



1. DATOS GENERALES DE LA ASIGNATURA

Nombre de la asignatura:	Minería de Datos y Aprendizaje Automático
Clave de la asignatura:	DAM-2403
SATCA¹:	2-4-6
Carrera:	Ingeniería en Sistemas Computacionales e Ingeniería en Tecnologías de la Información y Comunicaciones

2. PRESENTACIÓN

Caracterización de la asignatura

La asignatura aporta al perfil del egresado la capacidad de diseñar, desarrollar e implementar modelos matemáticos computacionales para resolver problemas de diversos contextos, mediante la selección e integración de diferentes herramientas tecnológicas.

Además, proporciona la capacidad de aplicar técnicas avanzadas de minería de datos y aprendizaje automático para extraer patrones y conocimientos, lo que conlleva al desarrollo de habilidades en la limpieza, preparación y transformación de datos. Esto se logra mediante la utilización de algoritmos de modelado predictivo para la toma de decisiones.

La importancia de la asignatura es permitir a los/las estudiantes analizar, interpretar y utilizar grandes volúmenes de información para identificar patrones, predecir tendencias y optimizar procesos. Esto impulsando la eficiencia y la innovación en diversas áreas tecnológicas a través de la minería de datos y el aprendizaje automático.

Adicionalmente, la asignatura se vincula con otras asignaturas como Probabilidad y Estadística, específicamente en competencias como Fundamentos de la Teoría de la Probabilidad, Estadística Descriptiva y Regresión Lineal; con Inteligencia Artificial, abordando temas como reglas de búsqueda y aplicaciones con técnicas de IA; y con el Taller de Bases de Datos, relacionándose con la competencia de Lenguaje de Manipulación de Datos.

Intención didáctica

La asignatura está dividida en cinco competencias, de modo que el estudiante, en las dos primeras, adquiera los conocimientos generales de inteligencia de negocios y big analytics, para luego abordar los temas de minería de datos, aprendizaje automático y finalmente presentar la información en un visualizador de datos.

En el primer tema, se abordan conocimientos esenciales que el/la estudiante debe poseer para comprender el origen de la inteligencia de negocios, abordando los temas Business Analytics y

¹ Sistema de Asignación y Transferencia de Créditos Académicos



Big Data, así como la arquitectura de un sistema de negocios, Inteligencia de negocios en Big Data, móvil y en la nube. Finalmente, el diseño de modelos Blockchain respaldará la toma de decisiones hacia el desarrollo de tecnologías emergentes de negocios.

En el segundo tema Big Analytics, se conocen y comprenden el uso de los temas de análisis predictivo, minería de datos, análisis de texto, análisis estadístico y adquisición de datos.

En el tercer tema, minería de datos, se introduce y profundiza en el Proceso Knowledge Discovery in Databases (KDD), Modelos y Técnicas de Data Mining, técnicas estadísticas utilizadas en Data Mining, áreas de aplicación, minería de textos y web mining; y Data Mining en el marketing

En el tema cuatro trata el aprendizaje automático, donde se comprende el uso y aplicación de modelos de clasificación estadística, cloustering, regresión, detección de anomalías, además de conocer las reglas de asociación para resolver problemas en sistemas informáticos que permitan la eficiencia y precisión en los algoritmos.

Finalmente, en el quinto tema Visualización de datos, se abordan los Conceptos generales de visualización de datos, se describen e implementan, gráficos, mapas, informes y consultas, estructura de un Dashboard y narrativa de datos (Data Storytelling).

Durante el desarrollo de las actividades programadas con fechas de entrega preestablecidas, el/la estudiante aprende a valorar el trabajo en el aula y en casa, comprendiendo que construye su hacer futuro. Lo anterior, lo impulsa a actuar de una manera profesional, reconociendo la importancia del conocimiento y de los hábitos del trabajo individual y en equipo. El estudiante también desarrolla precisión, curiosidad, imaginación, puntualidad, entusiasmo, interés, tenacidad, flexibilidad y autonomía.

