos de la economía y las finanzas. La sección III describe los métodos empleados en este estudio con especial énfasis en modelos Probit Ordenados y el algoritmo de Redes Neuronales. La sección IV muestra las estimaciones, la clasificación de las calificaciones crediticias y la capacidad de pronóstico usando una evaluación con datos dentro y fuera de la muestra. La sección V muestra las conclusiones sobre la capacidad de pronóstico de los tres métodos empleados y las implicaciones para la evaluación de las finanzas públicas estatales en México.

Contenido

I. Introducción.....	6
II. Revisión de Literatura.....	8
III. Redes Neuronales, Análisis Discriminante y Modelos Probit....	12
3.1 Red Neuronal.....	12
3.1.1 Estrategia de Entrenamiento.....	15
3.1.2 Algoritmo de Asignación de Calificación.....	17
3.2 Modelo Probit Ordenado.....	18
IV. Evaluación de los Métodos de Pronóstico de Calificación.....	21
4.1 Análisis de Datos y Definición de Variables.....	21
4.2 Implementación de la Red Neuronal.....	24
4.2.1 Clasificación vía Red Neuronal.....	25
4.3 Implementación del Modelo Probit Ordenado.....	26
4.4 Análisis Discriminante Múltiple.....	28
4.5 Bondad de Ajuste y Capacidad de Pronóstico.....	29
V. Conclusiones.....	32
 Bibliografía.....	 34
 Anexo.....	 36

I. Introducción

Las calificaciones crediticias se han constituido ya como un elemento más dentro del esquema fiscal y financiero de las entidades federativas en México. Desde su introducción en el año 2001 éstas son un requerimiento opcional para todo aquel gobierno estatal o municipal que busque financiamiento a través de la banca comercial en mejores condiciones o a través de la emisión bursátil de deuda.

Una calificación de crédito es una opinión sobre la solidez de las finanzas públicas del estado, que puede indicar a su vez la capacidad y voluntad de pago de las obligaciones financieras contraídas con la banca o bolsa de valores. Algunos autores consideran incluso que la calificación de riesgo es una aproximación de la probabilidad de incumplimiento y, en ese sentido, son un buen indicador de la salud financiera de las entidades federativas y de las finanzas públicas locales en su conjunto.¹

Las instituciones bancarias y acreedores por su parte toman como referencia a las calificaciones de crédito y su posición respecto a la calificación del país para calcular el monto reservas de capital para hacer frente al riesgo de incumplimiento. De esta manera, mientras mayor sea la brecha entre la calificación de la entidad federativa y la del país, mayor será el requerimiento de reservas de capital y, por tanto, mayor la tasa de interés que el estado tendría que pagar por los créditos que contrate.

Los órganos reguladores también están interesados en monitorear el desempeño de las finanzas públicas de las entidades federativas con el fin de detectar problemas de insolvencia e incumplimiento potencial que pueda poner en riesgo a todo el sistema financiero, sobre todo en condiciones de crisis como las que atraviesa

¹ Ver por ejemplo Rosch (2004) o a Stephen, et. al. (2008) quienes consideran que la calificación de riesgo y la probabilidad de incumplimiento se asocian fuertemente.

México ahora y desde mediados de 2008.² No obstante, el diagnóstico hecho al momento por diversos analistas—incluida la Secretaría de Hacienda y Crédito Público (SHCP)—es de corte histórico y descriptivo, no se hace un ejercicio de prospectiva o pruebas de resistencia (*stress testing*) sobre la condición crediticia de estas entidades.

Debido a la importancia que reviste para el sistema financiero y para el propio desempeño de las finanzas públicas estatales, es de interés contar con herramientas que permitan estimar la calificación crediticia con el fin de aproximar su probabilidad de incumplimiento y tener una idea clara de su desempeño, incluso antes de que pueda asignarse formalmente una baja en su calificación crediticia.

En la literatura financiera se han usado distintos métodos de clasificación para estimar la calificación crediticia de emisores bancarios y corporativos, tales como el Análisis Discriminante Múltiple y modelos de variable dependiente limitada (entre los que destacan los modelos Probit Ordenados y Logísticos). Sin embargo muy poco se ha hecho respecto al estudio de gobiernos locales o sub-soberanos y la literatura formal al respecto es prácticamente inexistente. Mucho menos se han implementado metodologías más sofisticadas para el estudio de las finanzas públicas locales como es el caso de algoritmos de redes neuronales.

Usando las calificaciones de crédito y datos sobre finanzas públicas de los estados en México desde el 2001, disponibles al público en la página de FitchRatings, en este artículo aplicamos tres métodos para clasificar y pronosticar la calificación de riesgo crediticio de los estados en México. Éste es de hecho el primer intento formal de realizar un pronóstico de las calificaciones de crédito de las entidades federativas en México usando tres métodos de comparación: Análisis Discriminante, Probit Ordenado y algoritmos de

² La Secretaría de Hacienda y Crédito Público (SHCP) elabora desde el año 2003 un diagnóstico anual sobre la situación de las haciendas públicas estatales y municipales en México que sin embargo no provee un análisis a futuro de las condiciones crediticias de estas entidades.

redes neuronales. De hecho es quizá el primer estudio formal acerca de la evaluación de métodos alternativos de pronóstico de las finanzas públicas de gobiernos locales para países emergentes.

La sección siguiente hace una revisión breve de la literatura sobre los métodos empleados comúnmente para el pronóstico de calificaciones crediticias y sus aplicaciones en distintos campos de la economía y las finanzas. La sección III describe los métodos empleados en este estudio con especial énfasis en modelos Probit Ordenados y el algoritmo de Redes Neuronales. La sección IV muestra las estimaciones, la clasificación de las calificaciones crediticias y la capacidad de pronóstico con una evaluación con datos dentro y fuera de la muestra. La sección V muestra las conclusiones sobre la capacidad de pronóstico de los tres métodos empleados y las implicaciones para la evaluación de las finanzas públicas estatales en México.

II. Revisión de Literatura

Diversos algoritmos de redes neuronales han sido aplicados a las ciencias sociales, la economía y los negocios con cierto éxito. Revisiones de literatura extensas³ ya han documentado la utilidad de las Redes Neuronales Artificiales (RNA) para el modelaje del mercado cambiario, mercados de capitales, evaluación de créditos, inversiones, problemas macroeconómicos y pronósticos de bancarrota y calidad crediticia. Se ha encontrado que estos algoritmos tienden a ofrecer en general mejores resultados que otras técnicas estadísticas o matemáticas usadas comúnmente en estas disciplinas. Entre las técnicas alternativas empleadas se encuentran el Análisis Discriminante Múltiple (ADM), modelos de variables categóricas o

³ Ver Wood y Bhaskar (1995) y Wong et. al. (1995) para aplicaciones en las finanzas, negocios y operaciones. Los primeros autores proveen cerca de 100 referencias con aplicaciones de redes neuronales en estudios publicados antes de 1995. La literatura desde entonces se ha multiplicado significativamente.

variable dependiente limitada (Probit y Logit), algoritmos genéticos, programación lineal (PL), entre muchos otros.

Por su parte el estudio de calificaciones crediticias se ha realizado desde distintos ángulos, desde el ADM hasta modelos Probit y Logit Ordenados y más recientemente, con un éxito notable, se han aplicado algoritmos de Redes Neuronales Artificiales (RNA). En esta línea de investigación para el pronóstico de calificaciones de crédito soberanas, por ejemplo Benewell, et. al. (2006) desarrollaron dos algoritmos de RNA y los comparan con modelos Probit Ordenados. Estos autores encuentran que los modelos de RNA dan resultados de pronóstico y precisión superiores. Por su parte, al estudiar la probabilidad de incumplimiento de asociaciones de ahorro y crédito, Salchenberger et. al. (1992) encuentran que las RNA presentan un mejor desempeño que los modelos logit ordenados. Benell, et. al. (2005) examinan a su vez las calificaciones crediticias de gobiernos soberanos para el periodo 1989-1999, comparan a las redes neuronales contra modelos probit ordenados y concluyen finalmente que redes neuronales artificiales arrojan mejores pronósticos.

