

<p>UNIVERSIDAD AUTÓNOMA DE CHIHUAHUA</p>  <p>UNIDAD ACADÉMICA:</p> <p>FACULTAD DE INGENIERÍA</p> <p>PROGRAMA ANALÍTICO DE LA UNIDAD DE APRENDIZAJE:</p> <p><u>INGENIERÍA DE DATOS</u></p>	DES:	INGENIERÍA
	Programa académico	Ingeniería en Computación
	Tipo de materia (Obli/Opta):	Obligatoria
	Clave de la materia:	CO605
	Semestre:	Sexto
	Área en plan de estudios:	Específica
	Total de horas por semana:	5
	<i>Teoría: Presencial o Virtual</i>	0
	<i>Laboratorio o Taller:</i>	4
	<i>Prácticas:</i>	0
	<i>Trabajo extra-clase:</i>	1
	Créditos Totales:	5
	Total de horas semestre (x sem):	80
	Fecha de actualización:	Octubre 2024
<i>Prerrequisito (s):</i>	CO504 Base de datos avanzadas	

DESCRIPCIÓN:

El curso permite obtener conocimientos y habilidades para el uso técnicas, modelos y herramientas para implementar soluciones de procesamiento y análisis de datos, mediante la creación una infraestructura de datos que sirva como base para procesos de Data Science, Machine Learning y Big Data Analytics.

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.

P1. CIENCIAS E INGENIERÍA

Aplica los conocimientos y metodologías para el planteamiento y resolución de problemas complejos de las ciencias naturales y de la ingeniería, para la toma de decisiones en un contexto de responsabilidad social y del medio ambiente.

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	OBJETOS DE ESTUDIO	RESULTADOS DE APRENDIZAJE	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>P1. CIENCIAS E INGENIERÍA</p> <p>Utiliza conceptos, métodos y leyes fundamentales de las ciencias básicas para dar soluciones a problemas complejos de ciencias e ingeniería a través del uso de tecnologías de información, analizando los resultados para emitir conclusiones en contextos reales</p> <p>E3 INTELIGENCIA ARTIFICIAL Y ANÁLISIS DE DATOS</p> <p>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 Ingeniería de Datos</p> <p>1.1. Definición y conceptos básicos</p> <p>1.2. Data Engineering vs. Data Science vs. Data Analytics</p> <p>1.3. Rol y responsabilidades del ingeniero de datos</p> <p>1.4. Ciclo de Vida de los Datos: Ingesta, almacenamiento, procesamiento, análisis y visualización</p> <p>1.5. ETL vs ELT,</p> <p>1.6. Procesamiento Batch vs Streaming</p> <p>1.7. On-premise vs Cloud</p> <p>1.8. Aplicaciones y plataformas</p> <p>1.9. Casos de uso</p>	<p>Identifica los componentes básicos de un ecosistema data engineering</p>	<ul style="list-style-type: none"> ○ Clase interactiva Maestro – Alumno. ○ Uso de recursos tecnológicos institucionales 	<p>Trabajos de investigación a entregar de forma presencial o a través del recurso tecnológico institucional.</p>

	<p>2. Arquitectura de Datos</p> <p>2.1. Introducción a la Arquitectura de Datos</p> <p>2.2. Tipos de Arquitectura de Datos</p> <p>2.3. Tecnología de almacenamiento de Datos</p> <p>2.3.1. Bases de Datos Relacionales</p> <p>2.3.2. Bases de Datos NoSQL</p> <p>2.3.3. Data Warehouses</p> <p>2.3.4. Data Lakes</p> <p>2.3.5. Distributed File Systems</p> <p>2.4. Gestión Raw Data</p>	<p>Identifica las principales arquitecturas de datos</p>	<ul style="list-style-type: none"> ○ Clase interactiva Maestro Alumno. ○ Uso de recursos tecnológicos institucionales ○ Ejercicios de los temas, desarrollados en clase de forma individual o grupal. ○ 	<p>Resolución de ejercicios a través de expresión escrita en clase y evidencia en el recurso tecnológico institucional.</p> <p>Trabajos de investigación a entregar de forma presencial o a través del recurso tecnológico institucional.</p> <p>Exámenes escritos y/o en el recurso tecnológico institucional.</p> <p>Proyecto</p>
	<p>3. Procesamiento de Datos</p> <p>3.1. Conexión e ingesta de datos desde relational databases, NosQL databases, web APIs, log files, etc.</p> <p>3.2. Herramientas ETL</p> <p>3.3. Herramientas ELT</p> <p>3.4. Transformación de datos</p> <p>3.4.1. Validación, limpieza, normalización</p>	<p>Aplica técnicas y herramientas para la construcción de ETL y ELT para el procesamiento de datos</p>	<ul style="list-style-type: none"> ○ Clase interactiva Maestro Alumno. ○ Uso de recursos tecnológicos institucionales ○ Ejercicios de los temas, desarrollados en clase de forma individual o grupal. ○ Desarrollo de Proyecto 	<p>Resolución de ejercicios a través de expresión escrita en clase y evidencia en el recurso tecnológico institucional.</p> <p>Trabajos de investigación a entregar de forma presencial o a través del recurso tecnológico institucional.</p> <p>Exámenes escritos y/o en el</p>