3. PARTICIPANTES EN EL DISEÑO Y SEGUIMIENTO CURRICULAR DEL PROGRAMA

Lugar y fecha de elaboración o revisión	Participantes	Observaciones
Tecnológico Nacional de México, Campus San Martín Texmelucan, abril/mayo de 2024	Docentes de la academia de ciencias de la computación del Tecnológico Nacional de México, Campus San Martín Texmelucan	Generación de diseño, consolidación y/o seguimiento curricular, basado en reuniones de Academia de Ciencias de la Computación correspondientes a los meses de abril y mayo de 2024.

4. COMPETENCIA(S) A DESARROLLAR

Competencia(s) específica(s) de la asignatura
Aplicar técnicas y modelos de ciencia de datos que permitan analizar información para la toma de decisiones en programas de software.

5. COMPETENCIAS PREVIAS

<ul style="list-style-type: none">• Diseña y desarrolla programas para dar solución a diversos problemas computacionales• Aplica los conceptos de la teoría de la probabilidad y estadística para organizar, clasificar, analizar e interpretar datos para la toma de decisiones.• Implementa bases de datos para apoyar la toma de decisiones considerando las reglas de negocio

6. TEMARIO

No.	Temas	Subtemas
1.	Inteligencia de negocios	1.1 Definición de Inteligencia de negocios 1.2 Business Analytics y Big Data 1.3 Arquitectura de un Sistema de negocios 1.4 Inteligencia de negocios en Big Data, móvil y en la nube 1.5 Blockchain
2.	Big Analytics	2.1 Introducción a Big Analytics 2.2 Análisis predictivo 2.3 Minería de Datos 2.4 Análisis de texto 2.5 Análisis estadístico 2.6 Adquisición de datos
3.	Minería de datos	3.1 Introducción a la minería de datos 3.2 Proceso KDD 3.3 Modelos y Técnicas de Data Mining 3.4 Algunas técnicas estadísticas utilizadas en data mining 3.5 Áreas de aplicación 3.6 Minería de Textos y Web Mining 3.7 Data mining y marketing
4.	Aprendizaje automático	4.1 Introducción a aprendizaje automático 4.2 Aprendizaje supervisado 4.2.1 Regresión lineal y regresión logística 4.2.2 Árboles de decisión y bosques aleatorios. 4.2.3 Máquinas de vectores de soporte (SVM). 4.2.4 Métodos de ensamble 4.3 Aprendizaje No Supervisado 4.3.1 Clustering 4.3.2 Análisis de asociación 4.3.3 Reducción de dimensionalidad no supervisada 4.3.4 Estimación de densidades de probabilidad

7. ACTIVIDADES DE APRENDIZAJE DE LOS TEMAS

1. Inteligencia de negocios	
Competencias	Actividades de aprendizaje
<p>Específica(s): Identificar y comprender los conceptos básicos de la inteligencia de negocios y su impacto en la toma de decisiones en las organizaciones</p> <p>Genéricas: <i>Instrumentales:</i></p> <ul style="list-style-type: none"> • Capacidad de análisis y síntesis. • Conocimientos básicos de la carrera. • Comunicación oral y escrita. • Habilidad para buscar y analizar información proveniente de fuentes diversas. • Solución de problemas. • Toma de decisiones <p><i>Interpersonales:</i></p> <ul style="list-style-type: none"> • Capacidad crítica y autocrítica. • Trabajo en equipo. • Habilidades interpersonales. <p><i>Sistémicas:</i></p> <ul style="list-style-type: none"> • Capacidad de aplicar conocimientos en la práctica. • Habilidades de investigación. • Capacidad de aprender. • Capacidad de generar nuevas ideas (creatividad). • Habilidad para trabajar en forma autónoma. 	<ul style="list-style-type: none"> • Investigar la definición y posibles aplicaciones de Business Intelligence, Business Analytics, y Big Data para distinguir sus diferencias y aportaciones. • Haciendo uso de una herramienta de software elaborar una arquitectura de un sistema de negocios. • Indagar sobre la inteligencia de negocios en Big Data, móvil y en la nube para analizar sus diferencias y semejanzas. • Construir un cuadro comparativo con las posibles características para el diseño de un modelo Blockchain
2. Big Analytics	
Competencias	Actividades de aprendizaje
<p>Específica(s): Conocer y comprender las diferentes técnicas y aplicaciones del análisis estadístico de textos de la minería de datos.</p> <p>Genéricas:</p>	<ul style="list-style-type: none"> • Solicitar al/la estudiante el análisis predictivo para procesar informáticamente los datos y prepararlos para el análisis siguiendo las bases de datos de uso más frecuente.