Respecto a los estudios que emplean el Análisis Discriminante Múltiple se encuentra Lee (2007) quien investiga las calificaciones de crédito corporativo usando también redes neuronales de retro-propagación y concluye que este método arroja mejores resultados que el ADM. Kumar y Bhattacharya (2006) por su parte comparan redes neuronales con ADM para el pronóstico de calificaciones de crédito corporativo. Estos autores también encuentran que las redes neuronales superan el desempeño de los modelos de ADM. Además, Patuwo (1993) concluye que el tamaño de la muestra (en contraste con la arquitectura de la red) ayuda a maximizar la tasa de clasificación y concluye que las redes neuronales arrojan resultados comparables o mejores que los modelos de ADM. Al comparar algoritmos de RNA con ADM, algoritmos genéticos y regresión logística en problemas de clasificación crediticia, Desai (1997)

encuentra que las RNA tienen un desempeño superior a estos métodos de clasificación. En contraste, al investigar el colapso de mercados accionarios y bancarios usando redes neuronales y algoritmos genéticos, Tan y Quek (2007) concluyen que los algoritmos genéticos proveen un mejor pronóstico con una implementación más ágil.

Dada la preferencia evidente por los métodos de RNA y ADM, autores como Markham (1995) han combinado ambas técnicas para generar pronósticos compuestos que pudieran ser más certeros que los obtenidos por las técnicas individuales. Otros autores han propuesto integrar redes neuronales, simulación computacional y optimización con resultados que superan el desempeño individual de estas técnicas—ver Ting Peng (1992). Finalmente técnicas de extracción formales basadas en programación genética han sido exploradas por Tokinaga (2005) donde se obtienen redes neuronales para clasificaciones binarias. En esta aplicación se estudió por ejemplo la posibilidad de incumplimiento del sector industrial Japonés para el periodo 1970-1986. Al igual que en otros estudios se encuentra que las redes neuronales arrojan mucho mejores ajustes y pronósticos que ejercicios con análisis discriminantes multivariados.

Otras aplicaciones de las RNA pueden encontrarse en trabajos como los de Kotsiantis, et. al. (2006) quienes investigan declaraciones financieras fraudulentas y pronósticos de bancarrota corporativa usando redes neuronales. Bharat y Barin (1995) por su parte usan RNA para modelar el precio de mercado de Oferta Públicas Primarias (IPOs) asociadas a un vector de variables económicas y concluyen que la sub-valoración comúnmente observada en IPOs es reducida en cerca de 8% usando una RNA.

En cuanto al uso de redes neuronales con el fin de clasificar observaciones de acuerdo a su calidad crediticia, Gutierrez y Serrano (2007) usan diversos indicadores financieros y sociales para pronosticar el riesgo de instituciones microfinancieras y la calificación

crediticia asignada. Los autores encuentran que en contraste con las variables sociales, los factores financieros explican de manera significativa la calificación crediticia. Finalmente Lee (2007) investiga las calificaciones de crédito corporativo.

La investigación sobre redes neuronales y calificaciones crediticias de gobiernos subsoberanos a nivel internacional es muy escasa y para México es prácticamente inexistente. La mayor parte de los estudios sobre calificación crediticia se enfoca en mercados corporativos o de deuda en países desarrollados y poco se ha dedicado al estudio de mercados emergentes. Para el caso de México sin embargo ha habido algunos intentos iniciales usando modelos Probit Ordenados en el análisis de las finanzas públicas estatales—ver García-Romo (2005) y Yorio (2006).

Este estudio pretende extender la literatura sobre finanzas locales y calificación crediticia usando un algoritmo de RNA y evaluar a su vez el desempeño contra modelos Probit Ordenados y ADM.

III. Redes Neuronales, Análisis Discriminante y Modelos Probit

En esta sección se presenta primero el algoritmo de retropropagación y la red neuronal usada en nuestro estudio y se concluye presentando las características del modelo probit ordenado.

3.1 Red Neuronal

Las redes neuronales artificiales (RNA), modelos matemáticos inspirados en el funcionamiento de las neuronas biológicas, han mostrado éxito en el modelaje de sistemas donde se desconocen las reglas de comportamiento pero donde se pueda contar con ejemplos descriptivos de éste. Este modelado se consigue debido a que las RNA son capaces de aprender, esto es, adecuar su comportamiento de acuerdo a las situaciones presentadas a éstas por medio de ejemplos. Asimismo, las RNA pueden, hasta cierto punto, abstraer y representar numéricamente información esencial de los ejemplos presentados, así como generalizar sus respuestas, de manera que, una vez entrenadas, presenten una respuesta tolerante a variaciones o ruido en la información.

Las RNA están formadas de componentes básicos de procesamiento, llamados neuronas, que se conectan unos a otros a través de valores numéricos reales llamados pesos. Las RNA reciben información del exterior, la cual procesan de acuerdo a una topología y regla de evaluación particular y producen salidas que pueden representar clasificaciones, pronósticos, evaluaciones, etc. Asimismo, las neuronas de una RNA pueden estar organizadas en diferentes grupos conectados entre sí, conocidos como niveles. Dependiendo de la estrategia de conexión de las neuronas y de los algoritmos utilizados para entrenar los pesos, se pueden crear diferentes modelos de RNA.

Existen un sinnúmero de modelos de RNA y algoritmos para entrenarlas. Uno de los más utilizados por su versatilidad es el de perceptrones de varios niveles entrenado con retropropagación

(Haykin 1999). Un perceptrón es una neurona j cuya salida está definida por:

$$y_j = \tau \left(\sum_{i=1}^m w_{ji} x_i + b_j \right) \quad (3.1)$$

donde:

w_{ji} son pesos que conectan hacia la neurona j desde m otras neuronas en un nivel anterior o desde m entradas externas a la red.

b_j es un peso adicional, referido comúnmente como "umbral," que se asocia a la j -ésima neurona.

x_i es la salida del i -ésima neurona que se encuentra en la capa anterior a la neurona y_j , y que conecta con ésta a través de w_{ji} . También puede corresponder a la j -ésima entrada externa a la red.

$\tau(\cdot)$ es una función continua y diferenciable que controla la activación de la neurona.

m es el número de conexiones hacia la neurona j .

Una red de perceptrones con m entradas, un solo nivel escondido con h neuronas y una neurona en el nivel de salida como la mostrada en la figura 1, define un sistema capaz de aproximar el valor de una función arbitraria $f(x_1, x_2, \dots, x_m)$ (Haykin 1999). Las entradas a la RNA corresponden a los valores de las variables independientes x_1, x_2, \dots, x_m ; su salida (aproximación de la función) está definida como:

$$F(x_1, x_2, \dots, x_m) = \sum_{j=1}^h \alpha_j \tau \left(\sum_{i=1}^m w_{ji} x_i + b_j \right) \quad (3.2)$$

Donde:

$w_{ji}, b_j \quad j = 1..h, i = 1..m$ son los pesos de las neuronas en el nivel escondido,

$\alpha_j \ j=1..h$ corresponden a los pesos de las neuronas en el nivel de salida,

$\tau(u) = \frac{1}{1 + e^{-\lambda u}}$ es la función de activación utilizada para las neuronas del nivel escondido; λ es un coeficiente de escalamiento que controla el uso de la función de activación en intervalos con $\tau'(u) \neq 0$, lo cual facilita el entrenamiento.

Nótese que, de acuerdo a la ecuación (3.2), la función de activación de la única neurona en la capa de salida de esta RNA es una función lineal $\mu(x) = x$.

