



SOMOS
ARTE, CIENCIA Y
DESARROLLO
CULTURAL



Materia: Ciencia de Datos

DATOS GENERALES:

Descripción:	En este curso se estudian las principales técnicas de recopilación, interpretación y procesamiento de datos para obtener información relevante y refinada. Se abordará la investigación científica de datos, la sintaxis sobre el lenguaje de programación Python y sus bibliotecas esenciales. También se estudian las redes neuronales artificiales, métodos exploratorios, clustering con aprendizaje no supervisado, agrupaciones jerárquicas y métodos predictivos y de regresión con aprendizaje supervisado.
Seriación y Correlación:	Subsecuentes:****
	Consecuentes:****
Objetivo:	Aprender los fundamentos esenciales de la ciencia de datos utilizando el lenguaje de programación Python. Englobando herramientas estadísticas y el aprendizaje autónomo, para la resolución de problemas y hacer toma de decisiones en base a grandes cantidades de datos los cuales se analizarán por agrupamiento, clasificación, regresión y filtrado.
Objetivos específicos:	<ul style="list-style-type: none"> • Programar en lenguaje Python básico y orientada a objetos. • Adquisición de datos de distintas bases de datos asociadas a la aplicación Python • Analizar tablas de datos reales utilizando códigos propios • Realizar modelos exploratorios y modelos predictivos con las herramientas del lenguaje de programación
Horas totales del curso:	(**) horas presenciales + (**) horas de autoestudio=** horas totales
Créditos:	** créditos

REVISIONES Y ACTUALIZACIONES:

Academia:	**
Autores o Revisores:	**
Fecha de actualización por academia:	15 de junio de 2022



SOMOS
ARTE, CIENCIA Y
DESARROLLO
CULTURAL



Sinopsis de la revisión y/o actualización:

PERFIL DESEABLE DEL DOCENTE:

Disciplina profesional:	Doctorado en ciencias
Experiencia docente:	Experiencia profesional docente mínima de dos años

ÍNDICE TEMÁTICO:

TEMA:	SUBTEMA
Lenguaje de programación Python y librerías	<ul style="list-style-type: none"> · Herramientas de trabajo · Jupyter · Modelado Predictivo · Asignaciones · Control de Flujo · Estructuras de datos · NumPy · Matplotlib · Pandas · Google Colab
Análisis exploratorio de datos	<ul style="list-style-type: none"> · Fuentes de Datos · Estadística Descriptiva · Visualización de Datos · Tidy Data · Cargar un conjunto de datos · Entender nuestros datos · Histogramas · Diagrama de Densidad y Boxplot. · Matriz de correlación y dispersión · Dispersión y Boxplot por clase ·



SOMOS
ARTE, CIENCIA Y
DESARROLLO
CULTURAL



Procesamiento de datos	<ul style="list-style-type: none">· Preprocesamiento de datos· Métodos de transformación de datos· Métodos de remuestreo· Detección y análisis de outliers· Preprocesamiento de datos.· Escalamiento y estandarización· Normalización y Binarización· BoxCox y YeoJohnson· Métodos de remuestreo· Validación cruzada.· División por porcentaje
Tratamiento de datos	<ul style="list-style-type: none">· Evaluación de algoritmos· Métricas Accuracy y Kappa· Métricas ROC y Matriz de confusión· Reporte de clasificación· Métrica MAE· Métricas MSE y R2· Feature selection· Correlación entre características· Backward y Univariable· Reducción de dimensiones con PCA
Modelado de datos	<ul style="list-style-type: none">· Algoritmos de Machine Learning· Algoritmos de Taxonomía lineal· Algoritmos de Taxonomía no lineal· Algoritmos Ensemble· Algoritmos lineales - Regresión· Algoritmos lineales - Clasificación· Algoritmos No lineales - Clasificación· Algoritmos No lineales - Regresión· Comparación de Algoritmos - Simple· Comparación de Algoritmos - Visualización· Algoritmos de conjunto
Aprendizaje no Supervisado	<ul style="list-style-type: none">· Algoritmos de Aprendizaje No Supervisado· Algoritmos de clustering· Implementación Algoritmos clustering· k-Means· Clustering Jerárquico· Métodos basados en densidad· Determinar clusters



