

Red neuronal para clasificación de riesgo en cooperativas de ahorro y crédito

Mario Peña^a, José Orellana^b

^a Grupo de investigación de Ingeniería Industrial/Ciencias Químicas/Ingeniería Industrial, Universidad de Cuenca

^b Ciencias Químicas/Ingeniería Industrial, Universidad de Cuenca

mario.pena@ucuenca.edu.ec, jose.orellana0106@gmail.com

Resumen—En Ecuador existe una gran cantidad de cooperativas de ahorro y crédito (COAC), específicamente un total de 852, las cuales están divididas en 5 segmentos en función de la cantidad de sus activos. En la actualidad el 66% del microcrédito obtenido dentro del país corresponde al sistema cooperativo. Sin embargo, sólo 35 de 58 COAC (segmento 1 y 2) presentan historiales de calificaciones de riesgo. El propósito de esta investigación es crear una red neuronal que presente un porcentaje de precisión aceptable, para clasificar una COAC dentro de una escala de riesgo con base al valor de sus índices financieros; ayudando a la detección temprana de problemas futuros. La red neuronal artificial (ANN, del inglés *Artificial Neural Network*) fue ajustada a partir de datos obtenidos a través de la Superintendencia de Economía Pública y Solidaria para COAC que presentaban un alto índice de activos. Adicionalmente, se usó el historial de calificaciones de riesgo trimestrales generadas por entes calificadoras en el mismo lapso: enero 2015 - septiembre 2017. Se obtuvo una ANN con una precisión de clasificación del 79.59%, porcentaje que está dentro del rango de precisión obtenido por estudios revisados para actividades de clasificación en entidades financieras. Se podría mejorar aún más los resultados de clasificación con la utilización de una estructura de clasificación jerárquica.

Palabras Claves— Cooperativas de ahorro y crédito, indicadores financieros, perceptrón multicapa, red neuronal artificial.

Abstract—In Ecuador exists a great number of credit unions (COAC) specifically 852, which are divided into 5 segments depending on their amount of assets. Nowadays, 66% of the microcredit obtained within the country corresponds to the cooperative system. Nevertheless, just 35 of 58 COAC (segments 1 and 2) present risk rating histories. The purpose of this research is create a neural network that achieves an acceptable percentage of accuracy, to classify a COAC within a scale of risk based on the value of its financial indexes; in order to helping the early detection of future problems. The artificial neural network (ANN) was fitted from data obtained through the Public and solidary economy Superintendence for COAC that presented a high index of assets. In addition, the history of quarterly risk ratings generated by rating agencies in the same period was used: January 2015 - September 2017. An ANN with a classification accuracy of 79.59% was obtained, percentage that is within the range of precision obtained by studies reviewed for classification activities in financial entities. The classification results could be further improved with the use of a hierarchical classification structure.

Keywords— Credit union, financial ratios, multilayer perceptron, neural network.

I. INTRODUCCIÓN

Desde los años noventa la tecnología de redes neuronales artificiales ha comenzado a ser usado para pronósticos de bancarrota en entidades financieras [1], enfatizándose en la

predicción y el análisis de información clave para presentar posibles escenarios futuros.

Las redes neuronales artificiales simulan el funcionamiento de redes neuronales biológicas presentes en el cerebro humano, permitiendo aprender a partir de experiencias [2]. El uso de redes neuronales artificiales ha proliferado en diversos campos como: cinética química [3], predicciones de reactores [4], control de sustancias ilegales [5], optimizaciones [6], clasificación de productos de agricultura [7], determinación de especies animales [8], en el campo de la medicina y diagnósticos médicos [3], entre otros, expandiendo su uso a diferentes negocios, comercios e industrias; especialmente, ha comenzado a usarse junto a herramientas tradicionales tales como: regresión logística y análisis discriminante [9]. El impacto de la era tecnológica e informática ha obligado a las organizaciones a innovar en sus métodos de análisis de datos [10]. Además, el uso de redes neuronales presenta la posibilidad de trabajar con relaciones y sistemas no lineales [11].

