



PONTIFICIA UNIVERSIDAD CATÓLICA DE CHILE
FACULTAD DE MATEMÁTICAS
DEPARTAMENTO DE ESTADÍSTICA
PROFESOR: DANILO ÁLVARES - AYUDANTE: RUBÉN SOZA

Ayudantía Introducción al Machine Learning

Diplomado DataScience UC

24 de Octubre 2020

```
# Librerías de Hoy  
library(tidyverse);library(factoextra); library(cluster)
```

Modelos no supervisados en R y Python

Modelos para datos continuos.

Cargue la base de datos `World.xlsx`, la cual contiene variables que describen a los países en diferentes ámbitos.

1. Realice una agrupación de los datos utilizando el algoritmo de K -means.
2. Realice una agrupación de los datos utilizando un clustering jerárquico.
3. Realice una reducción de dimension de los datos utilizando PCA.
4. Realice una reducción de dimension de los datos utilizando Análisis Factorial.

Modelos para datos discretos

Cargue la base de datos `mushrooms.csv`, la cual contiene información de 8124 tipos de hongos.

1. Realice una agrupación de los datos utilizando el algoritmo de K -modes.
2. Realice una agrupación de los datos utilizando un clustering jerárquico.