

### 1.- DATOS DE LA ASIGNATURA

<p><b>Nombre de la asignatura:</b> Minería de Datos y Aprendizaje Automático</p> <p><b>Carrera:</b> Lic. en Informática</p> <p><b>Clave de la asignatura:</b></p> <p><b>Horas teoría-horas práctica- 1-3-5 créditos</b></p>
---

### 2.- HISTORIA DEL PROGRAMA

<b>Lugar y fecha de elaboración o revisión</b>	<b>Participantes</b>	<b>Observaciones (cambios y justificación)</b>
Apizaco, Tlax a 24/11/06	Dr. José Federico Ramírez Cruz M.C. P. Malaquías Quintero Flores M.C. Orion F. Reyes Galaviz M.C. José Juan Hernández Mora M.I.A. Carlos Pérez Corona Ing. Roberto Acoltzi Nava	Propuesta Original

### 3.- UBICACIÓN DE LA ASIGNATURA

a). Relación con otras asignaturas del plan de estudio

Anteriores		Posteriores	
Asignaturas	Temas	Asignaturas	Temas
Probabilidad	Teoría de Probabilidad		
	Distribución de Probabilidad		
Estadística	Todos		
Estructura de datos	Estructura de datos		
Programación I y II	Lineales y No lineales		

**b). Aportación de la asignatura al perfil del egresado**

Contar con habilidades que le permitan identificar problemas y oportunidades donde se aplique el tratamiento de información para proponer soluciones por medio de modelos de aprendizaje automático y minería de datos con el fin de facilitar con ello la toma de decisiones.

Seleccionar y utilizar de manera optima las herramientas computacionales actuales y emergentes para desarrollar y administrar sistemas de información con el propósito de aumentar la productividad y competitividad en sus propuestas de solución dentro de las organizaciones

**4.- OBJETIVO(S) GENERAL(ES) DEL CURSO**

Conocer las técnicas, métodos, algoritmos y herramientas que hacen que los programas de computadora aprendan a mejorar su rendimiento y adquieran conocimiento de la información disponible con el propósito de dar solución a problemas del mundo real.

**5.- TEMARIO**

Unidad	Temas	Subtemas
I	Introducción a la minería de datos y	1.1. ¿Qué es el aprendizaje automático?

	el aprendizaje automático	<p>1.2 ¿Qué es la minería de datos?</p> <p>1.3 Tareas del aprendizaje y la minería de datos</p> <p>1.4 Campos de aplicación del aprendizaje automático y la minería de datos.</p> <p>1.2. Tipos de Aprendizaje</p>
II	Preparación de los datos de entrada	<p>Descripción de conjuntos de datos</p> <p>Conceptos, ejemplos y atributos</p> <p>Limpiado de datos</p> <p>Datos ruidosos</p> <p>Valores de atributos faltantes</p> <p>Datos Inconsistentes</p> <p>Normalización de datos</p> <p>Discretización de datos continuos</p> <p>Reducción de datos</p> <p>Reducción de dimensionalidad</p> <p>Selección de instancias y/o atributos</p>
III	Representación de conocimiento	<p>3.1 Conceptos básicos</p> <p>3.2 Tablas de decisión</p> <p>3.3 Árboles de decisión</p> <p>3.3 Reglas de Clasificación</p> <p>3.4 Reglas de Asociación</p> <p>3.5 Árboles para predicción numérica</p> <p>3.6 Representación basada en instancias</p> <p>3.7 Grupos</p>
IV	Algoritmos y métodos de clasificación, predicción y minado de datos.	<p>4.1 Inducción de árboles de decisión</p> <p>4.2 Construcción de reglas</p> <p>4.3 Depuración de reglas</p> <p>4.3 Minando reglas de asociación</p>
V	Métodos alternativos de clasificación, predicción y minado de datos.	<p>5.1 Métodos lineales y redes neuronales</p> <p>5.2 Métodos basados en instancias</p> <p>5.3 Clasificador Bayesiano</p> <p>5.4 Métodos evolutivos</p> <p>5.5 Métodos híbridos</p> <p>5.6 Algoritmos de agrupamiento</p> <p>5.7 Maquinas de vectores de soporte</p>

VI	Evaluación de los algoritmos	5.8 Métodos de ensambles de clasificadores 6.1 Entrenamiento y prueba 6.2 Métodos para evaluación del rendimiento 6.3 Comparación de modelos de minería
----	------------------------------	--

#### 6.- APRENDIZAJES REQUERIDOS

- Conocimientos de matemáticas básicas: logaritmos, álgebra, geometría analítica, cálculo diferencial e integral, álgebra lineal, entre otros.
- Conocimientos sobre probabilidades y estadística
- Conocimientos sobre bases de datos.
- Conocimientos sobre programación

#### 7.- SUGERENCIAS DIDÁCTICAS

- Se sugiere que alumno realice prácticas con alguna herramienta de minería de datos y aprendizaje automático
- Se sugiere el uso de herramientas de software para acceder a bases de datos.
- Se sugiere el uso de algún lenguaje de programación o herramienta matemática para implementar algunos de los algoritmos de aprendizaje automático.
- Se sugiere que el alumno realice algunos de los algoritmos de aprendizaje automático para su mejor comprensión
- Se sugiere que el alumno experimente minar diferentes bases de datos con atributos y/o clases discretas y continuas

