

# LAS REDES NEURONALES Y LA EVALUACIÓN DEL RIESGO DE CRÉDITO\*

---

Fredy Ocaris Pérez Ramírez\*\*  
Horacio Fernández Castaño\*\*\*

Recibido: 01/03/2007

Aceptado: 27/04/2007

## RESUMEN

A pesar del escepticismo del mundo académico sobre los avances de la inteligencia artificial, las redes neuronales han abierto un campo de exploración bursátil que aún tiene mucho por investigar. Atendiendo a las ventajas del uso de las redes neuronales artificiales (ANN, por sus siglas en inglés) y a su capacidad para estimar modelos no lineales, en este artículo se muestra la aplicación de las redes neuronales a la cuantificación del riesgo de crédito. Además, se hace el desarrollo teórico de los fundamentos básicos de las redes neuronales.

Para presentar las metodologías de medición de riesgo de crédito basados en redes neuronales, y aplicarlas a la base de datos de una cartera comercial, fue necesario elaborar un análisis exploratorio de cada una de las variables e investigar la correlación entre ellas. El objetivo del análisis es encontrar algunas relaciones para grupos determinados de la población, de acuerdo con sus características particulares.

Por tanto, se cruzan variables de cada cliente, del crédito y del comportamiento contra la variable default (fallidos y no fallidos). Variable que establece un procedimiento de clasificación, y permite determinar las ponderaciones necesarias y, además, establece la probabilidad de fallido.

**PALABRAS CLAVE:** Riesgo, Comité de Basilea, crédito, redes neuronales

---

\* Este artículo es derivado del proyecto de investigación Clasificación de riesgo en carteras de crédito aplicando metodología de redes neuronales, financiado por la Vicerrectora de Investigaciones de la Universidad de Medellín.

\*\* Matemático de la Universidad de Antioquia, Magíster en Matemáticas Aplicadas de la Universidad EAFIT y Estudios de Especialización en Estadística de la Universidad Nacional sede Medellín. Profesor de tiempo completo, Facultad de Ingenierías. Programa de Ingeniería Financiera. Universidad de Medellín. e-mail: foperez@udem.edu.co

\*\*\* Maestrando en Matemáticas Aplicadas de la Universidad EAFIT, Ingeniero Civil de la Escuela de Ingeniería de Antioquia, Especialista en Sistemas de Administración de la Calidad ISO 9000, Especialista en Gerencia de Construcciones y Licenciado en Matemáticas de la Universidad de Medellín. Profesor de tiempo completo, Facultad de Ingenierías. Programa de Ingeniería Financiera. Universidad de Medellín. e-mail: hfernandez@udem.edu.co

## ABSTRACT

In spite of the skepticism of the academic world on the advances of artificial intelligence, the neuronal networks have opened up a field of stock-exchange exploration that has still so much to research. Upon expounding the advantages of the usage of artificial neuronal network (ANN) and its capacity to estimate nonlinear models this article shows the application of the neuronal networks on the quantification of the credit risk. Furthermore, the article carries out a theoretical development of the basic foundations of neuronal networks.

In order to present the methodologies of measurements of credit risk, based upon the neuronal networks and to apply the to the data base of a commercial portfolio, it became necessary to elaborate an exploratory analysis of each of the variables and to research the correlation amongst them. The objective of the analysis is to pinpoint some relations for predetermined population groups according to their particular characteristics. Therefore, variables of each client, the credit and the behavior against the variable are crossed default (insolvent and non insolvent): a variable that establishes the classification procedure and determines the necessary averages in addition to establishing the probability of insolvency.

**KEY WORDS:** Neuronal risk, Basle Committee, credit, networks.

## INTRODUCCIÓN

Las entidades financieras operan en la actualidad bajo entornos macroeconómicos altamente volátiles, en donde las tasas de interés, tasas de cambio y algunas otras variables (exógenas y endógenas) presentan fluctuaciones considerables y de difícil predicción. El negocio bancario se caracteriza por la transformación de plazos, volúmenes y monedas, y generalmente la mayor parte de las posiciones de sus balances tienen una relación directa con las variables macroeconómicas. Dichas fluctuaciones implican riesgos para la entidad, los cuales deben ser administrados, es decir, debe existir un sistema de administración del riesgo como un procedimiento para identificarlos, medirlos y tomar las medidas preventivas que permitan enfrentarlos de la mejor manera posible.

Las técnicas para medir el riesgo de crédito son hoy en día muy variadas, y abarcan procedimientos que van desde simples cálculos, hasta sofisticadas metodologías con simulaciones dinámicas del futuro más próximo. Estos procedimientos se han

desarrollado tratando de representar cómo varía la capacidad de pago y qué efectos tienen estas variaciones sobre las finanzas de las instituciones.

Con base en la complejidad de las actividades de cada institución, se deben desarrollar modelos internos o implantar modelos estándar, los cuales, a su vez, requieren de sistemas de información apoyados en datos históricos. Es entonces esencial que la institución, además de dar cumplimiento a los requerimientos del Acuerdo de Basilea, reconozca primero qué tan riguroso debe ser el proceso de medición del riesgo de crédito, con base en las políticas internas de la organización. El Comité de Basilea recomienda metodologías que implican tener sistemas de información que permitan conocer la maduración de activos y pasivos y los flujos generados por éstos, con determinada periodicidad.

Básicamente, en este artículo, se expone la aplicación de un modelo de cuantificación del riesgo para una cartera comercial, calculando las probabilidades de impago para establecer el capital

mínimo requerido y las provisiones de cartera, a través de las metodologías de redes neuronales.

La complejidad de las organizaciones y de los mercados, y la rapidez con la que se deben tomar buenas decisiones dentro del sector financiero exigen que sean estudiadas las metodologías que permitan dar solución de manera óptima a un sinnúmero de problemas. En este sentido, la metodología de las redes neuronales es bondadosa en la solución de problemas financieros, mediante su aplicación en una de las necesidades reales dentro del sector, como lo es el cálculo de la probabilidad del incumplimiento de pago dentro de la cartera de cada entidad, más aún, con el dinamismo que han presentado en los últimos años, instrumentos financieros como el préstamo, tasa de interés, entre otros.

La metodología de redes neuronales es muy empleada en la solución de problemas financieros en el ámbito internacional y puede resultar muy útil en el mercado de capitales colombiano debido a la escasez de técnicas de estimación y predicción del comportamiento futuro de las variables más influyentes.

