



### PROGRAMA DE ESTUDIOS

#### NOMBRE DE LA ASIGNATURA

Minería de datos estáticos

SEMESTRE	CLAVE DE LA ASIGNATURA	TOTAL DE HORAS
Noveno semestre	075097	80

#### OBJETIVO(S) GENERAL(ES) DE LA ASIGNATURA

Al finalizar el curso, el alumno será capaz de tomar muestras de grupos de datos estáticos, extraer las características principales y reducir su dimensionalidad. También, identificará la distancia en los datos de tipo cuantitativos y categóricos, aplicará algoritmos para identificar los patrones de asociación requeridos y realizará el análisis de los racimos obtenidos. Finalmente, logrará realizar las clasificaciones mediante algún método de aprendizaje máquina y de aprendizaje profundo.

#### TEMAS Y SUBTEMAS

##### 1. Reducción de datos y transformación

- 1.1.1. Muestreo para datos estáticos
- 1.1.2. Extracción de subconjuntos de características
- 1.1.3. Reducción de dimensionalidad con rotación de ejes
  - 1.1.3.1. Análisis de componentes principales
  - 1.1.3.2. Descomposición en valores singulares
- 1.1.4. Reducción de dimensionalidad con transformaciones
  - 1.1.4.1. Transformadas de ondas de Haar
  - 1.1.4.2. Escalado multidimensional
  - 1.1.4.3. Transformación espectral e incrustación de gráficas

##### 2. Similitud y distancias

- 2.1. Datos multidimensionales
  - 2.1.1. Datos cuantitativos
    - 2.1.1.1. Distancias con normas  $L_p$
  - 2.1.2. Datos categóricos
  - 2.1.3. Datos cuantitativos y categóricos mezclados
- 2.2. Medidas de similitud de grafos

##### 3. Minería de patrones de asociación

- 3.1. Modelo frecuente de minería de patrones y sus reglas de asociación
  - 3.1.1. Conjuntos frecuentes de ítems
    - 3.1.1.1. Algoritmo de fuerza bruta
    - 3.1.1.2. Algoritmo a priori
    - 3.1.1.3. Algoritmos de enumeración de árboles
    - 3.1.1.4. Métodos recursivos basados en crecimiento de patrones en sufijos
- 3.2. Resumen de patrones

##### 4. Análisis de racimos

- 4.1. Selección de características para racimos
  - 4.1.1. Modelos de filtro
  - 4.1.2. Modelos envueltos
- 4.2. Algoritmo de k-medias
- 4.3. Algoritmos jerárquicos
- 4.4. Algoritmos basados en grafos
- 4.5. Validación de racimos

##### 5. Clasificación de datos

- 5.1. Selección de características para clasificación
  - 5.1.1. Modelos de filtrado
    - 5.1.1.1. Índice de Gini



VICE-RECTORIA  
ACADÉMICA

### PROGRAMA DE ESTUDIOS

- 5.1.1.2. Entropía
- 5.1.1.3. Puntuación de Fisher
- 5.2. Árboles de decisión
- 5.3. Clasificadores basados en reglas
- 5.4. Clasificadores probabilísticos
- 5.5. Redes neuronales

### ACTIVIDADES DE APRENDIZAJE

Sesiones dirigidas por el profesor en las que presenta los conceptos y, al mismo tiempo, se realizarán programas que los ilustrarán, se sugiere utilizar algún Notebook como Collaboratory o Jupyter para realizar programas con el lenguaje Python, así como GitHub para compartir y descargar algoritmos programables. Se recomienda ampliamente impartir el curso en un laboratorio con equipo de cómputo disponible para cada estudiante. Se iniciará con un proyecto desde el inicio del curso con fines didácticos y de práctica de cada uno de los temas.

### CRITERIOS Y PROCEDIMIENTOS DE EVALUACIÓN Y ACREDITACIÓN

En términos de los artículos 25 incisos (b), (e), (f) y (g); del 48 al 62, del Reglamento de alumnos de licenciatura aprobado por el H. Consejo Académico el 19 de mayo del 2016, los lineamientos que habrán de observarse en lo relativo a los criterios y procedimientos de evaluación y acreditación, entre lo más importante:

- i. Al inicio del curso el profesor deberá indicar el procedimiento de evaluación que deberá comprender, al menos tres evaluaciones parciales que tendrán una equivalencia del 50% de la calificación final y un examen ordinario que equivaldrá al restante 50%.
- ii. Las evaluaciones podrán ser escritas y/o prácticas y cada una consta de un examen teórico-práctico, tareas y proyectos. La parte práctica de cada evaluación deberá estar relacionada con la ejecución exitosa y la documentación de la solución de problemas sobre temas del curso.
- iii. Además, pueden ser consideradas otras actividades como: el trabajo extra-clase, la participación durante las sesiones del curso y la asistencia a las asesorías.
- iv. El examen tendrá un valor mínimo de 50%; las tareas, proyectos y otras actividades, un valor máximo de 50%.

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

#### Libros Básicos:

1. **Data Mining: The Textbook**, Aggarwal, Charu, Cham, Switzerland: Springer, 2015.
2. **Data Mining: Concepts, Models, Methods, and Algorithms**, Kantardzic, Mehmed Hoboken, NJ: John Wiley & Sons, 2020.
3. **Statistical and Machine-Learning Data Mining, Techniques for Better Predictive Modeling and Analysis of Big Data**, Bruce Ratner, second edition, CRC Press, Taylor & Francis Group, 2011.

#### Libros de Consulta:

1. **Python for data Science**, Luca Massaron and John Paul Mueller, John Wiley & Sons, 2015.
2. **Introducción a la minería de datos**, José Hernandez Orallo, Ma. José Ramírez Quintana, Cesar Ferri Ramirez, Pearson, Prentice Hall.
3. **Data Science strategy**, Ulrika Jägar, John Wiley & Sons, 2019.

### PERFIL PROFESIONAL DEL DOCENTE

Estudios de maestría o doctorado en matemáticas, matemáticas aplicadas o computación con conocimientos en ciencia de datos, Machine learning y Deep learning.

Vo.Bo.

DR. FRANCO BARRAGÁN MENDOZA  
JEFE DE CARRERA

AUTORIZÓ

DR. AGUSTÍN SANTIAGO ALVARADO  
VICE-RECTOR ACADÉMICO