

## **Métodos Estadísticos para Machine Learning**

### **1.-Introducción**

Introducción a Machine Learning

Qué es Machine Learning

Técnicas estadísticas usuales en Machine Learning

### **2.-Preparación de datos para Machine Learning**

Obtención de datos

Tipos de datos

Preparación de los datos para las técnicas de Machine Learning

Análisis y evaluación de datos

Limpieza de datos

Transformación, descripción y visualización los datos

Ejemplos prácticos

### **3.- Modelos de clasificación**

Conceptos generales sobre aprendizaje supervisado

Conceptos generales sobre aprendizaje no supervisado

Introducción al problema de clasificación

Clasificación a través de vecinos cercanos (Narest Neighbors)

Comprender la clasificación usando el algoritmo de vecinos cercanos

El algoritmo kNN

Cálculo de la distancia

Elegir un k apropiado

Preparación de datos para usar con kNN

Ejemplos prácticos

Classificación usando Naive Bayes

Comprender el algoritmo Naive Bayes

Conceptos básicos de los métodos bayesianos

Probabilidad

Probabilidad conjunta

Probabilidad condicional y teorema de Bayes

El algoritmo de Naive Bayes

Clasificación por medio de Naive Bayes

Ejemplos prácticos

#### **4.- Árboles de decisión, árboles de regresión, bosques aleatorios y boosting**

Introducción a los árboles de decisión

Diferencia entre árboles de clasificación y árboles de regresión

Árboles de clasificación

Índice de Gini

Entropía

Ganancia de información

Ji-cuadrada (CHAID)

Control del tamaño del árbol

Poda de un árbol

Extracción de reglas de un árbol

Ejemplos prácticos

Predicción

Bagging

Bosques aleatorios (Random forest)

Boosting

AdaBoost

Gradient Boosting

Stochastic Gradient Boosting

Ejemplos prácticos

#### **5.- Redes Neuronales**

Introducción a las redes neuronales

Perceptrón

Perceptrón multicapa

Funciones de activación

Medias de error

Ejemplos prácticos

## **6.- Maquinas de soporte vectorial**

Introducción a las máquinas de soporte vectorial

Kernels

Kernel Gaussiano o radial

Kernel polinomial

Kernel sigmoidal

Máquinas de relevancia vectorial

Ejemplos prácticos

## **7.- Minería de textos**

Introducción a la minería de texto

Procesamiento de texto

Representación documento-término

Normalización y medidas de similaridad

Nubes de palabras

Análisis semántico latente

Ejemplos prácticos

## **8.- Técnicas para evaluación de modelos**

Conjunto de entrenamiento, validación y prueba

Validación cruzada

La matriz de confusión

Métricas de exactitud y precisión

Curvas de ROC y área bajo la curva

Los ejemplos prácticos se realizarán utilizando el paquete **R**

## **Bibliografía**

Burger V, Scott (2018) *Introduction to Machine Learning (Rigorous Mathematical Analysis)*. O'REILLY

Ciaburro, G and Venkateswaran (2017) *Neural Network with R*. Packt Publishing

Dangeti, Pratap (2017) *Statistics for Machine Learning*. Packt Publishing

Lantz, B (2013) *Machine Learning with R*. Packt Publishing

Ramasubramanian,k and Singh, A (2019). *Machine Learning Using R With Time Series and Industry-Based Use Cases in R*. Apress