## COMITÉ DE BASILEA

Desde comienzos de la década de los setenta existen nuevas condiciones en el manejo de las entidades financieras, debido básicamente a la alta volatilidad de diversas variables exógenas y endógenas que las afectan. Esta volatilidad se debe, en gran parte, a factores como la globalización de los mercados financieros, expectativas inflacionarias crecientes, aumento en el déficit público y desregulación financiera.

Un riesgo importante que los bancos y las entidades comerciales enfrentan es el riesgo de crédito o el incumplimiento de la contraparte con los acuerdos contractuales. Ese riesgo no solamente se aplica a los créditos sino también a otras exposiciones dentro y fuera del balance. La cuantificación y la

evaluación del riesgo crediticio, en Colombia, se encuentra todavía en su infancia y está creciendo, como consecuencia de las instituciones reguladoras como el Comité de Basilea en el ámbito internacional.

El Banco Internacional de Pagos, con sede en Basilea, ciudad del norte de Suiza, ha sido una institución que desde los años treinta ha servido como centro de discusión de temas financieros. Este banco, por medio de algunos comités conformados por expertos internacionales, formula frecuentemente recomendaciones en temas financieros y bancarios; uno de estos comités es el denominado Comité de Supervisión Bancaria de Basilea, organismo fundado en 1974 y conformado por los presidentes de los bancos centrales y las autoridades de supervisión, en su momento, del grupo de los diez.

El objetivo de este comité es formular estándares de supervisión y recomendaciones de práctica de un buen gobierno de las entidades financieras, y en ningún momento es tener injerencias sobre las regulaciones financieras de los países. Sin embargo, sus recomendaciones son acatadas y generalmente se plasman en legislaciones. El Acuerdo de Basilea parte de unos principios establecidos en cuanto al riesgo de crédito, con el fin de establecer en el ámbito internacional los requisitos para que los bancos cuantifiquen y monitoreen ese riesgo.

En 1988, este comité publicó el denominado Acuerdo de Capitales de Basilea, con el fin de regular el nivel de solvencia y disminuir el riesgo de crédito; es decir, el acuerdo consiste en solicitar el desarrollo de una metodología para obtener un capital mínimo requerido, basado principalmente en el riesgo de crédito. Este acuerdo ha estado sujeto a múltiples reformas; así, en el año 2004, fue publicado el llamado Nuevo Acuerdo de Basilea (CSBB, 2004).

El nuevo acuerdo de capitales es un documento que se basa en tres pilares fundamentales, que son:

**Requisitos de capital mínimo:** esta parte trata sobre el cálculo de los requisitos de capital mínimo

por riesgo crediticio, riesgo de mercado y riesgo operativo a los cuales está expuesta una entidad financiera. Este capital mínimo está compuesto por tres elementos fundamentales: el primero es una definición de capital regulador, el segundo son activos ponderados por riesgo y el tercero es la relación mínima de capital-activos ponderados por riesgo. Este capital mínimo se explica mediante el cálculo de un coeficiente que no debe ser inferior al 8% del capital total.

**Proceso del examen supervisor:** este pilar apunta a la necesidad que tienen los bancos de evaluar sus exposiciones de suficiencia de capital con respecto a sus ingresos globales, así como la de los supervisores para examinar y adoptar las medidas adecuadas como respuesta a dichas evaluaciones. Estos elementos se consideran cada vez más necesarios para la eficiencia en la gestión de las entidades bancarias y la supervisión bancaria, respectivamente.

**Disciplina del mercado:** mediante este pilar se intenta desarrollar un conjunto de requisitos de divulgación que permitan a los participantes en el mercado evaluar las principales informaciones sobre el perfil de riesgo de un banco y sobre su nivel de capitalización. Al proporcionar una mayor disciplina de mercado a través de divulgaciones mejoradas, se pueden aportar ventajas significativas que ayuden a los bancos y supervisores a gestionar sus riesgos y a mejorar la estabilidad.

## METODOLOGÍAS EMPLEADAS PARA LA MEDICIÓN DEL RIESGO CREDITICIO

Los esquemas de medición de riesgo expuestos en el Nuevo Acuerdo de Basilea pueden agruparse en dos familias:

**Los tradicionales o estáticos:** Estos son modelos estándar que determinan el riesgo de crédito a cubrir aplicando un determinado porcentaje; consideran estático e independiente el comportamiento de las variables que intervienen en la estimación, lo cual

dista mucho de la realidad, dado que las variables a tener en cuenta son aleatorias y dependientes de otras variables.

**Los internos o dinámicos:** Estos modelos (conocidos como métodos fundamentados en la calificación interna) fueron desarrollados con mayor propiedad en el nuevo acuerdo, en el cual *se dan facultades a las entidades financieras para crear y diseñar sus propios modelos internos bajo ciertas características*. Estas metodologías permiten realizar un análisis dentro del contexto global de la gestión del riesgo de crédito, utilizando un modelo ajustado a las políticas de riesgo y a la infraestructura de la entidad financiera; igualmente, a las condiciones socioeconómicas y culturales del mercado. Estos modelos permiten descubrir concentraciones de riesgo, y además, realizar una asignación más óptima de los recursos según el riesgo, ya que la inversión crediticia atenderá a criterios objetivos y no a opiniones subjetivas.

Una metodología utilizada para estos modelos es el Internal Rating Based (IRB) que exige a las entidades diseñar un sistema de ratings que les permita clasificar y ordenar a sus acreditados en función de sus grados de riesgo y estimar las variables relevantes para la medición de los riesgos.

## CASO COLOMBIANO

Según la Asobancaria, el entorno macroeconómico en el cual se ha desarrollado el sector financiero colombiano en los últimos años ha estado enmarcado por significativas fluctuaciones, empezando por las crisis financieras de los años 70 y los 80, con aumentos del gasto público y la deuda externa; luego, entrada la segunda mitad de la década de los noventa, el crecimiento y la estabilidad económica mundial propiciaron unas condiciones más favorables. Sin embargo, a partir de 1997 con la crisis asiática y sus efectos en Rusia, las cosas cambiaron radicalmente para buena parte de las economías latinoamericanas.

