



UNIVERSIDAD DE CHILE
Facultad de Ciencias Físicas y Matemáticas
Departamento de Ingeniería Industrial

IN65A Aplicación de Bases de Datos y **Minería de Datos**

U.D. 10

REQUISITOS	:	CC42A/AD
CARÁCTER	:	Electivo de la Carrera de Ingeniería Civil Industrial
PROFESOR	:	RICHARD WEBER
SEMESTRE	:	PRIMAVERA 1999

OBJETIVOS:

Estudiar, analizar, diseñar y aplicar las tecnologías de Bases de Datos y Minería de Datos (Data Mining) en empresas u organizaciones.

ESPECIFICOS:

1. Entender modelos de Bases de Datos y sus (potenciales) usos en aplicaciones en las organizaciones.
2. Diferenciar el procesamiento transaccional del procesamiento analítico de datos de la información.
3. Entender los principales objetivos y métodos de Minería de Datos.
4. Llegar a conocer típicas aplicaciones de Minería de Datos y herramientas para realizarlas.
5. Poner a los alumnos en el lugar de decidir cual método es mejor para resolver un problema real donde hay datos para analizar.
6. Resolver un problema real usando los métodos vistos en clase y presentar la solución.

TEMARIO:

1. Revisión del modelo y de la teoría de Bases de Datos. Evolución de las tecnologías asociadas a Bases de Datos. Análisis transaccional versus analítico de los datos organizacionales.
2. Base de Datos analítica. Conceptos de Bodega de Datos, Data Marts, Base de Datos multi-dimensionales, y Bodega de Datos Virtual.
3. Base de Datos Inteligente. Representación y coordinación del conocimiento en la tecnología de Bases de Datos. Integración de datos e información.
4. Calidad de datos y diseño. Enfoques orientados a establecer políticas de calidad. Medición de calidad de Bases de Datos, y mecanismos de homogeneización de la calidad de múltiples Bases de Datos.
5. Conceptos y métodos de Minería de Datos. Presentación de los conceptos y métodos principales y discusión de sus fuerzas y debilidades. Métodos por presentar serán redes neuronales, métodos para la generación automática de reglas y métodos de estadística.
6. Aplicaciones de Minería de Datos. Demostración de aplicaciones reales y desarrollo de aplicaciones en la clase. Algunas áreas de aplicaciones serán database marketing, detección de fraude y evaluación de solvencia.
7. Perspectiva del desarrollo futuro. Descripción de investigaciones necesarias y desafíos futuros con respecto a Bases de Datos y Minería de Datos.

ACTIVIDADES:

Los alumnos, en forma adicional a las clases de cátedra, desarrollarán tres tareas en el Semestre.

Tarea 1: Creación de Bases de Datos

Desarrollar el diseño de una Base de Datos para modelar datos de un problema real. El objetivo de esta tarea es crear el conocimiento necesario para establecer Bases de Datos como requisito para Minería de Datos.

Tarea 2: Aplicación de técnicas de Minería de Datos

Modelación de un problema y aplicación de diferentes técnicas del área de Minería de Datos. El objetivo de esta tarea es adquirir el conocimiento necesario para aplicar técnicas de Minería de Datos.

Tarea 3: Presentación de una solución (Estudio de casos)

Los alumnos resuelven un problema real desde la fase de la definición del problema hasta la presentación final de la solución. El objetivo de esta tarea es la simulación de un proyecto real.

Las tareas deberán ser realizadas en grupos de no más de tres personas. El profesor se guarda el derecho de interrogar a cualquier integrante de un grupo para evaluar el trabajo del grupo completo.

EVALUACION:

La evaluación del curso corresponde a un control y un examen, 2 CTP's (controles de lectura) y las 3 tareas. La ponderación es la siguiente:

Control y Examen	50%
Tareas	40%
CTP's	10%

REFERENCIAS:

- Adriaans, P., Zantinge, D. (1996): Data Mining. Addison-Wesley, Harlow
- Berry, M. J. A., Linoff, G. (1997): Data Mining Techniques – For Marketing, Sales, and Customer Support. John Wiley & Sons, New York
- Cattell R. (1994): Object Data Management: Object oriented and Extended relational Database Systems, Addison Wesley
- Edelstein H. (1998): Introduction to Data Mining and Knowledge Discovery
- Famili, A., Shen, W.-M., Weber, R., Simoudis, E. (1997): Data Preprocessing and Intelligent Data Analysis. Intelligent Data Analysis Vol. 1, No. 1, <http://www.elsevier.com/locate/ida>
- Hammergren T., Hammergren T.C. (1997): Data Warehousing: Building the Corporate Knowledge Base
- Jajodia, S., Kerschberg, L. (eds.) (1997): Advanced Transaction Models and Architectures, Kluwer Academic Publishers, Boston
- Kilov H. Ross, J. (1994): Information Modeling: an object-oriented approach, Prentice Hall, Object Oriented Series
- Kimball, R. (1996): The Data Warehouse Toolkit: Practical Techniques for Building Dimensional Data Warehouse, John Wiley & Son
- Mattison, R. (1997): Data Warehousing and Data Mining for Telecommunications. Artech House, Boston, London
- Motro, A., Smets, P. (1997): Uncertainty Management in Information Systems: From Needs to Solutions, Kluwer Academic Publishers
- Quinlan, J. R., C4.5: Programs for Machine Learning. Morgan Kaufmann Publishers, San Mateo, 1993
- Rojas, R. (1996): Neural Networks – A Systematic Introduction. Springer-Verlag, Berlin Heidelberg
- Weiss, S. M., Kulikowski, C. A. (1991): Computer Systems that learn. Morgan Kaufmann, San Mateo

Zimmermann, H.-J. (1996): Fuzzy Set Theory - and Its Applications. Third Edition. Kluwer Academic Publishers, Boston Dordrecht London

Más información sobre los temas del curso es disponible en el siguiente sitio del internet: www.kdd.org