	<p>n y enriquecimiento de datos</p> <p>3.5. Data Wrangling: SQL y Python</p> <p>3.6. Procesamiento Batch</p> <p>3.7. Procesamiento Streaming</p>			<p>recurso tecnológico institucional.</p> <p>Desarrollo de Proyecto</p>
	<p>4. Data Pipelines y Orquestación</p> <p>4.1. Diseño y modelado de Data Pipeline</p> <p>4.2. Orquestación de Workflow</p> <p>4.3. Automatización de Data Pipelines</p> <p>4.4. Monitoreo de Data Pipeline</p> <p>4.5. Pruebas y resolución de problemas de Data Pipeline</p>	<p>Aplica herramientas para la construcción de Data Pipeline y su Orquestación</p>	<ul style="list-style-type: none"> ○ Clase interactiva Maestro Alumno. ○ Uso de recursos tecnológicos institucionales ○ Ejercicios de los temas, desarrollados en clase de forma individual o grupal. ○ Desarrollo de proyecto 	<p>Resolución de ejercicios a través de expresión escrita en clase y evidencia en el recurso tecnológico institucional.</p> <p>Trabajos de investigación a entregar de forma presencial o a través del recurso tecnológico institucional.</p> <p>Exámenes escritos y/o en el recurso tecnológico institucional.</p> <p>Desarrollo de Proyecto</p>
	<p>5. Deployment y Monitoreo</p> <p>5.1. Deployment de sistemas de ingeniería de datos</p> <p>5.2. Implementación y configuración</p> <p>5.3. Escalabilidad y tolerancia a fallos</p> <p>5.4. Monitoreo y mantenimiento</p>	<p>Aplica técnicas de despliegue e implementación de soluciones de data engineering</p>	<ul style="list-style-type: none"> ○ Clase interactiva Maestro Alumno. ○ Uso de recursos tecnológicos institucionales ○ Ejercicios de los temas, desarrollados en clase de forma individual o grupal. 	<p>Resolución de ejercicios a través de expresión escrita en clase y evidencia en el recurso tecnológico institucional.</p> <p>Trabajos de investigación a entregar de forma presencial o a través del recurso</p>

	<p>5.5. Resolución de problemas y optimización</p> <p>5.6. CI/CD para proyectos de Datos</p>			<p>tecnológico institucional.</p> <p>Exámenes escritos y/o en el recurso tecnológico institucional.</p>
	<p>6. Calidad y Gobernanza de Datos</p> <p>6.1. Calidad de datos</p> <p>6.2. Seguridad de Datos</p> <p>6.3. Eficiencia y Escalabilidad</p> <p>6.4. Gobernanza de Datos</p>	<p>Aplica técnicas y herramientas para asegurar la calidad y gobernanza de los datos</p>	<ul style="list-style-type: none"> ○ Clase interactiva Maestro Alumno. ○ Uso de recursos tecnológicos institucionales ○ Proyecto 	<p>Trabajos de investigación a entregar de forma presencial o a través del recurso tecnológico institucional.</p> <p>Exámenes escritos y/o en el recurso tecnológico institucional.</p> <p>Desarrollo de Proyecto</p>

FUENTES DE INFORMACIÓN (Bibliografía, direcciones electrónicas)	EVALUACIÓN DE LOS APRENDIZAJES (Criterios, ponderación e instrumentos)
<ul style="list-style-type: none"> • J. Reis, M. Housley. Marcombo S.L. (2023). Fundamentos de ingeniería de Datos. ISBN: 9788426736888 • G. Eagar. Packt Publishing (2023). Data Engineering with AWS. Second Edition. ISBN 9781804614426 • B. Lipp. Packt Publishing (2023). Modern Data Architectures with Python. ISBN: 9781801070492 • M. Malhotra. Orange Education Pvt Ltd (2023). Ultimate Data Engineering with Databricks. ISBN: 8196994788 	<p>Tres evaluaciones parciales durante el semestre.</p> <p>Ponderaciones:</p> <p>Primer parcial 30%</p> <ul style="list-style-type: none"> • Actividades 40% • Examen 40% • Actividad Integradora 20% <p>Segundo parcial: 30%</p> <ul style="list-style-type: none"> • Actividades 40% • Examen 40% • Actividad Integradora 20% <p>Tercer Parcial: 40%</p> <ul style="list-style-type: none"> • Actividades 30% • Examen 20% • Proyecto Final 50% <p>Se usará rúbrica para la entrega de actividades o tareas a realizar.</p> <p>La calificación mínima aprobatoria es 7.0.</p>

CRONOGRAMA

Objetos de estudio	Semanas															
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
UNIDAD I Introducción a la Ingeniería de Datos																
UNIDAD II. Arquitectura de Datos																
UNIDAD III. Procesamiento de Datos																
UNIDAD IV Data Pipelines y Orquestación																
UNIDAD V. Deployment y Monitoreo																
UNIDAD VI. Calidad y Gobernanza de Datos																