Llegada la crisis financiera a Colombia, uno de los indicadores más afectados fue la cartera del sistema financiero, con una reducción alrededor de \$8 billones de pesos, un deterioro en el nivel de actividad crediticia cercano al 15%, y, pese a que recientemente esta actividad crediticia ha presentado un ligero repunte, es necesario precisar que el actual ritmo de crecimiento no es suficiente para producir un despegue definitivo del crédito.

En Colombia, la Superintendencia Financiera, a través de sus circulares externas, ha desarrollado instructivos donde se establecen los principios y criterios generales que las entidades vigiladas deben adoptar para mantener adecuadamente evaluados los riesgos crediticios implícitos en los activos de crédito; define las modalidades de crédito, determina las calificaciones que se deben otorgar a tales operaciones según la percepción de riesgo que de las mismas se tenga, establece la periodicidad con que se deben efectuar tales calificaciones, consagra los mecanismos de recalificación, instruye sobre los criterios contables y las provisiones a realizar, sobre el contenido y frecuencia de los reportes a suministrar, y fija los mecanismos de control interno que las entidades deben adoptar para asegurar el adecuado cumplimiento de este instructivo.

La evaluación del riesgo crediticio de cada modalidad de crédito o contrato se realizará de acuerdo con una metodología que fije el respectivo organismo de dirección de la entidad vigilada, atendiendo para ello los parámetros mínimos establecidos. Esta metodología interna debe contemplar la estimación de, por lo menos, dos elementos, a saber:

- a) La probabilidad de deterioro o de cambio en la calificación de riesgo del crédito (probabilidad de no pago o tasa de morosidad esperada).
- b) La estimación o cuantificación de la pérdida esperada en que incurriría la entidad en caso de que se produzca el incumplimiento, durante un horizonte de tiempo determinado (por ejemplo, 12 meses). Para esta estimación es importante, entre otros aspectos, calcular el valor o tasa de

recuperación del valor del activo en el evento de que el crédito se vuelva irrecuperable. La existencia e idoneidad de las garantías que respaldan los créditos son un factor determinante a considerar en este contexto.

## REDES NEURONALES

Entre las definiciones más recientes de inteligencia artificial se expresa, en forma general, la inteligencia artificial como la capacidad que tienen las máquinas para realizar tareas que en el momento son realizadas por seres humanos; otros autores como Nebendah (1988) y Delgado (1998) dan definiciones más completas y las definen como el campo de estudio que se enfoca en la explicación y emulación de la conducta inteligente en función de procesos computacionales basados en la experiencia y el conocimiento continuo del ambiente. Autores como Marr (1977), Mompin (1987), Rolston (1992), en sus definiciones involucran los términos de soluciones a problemas muy complejos.

El nacimiento de la inteligencia artificial se sitúa en los años cincuenta; en esa fecha la informática apenas se había desarrollado, y ya se planteaba la posibilidad de diseñar máquinas inteligentes. Hoy en día se habla de vida artificial, algoritmos genéticos, computación molecular o redes neuronales. En algunas de estas ramas los resultados teóricos van muy por encima de las realizaciones prácticas.

A través de los años, se han utilizado diversas técnicas de inteligencia artificial para emular "comportamientos inteligentes". Al software que hace uso de dichas técnicas se le denomina, de forma genérica, "sistema inteligente", y es cada vez más amplia la gama de aplicaciones financieras donde incide la inteligencia artificial.

Un ejemplo de esto es que al usarse una tarjeta de crédito, suelen acumularse datos sobre patrones de consumo que después se venderán a diversas empresas. Sobre la base de los pagos efectuados en dicha tarjeta de crédito, los bancos e instituciones de crédito irán elaborando un historial del usuario,

el cual se utilizará para autorizar una transacción, para decidir cuándo extender el crédito y para detectar fraudes. Este tipo de procesos requiere de chequeos que suelen resultar bastante complejos, además del uso de criterios variables para poder tomar una decisión final en torno a la autorización de una cierta transacción. Claro que, al manejar enormes volúmenes de información, como los aproximadamente 16 millones de transacciones que Visa Internacional debe verificar diariamente, no resulta nada fácil poder detectar un fraude. Aunque es evidente la necesidad de automatizar procesos como éste, no es del todo obvio incorporar el comportamiento inteligente del ser humano a un programa de computadora que reemplace a un evaluador humano, ya que los sistemas de inteligencia artificial se toman como herramientas de apoyo analítico para el evaluador, mas no como una unidad autosuficiente que por sí sola pueda tomar decisiones.

Las redes neuronales artificiales son eficientes en tareas tales como el reconocimiento de patrones, problemas de optimización o clasificación, y se pueden integrar en un sistema de ayuda a la toma de decisiones, pero no son una panacea capaz de resolver todos los problemas: todo lo contrario, son modelos muy especializados que pueden aplicarse en dominios muy concretos.

## Definición

Las redes neuronales emulan la estructura y el comportamiento del cerebro, utilizando los procesos de aprendizaje para buscar una solución a diferentes problemas; son un conjunto de algoritmos matemáticos que encuentran las relaciones no lineales entre conjuntos de datos; suelen ser utilizadas como herramientas para la predicción de tendencias y como clasificadoras de conjuntos de datos.

Se denominan neuronales porque están basadas en el funcionamiento de una neurona biológica cuando procesa información.

## Las redes supervisadas: El perceptrón multicapa o MLP

Las redes neuronales supervisadas son técnicas para extraer datos a partir de las relaciones de entrada-salida y para almacenar tales relaciones en ecuaciones matemáticas que pueden utilizarse en actividades de pronóstico o en la toma de decisiones. Este tipo de red requiere que el usuario especifique la salida deseada y, por tanto, la red aprende a detectar la relación entre las entradas y las salidas suministradas, mediante un proceso iterativo y adaptativo; una vez que la red se ha entrenado, se puede utilizar con datos que nunca haya visto o puede ser incluida en un programa para el apoyo a las decisiones.

Para que una red neuronal descubra las conexiones no lineales entre dos conjuntos de datos, hay que entrenarla. Para esto se le presentan a la red los datos de entrada y los resultados requeridos por el analista. La red, utilizando de manera reiterada un algoritmo, denominado de aprendizaje, irá modificando los pesos, (que en el inicio de la red tienen un valor aleatorio), una y otra vez hasta encontrar el conjunto de ellos que consigue que con los datos suministrados se obtengan los resultados requeridos.

