UPIC: User and Position-Independent Classical Approach for Locomotion and Transportation Modes Recognition

Md. Sadman Siraj
University of Dhaka
sadman.siraj@ieee.org

Md. Ahasan Atick Faisal
University of Dhaka
atickfaisal@gmail.com

Omar Shahid
University of Dhaka
omarshahid232@gmail.com

Farhan Fuad Abir
University of Dhaka
farhan.fuad@ieee.org

Tahera Hossain
Kyushu Institute of Technology
taheramoni@gmail.com

Sozo Inoue
Kyushu Institute of Technology
sozo@brain.kyutech.ac.jp

M.A R. Ahad
University of Dhaka
atiqahad@du.ac.bd

ABSTRACT

The Sussex-Huawei Locomotion-Transportation (SHL) Challenge 2020 was an open competition of recognizing eight different activities that had been performed by three individual users and participants of this competition were tasked to classify these eight different activities with modes of locomotion and transportation. This year's data was recorded with a smartphone which was located in four different body positions. The primary challenge was to make a user-invariant as well as position-invariant classification model. The train set consisted of data from only user-1 with all positions whereas the test set consisted of data from user 2 and 3 with unspeicified sensor position. Moreover, a small validation with the same charecteristics of the test set was given to validate the classifier. In this paper, we have described our (Team Red Circle) approach in which we have used previous year's challenge data as well as this year's provided data to make our training dataset and validation set that have helped us to make our model generative. In our approach, we have extracted various types of features to make our model user independent and position invariant, we have applied Random Forest classifier which is a classical machine learning algorithm and achieved **92.69%** accuracy on our customized train set and **77.04%** accuracy on our customized validation set.

CCS CONCEPTS

• **Computing method** → **Feature extraction**; • **Learning system** → *Supervised*; • **Algorithm** → Random Forest.

KEYWORDS

Feature extraction, feature selection, classical approach, user invariant, position independent, classifier, SHL recognition challenge

UbiComp/ISWC '20 Adjunct, September 12-16, 2020,

ACM ISBN 978-1-4503-XXXX-X/18/06…$15.00
https://doi.org/10.1145/1122445.1122456

ACM Reference Format:
Md. Sadman Siraj, Md. Ahasan Atick Faisal, Omar Shahid, Farhan Fuad Abir, Tahera Hossain, Sozo Inoue, and M.A R. Ahad. 2020. UPIC: User and Position-Independent Classical Approach for Locomotion and Transportation Modes Recognition. In *Adjunct Proceedings of the 2019 ACM International Joint Conference on Pervasive and Ubiquitous Computing and the 2020 International Symposium on Wearable Computers (Ubi- Comp/ISWC '20 Adjunct), September 13-16, 2020.* ACM, New York, NY, USA, 6 pages. https://doi.org/10.1145/1122445.1122456

1 INTRODUCTION

The modern technological revolution in smartphones allow us to collect robust sensor data with greater accuracy, which can be utilized in a comprehensive analysis about the user's context and locomotion. State-of-the-art researches on human activity recognition mostly focus on walking, running, cycling, upstairs, downstairs and other daily activities [1, 2]. However, these robust smartphone data should not be constrained within the analysis of casual daily activities. To provide better personalized service, user's locomotion data in transportation is very important. The user's data in transportation mode is a significant contextual source of information which opens up opportunities for adaptive services like traffic route monitoring, parking spot detection, route or parking suggestion, proactive recommendation about transportation timetable, make faster content delivery. Besides, research on transportation modes render a great service in road condition analysis, providing probabilistic mobility and creation of locomotion mode, designing novel techniques related to localization and so on [9].

The typical approach to analyze activity based on wearable devices like smartphone or wrist watch is by applying supervised learning to inertial sensor data [3, 6]. However, an important, but most often neglected fact is that activity classification models are generally location-dependent [21], in other words, prediction model that was trained in a particular body location (e.g., the wrist) to predict the user's activity may perform quite poorly when they are evaluated in another body position (e.g., a pocket on the hip) [5]. The issue can be more crucial with smartphones, since they can be carried by the user in many different body locations. This scenario can be compensated by location-specific models [10, 17], but to train such location specific model, it is necessary to make the most of cross-location data [14].