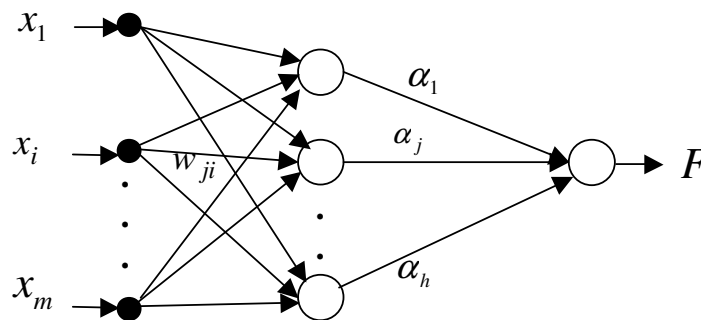


Figura 3.1. Red de perceptrones con m entradas y un nivel escondido con h neuronas, capaz de aproximar la función $f(x_1, \dots, x_m)$.

Utilizando la red neuronal definida en (3.2), puede construirse un modelo para aproximar la variable inobservada que define las calificaciones crediticias de gobiernos sub-soberanos de los estados de la república Mexicana, donde las variables observadas son la entrada a la RNA, y la calificación asignada es la salida, esto es, la función aproximada. De acuerdo a (3.2), la salida de la red neuronal es una función no lineal de las variables explicativas. La capacidad para aproximar de esta RNA está apoyada en el teorema de aproximación

universal (Cybenko 1989), el cual asegura que existen valores w_{ji}, α_j , y b_j tal que:

$$|F(x_1, \dots, x_m) - f(x_1, \dots, x_m)| < \varepsilon \text{ para } \varepsilon > 0, \quad (3.3)$$

F definida según (3.2).

3.1.1 Estrategia de entrenamiento

Para ajustar los pesos de la RNA se utilizó el algoritmo de retropropagación (Werbos 1990) según la derivación descrita en (Rumelhart et al. 1986). Retropropagación es un método de aprendizaje supervisado basando en ejemplos, derivado de una minimización de gradiente descendente en una función de costo. Este algoritmo busca disminuir progresivamente el error de salida de la red generado por un conjunto de P ejemplos de entrenamiento, de los que se conoce la calificación crediticia D_p , $p = 1..P$; esto es, se busca minimizar:

$$E_{total} = \sum_{p=1}^P E_p \quad (3.4)$$

donde:

$$E_p = |F_p - D_p| \quad (3.5)$$

F_p es la salida de la RNA al evaluarse con el p -ésimo ejemplo de entrenamiento. La modificación de los pesos de la red se realiza iterativamente hasta alcanzar un E_{total} mínimo deseado o hasta que se alcanza un máximo de barridas sobre el conjunto de entrenamiento.

El algoritmo para encontrar los pesos w_{ji}, b_j y α_j $j = 1..h, i = 1..m$ de la RNA definida por (3.2) que será capaz de determinar la calificación inobservada de un año t , basado en P ejemplos de entrenamiento se describe a continuación. Este algoritmo está basado en el definido en (Rumelhart et al. 1986) donde se encuentra una descripción detallada de la derivación de retropropagación.

Paso 1.

$$w_{ji} = \text{random}(-0.01, 0.01)$$

$$b_j = \text{random}(-0.01, 0.01) \quad i = 1..m, j = 1..h$$

m es el número de variables observadas o explicativas,
 h es el número de neuronas escondidas seleccionados para la red, definido de manera experimental según se describe en la sección 4.3

$\text{random}(a, b)$ es una función generadora de valores al azar distribuidos uniformemente en $[a, b]$.

$$\text{Barridas} = 0$$

Paso 2. Repite:

$$2.1 \quad E_{total} = 0$$

$$2.2 \quad \text{Para } \forall p \in P$$

$$O_j = \sum_{i=1}^m w_{ji} x_{ip} + b_j, \text{ para cada } j = 1..h$$

x_{ip} es la i -ésima variable explicativa del p -ésimo ejemplo en el conjunto de entrenamiento P

$$F_p = \sum_{j=1}^h \alpha_j \tau(O_j)$$

$$\tau(u) = \frac{1}{1 + e^{-\lambda u}}, \quad \lambda \text{ es el coeficiente de escalamiento}$$

$$E_p = |F_p - D_p|$$

$$E_{total} = E_{total} + E_p$$

$$\delta_1 = (D_p - F_p) F_p (1 - F_p)$$

$$\alpha_j = \alpha_j + \eta \delta_1 O_j, \quad j = 1..h$$

$$\delta_{2j} = \tau'(O_j) \delta_1 \alpha_j, \quad j = 1..h$$

$$b_j = b_j - \eta \delta_{2j}$$

$$w_{ji} = w_{ji} + \eta \delta_{2j} x_i, \quad j = 1..h, i = 1..m$$

$$2.3 \quad \text{Barridas} = \text{Barridas} + 1$$

Hasta que $E_{total} \leq 0.05$ ó $Barridas > MáximoBarridas$ (número máximo de iteraciones deseadas sobre el conjunto de entrenamiento)

Paso 3. Fin

3.1.2 Algoritmo de Asignación de Calificación Crediticia

Una vez entrenada, la red está lista para asignar una calificación de acuerdo a la escala de calificaciones otorgada por FichRatings (tabla A.1 en el anexo). Debido a que se está utilizando la red como un aproximador a un función $f(\mathbf{x}) : \mathfrak{R}^m \rightarrow \mathfrak{R}$, es necesario transformar la salida de la red neuronal, que es un valor real, al mejor valor entero correspondiente a una categoría en la escala de la tabla A.1. Para esto, se siguió el siguiente algoritmo:

Paso 1. Leer los valores explicativos x_k $k = 1..m$

m es el número de variables explicativas

Paso 2. Calcular $F = \sum_{j=1}^h \alpha_j \tau(\sum_{i=1}^m w_{ji} x_i + b_j)$

Paso 3. $CategoriaPosible = floor(F + 0.5)$

donde $floor(x)$ es una función que entrega el máximo entero menor o igual a a su argumento x

Paso 4. Si $CategoriaPosible < 1$ ó $CategoriaPosible > MAX_CATEGORIA$

$CategoriaAsignada = null$

sino

$CategoriaAsignada = CategoríaPosible$

donde $MAX_CATEGORIA$ corresponde al máximo valor numérico permitido de acuerdo a la tabla A.1. Si la red generó un valor más allá del rango permitido o menor que 1, se designa a la red incapaz de clasificar.

3.2 Modelo Probit Ordenado

Los modelos Probit Ordenados se construyen a partir de un *modelo de variable latente subyacente* o inobservada que satisface las suposiciones del modelo lineal clásico (Wooldridge, 2001):

$$y_t^* = \sum_{k=1}^K \beta_k x_{kt} + \varepsilon_t \quad (3.6)$$

Se asume que la calificación inobservada del año t , y_t^* , es una función lineal de una serie de variables explicativas, x_k , observadas ese mismo año, y de un término de error, ε_t , que en el caso de los modelos Probit, se asume sigue una distribución de probabilidad normal.

Siguiendo el procedimiento propuesto por García-Romo (2005) durante el periodo de análisis, se observa un rango de calificaciones que va de AAA a BB. De modo que se asigna un número discreto ordenado a cada calificación: 1 para AAA, 2 para AA+, ..., y 12 para BB.⁴ Entonces, la relación entre el índice inobservado, y^* , y la calificación observada, y_t , se establece como sigue:

$$y_t = \left\{ \begin{array}{ll} \text{AAA} = 1 & \text{si } y^* \leq \theta_1 \\ \text{AA+} = 2 & \text{si } \theta_1 < y^* \leq \theta_2 \\ \text{AA} = 3 & \text{si } \theta_2 < y^* \leq \theta_3 \\ \dots & \dots \\ \text{BB} = 12 & \text{si } \theta_{11} < y^* \leq \theta_{12} \end{array} \right\} \quad (3.7)$$

Los parámetros θ_i representan los puntos de quiebre del índice y marcan las fronteras entre cada calificación. Estos parámetros son estimados en conjunto con los coeficientes β_k de las variables explicativas a través del método de Máxima Verosimilitud.⁵

⁴ Ver tabla A.1 en el anexo.

