

MachineLearning and DeepLearning courses project

MOTION SENSE DATASET PROPOSAL

MIRCO DE MARCHI

UNIVERSITY OF VERONA

28/11/20

Dataset Description

DATASET DESCRIPTION

Time-series data generated by accelerometer and gyroscope sensors, collected with an iPhone 6s kept in the participant's front pocket using SensingKit app which collects information.

A total of 24 participants in a range of gender, age, weight, and height performed 6 activities in 15 trials of different duration in the same environment and conditions: downstairs, upstairs, walking, jogging, sitting, and standing.

The project aim is to predict the activity carried out by a man with a smartphone in his pocket through sensing data and some subject information.

Reference: www.kaggle.com/motionsense-dataset;

FEATURES

Number of features: **18**;

Target feature name: **trial type**;

Participant related features

(subject id), weight, height, age, gender

Motion related features

attitude.roll, attitude.pitch, attitude.yaw, gravity.x, gravity.y, gravity.z, rotationRate.x, rotationRate.y, rotationRate.z, userAcceleration.x, userAcceleration.y, userAcceleration.z, is long, trial type

PREDICTION TECHNIQUES CHOSEN

Feature selection:

- PCA

Models:

- Bayes decision model (not expected great result)
- K-NN with different K and distance metrics
- Parzen Window with different kernel and window size
- SVM for multiple class
- Markov model on time series
- RNN, CNN, ...

Dataset Analysis

DIMENSIONS AND TARGET FEATURE

Dataset size: **1412865** entry;

Number of *NULL* values: **0**;

The following is the list of classes of *trial type* target feature and the related occurrence of sensing data:

Class	Occurrence in dataset
walking	344288
sitting	338778
standing	306427
upstairs	157285
jogging	134231
downstairs	131856

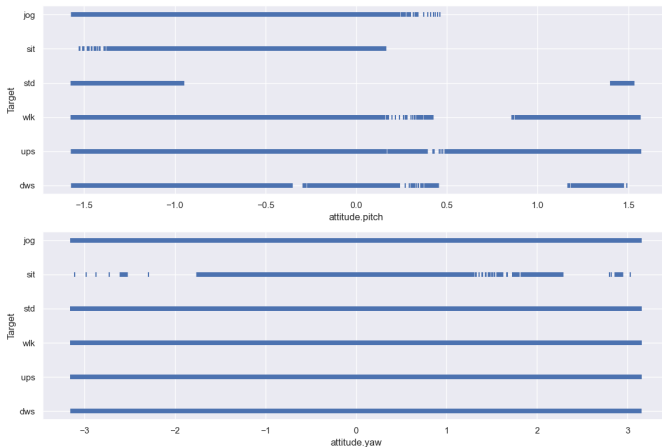
DATA COLLECTION PARTICIPANTS

The following is the number of data collected during the trials of each subject:

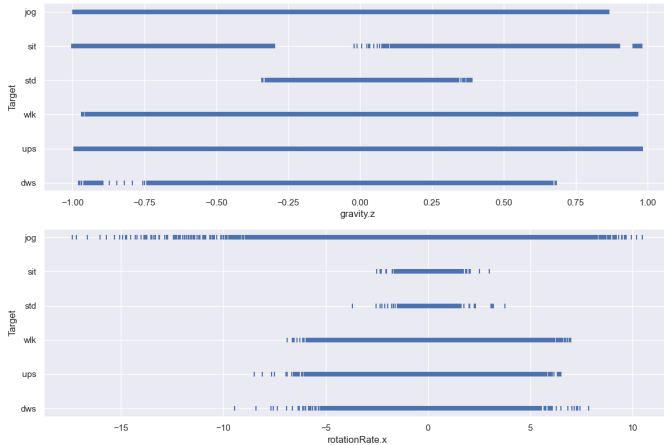
Subject	Occurrence
1	62312
2	62339
3	62553
4	56066
5	52283
6	57328
7	61276
8	60727
9	57136
10	59796
11	59292
12	53293

Subject	Occurrence
13	51147
14	56487
15	59443
16	65417
17	55539
18	62808
19	71949
20	55153
21	68932
22	56028
23	54556
24	51005

RELATION BETWEEN FEATURE AND TARGET 1



RELATION BETWEEN FEATURE AND TARGET 2



CORRELATION MATRIX

