



UNIVERSIDAD NACIONAL DE LA PLATA
FACULTAD DE INFORMÁTICA

Minería de Datos Usando Sistemas Inteligentes

Carrera: Ingeniería en Computación
Profesor Responsable: Lanzarini, Laura
Año: Optativa
Duración: Semestral
Carga Horaria Semanal: 6hs
Carga Horaria Total: 96hs

OBJETIVOS GENERALES

Introducir al alumno en las técnicas de Minería de Datos. Se analizarán modelos basados en regresión, árboles, reglas, redes neuronales y técnicas de agrupamiento. Se cubrirán las distintas etapas del proceso de Extracción de Conocimiento como herramienta de ayuda a la toma de decisiones. El énfasis está puesto en la resolución de problemas de clasificación y predicción.

PROGRAMA

Introducción

- El proceso KDD. Introducción. Fases del proceso de extracción del conocimiento. La Minería de Datos como fase del proceso KDD. Relación con otras disciplinas.
- Recuperación de información vs recuperación de datos. Proceso de recuperación de información.

Preparación de Datos

- Metadatos. Representaciones gráficas de los atributos.
- Limpieza y transformación. Transformación y creación de atributos. Discretización y Numerización, Normalización de rango, escalado y centrado.
- Exploración mediante visualización y selección de datos.

Técnicas de Minería de Datos

- Extracción de Patrones. Introducción. Tareas y Métodos. La Minería de Datos y el aprendizaje inductivo. Comparación de las técnicas de Minería de Datos.
- Técnicas estadísticas. Modelo de regresión.
- Métodos Bayesianos. Naive Bayes. Clasificadores bayesianos.
- Reglas de clasificación y asociación. Distintos métodos de obtención de reglas.
- Árboles de decisión. Partición o cobertura de ejemplos. Poda y reestructuración. Adaptación para grandes volúmenes de datos.
- Redes Neuronales. Introducción. Red Neuronal Perceptrón. Regla Delta. Combinador Lineal. Regla Delta Generalizada. Perceptrón multicapa. Resolución de problemas de clasificación y predicción.
- Técnicas de Agrupamiento: Jerárquicas y basadas en centroides (k-medias, SOM).



UNIVERSIDAD NACIONAL DE LA PLATA
FACULTAD DE INFORMÁTICA

- Otras Técnicas de agrupamiento.

Análisis y difusión del modelo obtenido

- Evaluación de modelos. Comparación de técnicas de aprendizaje.

BIBLIOGRAFIA

Título: Data Mining: Practical Machine Learning Tools and Techniques, (Third Edition)

Autores: Ian H. Witten, Eibe Frank, Mark A. Hall.

Editorial: Morgan Kaufmann.

Año de edición: 2013.

ISBN-13: 978-0123748560.

Título: Introducción a la Minería de Datos.

Autores: Hernández Orallo, Ramírez Quintana, Ferri Ramírez.

Editorial: Prentice Hall.

Año de edición: 2004.

ISBN 84-205-4091-9.

Título: Data Mining: Theories, Algorithms, and Examples.

Autor: Nong Ye.

Editorial: CRC Press.

Año de Edición: 2013.

ISBN 9781439808382.

DESCRIPCIÓN DE LAS ACTIVIDADES TEÓRICAS Y PRÁCTICAS

Dado que se trata de una materia con un fuerte contenido teórico, durante todo el proceso de aprendizaje se enfatiza en la relación existente entre dichos contenidos y su aplicación en distintas áreas tales como: Reconocimiento de Patrones y Minería de Datos. Esto incrementa la valoración del alumno por el tema y clarifica su comprensión.

Durante las clases teóricas se resuelven problemas concretos en forma grupal fomentando el análisis, modelado e implementación de distintas soluciones. Además, se sugieren lecturas complementarias en diferentes textos que se citan en la bibliografía de la materia.

Las actividades prácticas se realizan principalmente en grupos de 2 alumnos para fomentar la discusión y el intercambio de ideas.



UNIVERSIDAD NACIONAL DE LA PLATA
FACULTAD DE INFORMÁTICA

Durante el desarrollo del curso los alumnos resuelven en clase problemas concretos asociados a los temas vistos en la teoría. Además se publican autoevaluaciones, en forma periódica, con el objetivo de que los alumnos comprueben los conocimientos adquiridos en forma voluntaria.

También se define un trabajo integrador que se comienza a desarrollar a partir de la cuarta semana de clase y que debe ser entregado al finalizar el curso.

METODOLOGÍA DE ENSEÑANZA Y EVALUACIÓN

METODOLOGÍA DE ENSEÑANZA

El dictado de la asignatura tiene modalidad de Taller lo que permite a los alumnos aplicar las estrategias propuestas en la resolución de problemas concretos sencillos a medida que se desarrolla la teoría. Las clases son guiadas a través de la proyección de transparencias utilizando el cañón y la PC disponibles en el aula.

Muchos temas tienen una fuerte justificación matemática cuya comprensión puede facilitarse a través de representaciones gráficas o de algoritmos de aproximación. Por esta razón, la materia se dicta íntegramente en la Sala de PC.

METODOLOGÍA DE EVALUACIÓN

Cada alumno puede optar por una de las siguientes formas de aprobación:

a) Régimen de promoción

Se ofrecen dos modalidades para promocionar la materia. El alumno deberá optar por una de ellas al inicio del curso. Los requisitos para cada modalidad son los siguientes:

Modalidad Presencial:

- Asistir al 70% de las clases teóricas y prácticas.
- Entregar al final del curso la resolución del trabajo práctico integrador. Este trabajo podrá ser consultado en los horarios de práctica.
- Aprobar el examen que se tomará al finalizar el curso. Este examen cuenta con dos recuperatorios. Podrán utilizarse las autoevaluaciones teórico-prácticas para acumular hasta 2 (dos) puntos extras en la calificación final a razón de 1 (punto) por cada 3 (tres) autoevaluaciones aprobadas con nota igual o mayor a 6 (seis). Estos puntos sólo podrán utilizarse cuando la calificación del examen sea mayor o igual a 5 (cinco) puntos.
- Para promocionar la materia deberá obtener una calificación mayor o igual a 6 (seis) puntos.



UNIVERSIDAD NACIONAL DE LA PLATA
FACULTAD DE INFORMÁTICA

Modalidad semi-presencial:

- Realizar el 70% de las autoevaluaciones teórico-prácticas (5 autoevaluaciones de un total de 8).
- Entregar al final del curso la resolución del trabajo práctico integrador. De ser necesario se fijarán dos encuentros presenciales para consulta por parte del alumno.
- Aprobar el examen que se tomará al finalizar el curso. Este examen cuenta con dos recuperatorios. Podrán utilizarse las autoevaluaciones teórico-prácticas para incrementar la calificación final en 1 (un) punto en caso de aprobar 3 (tres) autoevaluaciones con nota igual o mayor a 6 (seis). Este punto adicional sólo podrá utilizarse cuando la calificación del examen sea mayor o igual a 5 (cinco) puntos.
- Para promocionar la materia deberá obtener una calificación mayor o igual a 6 (seis) puntos.

b) Régimen convencional

Los alumnos que opten por el régimen convencional no tendrán la obligación de cumplir con ningún requisito de asistencia ni de realización de autoevaluaciones.

Al finalizar el curso el alumno deberá rendir un examen escrito referido a los aspectos prácticos de la materia. Este examen cuenta con dos recuperatorios. Quienes lo aprueben con nota mayor o igual a 4 (cuatro) puntos obtendrán la cursada de la asignatura debiendo luego rendir examen final.