



Universidad Tecnológica de la Mixteca

Clave DGP:

Doctorado en Inteligencia Artificial

00034

PROGRAMA DE ESTUDIOS

NOMBRE DE LA ASIGNATURA

ANÁLISIS DE DATOS Y MODELOS DE APRENDIZAJE MÁQUINA

SEMESTRE	CLAVE DE LA ASIGNATURA	TOTAL DE HORAS
Segundo	351205RA	80

OBJETIVO(S) GENERAL(ES) DE LA ASIGNATURA

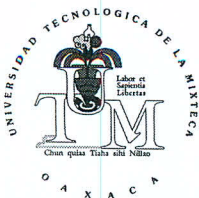
Conocer algunas de las técnicas para agrupar, clasificar y detectar valores atípicos en una base relativamente grande de información. Sea capaz de determinar y aplicar el modelo de aprendizaje máquina adecuado para cada fin y, además, logre hacerlo para problemas que presentan una mayor dificultad, tales como con bases que presentan un cambio constante (data stream) y con datos de texto.

TEMAS Y SUBTEMAS

- 1. Análisis de agrupación**
 - 1.1. Modelos para la selección de características.
 - 1.2. Algoritmos basados en representantes (k-Means, k-Medians y k-Medoids).
 - 1.3. Algoritmos de agrupación jerárquica.
 - 1.4. Algoritmos basados en modelos probabilísticos.
 - 1.5. Algoritmos basados en cuadrícula y en densidad.
 - 1.6. Algoritmos basados en grafos.
 - 1.7. Factorización de matrices no negativas
 - 1.8. Validación.
- 2. Análisis de valores atípicos**
 - 2.1. Introducción.
 - 2.2. Análisis de valores extremos.
 - 2.3. Modelos probabilísticos.
 - 2.4. Agrupación para detección de valores atípicos.
 - 2.5. Detección basada en distancia.
 - 2.6. Métodos basados en densidad.
 - 2.7. Modelos de información teórica.
 - 2.8. Validación de valores atípicos.
- 3. Clasificación de datos**
 - 3.1. Introducción.
 - 3.2. Selección de características para clasificación.
 - 3.3. Árboles de decisión.
 - 3.4. Clasificadores basados en reglas.
 - 3.5. Clasificadores probabilísticos.
 - 3.6. Máquinas de soporte vectorial.
 - 3.7. Redes neuronales.
 - 3.8. Aprendizaje basado en instancias
 - 3.9. Evaluación del clasificador.
- 4. Minería de flujo de datos (Data streams)**
 - 4.1. Introducción.
 - 4.2. Estructuras de datos para streams.
 - 4.3. Patrones frecuentes.
 - 4.4. Agrupación.
 - 4.5. Detección de valores atípicos.
 - 4.6. Clasificación.
- 5. Minería de datos de texto**
 - 5.1. Introducción.
 - 5.2. Preparación de documentación y cálculo de similitud.



VICE-RECTORIA
ACADÉMICA



Universidad Tecnológica de la Mixteca

Clave DGP:

Doctorado en Inteligencia Artificial

00035

PROGRAMA DE ESTUDIOS

- 5.3. Métodos especializados de agrupación para texto.
- 5.4. Modelación.
- 5.5. Métodos especializados de clasificación para texto.
- 5.6. Detección de novedades y primeras historias.

ACTIVIDADES DE APRENDIZAJE

Sesiones dirigidas por el profesor en las que se presentan los conceptos y, al mismo tiempo, se realizarán programas que ilustran cada uno de los temas. Se utilizará algún Notebook, por ejemplo, Collaboratory o Jupyter para realizar la programación con el lenguaje Python mediante las librerías SciKit-Learn y Tensor Flow; así como las redes Kaggle y GitHub para compartir o descargar algoritmos. El contenido se abordará a profundidad y se iniciará con un proyecto desde su inicio con fines didácticos y de práctica.

CRITERIOS Y PROCEDIMIENTOS DE EVALUACIÓN Y ACREDITACIÓN

El Capítulo II, De las Evaluaciones, del Reglamento General de Posgrado establece que, Artículo 33, la calificación final del alumno se obtendrá de tres evaluaciones parciales (50%) y un examen ordinario (50%), Artículo 32. Para cada evaluación parcial se indicará al inicio de semestre la modalidad de evaluación a utilizar, Artículo 24.

BIBLIOGRAFÍA (TIPO, TÍTULO, AUTOR, EDITORIAL Y AÑO)

Básica:

1. **Data mining: the textbook.** AGGARWAL, Charu C., et al. New York: Springer, 2015.
2. **Data mining: concepts, models, methods, and algorithms.** KANTARDZIC, Mehmed. John Wiley & Sons, 2011.
3. **Statistical and machine-learning data mining:: Techniques for better predictive modeling and analysis of big data.** RATNER, Bruce. CRC Press, 2017.

Consulta:

1. **Python for data science for dummies.** MUELLER, John Paul; MASSARON, Luca. John Wiley & Sons, 2019.
2. **Introducción a la Minería de Datos.** ORALLO, José Hernández; QUINTANA, María José Ramírez; RAMÍREZ, César Ferri. Pearson Educación, 2004.
3. **Data science strategy for dummies.** JÄGARE, Ulrika. John Wiley & Sons, 2019.

PERFIL PROFESIONAL DEL DOCENTE

Estudios mínimos de Doctorado con conocimientos en ciencia de datos o minería de datos y aprendizaje máquina.

Vo.Bo
DR. JOSÉ ANÍBAL ARIAS AGUILAR
JEFE DE LA DIVISIÓN DE ESTUDIOS DE POSGRADO



DIVISIÓN DE ESTUDIOS DE POSGRADO

AUTORIZÓ
DR. AGUSTÍN SANTIAGO ALVÁRADO
VICE-RECTOR ACADÉMICO

VICE-RECTORIA ACADÉMICA