

**UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO****Licenciatura en Ciencias de la Computación****Facultad de Ciencias**

Programa de la asignatura

**Denominación de la asignatura:*****Almacenes y Minería de Datos***

Clave:	Semestre: 7-8	Eje temático: Bases de Datos y Manejo de la Información			No. Créditos: 10
Carácter: Optativa		Horas		Horas por semana	Total de Horas
Tipo: Teórico-Práctica		Teoría:	Práctica:	7	112
		3	4		
Modalidad: Curso		Duración del programa: Semestral			

**Asignatura con seriación indicativa antecedente:** Sistemas Manejadores de Bases de Datos**Asignatura con seriación indicativa subsecuente:** Ninguna**Objetivo general:**

Conocer los conceptos generales de la tecnología de almacenes de datos y el minado de datos.

**Índice temático**

Unidad	Temas	Horas	
		Teóricas	Prácticas
I	Almacenes de Datos	9	12
II	Pre-procesamiento de datos	9	12
III	Computación de Cubos de datos <i>Data cubes</i> ) y generalizaciones de datos	6	8
IV	Minado de Patrones Frecuentes, Asociaciones, y Correlaciones	6	8
V	Modelos de clasificación y predicción	9	12
VI	Análisis de agrupamiento	9	12
Total de horas:		48	64
<b>Suma total de horas:</b>		<b>112</b>	

**Contenido temático**

Unidad	Tema
I Almacenes de Datos	
I.1	Introducción.

I.2	Un modelo de datos multidimensional.
I.3	Arquitectura de un almacén de datos.
I.4	Implementación de un almacén de datos.
I.5	Relación entre los almacenes de datos y la minería de datos.
II Pre-procesamiento de datos	
II.1	Introducción.
II.2	El pre-procesamiento de datos.
II.3	Resumen descriptivo de datos.
II.4	Limpieza de datos.
II.5	Integración y transformación de datos.
II.6	Reducción de datos.
II.7	Generación de jerarquías.
III Computación de Cubos de datos ( <i>Data cubes</i> ) y generalizaciones de datos	
III.1	Cálculo eficiente de cubos de datos.
III.2	Exploración y descubrimiento de información en bases de datos multidimensionales.
III.3	Construcción y explotación de cubos
IV Minado de Patrones Frecuentes, Asociaciones, y Correlaciones	
IV.1	Conceptos básicos.
IV.2	Minado eficiente y escalable de conjuntos de elementos frecuentes.
IV.3	Minando varios tipos de reglas de asociación.
IV.4	Del minado de reglas de asociación al análisis de correlaciones.
V Modelos de clasificación y predicción	
V.1	Aprendizaje supervisado.
V.2	Árboles de decisión.
V.3	Clasificador Bayesiano.
V.5	Clasificación basada en reglas.
V.5	Redes neuronales.
V.6	Otros métodos de clasificación.
VI Análisis de agrupamiento	
VI.1	Introducción.
VI.2	Tipos de Datos.
VI.3	Categorización de métodos de agrupamiento.
VI.4	``Agrupando" datos de alta dimensionalidad.
VI.5	Análisis de valores atípicos ( <i>Outliers</i> ) o información ruidosa.
VI.6	Métodos Jerárquicos.
VI.7	Métodos basados en densidad.
VI.8	Modelos basados en métodos de agrupamiento.
<b><u>Bibliografía básica:</u></b>	
1. Jiawei-Han and Micheline Kamber, ed. <i>Data Mining: Concepts and Techniques, 2ed</i> , Morgan Kaufman, 2006.	
2. Witten, Ian H., Frank, Eibe, <i>Data Mining: Practical Machine Learning Tools and Techniques, Second Edition</i> , Morgan Kaufman, 2005.	

**Bibliografía complementaria:**

1. Nisbet, R; Elder IV, J., *Handbook of Statistical Analysis and Data Mining Applications*, Elsevier, Academic Press, 2009.
2. Hastie, T.; Tibshirani, R\$.; Friedman, J.; *The Elements of Statistical Learning, Data Mining, Inference, and Prediction*, Springer, 2009.
3. Artículos escogidos de ACM SIGMOD y similares.

**Sugerencias didácticas:**

Exposición oral	(X)
Exposición audiovisual	(X)
Ejercicios dentro de clase	(X)
Ejercicios fuera del aula	( )
Seminarios	(X)
Lecturas obligatorias	(X)
Trabajo de investigación	( )
Prácticas de taller o laboratorio	(X)
Prácticas de campo	( )

Otras: \_\_\_\_\_

**Métodos de evaluación:**

Exámenes parciales	(X)
Examen final escrito	( )
Trabajos y tareas fuera del aula	(X)
Prácticas de laboratorio	( )
Exposición de seminarios por los alumnos	( )
Participación en clase	(X)
Asistencia	( )
Proyectos de programación	(X)
Proyecto final	( )
Seminario	( )

Otras: \_\_\_\_\_

**Perfil profesiográfico:**

Egresado preferentemente de la Licenciatura en Ciencias de la Computación o matemático con especialidad en computación con amplia experiencia de programación. Es conveniente que posea un posgrado en la disciplina. Con experiencia docente.