

UNIVERSIDAD AUTÓNOMA DE
CHIHUAHUA



UNIDAD ACADÉMICA:
FACULTAD DE INGENIERÍA
PROGRAMA ANALÍTICO DE LA
UNIDAD DE APRENDIZAJE:
CIENCIA DE DATOS

DES: INGENIERÍA	INGENIERÍA
Programa académico	Ingeniería en Computación Ingeniería en Ciencias de datos y Matemáticas Aplicadas.
Tipo de materia (Obli/Opta):	Optativa
Clave de la materia:	MC512
Semestre:	Octavo
Área en plan de estudios:	Específica
Total de horas por semana:	6
<i>Teoría: Presencial o Virtual</i>	0
<i>Laboratorio o Taller:</i>	4
<i>Prácticas:</i>	0
<i>Trabajo extra-clase:</i>	2
Créditos Totales:	6
Total de horas semestre (x sem):	96
Fecha de actualización:	Octubre 2024
<i>Prerrequisito (s):</i>	CO605 Ingeniería de datos

DESCRIPCIÓN:

La materia sensibiliza al estudiante sobre la importancia de la información en la vida actual. Debido a la enorme generación de datos diariamente, es importante aprender técnicas modernas para su procesamiento y generación de conocimiento y soporte a la toma de decisiones en áreas que incluyen, pero no se limitan a: servicio al cliente, medicina, política, educación. La asignatura le aporta al alumno entendimiento y capacidad de aplicación en áreas tales como: procesamiento y limpieza de datos, visualización, aprendizaje supervisado y no supervisado, aprendizaje máquina tradicional y profundo, análisis de series de tiempo entre otras.

COMPETENCIAS PARA DESARROLLAR:

B4. Transformación Digital

Transforma la cultura digital en la sociedad, en las organizaciones e instituciones educativas para aprovechar al máximo el potencial de las tecnologías y herramientas digitales; propiciar su uso responsable y ético que estimule la creatividad, innovación, la comunicación efectiva y el trabajo colaborativo e interdisciplinar en la solución de problemas de la sociedad digital; promoviendo la privacidad y la seguridad, así como el respeto a los derechos de autor y la propiedad intelectual.

E3. INTELIGENCIA ARTIFICIAL Y ANÁLISIS DE DATOS

Se desarrolla en el campo de la inteligencia artificial y el análisis de datos, aplicando algoritmos de aprendizaje automático y herramientas de Big Data para generar perspectivas significativas y soluciones innovadoras. Engloba la habilidad para manejar y analizar grandes volúmenes de datos, así como la creatividad para aplicar la IA en la solución de problemas y desafíos contemporáneos en diversos sectores.

DOMINIOS (Se toman de las competencias)	OBJETOS DE ESTUDIO (Contenidos necesarios para desarrollar cada uno de los dominios)	RESULTADOS DE APRENDIZAJE (Se plantean de los dominios y contenidos)	METODOLOGÍA (Estrategias, secuencias, recursos didácticos)	EVIDENCIAS (Productos tangibles que permiten valorar los resultados de aprendizaje)
<p>B4.3 Aplica de forma ética diferentes herramientas digitales que favorezcan el trabajo colaborativo e interprofesional, considerando las principales innovaciones científicas y tecnológicas, relacionadas con la profesión.</p> <p>E3. INTELIGENCIA ARTIFICIAL Y ANÁLISIS DE DATOS</p> <p>Dominios</p> <p>1. Aplicar técnicas de aprendizaje automático y procesamiento de datos para desarrollar soluciones en el campo de la inteligencia artificial.</p> <p>2. Analizar grandes conjuntos de datos utilizando herramientas, métodos de Big Data para extraer información valiosa y tomar decisiones basadas en datos.</p>	<p>1. INTRODUCCIÓN A LA CIENCIA DE DATOS</p> <p>1.1. Definición de Ciencia de Datos e importancia en el contexto actual.</p> <p>1.2. Descripción de los procesos de extracción, transformación y carga (ETL).</p> <p>1.3. Web scraping.</p> <p>1.4. Repaso de Python</p> <p>1.4.1. Estructuras básicas: listas, tuplas, diccionarios, etc.</p> <p>1.4.2. Iteradores vs ciclos.</p> <p>1.4.3. Librerías básicas: Pandas, NumPy, Scikit-Learn, etc.</p> <p>1.4.4. POO en Python.</p>	<p>Describe la importancia de la ingeniería de datos en el contexto actual utilizando términos técnicos y ejemplos prácticos.</p>	<p>Clase interactiva Maestro - Alumno.</p> <p>Recursos tecnológicos institucionales.</p>	<p>Tareas y/o mini proyectos.</p>