The Sussex-Huawei Locomotion-Transportation (SHL) Recognition Challenge 2020 [12, 20] dealt with almost the same types of problems. This challenge provided us with the dataset to recognize eight modes of locomotion and transportation activities using the inertial sensor data (i.e., accelerometer) of a smartphone. The ultimate goal of this challenge was to recognize the user's activity from the data collected from a smartphone placed on an unknown position of a different user as the provided train set is consisted of user-1 data with all four positions (hand, torso, hips, bag), but test set is consisted of data from an unknown body location of user-2 and user-3.

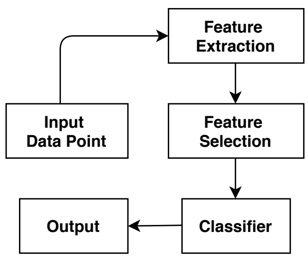

Figure 1: General description of our method

In this challenge, participants are allowed to use previous years' challenge data [18]. Therefore, we have tried to find correlation between dataset to figure out the user and position's of different sets (i.e., validation 2019, test 2019). In our approach, we have applied a very basic method (showed in Figure 1) for classification which consists of feature extraction, feature selection and classifier to predict output. We have used the combination of Validation 2020, Validation 2019, Test 2019 as train set, and Train 2020 as our validation set. Moreover, we have applied smart feature extraction algorithm with different feature selection algorithms along with different feature selection techniques to find out the optimum features for user and position invariant.

In rest of the sections of the paper, we have described our approach as follows: in the section 2 we have described this year's challenge dataset with their position, user and instances information based on different classes. In section 3 we have discussed about our method and motivation behind it briefly. Data visualization, analysis, feature extraction along with feature selection steps of experimentation have been discussed in section 4. Evaluation of our result on individual positions have been described in section 5. Section 6 have addressed strong side of our approach as well as some limitations. Finally, we have drawn the conclusion in section 7.

2 DATASET DESCRIPTION

The main purpose of the SHL challenge was to make a prediction model to recognize eight modes of locomotion and transportation – i) Walk, ii) Run, iii) Bike, iv) Still, v) Car, vi) Bus, vii) Train, viii) Subway – using data from the inertial sensor of a smartphone. The dataset provided by the challenge organizer originally recorded with four smartphone that had been worn at four different body location (hand, hips, torso, bag).

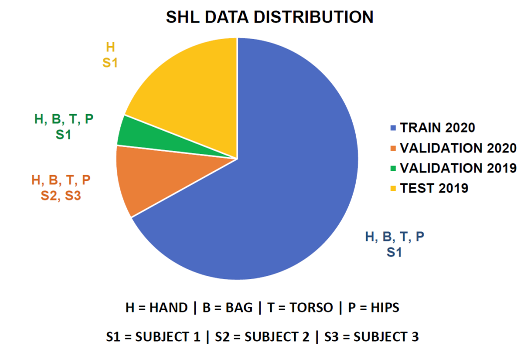

Figure 2: Distribution of dataset according to sample number, sensor location and subject

Challenge organizer provided data in three different sets this year. The SHL-Train Set 2020 contained data from all four positions of user-1 only. SHL Validation Set 2020 provided a small set compared to train set which have data from all positions of user-2 and 3. SHL Test Set 2020 contained data of user-2 and 3 from an unknown position. On another note, organizer's gave us permission to use previous year's challenge data for our training and validation. SHL Validation 2019 contained data from all four positions of user-1 and SHL Test 2019 provided data of only the hand position of user-1. In Figure 2, we have reviewed the dataset distribution with respect to all positions and subjects. Here, one thing is notable that we are using prefix 'SHL' before mentioning dataset as we have not used exactly same train set for training and validation set to validate. We have made individual customized set for training, validation and testing that we have described in later sections.

