

CSC8635 Machine Learning – Extended Technical Project

Issy Middleton C1000051

Project Overview

This is the summary of the project. For instructions on how to run the analysis see the ReadMe file.

Context

Smart phones can record motion sensor data, e.g., accelerometer and gyroscope, linear acceleration and angular velocity.

Objective

The purpose of the project is to determine whether smartphone activity data can be used to infer individual characteristics about weight, height, gender, age, and if so, to what degree of accuracy.

Method

Time-series data was generated by an experiment at Queen Mary University of London. An iPhone 6 was used by 24 participants to perform 6 activities in 15 trials, of varying length.

Linear and non-linear, deep and machine learning models were evaluated. These were Multi Layer Perceptron, Support Vector Machine Linear, Support Vector Machine - Radial Basis Function (RBF), Logistic Regression (L1), Logistic Regression (L2), Decision tree, Random Forest, K- Nearest Neighbour.

Results

The average accuracy rate of the Decision Tree model, at 99%, is worthy of note. Particularly as the model was trained and tested without feature engineering, on activity types indicating no movement, like sitting and standing. Even when the limitations of the dataset are considered, it is reasonable to conclude that characteristics of weight, height, age and gender can be predicted to a high degree of accuracy. Computational time is not linked to accuracy.

Novelty

This work supplements previous publications in the field by increasing the number of individual characteristics being predicted, extending the number of classification machine learning models utilised, varying the feature engineering applied then evaluating performance for accuracy, mean f1 score and time to compute.

PART 2B

Key Images

