Your Mobile Phone Can Recognize Your Physical Activity!

1 MOTIVATION

Being motivated in the S-Health application of my Samsung Note 4 mobile phone, I got interested to learn more about data processing of this healthcare application. One step further, I was eager to see do we can use gathered data by phones' sensors to predict our physical activities.

In this project, we observe that using collected data by accelerometer and gyroscope sensors in smart phones, one can recognize our activities such as walking or running with high accuracy.

S Health

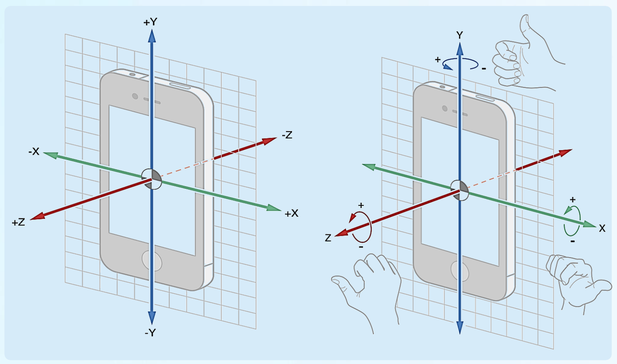
2 INTRODUCTION

As we observed during the course, in any type of data science analysis, the initial and essential part is data which based on the main target it could have various structures.

For our project, based on official documentation of the software and additionally by searching on the internet, we got that to recognize physical activity of a person one can use data of accelerometer and gyroscope sensors.

Accelerometer gives changes in phones velocity in three different axes including (X, Y, Z) which can model movement of a part of body.

A gyroscope measures rotation speed of the phone around a particular axis in three axes which can be used for recognizing quick or slow movements.



Left) Accelerometer Sensor **Right)** gyroscope Sensor.

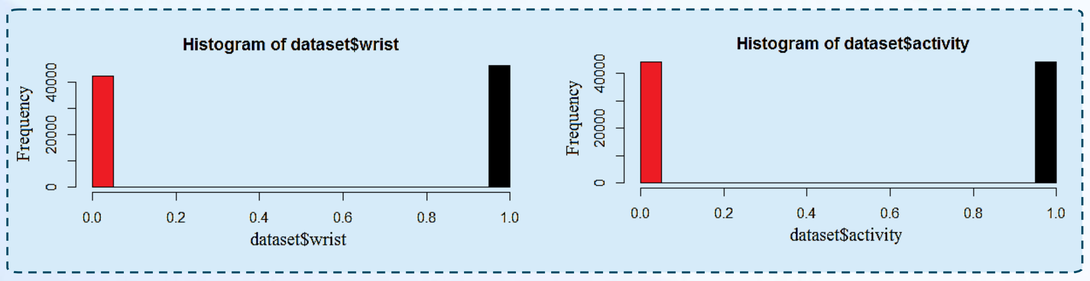
3 CHALLENGES with DATA

During the project initially we had some challenges with data:

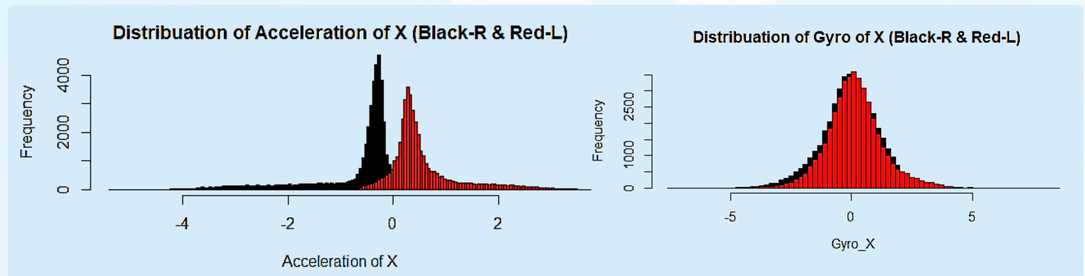
- What sort of data?
- Collect data (One is selected among several datasets from Kaggle)
- Exploring data and verifying data quality
- Finalize the dataset

