

La Gaceta

ÓRGANO OFICIAL

DE LA UNIVERSIDAD NACIONAL DE INGENIERÍA

""""""""""
AÑO LVII LIMA 10 DE FEBRERO DE 2022 NÚMERO 021
""""""""""

UNIVERSIDAD NACIONAL DE INGENIERÍA Escuela Central de Posgrado

Se invita a la comunidad universitaria a participar de la videoconferencia de la defensa pública virtual de la Tesis de **DOCTORADO EN INGENIERÍA DE SISTEMAS**, del Mag. César Aldo Canelo Sotelo, a realizarse el día martes 15 de febrero, a las 09 h 00.

TÍTULO DE LA TESIS:

“REDES NEURONALES ARTIFICIALES Y MÁQUINA CON SOPORTE VECTORIAL PARA CLASIFICAR A LOS SOLICITANTES DE MICROCRÉDITO”.

ASESOR:

Dr. PEDRO CELINO ESPINOZA HARO
Docente de la Facultad de Ingeniería Industrial y de Sistemas
Universidad Nacional de Ingeniería

RESUMEN

Las entidades crediticias constantemente se enfrentan al problema de controlar el riesgo de crédito al que se exponen al desarrollar sus operaciones crediticias, en tal sentido, siempre han requerido apoyarse en modelos predictivos que les ayuden a tomar decisiones acertadas para la aceptación o el rechazo de una solicitud de crédito. Los modelos predictivos que emplean las instituciones financieras para calificar a los solicitantes de crédito son los conocidos modelos clásicos basados en técnicas estadísticas y los modelos basados en técnicas de máquinas de aprendizaje.

En esta investigación, con el objetivo de clasificar a los solicitantes de microcrédito y contribuir a la mejora de la gestión del riesgo crediticio, se trabaja con la Base de datos de una Caja Municipal de Ahorro y Crédito (CMAC) que contiene 15,569 registros, cada uno con 27 variables, en donde las primeras 26 variables son los datos del cliente, y la última es la variable de aceptación o rechazo del crédito (V27).

Al estudiar la Base de datos, se determinó que la variable Días de atraso de la última cuota pagada (V8) tiene la más alta correlación (0.78) con la Aceptación o el rechazo del crédito, y luego de un estudio más detenido de la Base de datos se descubrió que aquellos que tenían días de atraso de la última cuota superior a los 30 días se constituían en los clientes rechazados y los que no, en aceptados, por esta razón, las pruebas empleando las técnicas de máquinas de aprendizaje, se hacen con la Base de datos que, en unos casos incluyen esta variable y en otros casos la excluyen. En primer lugar, se emplean las Redes Neuronales Artificiales (RNA).

Backpropagation, para predecir el comportamiento crediticio de los prestatarios ante una entidad de

microfinanzas. En segundo lugar, se emplean las Redes SelfOrganizing-Maps (RNA-SOM) para agrupar los prestatarios en clústeres, y estudiar qué variables han influido en la conformación de los clústeres; y, en tercer lugar, se emplean las Máquinas con Soporte Vectorial (MSV) para separar los registros de la Base de datos

Con la RNA Backpropagation se hicieron pruebas con diversas arquitecturas de la red, y se determinó que con una red de 4 capas con 14, 10, 8 y 1 neuronas respectivamente, se obtuvo una precisión 0.97682 que fue la mayor obtenida en todas las pruebas hechas con la Base de datos completa. Luego se han hecho pruebas con la Base de datos en la que la variable V8 sustituye a la variable V27, y la precisión obtenida es menor, pero sigue siendo una buena precisión de la red, y finalmente, se excluye de la Base de datos a la variable V8, obteniéndose una precisión menor, y de esta manera se comprueba que la variable V8 es la más realista.

Con Redes Self-Organizing-Maps (RNA-SOM) de dos neuronas se han hecho dos pruebas, una con topología Gridtop y métrica Dist, y otra prueba cambiando a la topología Hextop y a la métrica Linkdist. Los resultados obtenidos que están respaldados por las medidas estadísticas de las variables en cada clúster y los coeficientes de correlación en la formación de los clústeres, concluyen que, con el cambio de topología y métrica, no hay diferencias sustanciales en la composición de los clústeres, sólo ha habido el desplazamiento de un pequeño grupo de prestatarios entre los clústeres.

Con Redes Self-Organizing-Maps (RNA-SOM) de tres neuronas se han hecho dos pruebas, una con topología Gridtop y métrica Dist, y otra prueba cambiando a la topología Hextop y a la métrica Linkdist. Las medidas estadísticas de las variables en cada clúster y los coeficientes de correlación en la formación de los clústeres, permiten concluir que, con el cambio de topología y métrica, no hay diferencias sustanciales en la composición de los clústeres, sólo se ha producido el desplazamiento de un pequeño grupo de prestatarios entre los clústeres contiguos.

Finalmente, con la Máquina con Soporte Vectorial con núcleo lineal, se han separado a los prestatarios en dos grupos: aceptados y rechazados. Se han hecho dos pruebas, una con la Base de datos completa, y otra prueba eliminando la variable V8. En ambas pruebas se ha logrado la separación con un reducido número de vectores soporte en las fronteras, es decir se ha encontrado un hiper-plano de separación óptima que ha dado lugar a la separación de dos grupos de clientes bien definidos.

ENLACE

La Escuela Central Posgrado UNI le está invitando a una reunión de Zoom programada.

Tema: SUSTENTACION DE TESIS DOCTORAL

FECHA: 15 febrero 2022

HORA: 09:00 a. m. Lima

Unirse a la reunión Zoom

<https://us02web.zoom.us/j/85824232190>

ID de reunión: 858 2423 2190



**EDITOR: SECRETARIA GENERAL UNI
IMPRESA DE LA EDUNI**