

Data Science, Matemáticas y BigData
Facultad de Ciencias

Impartido por:

Dr. Gabriel Guerrero (Matemático de la Fac. de Ciencias UNAM/ Director General en saXsa)

Objetivo

Ofrecer los conocimientos matemáticos y computacionales para implementar en la organización la Ciencia de Datos para la toma de decisiones.

Beneficios

1. Adquirir un conocimiento profundo de la Ciencia de Datos con el estudio de matemáticas, estadística, cómputo distribuido y algoritmos en Python en un ambiente Spark
2. Conocer las técnicas y algoritmos del Aprendizaje Automatizado que promueven la extracción de valor de los datos de las empresas
3. Tener los elementos para la instrumentación inmediata de los algoritmos del Aprendizaje Automatizado (Machine Learning) en una plataforma de cómputo distribuido y tolerante a fallas

Tema 1 Introducción a la Ciencia de Datos con Big Data

Tema 1_1 Identificar los conceptos de la Ciencia de Datos

Tema 1_2 Panorama de Estadística, el Aprendizaje Automatizado (ML Machine Learning) y los modelos matemáticos de la Ciencia de Datos en un ambiente Big Data

Tema 1_3 Casos de aplicación de la Ciencia de Datos con ML en un ambiente Big Data

Tema 1_4 Aspectos matemáticos de Regresión Logística

Tema 1_5 Aspectos matemáticos de Descomposición en Valores Singulares (SVD Singular Value Decomposition)

Tema 1_6 Aspectos matemáticos de algoritmos kNN, los k vecinos más próximos

Tema 2 Gobernanza de Grandes Volúmenes de Datos y visualización de Datos

Tema 2_1 Identificación de tipos de datos por su temporalidad

Tema 2_2 Mapeo del problema de negocio en una problemática de Datos.

Tema 2_3 Análisis Exploratorio de Datos (EPA).

Tema 2_4 Utilización de la Interfaz Jupyter para visualización

Tema 2_5 Utilización de biblioteca Bokeh para graficación

Tema 2_6 Algoritmos matemáticos de la graficación y su aplicación en Mapas

Tema 3 Algoritmos matemáticos del Aprendizaje Automatizado (Machine Learning) y su instrumentación en un ambiente de cómputo distribuido

Tema 3_1 Análisis descriptivo y Análisis predictivo con Spark

Tema 3_2 Aspectos matemáticos de Regresión Logística y algoritmos existentes en un ambiente Big Data

Tema 3_3 Análisis de textos con LSA (Latent Semantic Analysis) y su instrumentación en ambiente Big Data

Tema 3_4 Descomposición por Valores Singulares (SVD Singular Value Decomposition) y su instrumentación en ambiente Big Data

Tema 3_5 Agrupación (Clustering) con algoritmos de tipo K-means

Tema 3_6 Ejemplos de aplicación de los principales algoritmos del Aprendizaje Automatizado, programando en ambiente Big Data

Requisitos

Conocimiento básico de algún lenguaje de programación.

Conocimiento básico de conceptos de estadística y matemáticas.

Duración: 21 horas, divididas en 3 sesiones sabatinas de 7hrs cada una

Fechas: 26 de Agosto, 2 y 9 de septiembre 2017

Costos:

Público en Gral: \$14,000 pesos (más IVA)

Comunidad UNAM: \$ 9,500 pesos (más IVA)