<p><i>Instrumentales:</i></p> <ul style="list-style-type: none"> • Capacidad de análisis y síntesis. • Conocimientos básicos de la carrera. • Comunicación oral y escrita. • Habilidad para buscar y analizar información proveniente de fuentes diversas. • Solución de problemas. • Toma de decisiones. <p><i>Interpersonales:</i></p> <ul style="list-style-type: none"> • Capacidad crítica y autocrítica. • Trabajo en equipo. • Habilidades interpersonales. <p><i>Sistémicas:</i></p> <ul style="list-style-type: none"> • Capacidad de aplicar conocimientos en la práctica. • Habilidades de investigación. • Capacidad de aprender. • Capacidad de generar nuevas ideas (creatividad). • Habilidad para trabajar en forma autónoma. 	<ul style="list-style-type: none"> • Analizar datos mediante la aplicación de las herramientas de análisis estadístico avanzado. • Desarrollar una práctica sobre el proceso de instalaciones Oracle Analytics Desktop, Machine Learning y Advanced Analytics en Sistemas Operativos Windows o Mac OS. • Desarrollar una práctica para la gestión de conexiones a los orígenes y juego de los datos en Oracle. • Aplicar conocimientos de Base de Datos para comprender el alcance y áreas de oportunidad de la Minería de Datos.
---	---

3. Minería de datos

Competencias	Actividades de aprendizaje
<p>Específica(s):</p> <p>Aplicar los conceptos, objetivos y áreas de implementación de la minería de datos en el sector tecnológico.</p> <p>Genéricas:</p> <p><i>Instrumentales:</i></p> <ul style="list-style-type: none"> • Capacidad de análisis y síntesis. • Conocimientos básicos de la carrera. • Comunicación oral y escrita. • Habilidad para buscar y analizar información proveniente de fuentes diversas. • Solución de problemas. 	<ul style="list-style-type: none"> • Proporcionar a los estudiantes conjuntos de datos desordenados y solicitarles que identifiquen y corrijan problemas como valores atípicos, datos faltantes y errores de formato. • Solicitar un script en algún lenguaje de programación para automatizar la limpieza de datos. • Presentar a los estudiantes un problema de clasificación o regresión. • Solicitar a los estudiantes que construyan y evalúen diferentes modelos predictivos utilizando técnicas como: árboles de decisión, regresión logística o máquinas de vectores de soporte, etc.