⁵ La estimación de los modelos Probit ordenados vía Máxima Verosimilitud arroja mejores estimadores y pronósticos que estimaciones vía Mínimos Cuadrados

Siguiendo a García-Romo (2005), la probabilidad de obtener cada calificación es

$$\begin{aligned}
 P(y_t = 1) &= F\left(-\sum_{k=1}^K \beta_k x_{kt}\right) \\
 P(y_t = 2) &= F\left(\theta_1 - \sum_{k=1}^K \beta_k x_{kt}\right) - F\left(-\sum_{k=1}^K \beta_k x_{kt}\right) \\
 P(y_t = 3) &= F\left(\theta_2 - \sum_{k=1}^K \beta_k x_{kt}\right) - F\left(\theta_1 - \sum_{k=1}^K \beta_k x_{kt}\right) \\
 &\dots \\
 P(y_t = 12) &= 1 - F\left(\theta_{11} - \sum_{k=1}^K \beta_k x_{kt}\right)
 \end{aligned} \tag{3.8}$$

donde $F(\dots)$ es la función de distribución acumulada. Es decir, la probabilidad de obtener una calificación AAA es igual a la probabilidad de que y_t sea menor o igual a 1, la probabilidad de obtener AA+ es igual a la probabilidad de que y_t sea menor o igual a 2 menos la probabilidad de que sea menor o igual a 1, y así sucesivamente hasta encontrar que la probabilidad de obtener una calificación BB es igual a uno menos la probabilidad de que y_t sea menor o igual a 11.

La naturaleza no lineal de la función de distribución normal acumulada $F(z)$ provoca que los coeficientes de las variables dependientes en el *modelo de variable latente subyacente* no representen los efectos marginales correspondientes. Por tanto, el *efecto parcial* resulta más bien la derivada parcial de la probabilidad de obtener una calificación j con respecto a la variable x_k :

$$\begin{aligned}
 \frac{\partial P(y = 1)}{\partial x_k} &= -\left[f\left(-\sum_{k=1}^K \beta_k x_{kt}\right)\right] \beta_k \\
 \frac{\partial P(y = 2)}{\partial x_k} &= \left[f\left(-\sum_{k=1}^K \beta_k x_{kt}\right) - f\left(\theta_1 - \sum_{k=1}^K \beta_k x_{kt}\right)\right] \beta_k
 \end{aligned} \tag{3.13}$$

Ordinarios (MCO) y con esto se permite superar las dos principales desventajas de los Modelos de Probabilidad Lineal (MPL): las probabilidades ajustadas de los modelos Probit se encuentran estrictamente entre cero y uno debido al supuesto de normalidad en la función de distribución, y el modelo permite que los efectos parciales de las variables independientes, medidas en niveles, no sean constantes (ver Wooldridge, 2001).

$$\frac{\partial P(y=3)}{\partial x_k} = [f(\theta_1 - \sum_{k=1}^K \beta_k x_{kt}) - f(\theta_2 - \sum_{k=1}^K \beta_k x_{kt})] \beta_k$$

$$\frac{\partial P(y=12)}{\partial x_k} = [f(\theta_{11} - \sum_{k=1}^K \beta_k x_{kt})] \beta_k$$

donde $f(\dots)$ es la función de densidad de probabilidad normal. De este modo, se puede estimar el cambio en la probabilidad de que cierto Estado sea valorado con cualquiera de las calificaciones debido a la variación de una unidad en la variable x_k . Como podemos ver, el efecto marginal de una variable dada tendrá el mismo signo de su coeficiente sobre la probabilidad de $y=12$, y un signo contrario sobre la probabilidad de $y=1$. Para todas las demás variables la concordancia de los signos es ambigua y sólo se puede determinar después de su estimación, ya que depende de los valores puntuales de las demás variables.

De esta forma, si el signo del coeficiente β_k es negativo, inequívocamente el efecto marginal de la variable x_k será el de aumentar la probabilidad de obtener la calificación $y=1$ y de disminuir la de $y=12$, y lo contrario si β_k es positivo, mientras que para las demás calificaciones el signo del efecto marginal sólo se puede determinar tras su estimación en los valores puntuales de las variables para un Estado específico. Así, el signo del efecto marginal determina si variaciones en la variable explicativa se asocian a incrementos o disminuciones en la probabilidad que tiene un gobierno local de obtener la calificación que se evalúa.

IV. Evaluación de los Métodos de Pronóstico de Calificación Crediticia

En esta sección se presentan los resultados de las aplicaciones del algoritmo de Red Neuronal Artificial (RNA) y el modelo Probit Ordenado haciendo énfasis en la capacidad de pronóstico y clasificación de los dos métodos para datos dentro y fuera de la muestra. Como ejercicio adicional se compara también el desempeño de la RNA con la clasificación obtenida de un Análisis de Discriminante Múltiple (ADM).

4.1 Análisis de Datos y definición de variables

Información sobre finanzas públicas de las entidades federativas en México fue recolectada de la base de datos provista por la calificadora FitchRatings disponible en la sección de finanzas públicas de su página web (www.fitchmexico.com). Se prefiere esta información a la proporcionada por otras fuentes como el INEGI o la SHCP con el fin de mantener la consistencia de las variables usadas por la agencia calificadora y por ser la información primaria obtenida por esta FitchRatings directamente de las tesorerías de los estados en el proceso de calificación.

Tabla 1. Definición de variables financieras estatales

Nombre	Definición	Información específica
<i>Dimensión Estatal</i>		
IT	Ingresos Totales	Ingresos Propios + Ingresos Federales (y Estatales, en el caso de los municipios)
IFOS	Ingresos Fiscales Ordinarios	Ingresos Propios (impuestos, derechos, productos, etc.) + participaciones federales estatales (excluyendo las correspondientes a los municipios) + otros ingresos federales disponibles (Aportaciones Federales Ramo 33 y otros)
GPRI	Gasto primario	Servicios personales, materiales y servicios generales
GCR	Gasto corriente	Niveles a millones de pesos del 2006.
AHOIN	Ahorro Interno	Ingreso total menos gasto primario
TRIB	Participaciones Federales / Tansferencias Totales	Las transferencias incluyen las transferencias a municipios etiquedas, no etiquetadas y otras.
<i>Generación de Ingreso, Ahorro e Inversión</i>		
IEIT	Ingresos Propios/Ingresos Totales	Ingresos Propios (impuestos, derechos, productos, etc.) + Participaciones Federales al Municipio + Fondo de Fortalecimiento Municipal (F-IV Ramo 33)/Ingresos Propios
IEGO	Ingresos Propios/Gasto Corriente	Ingreso propio por cada peso gastado
INVI	Gasto en inversión/Ingresos propios	
INVB	Gasto en inversión/Producto Interno Bruto	
INVP	Gasto en inversión/Gasto primario	Inversión Estatal (excluye inversión del Ramo 33)/Ingresos Propios
AHOINIFO	Ahorro Interno/IFOS	Balance primario menos el pago de intereses
<i>Gasto Ordinario</i>		
GOIFO	Gasto corriente/IFOS (Gasto corriente +	
GOTNEIFO	Transferencias No Etiquetadas)/IFOS	
CORP	Gasto corriente/gasto primario	Gasto corriente, Transferencias, Inversión Total y Adeudos de Ejercicios Fiscales Anteriores (ADEFAS)
<i>Apalancamiento</i>		
DEU	Balance de la deuda	SDEU/Participaciones Federales (excluyendo las correspondientes a los municipios)
DAH	DEU/Ahorro Interno	SDEU/(Participaciones Federales al municipio + Fondo de Infraestructura Social (Ramo 33) + Fondo de Fortalecimiento Municipal (Ramo 33) + Otras Aportaciones)
DPIB	DEU/PIB	
DPAR	DEU/Ingresos Federales	
DIFOS	DEU/IFOS	Deuda pública directa + deuda organismos + deuda pública municipal
DD	Deuda Directa	Niveles en millones de pesos del 2006
DIOD	Deuda Indirecta por Organismos descentralizados	Niveles en millones de pesos del 2006
DIM	Deuda Indirecta por Municipios	Niveles en millones de pesos del 2006
DDIFO	DD / IFO	
DDAI	DD /AI	
<i>Sostenibilidad de la Deuda</i>		
SDEU	Servicio de la deuda	IFOS—gasto operativo—transferencias no etiquetadas y otros
SDEUAI	SDEU/Ahorro Interno	
SAHO	SDEU/IT-GPRI + INV	INV representa el gasto de inversión
SPAR	SDEU/Ingresos Federales	Pago de intereses + amortizaciones de deuda
SIFOS	SDEU/IFOS	
<i>Resultados</i>		
BPRI	Balance primario	Niveles en millones de pesos del 2006
BFIN	Balance financiero	Niveles en millones de pesos del 2006
PIB	Producto Interno Bruto	Niveles en millones de pesos del 2006
PIBPER	PIB Per capita	Niveles en pesos del 2006

