

1. DATOS INFORMATIVOS

DOMINIO: Hábitat, Infraestructura y Movilidad	
CARRERA: Maestría virtual en Sistemas de Información mención Data Science	
Asignatura/Módulo: Adquisición, gestión y gobernanza de datos	
Paralelo: N° horas 96	
Plan de estudios: 2024-02	H. aprendizaje en contacto con el docente: 11
Prerrequisitos:	
H. aprendizaje autónomo: 74	
H. aprendizaje práctico-experimental: 11	
Periodo académico:	2024-02
Docente o Co-Docente 1:	Grado académico y título profesional:
Eduardo José Montero Bermúdez	Master en Ciencia de datos. Científico de datos
Co-Docente 2:	Grado académico y título profesional:
Breve reseña de la actividad	
Breve reseña de la actividad académica y/o profesional:	
Indicación de horario de atención al estudiante:	
Tutoría presencial:	Teléfono:
Tutoría virtual:	Correo electrónico: ejmontero@puce.edu.ec

2. DESCRIPCIÓN DE LA ASIGNATURA

DESCRIPCIÓN DE COMPETENCIAS Y RESULTADOS DE APRENDIZAJE

COMPETENCIAS TRANSVERSALES	COMPETENCIAS DISCIPLINARES DEL DOMINIO	RESULTADOS DE APRENDIZAJE DE LA CARRERA	RESULTADOS DE APRENDIZAJE DE LA ASIGNATURA
1. Humanista y con p	Evaluación del entorno complejo Transformación pertinente y sostenible del hábitat	1. Recopilar datos de diversas fuentes de información, alineado con el objetivo específico planteado o el problema a resolver.	Utilizar cada una de las etapas del proceso de Descubrimiento de Conocimiento en el contexto de Big Data
2. Comprometido soci	Innovación responsable con el medio ambiente Transformación pertinente y sostenible del hábitat	2. Analizar datos estructurados y no estructurados utilizando técnicas y métodos estadísticos, asegurando el rigor y la precisión del análisis mediante el método científico	
3. Crítico y analíti	Evaluación del entorno complejo Transformación pertinente y sostenible del hábitat	3. Desarrollar modelos de Inteligencia Artificial (IA) inductivos (predictivos) o deductivos (descriptivos), que apoyen a las instituciones en procesos de toma de decisiones estratégicas.	Emplear metodologías apropiadas en la obtención de modelos inductivos (predictivos) y deductivos (descriptivos), que faciliten los procesos de toma de decisiones basadas en datos en las Instituciones.
4. Creativo, innovad	Innovación responsable con el medio ambiente Aplicación de lenguajes espaciales, objetuales, visuales y verbales	4. Crear visualizaciones de datos que faciliten la comprensión y apoyen la toma de decisiones	
5. Motivado a ser má	Evaluación del entorno complejo Transformación pertinente y sostenible del hábitat	5. Aplicar principios éticos en la gestión y análisis de datos, con responsabilidad social empresarial.	Emplear técnicas y métodos analíticos basados en la estadística, la inteligencia artificial, y el procesamiento de datos para la identificación de patrones y/o relaciones en el descubrimiento de comportamientos relevantes en las Instituciones
6. Comunicador asert	Evaluación del entorno complejo Aplicación de lenguajes espaciales, objetuales, visuales y verbales		

