



UNIVERSIDAD AUTÓNOMA DE CHIAPAS
Escuela de Tecnologías Digitales Aplicadas
Licenciatura en Inteligencia Artificial y Ciencia de Datos



Unidad de Competencia							
Ingeniería de Características							
Cuatrimestre	Fecha de elaboración			Modalidad	Tipo	Área de formación	
5	DD	MM	AÑO	• Escolarizada Opción: ○ Presencial	• Curso / Laboratorio	Ciencia de datos, aprendizaje automático y modelos avanzados	
	27	11	2025				
Perfil docente							
• Formación: ○ Estudios de licenciatura (deseable Ingeniería, Ciencia de Datos, Computación, Matemáticas, Estadística o afines) y posgrado (deseable en Ciencia de Datos, Machine Learning, Estadística o Ingeniería de Datos). • Experiencia profesional: ○ Experiencia (mínimo 3 años) en proyectos con preprocesamiento de datos y ML, dominio de Python (Pandas, NumPy), scikit-learn, técnicas de feature engineering y pipelines, experiencia con datasets reales y procesos ETL • Dominio de inglés: ○ B2 (lectura de literatura técnica, documentación y papers). • Competencias digitales: ○ Python científico, scikit-learn, herramientas de visualización, Git/GitHub, notebooks.							
Horas con Mediación Docente Cuatrimestrales		Horas con Mediación Docente semana/mes		Horas de Actividad Independiente	Horas de Práctica de Campo	Créditos	Academia
HT	HP	HT	HP	64	0	8	
32	32	2	2				
Horas de docencia total de la unidad de competencia: 64							
Propósito general:				Presentación:			
El estudiantado diseña, implementa y evalúa técnicas de ingeniería de				La ingeniería de características es uno de los procesos más críticos en			



UNIVERSIDAD AUTÓNOMA DE CHIAPAS
Escuela de Tecnologías Digitales Aplicadas
Licenciatura en Inteligencia Artificial y Ciencia de Datos



características para transformar datos brutos en representaciones útiles, mejorando el rendimiento y la interpretabilidad de modelos predictivos mediante preprocesamiento, selección y creación de características	ciencia de datos y aprendizaje automático. Consiste en seleccionar, transformar, escalar y crear nuevas variables que capturen relaciones relevantes en los datos. Esta asignatura dota al estudiantado de habilidades avanzadas para preparar datos de manera eficiente, construir pipelines reproducibles y comprender cómo la calidad de las características afecta la precisión de modelos. Se articula con Machine Learning I, Fundamentos de Data Science, Inferencia Estadística, Optimización y Minería de Datos
Competencias	
Genéricas	Disciplinares
• Pensamiento crítico y analítico. • Resolución de problemas complejos. • Comunicación técnica fundamentada. • Trabajo colaborativo	• Aplica técnicas sistemáticas de preprocesamiento. • Selecciona y transforma características relevantes. • Evalúa impacto de transformaciones en modelos predictivos.
Profesionales	
• Implementa pipelines reproducibles. • Documenta decisiones de transformación. • Asegura calidad y consistencia de los datos	
Producto final	
Pipeline profesional de ingeniería de características, que incluya: Limpieza, imputación y escalamiento, Codificación categórica, Selección y creación de características, Métricas comparativas antes/después, Integración con modelos supervisados y Código reproducible y documentación	
Enseñanza con perspectiva de género, interculturalidad y Derechos Humanos	
Con la intención de propiciar una enseñanza con perspectiva de género, interculturalidad y derechos humanos, la o el docente que imparta esta unidad de competencia deberá: • Eliminar y rechazar actitudes y comportamientos sexistas, clasistas, racistas y todo tipo de estereotipos que supongan discriminación y violencia en los ámbitos de enseñanza-aprendizaje (virtual y presencial), a partir del ejemplo en su práctica docente. • Promover que el desarrollo del programa educativo y material a utilizar se base en el respeto, protección, promoción y garantía de los	



UNIVERSIDAD AUTÓNOMA DE CHIAPAS
Escuela de Tecnologías Digitales Aplicadas
Licenciatura en Inteligencia Artificial y Ciencia de Datos



derechos humanos, de la dignidad e integridad de todas las personas.

