

Introducción al cálculo numérico y ciencia de datos - V2024

Introducción

En la Ciencias de Datos confluyen varios métodos científicos, procesos y sistemas para que, a partir de cierta información, poder extraer conocimiento o un mejor entendimiento de lo datos en sus diferentes formas. Algunas técnicas de análisis empleadas para esto provienen de la estadística, la minería de datos, el aprendizaje automático, y la analítica predictiva¹.

A fin de introducir en el tema, proponemos para estudiantes, docentes y profesionales de diferentes ramas de la ciencias e ingeniería, este curso de Educación Permanente. Curso donde se repasaran conceptos de calculo numérico que permitirá tener una base general para realizar los primeros análisis con técnicas aprendizaje automático.

Como herramienta principal para poder llevar acabo algunas de las aplicaciones que se verán, estará el lenguaje Python y el uso de algunas de las librerías mas populares para implementar redes neuronales y algoritmos de clasificación.

Se pretende introducir este tema muy en estudiado últimamente, explicando conceptos de forma amigable, acompañado de ejemplos varios en la industria y el agro.

Público Objetivo:

- Egresados, investigadores, docentes, estudiantes de grado y posgrado, de diferentes disciplinas universitarias que necesiten incorporar conocimientos de herramientas para análisis predictivos a partir de datos.
- Profesores de Ciencias Básicas, profesionales, emprendedores y empleados en relación de dependencia, que deseen incorporar estas herramienta en su formación.
- De forma excluyente, deberán poseer conocimiento de álgebra matricial, programación y manejo de Python.

Contenido:

Módulo	Tema	Exposiciones Teóricas	Actividades Prácticas
1	<ul style="list-style-type: none"> • Introducción al curso • Aprendizaje automático y Aprendizaje Profundo: definición, conceptos, diferencias y aplicaciones. • Presentación de entorno Google Colab y ejecución de scripts. 	1 clase teórica (3 hs)	Trabajo adicional durante la semana (3 hs).
2	<ul style="list-style-type: none"> • Resolución de sistemas de ecuaciones lineales. • Resolución de ecuaciones no lineales. • Optimización por métodos con y sin derivadas. Máximos y mínimos. • Ajuste de funciones. Mínimos cuadrados. 	2 clases teóricas (6 hs)	1 clase de práctica para realizar ejercicios (3 hs) Trabajo adicional durante la semana (3 hs).
3	<ul style="list-style-type: none"> • Redes Neuronales: estructuras, tipos, activación, costos y entrenamiento. • Modelos con Scikit-Learn y H2O • Ejemplos de aplicaciones. 	3 clases teóricas (9 hs)	2 clases de práctica para realizar ejercicios (6 hs) Trabajo adicional durante la semana (3 hs).
4	<ul style="list-style-type: none"> • Minería de Datos (Data Mining): definición, conceptos y aplicaciones. 	3 clases teóricas (9 hs)	2 clases de práctica para realizar ejercicios (6 hs)

¹ Liu, Alex. 2015. «Data Science and Data Scientist», <http://www.researchmethods.org/DataScienceDataScientists.pdf>



	<ul style="list-style-type: none"> Reducción de la dimensionalidad: análisis de componentes principales. Clasificación: arboles de decisión y SVM. Ejemplos de aplicaciones. 		Trabajo adicional durante la semana (6 hs).
--	---	--	---

Régimen de cursada:

14 clases de 3 horas cada una

- 9 clases teóricas (27 hs)
- 5 clases prácticas (15 hs)
- 18 hs de **autoestudio**, donde el estudiante deberá repasar conceptos vistos, leer material complementario y trabajar en la resolución de ejercicios.

Cronograma esquemático:

Horarios: Lunes de 17:00 a 20:00, miércoles y viernes de 18:00 a 21:00 hs.

Semana	Lunes	Martes	Miércoles	Jueves	Viernes	Sábado	Domingo
1	Teórico M1		Teórico M2		Teórico M2		
2	Práctico M2		Teórico M3		Teórico M3		
3	Práctico M3		Teórico M3		Práctico M3		
4	Teórico M4		Teórico M4		Práctico M4		
5	Teórico M4		Práctico M4		Evaluación		

Evaluación:

Prueba práctica de resolución de ejercicios. Aprobación con 60% o más de la pauta propuesta.

Asistencia:

75% de las clases.

Bibliografía:

- M. F. Fenner (2020) Machine Learning with Python for Everyone, Addison-Wesley
- C.M. Bishop (2006) Pattern Recognition and Machine Learning (Information Science and Statistics)
- J. D. Kelleher (2019) Deep Learning (The MIT Press Essential Knowledge series).
- González Duque. Python para todos ([https://launchpadlibrarian.net/18980633/Python para todos.pdf](https://launchpadlibrarian.net/18980633/Python%20para%20todos.pdf))
- Rojas (2018) Prealgebra via Python Programming: First steps to perform large scale computational tasks in the Sciences and Engineerings (<https://github.com/rojassergio/Prealgebra-via-Python-Programming>)