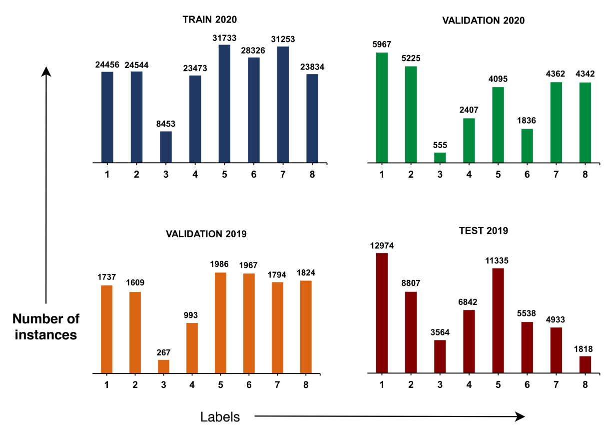

Figure 3: Class-wise data distribution

In summary, the challenge dataset comprised of 59 days of training data, 6 days of validation data and 40 days of test data. Train, test and validation - all three sets of data were generated with a

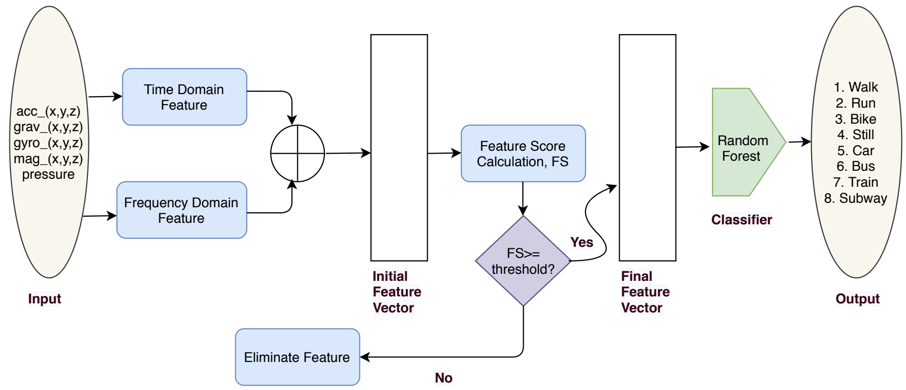

Figure 4: Block diagram of our proposed model

non-overlap sliding window of 5 seconds. SHL Train 2020 consists of 196072 frames, SHL Validation 2020 contains 28789 frames, SHL Test 2020 contains 57573 frames where each frame containing 500 samples (5 seconds at the sampling rate 100 Hz) for all sets. Each of the dataset has eight labelled activities. In Figure 3, we have demonstrated the distribution of dataset in accordance with the class-wise instances.

3 PROPOSED METHOD

As we have mentioned before, this challenge focused on recognising modes of transportation in a user-independent manner with an unknown phone position. In our approach, firstly, we have tried to figure out the position of SHL Test 2020 by creating a position classifier. In this case, we have used only selected features, which varies with different positions. Our position classifier predicted most of the samples as hand data. Therefore, we have given more emphasize on hand data when we build our customized train and validation set.

Figure 4 demonstrates our whole method in a block diagram. In our method, we have extracted both time domain and frequency domain features. At first, we have aggregated all features to make our initial feature set with 789 features. After that we have applied feature selection method to get importance score for each feature to determine which feature is important. We have taken only those features which are equal or above a certain threshold in terms of feature score which give us 349 features. We have put those 349 features in a random forest classifier to make output prediction.

4 EXPERIMENT

We have done our experiment as following sequence:

- Position recognition
- Data preparation
- Feature extraction
- Feature selection
- Classification

Position recognition: One of the main challenges of SHL Dataset Challenge 2020 is that the location or position of the sensors of the provided test set is unknown. Hence, our initial notion was to build a position recognition model to predict the position of the test set sensors. Usually the orientation of the phone is different in different positions. The acceleration pattern is also position dependent. Features for position classification need to capture the orientation information. So, the rotation matrix was not applied before extracting features. Only 19 features were used for position classification. These features include-

- Min, max, correlation coef., and average of gravity X and Z axis
- Min, max, average of accelerometer X, Y and Z axis
- Min, average, correlation coef. of pitch angle
- Linear velocity along X and Y axis

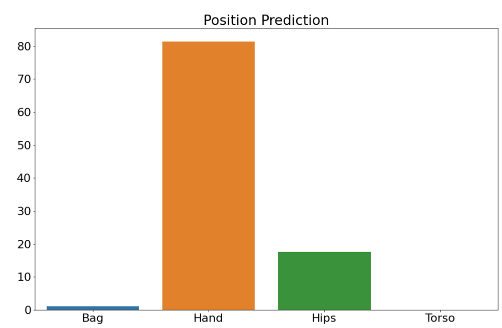

Figure 5: Prediction result of position recognition on test data 2020

We then put these generated features to the random forest classifier to predict position of SHL Test data 2020. Figure 5 is showing our predicted distribution where we can see most of the instances have been classified as hand data.

