Human Activity Recognition On Smartphones Using Multi-Class Classifiers

Fatmanur Yaman and Fatih Yusuf Çıra
fatmanur.yaman@boun.edu.tr, fatih.cira@boun.edu.tr

Abstract. Human activity recognition on smartphones using multi-class classifiers is a challenging research area that has gained significant attention in recent years. Accurately classifying and predicting various types of human activities has numerous applications, including personalized recommendations, user experience improvement, and better understanding of human behavior. This study explores the effectiveness of different machine learning algorithms for multi-class classification in the context of human activity recognition. It introduces and applies classifiers such as Naive Bayes, SVMs, KNNs, XGBoost, and Neural Networks to a comprehensive dataset collected from smartphones equipped with inertial sensors. The results demonstrate the superior performance of the XGBoost algorithm, with high accuracy and F1 scores, in accurately classifying and predicting human activities. The hyperparameter selection is built on XGBoost to provide a better model. Furthermore, explainable AI techniques are employed to identify discriminative features contributing to activity recognition. The findings of this research provide valuable insights into selecting appropriate classifiers and enhancing the accuracy and reliability of human activity recognition on smartphones.

Keywords: Human Activity Recognition, XGBoost, Smartphones, Explainable AI.

1 Introduction

The collection of human activity data through smartphones has become increasingly prevalent in recent years. This data, when properly classified, holds significant importance as it provides valuable insights into the actions and behaviors of individuals within specific time ranges. Such information not only sheds light on the users' activities but also offers clues about their personality traits, which has great significance in our modern world where accurate and successful recommendations play a crucial role. Consequently, the problem of human activity recognition has gained substantial attention and is considered one of the prominent challenges in the field.

2 Related Work

One notable study that addresses this challenge is the research titled "Human Activity Recognition on Smartphones using Multiclass Hardware-Friendly Support Vector Machines." [1]In this paper, the authors propose the utilization of support vector machines (SVMs) as a means to classify and predict various types of human activities. SVMs are popular machine learning algorithms that excel in handling multi-class prediction tasks. The study introduces and applies different types of SVMs to tackle the human activity recognition problem on smartphones.

[1] Davide Anguita, Alessandro Ghio, Luca Oneto, Xavier Parra and Jorge L. Reyes-Ortiz. Human Activity Recognition on Smartphones using a Multiclass Hardware-Friendly Support Vector Machine. International Workshop of Ambient Assisted Living (IWAAL 2012). Vitoria-Gasteiz, Spain. Dec 2012

In addition to SVMs, previous works have explored a range of machine learning methods for multi-class classification problems. Some notable approaches include Naive Bayes, which is based on probabilistic reasoning and often used for its simplicity and efficiency. Another widely used method is Threshold-based classification, where decision thresholds are set to distinguish between different activity classes. Markov Chains have also been employed in this context, leveraging the sequential nature of human activities to make predictions.

Recognizing the importance of exploring diverse classifiers, our study initially incorporates various types of classifiers to tackle the multi-class classification problem in human activity recognition. These classifiers are carefully chosen to leverage their individual strengths and capabilities. Subsequently, we introduce the widely acclaimed XGBoost algorithm, which has gained significant popularity in recent years due to its exceptional performance in various domains. Furthermore, we employ a rigorous process for hyperparameter selection to ensure optimal model configuration and maximize prediction accuracy.

By employing a comprehensive range of classifiers and selecting the most suitable one, our research aims to enhance the accuracy and reliability of human activity recognition on smartphones. The findings of this study have the potential to contribute significantly to the development of personalized recommendations, as well as improving user experience and facilitating better understanding of human behavior in different contexts.

In conclusion, human activity recognition on smartphones using multi-class classifiers is an area of active research. Various machine learning methods, including SVMs, Naive Bayes, Threshold-based classification, and Markov Chains, have been explored in previous works. Our study extends the research in this field by utilizing a combination of classifiers and introducing the XGBoost algorithm with careful hyperparameter selection. The outcomes of this research hold promise for enabling more accurate and successful recommendations based on users' activities, ultimately enhancing the overall user experience and understanding of human behavior.

