

Almacenes y Minería de Datos (Curso 2003-2004)

Código Sigma: 1218

Carácter: Libre Configuración

Curso: 2º

Ciclo: 2º

Cuatrimestre: Anual

Créditos: 9 (3 T + 6 P)

Profesor: [Gerardo Pastor Vegas](#) y [M^a Angeles Martín Mínguez](#) de la empresa ALGOR Consultoría y Sistemas S.L.

Departamento: [Estadística e Investigación Operativa](#)

Objetivos:

Evaluación:

Se valorarán las prácticas realizadas, así como los resultados de dos exámenes cuatrimestrales, los cuales constarán de una parte teórica y otra práctica, proporcionando cada una de dichas partes hasta un 50% de la calificación de cada examen.

Prerrequisitos: Es aconsejable tener conocimientos básicos de Bases de Datos y de Análisis de Datos Multivariantes.

Dirigido a estudiantes de la titulación de Licenciado en Ciencias y Técnicas Estadísticas, tercer curso de Diplomado en Estadística, o de segundo ciclo de Matemáticas.

Descriptor: Construcción y manejo de almacenes de datos. Herramientas para su explotación estadística. Aplicaciones.

Programa:

PARTE 1: BASES DE DATOS

1.1 - Características generales .Objetos de los que consta una BD.

- 1.2 - Modelos de datos
 - 1.2.1 Entidad-Relación
 - 1.2.2 Jerárquico
 - 1.2.3 De red
 - 1.2.4 Relacional
 - 1.2.5 Relacional extendido
 - 1.2.6 Orientado a objetos

1.3 - Modelo Relacional

- 1.3.1 Conceptos básicos: relación, tabla, dominio, tupla, esquemas
- 1.3.2 Algebra relacional
- 1.3.3 Lenguajes comerciales: SQL, Quel, QBE

1.4 - Diseño de Bases de Datos relacionales

- 1.5 - Arquitecturas para un sistema de Bases de Datos.
- 1.6 - Bases de datos de interés

PARTE 2 : BASES DE DATOS DE SOPORTE A LA DECISIÓN: DATAWAREHOUSE.

2.1 - Evolución de los sistemas de soporte a la decisión: Datawarehouse.

- 2.1.1 Concepto de Datawarehousing.
- 2.1.2 Características de un Datawarehousing.
- 2.1.3 Estructura del Datawarehouse. Metadata.
- 2.1.4 Arquitectura de un Datawarehouse

2.1.5 Transformación de datos y metadatos.Flujo de datos

2.1.6 Usos del Datawarehouse

2.2 - Proyecto de elaboración de un Datawarehouse.

2.2.1 Organización:Factores en la planificación de un Datawarehouse, estrategias de diseño , estrategias para la gestión.

2.2.2 Desarrollo: Bloques de Datawarehouse, alcance, redundancia de datos, Diseño de la arquitectura.

2.2.3 Implementación.

2.2.4 Evaluación del rendimiento del Datawarehouse.

2.3 - Técnicas avanzadas de diseño de un Datawarehouse.

PARTE 3:EL CICLO DE DATA MINING: FASES Y TIPOS DE PROBLEMAS

3.1. Fases del proceso de Data Mining.

3.2. La fase de preprocesamiento de los datos: errores, outliers, missing.

3.3. Tipos de problemas de Data Mining y tipos de enfoques.

PARTE 4: TÉCNICAS DE DATA MINING

4.1 Problemas de Clasificación y Predicción.

4.1.1. Regresión.

4.1.2. Redes neuronales.

4.1.3. Árboles de decisión.

4.2. Evaluación e implantación del modelo.

4.3. Análisis de segmentación: Clustering.

4.4. Reglas de asociación en bases de datos.

Actividades:

Bibliografía:

- Everitt, B.S. (1994). A Handbook of Statistical Analyses using S-Plus. Chapman and Hall.
- Inmon, W.H. (1996). Building the Datawarehouse. J.Wiley & Sons.
- Jiawei Han & Micheline Kamber. Data Mining: Concepts and Techniques. Morgan Kaufman Publishers.
- Kimball, R (1996). The Data Warehouse Toolkit. John Wiley & Sons.
- T. Hastie, R. Tibshirani, J. H. Friedmanm. The Elements of Statistical Learning: Data Mining, Inference and Prediction. (Springer Series in Statistics), Springer.