Data preparation: From Figure 2, we can see that Validation 2020 contained data of user-2 and 3. By applying our position classifier, we have found that our test data is recorded on hand position. Therefore, we have made our train set by combining VALIDATION2020, VALIDATION2019 and TEST2019. In this case, we have included VALIDATION2019 as it includes user-1 data, therefore, our generated model would learn features from user-1 data as well. And we have taken TRAIN2020 as our validation set for

Table 1: Data preparation

Train Set	Validation Set
VALIDATION2020 VALIDATION2019 TEST2019	TRAIN2020

our prediction model. TRAIN2020 contains data of all positions of use-1, so validating our model on user-1 will make our model more generative. Table 1 represents our data customized train and validation set.

Feature extraction: Firstly, we have taken the accelerometer and magnetometer sensors data from all provided body location (hand, bag, torso, hips) and de-rotated the data to earth axis using the rotation matrix calculated from orientation sensor values [4]. It ensures our data to be position invariant. Afterwards, we have generated 12 more data channels as accelerometer magnitude, linear accelerometer magnitude, gravity magnitude, gyroscope magnitude, magnetometer magnitude, vertical acceleration, horizontal acceleration, total acceleration jerk, body acceleration jerk, vertical acceleration jerk, horizontal acceleration jerk and pressure derivative. Then, we have used the provided tri-axial (x, y, z) data of sensor channels of SHL Dataset 2020, accelerometer, linear accelerometer, gravity, gyroscope, magnetometer and pressure data along with the data of the 12 generated data channels to extract time domain and frequency domain features which lead to a total of 789 features. Table 2 provides a complete list of features that we have used in our experiment.

Feature Selection: All extracted features might not be useful for classification, some may have a negative effect on classification performance. In other words, feeding those features to the classifier can decrease the accuracy. Furthermore, there exist some redundant features that have no impact on performance. A feature selection framework is necessary to eliminate these negative and redundant features effectively. We have used six different feature scoring technique to calculate the importance of features.

- **Mutual Information [8]:** This metric provide the information that if two features are independent or not. A higher value denotes higher dependency to other one.
- **Chi-square(χ^2) Test [15]:** It measures the χ^2 statistics between two non-negative features which helps to eliminate irrelevant features.
- **Tree-based Selection [13]:** This approach uses Decision Tree-based classifiers to calculate scores. The higher value indicates the higher importance.
- **Pearson Correlation Coefficient [7]:** This approach eliminates the highly correlated features or redundant features by calculating correlation coefficient and p-value.
- **Spearman Correlation Coefficient [16]:** It measures the monotonicity between two features and spearman correlation does not consider the feature vector to be normally distributed.
- **ANOVA F-value [11]:** An ANOVA test is performed on the dataset to calculate the f-value and p-value for the features and importance values are calculated based on that.

Table 2: Feature List

Channels	Time Domain Features	Frequency Domain Feature
Acceleration (x, y, z, mag), Linear Acceleration (x, y, z, mag), Gravity (x, y, z, mag), Magnetometer (x, y, z, mag), Gyroscope (x, y, z, mag), Vertical Acceleration, Horizontal Acceleration, Jerk (Total, Body, Horizontal and Vertical), Pressure, Derivative of pressure	Min, Max, Peak to Peak Range, Average, Standard Deviation, Variance, Max Rate of Change, Average Rate of Change, Mean Absolute Deviation, Interquartile Range, Correlation Coefficient, Mean Crossing Rate, Mutual Correlation (X-y), Covariance (X-y), Signal Magnitude Area, Root Mean Square, Energy, Linear Velocity	Max Spectral Power, Center Frequency, Dominant Frequency, Entropy, Spectral Energy, Skewness, Kurtosis, Number of Peaks, First 10 FFT Coefficients

Table 3: Results for different positions

Position	Training Accuracy	Validation Accuracy
Bag		77.81%
Hand		77.57%
Hips	92.69%	74.91%
Torso		77.85%
All		77.04%

After calculating the feature score, we have taken the average of six scores for individual feature and sort them according to the value. We set a certain threshold and find that 349 features are equal or above that threshold score. Finally, we have taken that 349 features as final feature vector.