Las redes neuronales son usadas para la simulación mediante el aprendizaje del reconocimiento de patrones presentes en la muestra. Una red neuronal entrenada, sería capaz de buscar patrones en grupos de datos y mostrar posibles escenarios de comportamiento con base a ciertas variables de entrada [12]. Por lo cual, uno de los requerimientos para el uso adecuado de una red neuronal es disponer de una gran cantidad de datos, para cubrir todos los escenarios posibles.

Una red neuronal, específicamente un perceptrón multicapa, se compone de varias capas generalmente divididos en tres grupos: una capa de entrada, una o varias capas ocultas y una capa de salida, además consta de pesos entre las conexiones, una función de activación, un valor para el sesgo (*bias*), y generalmente para el ajuste de los pesos, se realiza el cálculo del error entre la salida deseada y la real, ajustando los pesos de manera inversa (*backpropagation*).

La facilidad con la que estas redes pueden ser integradas al entorno empresarial, incluyendo pequeñas y medianas compañías [13] es una ventaja considerable para su uso a nivel organizacional. Las redes neuronales pueden ser diseñadas para varios entornos ayudando a la gestión y toma de decisiones; sin embargo, una clase de compañías con más riesgos son las cooperativas de crédito, debido a que manejan ingentes cantidades de dinero, razón por la cual se han creado instituciones especiales para su control en cada país, como, por ejemplo, en Ecuador la Superintendencia de Economía Pública y Solidaria (SEPS) la cual define políticas y reglamentos para las COAC [14].

El propósito de este estudio se centra en el uso de una red

neuronal artificial (perceptrón multicapa) para clasificación de riesgo usando indicadores financieros, pudiendo determinar problemas financieros presentes y futuros en la organización. El estudio se centra en las cooperativas de ahorro y crédito de Ecuador, usando datos desde enero 2015 hasta septiembre 2017.

El presente trabajo está dividido en IV secciones. La sección II presenta el método utilizado para definir los parámetros y la estructura general de la red neuronal. La sección III presenta los resultados obtenidos. Finalmente, la sección IV contiene las conclusiones del presente estudio.

II. MÉTODO

A. Recopilación de datos y establecimiento de las entradas y salidas de la red neuronal.

Para este estudio se usaron indicadores financieros obtenidos de la página oficial de la SEPS, la misma que cuenta con una publicación mensual de estados financieros de las COAC del segmento 1 y 2, titulada “Boletines financieros mensuales”. Los sectores 1 y 2 están integrados por COAC con mayor valor de activos de acuerdo a la norma de segmentación de entidades financieras del sector público y solidario (Tabla I).

TABLA I
VALORES PARA SEGMENTACIÓN DE ENTIDADES FINANCIERAS

Segmento	Activos (USD)
1	Mayor a 80'000.000,00
2	Mayor a 20'000.000,00 hasta 80'000.000,00
3	Mayor a 5'000.000,00 hasta 20'000.000,00
4	Mayor a 1'000.000,00 hasta 5'000.000,00
5	Hasta 1'000.000,00. Cajas de Ahorro, bancos comunales y cajas comunales

En Ecuador existen un total de 852 COAC activas, de las cuales, la existencia de registros de calificaciones de riesgo está concentrado en los segmentos 1 y 2, siendo casi inexistentes para los demás segmentos (Tabla II).

TABLA II
COACTIVAS EN ECUADOR (2016)

Segmento	Número total de COAC	Número de COAC con calificación de riesgo
1	24	24
2	34	11
3	87	5
4	188	0
5	519	0

Para la creación de la red neuronal artificial (ANN del inglés *Artificial Neural Network*), se definió las entradas (*inputs*) a través de investigaciones bibliográficas y diálogos con expertos. Las entradas seleccionadas son indicadores financieros usados para conocer la situación interna de la organización, los cuales permiten conocer el estado real de la misma, además de permitir realizar comparaciones (benchmarking) con las demás organizaciones del sector. Se definieron un total de 9 indicadores (Tabla III).