4. EVALUACIÓN DE LOGROS DE APRENDIZAJE

Resultado de aprendizaje	Definición del criterio de evaluación del EdA	Ponderación (en porcentaje)
<p>Utilizar cada una de las etapas del proceso de Descubrimiento de Conocimiento en el contexto de Big Data</p>	<p>Criterio 1: Identifica las etapas del proceso de descubrimiento del conocimiento</p>	33
	<p>Criterio 2: Asocia las subtareas que tiene cada etapa del procesos de descubrimiento del conocimiento</p>	33
	<p>Criterio 3: Implementa el proceso de descubrimiento de conocimiento en un caso práctico</p>	33
<p>Emplear metodologías apropiadas en la obtención de modelos inductivos (predictivos) y deductivos (descriptivos), que faciliten los procesos de toma de decisiones basadas en datos en las Instituciones.</p>	<p>Criterio 1: Relaciona las metodologías de modelos predictivos y descriptivos con los objetivos del negocio en una organización</p>	50
	<p>Criterio 2: Identifica las etapas de cada una de las metodologías en la obtención de modelos inductivos y deductivos</p>	50
<p>Emplear técnicas y métodos analíticos basados en la estadística, la inteligencia artificial, y el procesamiento de datos para la identificación de patrones y/o relaciones en el descubrimiento de comportamientos relevantes en las Instituciones</p>	<p>Criterio 1: Aplica métodos estadísticos y de inteligencia artificial adecuados en la identificación de patrones</p>	33
	<p>Criterio 2: Interpreta correctamente los resultados de la aplicación de algoritmos de inteligencia artificial</p>	33
	<p>Criterio 3: Desarrolla correctamente procesos que permitan garantizar la calidad en el preprocesamiento de datos</p>	33

5. METODOLOGÍA

6. RELACIÓN RESULTADOS DE APRENDIZAJE, EXPERIENCIAS DE APRENDIZAJES Y DIMENSIÓN DEL CONOCIMIENTO

RESULTADOS DE APRENDIZAJE DE LA ASIGNATURA	Semana	DIMENSIÓN DEL CONOCIMIENTO (conceptos, hechos, procedimientos o principios)
<p>Entender cada una de las etapas del proceso de Descubrimiento de Conocimiento en el contexto de Big Data</p>		<p>1. Datos 2. Tipos de Atributos 3. Tipos de datos según su estructura</p>
		<p>4. Proceso de Descubrimiento del Conocimiento. 4.1. Selección 4.2. Preprocesamiento 4.3. Transformación 4.4. Minería 4.5. Interpretación</p>
		<p>5. Técnicas de Preprocesamiento de datos 5.1. Limpieza 5.2. Integración</p>
		<p>5.3. Detección de Outliers 5.4. Reducción 5.5. Transformación</p>
		<p>6. Características del modelo de datos 7. Tipos de modelos de datos 7.1. Modelo Lógico 7.2. Modelo Conceptual 7.3. Modelo Físico</p>
		<p>8. Conceptos de Business Intelligence 9. Normalización 10. Modelos Dimensionales</p>
<p>Esquematar metodologías apropiadas en la obtención de modelos inductivos (predictivos) y deductivos (descriptivos), que faciliten los procesos de toma de decisiones basadas en datos en las Instituciones.</p>		<p>11. Recuperación de información 12.1. Definiciones</p>
		<p>13. Sistemas de recuperación de información 13.1. Modelo Booleano 13.1.1. Índice Invertido 14. Análisis de Sentimientos 14.1. Normalización</p>
		<p>15. Teoremas de Bases de datos 15.1. Teorema ACID 15.2. Teorema BASE</p>
		<p>16. NO SQL Databases 16.1. Mongo DB 16.2. Cassandra</p>
		<p>17. Data Governance Program (DGP) 17.1. Roles del Data Governance Program (DGP) 17.2. Implementación del DGP</p>
		<p>17.3. Formas de implementación del DGP 17.4. Ejecución del DGP 17.5. Medición del DGP</p>
<p>Emplear técnicas y métodos analíticos basados en la estadística, la inteligencia artificial, y el procesamiento de datos para la identificación de patrones y/o relaciones en el descubrimiento de comportamientos relevantes en las Instituciones</p>		<p>18. CRISP 18.1. Metodología 18.2. Modelo de Proceso 19. SEMMA</p>
		<p>20. Elementos de la ciencia de datos 20.1. Análisis de correlaciones 20.2. Métricas de performance 20.2.1. Curva RoC 20.2.2. Matriz de Confusión 21. Algoritmos de aprendizaje supervisado 21.1. Decision Tree 21.2. Random Forest</p>
		<p>22. Introducción a python para ciencia de datos 22.1. Tipos de datos y estructuras 22.2. Joins y funciones básicas</p>
		<p>22.3. Analíticos descriptivos 22.4. Librerías de visualización 22.5. Ejercicios Prácticos</p>