Fuente: García-Romo (2005) y cálculos propios.

Se calcularon posteriormente diversas razones financieras de acuerdo a las definiciones propuestas por García-Romo (2005) y las

propuestas por FitchRatings en sus análisis de crédito—ver tabla 1 enseguida. La base de datos final se compone de 22 entidades federativas calificadas en México por FitchRatings desde el año 2001 al año 2007. La información financiera original por año para estos estados consiste de 133 casos estado/año con 35 variables financieras para cada uno de ellos incluyendo la calificación crediticia.

Tabla 2. Estadística descriptiva de las finanzas públicas estatales 2001-2007.

	Desviación						
	Promedio	Estándar	Sesgo	Curtosis	Máximo	Mínimo	Normalidad
IT ^a	18,650	10,773	0.7982	0.1199	48,695	4,180	0.9421
IFOS	7,542	4,861	0.9003	0.0724	21,501	393	0.9156
GPRI	18,640	11,026	0.7913	0.1106	50,051	1,145	0.9406
GCR	3,690	2,694	0.5567	-0.8519	10,378	126	0.9031
AHOIN	1,777	1,225	1.0189	0.6485	5,571	0.0000	0.9102
TRIB	1.9519	0.4360	0.7343	1.6336	3.5658	1.0400	0.9472
IEIT	0.0673	0.0372	1.2265	1.3481	0.1909	0.0050	0.8729
INVI	1.7181	0.9882	1.8865	4.6434	6.0981	0.2831	0.8322
INVB	0.0131	0.0160	6.7553	56.5614	0.1599	0.0000	0.8084
INVP	0.1091	0.0634	2.6699	10.3917	0.4725	0.0157	0.8045
IEGO	0.3888	0.1883	0.9052	0.2587	0.9616	0.0906	0.9217
AHOINIFO	0.2458	0.0906	-0.4274	-0.0302	0.4440	0.0000	0.9736
GOIFO	0.4570	0.1302	-0.4374	-0.3444	0.7261	0.1200	0.9802**
GOTNEIFO	0.9709	0.5145	1.8915	2.2664	2.6904	0.5930	0.6416
CORP	0.1834	0.0627	0.0692	-0.0289	0.3332	0.0400	0.9864*
DEU	2,100	2,074	2.0000	4.0000	10,658	0.0000	0.7902
DAH	1.441	1.460	2.958	13.049	10.344	0.0000	0.7144
DPIB	0.022	0.075	8.118	67.849	0.708	0.0000	0.8064
DPAR	0.396	0.307	0.267	-0.193	1.181	-0.539	0.9438
DIFOS	6.575	72.466	11.53	132.998	836.010	0.0000	0.9594
DD	1,513.20	1,473.66	1.930	5.630	8,181.20	0.0000	0.8117
DIOD	401.91	738.44	2.500	6.770	4,310.50	0.0000	0.5942
DIM	145.91	232.83	2.010	3.810	1,121.88	0.0000	0.7173
DDIFO	0.2172	0.1484	0.3189	-0.4432	0.6063	0.0000	0.9699
DDAI	1.0317	1.0554	3.3378	17.3855	8.1009	0.0000	0.6998
SDEU	359.60	488.08	3.2200	14.3600	3528.67	0.0000	0.6524
SDEUAI	0.2290	0.2372	1.7458	3.1734	1.2037	0.0000	0.8194
SAHO	0.2586	0.3470	3.4530	15.9867	2.4861	0.0000	0.6417
SPAR	0.0595	0.0896	0.9336	6.7695	0.4921	-0.2641	0.8741
SIFOS	0.0515	0.0524	2.0684	5.4257	0.2918	0.0000	0.8160
BPRI	-68.7684	786.1661	-0.9761	2.6892	1,921.20	-3,003.10	0.9476
BFIN	-194.2367	788.6629	-1.1801	3.0117	1,604.90	-3,227.40	0.9552
PIB	2,140,141	11,309,570	7	59	107,092e	1,188	0.1970
PIBPERC	6.0762	2.4007	0.5830	-0.4994	12.7854	2.6000	0.9385
Calificación	6	1	-0.3810	0.3911	9	2	0.9893*

^aVer tabla 1 para la definición de variables.

*, ** y *** Significativo al 1%, 5% y 10% respectivamente.

La tabla 2 (arriba) muestra estadística descriptiva básica de las variables financieras, tanto los valores en niveles como las razones financieras. Los valores del sesgo y la kurtosis indican que las variables no parecen seguir individualmente una distribución normal y esto se confirma con el valor de los estadísticos Jarque-Bera y Shapiro Wilk. En adición a esto se observa evidentemente una gran disparidad entre las variables, los rangos estadísticos y la varianza de las variables en nivel son altos, lo que finalmente motiva la normalización de las variables en este estudio.⁶

4.2 Implementación de la Red Neuronal

La topología de la Red descrita en la sección 3.1 fue entrenada utilizando un conjunto de 112 ejemplos de asignación crediticia estatales, cada uno con 34 valores de entrada, correspondientes al número de variables financieras explicativas. Este conjunto será referenciado en el resto de artículo como "datos dentro de la muestra." Para determinar el número óptimo de neuronas a utilizar en el nivel escondido, se entrenaron redes conteniendo desde 3 hasta 40 nodos escondidos, y se probó el desempeño de cada uno sobre un conjunto de validación de 21 casos. Este conjunto será referenciado en el resto del artículo como "datos fuera de la muestra." El desempeño del aproximador se midió en 4 vertientes: porcentaje de aciertos exactos a la calificación y porcentaje de aciertos entre uno, dos y tres *notches* hacia arriba o hacia abajo de la calificación real. Para evitar un sobre ajuste del aproximador neuronal a los datos de entrenamiento, se detuvo el entrenamiento cuando $E_{total} \leq 0.05$. Los pesos se inicializaron al azar con valores en un rango de $[-0.01, 0.01]$, se utilizó un coeficiente de aprendizaje $\eta = 0.01$ y un factor de escalamiento de la función de activación de $\lambda = 0.01$. Los valores iniciales de los pesos, coeficientes de aprendizaje y factor de

⁶ La alta variabilidad observada se debe en buena medida al efecto que ejercen tanto el Distrito Federal como el Estado de México.

escalamiento se definieron empíricamente a través de observar los resultados de entrenamiento obtenidos en las primeras topologías. La figura 4.1 muestra el porcentaje de aciertos exactos obtenido por las redes probadas fuera de la muestra, donde puede notarse que el mejor desempeño se obtuvo para una red con 10 nodos escondidos.