- Promover la eliminación de estereotipos o cualquier otra práctica de discriminación o denigración de la persona, que pueda presentarse mediante el uso del lenguaje, el empleo de ejemplos, la diferenciación entre grupos, origen étnico o cualquier otra circunstancia.
- Respetar la identidad de género de las personas en todo momento y nombrarlas tal y como ellas lo deseen.
- Respetar y promover la igualdad de trato y de oportunidades entre las y los estudiantes, favoreciendo acciones afirmativas en beneficio de los grupos en situación de vulnerabilidad, incluso en las afectaciones personales provocadas por la pandemia COVID-19
- Promover y usar lenguaje incluyente y no sexista tanto en el salón de clases como en los trabajos, tesis y otros documentos que elabore la comunidad estudiantil.

Estos elementos son motivados por la información asequible en los programas de estudio de la Licenciatura en Derecho, Modalidad Educativa Presencial (sistema escolarizado), Tomo II, de la Universidad Nacional Autónoma de México (UNAM 2020).



UNIVERSIDAD AUTÓNOMA DE CHIAPAS
Escuela de Tecnologías Digitales Aplicadas
Licenciatura en Inteligencia Artificial y Ciencia de Datos



Nivel Unidad/Subcompetencias	
Nombre de la Subcompetencia 1	Elementos de la subcompetencia 1
Preprocesamiento y Preparación de Datos	Conocimientos: • Limpieza y depuración • Detección y tratamiento de valores faltantes • Eliminación de duplicados • Transformaciones básicas
Número de semanas	
4 semanas	
Propósito de la subcompetencia	Habilidades: • Imputación de valores faltantes • Normalización de tipologías • Manejo de inconsistencias Valores y actitudes: • Honestidad en tratamiento de datos • Rigor • Responsabilidad técnica
El estudiantado aplica técnicas de limpieza y preprocesamiento para garantizar la calidad de los datos antes del modelado.	
Estrategias Didácticas	Recursos didácticos
<i>Las estrategias y recursos didácticos consideran la inclusión y perspectiva de género, así como la diversidad cultural.</i>	
• Laboratorio con datasets reales • Ejercicios guiados • Mini-proyecto	• Python, Pandas • Notebooks • Data profiling

Nombre de la Subcompetencia 2	Elementos de la subcompetencia 2
Transformación y Codificación de Variables	Conocimientos: • Escalamiento (standard, min-max, robust)
Número de semanas	



UNIVERSIDAD AUTÓNOMA DE CHIAPAS
Escuela de Tecnologías Digitales Aplicadas
Licenciatura en Inteligencia Artificial y Ciencia de Datos



4 semanas	• Codificación categórica (one-hot, ordinal, target encoding básico) • Transformaciones logarítmicas, polinomiales • Tratamiento de outliers Habilidades: • Usa escaladores y codificadores • Identifica variables que requieren transformación • Evalúa efectos de las transformaciones Valores y actitudes: • Ética en decisiones de transformación • Precisión • Respeto a integridad de los datos
Propósito de la subcompetencia	
El estudiantado transforma variables mediante técnicas que optimizan el desempeño y estabilidad de modelos de machine learning.	
Estrategias Didácticas	Recursos didácticos
<i>Las estrategias y recursos didácticos consideran la inclusión y perspectiva de género, así como la diversidad cultural.</i>	
• Visualización de efectos • Laboratorio con sklearn.preprocessing • Comparativas de modelos	• Scikit-learn • Python científico • Visualización

Nombre de la Subcompetencia 3	Elementos de la subcompetencia 3
Selección de Características	Conocimientos: • Métodos filtro: correlación, chi-cuadrada • Métodos wrapper: RFE • Métodos embedded: regularización (Lasso, Ridge) • Importancia de características Habilidades: • Aplica filtros y regularización
Número de semanas	
4 semanas	
Propósito de la subcompetencia	
El estudiantado selecciona características relevantes para mejorar la eficiencia, interpretabilidad y desempeño de modelos.	



