

UNIVERSIDAD DE LOS ANDES.
FACULTAD DE CIENCIAS ECONÓMICAS Y SOCIALES.
ESCUELA DE ESTADÍSTICA.

Licenciatura de Estadística.

MINERIA DE DATOS

DATOS DE IDENTIFICACIÓN:

Código	Semestre	Créditos	Horas	Condición	Vigencia	Cátedra Responsable
	VIII	4	T:4 P:0	Electiva	B2015	Computación
Prelaciones: Métodos Estadísticos III, Base de Datos						

JUSTIFICACION:

La minería de datos surge como una solución al problema que implica la explosión de datos producto de los progresos en la tecnología de adquisición y almacenamiento de grandes volúmenes de información a partir de una diversidad de fuentes.

Dada la gran cantidad de información que manejan las empresas y organizaciones hoy en día, se hace cada vez más necesario el uso de métodos automáticos para describir y analizar la información más valiosa que se encuentra escondida en los datos.

La minería de datos se define como el proceso no trivial de extracción de conocimiento válido, novedoso, potencialmente útil y comprensible, a partir de grandes volúmenes de datos.

La minería de datos se utiliza en diferentes campos de la ciencia: aplicaciones financieras y de banca, en análisis de mercado, en seguros y salud privada, en educación, procesos industriales, en medicina, telecomunicaciones y un sinnúmero de áreas más.

Las herramientas de la Minería de Datos permitirá afianzar el área de conocimientos del Análisis de Datos, pilar fundamental del Perfil del Licenciado en Estadística.

REQUERIMIENTOS:

Para alcanzar los objetivos que se plantean el estudiante debe tener conocimientos en Métodos Estadísticos y Computación, con especial énfasis en la programación de computadores y las Bases de Datos.

OBJETIVO GENERAL:

Promover las competencias necesarias para la aplicación efectiva del proceso de minería de datos en diferentes dominios o áreas.

OBJETIVOS ESPECIFICOS:

- Aprender los fundamentos teóricos y prácticos relacionados con la minería de datos.
- Comprender y Estudiar las técnicas de minería de datos para resolver tareas predictivas y descriptivas.
- Aplicar las técnicas de minería de datos para descubrir conocimiento en diversas áreas.
- Aprender algún software para aplicar el proceso de minería de datos.

CONTENIDO PROGRAMATICO:

1. Introducción a la minería de datos.

Concepto de Minería de Datos (MD).
Datos, patrones y tipos de patrones.
Tareas de la Minería de Datos.
Proceso de descubrimiento de conocimiento a partir de datos.
Técnicas de Minería de Datos.
Aplicaciones de la Minería de Datos.

2. Preparación de los datos.

Limpieza y depuración de los datos.
Transformaciones.
Reducción de la dimensionalidad y selección de variables.

3. Análisis exploratorio de datos.

Conceptos básicos.
Resúmenes estadísticos.
Técnicas de visualización para análisis de datos.

4. Técnicas de minería de datos para tareas predictivas.

Técnicas para clasificación:
Introducción.
Aprendizaje de árboles de clasificación.
Aprendizaje de reglas de clasificación.
Métodos basados en vecindad.

Técnicas para regresión:
Introducción.
Aprendizaje de modelos de regresión.

5. Técnicas de minería de datos para tareas descriptivas.

Técnicas para agrupación:

Introducción.

Métodos basados en particiones.

Métodos jerárquicos.

Técnicas para análisis de asociación:

Introducción.

Soporte y confianza de una regla.

Aprendizaje de reglas de asociación.

6. Evaluación de modelos.

Medidas de evaluación.

Técnicas de evaluación.

ESTRATEGIAS METODÓLOGICAS:

Clases presenciales donde se desarrollaran los aspectos teóricos, la resolución de algunos ejemplos. El docente podrá usar diapositivas y/o pizarrón y marcadores para impartir sus clases. Incentivará al estudiante sobre la consulta de bibliografía recomendada, así como el desarrollo de los problemas usando algún software de programación y/o minería de datos.

EVALUACION:

Realización de dos exámenes escritos. Realización de Ejercicios en Clases. Tareas y Ejercicios para casa. Trabajo o Proyecto Final.

BIBLIOGRAFÍA:

- César Pérez López; Daniel Santin González. **Minería de datos. Técnicas y Herramientas**. Thomson Editorial Paraninfo. 2007.
- José Hernández Orallo, Ma José Ramírez Quintana, César Ferri Ramírez. **Introducción a la Minería de Datos**. Prentice-Hall. 2004.
- Jiawei Han, Micheline Kamber, Jian Pei. **Data Mining. Concepts and Techniques**. Morgan Kaufmann. Elsevier. Tercera Edición. 2011.