

Codigo de asignatura: 4785B4

Nombre del programa académico	Maestría en Ingeniería Eléctrica		
Nombre completo de la asignatura	Representación de Datos		
Número de créditos ECTS por categoría	Ciencias naturales y matemáticas	Módulos profesionales y especiales	Humanidades y ciencias sociales y económicas
Semestre y año de actualización	2026-1		
Semestre y año en que se imparte	2026-1		
Tipo de asignatura	[] Obligatoria [X] Electiva		
Director o contacto del programa	Andrés Escobar Mejía		
Coordinador o contacto de la asignatura	Walter Serna Serna		
Descripción y contenidos			
1. Breve descripción			
Este curso aborda la representación de datos como un problema central de la analítica moderna, orientado a la comprensión y construcción de representaciones compactas que capturen la estructura latente de datos complejos. Se estudian fundamentos y métodos para la representación de datos mediante embeddings y espacios latentes, integrando enfoques basados en agrupamiento, aprendizaje de manifolds y aprendizaje de representaciones inductivas. El curso enfatiza el análisis crítico, la interpretación y la evaluación de representaciones de datos para la extracción de información relevante y la visualización, promoviendo la selección fundamentada de métodos según la estructura de los datos y la tarea de análisis.			
2. Objetivo del curso:			
Desarrollar en el estudiante la capacidad de comprender y analizar las estructuras latentes presentes en datos de alta dimensionalidad, mediante el estudio y aplicación de técnicas de reducción de dimensión basadas en paradigmas de clustering, manifolds y modelos de aprendizaje profundo, con el fin de extraer información relevante y facilitar su interpretación y visualización.			
3. Resultados de aprendizaje. Los propósitos de formación en el estudiante de posgrado son:			
Al finalizar el curso, el estudiante será capaz de analizar, seleccionar, implementar y evaluar técnicas de reducción de dimensión para la extracción de información relevante y la visualización de datos de alta dimensionalidad, comprendiendo los fundamentos geométricos, estadísticos y computacionales de dichos métodos, y justificando su aplicación según la estructura de los datos y la tarea de análisis o aprendizaje considerada.			
4. Contenido			
T1. Fundamentos del análisis de datos de alta dimensionalidad: Características, desafíos y limitaciones del análisis y la visualización de datos de alta dimensionalidad.			
T2. Marco conceptual y paradigmas de la representación de datos: Objetivos, categorizaciones y paradigmas fundamentales de la reducción de dimensión en analítica de datos.			
T3. Métodos de reducción de dimensión basados en el paradigma Cluster: Técnicas orientadas a la identificación de estructuras de agrupamiento y subespacios relevantes.			
T4. Métodos de reducción de dimensión basados en el paradigma Manifold: Métodos no lineales de embedding basados en la hipótesis de manifold y preservación de estructuras.			
T5. Aprendizaje de representaciones y autoencoders para reducción de dimensión: Enfoques inductivos de aprendizaje de representaciones mediante autoencoders y autosupervisión.			
T6. Evaluación, estabilidad y confiabilidad de técnicas de reducción de dimensión: Criterios y métricas para evaluar desempeño, estabilidad y confiabilidad de las proyecciones.			
T7. Tendencias actuales y aplicaciones avanzadas en la representación de datos: Tendencias, escalabilidad y aplicaciones emergentes en reducción de dimensión para ingeniería.			
5. Requisitos. Los definidos en requisito de admisión de la IES.			
6. Recursos			
Aula con proyector multimedia y acceso a internet.			
Sala de cómputo o equipos portátiles para los estudiantes.			
Lenguajes de programación para análisis de datos (Python o equivalente).			
Entornos de desarrollo interactivo (Jupyter Notebook o similares).			
Librerías para análisis numérico y científico.			
Librerías para aprendizaje automático y reducción de dimensión.			
Librerías para visualización de datos.			
Recursos computacionales para entrenamiento de modelos (CPU/GPU según disponibilidad).			
Repositorios de datos públicos para prácticas y estudios de caso.			
Plataforma institucional para gestión de contenidos y entregas académicas.			

Bibliografía

- Murphy, K. P. (2022). Probabilistic machine learning: an introduction. MIT press.
- Lee, J. A., & Verleysen, M. (2007). Nonlinear dimensionality reduction (Vol. 1). New York: Springer.
- Bishop, C. M., & Nasrabadi, N. M. (2006). Pattern recognition and machine learning (Vol. 4, No. 4, p. 738). New York: springer.

7. Herramientas técnicas de soporte para la enseñanza

- Clase magistral y presentaciones en clase
- Estudio y análisis de artículos científicos
- Trabajo sobre material audiovisual y trabajos en grupo
- Laboratorios de programación en Python y Matlab.

8. Trabajos en laboratorio y proyectos

Laboratorio 1: Exploración y caracterización de datos de alta dimensionalidad.

Laboratorio 2: Reducción de dimensión basada en estructuras de agrupamiento (Paradigma Cluster).

Laboratorio 3: Aprendizaje de embeddings no lineales (Paradigma Manifold).

Laboratorio 4: Evaluación, estabilidad y confiabilidad de métodos de reducción de dimensión.

9. Métodos de aprendizaje

El curso se desarrollará mediante clases magistrales orientadas a la discusión conceptual, laboratorios de programación enfocados en el análisis crítico de datos reales, y estudios de caso que integren teoría y práctica. Se promoverá el aprendizaje activo a través de la comparación de métodos, la interpretación de resultados y la justificación técnica de decisiones, así como el trabajo autónomo y colaborativo en actividades de análisis y evaluación.

10. Métodos de evaluación

La evaluación del curso se realizará mediante laboratorios de programación con informes técnicos, ejercicios de análisis comparativo de métodos y un proyecto integrador final orientado a la evaluación crítica y justificada de técnicas de reducción de dimensión. Se valorará la comprensión conceptual, la correcta interpretación de resultados, la solidez de los criterios de evaluación utilizados y la capacidad de argumentación técnica más allá del desempeño instrumental de los algoritmos.