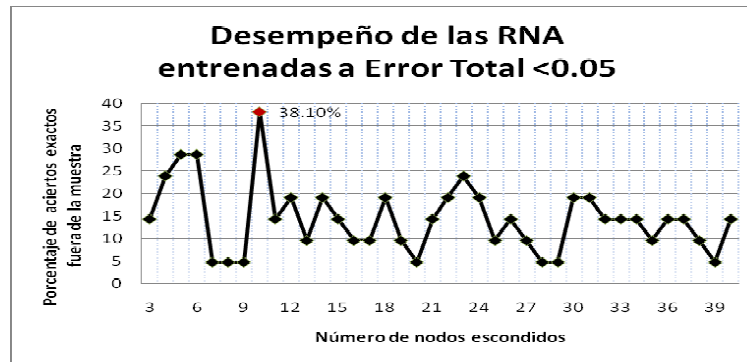


Figura 4.1 Desempeño obtenido por RNA con diferente número de nodos escondidos para datos fuera de la muestra

4.2.1 Análisis de Clasificación vía Red Neuronal Artificial

En esta sección se muestra la clasificación de la calificación crediticia de los Estados usando la red neuronal descrita en la sección anterior. La matriz de clasificación en la tabla 3 abajo muestra un porcentaje de aciertos en el cien por ciento de los casos dentro de la muestra. Este resultado en principio resulta sorprendente al mostrar la eficiencia de la Red Neuronal como algoritmo de clasificación para datos *dentro de la muestra* y es consistente con lo encontrado en la literatura. El aprendizaje del algoritmo en la etapa de entrenamiento permite mejorar iterativamente el desempeño de la red hasta encontrar este resultado. Sin embargo la verdadera eficiencia de la red se pone a prueba en las siguientes secciones donde se examina de manera directa la capacidad de pronóstico que tiene este algoritmo de red neuronal artificial con datos *fuera de la muestra*.

Tabla 3. Matriz de clasificación de RNA's con 10 nodos, entrenada a un $E_{total} \leq 0.05$ con datos dentro de la muestra

		Calificación Real ^a							
		AA+	AA	AA-	A+	A	A-	BBB+	BBB
Calificación	AA+	3	0	0	0	0	0	0	0
Pronosticada	AA	0	7	0	0	0	0	0	0
	AA-	0	0	3	0	0	0	0	0
	A+	0	0	0	35	0	0	0	0
	A	0	0	0	0	26	0	0	0
	A-	0	0	0	0	0	28	0	0
	BBB+	0	0	0	0	0	0	6	0
	BBB	0	0	0	0	0	0	0	4
	Observaciones Totales		3	7	3	35	26	28	6
Pronósticos Acertados		3	7	3	35	26	28	6	4
Proporción		100%	100%	100%	100%	100%	100%	100%	100%

a La equivalencia en calificaciones se muestra en la tabla A.1

4.3 Implementación del Modelo Probit Ordenado

La tabla A.2 en el anexo muestra la matriz de correlaciones entre las variables y también resulta evidente la asociación lineal existente entre muchas de ellas para distintos niveles, lo que sugiere reducir la dimensión de la base mediante métodos multivariados. La asociación lineal entre las variables explicativas justifica el uso de componentes principales y análisis factorial para explicar la variación total de la muestra con un número significativamente menor de indicadores de una manera más eficiente.

Del análisis multivariado aplicado a esta muestra se concluye que seis factores explican el comportamiento de las finanzas públicas estatales en México.⁷ Estos seis factores son usados en la estimación de los modelos probit ordenados para reducir el problema de multicolinealidad en la muestra. Estos factores reportan una varianza explicada total del 80.1 por ciento. Una relación de estos índices, la

⁷ Estos factores se reportan formalmente en un artículo de Mendoza (2010) por publicarse en la Revista de El Trimestre Económico donde el lector puede consultar más detalles acerca de los métodos multivariados usados y una mejor descripción de las implicaciones. En tal estudio se proponen índices para el monitoreo de la presión y la vulnerabilidad financiera de los gobiernos estatales en México.

variación explicada por cada uno y las variables que los componen se presenta en la tabla 4.⁸

Tabla 4. Factores Financieros Estatales.

Factor ^a	Componentes	Impacto Esperado ^b	Varianza Explicada (%)
1. Dimensión	IT, IFOS, GPRI, GCR, AHOIN, DIOD	(?) ^c	20.5
2. Apalancamiento	DDIFO(-) ^d , DPAR(-), DDAI(-), DD(-), DAH(-), DEU(-),	(+) ^e	14.1
3. Sostenibilidad	SIFOS(-), SAHO(-), SDEUAI(-), SDEU(-), SPAR(-)	(+)	13.4
4. Gasto Ordinario	IEGO(-), AHOINIFO(-), GOIFO, GOTNEIFO	(+)	12.0
5. Autonomía	TRIB(-), CORP, IEIT, INVI(-), DIM	(-)	10.5
6. Resultados	BPRI, BFIN, INVP(-)	(-)	9.7

^aEl nombre del factor se asigna de acuerdo al significado de las variables que lo componen. ^bSe refiere al impacto sobre la probabilidad de obtener la calificación más alta. ^cImpacto indeterminado. ^dEl signo entre paréntesis indica la dirección del impacto de cada variable. Las variables sin (-) indican impactos positivos. ^eEntre mayor sea el apalancamiento menor será la probabilidad de recibir la calificación más alta—vea la escala definida en la tabla A.1.

La tabla 5 siguiente muestra la matriz de clasificación que resulta de emplear el modelo Probit Ordenado como método de clasificación y pronóstico de las calificaciones dentro de la muestra. Es evidente que para este caso la capacidad de pronóstico del Modelo Probit ordenado es mucho menor que la observada con el algoritmo de Red Neuronal empleado en la sección anterior. Esto debido a que el máximo porcentaje de aciertos (88.6%) se da cuando la calificación es A+.

⁸ Para una descripción más detallada del método de componentes principales y análisis de factores con aplicación a las Finanzas Públicas estatales en México refiérase a Mendoza (2010).

Tabla 5. Matriz de clasificación del Modelo Probit Ordenado

		Calificación Real							
		AA+	AA	AA-	A+	A	A-	BBB+	BBB
Calificación Pronosticada	AA+	0	0	0	0	0	0	0	0
	AA	0	0	0	0	0	0	0	0
	AA-	0	0	0	0	0	0	0	0
	A+	3	7	3	31	10	5	0	0
	A	0	0	0	2	4	2	1	0
	A-	0	0	0	2	11	21	5	4
	BBB+	0	0	0	0	0	0	0	0
	BBB	0	0	0	0	1	0	0	0
Observaciones Totales		3	7	3	35	26	28	6	4
Pronósticos Acertados		0	0	0	31	4	21	0	0
Proporción		0.0%	0.0%	0.0%	88.6%	15.4%	75.0%	0%	0.0%

4.4 Análisis Discriminante Múltiple

Como un ejercicio adicional la tabla 6 siguiente muestra los resultados de clasificación derivados usando funciones discriminantes (AD). Este método es ampliamente usado en la práctica para realizar clasificación de observaciones. Los aciertos del algoritmo de clasificación se presentan en la diagonal principal. El mayor porcentaje de aciertos se ubica en calificaciones **A**. De acuerdo con la distribución de los *desaciertos* se observa en general que el algoritmo AD tiende a sobreestimar las calificaciones de FitchRatings. Es decir, el algoritmo pronostica que la calificación es en general mayor a la originalmente otorgada. Aunque se observa que la capacidad de clasificación es de hecho menor que la obtenida por el método de RNA los resultados son mucho más satisfactorios que los arrojados por el modelo Probit Ordenado.

