

IDENTIFICACIÓN DE LA UNIDAD DE APRENDIZAJE

Unidad académica: Instituto de Investigación en Ciencias Básicas y Aplicadas							
Plan de estudios: Licenciatura en Inteligencia Artificial							
Unidad de aprendizaje: APRENDIZAJE AUTOMÁTICO				Ciclo de formación: Profesional Eje general de formación: Teórico-Técnica Área de conocimiento: Bases de la Inteligencia Artificial y la Ciencia de Datos Semestre: 4º			
Elaborada por: Dr. Jorge Hermosillo Valadez				Fecha de elaboración: Abril, 2021			
Clave:	Horas teóricas :	Horas prácticas :	Horas totales :	Créditos :	Tipo de unidad de aprendizaje :	Carácter de la unidad de aprendizaje :	Modalidad:
AA25CP03020 8	03	02	05	08	Obligatoria	Teórico - Práctica	Escolarizada
Plan (es) de estudio en los que se imparte: A partir de todos los programas impartidos por el Instituto de Investigación en Ciencias Básicas y Aplicadas.							

ESTRUCTURA DE LA UNIDAD DE APRENDIZAJE

<p>Presentación:</p> <p>El Aprendizaje Automático (<i>Machine Learning</i>) es una rama de la Inteligencia Artificial (IA) cuyo propósito es proporcionar a las computadoras la habilidad de aprender sin ser explícitamente programadas. Debido al crecimiento exponencial que ha tenido la capacidad de cómputo de las máquinas, el aprendizaje automático se ha convertido en un pilar de la ciencia de datos y es uno de los campos de aplicación de la IA donde la teoría de la información y las matemáticas, puras y aplicadas, han tenido mayor impacto.</p>



Propósito:

Conozca, describa y aplique los métodos de aprendizaje supervisado y no supervisado, utilizando modelos de regresión, clasificación, y métodos de agrupamiento y reducción de dimensionalidad para resolver problemas reales de representación, clasificación y análisis de datos numéricos y categóricos, al finalizar la unidad de aprendizaje, mediante la programación en lenguaje Python, utilizando plataformas de programación estándares en la industria y el cómputo científico, para analizar el impacto de sus soluciones, justificando sus elecciones, con responsabilidad social y ética profesional.

Competencias que contribuyen al perfil de egreso

Competencias genéricas:

- CG8. Capacidad creativa.
- CG14. Capacidad de aplicar los conocimientos en la práctica.
- CG19. Conocimiento sobre el área de estudio y la profesión.

Competencias específicas:

- CE11. Desarrolla sistemas computacionales inteligentes utilizando una computadora con la arquitectura y lenguaje de programación adecuados para la resolución de problemas con una actitud investigativa y socialmente responsable.
- CE12. Implementa, prueba y mantiene proyectos de sistemas inteligentes empleando criterios de cumplimiento según estándares de calidad establecidos y aprovechando al máximo sus recursos, para resolver problemas científicos y tecnológicos y tomar decisiones que generen bienestar para la sociedad en su conjunto.

CONTENIDOS

Bloques	Temas
1. Introducción al curso	1.1 Nociones básicas 1.1.1 Vocabulario y nociones básicas 1.1.2 Elementos del aprendizaje automático 1.1.3 El problema de la dimensionalidad 1.1.4 Preguntas clave en aprendizaje automático 1.1.5 Regresión versus Clasificación 1.1.6 Clasificación supervisada versus no supervisada



	<p>1.2 Distinguiendo el tipo de problema</p> <p>1.2.1 Características de un problema de regresión</p> <p>1.2.2 Características de un problema de clasificación</p> <p>1.2.3 Contextos de uso de la clasificación supervisada</p> <p>1.2.4 Contextos de uso de modelos no supervisados</p> <p>1.3 Plataformas y herramientas para desarrollo de soluciones</p> <p>1.3.1 Plataformas para desarrollo local</p> <p>1.3.2 Plataformas de desarrollo en la nube</p> <p>1.3.3 Gestión de proyectos colaborativos</p> <p>1.3.4 Introducción a Scikit-learn</p> <p>1.3.5 Herramientas de extracción de datos en web</p>
2. Regresión lineal	<p>2.1 Conceptos básicos en regresión lineal</p> <p>2.1.1 Funciones base</p> <p>2.1.2 Noción de características (features)</p> <p>2.1.3 Características polinomiales</p> <p>2.1.4 Noción de función de pérdida/error</p> <p>2.2 Ajuste de un modelo</p> <p>2.2.1 Complejidad y sobre-ajuste</p> <p>2.2.2 Métricas de rendimiento R2 MSE</p> <p>2.2.3 Regularización</p> <p>2.2.4 Mínimos Cuadrados Ordinarios</p> <p>2.3 Resolución de problemas de regresión</p> <p>2.3.1 Pre-procesamiento y normalización de datos</p> <p>2.3.2 Problemas de regresión uni-variada</p> <p>2.3.2 Problemas de regresión multi-variada</p> <p>2.3.3 Implementación de modelos con Python</p>
3. Clasificación lineal	<p>3.1 Conceptos básicos en clasificación supervisada</p> <p>3.1.1 Función de pérdida en clasificación</p> <p>3.1.2 Descenso de gradiente</p> <p>3.1.3 Funciones discriminantes</p> <p>3.1.4 Mínimos cuadrados para clasificación</p>