UNIVERSIDAD AUTÓNOMA DE CHIAPAS
Escuela de Tecnologías Digitales Aplicadas
Licenciatura en Inteligencia Artificial y Ciencia de Datos



	• Compara modelos con y sin selección • Identifica variables irrelevantes o redundantes
	Valores y actitudes: • Responsabilidad • Profesionalismo técnico • Ética en selección (no manipulación engañosa)
Estrategias Didácticas	Recursos didácticos
<i>Las estrategias y recursos didácticos consideran la inclusión y perspectiva de género, así como la diversidad cultural.</i>	
• Taller de selección • Evaluación antes/después • Pruebas comparativas	• sklearn.feature_selection • Notebooks • Datasets reales

Nombre de la Subcompetencia 4	Elementos de la subcompetencia 4
Creación de Características y Pipelines Reproducibles	Conocimientos: • Feature engineering manual (fechas, texto, interacción de variables) • Generación de métricas derivadas • Pipelines y ColumnTransformer • Buenas prácticas de reproducibilidad Habilidades: • Crea nuevas características útiles • Automatiza procesos • Evalúa contribuciones de características nuevas Valores y actitudes: • Creatividad
Número de semanas	
4 semanas	
Propósito de la subcompetencia	
El estudiantado construye características nuevas y diseña pipelines automatizados para integrarlas de forma consistente en los modelos.	



UNIVERSIDAD AUTÓNOMA DE CHIAPAS
Escuela de Tecnologías Digitales Aplicadas
Licenciatura en Inteligencia Artificial y Ciencia de Datos



	• Rigor en validación • Ética en la manipulación de datos
Estrategias Didácticas	Recursos didácticos
<i>Las estrategias y recursos didácticos consideran la inclusión y perspectiva de género, así como la diversidad cultural.</i>	
• Miniproyectos • Laboratorio intensivo • Evaluación comparativa con modelos	• Sklearn.pipeline • Python científico • Kit de minería de datos

Evaluación	
La evaluación de los aprendizajes se desarrollará de forma continua durante el proceso de enseñanza-aprendizaje a través de los siguientes momentos: Evaluación diagnóstica: Diagnóstico de fundamentos de Python, Pandas y ML. Evaluación formativa: Prácticas semanales, notebooks de transformación, actividades de selección y creación de características y retroalimentación continua. Evaluación sumativa: Pipeline de ingeniería de características, examen teórico–práctico y portafolio digital.	
Estrategia de evaluación de los aprendizajes:	Instrumentos de Evaluación:
Estrategias que el estudiante desarrolla para la elaboración de las evidencias de aprendizaje. • Transformaciones aplicadas • Comparativas de desempeño • Pipelines reproducibles • Documentación técnica • Reflexión analítica	• Rúbricas • Listas de cotejo • Revisión de notebooks • Reportes comparativos
Referencias	
<i>Las referencias consideran la inclusión y perspectiva de género, así como la diversidad cultural.</i>	



UNIVERSIDAD AUTÓNOMA DE CHIAPAS
Escuela de Tecnologías Digitales Aplicadas
Licenciatura en Inteligencia Artificial y Ciencia de Datos



Básicas:		Complementarias:	
• Géron, A. (2022). Hands-On Machine Learning. O'Reilly. • Kuhn, M. & Johnson, K. (2021). Feature Engineering and Selection. CRC. • scikit-learn Documentation.		• VanderPlas, J. (2020). Python Data Science Handbook. • Kaggle Learn – Feature Engineering.	
Videos:			
Bases de Datos:			
• CONRICYT • UCI Repository • Kaggle Datasets			