TABLA III
INDICADORES FINANCIEROS SELECCIONADOS

Indicador	Fórmula
1. Suficiencia patrimonial	(Patrimonio + Resultados) / Activos inmovilizados
2. Estructura y calidad de estructura de activos	Activos productivos / Total activos
3. Cobertura de provisiones	Cobertura de la cartera problemática
4. Eficiencia microeconómica	Gastos de operación / Margen financiero
5. ROE	Resultados del ejercicio / Patrimonio promedio
6. ROA	Resultados del ejercicio / Activo promedio
7. Margen eficiencia financiera	Margen de intermediación estimado / Patrimonio promedio
8. Rendimiento de cartera	Cartera por vencer total
9. Liquidez	Fondos disponibles / total depósitos a corto plazo

Los elementos de clasificación (*outputs*), están formados por las categorías o grupos de calificación que se otorga a las entidades financieras (Tabla IV). La clasificación de riesgo obedece a una escala de 7 grupos de acuerdo con un análisis financiero y organizacional; además, cada nivel puede tener un grado de solidez representado por “+” o “-”. Para el estudio propuesto sólo se realizó la clasificación de grupos independientemente del grado de solidez.

Para realizar el estudio, se ha usado información disponible de 35 cooperativas de ahorro y crédito, pertenecientes a los segmentos 1 y 2, las cuales han recibido calificaciones de riesgo dentro del periodo marzo 2015 – septiembre 2017.

TABLA IV
ESCALA DE CALIFICACIÓN DE RIESGO PARA ENTIDADES FINANCIERAS (SCORE DE RIESGO)

Riesgo de crédito	
AAA	Excelente
AA	Muy buena
A	Buena
BBB	Suficiente
BB	Muy baja
B	Vulnerable
C	Extremadamente vulnerables

La calificación de riesgo se realiza en cuatro fechas de corte al año: marzo, junio, septiembre y diciembre. Esta clasificación dada por entidades acreditadas, depende del valor de los índices financieros y de factores organizacionales. Sin embargo, para este estudio sólo se tomaron en cuenta índices financieros. El rango de tiempo y el número de muestras fueron definidos en función a la disponibilidad de datos y registros. El total de muestras es el conteo de calificaciones disponibles de las COAC respecto a la correspondiente categoría en todo el periodo de análisis (Tabla V). Dentro del rango de tiempo definido para el estudio, algunas cooperativas habían recibido la calificación de riesgo de manera consecutiva (en todas las fechas de corte), mientras otras habían recibido de forma no consecutivas (saltándose fechas de corte o suspendiendo el servicio en los últimos periodos del periodo de análisis).

TABLA V
CANTIDAD DE MUESTRAS USADAS EN EL ESTUDIO

Categoría	Total de muestras
AA	33
A	154
BBB	54
BB	72
B	13

B. Estructura de la red neuronal.

Los parámetros de la red que fueron usados para el presente estudio se definieron a partir de la revisión bibliográfica de documentación relacionada al uso de redes neuronales para entidades financieras (bancos), especialmente como medio de detección de bancarrota y fraude. Luego se seleccionó y se creó una base de datos que contenga las variables de entrada (Tabla III) y salida (Tabla IV), junto con sus respectivos valores para cada categoría.

Los parámetros para la ANN se definieron mediante la variación progresiva de los mismos hasta encontrar la combinación óptima que proporcione la mayor precisión de clasificación. Este método es conocido como ajuste aleatorio. A pesar de que existen métodos propuestos para la correcta definición de parámetros (como el algoritmo genético, método Taguchi) [15], el método de variación aleatoria ha sido usado en los estudios revisados. De este modo, el ratio de aprendizaje, el cual determina que tanto cambia los pesos en cada paso (step), varió desde 0.1: 0.1: 0.5; con un valor final de 0.3. El momento, el cual añade una fracción del peso previo previniendo el máximo local, varió desde 0.1: 0.1: 0.9; con un valor final de 0.2. Para determinar el número de nodos en la capa oculta se varió la cantidad de nodos desde 5: 1: 18; creándose un total de dos capas ocultas con 15 nodos cada una.

C. Desarrollo del modelo.

Para la creación de la ANN, se realizó la separación de muestras para la etapa de entrenamiento y prueba; la separación se realizó en una proporción de 70% para entrenamiento y 30% para prueba. En la separación de muestras se usó la condición de muestras aleatorias (*shuffled sampling*), el cual escoge muestras con el fin de crear los subgrupos aleatorios. Luego, se procedió a normalizar la muestra de entrenamiento usando el método de Z-transformación, del cual se obtuvieron los valores de media y desviación estándar que fueron usados luego para normalizar la muestra de prueba. Finalmente se aplicó el modelo (Fig. 1).