8. BIBLIOGRAFÍA

a. BÁSICA / LIBROS

Bibliografía (basarse en normas APA)	Código Biblioteca PUCESI	Nro. de ejemplares	Link
Piattini Velthuis, M. G. (2018). Calidad de datos y gobierno de la información: Aspectos, técnicas y modelos (eBook). RA-MA Editorial. https://doi.org/10.1234/56789	2498292		https://eds.p.ebscohost.com/eds/detail/detail?vid=1&sid=cf4ec0a0-f32c-40da-b8e7-4af54971bab2%40redis&bdata=Jmxhbm9ZXMmc2l0ZT1lZHMtbGl2ZQ%3d%3d#AN=2498292&db=nlebk
Dinesh, L., & Devi, K. G. (2024). An efficient hybrid optimization of ETL process in data warehouse of cloud architecture. Journal of Cloud Computing: Advances, Systems and Applications, 13(1), 12. https://doi.org/10.1186/s13677-023-00571-y	edssjs.79B4C7EA		https://edsp.puce.elogim.com/eds/detail/detail?vid=7&sid=cf4ec0a0-f32c-40da-b8e7-4af54971bab2%40redis&bdata=Jmxhbm9ZXMmc2l0ZT1lZHMtbGl2ZQ%3d%3d#AN=edssjs.79B4C7EA&db=edssjs
Ramos, J. A., & Silva, A. M. (2018). Relational models and work-family conflict: A study with samples from Brazil and the United States of America. Ciencias Psicológicas, 12(2), 167-176. https://doi.org/10.22235/cp.v12i2.1679	10.22235/cp.v12i2.1679		https://edsp.puce.elogim.com/eds/detail/detail?vid=10&sid=cf4ec0a0-f32c-40da-b8e7-4af54971bab2%40redis&bdata=Jmxhbm9ZXMmc2l0ZT1lZHMtbGl2ZQ%3d%3d#AN=155131274&db=pbh

b. COMPLEMENTARIA / CAPÍTULOS EN LIBROS

Bibliografía (basarse en normas APA)	Código Biblioteca PUCESI	Nro. de ejemplares	Link
Leavitt, K., Schabram, K., Hariharan, P., & Barnes, C. M. (2024). The machine hums! Addressing ontological and normative concerns regarding machine learning applications in organizational scholarship. Academy of Management Review, 49(2), 448-451. https://doi.org/10.5465/amr.2021.0166	176926381		https://edsp.puce.elogim.com/eds/detail/detail?vid=12&sid=cf4ec0a0-f32c-40da-b8e7-4af54971bab2%40redis&bdata=Jmxhbm9ZXMmc2l0ZT1lZHMtbGl2ZQ%3d%3d#AN=176926381&db=bth
Hoare, J., & Bock, T. (2019). A brand's eye view of correspondence analysis. International Journal of Market Research, 61(1), 12-21. https://doi.org/10.1177/1470785318801480	134679382		https://edsp.puce.elogim.com/eds/detail/detail?vid=14&sid=cf4ec0a0-f32c-40da-b8e7-4af54971bab2%40redis&bdata=Jmxhbm9ZXMmc2l0ZT1lZHMtbGl2ZQ%3d%3d#AN=134679382&db=bth
CheshmehSohrabi, M., & Mashhadi, A. (2023). Using data mining, text mining, and bibliometric techniques to the research trends and gaps in the field of language and linguistics. Journal of Psycholinguistic Research, 52(2), 607-630. https://doi.org/10.1007/s10936-022-09911-6	10.1007/s10936-022-09911-6		https://edsp.puce.elogim.com/eds/detail/detail?vid=17&sid=cf4ec0a0-f32c-40da-b8e7-4af54971bab2%40redis&bdata=Jmxhbm9ZXMmc2l0ZT1lZHMtbGl2ZQ%3d%3d#AN=edssjs.A1FF1BBC&db=edssjs



MSc. Eduardo Montero