Classifier: We have applied Random Forest (RF) classifier on selected 349 feature vector.

5 RESULTS:

From the Figure 5 it can be seen that the model predicted the device position of the test data to be of hand position. So, initially we have decided to train a classifier using data from sensors of hand position only. The main reason behind this approach was that a model trained only on data from hand position would be most likely to achieve best results on the provided test set which was also predicted to be from hand position. In this case, we have got 76.89% accuracy on our validation set.

However, there remains a certain degree of uncertainty about this position recognition approach and so, we decided to have a more generalized approach of building a position independent model that would be able to take data from any sensor position and provide comparable results on the provided test set.

Table 3 shows overall results of our method on different positions in terms of training and validation accuracy. Furthermore, Figure 6-9 shows the confusion matrices of our result on four individual

Figure 6: Confusion matrix for bag

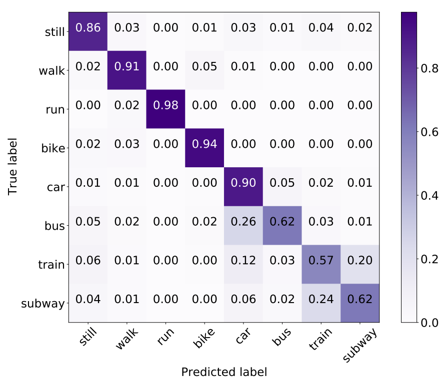

Figure 7: Confusion matrix for hand

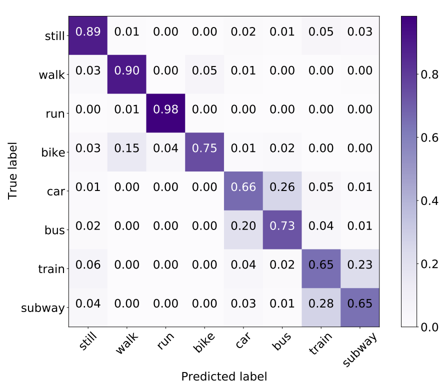

Figure 8: Confusion matrix for hips

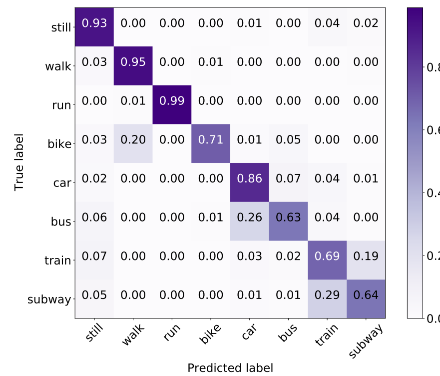

Figure 9: Confusion matrix for torso

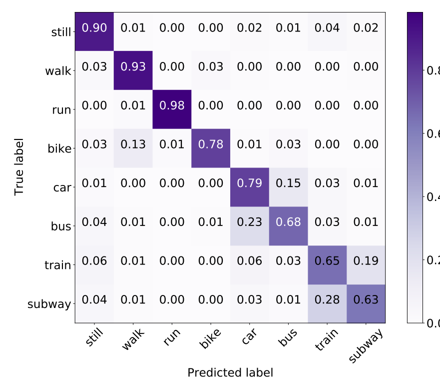

Figure 10: Confusion matrix for all position

positions. And finally, Figure 10 represents our result on all positions of our validation set. We have reported this result in our final submission.