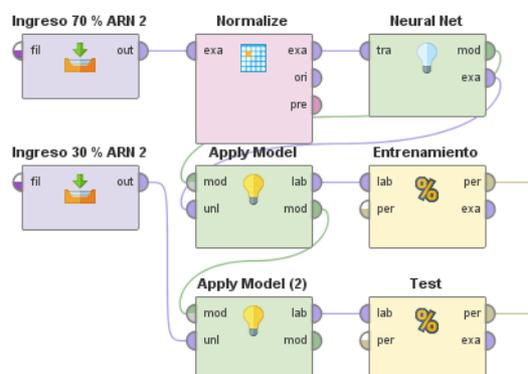


Fig. 1. ANN para clasificación

Para el desarrollo del modelo se usó el software RapidMiner Studio versión 8.1, el cual incluye herramientas para análisis y clasificación de datos.

III. ANÁLISIS DE RESULTADOS

La construcción de la red neuronal se realizó en dos etapas, primero, el entrenamiento, donde la red neuronal debe ajustar los pesos de las conexiones de cada neurona dependiendo de los resultados obtenidos y del error de clasificación; y segundo, la prueba, donde se obtiene el porcentaje de precisión de clasificación de muestras “desconocidas” para la red. La matriz de confusión presenta la clasificación para las muestras de prueba (Tabla VI).

TABLA VI
MATRIZ DE CONFUSIÓN

		Actual Class					Precisión
		AA	A	BBB	BB	B	
Predicted Class	AA	5	3	0	0	0	63%
	A	7	41	2	0	0	82%
	BBB	0	2	10	1	0	77%
	BB	0	3	1	21	0	84%
	B	0	0	0	1	1	50%
Recall		42%	84%	77%	91%	100%	

El rendimiento de clasificación obtenido por la ANN es 79.59%. Aunque no se han encontrado estudios que realicen el mismo tipo de clasificación, el valor de precisión de este trabajo está dentro del rango de resultados obtenidos por estudios realizados para detección de bancarrota y fraude mediante ANN. De un total de 5 estudios revisados, el mínimo valor obtenido de clasificación es de 78% [16] y el máximo de 93.58% [17]. Además, la precisión de clasificación pudo verse afectada por nivel de correlación entre variables, el grado de separabilidad lineal de los datos y el nivel de ruido, de la misma manera que concluyó el estudio realizado por Philippe du Jardin [17].

IV. CONCLUSIONES

En el presente estudio se ha demostrado la posibilidad de usar ANN para actividades de clasificación de riesgo dentro de cooperativas de ahorro y crédito (COAC). Actualmente, esta actividad se realiza mediante *outsourcing*; sin embargo, el presente estudio ha explorado el uso de una ANN para tal actividad a partir de indicadores financieros, obteniéndose una ANN con 79.59% de precisión, el cual está dentro del rango de precisión obtenido por estudios similares para clasificación de bancarrota y fraude.

La ANN obtenida podría proporcionar alertas tempranas sobre el desenvolvimiento financiero actual y predecir posibles problemas institucionales. Además, debido a que solo 35 de 852 COAC cuentan con la calificación de riesgo, existe una amplia cantidad de COAC que pueden usar la metodología de redes neuronales para conocer su grado de riesgo, ayudándolas a ofrecer una mayor solidez en sus servicios, además de incentivar a organizaciones de todo tamaño a implementar nuevas herramientas para actividades de control y apoyo en la toma de decisiones.

Aunque la clasificación de riesgo involucra elementos financieros y organizaciones, el estudio se ha centrado solo en la parte financiera, debido a que, es un eje transversal entre las distintas entidades del sector financiero.

AGRADECIMIENTOS

Agradecemos al grupo de investigación de la carrera de Ingeniería Industrial de la Facultad de Ciencias Químicas de la Universidad de Cuenca por su apoyo en la realización de esta investigación.

