

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: ANÁLISIS ESTADÍSTICO DE DATOS				Ciclo de formación: Profesional Eje general de formación: Para la Generación y Aplicación del Conocimiento Semestre: 6º			
Elaborada por: Dr. Mauricio Rosales Rivera				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:
	03	02	05	08	Optativa	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

Presentación: En esta unidad de aprendizaje se estudian habilidades fundamentales que le permitirán a la y el estudiante, comprender el análisis estadístico directamente aplicable a situaciones de la vida real.
Propósito: Adquiera conocimientos sobre los fundamentos de la estadística, al finalizar la unidad de aprendizaje, mediante el entendimiento y visualización de diferentes tipos de datos con diferentes lenguajes de programación, para resolver problemas de la vida real, con responsabilidad sustentable.
Competencias que contribuyen al perfil de egreso
Competencias genéricas:



- CG1. Capacidad para el aprendizaje de forma autónoma.
- CG5. Capacidad de aprender y actualizarse permanentemente.
- CG6. Capacidad para la investigación.

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.
- CE13. Analiza impactos locales y globales de la Inteligencia Artificial mediante el uso de criterios objetivos utilizando lenguaje técnico apropiado, comunicando efectivamente conceptos, métodos y resultados en forma oral y escrita, para presentar propuestas y proyectos de una manera ética y responsable.

CONTENIDOS

Bloques	Temas
1. Estadística descriptiva: Fundamentos	1.12. Entendimiento de la diferencia entre muestra y población. Simulación numérica 1.13. Tipos de datos 1.14. Niveles de medición 1.15. Variables categóricas y numéricas 1.16. Tablas de frecuencias para diferentes tipos de variables
2. Obtención, visualización y mediciones de datos	2.17. Análisis de conjuntos de datos: limpieza, detección de anomalías 2.18. Histogramas, gráficos de dispersión, tablas de contingencia 2.19. Estadística paramétrica y robusta: Aplicación 2.20. Calculando y entendiendo la covarianza 2.21. Coeficiente de correlación 2.22. Ejercicios
3. Estadística inferencial de los datos	3.11 ¿Qué es una distribución? Manejando la distribución normal 3.12 Uso del teorema del límite central para comprender tipos de datos 3.13 Errores estadísticos: trabajando con estimadores y estimaciones 3.14 Intervalos de confianza/significancia: herramienta indispensable para la toma de decisiones 3.15 Planteamiento de hipótesis estadísticas de los datos 3.16 Errores estadísticos: Tipo 1 y 2 3.17 Matriz de confusión 3.18 Pruebas de potencia, bootstrapping y permutación 3.19 Ejercicios



4. Modelación con simulación numérica	4.1. Introducción a los modelos probabilistas 4.2. Principios de los métodos Monte Carlo 4.3. Simulación numérica: Monte Carlo 4.4. Construyendo modelos probabilistas 4.5. Simulación de datos: Estacionario y Transitorio 4.6. Modelando datos reales con MCMC (Cadenas Markov / Monte Carlo) 4.7. Ejemplos
5. Metodología de ciencia de datos	5.1. Del problema al enfoque y de los requisitos a la recopilación de datos 5.2. De la comprensión a la preparación de datos 5.3. Análisis exploratorio de los datos 5.4. Panorama de herramientas de aprendizaje máquina 5.4. Introducción de la implementación y comprobación de hipótesis 5.5. Comunicación de resultados

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)	()
Debate o Panel	()	Trabajos de investigación documental	()
Lectura comentada	(X)	Anteproyectos de investigación	()
Seminario de investigación	()	Discusión guiada	(X)
Estudio de Casos	(X)	Organizadores gráficos (Diagramas, etc.)	()
Foro	()	Actividad focal	()
Demostraciones	()	Analogías	()
Ejercicios prácticos (series de problemas)	(X)	Método de proyectos	()
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	()

CRITERIOS DE EVALUACIÓN

Criterios	Porcentaje
<ul style="list-style-type: none"> • Exámenes parciales • Realización de práctica • Búsqueda de información • Participación en clase • Reportes 	30% 30% 20% 10% 10%
Total	100 %



PERFIL DEL PROFESORADO

Licenciatura, Maestría o Doctorado en el área afín a la disciplina de la unidad de aprendizaje (Ciencias Computacionales)

REFERENCIAS

Básicas:

- Bruce, P., Bruce, A., & Gedeck, P. (2020). *Practical Statistics for Data Scientists: 50+ Essential Concepts Using R and Python*. O'Reilly Media.
- Müller, A. C., & Guido, S. (2016). *Introduction to machine learning with Python: a guide for data scientists*. " O'Reilly Media, Inc."
- Chen, D. G., Jin, Z., Li, G., Li, Y., Liu, A., & Zhao, Y. (Eds.). (2017). *New Advances in Statistics and Data Science*. Cham: Springer International Publishing.

Complementarias:

- Simulation and modeling of natural processes: <https://www.coursera.org/learn/modeling-simulation-natural-processes>