6 DISCUSSION

This year's challenge demanded the approach to be independent of the sensor position or location and subject or user invariant. So, the core idea behind our approach was to train and validate the model in such a way so that it can generate satisfactory results on data across all the sensor positions and simultaneously for different users or subjects of the dataset. Consequently, this approach led us to train our model on data from all sensor positions and from all available subjects. In other words, our model is able to learn the 8 classes of the dataset from features extracted from the data of sensors irrespective of the sensor position and the subject to whom the sensor is attached to. This is why we believe our model is able to provide good results on the provided test set as it's sensors' position is unknown and also the data can be of either subject 2 or subject 3. On another note, in section 4, we have mentioned the position recognition approach. Through this approach, we have trained the model by having a train set of data only from hand position and get the validation accuracy of 76.89%. It should be mentioned here that our proposed approach of a position independent model trained on data from all sensor positions achieved better validation accuracy on our customized validation data of only hand position, 77.57%. So, our proposed approach provides a better result compared to the approach of using only data of hand position to train the model.

One fundamental limitation of our approach is that the model becomes very large as it is trained on data from all sensor positions and from all subjects. The training time also becomes sufficiently long with such a large amount of data. In order to solve this problem, we chose to build our own train set using validation data and test data of 2019 instead of using the train data of 2019. The sample size of validation data and test data of 2019 combinedly is around 2.88 times smaller than the sample size of train data of 2020. In this way, the model size and training time are brought down to a considerable range. We also performed feature selection to remove redundant features and features which are highly uncorrelated for the different sensor positions and subjects. Thereby, the model size and training time was further reduced.

7 CONCLUSION

This year's challenge was quiet interesting to work as it aims to solve user dependency as well as position dependency problem. We have tried to generalize our method as much we could and developed a simple learning method which provides better accuracy. We have applied a position recognition classifier on test set to figure out the device position, so that we can cross check model prediction with specific position of test data that how much our model generalized. Our feature selection technique makes our method robust for sensor based activity analysis, as we have applied six different approaches of feature selection, which makes our feature vector more problem specific and accurate. If all our intuitions are true, then our model will perform much better than our validation performance and will achieve better accuracy on test set. The recognition result for the testing dataset will be presented in the summary paper of the challenge [19].

REFERENCES

[1] MAR Ahad, AD Antar, and M Ahmed. 2019. IoT Sensor-Based Activity Recognition-Human Activity Recognition.

[2] Md Atiqur Rahman Ahad. 2011. *Computer vision and action recognition: A guide for image processing and computer vision community for action understanding.* Vol. 5. Springer Science & Business Media.

[3] Md Atiqur Rahman Ahad. 2012. *Motion history images for action recognition and understanding.* Springer Science & Business Media.

[4] Masud Ahmed, Anindya Das Antar, Tahera Hossain, Sozo Inoue, and Md Atiqur Rahman Ahad. 2019. POIDEN: position and orientation independent deep ensemble network for the classification of locomotion and transportation modes. In *Adjunct Proceedings of the 2019 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2019 ACM International Symposium on Wearable Computers.* 674–679.

[5] Anindya Das Antar, Masud Ahmed, and Md Atiqur Rahman Ahad. 2019. Challenges in Sensor-based Human Activity Recognition and a Comparative Analysis of Benchmark Datasets: A Review. In *2019 Joint 8th International Conference on Informatics, Electronics & Vision (ICIEV) and 2019 3rd International Conference on Imaging, Vision & Pattern Recognition (icIVPR).* IEEE, 134–139.

[6] Anindya Das Antar, Masud Ahmed, Mohammad Shadman Ishrak, and Md Atiqur Rahman Ahad. 2018. A comparative approach to classification of locomotion and transportation modes using smartphone sensor data. In *Proceedings of the 2018 ACM International Joint Conference and 2018 International Symposium on Pervasive and Ubiquitous Computing and Wearable Computers.* 1497–1502.

[7] Jacob Benesty, Jingdong Chen, Yiteng Huang, and Israel Cohen. 2009. Pearson correlation coefficient. In *Noise reduction in speech processing*. Springer, 1–4.

[8] Alexander Bulinski and Denis Dimitrov. [n.d.]. Statistical estimation of the Shannon entropy. ([n. d.]). arXiv:math.ST/http://arxiv.org/abs/1801.02050v1

[9] Ricardo Chavarriaga, Hesam Sagha, Alberto Calatroni, Sundara Tejaswi Digumarti, Gerhard Tröster, José del R Millán, and Daniel Roggen. 2013. The Opportunity challenge: A benchmark database for on-body sensor-based activity recognition. *Pattern Recognition Letters* 34, 15 (2013), 2033–2042.