Una vez que la red está entrenada se le presentan nuevos datos y se hace un test para comprobar la bondad del conjunto de pesos. Cuando la red es analizada y ofrece un rendimiento óptimo, ya está lista para trabajar. A la red se le pueden introducir datos nuevos y ofrecerá resultados del problema en el que está trabajando. A este tipo de redes neuronales se les denomina supervisadas, debido a que al introducir los datos iniciales, la red en la fase de entrenamiento, puede calcular el error que comete y modificar los pesos sinápticos con el objetivo de disminuir este error.

La red neuronal más utilizada se denomina Perceptrón Multicapa o MLP (Multi-Layer Perceptrón). Esta es una red de varias capas, usualmente tres (entrada, oculta y salida) que utiliza funciones sig-

moideas como función de transferencia en la capa oculta. Las funciones de la capa de salida pueden ser lineales o sigmoideas, dependiendo del tipo de salida que se quiera. Pero la característica más importante de la MLP es que utiliza como función de aprendizaje la Retropropagación o regla Back Propagation (BP).

Existen diferentes variantes de la Regla BP que se utilizan en función del problema que se quiere resolver. La regla BP puede trabajar de dos maneras: aprendizaje por lotes o aprendizaje en serie.

El aprendizaje por lotes acumula las variaciones de los pesos, y al final de cada ciclo actualiza, a la vez, todos los pesos. El aprendizaje en serie va a la vez actualizando los pesos cada vez que se presenta un dato; una desventaja del aprendizaje en serie es que se debe respetar el orden de presentación de las entradas, mientras que en el aprendizaje por lotes, el orden no tiene importancia. La ventaja del aprendizaje en serie es su velocidad, mientras que la ventaja del aprendizaje por lotes es que se puede aplicar a la mayoría de los problemas.

Una de las decisiones que debe tomar el analista a la hora de diseñar una red neuronal es elegir un algoritmo de aprendizaje. Sin embargo, se considera difícil elegir el algoritmo adecuado para cada problema, ya que un algoritmo que es rápido y consigue soluciones satisfactorias para un problema no es eficaz para otro, o se vuelve muy lento; en otras ocasiones es preferible un algoritmo que trabaje de manera lenta, debido al resultado que se quiere obtener, y otras veces se necesita rapidez, y la precisión de los datos no es tan importante.

La red debe ser entrenada; para ello se establece que se presenten los resultados del entrenamiento, y el error permisible de los resultados obtenidos

## Redes neuronales no supervisadas

Las redes neuronales no supervisadas son técnicas para clasificar, organizar y visualizar grandes conjuntos de datos. Los mapas autoorganizados

(SOM por sus siglas en inglés) son un ejemplo del enfoque de las redes neuronales no supervisadas. Este enfoque, con origen a comienzos de los años 80, se ha utilizado ampliamente en ingeniería y en muchos otros campos.

Los mapas autoorganizados son un enfoque de red neuronal con alimentación directa que utiliza un algoritmo de entrenamiento no supervisado, que mediante un proceso conocido como autoorganización, configura las unidades de salida en una representación topológica de los datos originales. Los SOM son una técnica de red neuronal que aprende sin supervisión. En contraste con las técnicas de las redes neuronales supervisadas que requieren que se especifique una o más salidas, conjuntamente con una o más entradas a fin de encontrar patrones o relaciones entre los datos, los SOM reducen datos multidimensionales a un mapa de menores dimensiones o una rejilla de neuronas.

El algoritmo de los SOM se basa en el aprendizaje competitivo, proporciona un mapeo de preservación topológica a partir de un espacio multidimensional para mapear las unidades. Las unidades de mapeo o neuronas usualmente forman una rejilla de dos dimensiones y el mapeo es de un espacio multidimensional en un plano o cubo. La propiedad de preservación topológica significa que en los SOM se agrupan vectores de datos de entradas similares en neuronas: los puntos que se encuentran cercanos unos de otros en el espacio de entrada son mapeados en el SOM, en unidades del mapa que son cercanas; por ello el SOM puede servir como una herramienta de agrupamiento (clustering) así como para visualizar datos multidimensionales.

Los mapas autoorganizados consisten de dos capas de unidades de procesamiento: la primera es una capa de entrada que contiene unidades de procesamiento para cada elemento en el vector de entrada; en segundo lugar está la capa de salida o rejilla de unidades de procesamiento, que está completamente conectada con la capa de entrada. El número de unidades de procesamiento en la capa de salida es determinado por el usuario, a partir de la forma y

el tamaño inicial del mapa deseado. A diferencia de otras redes neuronales, no existe ninguna capa ni unidad de procesamiento oculto.

En contraste con los métodos clásicos, los mapas autoorganizados ofrecen una fácil visualización, imponen pocos presupuestos y restricciones, y son capaces de manipular grandes conjuntos de datos para detectar, en éstos, estructuras y patrones aislados. Por ello, los mapas autoorganizados han cobrado creciente interés para el análisis exploratorio de datos y la minería de datos en diferentes campos entre los que se destacan las finanzas y la economía, en donde sobresalen los procedimientos sistemáticos aplicados a la toma de decisiones gerenciales. Se han basado históricamente en medios cuya evolución ha progresado a medida que aumentaba la complejidad de las relaciones empresariales. Con anterioridad a la aplicación del método científico en la resolución de problemas relativos a la actividad económica y empresarial, las decisiones se adoptaban de forma individual por el propio decisor, sobre la base del binomio experiencia-intuición.

## ALGORITMO DE APRENDIZAJE

El aprendizaje o proceso de entrenamiento de una red neuronal se entiende como la etapa mediante la cual la red neuronal adquiere la capacidad de desempeñar las funciones específicas que requiere el problema que se pretende abordar. Los distintos modelos se clasifican de acuerdo con el paradigma básico de aprendizaje que estos usan. Los tres principales paradigmas de aprendizaje son: supervisado, no supervisado y de refuerzo.

El aprendizaje supervisado es el paradigma de entrenamiento más común para realizar tareas de clasificación y aplicaciones de predicción. En este enfoque el algoritmo de aprendizaje ajusta adaptativamente la diferencia entre el output deseado y el que la red neuronal genera. Así, la red aprende a realizar automáticamente una tarea prescrita. Esto requiere que el diseñador de la red conozca

de antemano la respuesta de la red en una muestra de casos que son usados para el entrenamiento.

