

1. Datos Generales de la asignatura

Nombre de la asignatura:	MINERÍA DE DATOS
Clave de la asignatura:	CDF - 2304
SATCA¹:	3 - 2 - 5
Carrera:	Ingeniería en Tecnologías de Información y Comunicaciones

2. Presentación**Caracterización de la asignatura**

Esta asignatura aporta al perfil del Ingeniero en Tecnologías de la Información y Comunicaciones la posibilidad de utilizar tecnologías y herramientas de la Minería de datos e Inteligencia Artificial acordes a las necesidades del entorno, además de desarrollar e implementar sistemas para la toma de decisiones.

- Su importancia redundante en que en las organizaciones cuentan con grandes cantidades de información resultante de sus transacciones diarias, que requieren analizar en la extracción de conocimiento nuevo, para una buena toma de decisiones.
- La asignatura se enfoca a la extracción y análisis de la información utilizando herramientas de la Minería de datos e inteligencia artificial, que permita apoyar a las organizaciones haciendo un análisis del pasado para realizar predicciones en el futuro.
 - Debe ser cursada después de las asignaturas de base de datos, taller de Ingeniería de Software e ingeniería del conocimiento

Intención didáctica

Esta asignatura está organizada en cinco temas relacionados con los conceptos y tecnologías para realizar Minería de Datos en las organizaciones.

- En el tema uno, se introduce al estudiante a la Minería de datos, donde se analizan los objetivos, la importancia, así como se proporcionan algunos ejemplos de minería de datos para la toma de decisiones.
- En el tema dos, se estudia las metodologías CRISP-DM y KDD para hacer minería de datos.
- En el tema tres se estudia los métodos (selección, limpieza) para organizar los datos para la extracción del conocimiento. El docente debe fomentar que el estudiante proponga una base de datos donde pueda aplicar los métodos estudiados en este tema.
- En el tema cuatro se estudia los algoritmos para el aprendizaje no supervisado, así como definiciones básicas para un mejor entendimiento. Es conveniente que el docente promueva el diseño y aplicación de estos algoritmos en un caso real.
- En el tema cinco se estudian los métodos básicos predictivos, como los árboles

¹ Sistema de Asignación y Transferencia de Créditos Académicos

de decisión y la regla KNN. El docente debe promover el uso de algún software libre para aplicar estos métodos en una aplicación.

- Sugiere que el docente solicite a los estudiantes realizar un proyecto completo de Minería de Datos donde apliquen las competencias adquiridas en las asignaturas de Taller de base de datos, Ingeniería de Software y Taller de Ingeniería de Software.
- El docente deberá promover que el estudiante desarrolle las competencias genéricas para trabajar en equipo en el planteamiento, análisis y resolución de problemas complejos reales, así como las discusiones grupales y exposiciones que fomente la competencia de expresión oral.

3. Participantes en el diseño y seguimiento curricular del programa

Lugar y fecha de elaboración o revisión	Participantes	Observaciones
Instituto Tecnológico de Toluca, septiembre de 2022	<ul style="list-style-type: none"> • Dra. Eréndira Rendón Lara • M. en C. Itzel María Abúndez Barrera 	Miembros de la Academia de Sistemas Computacionales del Instituto Tecnológico de Toluca

4. Competencias a desarrollar

Competencias específicas de la asignatura
<ul style="list-style-type: none"> • Conoce los métodos, técnicas y tecnología para el diseño de sistemas utilizando la minería de datos para la toma de decisiones en las organizaciones.

5. Competencias previas

Las adquiridas en las asignaturas de:

- Ingeniería del conocimiento
- Taller de bases de datos
- Base de datos distribuidas
- Programación II
- Ingeniería de software
- Taller de Ingeniería de Software

