

RESUMEN

La deserción académica dentro del sistema universitario constituye un desafío para las Instituciones de Educación Superior dentro y fuera del Ecuador. En la Escuela de Formación de Tecnólogos (ESFOT) de la Escuela Politécnica Nacional (EPN) continuamente se busca aumentar la retención de estudiantes, mediante alternativas que suban las tasas de deserción académica. Mediante este trabajo se analizó la información de los estudiantes de la ESFOT, armando una bodega de datos, que luego de ser sometido a un proceso ETL, se agrupó a factores del alumnado con base al rendimiento académico. Esta bodega de datos, permitió crear un modelo de minería de datos que determinó patrones y características que generan repitencia y deserción, para guiar adecuadamente al alumnado con mayor tendencia de abandono o fallo. El proyecto se enmarcó a la metodología de investigación científica basada en el diseño en conjunto con la metodología de minería de datos CRISP-DM, que proporcionó una guía de referencia normalizada del ciclo de vida de un proyecto de análisis de datos. Utilizando el algoritmo DECISION TREE, y Auto Model de RapidMiner se obtuvo un modelo que identifica los patrones de comportamiento que afectan el rendimiento escolar de los estudiantes de la ESFOT con un nivel de confianza del 96,9%. Se detectaron algunos patrones de comportamiento que influyen en su aprovechamiento académico.

PALABRAS CLAVES:

- **MODELO DE MINERÍA DE DATOS**
- **RENDIMIENTO ESCOLAR**
- **PATRONES DE COMPORTAMIENTO**
- **BODEGA DE DATOS**

ABSTRACT

The academic desertion within the university system constitutes a challenge for Higher Education Institutions inside and outside of Ecuador. In the School of Training of Technologists (ESFOT) of the National Polytechnic School (EPN) is continuously seeking to increase the retention of students, through alternatives that increase the rates of academic desertion. Through this work the information of ESFOT students was analyzed, setting up a Data Warehouse, which after being subjected to an ETL process, groups student factors based on academic performance. This Data Warehouse allowed the creation of a data mining model that determined characteristics and patterns that cause repetition and desertion, to adequately guide students with a higher risk of failure or abandonment. The project was framed to the scientific research methodology based on the design in conjunction with the data mining methodology CRISP-DM, which provided a standard reference guide for the life cycle of a data analysis project. Using the DECISION TREE algorithm and RapidMiner's Auto Model, a model was obtained that identifies behavior patterns that affect the school performance of ESFOT students with a confidence level of 96.9%. Some patterns of behavior that influence their academic performance were detected.

KEYWORDS:

- **DATA MINING MODEL**
- **SCHOOL PERFORMANCE**
- **BEHAVIOR PATTERNS**
- **DATA WAREHOUSE**