<ul style="list-style-type: none"> • Toma de decisiones. <p><i>Interpersonales:</i></p> <ul style="list-style-type: none"> • Capacidad crítica y autocrítica. • Trabajo en equipo. • Habilidades interpersonales. <p><i>Sistémicas:</i></p> <ul style="list-style-type: none"> • Capacidad de aplicar conocimientos en la práctica. • Habilidades de investigación. • Capacidad de aprender. • Capacidad de generar nuevas ideas (creatividad). • Habilidad para trabajar en forma autónoma. 	<ul style="list-style-type: none"> • Solicitar a los estudiantes que preparen una presentación donde expliquen los resultados de un proyecto de minería de datos, incluyendo la metodología utilizada y las recomendaciones derivadas del análisis.
<p>4. Aprendizaje Automático</p>	
<p>Competencias</p>	<p>Actividades de aprendizaje</p>
<p>Específica(s):</p> <p>Diseñar y aplicar un modelo de aprendizaje automático para resolver problemas en sistemas informáticos que permitan la eficiencia y precisión en la toma de decisiones</p> <p>Genéricas:</p> <p><i>Instrumentales:</i></p> <ul style="list-style-type: none"> • Capacidad de análisis y síntesis. • Conocimientos básicos de la carrera. • Comunicación oral y escrita. • Habilidad para buscar y analizar información proveniente de fuentes diversas. • Solución de problemas. • Toma de decisiones <p><i>Interpersonales:</i></p> <ul style="list-style-type: none"> • Capacidad crítica y autocrítica. • Trabajo en equipo. • Habilidades interpersonales. <p><i>Sistémicas:</i></p>	<ul style="list-style-type: none"> • Identificar y clasificar los tipos de aprendizaje automático existentes en sistemas informáticos • Analizar diferentes casos de estudios reales e identificar el tipo de aprendizaje automático utilizado en cada uno. • Generar una comparación entre los diferentes tipos de algoritmos de aprendizaje automático • Investigar las herramientas o bibliotecas que son utilizadas para la integración de aprendizaje automático • Integrar un modelo de aprendizaje automático en un sitio web o proyecto integrador



<ul style="list-style-type: none"> • Capacidad de aplicar conocimientos en la práctica. • Habilidades de investigación. • Capacidad de aprender. • Capacidad de generar nuevas ideas (creatividad). • Habilidad para trabajar en forma autónoma. 	
5. Visualización de datos	
Competencias	Actividades de aprendizaje
<p>Específica(s): Conocer y aplicar la visualización de datos.</p> <p>Genéricas:</p> <p><i>Instrumentales:</i></p> <ul style="list-style-type: none"> • Capacidad de análisis y síntesis. • Conocimientos básicos de la carrera. • Comunicación oral y escrita. • Habilidad para buscar y analizar información proveniente de fuentes diversas. • Solución de problemas. • Toma de decisiones. <p><i>Interpersonales:</i></p> <ul style="list-style-type: none"> • Capacidad crítica y autocrítica. • Trabajo en equipo. • Habilidades interpersonales. <p><i>Sistémicas:</i></p> <ul style="list-style-type: none"> • Capacidad de aplicar conocimientos en la práctica. • Habilidades de investigación. • Capacidad de aprender. • Capacidad de generar nuevas ideas (creatividad). • Habilidad para trabajar en forma autónoma. 	<ul style="list-style-type: none"> • Investigar los conceptos generales de visualización de datos (indicadores clave de desempeño - KPI) • Crear un módulo de visualización de Big Data, basado en gráficos de dispersión y barras con detalle. • Realizar la estructura de un Dashboard. • Elaborar un módulo de gráfico en tiempo real.

8. PRÁCTICA(S)

- Determinar las características, tipo y diseño de un modelo Blockchain aplicable a su proyecto integrador.
- Generar un clasificador de imágenes que utilice aprendizaje supervisado para identificar objetos o características específicas.
- Generar agrupamiento de datos utilizando el algoritmo K- Means.
- Utilizando una herramienta de software, identificar anomalías, patrones y/o correlaciones en grandes conjuntos de datos para predecir resultados.
- Crear un chatbot inteligente para resolver preguntas y proporcionar respuestas precisas y útiles para diversas consultas.
- Generar un visualizador de datos de su proyecto integrador.
- Instalar Oracle Analytics Desktop.
- Gestionar las conexiones a los orígenes y juego de los datos en Oracle.
- Visita a empresa desarrolladora de software para conocer la forma en la que están utilizando los diferentes modelos de análisis de los datos en los sistemas que están implementando.

