

## **RESUMEN**

El abandono o deserción escolar es una problemática actual que afecta directamente al sistema educativo y a los involucrados en la educación de un individuo como la familia, la institución o el país, por tanto, la detección temprana del abandono escolar es una estrategia clave para las instituciones académicas a todo nivel. Como medida actual para esta detección se encuentra el uso de técnicas de Minería de Datos Educativa (EDM) que permite descubrir patrones en los datos demográficos y académicos de los alumnos; siendo uno de los casos el uso de modelos de clasificación para predecir los alumnos que abandonen la institución; en este estudio se utilizó este marco conceptual con un conjunto de datos de tres años lectivos de la Unidad Educativa Lenin School ubicada en Latacunga-Cotopaxi para evidenciar la efectividad de los modelos de clasificación de EDM. Por medio del software estadístico R se ejecutaron 8 modelos de clasificación sobre los datos de los alumnos y se realizó la evaluación de estos por medio de la matriz de confusión y el área bajo la curva ROC. Como resultado se obtuvo que el modelo NNET tiene una exactitud del 98% sobre los datos recogidos del alumnado para la predicción de los posibles abandonos en un año lectivo. Adicionalmente también se comprobó que el mismo conjunto de datos puede servir para explicar la retención de alumnos de la institución, aunque en un menor porcentaje de exactitud.

### **Palabras Clave:**

**ABANDONO ESCOLAR**

**MINERÍA DE DATOS EDUCACIONAL**

**CLASIFICACIÓN**

**NNET**

**ROC**

## **ABSTRACT**

Early Leaving or drop-out school is a current problem that directly affects the education system and people who are involved in the education process such as family, institution or country, so early detection of school drop-out is a key strategy for academic institutions at all levels. A current solve for this detection is the use of Educational Data Mining (EDM) techniques that allows to discover patterns in demographic and academic data of the students; being one of the cases the use of classification models to predict the students that drop-out the institution. In this study, this conceptual framework was used with a data set of three academic years of the “Lenin School” Educational Unit located in Latacunga-Cotopaxi to demonstrate the effectiveness of the EDM classification models. Using the statistical software R, 8 models of classification were executed on the data of the students and the evaluation of these was made by the confusion matrix and the area under the ROC curve. As a result of the study was obtained the NNET model which has an accuracy of 98% on the data collected from the students for the prediction of possible drop-outs in a school year. In addition, it was also verified that the same data set can be used to explain the retention of students in the institution, even though in a lower percentage of accuracy.

### **Keywords:**

**DROPOUT SCHOOL**

**EDUCATIONAL DATA MINING**

**CLASSIFICATION**

**NNET**

**ROC**