	<p>2. PREPROCESAMIENTO DE LOS DATOS</p> <p>2.1. Importancia del preprocesamiento de los datos</p> <p>2.2. Limpieza: identificación y manejo de valores faltantes, <i>outliers</i> y datos erróneos.</p> <p>2.3. Normalización y estandarización de datos.</p> <p>2.4. Codificación de variables categóricas.</p> <p>2.5. Estratificación de datos.</p>	<p>Analiza la importancia del preprocesamiento de los datos y su impacto en la calidad de los análisis posteriores.</p> <p>Diseña estrategias de normalización, estandarización y codificación</p>	<p>Clase interactiva Maestro - Alumno.</p> <p>Recursos tecnológicos institucionales.</p>	<p>Tareas y/o mini proyectos.</p>
	<p>2.5.1. Integración de datos de múltiples fuentes.</p> <p>2.6. Visualización de datos.</p> <p>2.6.1. La librería Matplotlib,</p> <p>2.6.2. Reduciendo dimensionalidad con PCA (Principal Components Analysis) y t-SNE (t-distributed Stochastic Neighbor Embedding) para visualización de datos n-dimensionales.</p>	<p>de variables categóricas para mejorar la calidad y la interpretación de los datos.</p> <p>Domina estrategias de visualización de datos.</p>		
	<p>3. INGENIERÍA DE DATOS</p> <p>3.1. Selección de características: métodos univariados, multivariados y basados en modelos.</p> <p>3.2. Creación de nuevas características: transformaciones, combinaciones y derivaciones.</p> <p>3.3. Fundamentos de probabilidad y estadística.</p> <p>3.3.1. Estadística descriptiva.</p> <p>3.3.2. Distribuciones de probabilidad (binomial, poisson, etc.).</p> <p>3.3.3. Teorema de Bayes.</p> <p>3.3.4. Probabilidad Condicional.</p>	<p>Implementa técnicas de selección de características para mejorar la eficiencia y la interpretación de los modelos de machine learning.</p>	<p>Clase interactiva Maestro - Alumno.</p> <p>Recursos tecnológicos institucionales.</p>	<p>Tareas y/o mini proyectos.</p>

	<p>4. PROCESAMIENTO AVANZADO DE DATOS</p> <p>4.1. Tecnologías y herramientas para procesamiento de grandes volúmenes de datos.</p> <p>4.2. Optimización de consultas y procesos de ETL.</p> <p>4.3. Procesamiento en tiempo real: streaming de datos.</p> <p>4.4. Aprendizaje supervisado vs no supervisado.</p> <p>4.5. Validación cruzada de datos.</p>	<p>Optimiza consultas y procesos de ETL para mejorar el rendimiento y la eficiencia en entornos de big data.</p> <p>Conoce sobre aprendizaje máquina supervisado y no supervisado,</p>	<p>Clase interactiva Maestro - Alumno.</p> <p>Recursos tecnológicos institucionales.</p>	<p>Tareas y/o mini proyectos.</p>
	<p>4.6. Series de tiempo.</p>	<p>además de técnicas de validación de modelos tales como validación cruzada.</p>		
	<p>5. ALGORITMOS TRADICIONALES</p> <p>5.1. Logistic Regression.</p> <p>5.2. Naive Bayes.</p> <p>5.3. K Vecinos más cercanos.</p> <p>5.4. Árboles de decisión.</p> <p>5.5. Algoritmos por ensamble.</p> <p>5.6. Máquinas de vectores de soporte.</p>	<p>Entiende algoritmos de aprendizaje máquina tradicionales y su importancia actual.</p>	<p>Clase interactiva Maestro - Alumno.</p> <p>Recursos tecnológicos institucionales.</p>	<p>Tareas y/o mini proyectos.</p>
	<p>6. ALGORITMOS MODERNOS</p> <p>6.1. Redes neuronales recurrentes.</p> <p>6.2. Redes neuronales convolucionales.</p> <p>6.3. Redes neuronales Transformer.</p>	<p>Conoce los algoritmos actuales de Deep learning para abordar problemas de ciencia de datos.</p>	<p>Clase interactiva Maestro - Alumno.</p> <p>Recursos tecnológicos institucionales.</p>	<p>Tareas y/o mini proyectos.</p>

FUENTES DE INFORMACIÓN (Bibliografía, direcciones electrónicas)	EVALUACIÓN DE LOS APRENDIZAJES (Criterios, ponderación e instrumentos)
<ul style="list-style-type: none"> • Grus, J. (2019). <i>Data Science from Scratch: First Principles with Python</i> (2nd ed.). O'Reilly Media, Inc. ISBN 9781492041139. • Bruce, P., Bruce, A., & Gedeck, P. (2020). <i>Practical statistics for data scientists</i> (2nd ed.). O'Reilly Media, Inc. ISBN 978-1492072942. • Nield, T. (2022). <i>Essential math for data science</i>. O'Reilly Media, Inc. ISBN 978-1098102937. • Géron, A. (2022). <i>Hands-on machine learning with Scikit-Learn, Keras, and TensorFlow</i> (3rd ed.). O'Reilly Media, Inc. ISBN 978-1098125974. • Goodfellow, I., Bengio, Y., & Courville, A. (2016). <i>Deep learning</i>. The MIT Press. ISBN 978-0262035613. 	<p>Parcial 1 (30%)</p> <ul style="list-style-type: none"> • Tareas y/o Mini proyecto. (100%) <p>Parcial 2 (30%)</p> <ul style="list-style-type: none"> • Tareas y/o Mini proyecto. (100%) <p>Parcial 3 (40%)</p> <ul style="list-style-type: none"> • Tareas y/o Mini proyecto. (100%) <p>La calificación mínima es 7.0.</p> <p>Se usará rúbrica para la entrega de actividades o tareas a realizar.</p>

CRONOGRAMA

Objetos de estudio	Semanas																
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	
1. INTRODUCCIÓN A LA CIENCIA DE DATOS																	
2. PREPROCESAMIENTO DE LOS DATOS																	
3. INGENIERÍA DE DATOS																	
4. PROCESAMIENTO AVANZADO DE DATOS																	
5. ALGORITMOS TRADICIONALES																	
6. ALGORITMOS MODERNOS																	