6. Temario

No.	Temas	Subtemas
1	Introducción a la Minería de Datos	1.1 ¿Qué es la minería de datos? 1.1.1 Usos e importancia de la minería de datos 1.1.2 La minería de datos y el proceso de descubrimiento de conocimiento en bases de datos 1.1.3 Diferencia de la minería de datos y el Aprendizaje automático (Machine Learning) 1.1.4 Aplicaciones 1.2 El proceso de extracción de conocimiento
2	Metodologías para la Minería de Datos	2.1 Metodología CRISP-DM 2.1.1 Introducción 2.1.2 Comprensión del negocio 2.1.3 Comprensión de los datos 2.1.4 Preparación de los datos 2.1.5 Elaboración del modelo 2.1.6 Implementación y despliegue de la solución 2.2 Descubrimiento de Conocimiento en bases de datos (KDD)
3	Preparación de Datos	3.1 Recopilación. Almacenes de datos 3.2 Limpieza y transformación 3.3 Exploración y selección
4	Aprendizaje no supervisado	4.1 Introducción 4.1.1 Definición 4.1.2 Conceptos generales (objeto, tipos de datos, distancias y medidas de similitud, índices de validación) 4.2 Algoritmos de agrupamiento 4.2.1 K-medias 4.2.2 K-prototypes 4.2.3 DBSCAN
5	Aprendizaje supervisado	5.1 Definición 5.2 Matriz de confusión 5.3 Algoritmo de la regla del vecino más cercano (KNN) 5.4 Árboles de decisión 5.4.1 Definición 5.4.2 Conceptos de entropía y ganancia 5.4.3 Algoritmo ID3

7. Actividades de aprendizaje de los temas

Tema 1: Introducción a la Minería de Datos.	
Competencias	Actividades de aprendizaje
<p>Específicas: Identifica las aplicaciones prácticas y la importancia de la Minería de Datos para analizar el pasado y predecir el futuro analizando la información histórica de las organizaciones para la toma de decisiones.</p> <p>Genéricas:</p> <ul style="list-style-type: none"> • Habilidades de investigación. • Capacidad de análisis y de síntesis. • Habilidad para trabajar en forma autónoma y en equipo. • Comunicación oral y escrita. • Flexibilidad. • Gestión de tiempo. • Colaboración. 	<ul style="list-style-type: none"> • Investigar y analizar la importancia de la Minería de Datos en las organizaciones. • Analizar e investigar el origen de la Minería de Datos • Analizar y realizar resúmenes de diferentes fuentes de las aplicaciones de la Minería de datos.
Tema 2: Metodologías para la minería de datos	
Competencias	Actividades de aprendizaje
<p>Específicas: Desarrolla las habilidades para aplicar las metodologías aplicaciones reales de minería de datos.</p> <p>Genéricas:</p> <ul style="list-style-type: none"> • Habilidades de investigación. • Capacidad de análisis y de síntesis. • Habilidad para trabajar en forma autónoma y en equipo. • Comunicación oral y escrita. • Flexibilidad. • Gestión de tiempo. • Colaboración. 	<ul style="list-style-type: none"> • Investigar y analizar las etapas de cada una de las metodologías. • Analizar e investigar como se aplican las metodologías. • Desarrollar un caso práctico con las dos metodologías.

Tema 3: Preparación de datos	
Competencias	Actividades de aprendizaje
<p>Específicas: Conoce, comprende y aplica la organización y extracción del conocimiento</p> <p>Genéricas:</p> <ul style="list-style-type: none"> • Habilidades de investigación. • Capacidad de análisis y de síntesis. • Habilidad para trabajar en forma autónoma y en equipo. • Comunicación oral y escrita. • Flexibilidad. • Gestión de tiempo. • Colaboración. 	<ul style="list-style-type: none"> • Investiga y conceptualiza las técnicas de selección y limpieza de datos. • Aplica las técnicas de selección y limpieza de datos a la base de datos de su proyecto. • Presenta ante el grupo sus avances del proyecto.
Tema 4: Aprendizaje no supervisado	
Competencias	Actividades de aprendizaje
<p>Específicas: Identifica, comprende y aplica los modelos descriptivos de Aprendizaje no supervisado para la Minería de Datos.</p> <p>Genéricas:</p> <ul style="list-style-type: none"> • Habilidades de investigación. • Capacidad de análisis y de síntesis. • Habilidad para trabajar en forma autónoma y en equipo. • Comunicación oral y escrita. • Flexibilidad. • Gestión de tiempo. • Colaboración. 	<p>Analiza y comprende los algoritmos descriptivos. Realizando pruebas de escritorio.</p> <ul style="list-style-type: none"> • Practica los algoritmos descriptivos utilizando algún software libre (WEKA). • Aplica los algoritmos descriptivos a su proyecto y los presenta en clase.
Tema 5: Aprendizaje supervisado	
Competencias	Actividades de aprendizaje
<p>Específicas: Identifica, comprende y aplica las técnicas de aprendizaje supervisado a la Minería de Datos.</p> <p>Genéricas:</p> <ul style="list-style-type: none"> • Habilidades de investigación. • Capacidad de análisis y de síntesis. 	<ul style="list-style-type: none"> • Analiza y comprende los algoritmos de clasificación. Redes neuronales y Bayesianas. Realizando pruebas de escritorio. • Practica los algoritmos predictivos utilizando algún software libre (WEKA o la librería SKLEARN DE PYTHON). • Aplica los algoritmos predictivos a su proyecto y los presenta en clase.