	<p>3.1.5 Métricas de rendimiento y Curvas ROC</p> <p>3.2 Clasificadores lineales</p> <p>3.2.1 Regresión logística</p> <p>3.2.2 Algoritmo perceptron</p> <p>3.2.3 SVM</p> <p>3.2.4 Métricas de rendimiento y Curvas ROC</p> <p>3.2.5 Clasificación multi-variada</p> <p>3.3 Resolución de problemas de clasificación lineal</p> <p>3.3.1 Validación cruzada</p> <p>3.3.2 Implementación de modelos con Python</p>
4. Árboles de decisión	<p>4.1 Conceptos básicos</p> <p>4.1.1 Entropía de Shannon</p> <p>4.1.2 Ganancia de información</p> <p>4.1.3 Métodos de decisión</p> <p>4.2 Árboles de decisión</p> <p>4.2.1 Árboles de regresión</p> <p>4.2.2 Árboles de clasificación</p>
5. Agrupamiento	<p>5.1 Agrupamiento por particiones</p> <p>5.1.1 Dispersión Intra-cluster e inter-cluster</p> <p>5.1.2 Algoritmo K-Means</p> <p>5.1.3 Métodos de selección de K</p> <p>5.1.4 Métricas de evaluación de cúmulos</p> <p>5.2 Agrupamiento jerárquico</p> <p>5.2.1 Método aglomerativo</p> <p>5.2.2 Método divisivo</p> <p>5.2.3 Métodos de aglomeración/división</p> <p>5.2.4 Dendogramas</p> <p>5.3 Agrupamiento basado en densidad</p> <p>5.3.1 Nociones básicas sobre DBSCAN</p>
6. Reducción de dimensionalidad	<p>6.1 Descomposición en valores singulares</p> <p>6.1.1 Factorización de matrices</p>



	6.1.2 Valores singulares
	6.1.3 EL método SVD
	6.1.4 Interpretaciones intuitivas y geométricas
	6.2 Análisis de Componentes Principales — PCA
	6.2.1 Usos de PCA
	6.2.2 Noción de cambio de base
	6.2.3 Proyección a nueva base
	6.2.4 Matriz de covarianza
	6.2.5 Interpretaciones intuitivas y geométricas
	6.3 Representación de datos
	6.3.1 Uso de PCA para representaciones latentes
	6.3.2 Mapas Auto-organizables
	6.3.3 Modelo t-SNE

ESTRATEGIAS DE ENSEÑANZA - APRENDIZAJE

Estrategias de aprendizaje sugeridas (Marque X)			
Aprendizaje basado en problemas	(X)	Nemotecnia	()
Estudios de caso	()	Análisis de textos	()
Trabajo colaborativo	(X)	Seminarios	()
Plenaria	()	Debate	(X)
Ensayo	()	Taller	()
Mapas conceptuales	()	Ponencia científica	()
Diseño de proyectos	(X)	Elaboración de síntesis	()
Mapa mental	()	Monografía	()
Práctica reflexiva	()	Reporte de lectura	()
Trípticos	()	Exposición oral	(X)
Otros			



Estrategias de enseñanza sugeridas (Marque X)			
Presentación oral (conferencia o exposición) por parte del docente	(X)	Experimentación (prácticas)	(X)
Debate o Panel	()	Trabajos de investigación documental	()
Lectura comentada	(X)	Anteproyectos de investigación	()
Seminario de investigación	()	Discusión guiada	(X)
Estudio de Casos	()	Organizadores gráficos (Diagramas, etc.)	()
Foro	()	Actividad focal	()
Demostraciones	()	Analogías	()
Ejercicios prácticos (series de problemas)	(X)	Método de proyectos	(X)
Interacción la realidad (a través de videos, fotografías, dibujos y software especialmente diseñado).	()	Actividades generadoras de información previa	()
Organizadores previos	()	Exploración de la web	()
Archivo	()	Portafolio de evidencias	()
Ambiente virtual (foros, chat, correos, ligas a otros sitios web, otros)	()	Enunciado de objetivo o intenciones	(X)

CRITERIOS DE EVALUACIÓN

Criterios	Porcentaje
<ul style="list-style-type: none"> • Exámenes parciales • Proyectos en clase • Proyecto final • Tareas 	30% 30% 30% 10%
Total	100 %



PERFIL DEL PROFESORADO

Licenciatura, Maestría o Doctorado en ciencias computacionales, matemáticas o ingeniería en áreas afines a las ciencias computacionales, con experiencia docente en el área.

REFERENCIAS

Básicas:

- Salvador Garca, Julin Luengo, and Francisco Herrera. (2014). *Data Preprocessing in Data Mining*. Springer Publishing Company, Incorporated.
- Kevin P. Murphy. (2012). *Machine Learning A Probabilistic Perspective*. The MIT Press Cambridge.
- Flach, Peter. (2012). *Machine Learning: The Art and Science of Algorithms that Make Sense of Data*. Cambridge University Press.
- Trevor Hastie, Robert Tibshirani, Jerome H. Friedman. (2009). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. 2da edición. Springer series in statistics.
- Bishop, Christopher M. (2006). *Pattern recognition and machine learning*. New York. Springer.
- Jaynes, E.T. (2003). *Probability Theory - The Logic of Science*. Cambridge University Press.

Complementarias:

- Geron Aurelien. (2017). *Hands-On Machine Learning with Scikit-Learn and Tensorflow: Concepts, Tools, and Techniques to Build Intelligent Systems*. O'Reilly
- Pedro Domingos. (2018). *The Master Algorithm: How the Quest for the Ultimate Learning Machine Will Remake Our World*. Basic Books, Inc.
- Página de GitHub de Sebastian Raschka: <https://github.com/dkryadav/python-machine-learning-by-sebastian-raschka-and-vahid-mirjalili-pdf-3rd-edition>

