

UNIVERSIDAD NACIONAL MAYOR DE SAN MARCOS

FACULTAD DE CIENCIAS MATEMÁTICAS

ESCUELA PROFESIONAL DE ESTADÍSTICA

**Comparación de árboles de clasificación Cart y redes
neuronales perceptrón multicapa para la clasificación
de estudiantes de secundaria que consumen o no
alcohol**

TESIS

Para optar el Título Profesional de Licenciada en Estadística

AUTOR

Ingrid VILLANUEVA GALVÁN

ASESOR

Geraldine Judith VIGO CHACÓN

Lima - Perú

2017

RESUMEN

El propósito del presente estudio es mostrar que las técnicas de minería de datos pueden ser una alternativa para la clasificación de los estudiantes de secundaria que consumen o no alcohol. Las técnicas que se usarán son Árboles de clasificación CART y Redes Neuronales de tipo Perceptrón Multicapa.

Se aplicarán las dos técnicas en una población de 1044 estudiantes de secundaria, y esta, será dividida en dos muestras, el 70% para el entrenamiento y 30% para la validación.

Para la comparación de los modelos que mejor discrimina a los estudiantes de secundaria que consumen o no alcohol, se utilizará la Sensibilidad y el Coeficiente de Gini. Adicionalmente, se hará la comparación aplicando la validación cruzada en ambos modelos.

Palabras clave: Redes Neuronales Perceptrón Multicapa, Arboles clasificación CART, Gini, Sensibilidad.

ABSTRACT

The purpose of this study is to demonstrate that data mining techniques are an alternative for classification of high school students who drink and those who do not drink alcohol. The techniques to be used are Classification and Regression Trees (CART) and Multi-layer Perceptron Neural Networks.

The two techniques will be applied in a population of 1044 secondary students and it will be divided into two samples, 70% for training and 30% for validation. Sensitivity and Gini coefficient will be used to determine the model that best discriminates the high school students who drink and those who do not drink alcohol. Additionally, comparison will be performed using cross validation in both models

Key words: Neural networks Multi-Layer Perceptron, CART classification tree, Gini, Sensitivity.