#### 8.- SUGERENCIAS DE EVALUACIÓN

Se sugiere evaluar los conceptos adquiridos en clase por medio de los siguientes criterios:

- Exámenes parciales
- Prácticas de laboratorio
- Tareas y ensayos
- Proyectos de curso

#### 9.- UNIDADES DE APRENDIZAJE

**UNIDAD 1.- Introducción al la minería de datos y el aprendizaje automático**

<b>Objetivo Educativo</b>	<b>Actividades de Aprendizaje</b>	<b>Fuentes de Información</b>
Conocer la definición, objetivos y aplicaciones de la minería de datos y el aprendizaje automático	Clases teóricas Investigación sobre otras aplicaciones de la minería no vistas en clase	[1], [2], [3]

**UNIDAD 2.- Preparación de los datos de entrada**

<b>Objetivo Educativo</b>	<b>Actividades de Aprendizaje</b>	<b>Fuentes de Información</b>
El alumno conocerá los métodos del pre-procesamiento de datos para poder tratarlos con los algoritmos de aprendizaje	Clases teóricas Implementación de algunos algoritmos de procesamiento 2.3 Uso de herramienta(s) de minería con la opción de preprocesamiento	[1]

**UNIDAD 3.- Representación de conocimiento**

<b>Objetivo Educativo</b>	<b>Actividades de Aprendizaje</b>	<b>Fuentes de Información</b>
Que el alumno conozca las posibles formas de representar conocimiento	3.1 Clases teóricas 3.2 Trabajos en clase y tareas en casa	

**UNIDAD 4.- Algoritmos y métodos de clasificación, predicción y minado de datos**

<b>Objetivo</b>	<b>Actividades de Aprendizaje</b>	<b>Fuentes de</b>
-----------------	-----------------------------------	-------------------

<b>Educacional</b>		<b>Información</b>
Que el alumno conozca el funcionamiento de los algoritmos de aprendizaje automático para la minería de datos	3.1 Clases teóricas 3.2 Trabajos en clase y tareas en casa 3.3 Uso de herramienta(s) de minería con la opción de algoritmos de clasificación 3.4 Implementación de algunos algoritmos de aprendizaje automático	

**UNIDAD 5.- Métodos alternativos de clasificación, predicción y minado de datos**

<b>Objetivo Educacional</b>	<b>Actividades de Aprendizaje</b>	<b>Fuentes de Información</b>
Que el alumno conozca el funcionamiento de otros algoritmos alternativos de aprendizaje automático para la minería de datos	3.1 Clases teóricas 3.2 Trabajos en clase y tareas en casa 3.3 Uso de herramienta(s) de minería con la opción de algoritmos de clasificación 3.4 Implementación de algunos algoritmos de aprendizaje automático	

**UNIDAD 6.- Evaluación de los algoritmos**

<b>Objetivo Educacional</b>	<b>Actividades de Aprendizaje</b>	<b>Fuentes de Información</b>
Que el alumno conozca el funcionamiento de algunos métodos de validación de los algoritmos de aprendizaje automático	3.1 Clases teóricas 3.2 Trabajos en clase y tareas en casa 3.3 Uso de herramienta(s) de minería con la opción de algoritmos de clasificación 3.4 Implementación de algunos métodos de validación con algún o algunos algoritmos de aprendizaje automático	

**10. FUENTES DE INFORMACIÓN**

- [1] **Ian Witten, Eibe Frank**  
**Data Mining**: Practical Machine Learning Tools and Techniques  
Fecha de Pub.: Junio 2005  
ISBN: 0120884070  
Editorial: Elsevier Science & Technology Books  
Número de Edición: 2  
Serie: [Morgan Kaufmann Series in Data Managemen](#)
- [2] **Vipin Kumar, Vipin Kumar, Michael Steinbach, Pang-Ning Tan, Michael Steinbach**  
**Introduction to Data Mining**  
ISBN: 0321321367  
Fecha de Pub.: Mayo 2005  
Editorial: Pearson Education  
Numero de Edición: 1
- [3] **Tom M. Mitchell**  
**Machine Learning**  
ISBN: 0070428077  
Fecha de Pub.: March 1997  
Editorial: McGraw-Hill Companies, The  
Número de edición: 1  
Series: [McGraw-Hill Series in Computer Science](#)
- [4] **Ethem Alpaydin**  
**Introduction to Machine Learning**  
ISBN: 0262012111  
Fecha Pub.: August 2004  
Editorial: MIT Press  
Número de Edición: 1  
Series: [Adaptive Computation and Machine Learning Series](#)

## 11. PRÁCTICAS

- Realizar el procesamiento de los datos (base de datos) utilizando el método de selección de atributos Relief.
- Realizar el procesamiento de los datos (base de datos) utilizando el método de selección de atributos Análisis de Componentes Principales.
- Realizar la clasificación de una base de datos con el algoritmo de ID3
- Extraer las reglas de clasificación de una base de datos utilizando el algoritmo de C4.5
- Realizar la clasificación de una base de datos utilizando los algoritmos de:

- Naive Bayes
  - K-Vecinos cercanos
  - Máquinas de vectores de soporte
- 
- Realizar la predicción de una base de datos utilizando los métodos de:
    - Redes neuronales
    - Regresión Lineal Ponderada Localmente
- 
- Realizar la clasificación de una base de datos utilizando los métodos de agrupamiento y redes competitivas