Ensemble Approach for Sensor-Based Human Activity Recognition

Sunidhi Brajesh[1], Indraneel Ray[2]

[1]sb7253@nyu.edu, [2]ir944@nyu.edu

Computer Science, New York University

Abstract

This poster discussed our ensemble based approach to detect Human Activity for the Sussex-Huawei Locomotion Transportation(SHL) recognition challenge. The objective was to recognize 8 modes of transportation: *1-Still, 2- Walk, 3-Run, 4-Bike, 5-Car, 6-Bus, 7-Train, 8-Subway* in user and position independent manner. We trained Random Forest model on combined and shuffled SHL-Training and SHL-Validation data. To improve accuracy and prevent over-fitting, hyperparameter tuning was done with 10-fold cross validation. The model with best oob_score of 85.12% was selected as final one. The estimators were trained parallely to reduce training time using all logical cores.

Dataset and Features

- The SHL Recognition Challenge 2020 focused on identifying 8 modes of transportation - Still, Walk, Run, Bike, Car, Bus, Train, Subway using inertial sensor data of a smartphone. The dataset was acquired from three users wearing four smartphones at positions - Hips, Bag, Hand and Torso simultaneously.
- The Dataset [1][2] was divided into 3 parts - SHL- Training, SHL-Validation and SHL-Test dataset. The SHL-Training dataset contained labeled data over 59 days, SHL-Validation data over 6 days and SHL-Test data over 40 days. SHL-Training data was the largest, containing raw sensor data at four positions from User 1. The SHL-Validation contained data for User 2 and User 3 at four positions while, SHL-Test contained data for User 2 and User 3 at a position, which was kept unknown. The raw sensor data was sampled at frequency of 100Hz and included data from following sensors: acceleration (x, y and z), linear acceleration (x, y and z), magnetic field (x, y and z), gravity (x, y and z), gyroscope (x, y and z), Orientation (x, y, z and w) and Pressure. The data was segmented with a non-overlapping sliding window of 5 seconds and labels were provided per sample.
- Acceleration, Linear Acceleration, Magnetic field, Gyroscope and Pressure were used as raw features.

Methodology

- Missing values in every sample of 5 seconds replaced by mean in the frame
- To overcome the impact of position and orientation, magnitude for each sensor was calculated. $\sqrt{x^2+y^2+z^2}$
- SHL-Training and SHL-validation dataset was combined and shuffled then split into 70% training and 30% test data
- 65 statistical features derived in Time-domain and Frequency domain for each frame of 5 seconds
    - Time-domain: mean, median, max, min, variance, standard deviation, interquartile range
    - Frequency-domain: mean frequency, energy, entropy, kurtosis, skewness, peak frequency
- For each frame of 5 seconds, corresponding label determined using mode of labels provided in that frame
- Out of SVM, KNN, XGBoost and Random Forest, we found Random Forest to have the best combination of accuracy and training time, making it suitable for our analysis.
- Ensemble of Decision Trees-Random forest model was trained parallelly thus reducing training time.
- Output of each decision tree combined by averaging the probabilistic prediction instead of classifier vote for one class

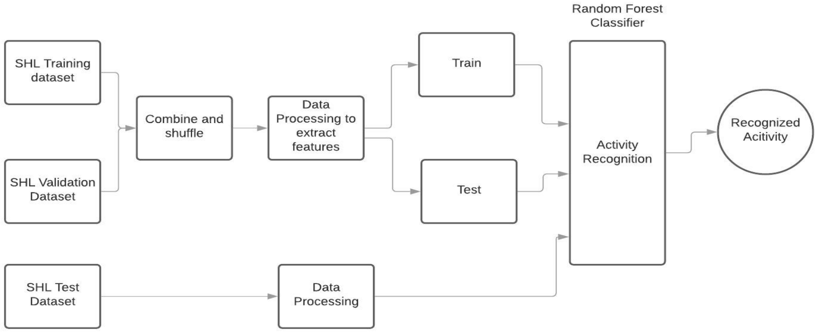

Fig 1: Structure of Activity Recognition System

- Randomized Grid Search used for hyperparameter tuning with 10-fold cross validation for 20 iterations. Final parameters:
    - Criterion : 'Gini', max_depth: 46 , n_estimators:125
- oob_score increased to 85.12% from 80.64% after hyperparameter tuning
- Computation methods: *2.3 GHz CPU i9 16GB RAM* Language: Python Library: Scikit learn 0.22.1
- Model Training Time : 5 minutes 3 seconds , HyperParameter Tuning time : 20 hours 30 minutes
- Final Model size : 278.3 MB

Conclusions

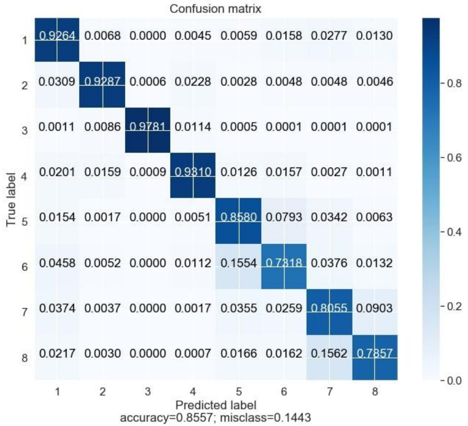

Fig 2: Normalized confusion matrix on test data : Still-1 Walk-2, Run- 3, Bike-4, Car-5, Bus-6, Train-7, Subway-8

- The F1-Score for the final model on Test dataset is: Still-88%, Walk-94%, Run-99%, Bike-94%, Car-84%, Bus-77%, Train- 80%, Subway-81%
- To overcome the issue of imbalanced dataset, weighted random forest model can be evaluated on this dataset. Furthermore, the window size has been considered of a fixed length of 5 seconds.It is possible that adjusting the window size to be of a shorter duration 2-3 seconds might yield better results.

References

[1] L.Wang, H. Gjoreski, M. Ciliberto, S. Mekki, S. Valentin, and D. Roggen. 2019. Enabling Reproducible Research in Sensor-Based Transportation Mode Recognition With the Sussex-Huawei Dataset. IEEE Access 7 (2019), 10870–10891. https://doi.org/10.1109/access.2019.2890793

[2] L. Wang, H. Gjoreski, M. Ciliberto, P. Lago, K. Murao, T. Okita, and D. Roggen. 2020. Summary of the Sussex-Huawei locomotion transportation recognition challenge 2020. Proceedings of the 2020 ACM International Joint Conference and 2020 International Symposium on Pervasive and Ubiquitous Computing and Wearable Computers (2020).