En cambio, el aprendizaje no supervisado es apropiado para resolver problemas en los que se tiene clara la pregunta pero no se conoce la respuesta. Por ejemplo la pregunta puede ser: ¿Cómo se relacionan los datos? ¿Cuáles datos son similares y en qué consiste esta similitud? ¿Qué patrones o estructuras globales subyacen en la colección de datos? Por lo tanto, el cálculo que las redes neuronales no supervisadas realizan automáticamente termina revelando una organización o estructura final. Por esta razón, este tipo de redes neuronales son llamadas redes “auto-organizadas”. Estas redes son muy utilizadas para realizar tareas de cluster y segmentación jerárquica de datos, que son funciones de gran utilidad para descubrir conocimiento en grandes bases de datos.

Mediante el aprendizaje, los pesos de las conexiones a la red se modifican, y se dice que la red ha aprendido cuando los valores de los pesos permanecen estables o tienen un error mínimo. Los criterios para cambiar el valor de los pesos de conexión son determinados por la regla de aprendizaje. Algunas redes pueden aprender durante su funcionamiento habitual, denominado aprendizaje ON-LINE; en cambio si el aprendizaje requiere la desconexión de la red, hasta que el proceso termine, se denomina aprendizaje OFF-LINE.

Se puede tener también aprendizaje supervisado y no supervisado. La forma más común es el supervisado. En este caso, como se dijo anteriormente, se especifican todas las entradas, así como las salidas. Para cada ejemplo que se propone en la red, se comparan las salidas que producen, con las salidas deseadas. Antes de que el subconjunto entero de ejemplos de entrenamiento haya sido procesado, se han debido actualizar los pesos de conexión de la red. El entrenamiento o la actualización de pesos se hace de manera que se va reduciendo la medida del error.

Si a la red se le proporcionan tanto la entrada como la salida correcta, puede aprender, porque se dedicará a ajustar sus pesos tratando de minimizar el error de su salida calculada. Pero, incluso, la red puede aprender también proporcionándole solo la entrada, siempre que tenga definida una ley adecuada de aprendizaje.

En el aprendizaje supervisado podemos tener varias categorías como:

**Aprendizaje por corrección de error:** se ajustan los pesos de la red, en función de la diferencia entre los valores deseados y los obtenidos en la salida de la red, minimizando el error cometido.

**Aprendizaje por refuerzo:** la función se reduce a indicar mediante una señal de refuerzo si la salida obtenida en la red se ajusta a la deseada, teniendo en cuenta que éxito = +1 y fracaso = -1 y con ello se ajustan los pesos basándose en un mecanismo de probabilidades.

**Aprendizaje estocástico:** se suele hacer de acuerdo con un símil termodinámico, asociando la red con un sólido físico, el cual tiene un cierto estado energético. La energía de la red representará el grado de estabilidad del sistema. El estado de mínima energía corresponderá a una situación en la que los pesos de las conexiones hacen que su funcionamiento sea el que más se ajusta con el objetivo deseado. El aprendizaje consistirá en realizar un cambio aleatorio de los valores de los pesos y en determinar la energía de la red (habitualmente la función de energía es una función de Lyapunov). Si la energía es menor después del cambio, se acepta el cambio. Si, por el contrario, la energía no es menor, se aceptaría el cambio en función de una determinada y preestablecida distribución de probabilidades.

En el aprendizaje no supervisado se tiene igualmente una colección de ejemplos de entrada, pero no se provee a la red de las salidas para esos ejemplos. Se asume típicamente que cada entrada surge de una o varias clases, y las salidas de la red hacen una identificación de la clase a la cual pertenece

su entrada. El proceso de entrenamiento consiste en descubrir las características sobresalientes del conjunto de entrenamiento y usar estas características para agrupar las salidas en clases que la red encuentra distintas.

Otro tipo de entrenamiento, que es un híbrido entre los anteriores, es el aprendizaje por refuerzo, el cual no es supervisado, en el sentido que la forma exacta de las salidas no es especificada.

Si se considera el error presentado en una superficie n-dimensional, como un gran valle geográfico, con hondonadas y sinuosidades unas más bajas que otras y ubicadas a lo largo de una pendiente que va hacia una concavidad máxima, se usa el procedimiento de perturbar los resultados con ruido para "hacer salir" las soluciones parciales de los mínimos locales, lo cual se denomina temple simulado (simulated annealing); cuando se hace su combinación con la asignación probabilística mediante la capa oculta, se le conoce como aprendizaje estocástico.

A continuación se presentan diferentes tipos de mecanismos de aprendizaje:

- a) La regla Widrow-Hoff o LMS. Es muy importante porque es la base de la mayoría de los algoritmos de aprendizaje de un amplio conjunto de redes neuronales. Básicamente la metodología para deducir una regla de aprendizaje de una red neuronal se resume en
  1. Definición de la función de error, también denominada función de costo. Mide la bondad del modelo. Cuanto más pequeño sea el error, más eficiente es el modelo. Depende de los pesos sinápticos, que son las incógnitas a resolver en el entrenamiento.
  2. Optimización de la función de error. Se busca un conjunto de pesos sinápticos que minimice la función de error. Esta búsqueda se realiza mediante un proceso iterativo denominado Descenso por el Gradiente.

b) Regla Back Propagation. Este tipo de red neuronal se aplica en variados problemas de predicción, y funciona especialmente en los casos en los que se presentan datos de entrada con ruido y en los que se requiere que una señal de salida tenga una alta aproximación con valores reales, o se requiera traducir señales de entrada obteniendo otras de salida. La red de propagación hacia atrás, back propagation, está diseñada para funcionar como una red multicapa, con propagación hacia delante, empleando modo supervisado de aprendizaje.

**Redes y mapas auto-organizados:** las redes que ocupan un paradigma de entrenamiento no supervisado se basan en un enfoque competitivo: durante el proceso de entrenamiento, cada uno de las neuronas o nodos de la red compiten entre sí para determinar cuál de ellos tiene mayor capacidad de representar cada uno de los datos que se le van presentando a la red. Esta competencia va generando progresivamente una reorganización (interna) de las neuronas que finalmente puede ser representada en diversos tipos de mapas bidimensionales que permiten revelar visualmente la organización global (del conjunto de datos de entrada) encontrada por la red. Dentro de las redes neuronales de este tipo más usadas se encuentran: “Adaptive Resonance Theory” (ART) y los “Features Maps” de Kohonen.

