

Destinatarios

Estudiantes de la carrera de Licenciatura en Economía, Licenciatura en Administración y Contador Público.

Objetivos

Este seminario pretende que el alumno:

- 1) logre un dominio básico de las principales técnicas de ciencia de datos en Python.
- 2) Comprenda las bases conceptuales y los desafíos más importantes en machine learning.
- 3) Conozca la relevancia de la ciencia de datos para el desarrollo de un negocio, de un proyecto gubernamental o de investigación.
- 4) tenga la posibilidad de una rápida inserción en el mercado laboral, en tanto se trata de una de las disciplinas más demandadas en la actualidad.
- 5) culmine su carrera universitaria con una visión más acabada del tipo de problema que se puede abordar con soluciones de machine learning y de las habilidades necesarias para hacerlo.

Contenidos**Unidad 1**

Introducción a la ciencia de datos: ¿qué es y qué problemas aborda? ¿Qué es inteligencia artificial, machine learning y deep learning? ¿Qué profesionales trabajan junto un científico de datos? Ciencia de datos en el sector privado, público y académico.

Introducción a Python: Notebooks y Colab, librerías, virtual environments, terminal. Instalación del software necesario. Dominio básico: tipos de datos, listas, diccionarios, loops, numpy, booleans, estructuras de control y funciones.

Unidad 2

Manipulación y análisis de datos en Python: estructura de un dataset, estadística descriptiva y visualización de datos con Pandas, Matplotlib y Seaborn (media, varianza, histogramas, cuantiles, boxplots, correlaciones, scatterplots), limpieza de outliers y valores nulos.

Unidad 3

Introducción a machine learning: paradigma supervisado y no supervisado, target y features, regresión y clasificación. Inferencia versus predicción: concepto de entrenamiento, train y test sets, parámetros e hiperparámetros, trade off sesgo y varianza, complejidad y overfitting. ScikitLearn. Árboles de decisión para problemas de clasificación. Búsqueda de hiperparametría. Medidas de performance: matriz de confusión, accuracy, precision, recall, curva ROC.

Docentes

- Natalia Fonzo
- Alejandro Parra

Días y Horarios

Lunes de 19 a 22 hs

Metodología

Virtual