

BUSINESS INTELLIGENCE

PROFESOR DE CÁTEDRA: SEBASTIÁN MALDONADO

OBJETIVO:

Diseñar y aplicar las tecnologías de Bases de Datos y Minería de Datos (*Data Mining*) al mundo de la Inteligencia de Negocios. Al terminar este curso el alumno debe ser capaz de utilizar la metodología de *Knowledge Discovery in Databases* (KDD), eligiendo el método más adecuado para una aplicación real en BI.

ESPECIFICOS:

1. Comprender las aplicaciones de modelos de Bases de Datos, Bodegas de Datos y tecnología OLAP, entendiendo sus usos y en aplicaciones corporativas de BI.
2. Entender la operación de los principales métodos de Data Mining, su clasificación y rango de aplicación.
3. Conocer aplicaciones clásicas de Data Mining en BI y tener nociones sobre el uso de herramientas computacionales para la solución y aplicación de distintos métodos de Minería de Datos.

TEMARIO:

1. Introducción y Proceso KDD (Knowledge Discovery Databases).
2. Diseño de Bases de Datos. Consolidación de bases de datos y herramientas de consulta (SQL).
3. Preprocesamiento y Limpieza de Datos. Aplicación de limpieza, imputación, transformación y selección de atributos.
4. Conceptos y métodos principales de Data Mining supervisado (clasificación y regresión) centrándose en la aplicación y diseño particular de experimentos. Los métodos a presentar incluyen:
 - Árboles de Decisión.
 - Regresión Logística.
 - Redes Neuronales Artificiales.
 - Support Vector Machines.
5. Clustering de datos. Aplicación de herramientas para segmentar observaciones e interpretación de grupos obtenidos. Métodos:
 - K-Medias.
 - Métodos jerárquicos.
6. Aplicaciones de Data Mining. Demostración de aplicaciones reales y desarrollo de aplicaciones. Ej: Asignación de Créditos, Detección de Fraude, entre otras.

ACTIVIDADES:

Se realizarán **siete** sesiones de tres horas. Los alumnos, en forma adicional a las clases de cátedra, desarrollarán una tarea (grupal) en el curso. Durante la última sesión se realizará un examen individual.

EVALUACION:

Para aprobar el curso se necesita una nota final igual o superior a 4.0. No se contemplan actividades adicionales. La nota final del curso se determina en base a la siguiente ponderación:

- Nota Tarea: 40% Nota final .
- Nota Examen: 60% Nota final.

La tarea contemplará un informe escrito donde se deberán reportar los resultados y procedimientos obtenidos. La calidad y presentación del informe serán consideradas en la nota final de la tarea. La tarea contempla, en forma adicional, un bono o castigo asociado a la *performance* del modelo en una muestra ciega.

BIBLIOGRAFÍA DEL CURSO:

- Anahory, S., Murray, D. (1997): Data Warehousing in the Real World. Addison-Wesley, Harlow; capítulo 4, páginas 29-53
- 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
- Edelstein H. (1998): Introduction to Data Mining and Knowledge Discovery
- Elmasri, R., Navathe, S.B. (1997): Sistemas de Bases de Datos – Conceptos fundamentales. Seg. Edición, Addison-Wesley Iberoamericana, Wilmington, Delaware; capítulo 6, páginas 139-167, capítulo 7, página 188-205.
- Freeman, J.A., Skapura, D.M. (1993): Redes Neuronales – Algoritmos, Aplicaciones y Técnicas de Programación. Addison-Wesley Iberoamericana/ Diaz de Santos, Wilmington, Delaware
- Famili, A., Shen, W.-M., Weber, R., Simoudis, E. (1997): Data Preprocessing and Intelligent Data Analysis. Intelligent Data Analysis Vol. 1, No. 1, 3-23
- Hammergren T., Hammergren T.C. (1997): Data Warehousing: Building the Corporate Knowledge Base
- Han, J., Kamber, M. (2001): Data Mining – Concepts and Techniques. Morgan Kaufmann Publishers, San Francisco
- Mattison, R. (1997): Data Warehousing and Data Mining for Telecommunications. Artech House, Boston, London
- 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
- Weber, R. (2000): Data Mining en la Empresa y en las Finanzas Utilizando Tecnologías Inteligentes. Revista Ingeniería de Sistemas XIV, No 1, 61-78
- Zhang, P. (2007): Avoiding Pitfalls in Neural Networks Research. IEEE Transactions on Systems, Man and Cybernetics – Part C: Applications and Reviews. Vol. 37, No. 1, 3-16.