

Big Data. Machine Learning

BGD-1009. 15 horas

Objetivos

- Introducción al aprendizaje automático, la minería de datos y el reconocimiento de patrones estadísticos.
- Afrontar un problema de explotación de los datos basándose en Machine Learning,
- Desarrollar o modificar algoritmos existentes para la mejor adecuación a la solución del problema

Dirigido

- Programadores
- Estadísticos.
- Toda persona que sepa programar y tenga un conocimiento medio de matemáticas y quiera desarrollarse en el campo del Machine Learning

Requisitos Iniciales

- Conocimientos de Sistema Operativo Linux
- Se requieren conocimientos de programación matemáticos y estadísticos, aunque sea a nivel básico

Contenidos

- Introducción
 - Origen e historia del ML
 - Introducción a Python
 - Preparación del entorno
- Estadística 101
 - Uso de DataFrames
 - Media, Mediana y Moda
 - Varianza y desviación estándar
 - Tipos de distribuciones
 - Probabilidad condicional
 - Teorema de Bayes
- Modelos predictivos
 - Regresión Lineal
 - Regresión Polinómica
 - Regresión Multivariable (OLS)

- Machine Learning I
 - Aprendizaje supervisado y aprendizaje sin supervisar
 - Clasificador Naive Bayes
 - K-Means clustering
 - Árboles de decisión
 - Random Forests
 - Support Vector Machines (SVM)
- Machine Learning II
 - KNN (K-Nearest-Neighbors)
 - Reducción de dimensionalidad
 - Análisis de componentes principales (PCA)
 - Data Warehousing ETL vs ELT
- K-fold cross-validation