

RESUMEN

La minería de datos es ampliamente utilizada en diversos campos: educación, computación móvil, minería web, análisis financiero, análisis de delitos, ingeniería, gestión, medicina, etc. Naturalmente, en la dinámica de la sociedad actual, se ha dado, cada vez, mayor importancia a la educación y la investigación. Adicionalmente, las instituciones de educación generan y almacenan datos sobre los estudiantes, que con un procesamiento subsecuente adecuado pueden resultar útiles para tomar decisiones estratégicas en pro de todos los procesos académicos internos de las mismas instituciones. En este sentido, las técnicas computarizadas, particularmente las técnicas de minería de datos, han tomado importancia dado que permiten comprender mejor a los estudiantes y los entornos en los que aprenden, orientando a las instituciones en cómo proceder para brindar mejoras continuas en la calidad de educación. El presente trabajo presenta un estudio de técnicas de minería de datos, aplicada a datos de estudiantes de educación media, para proponer una base conceptual y algorítmica para diseñar una herramienta que tenga un margen mayor de análisis de la información, a través de los datos generados por la aplicación, apoyándose de igual manera en algoritmos de inteligencia artificial, los cuales son: método ponderado de vecinos cercanos, árboles de decisión, máquinas de vectores de soporte, y enfoques de multi-clasificadores.

PALABRAS CLAVE:

- **MINERÍA DE DATOS**
- **PATRONES DE DESEMPEÑO ACADEMICO**
- **SELECCIÓN DE CARACTERÍSTICAS**
- **CLASIFICADORES**
- **MULTI-CLASIFICADORES**
- **MATLAB**

ABSTRACT

Data mining is widely used in several fields, such as: education, mobile computing, web mining, financial analysis, crime analysis, engineering, management, and medicine, among others. Naturally, in the dynamics of today's society, education and research have become increasingly important. Furthermore, educational institutions are generating and storing students' data, which – undergoing an appropriate, subsequent processing- can be useful to make strategic decisions in favor of all the internal academic processes thereof. In this sense, computerized techniques, particularly data mining techniques, have become a key tool since they allow students for a better understanding, as well as identifying the environments in which they learn, by guiding the institutions on how to proceed to provide continuous improvements in reaching the high quality in education. This master's thesis presents a study of data mining techniques -applied to data of students from secondary education- to propose a conceptual and algorithmic basis to design a tool having a greater margin of analysis of the information. Specifically, the explored methods are: weighted nearest neighbors, decision trees, support vector machines, and multi-classifier approaches.

KEY WORDS:

- **DATA MINING**
- **ACADEMIC PERFORMANCE PATTERNS**
- **FEATURE SELECTION**
- **CLASSIFIERS**
- **MULTIPLE CLASIFIERS**
- **MATLAB**