## APLICACIONES DE LAS REDES NEURONALES EN EL RIESGO DE CRÉDITO

**Estudios de clasificación.** En los problemas de clasificación se trata de asignar a cada caso su clase correspondiente, a partir de un conjunto de ejemplos: abarca problemas como el estudio del fracaso empresarial, la concesión de préstamos, la calificación de obligaciones, etc.

**Concesión de préstamos.** Metodológicamente son muy similares los trabajos de concesión de préstamos a los de predicción de la quiebra. El conceder o no un préstamo es también una deci-

sión no estructurada y la diferencia radica en que la información disponible no se extrae de bases de datos comerciales o de los registros, sino que es información que suministra el propio banco o entidad financiera que encarga el estudio. En este caso la información no se compone exclusivamente de ratios financieros, sino que también hay otro tipo de datos, como por ejemplo quién es el director de la compañía, si es un cliente nuevo, etc.

Una aplicación del Chase Manhattan Bank para la concesión de préstamos, es un sistema mixto que incorpora herramientas estadísticas y un perceptrón multicapa. El Chase Manhattan Bank concede préstamos nuevos cada año por valor de trescientos millones de dólares e hizo una apuesta muy fuerte en el desarrollo de un modelo informatizado basado en el reconocimiento de patrones para las decisiones de concesión de préstamos. El programa se encuentra en un ordenador central al que los usuarios pueden acceder desde ordenadores compatibles, vía módem. Lógicamente, algunos procedimientos se encuentran patentados por lo que es difícil conocer todos los módulos que integran el sistema. PCLM, que son las siglas de Public Company Loan Model, proporciona extensos informes, gráficos, puntos fuertes y débiles de la compañía, así como una clasificación de las empresas en buenas, malas y críticas.

En Colombia se conoce el desarrollo de un trabajo para una importante corporación de ahorro, en el cual empleando la metodología de redes neuronales se construyó un Scoring de crédito.

## VENTAJAS DE LAS REDES NEURONALES

En general, las redes neuronales son superiores a otras técnicas porque:

- Se entrenan, auto organizan, aprenden y olvidan.
- Son robustas y tolerantes a fallas; la falla de una o varias neuronas no implica un fallo total en la red neuronal.

- Son flexibles, lo que les permite adaptarse fácilmente a nuevos ambientes, ya que pueden catalogarse como sistemas independientes.
- Se emplean en datos en los cuales el patrón es oscuro e imperceptible, que exhiben comportamiento impredecible o no lineal, como en modelos tradicionales de series temporales y datos caóticos.
- La velocidad de respuesta es menor que la del cerebro humano.
- Son hábiles en el proceso de asociar, evaluar o reconocer patrones

## DESVENTAJAS DE LAS REDES NEURONALES

- Funcionan como una caja negra, resuelven un problema, pero es difícil saber cómo lo han hecho. ¿Cómo hace, por ejemplo, una red neuronal un algoritmo de scoring?. En lugar de ser un sistema de apoyo a la decisión, la caja negra se puede convertir en el "tomador" de la decisión. Puede ocurrir que un director de riesgo niegue un crédito sólo porque se lo dice la caja negra, sin que él pueda argumentar esta decisión ya que no entiende el funcionamiento de la red neuronal.

En este contexto, el problema de quién controla una red neuronal no siempre tiene una respuesta fácil. Un caso interesante, al respecto, es un proyecto llevado a cabo por el Pentágono. El aprendizaje de la red neuronal consistía en distinguir entre tanques "amigos" y "enemigos". El aprendizaje de la red fue una cuestión fácil que consistía en introducir datos sobre los diferentes tanques. Para los tanques de la OTAN tenían acceso a especificaciones detalladas de los fabricantes, mientras que para los tanques "enemigos" se introdujeron fotos y diapositivas. Al probar el sistema en la realidad, éste fracasó por completo, porque estaba basado en el "color verde": hay color verde en el entorno, entonces debe ser un tanque enemigo. Analizar y explicar este tipo de toma de decisiones no es fácil

para el caso de las redes neuronales, donde las decisiones son generadas por una caja negra.

Con los métodos tradicionales, como las tablas de puntuación de un sistema de scoring, es más fácil analizar y explicar el funcionamiento de un sistema de apoyo a la decisión.

No obstante, este punto débil puede convertirse en un factor positivo. Por ejemplo, en el caso de los modelos para predecir el fraude en solicitudes de crédito, cuando es importante minimizar la facilidad del sistema para ser manipulado por un solicitante. Cuanto más difícil es de entender la lógica del sistema, más difícil es su manipulación.

- No resuelven todos los problemas ni siempre los resuelven de la mejor manera.
- Tienen problemas en la estimación de cálculos precisos. Funcionan bien con problemas complejos de difícil cálculo pero que no requieren respuestas perfectas, sino solo respuestas rápidas y buenas, como ocurre en la bolsa, en donde se necesita saber con rapidez si conviene comprar, vender o mantener.
- En predicción bursátil se emplean con frecuencia en boletines divulgativos de las entidades financieras, lo que significa que su éxito es relativo, puesto que en otro caso su predicción no se publicaría, sino que se aprovecharía.
- Las redes neuronales se están comportando bien en predicción a largo plazo con componentes no lineales; en cambio, no están claras las mejoras observadas en series cortas y estacionales, como son típicas en las predicciones de ventas.

## APLICACIÓN

### DESCRIPCIÓN

Un aspecto muy importante sobre el cual enfatiza el Sistema de Administración del Riesgo de Crédito (SARC) es el seguimiento y control de procesos que

tengan relación directa con el riesgo crediticio. Por lo tanto, se hace necesario el monitoreo de procesos tales como: otorgamientos, comportamiento y provisiones. Estos procesos sintetizan las diferentes etapas de la vida de una obligación, razón por la cual las variables contempladas en cada uno deben tener relación directa con el objeto mismo del crédito, así como el análisis y seguimiento a las mismas.

Desde hace algunos años en el país se ha estado estudiando la forma de desarrollar el SARC, el cual implica menor regulación y mayor supervisión. Esto incluye un aumento en la capacitación tanto de las entidades vigiladas como de la propia Superintendencia Financiera para mejorar los procesos de evaluación. El objetivo principal de las entidades financieras es lograr por medio del SARC estimar el nivel adecuado de provisiones para el crédito otorgado.