<ul style="list-style-type: none"> • Habilidad para trabajar en forma autónoma y en equipo. • Comunicación oral y escrita. • Flexibilidad. • Gestión de tiempo. • Colaboración. 	
--	--

8. Práctica(s)

<ul style="list-style-type: none"> • Propiciar actividades de búsqueda, selección y análisis de información de distintas fuentes bibliográficas. • Desarrollar actividades de aprendizaje que propicien la aplicación de los conceptos, modelos y metodologías que se vayan adquiriendo en el transcurso de la asignatura. • Utilizar herramientas de docencia no presencial, como soporte de los contenidos teóricos y prácticos, incorporando documentación adicional, guiones de prácticas y herramientas de apoyo. • Proponer problemas reales, que permitan al estudiante la integración de contenidos en y entre distintas asignaturas, para su correspondiente desarrollo. • Proponer un caso real para que el estudiante aplique los conceptos metodológicos y teóricos aprendidos

9. Proyecto de asignatura

<p>El objetivo del proyecto: El estudiante desarrolla una aplicación que proporcione solución a una problemática de su entorno, empleando Minería de datos.</p> <ul style="list-style-type: none"> • Fundamentación: Empleando programación en Python poniendo en práctica lo aprendido desarrollando una solución a una problemática de su entorno. • Planeación: Se recomienda realizar la solicitud desde el inicio de la unidad Temática uno, generando la documentación asociada y desarrollando conforme se va avanzando. • Ejecución: La aplicación final debe de cubrir todos los aspectos vistos en la asignatura, por lo que debe ser desarrollada en un Python, apeándose a una metodología vista en el curso • Evaluación: La funcionalidad de la aplicación se sugiere que sea evaluada en su totalidad, con las rubricas que el docente considere pertinentes.

10. Evaluación por competencias

La evaluación debe ser continua y formativa por lo que se debe considerar el desempeño en cada una de las actividades de aprendizaje, haciendo especial énfasis en:

- Utilizar diferentes técnicas de evaluación con un criterio específico para cada una de ellas (teoría-práctica).
- Participación y desempeño en el aula-laboratorio.
- Desarrollar proyectos por cada unidad de aprendizaje que integren los tópicos abordados en las mismas.
- Desarrollar proyecto final que integre todas las unidades de aprendizaje.
- Producir un portafolio de evidencias (prácticas, investigaciones, documentos electrónicos, entre otros).

11. Fuentes de información

1. Hernández Orallo José, Ramírez Quintana Ma. José, Ferri Ramírez Cesar, Introducción a la Minería de Datos, Prentice Hall, 2007.
2. Howson Cindi, Business Intelligence, Mc. Graw Hill, 2008.
3. Mirkin Boris, Clustering for Data Mining A Data Recovery Approach, 2005.
4. Friedman Menahen, Kandel Abram, Introduction to Pattern Recognition, World Scientific, 1999.
5. HarjinderS. Gill, PrakashC.Rao Data Warehousing la integración de información para la mejor toma de decisiones, Prentice Hall.
6. Pérez César, Santín Daniel, DataMining Soluciones con Enterprise Miner, Alfa Omega.
7. Ramaktishnan, Gehrke Sistemas de Gestión de Bases de Datos, Mc Graw Hill.
8. Microsoft ASP.NET and AJAX: Architecting Web Applications (PRODeveloper) by Dino Eposit (1 Apr 2009)
9. Building Almacén de Datos by Milind D. Zodge (11 Jan 2011)
10. Der Aufbau von Almacén de Datos-Systemen: Konzeption - Datenmodellierung - Vorgehen (German Edition) by Jan Holthuis (1 May 1999)