

Capitolo 1

Data Mining

1.1 Dataset

In nostro possesso è stato dato un dataset in formato *.csv* contenente 62.584 osservazioni e 13 colonne riguardanti il valore dei sensori di uno smartphone android e l'attività dell'utente.

Le etichette di questo dataset assumono 5 valori differenti:

- **STILL**: l'utente non si muove.
- **WALKING**: l'utente cammina.
- **CAR**: l'utente è in auto.
- **BUS**: l'utente è in bus.
- **TRAIN**: l'utente è in treno.

Le feature del dataset sono 12 e sono le seguenti:

- **accelerometer#mean**: la media delle osservazioni dell'accelerometro.
- **accelerometer#min**: l'osservazione minima dell'accelerometro.
- **accelerometer#max**: l'osservazione massima dell'accelerometro.
- **accelerometer#std**: la deviazione standard delle osservazioni dell'accelerometro.
- **gyroscope#mean**: la media delle osservazioni del giroscopio.
- **gyroscope#min**: l'osservazione minima del giroscopio.
- **gyroscope#max**: l'osservazione massima del giroscopio.

- **gyroscope#std**: la deviazione standard delle osservazioni del giroscopio.
- **gyroscopeuncalibrated#mean**: la media delle osservazioni del giroscopio non calibrato.
- **gyroscopeuncalibrated#min**: l'osservazione minima del giroscopio non calibrato.
- **gyroscopeuncalibrated#max**: l'osservazione massima del giroscopio non calibrato.
- **gyroscopeuncalibrated#std**: la deviazione standard delle osservazioni del giroscopio non calibrato.