REFERENCIAS

- [1] Q. Yu, Y. Miche, E. Séverin and A. Lendasse, "Bankruptcy prediction using Extreme Learning Machine and financial expertise," *Neurocomputing*, vol. 128, pp. 296-302, 2014.
- [2] P. A. Maithili, "Neural network towards business forecasting," *IOSR Journal of Engineering*, vol. 02, no. 04, pp. 831-836, 2012.
- [3] F. Amato, A. López, E. Peña-Méndez, P. Vañhara, A. Hampl and J. Havel, "Artificial neural networks in medical diagnosis", *Journal of Applied Biomedicine*, vol. 11, no. 2, pp. 47-58, 2013.
- [4] E. Molga, B. van Woezik and K. Westerterp, "Neural networks for modelling of chemical reaction systems with complex kinetics: oxidation of 2-octanol with nitric acid," *Chemical Engineering and Processing: Process Intensification*, vol. 39, no. 4, pp. 323-334, 2000.
- [5] Y. Li, A. Rauth and X. Wu, "Prediction of kinetics of doxorubicin release from sulfopropyl dextran ion-exchange microspheres using artificial neural networks," *European Journal of Pharmaceutical Sciences*, vol. 24, no. 5, pp. 401-410, 2005.
- [6] E. Gueguim Kana, J. Oloke, A. Lateef and M. Adesiyan, "Modeling and optimization of biogas production on saw dust and other co-substrates using Artificial Neural network and Genetic Algorithm," *Renewable Energy*, vol. 46, pp. 276-281, 2012.
- [7] B. Rodríguez Galdón, E. Peña-Méndez, J. Havel, E. Rodríguez Rodríguez and C. Díaz Romero, "Cluster Analysis and Artificial Neural Networks Multivariate Classification of Onion Varieties," *Journal of Agricultural and Food Chemistry*, vol. 58, no. 21, pp. 11435-11440, 2010.
- [8] P. Fedor, I. Malenovský, J. Vañhara, W. Sierka and J. Havel, "Thrips (Thysanoptera) identification using artificial neural networks," *Bulletin of Entomological Research*, vol. 98, no. 05, 2008.
- [9] M. Khashei, A. Zeinal Hamadani and M. Bijari, "A novel hybrid classification model of artificial neural networks and multiple linear regression models," *Expert Systems with Applications*, vol. 39, no. 3, pp. 2606-2620, 2012.
- [10] H. Chen, R.H. Chiang, and V.C Storey, "Business intelligence and analytics: from big data to big impact," *MIS quarterly*, pp.1165-1188, 2012.
- [11] F. López and I. Sanz, "Bankruptcy visualization and prediction using neural networks: A study of U.S. commercial banks," *Expert Systems with Applications*, vol. 42, no. 6, pp. 2857-2869, 2015.
- [12] H. Li, S. Guo, C. Li and J. Sun, "A hybrid annual power load forecasting model based on generalized regression neural network with fruit fly optimization algorithm," *Knowledge-Based Systems*, vol. 37, pp. 378-387, 2013.
- [13] M. Zekic-Susac, N. Sarlija, and Bencic, "Small business credit scoring: a comparison of logistic regression, neural network, and decision tree models," in *26th International Conference on Information Technology Interfaces*, pp. 265-270, 2004.
- [14] SOLIDARIO. S. F. P. Y, "Ley Orgánica de Economía Popular y Solidaria," 2014.
- [15] M. Bashiri and A. Farshbaf Geranmayeh, "Tuning the parameters of an artificial neural network using central composite design and genetic algorithm," *Scientia Iranica*, vol. 18, no. 6, pp. 1600-1608, 2011.
- [16] M. Virág and T. Kristóf, "Neural Networks in Bankruptcy Prediction - A Comparative Study on the Basis of the First Hungarian Bankruptcy Model," *Acta Oeconomica*, vol. 55, no. 4, pp. 403-426, 2005.
- [17] P. du Jardin, "Predicting bankruptcy using neural networks and other classification methods: The influence of variable selection techniques on model accuracy," *Neurocomputing*, vol. 73, no. 10-12, pp. 2047-2060, 2010.