

**MINERIA DE DATOS USANDO SISTEMAS
INTELIGENTES**

Año 2026

Carrera/Plan:

Licenciatura en Sistemas Planes 2021/2015
Licenciatura en Informática Planes 2021/2015

Área: Algoritmos y Lenguajes

Año: 4º o 5º año

Régimen de Cursada: Semestral

Carácter: Optativa

Correlativas: Algoritmos y Estructuras de Datos
– Matemática 3

Profesora: Dra. *Laura Lanzarini*

Hs semanales: 2 hs de teoría – 4 hs de práctica

FUNDAMENTACIÓN

La Minería de Datos reúne un conjunto de técnicas orientadas a descubrir, describir y comprender patrones presentes en grandes volúmenes de datos, con el objetivo de generar conocimiento nuevo, potencialmente útil y accionable.

A diferencia de los enfoques tradicionales basados en la verificación de hipótesis predefinidas, la Minería de Datos propone un enfoque exploratorio e inductivo, en el cual los patrones emergen a partir del análisis sistemático de los datos.

Este curso pone especial énfasis en la interpretación de los modelos obtenidos, la calidad y validez de los patrones descubiertos, y su comunicación efectiva para la toma de decisiones, priorizando técnicas comprensibles e interpretables.

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 estará puesto en analizar los resultados obtenidos considerando sesgos, limitaciones y supuestos de los datos y modelos.

RESULTADOS DE APRENDIZAJE

Al finalizar el curso, los estudiantes serán capaces de:

- Comprender los fundamentos y alcances de las principales técnicas de Minería de Datos.
- Aplicar técnicas de clasificación, agrupamiento y reglas para descubrir y describir patrones en datos reales.
- Evaluar modelos no solo por su performance, sino también por su interpretabilidad, estabilidad y utilidad práctica.
- Identificar posibles sesgos en los datos y en los modelos y analizar su impacto.
- Comunicar los resultados de forma clara, justificando su validez y relevancia para la toma de decisiones.

COMPETENCIAS

- LI-CE4- Planificar, dirigir, realizar y/o evaluar proyectos de relevamiento de problemas del mundo real, especificación formal de los mismos, diseño, implementación, prueba, verificación, validación, mantenimiento y control de calidad de sistemas de software/sistemas de información que se ejecuten sobre equipos de procesamiento de datos, con capacidad de incorporación de tecnologías emergentes del cambio tecnológico. Capacidad de análisis, diseño y evaluación de interfases humano computador y computador-computador.

- LS-CE1- Planificar, dirigir, realizar y/o evaluar proyectos de relevamiento de problemas del mundo real. Especificación formal, diseño, implementación, prueba, verificación, validación, mantenimiento y control de calidad de sistemas de software que se ejecuten sobre sistemas de procesamiento de datos, con capacidad de incorporación de tecnologías emergentes del cambio tecnológico. Capacidad de análisis, diseño y evaluación de interfases humano computador y computador-computador.

CONTENIDOS MINIMOS (de acuerdo al Plan de Estudios)

Introducción a la Minería de Datos.

Técnicas

- Árboles de decisión.
- Reglas de clasificación y asociación
- Técnicas de Agrupamiento

Evaluación e interpretación de Modelos

PROGRAMA ANALÍTICO

- Introducción. Obtención de conocimiento a partir de los datos. El concepto de patrón. El proceso KDD. 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. Análisis de la información de entrada. Medidas estadísticas básicas. Construcción y análisis de representaciones gráficas. Medidas de similitud entre atributos y entre ejemplos. 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. Tareas predictivas y descriptivas. Aprendizaje supervisado y aprendizaje no supervisado. La Minería de Datos y el aprendizaje inductivo. Comparación de las técnicas de Minería de Datos.
- Técnicas de Agrupamiento. Métricas de calidad del agrupamiento. Tipos de agrupamiento: particitivo y jerárquico. Medidas de distancia y de conectividad. Proceso de agrupamiento. Clustering particitivo. Algoritmo k-medias. Algoritmos de clustering jerárquicos aglomerativos y divisivos. Dendrogramas.
- Árboles de decisión. Métricas de selección de atributos. Entropía. Ganancia de Información. Tasa de Ganancia. Índice Gini. Poda y Sobreajuste. Algoritmos Id3, C4.5 y CART.
- Clasificadores bayesianos. Teorema de Bayes. Hipótesis máxima a posteriori. Clasificador Naïve Bayes. Ejemplos.
- Reglas de clasificación. Partición vs cobertura. Métodos OneR, PRISM, PART.

- Reglas de asociación. Calidad de las reglas. Algoritmo A priori. Concepto de ítem frecuente. Mejoras del algoritmo a priori: FP-Growth y FP-Tree. Métricas de una regla: soporte, cobertura, confianza e interés.
- Análisis y difusión del modelo obtenido. Evaluación de modelos. Comparación de técnicas de aprendizaje. Evaluación y mejora del modelo obtenido. Performance del modelo. Matriz de confusión. Sensibilidad, especificidad, precisión y recall. F measure. Visualización utilizando curvas ROC.
- Interpretabilidad de modelos. Técnicas de interpretabilidad: importancia de atributos, visualización de reglas, SHAP/LIME.