SOMOS
ARTE, CIENCIA Y
DESARROLLO
CULTURAL



BIBLIOGRAFIA

Principal:	<ul style="list-style-type: none"> • Statistics in a nutshell, Sarah Boslaugh and Paul Andrew Watters, 2008 • Practical Statistics for Data Scientists, Peter Bruce, Andrew Bruce, and Peter Gedeck, 2020 • A Primer on Scientific Programming with Python Fourth Edition, Hans Petter Langtangen, 2014 • An Introduction to Statistics with Python With Applications in the Life Sciences, Thomas-Haslwanter, 2015 • Data Science from Scratch, Joel Grus, 2015 • Statistics for Python, Michiel de Hoon, Columbia University, 2010 • Mastering Machine Learning with Python in Six Steps, Manohar Swamynathan, 2019 • Numerical Python Scientific Computing and Data Science Applications with Numpy, SciPy and Matplotlib, Robert Johansson, 2019 • Python Data Analytics Data Analysis and Science Using Pandas, matplotlib, and the Python Programming Language, Fabio Nelli, 2015 • Python for Probability, Statistics, and Machine Learning, José Unpingco, 2016 • Python Data Science Handbook Essential Tools for Working with Data, Jake VanderPlas, 2017 • Statistics and Machine Learning in Python, Edouard Duchesnay, Tommy Löfstedt, Feki Younes, 2019 • Thoughtful Machine Learning with Python, Matthew Kirk, 2017
Enlaces digitales:	
Complementaria:	

PLANEACIÓN EDUCACIONAL:

Competencias generales:	<p>Desarrollar aplicaciones en el lenguaje de programación Python con herramientas para la ciencia de datos como modelos de exploración, modelos de predicción y redes neuronales, con el objetivo de resolver problemas reales en el ámbito educativo, científico o empresarial, con la finalidad de ahorrar tiempo, recursos y energía.</p>
-------------------------	---



SOMOS
ARTE, CIENCIA Y
DESARROLLO
CULTURAL



Competencias específicas:

- Aplicar técnicas analíticas y visuales de grandes cantidades de datos para resolver problemas con Python
- Entender las diferentes técnicas y mecanismos para la predicción de análisis de datos
- Desarrollar y analizar proyectos de aprendizaje no supervisado
- Aprender las características fundamentales de las redes neuronales artificiales y recurrentes
- Aplicar técnicas y métodos para el tratamiento y monitoreo de datos para mejorar la robustez de resultados
- Entender las características y diferencias de Machine Learning y Deep Learning

CONTRIBUCIÓN AL PERFIL DE EGRESO:

CONOCIMIENTO:	HABILIDADES:	VALORES:
Tener una comprensión profunda del lenguaje de programación Python y ser capaz de extraer el mayor conocimiento complejo e implícito, que sea potencialmente útil (anomalía, desviación, patrón, valor anómalo, tendencias etc.) a partir de una base de datos con un gran volumen de información.	Utilizar las técnicas, métodos y herramientas de la inteligencia artificial, aprendizaje autónomo, estadística y sistemas de base de datos con un enfoque al científico de datos, para la toma de direcciones, con el fundamento en los datos analizados con modelos matemáticos.	Tener hábitos de trabajo necesarios para el desarrollo de la profesión tales como el rigor científico, el autoaprendizaje y la persistencia. Actuar con responsabilidad, honradez y ética profesional, manifestando conciencia social de solidaridad y justicia. Mostrar tolerancia en su entorno social, aceptando la diversidad cultural, étnica y humana. Desarrollar un mayor interés por aquellos problemas cuya solución sea de beneficio social y el medio ambiente



SOMOS
ARTE, CIENCIA Y
DESARROLLO
CULTURAL



ESTRATEGIAS PEDAGÓGICAS:

Estrategias de enseñanza:	Estrategias de aprendizaje:
<ul style="list-style-type: none"> • El docente explicará la teoría y presentará ejemplos en las clases presenciales o virtuales. • El docente explicara los fundamentos del lenguaje de programación Python. • Se motivará con prácticas para el trabajo individual y en equipo. • Discusión de preguntas y problemas en clase. • Se motivará para realizar un proyecto final 	<ul style="list-style-type: none"> • El alumno asistirá al menos a un 80% de las clases principales o virtuales impartidas. • El estudiante trabajará en forma individual o por equipo en la resolución y programación de las aplicaciones. • El estudiante resolverá las practicas que se soliciten, además, generará un reporte de cada practica • El estudiante propondrá y ejecutará el análisis de un proyecto, además de un reporte de proyecto. • Asistirá a asesorías para resolver dudas sobre la teoría o sobre la solución de problemas.

PROPUESTA DE CRITERIOS DE EVALUACIÓN:

Criterio de evaluación:	Porcentaje:
Tres exámenes parciales	60%
Tareas y practicas	25%
Proyecto final	15%
Asistencia	0%