

Detección de anomalías

Laura del Pino Díaz

4/2/2017

Índice

Personas en el gimnasio del campus	2
Detección de anomalías	3
Paquetes necesarios	3

Personas en el gimnasio del campus

La base de datos que vamos a usar es *Crowdeness at campus gym* que traducido al español es la población que hay dentro de un gimnasio universitario.

Los atributos que tenemos son los siguientes:

- timestamp (entero): el número de segundos desde el comienzo del día.
- day_of_week (entero): día de la semana del 0 al 6.
- is_weekend (entero): si el día pertenece a un fin de semana.
- is_holiday (entero): si el día es un día de vacaciones.
- apparent_temperature(real): temperatura aparente en el momento en el que se realiza la observación, en grados Fahrenheit.
- temperature (real): temperatura medida.
- is_start_of_semester (entero): indica si es el comienzo del semestre.

Como salida se tiene el siguiente atributo: * Number_of_people (entero)

El objetivo sobre esta base de datos es detectar si existen anomalías en el número de personas que están en el gimnasio.

Base de datos obtenida de Kaggle

Detección de anomalías

Paquetes necesarios

Son varios los paquetes de detección de anomalías que podemos encontrar disponibles para R, pero los que vamos a utilizar se listan a continuación.