Para desarrollar la aplicación, se consideraron 63 variables, distribuidas entre cuantitativas y cualitativas, correspondientes a una cartera comercial.

El estudio se inicia analizando el nivel de correlación que existe entre dichas variables, debido a que pueden existir problemas de multicolinealidad. Para evitar estos problemas se realizó el análisis de componentes principales el cual permitió pasar de un espacio linealmente dependiente (LD) a un espacio linealmente independiente (LI), permitiendo estimar el modelo de redes neuronales. Para este trabajo, después de evaluar el comportamiento de varias redes neuronales, se determinó trabajar con la red neuronal probabilística.

## RED NEURONAL PROBABILISTICA

La red neuronal probabilística, acorde con nuestro problema, debe encargarse de efectuar una clasificación de los vectores de entrada en dos clases, cuyas regiones de decisión se acercan asintóticamente a las determinadas por el clasifi-

cador óptimo de Bayes. La primera capa cumple la tarea de distribuir los valores de entrada, y está formada por unidades denominadas unidades de vectores. Hay un nodo de éstos por cada vector de entrenamiento. Cada unidad de vector realiza un producto escalar del vector de entrada  $x$  con un vector de pesos  $w_i$ , dado por  $z_i = x \cdot w_i$ . La función de activación de estos nodos está dada por  $\exp\left(\frac{z_i - 1}{\sigma^2}\right)$ .

En esta capa existe una unidad que cumple la tarea de sumar las salidas de las unidades de vectores de la clase respectiva. Al sumar dichas salidas se está haciendo una “aproximación” a la función de densidad estimada correspondiente a cada clase. La última capa contiene un nodo, el cual se encarga de implementar la regla de decisión.

Este tipo de red podemos usarla para problemas de clasificación. Cuando presentamos una entrada, la primera capa calcula distancias desde el vector de entrada hasta los vectores de entrenamiento de entrada, y produce un vector cuyos elementos indican qué tan próxima está la entrada a una entrada de entrenamiento. La segunda capa suma estas contribuciones para que cada clase de entradas produzca la salida de la red como un vector de probabilidades. Finalmente, una función de transferencia competitiva ubicada en la salida de la segunda capa escoge el máximo de estas probabilidades, y produce un 1 para esa clase y un 0 para las otras clases. Conviene mencionar que para un vector de entrenamiento, las clases se conocen de antemano.

## ARQUITECTURA DE LA RED

Se trata de considerar que existen  $Q$  pares de vector de entrada v.s vector deseado. Cada uno de los vectores deseados tiene  $K$  elementos. Uno de esos elementos es uno y el resto son ceros. Entonces, cada vector de entrada está asociado con una de las  $K$  clases. La figura 1 muestra los componentes de la red.

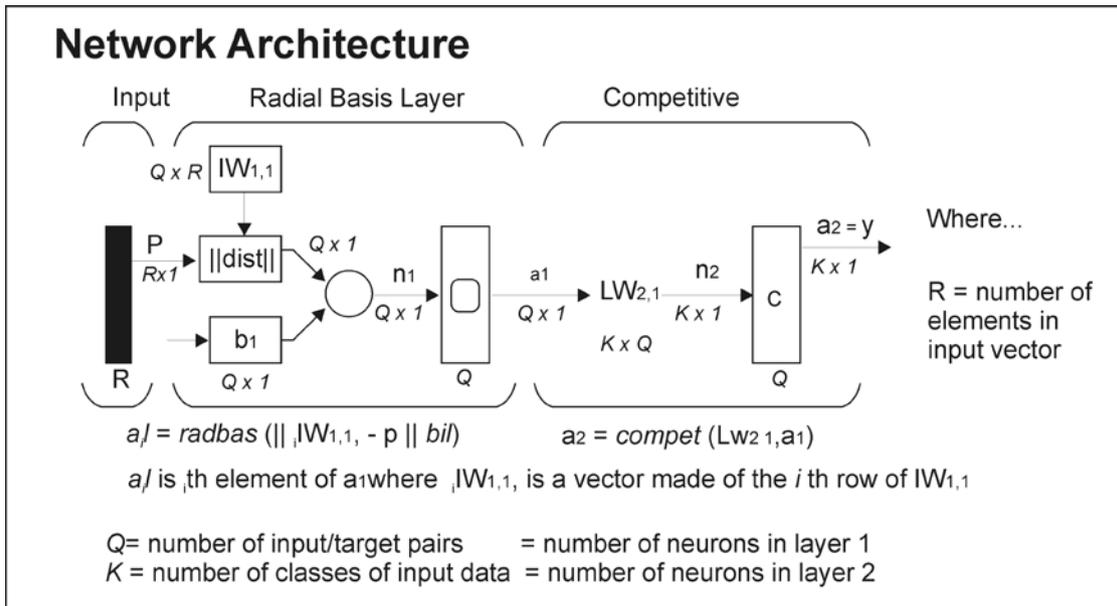


Figura 1 Componentes de la red

**$IW_{1,1}$  (net.IW {1,1}): Pesos de entrada de la primera capa:** los cuales se fijan a la transpuesta de la matriz formada de los  $Q$  pares de entrenamiento denotada por  $P$ . Cuando se presenta una entrada, la caja  $||dist||$  produce un vector cuyos elementos indican cuán próxima se encuentra la entrada de los vectores del conjunto de entrenamiento. Estos elementos se multiplican, elemento por elemento, y junto con el *bias*, se envían a la función de transferencia radbas. Un vector próximo a un vector de entrenamiento se representa por un número próximo a 1 en el vector de salida  $a_1$ . Si una entrada está próxima a varios vectores de entrenamiento de una sola clase, ella es representada mediante varios elementos de  $a_1$ , los cuales se encuentran próximos a 1.

**FUNCIÓN DE TRANSFERENCIA RADBAS**

La función de transferencia que se muestra en la figura 2 se encarga de calcular la salida de una capa a partir de sus entradas a la red.

