

MAT-477: Teoría de Minería de Datos

Identificación

Asignatura:	Teoría de Minería de Datos
Sigla:	MAT-477
Area Curricular:	Computación Científica
Modalidad:	Semestral
Nivel Semestral:	Segundo o Tercer Semestre de la Maestría
Horas Teóricas:	4 por semana en dos sesiones
Horas Prácticas:	6 por semana
Horas de Laboratorio	4 por semana
Créditos:	6
Carreras destinatarias:	Matemática y Area de Ciencia y Tecnología

Objetivos

Desarrollar los fundamentos matemáticos de reducción, sistetización, asociación de datos y de los diferentes métodos del aprendizaje de máquina con el objetivo de explicar teoricamente el nuevo conocimiento generado por los diversos métodos de la minería de datos.

Competencias

Explica con fundamentos de álgebra, cálculo, algebra lineal, teoremas fundamentale del análisis matemático los diferentes métodos o algoritmos que se desarrollan en el descubrimiento del nuevo conocimeinto a partir de base de datos multivariados de grande escala.

Contenido Mínimo

1. *Análisis de calidad de datos*
2. *Reducción y Sintetización*
3. *Asociación*
4. *Agrupamiento*
5. *Aprendizaje de Máquina.* 5.1 Selección. 5.2 Árbol de decisión y otros. 5.3 Redes neuronales. 5.4 Balanceamiento.

Estructura de Evaluación

La evaluación es la valoración de las competencias de conocimientos (saber), habilidades (saber hacer) y de valores (saber ser) alcanzadas mediante exámenes parciales periódicas (60%), prácticas e implementaciones de laboratorio (15%) y una evaluación final (25%) de todo el contenido de la asignatura. Sobre un total de 100%, la nota mínima de aprobación en el postgrado es de 66%. La distribución de temas por parciales, así como el cronograma de los exámenes se presenta en un plan de trabajo al inicio del semestre. También está prevista un examen de recuperación de cualquier examen parcial cuya nota reemplaza a la anterior.

Métodos y Medios Didácticos

Los métodos didácticos aplicados en el proceso de enseñanza y aprendizaje de la materia son las exposiciones magistrales del docente que utiliza recursos educativos y métodos de razonamiento *inductivo*, *deductivo*, *analógico* y *heurístico* para inducir el aprendizaje *por descubrimiento propio*, *dialogado*, *programado* y *demostrativo* que permita al estudiante desarrollar su potencialidad *creativa* con pensamiento crítico capaz de demostrar y presentar los teoremas con rigor lógico utilizando el lenguaje matemático formal. Los medios didácticos que dispone la Carrera son las aulas equipadas con medios audio visuales, laboratorio de computación con internet, aplicaciones computacionales, guías de practicas, material impreso o digital, mapas conceptuales y una Biblioteca especializada que facilita el desarrollo teórico y práctico de la asignatura.

Prácticas y Laboratorio

Las prácticas escritas como las implementaciones en laboratorio son monitoriadas por el mismo docente del postgrado.

Criterios de Evaluación

La evaluación de la asignatura consiste en pruebas escritas u orales, donde se valora la aplicación adecuada de *conceptos, teoremas y métodos* en la *demostración o resolución* de problemas planteados; y, en la calificación de prácticas o trabajos de laboratorios cuyo informe debe estar escrito en un *lenguaje matemático* adecuado con rigor lógico. Se valora de forma adicional la *creatividad* y la *simplicidad* en la presentación de sus resultados.

Cronograma de Avance

Semana	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Capítulos	1	2		3			4				5									

Bibliografía

- [1] Fayyad, U., Haussler, D. & Stolorz, P., Mining Scientific Data, Communications of the Acm, Nov. 1996, Vol. 39, No. 11, 1996, Pages 51-57.
- [2] Han, J. & Kamber, M. Data Mining: Concepts and Techniques, 2nd Edition, Morgan Kaufmann, 2006.
- [3] Hand, D. J.; Mannila, H. & Smyth, P. Principles of Data Mining, the Mit Press, 2001.
- [4] Liu, H.; Motoda, H. Computational Methods of Feature Selection, Chapman & Hall/crc, 2008. 419p.
- [5] Mardia, K.v.; Kent, J.t. & Bibby, J.m. Multivariate Analysis, Academic Press, 1979.
- [6] Piatetski-Shapiro, G. & Frawley, W.j., Knowledge Discovery in Databases. Aaai Press, Menlo Park, California, 1991.
- [7] Pyle, D., Data Preparation for Data Mining, Morgan Kaufmann, 1999. Rezende, S. O., Sistemas Inteligentes, Manole, 2003.
- [8] Tan, Pang-Ning; Steinbach, Michael; Kumar, Vipin. Introduction to Data Mining. Addison Wesley, 2006. 769p.
- [9] Witten, I.h. & Frank, E., Data Mining: Practical Machine Learning Tools and Techniques With Java Implementations, Morgan Kaufmann, 2011.
- [10] <http://www.cs.waikato.ac.nz/ml/weka/>