Smartphone-Based Recognition of Human Activities and Postural Transitions Data Set

My Solution and exploration

Sameh Muhammad Amin

6/Mar/2019

Project Describtion

Recognizing the human activity using smartphone signals captured from the Accelerometer and gyroscope sensors.

Experiment description

- The 61 experiments were carried out with a group of 30 volunteers.
- Activities composed of six basic activities
  - three static postures (standing, sitting, lying) and three dynamic activities (walking, walking downstairs and walking upstairs).
- The experiment also included postural transitions that occurred between the static postures. These are: stand-to-sit, sit-to-stand, sit-to-lie, lie-to-sit, stand-to-lie, and lie-to-stand.

Data Exploration

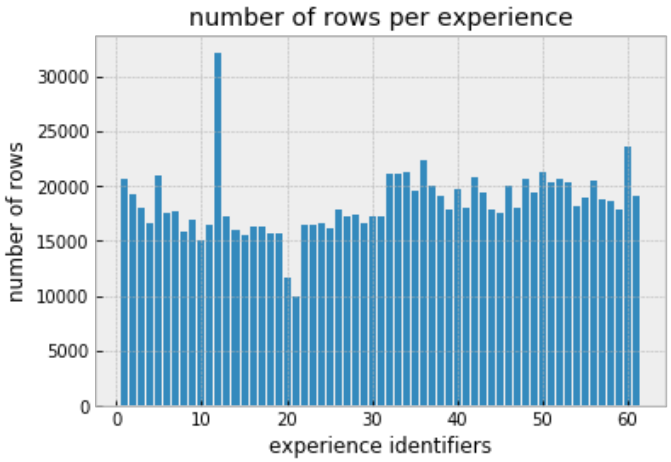
- 1214 experiments completed.

	experiment_number_ID	user_number_ID	activity_number_ID	Label_start_point	Label_end_point
count	1214.000000	1214.000000	1214.000000	1214.000000	1214.000000

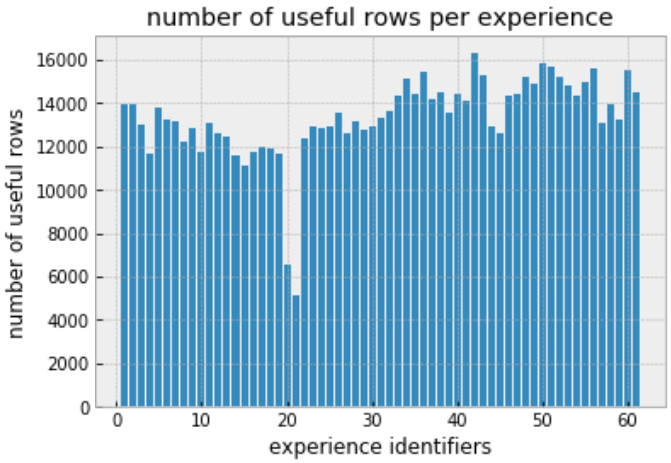
- Mean Activity signal start point is 7960.
- Mean Activity signal end point is 8630.

	experiment_number_ID	user_number_ID	activity_number_ID	Label_start_point	Label_end_point
count	1214.000000	1214.000000	1214.000000	1214.000000	1214.000000
mean	30.945634	15.390445	5.140857	7959.962932	8630.803130

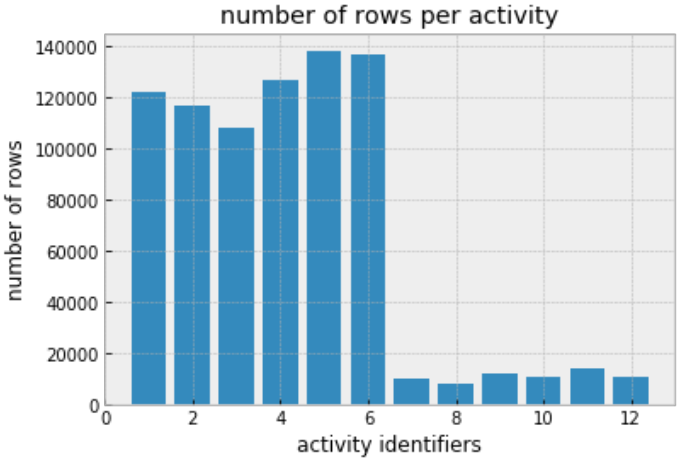
- Max number of signal readings was done on experiment #12



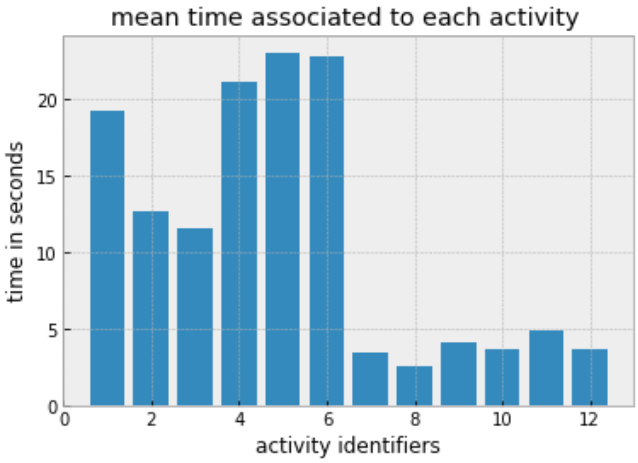
- Most of useful signal readings, which is related to experiment were related to experiments from 40 – 61,
- Most of unuseful signal readsins, were related to experiments from 18-25.



- Max number of signal readings per activity was related to **Standing** activity (activity label #5) - “Make sense!”



- Max Mean time per activity was related to activity #5, also **Standing** activity



Machine Learning Part

Data Splitting

I splitted the dataset to 80% for training and 20% for test.

Machine Learning Algorithms used

I tried to the following machine learning algorithms:

Algorithm	**Accuracy**
Logistic Regression with 10 folds cross validation	**0.94**
Linear Discriminant Analysis with 10 folds cross validation	**0.95**
Decision Tree	**0.86**
Random Forest	**0.88**
XGBoost	**0.92**
Neural Network (ANN)	**0.96**

Conclusion

- I got the best accuracy be implementing the Artificial nueral network algorithm with accuracy 96%.
- Best Linear classification algorithm accuracy was Linear discriminant analysis with cross validation, which i got 95% accuracy. With very good performance.
- I think i got those results because of the signals data nature, which becamse more to categorical than linear, because of the signal normalization implemented on the accelorometer and gyroscope signals to have more smooth and normalizaed signals. that’s why ANN and LDA were the best algorithms results.
- I found Decision Tree and Random Forest the lowest accuracy, and i think that becuase those algorithms fit more with categorical data, not linear data such as sensor signals.