

CC68N TALLER DE MINERIA DE DATOS
05 UD

Prof. Hugo Mora

Requisitos: (CC42A, Autor) / CC60K

Objetivos:

Este taller tiene por objetivo permitir a los alumnos ejercitar en forma práctica las técnicas de minería de datos que le permitan resolver problemas de negocio concretos.

En este curso se enseñarán los siguientes algoritmos: Clustering, Association Rules, Decisión Trees, Naive Bayes, entre otros, así como una metodología para enfrentar proyectos de minado de datos.

Carga académica:

Se estima un curso de una clase a la semana, con trabajo práctico de los alumnos para realizar un proyecto de minería durante el semestre.

Recursos computacionales:

El profesor usará Microsoft SQL Server 2005 Analysis Services para sus presentaciones y ejemplos. La sugerencia es que los alumnos usen lo mismo, aun cuando pueden usar lo que estimen conveniente como herramienta de minado.

Horario:

Martes de 8:30 a 10:00 am. Una vez por semana. Sin ayudantías.

Temario:

Parte 1:

- Fundamentos y objetivos de minería de datos
- Tareas que resuelve minería de datos
- Algunos algoritmos disponibles
- Aplicaciones de minería de datos
- Metodología para desarrollar proyectos de minería de datos

Parte 2:

- Uso de algoritmos de clustering
- Uso de algoritmos de association rules
- Uso de algoritmos de Decision trees

- Uso de algoritmos de Naive Bayes
- Visión general de otros algoritmos (neural nets, times series, regresión, etc.)

Parte 3:

- Proyecto de curso: cada alumno deberá hacer un modelo de minería en base a definición y requerimiento dados por el profesor.

Evaluación:

- 50% Nota de examen, que incluirá material presentado en clases por el profesor más lecturas que se darán durante el semestre
- 50% Trabajo de investigación, construcción y presentación de un problema de negocios resuelto con minería de datos (para esto es imprescindible que el alumno cuente con una herramienta de minería de datos)

Bibliografía:

- Data Mining Techniques, Second Edition – Michael Berry, Gordon Linoff, 2004, John Wiley & Sons. Inc.
- Data Mining with SQL Server 2005 – ZhaoHui Tang, Jamie MacLennan, september 2005, John Wiley & Sons. Inc.
- Mastering Data Mining – Michael Berry, Gordon Linoff, 2000, John Wiley & Sons. Inc.
- Corporate Information Factory by W. H. Inmon, Claudia Imhoff and Ryan Sousa (John Wiley & Sons, 1998)
- The Data Warehouse Lifecycle Toolkit by Ralph Kimball, Laura Reeves et al. (John Wiley & Sons, 1998)
- Building Data Mining Applications for CRM – Alex Berson, Stephen Smith, Kurt Thearling, 2000, McGraw-Hill
- Data Mining, Concepts and Techniques – Jiawei Han, Micheline Kamber, 2001, Morgan Kaufmann Publishers
- Improving Data Warehouse and Business Information Quality – Larry English, 1999, John Wiley & Sons. Inc.
- Recursos técnicos de DW en la Web: www.dw-institute.com, www.olapreport.com
- Recursos técnicos de DM en la Web: www.kdnuggets.com
- SQL Server 2005: www.microsoft.com/sql/2005/default.msp
- Difusión de DM por parte de Microsoft: <http://www.sqlserverdatamining.com/DMCommunity/default.aspx>