9. PROYECTO DE ASIGNATURA


El objetivo del proyecto que planteé el docente que imparta esta asignatura, es demostrar el desarrollo y alcance de la(s) competencia(s) de la asignatura, considerando las siguientes fases:

- **Fundamentación:** marco referencial (teórico, conceptual, contextual, legal) en el cual se fundamenta el proyecto de acuerdo con un diagnóstico realizado, mismo que permite a los y las estudiantes lograr la comprensión de la realidad o situación objeto de estudio para definir un proceso de intervención o hacer el diseño de un modelo.
- **Planeación:** con base en el diagnóstico en esta fase se realiza el diseño del proyecto por parte de los y las estudiantes con asesoría del/de la docente; implica planificar un proceso: de intervención empresarial, social o comunitario, el diseño de un modelo, entre otros, según el tipo de proyecto, las actividades a realizar los recursos requeridos y el cronograma de trabajo.
- **Ejecución:** consiste en el desarrollo de la planeación del proyecto realizada por parte de los y las estudiantes con asesoría del/de la docente, es decir en la intervención (social, empresarial), o construcción del modelo propuesto según el tipo de proyecto, es la fase de mayor duración que implica el desempeño de las competencias genéricas y específicas a desarrollar.

- Evaluación: es la fase final que aplica un juicio de valor en el contexto laboral-profesión, social e investigativo, ésta se debe realizar a través del reconocimiento de logros y aspectos a mejorar se estará promoviendo el concepto de “evaluación para la mejora continua”, la metacognición, el desarrollo del pensamiento crítico y reflexivo en los y las estudiantes.

10. EVALUACIÓN POR COMPETENCIAS

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

- Presentación de reportes de búsqueda de información en fuentes bibliográficas o digitales de reconocido valor.
- Participación en actividades para demostrar la comprensión de los conocimientos adquiridos a través de las investigaciones anteriores, tales como participación en foros, etc.
- Elaboración de proyectos de aplicación donde se incluyan e integren los algoritmos y modelos de minería de datos y aprendizaje automático.
- Reportes escritos de las prácticas realizadas, así como  de las conclusiones obtenidas de dichas prácticas.
- Solución de casos de estudio o problemas.
- Desarrollo de proyectos integradores.

Instrumentos y herramientas:

- Rúbricas.
- Lista de cotejo.
- Guías de observación.

11. FUENTES DE INFORMACIÓN

- DATAtab Team (2024). DATAtab: Online Statistics Calculator. DATAtab e.U. Graz, Austria. URL <https://datatab.es>
- Elrom E. The Blockchain Developer: A Practical Guide for Designing, Implementing, Publishing, Testing, and Securing Distributed Blockchain-Based Projects. Editorial Apres.
- Gironés Roig, J., Casas Roma, J., Minguillón Alfonso J., Caihuelas Quiles R. (2017). Minería de datos, modelos y algoritmos. Editorial UOC



- Hernández Orallo, J., Ramírez Quintana, M.J., Ferri Ramírez, C. (2004). Introducción a la Minería de Datos. Editorial Pearson
- Huyen C. Diseño de sistemas de machine learning. Marcombo.
- Joyanes Aguilar, L. (2019) Inteligencia de Negocios y Analítica de Datos. Una visión global de business intelligence & Analytics. Editorial Alfaomega
- Ortiz, D. (24 de noviembre de 2023). ¿Qué es un dashboard y para qué se usa? <https://www.cyberclick.es/numerical-blog/que-es-un-dashboard>
- Pertuz Pineda, C. (2022) Aprendizaje automático y profundo en Python. Ediciones de la U
- Salvador, F. (2014). Big Data: ¿La ruta o el destino? Editorial Oracle.
- Soria Olivas, E., Sánchez Montañes, M.A., Cruz Gamero, R., Castilla Caballero, B., Cano Michelena, P. (2023) Sistemas de Aprendizaje Automático. Editorial Ra-Ma
- Warburg B. Basics of blockchain: a guide for building literacy in the economics, technology, and business of blockchain