[10] Božidara Cvetković, Robert Szeklicki, Vito Janko, Przemyslaw Lutomski, and Mitja Luštrek. 2018. Real-time activity monitoring with a wristband and a smartphone. *Information Fusion* 43 (2018), 77–93.

[11] Nadir Omer Fadl Elssied, Othman Ibrahim, and Ahmed Hamza Osman. 2014. A Novel Feature Selection Based on One-Way ANOVA F-Test for E-Mail Spam Classification. *Research Journal of Applied Sciences, Engineering and Technology* 7, 3 (jan 2014), 625–638. https://doi.org/10.19026/rjaset.7.299

[12] Hristijan Gjoreski, Mathias Ciliberto, Lin Wang, Francisco Javier Ordonez Morales, Sami Mekki, Stefan Valentin, and Daniel Roggen. 2018. The university of sussex-huawei locomotion and transportation dataset for multimodal analytics with mobile devices. *IEEE Access* 6 (2018), 42592–42604.

[13] K. Grabczewski and N. Jankowski. 2005. Feature selection with decision tree criterion. In *Fifth International Conference on Hybrid Intelligent Systems (HIS'05).* IEEE. https://doi.org/10.1109/ichis.2005.43

[14] Vito Janko, Martin Gjoreski, Carlo Maria De Masi, Nina Reščič, Mitja Luštrek, and Matjaž Gams. 2019. Cross-location transfer learning for the sussex-huawei locomotion recognition challenge. In *Adjunct Proceedings of the 2019 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2019 ACM International Symposium on Wearable Computers.* 730–735.

[15] Xin Jin, Anbang Xu, Rongfang Bie, and Ping Guo. 2006. Machine Learning Techniques and Chi-Square Feature Selection for Cancer Classification Using SAGE Gene Expression Profiles. In *Lecture Notes in Computer Science.* Springer Berlin Heidelberg, 106–115. https://doi.org/10.1007/11691730_11

[16] Leann Myers and Maria J Sirois. 2004. S pearman Correlation Coefficients, Differences between. *Encyclopedia of statistical sciences* (2004).

[17] Swapnil Sayan Saha, Shafizur Rahman, Zarin Rezwana Ridita Haque, Tahera Hossain, Sozo Inoue, and Md Atiqur Rahman Ahad. 2019. Position independent activity recognition using shallow neural architecture and empirical modeling. In *Adjunct Proceedings of the 2019 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2019 ACM International Symposium on Wearable Computers.* 808–813.

[18] Lin Wang, Hristijan Gjoreski, Mathias Ciliberto, Paula Lago, Kazuya Murao, Tsuyoshi Okita, and Daniel Roggen. 2019. Summary of the Sussex-Huawei locomotion-transportation recognition challenge 2019. In *Adjunct Proceedings of the 2019 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2019 ACM International Symposium on Wearable Computers.* 849–856.

[19] Lin Wang, Hristijan Gjoreski, Mathias Ciliberto, Paula Lago, Kazuya Murao, Tsuyoshi Okita, and Daniel Roggen. 2020. Summary of the Sussex-Huawei locomotion-transportation recognition challenge 2020. In *Proceedings of the 2020 ACM International Joint Conference and 2020 International Symposium on Pervasive and Ubiquitous Computing and Wearable Computers, 2020.*

[20] Lin Wang, Hristijan Gjoreski, Mathias Ciliberto, Sami Mekki, Stefan Valentin, and Daniel Roggen. 2019. Enabling reproducible research in sensor-based transportation mode recognition with the Sussex-Huawei dataset. *IEEE Access* 7 (2019), 10870–10891.

[21] Peter Widhalm, Maximilian Leodolter, and Norbert Brändle. 2019. Ensemble-based domain adaptation for transport mode recognition with mobile sensors. In *Adjunct Proceedings of the 2019 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2019 ACM International Symposium on Wearable Computers.* 857–861.

A APPENDIX

Programming Language and Libraries

Programming language: Python

Libraries: Numpy, Pandas, Matplotlib, Scikit-learn, Sci-py

Machine Specification

- RAM: 13 GB
- CPU: Intel Xeon (2) @ 2.3GHz
- GPU: N/A

Training and testing time

Training: 58.8 minutes

Testing: 40.9 seconds