Tabla 6. Matriz de clasificación del Análisis Discriminante Múltiple

		Calificación Real							
		AA+	AA	AA-	A+	A	A-	BBB+	BBB
Calificación Pronosticada	AA+	3	2	1	0	0	0	0	0
	AA	0	5	1	4	5	1	0	0
	AA-	0	0	1	4	0	0	0	0
	A+	0	0	0	23	1	0	0	0
	A	0	0	0	4	20	1	0	0
	A-	0	0	0	0	0	21	1	0
	BBB+	0	0	0	0	0	4	4	1
	BBB	0	0	0	0	0	1	1	3
Observaciones Totales		3	7	3	35	26	28	6	4
Pronósticos Acertados		3	5	1	23	20	21	4	3
Proporción		100%	71.4%	33.0%	65.7%	76.9%	75.0%	66.7%	75.0%

4.5 Bondad de Ajuste y Capacidad de Pronóstico

La capacidad de pronóstico y la bondad de ajuste de los tres métodos presentados (i.e, Redes Neuronales, modelos Probit Ordenados y Análisis Discriminante) se ponen a prueba dividiendo la muestra en dos partes. La primera, definida ya en secciones anteriores y denominada como datos *dentro de la muestra*, integrada por todas aquellas observaciones hasta el año 2006 (es decir 112 obs.) y la segunda, *datos fuera de la muestra*, por todas las observaciones correspondientes al año 2007 (i.e., 21 obs.). La primera de estas muestras sirve como base para la estimación de los parámetros asociados al modelo probit ordenado y al análisis discriminante respectivamente. Esta muestra también se usa como insumo para definir la Red Neuronal Base. Con estas estimaciones se procede a clasificar ahora las observaciones de los *datos fuera de la muestra*, es decir se usan los modelos de estimación obtenidos con los datos dentro de la muestra y se usan como algoritmo de pronóstico de las observaciones correspondientes a los años 2007, datos que no fueron empleados originalmente en la estimación de los parámetros.

En la tabla 7 siguiente se muestra el porcentaje de aciertos totales por cada uno de los métodos así como algunas medidas para evaluar el error de pronóstico. Se observa que el método de AD puede

clasificar correctamente 71.4% de los casos *dentro de la muestra* (con un error absoluto medio de 0.4554 'notches') mientras que solo puede pronosticar 28.6% de los casos *fuera de la muestra* (con un error absoluto medio de 1.43 'notches'). Si el pronóstico se realiza para un intervalo de un 'notch' el porcentaje de aciertos se incrementa significativamente a 89.3% y 61.90% para datos dentro y fuera de la muestra respectivamente. Por su parte los aciertos se incrementan a 94.6% y a 99.1% para dos y tres 'notches' respectivamente en estimaciones dentro de la muestra y a 76.2% y 95.2% para datos fuera de la muestra. Es decir la capacidad de pronóstico de este algoritmo es satisfactoria en términos generales.

Por su parte, la capacidad de pronóstico del modelo Probit ordenado para clasificaciones dentro de la muestra es la más baja de todos los métodos examinados. Se reportan solo 50% de aciertos para datos dentro de la muestra. De forma interesante, los aciertos para datos fuera de la muestra ascienden a solo 28.6%, un porcentaje de aciertos similar al del método de Discriminantes. El nivel de variabilidad en el pronóstico es muy similar en los dos métodos.

De los tres métodos, las Redes Neuronales arrojan un mejor pronóstico para los datos fuera de la muestra al clasificar correctamente 38.1% de los casos. De manera interesante sin embargo se reporta un nivel de variabilidad mayor que con respecto al método de Discriminantes o al Probit Ordenado.

Respecto a los resultados obtenidos por la RNA en este estudio cabe especificar que la topología presentada corresponde a la red que obtuvo el mejor desempeño fuera de la muestra. Los aciertos obtenidos fuera de la muestra se presentan en la diagonal principal de la tabla 9, notándose que la RNA acierta de manera exacta en el 100% de los casos dentro de la muestra y en el 38.1% de los casos fuera de la muestra. Además, respecto al ejercicio fuera de la muestra, el error absoluto medio fue de 1.6191 *notches*, el error absoluto medio relativo fue de 0.2892 y el error absoluto máximo fue

de 5 *notches*. Los aciertos fuera de la muestra para intervalos entre 1, 2 y 3 *notches* fueron 61.90% 71.43% y 76.19% respectivamente (ver tabla 4.2).

Tabla 7. Desempeño dentro y fuera de la muestra.

Criterio	Análisis	Probit	Red
	Discriminante*	Ordenado	Neuronal
% Aciertos en la muestra	71.4%	50.0%	100%
% Aciertos entre 1 notch	89.3%	3.57%	100%
% Aciertos entre 2 notches	94.6%	20.54%	100%
% Aciertos entre 3 notches	99.1%	50.01%	100%
Error Absoluto Medio	0.4554	0.7411	0.0
Error Absoluto Medio Relativo	0.0826	0.1660	0.0
Error Absoluto Máximo	4	3.00	0
% Aciertos fuera de la muestra	28.57%	28.57%	38.10%
% Aciertos entre 1 notch	61.90%	47.62%	61.90%
% Aciertos entre 2 notches	76.19%	9.5%	71.43%
% Aciertos entre 3 notches	95.24%	9.5%	76.19%
Error Absoluto Medio	1.4300	1.1428571	1.6191
Error Absoluto Medio Relativo	0.2549	0.26564625	0.2892
Error Absoluto Máximo	4	4	5

Tabla 8. Matriz de clasificación de RNA con 10 nodos escondidos entrenadas a $E_{total} \leq 0.05$, datos en la muestra.

		Calificación Real							
		AA+	AA	AA-	A+	A	A-	BBB+	BBB
Calificación	2	3	0	0	0	0	0	0	0
Pronosticada	3	0	7	0	0	0	0	0	0
	4	0	0	3	0	0	0	0	0
	5	0	0	0	35	0	0	0	0
	6	0	0	0	0	26	0	0	0
	7	0	0	0	0	0	28	0	0
	8	0	0	0	0	0	0	6	0
	9	0	0	0	0	0	0	0	4
Observaciones Totales		3	7	3	35	26	28	6	4
Pronósticos Acertados		3	7	3	35	26	28	6	4
Proporción		100%	100%	100%	100%	100%	100%	100%	100%

Tabla 9. Matriz de clasificación de RNA's con 10 nodos escondidos entrenadas a $E_{total} \leq 0.05$, datos fuera de la muestra.

		Calificación Real							
		AA+	AA	AA-	A+	A	A-	BBB+	BBB
Calificación Pronosticada	1	0	0	0	0	1	0	0	0
	2	0	0	0	0	2	0	0	0
	3	1	0	0	1	0	1	0	0
	4	0	1	1	0	0	0	0	0
	5	0	0	0	3	1	0	1	0
	6	0	0	0	0	2	0	1	0
	7	0	0	0	0	1	2	0	0
	8	0	0	0	0	0	1	0	0
	9	0	0	0	0	0	0	0	0
	10	0	0	0	1	0	0	0	0
Observaciones Totales		1	1	1	5	7	4	2	0
Pronósticos Acertados		0	0	1	3	2	2	0	0
Proporción		0%	0%	100%	60%	29%	50%	0%	100%

5. Conclusiones

Las calificaciones de riesgo crediticio se han convertido en un insumo importante para mejorar la claridad en el proceso de evaluación del riesgo crediticio de las finanzas públicas en las entidades federativas en México y, también, como un indicador que permite a órganos reguladores monitorear el desempeño de las finanzas públicas estatales.

Resulta cada vez más apremiante, dada la crisis financiera en la que actualmente enfrentan los gobiernos locales (estatales y municipales) contar con herramientas de pronóstico de calificación de riesgo. En este trabajo comparamos tres métodos usados comúnmente en la literatura para pronosticar la calificación crediticia y evaluamos su desempeño para la clasificación de las finanzas públicas de los estados en México desde el año 2001. Se usan datos financieros y económicos proporcionados por FitchRatings como variables explicativas o de entrada en el caso de las Redes Neuronales Artificiales (RNA).