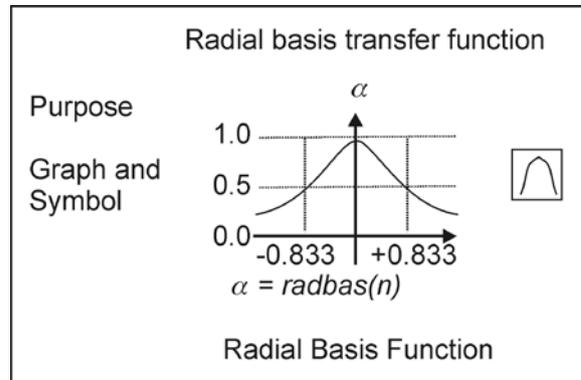


Figura 2. Función de transferencia RADBAS

**$LW_{2,1}$  (net.LW {2, 1}): La segunda capa de pesos.** Se asignan a la matriz  $T$  de vectores deseados. Cada vector tiene un *uno* solamente en la fila asociada con la clase particular de entrada, y *ceros* en otra parte. La función **ind2vec** se usa para crear los vectores apropiados. La multiplicación  $T \cdot a_1$  suma los elementos de  $a_1$  debido a cada una de las  $n$  clases de entradas. Finalmente, la función de transferencia de la segunda capa competitiva, produce un uno correspondiendo al elemento más grande de  $n_2$ , y ceros en otra parte.

De esta manera, la red ha clasificado el vector de la entrada en una clase específica de  $K$ , porque esa clase tenía la probabilidad máxima de ser correcta.

## CONSTRUCCIÓN DE LA RED EN MATLAB

1. Se importa la base de datos de Excel a MATLAB (4325 filas y 25 columnas), la cual fue capturada con la variable **data**.
2. **DISEÑO**: Se usa la función **newpnn** para crear una red neuronal probabilística.
3. Se ingresan los vectores de entrada:  $P=[data(1:4325,1:23)]'$  ;
4. Se ingresan los vectores deseados, la variable default (TARGETS):  $Tc=[data(1:4325,25)]$ ;
5. Se requiere de una matriz **target** con 1's en el lugar correcto. Esto es posible gracias a la función **ind2vec**. Da una matriz con 0's excepto en los puntos o sitios correctos. Para que pueda ejecutarse se hace  $T = ind2vec(Tc)$
6. El resultado del proceso anterior permite obtener las características ampliamente discriminantes, en las regiones de probabilidad de incumplimiento y probabilidad de no incumplimiento.
7. En este punto del procedimiento, se está en condiciones de crear la red para luego simularla. Para lograrlo se usó la entrada **P**:  $net = newpnn(P, T)$ , lo que garantiza obtener las clasificaciones correctas.
8. Se simula la red, haciendo  $Y = sim(net,P)$
9. Para obtener las clasificaciones correctas, se usó la función **vec2ind**, la cual convierte la salida **Y** en una fila **Yc**, haciendo  $Yc = vec2ind(Y)$
10. Se comparan los resultados, **Yc**, con la variable default **Tc**, para la cual se escribe en MATLAB la instrucción  $[Tc'; Yc]$ .
11. Coincidencia entre los resultados originales ( $Tc'$ ) y los de la red neuronal (**Yc**). Para obtener es-

tos resultados se empleó la siguiente instrucción de MATLAB: **COINCIDENCIA=find(Tc'==Yc)**. La coincidencia fue adecuada.

12. Todo este tratamiento se encuentra en el apéndice 1.
13. Visualización de la red neuronal. Estando en la ventana de comando de MATLAB, se da la instrucción **gensim (net)**.

## CONCLUSIONES

1. En el desarrollo del modelo se utilizó primero un perceptrón simple, el cual clasificó los buenos como buenos y los malos como malos con un porcentaje de aciertos del 43.22%. Luego se utilizó un perceptrón de dos capas con un porcentaje de aciertos del 94.19% y por último se trabajó con una red neuronal probabilística, siendo esta última la más adecuada para resolver el problema de clasificación, ya que arrojó el más alto porcentaje de aciertos.
2. La red neuronal probabilística es adecuada para resolver el problema planteado porque:
  - Dispone de los elementos necesarios para distribuir adecuadamente los datos de entrada; operación que ejecuta con eficiencia.
  - Contiene la estrategia de clasificación apropiada para minimizar el riesgo esperado.
  - Posee la habilidad para estimar las densidades de probabilidad de los vectores en las clases, desarrollada a partir de los vectores de entrenamiento y para los cuales la clasificación es conocida.
  - Permite discriminar ampliamente las regiones de probabilidad de incumplimiento y la probabilidad de no incumplimiento (regionalización del espacio de las clases).

## Apéndice I

% Solución del problema de crédito institucional: puntaje de clasificación (aprobados y no aprobados).

% La base de datos entregada originalmente posee las dimensiones siguiente: 4331 filas y 23 columnas

.

% Filas=número de clientes y columnas= variables de entrada.

% Los puntajes de clasificación (base de datos cartera comercial-spss):x36-default (ceros y unos).

% También se adicionó otra columna (los mismos puntajes pero convirtiendo los ceros en dos).

% Red neuronal probabilística: la red con la que se resolvió con acierto total el problema.

% Ingresar los datos de entrada.

p=[data(1:4325,1:23)];

% Ingresar sus correspondientes valores deseados (targets)

tc=[data(1:4325,25)];

% Ingresar la matriz con 1's en el lugar que es correcto

t=ind2vec(tc)

% Luego se creó la red correspondiente

net =newpnn(p,t);

% Luego la simulamos para que produzca las clasificaciones correctas

y=sim(net,p)

% Ahora se convierten las salidas (y) en las filas que hacen las clasificaciones correctas (y<sub>c</sub>)

y<sub>c</sub>=vec2ind(y)

% Los resultados nos muestran que la coincidencia se da entre todos los valores(4325):100% de efectividad

% Visualización de la red neuronal

gensim(net)

## BIBLIOGRAFÍA

- CSBB. 2004. (Comité de Supervisión Bancaria de Basilea). International Convergence of Capital Measurements and Capital Standards. A Revised Framework, Bank for International Settlements, Basel, June.
- DELGADO, A. 1998. Inteligencia artificial y mini robots. Segunda Edición. Ecoe Ediciones.
- MARR, D. C. 1977. Artificial Intelligence: a Personal View, Artificial Intelligence. EEUU.
- MOMPIN, J. 1987. Inteligencia Artificial: Conceptos, Técnicas y aplicaciones. Marcomobo. S. A Ediciones. España.
- NEBENDAH, D. 1988. Sistemas expertos. Ingeniería y comunicación. Editores Marcombo. Barcelona.
- ROLSTON, D. 1992. Principios de inteligencia artificial y sistemas expertos. Mc Graw Hill. México.