BIBLIOGRAFIA BASICA

- Hernández Orallo, Ramírez Quintana, Ferri Ramírez. *Introducción a la Minería de Datos*. Prentice Hall. 2004. ISBN 84-205-4091-9.
- Ian H. Witten, Eibe Frank, Mark A. Hall. *Data Mining: Practical Machine Learning Tools and Techniques*, (Fourth Edition). Morgan Kaufmann. 2017. ISBN 978-0-12-804291-5.

BIBLIOGRAFIA COMPLEMENTARIA

- Mehmed Kantardzic. *Data Mining. Concepts, Models, Methods, and Algorithms*. (Third Edition). Wiley. IEE Press. 2020. ISBN: 978-1-119-51604-0
- Mohammed Zaki, Wagner Meira, Jr. *Data Mining and Machine Learning. Fundamental Concepts and Algorithms* (Second Edition). Cambridge University Press. 2020. ISBN: 978-1-108-47398-9.
- Jiawei Han, Jian Pei, Hanghang Tong. *Data Mining: Concepts and Techniques*, (Fourth Edition). Morgan Kaufmann. 2022. ISBN- 978-0-128-11761-3.

METODOLOGÍA DE ENSEÑANZA

La asignatura tiene como objetivo central capacitar al alumno en el uso de técnicas de Minería de Datos para la modelización, análisis e interpretación de la información disponible. En función de este objetivo, el dictado del curso se organiza a partir de una estrecha articulación entre los contenidos teóricos y las actividades prácticas.

Esta articulación se lleva a cabo mediante la ejemplificación y resolución de problemas concretos, similares a los planteados en las actividades prácticas, favoreciendo la aplicación de los conceptos teóricos a situaciones reales y el análisis crítico de los resultados obtenidos.

Al inicio del curso se publica el cronograma de actividades, en el cual se detallan los temas a desarrollar en cada clase, tanto teórica como práctica, así como las instancias de evaluación previstas.

ACTIVIDADES PRACTICAS

Durante el desarrollo del curso se propondrán entregas parciales de ejercicios prácticos, orientadas a afianzar los conceptos abordados en clase, y la elaboración de un trabajo práctico integrador, cuyo objetivo es aplicar de manera conjunta las técnicas vistas a lo largo de la asignatura sobre un problema concreto de Minería de Datos.

EVALUACIÓN

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

a) **Régimen de promoción**

Los alumnos que elijan esta opción deberán cumplir con los siguientes requisitos:

- Entregar durante la cursada la resolución de algunos ejercicios seleccionados de cada enunciado de trabajos prácticos.
- Entregar y aprobar el trabajo práctico integrador al final del curso.
- Aprobar con una nota mayor o igual a 6 (seis) puntos el examen que se tomará al finalizar el curso. Este examen contará con dos (2) instancias de recuperación.

El incumplimiento de cualquiera de estos requisitos impedirá el acceso al régimen de promoción.

b) **Régimen convencional**

Los alumnos que elijan esta opción deberán:

- Entregar durante la cursada la resolución de algunos ejercicios seleccionados de cada enunciado de práctica.
- Aprobar con nota mayor o igual a 4 (cuatro) puntos el examen que se tomará al final del curso. Este examen cuenta con 2 (dos) recuperatorios.

Quienes aprueben con nota mayor o igual a 4 (cuatro) puntos obtendrán la cursada de la asignatura debiendo luego rendir examen final.

CRONOGRAMA DE CLASES Y EVALUACIONES

Semana		Contenidos/Actividades
1	11-mar	Introducción. Proceso KDD. Análisis y visualización de atributos..
2	18-mar	Preprocesamiento de atributos. Normalización.
3	25-mar	Agrupamiento Partitivo. Algoritmo k-means. Métricas de calidad.
4	01-abr	Agrupamiento Jerárquico.
5	08-abr	Clasificador Naive Bayes
6	15-abr	Análisis de modelos predictivos. Métricas para selección de atributos
7	22-abr	Arboles de Clasificación. Enfoque de Quinlan. Algoritmos ID3, C4.5
8	29-abr	Arboles CART. Enfoque de Breiman (clasificación y regresión).
9	06-may	Ensamblados basados en árboles. Random Forest. Boosting
10	13-may	Reglas de clasificación. Métodos OneR, PRISM y PART
11	20-may	Reglas de asociación. Conjuntos de ítems frecuente. Método A Priori
12	27-may	Análisis de resultados y modelos obtenidos
13	05-jun	Consultas – Trabajo Final Integrador
14	12-jun	1ra.Fecha de examen
15	19-jun	Consultas y muestra de exámenes de la 1ra. Fecha.
16	26-jun	2da. Fecha de parcial
17	03-jul	Consultas y muestra de exámenes de la 2da. Fecha.
18	10-jul	3ra. Fecha de parcial
18	17-jul	Muestra de exámenes de la 3ra. Fecha

Evaluaciones previstas	Fecha (a confirmar)
1ra.Fecha de examen	12/06/2026
2da. Fecha de examen	26/06/2026
3ra. Fecha de examen	17/07/2026

Contacto de la cátedra

- e-mail: midusi.unlp@gmail.com
- WEB : weblidi.info.unlp.edu.ar/catedras/md_si
- Plataforma virtual de gestión de cursos: IDEAS.

Dra. Laura Lanzarini