Se encuentra que el algoritmo de Redes Neuronales Artificiales empleado tiene una mejor capacidad de pronóstico puntual para datos

dentro y fuera de la muestra. Dentro de la muestra se puede tener una aproximación exacta de las calificaciones crediticias asignadas a los estados, mientras que la capacidad de pronóstico disminuye sustancialmente cuando se pronostican calificaciones fuera de la muestra. En este último caso, para datos fuera de la muestra, se clasifican correctamente 38.5% de las observaciones, en clara ventaja con el Análisis de Discriminante Múltiple (ADM) y el Probit Ordenado donde se acierta en 28.5% de las observaciones respectivamente. Sin embargo, la variación medida por el Error Absoluto Medio y el Error Absoluto Medio Relativo, es mayor para el caso de las redes neuronales que para los otros dos métodos de pronóstico de calificación crediticia.

También se observa que el Análisis Discriminante tiene un mejor pronóstico por intervalos que las Redes Neuronales o los modelos Probit. Aunque para uno y dos *notches* la clasificación correcta es similar entre el ADM y la RNA, dentro de tres *notches* el Análisis Discriminante clasifica correctamente el 95% de los casos comparado con solo 76% de las Redes Neuronales.

En términos generales sin embargo se muestra que las redes neuronales proveen una mejor estimación puntual, mientras que el Análisis Discriminante genera una mejor estimación por intervalos y para datos dentro de la muestra. El modelo Probit Ordenado no logra clasificar de manera satisfactoria la calidad crediticia de los estados en México. En otras palabras, si el analista busca una estimación de la calificación crediticia en intervalos de dos o tres *notches* el Análisis Discriminante supera sustancialmente al algoritmo de Redes neuronales aquí propuesto. Si el analista busca una estimar la calificación crediticia de manera puntual la Red Neuronal Artificial es sin duda la mejor herramienta para hacerlo.

Bibliografía

Bharat y Barin (1997). "Performance Evaluation of Neural Network Decision Models", *Journal of Management Information Systems*, Vol. 14, Issue 2, pp. 201-216.

Benell, Julia A., David Crabe, Stephen Thomas and Owain ap Gwilym (2006). "Modelling sovereign credit ratings: Neural networks versus ordered probit", *Expert Systems with Applications*, 30, pp. 415-425.

Cybenko, G. 1989 "Approximation by superpositions of a sigmoidal function," *Mathematics of Control, Signals and Systems*, vol. 2 pp. 303-314

Desai et al., 1996 V.S. Desai, J.N. Crook and G.A. Overstreet Jr., A comparison of neural networks and linear scoring models in the credit union environment, *European Journal of Operational Research* **95** (1) (1996), pp. 24-37 22.

García-Romo, G., Ibarra-Salazar, J. and Sotres-Cervantes, L. (2005). "Determinants of Mexican States Government Credit Ratings", borrador no publicado, Departamento de Economía, Tecnológico de Monterrey.

Gutiérrez Nieto, B.; Serrano Cinca, C. (2007). "Factors explaining the rating of Microfinance Institutions", *Nonprofit & Voluntary Sector Quarterly*, *forthcoming*.

Haykin, S. 1999. *Neural Networks*, Second Edition. Upper Saddle River, NJ: Prentice-Hall.

Kotsiantis, S., E. Koumanakos, D. Tzepelis and V. Tampakas (2006). "Financial Application of Neural Networks: Two Case Studies in Greece". *Lecture Notes in Computer Science, Artificial Neural Networks ICCAN 2006*, part II, pp.672-681.

Kumar, Kuldeep and Sukato Bhattacharya (2006). "Artificial Neural Network vs. Linear Discriminant Analysis in Credit Rating Forecasts". *Review of Accounting and Finance*, Vol. 5, No. 3, pp. 216-227.

Lee, Young-Chan (2007). "Application of support vector machines to corporate credit rating prediction". *Expert Systems with Applications*, 33, 67-74.

Markham, I.S. and C.T. Ragsdale (1995). "Combining Neural Networks and Statistical Predictions to Solve the classification problem in Discriminant Analysis" *Decision Sciences*, 26, p. 229.

Mendoza, Alfonso (2009). "The Information Content and Redistribution Effects of State and Municipal Rating Changes in Mexico". *Economics Discussion Papers*, No 2009-17. Available at <http://www.economics-ejournal.org/economics/discussionpapers/2009-17>

Mendoza, Alfonso (2010). "Indicadores de Desempeño, Presión y Vulnerabilidad de las Finanzas Públicas Estatales en México", *El Trimestre Económico*, por publicarse No 307, Julio Septiembre 2010.

Patuwo, M.Y. Hu and M.S. Hung (1993). "Two-group classification using neural networks", *Decision Sciences*, 24 p. 825.

Salchenberger, L.M., E.M. Cinar and N.A. Lash (1992), "Neural networks: A new tool for predicting thrift failures", *Decision Sciences* 23, pp. 899–916 Tan y Quek

Ting Peng Liang, Herbert Moskowitz and Yuehwern Yih (1992). "Integrating Neural Networks and Semi-Markov Processes for Automated Knowledge Acquisition: An Application to Real-time Scheduling", *Decision Sciences*, Vol. 23, Issue 6, pp. 1297-1314.

Tokinaga S., J. Lu and Y. Ikeda (2005) "Neural network rule extraction by Using Genetic Programming and its applications to exploratory clasificacions", *IEICE, Trans. Fundamentals*, Vol E88-A, No 20, pp. 2627-2635.

Rosch, Daniel (2004). "An Empirical Comparison of Default Risk Forecasts from Alternative Credit Rating Philosophies *International Journal of Forecasting*, Vol. 21, No. 1, (January-March 2005), pp. 37-51

Rumelhart, D.E., G. E. Hinton, and R.J. Williams, 1986. "Learning internal representations by error propagation," in D. E. Rumelhart and J.L. McClelland, eds. Vol 1, Chapter 8, Cambridge, MA: MIT Press.

Stephen Figlewski, Halina Frydman and Weijian Liang (2008). "Modeling the Effect of Macroeconomic Factors on Corporate Default and Credit Rating Transitions", *Default Risk.com*, disponible en:

URL: http://www.defaultrisk.com/pp_model146.htm [Ultimo acceso Mayo 2009]

Werbos, P.J. 1990. "Backpropagation through time: What it does and how to do it" *Proceedings of the IEEE*, vol. 78, pp. 1550-1560.

Wong B.K. and Y. Selvi, (1998) Neural network applications in finance: A review and analysis of literature (1990–1996), *Information and Management* **34** (3), pp. 129–139.

Wong B.K., T.A. Bodnovich and Y. Selvi, Neural network applications in business: A review and analysis of the literature (1988–1995), *Decision Support Systems* **19** (1997), pp. 301–320

Wooldridge, Jeffrey (2001). "Introductory Econometrics: A modern Approach", Southwestern College Pub., U.S.

Anexo

Tabla A.1. Escala de Calificaciones Otorgada por FitchRatings.

Escala	Categoría	Definición
AAA	1	La más alta calidad crediticia
AA+ ^a	2	
AA	3	Muy alta calidad crediticia
AA-	4	
A+	5	
A	6	Alta calidad crediticia
A-	7	
BBB+	8	
BBB	9	Adecuada calidad crediticia
BBB-	10	
BB+	11	
BB	12	Especulativa
BB-	13	
B+	14	
B	15	Altamente especulativa
B-	16	
CCC	17	Alto riesgo de incumplimiento
CC	18	Muy alto riesgo de incumplimiento
C	19	El más alto riesgo de incumplimiento
D	20	Incumplimiento
E	21	Calificación suspendida

Fuente: Elaborado a partir de información de Fitch Ratings. ^aLos signos + y - indican la fortaleza o posición relativa al interior de las calificaciones que para nuestra muestra de entidades federativas oscila entre B a AA.