1	date	time	username	wrist	activity	acceleration_x	acceleration_y	acceleration_z	gyro_x	gyro_y	gyro_z
2	########	13:51:15:847724020	viktor	1	0	0.265	-0.7814	-0.0076	-0.059	0.0325	-2.9296
3	########	13:51:16:246945023	viktor	1	0	0.6722	-1.1233	-0.2344	-0.1757	0.0208	0.1269
4	########	13:51:16:446233987	viktor	1	0	0.4399	-1.4817	0.0722	-0.9105	0.1063	-2.4367
5	########	13:51:16:646117985	viktor	1	0	0.3031	-0.8125	0.0888	0.1199	-0.4099	-2.9336
6	########	13:51:16:846738994	viktor	1	1	0.4814	-0.9312	0.0359	0.0527	0.4379	2.4922
7	########	13:51:17:46806991	viktor	0	1	0.4044	-0.8056	-0.0956	0.6925	-0.2179	2.575
8	########	13:51:17:246767997	viktor	0	1	0.632	-1.129	-0.2982	0.0548	-0.1896	0.4473
9	########	13:51:17:446569025	viktor	0	1	0.667	-1.3503	-0.088	-0.8094	-0.7938	-1.4348
10	########	13:51:17:646152973	viktor	0	1	0.2704	-0.8633	0.1293	-0.4173	-0.1904	-2.6759

Dataset: wrist: 0 (left wrist) & 1 (right wrist) activity: 0 (walking) & 1 (running) acceleration(x, y, z) gyro(x, y, z)



Left) Histogram of "wrist" to check that is binary **Right)** Histogram of "activity" to check that is binary



Numerical Data Distribution: Left) Acceleration_x **Right)** Gyro_y

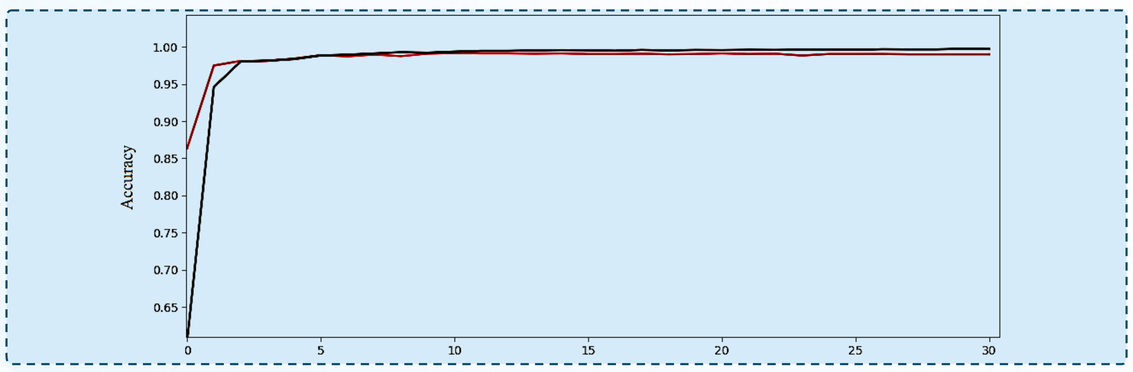
4 PROCEDURE

After finalizing the dataset, it was ready to use. The goal was to design and train a machine learning model that can predict whether the person walks or runs.

- **Framework and Model Selection:** Selecting a proper Machin learning framework was an important step which by searching on the internet finally I choose Keras that is designed with a focus for rapid experimenting.
- **Model Architecture:** input layer, after some investigations the dataset transformed into 6 separate ones, each one representing an axis for accelerometer or gyroscope and containing 12 sensor samples (~2 sec. obs. time).
- **Model Architecture:** hidden layer, 3 hidden layers produced at the end 99,2% accuracy.
- Data Size: the size of data which the network needs to perform accurately. By test and try: 12 samples, 99,23%.
- **Training data from one wrist:** It is observed that if one would have collected all training data from only one wrist, It turned out, the neural network is able to predict a correct activity only in 91% of all cases.

5 CLASSIFIER

- Configure model
- Assemble classifier and train it
- Perform 10-fold cross-validation on validation data
- Plot accuracy for train and validation data



Accuracy for train and validation data

6 CONCLUTION

- ✓ We got acquaintance with some of data mining analysis approaches behind healthcare applications, such as Samsung S-Health.
- ✓ Practical experience on dataset validity and quality check
- ✓ A classifier for recognizing physical activities (run & walk). The classifier can be extended and used in various applications (e.g. tracking,…) and environments.

Shahla Atapoor
Institute of Computer Science, University of Tartu
shahla.atapoor@ut.ee

Final Project of
Data Mining
Fall 2017

UNIVERSITY OF TARTU
1632
INSTITUTE OF COMPUTER SCIENCE