3 Methodology

3.1 Dataset Description & Experimental Setup

The Human Activity Recognition dataset is a valuable collection created from recordings of 30 participants engaging in their daily activities while wearing a waist-mounted smartphone embedded with inertial sensors. The primary goal of this dataset is to accurately classify the performed activities into six predefined categories.

The experiments involved a diverse group of 30 volunteers aged between 19 and 48 years. Each participant performed six distinct activities, including WALKING, WALKING_UPSTAIRS, WALKING_DOWNSTAIRS, SITTING, STANDING, and LAYING. Throughout the experiments, a Samsung Galaxy S II smartphone was securely attached to their waist, equipped with an embedded accelerometer and gyroscope. These sensors allowed for the capture of data on 3-axial linear acceleration and 3-axial angular velocity at a constant rate of 50Hz. To facilitate accurate labeling, the experiments were also recorded on video.

To ensure the dataset's reliability, the collected data was randomly split into two sets: a training set and a test set. Around 70% of the volunteers were included in the training data, while the remaining 30% were assigned to the test data. This division enables the models developed using the training data to be effectively evaluated using the separate test data.

To enhance the quality of the sensor signals, pre-processing techniques were employed. Noise filters were applied to eliminate unwanted noise from the accelerometer and gyroscope readings. The signals were then sampled using fixed-width sliding windows of 2.56 seconds with a 50% overlap, resulting in 128 readings per window. Additionally, a Butterworth low-pass filter with a cutoff frequency of 0.3 Hz was utilized to separate the accelerometer signal into two components: total acceleration and estimated body acceleration. This process isolated the low-frequency components associated with gravitational force.

From each window of sensor data, a comprehensive feature vector was derived by calculating various variables in the time and frequency domains. This feature vector consists of 561 attributes that provide informative insights into the observed activities. Each record in the dataset includes an activity label, indicating the specific activity performed, as well as an identifier for the participant who conducted the experiment.

The availability of this dataset, along with its attribute information, presents numerous opportunities for developing robust activity recognition models and conducting further analysis. Researchers and practitioners can explore the triaxial acceleration and angular velocity data, as well as the extensive feature vector, to devise innovative approaches and evaluate their effectiveness in human activity recognition. This dataset serves as a valuable resource for advancing the understanding and application of activity recognition across various domains, including healthcare, fitness tracking, and human-computer interaction.

3.2 ML Methods Used

The forecasts are made using various machine learning (ML) models, with the goal of finding the model that provides the best predictions. Initially, the models are employed with their default values to determine the top-performing predictor. Once the best model is identified, hyperparameter tuning is conducted to further enhance the results. Before training the model, the features were selected by using Pearson Correlation Coefficient with threshold 0.1 value.

- **Naive Bayes Classifier**

One of the ML models employed in this analysis is the Naive Bayes Classifier, a probabilistic algorithm based on Bayes' theorem. It makes predictions by leveraging the probability of different categories and operates under the assumption of conditional independence among predictors given the target variable. This assumption allows the Naive Bayes Classifier to perform well when the independence assumption holds reasonably well. It is particularly suited for multi-class prediction tasks where the features are considered independent. In the context of this dataset, the Naive Bayes Classifier achieves an accuracy of 0.57 and an accuracy of 0.53. By leveraging probability and the assumption of independence, the Naive Bayes Classifier provides valuable insights for classifying activities in this study.

- **Support Vector Machine**

Furthermore, the analysis employed Support Vector Machine (SVM), a powerful algorithm designed to handle complex datasets by creating margin boundaries between classes. SVM strives to find optimal hyperplanes that effectively separate different classes within the dataset. Its objective is to maximize the margin between classes, ensuring clear and distinct boundaries. One of the key advantages of SVM is its ability to handle both linear and nonlinear separable datasets through the utilization of the kernel trick. This technique allows SVM to transform the input space into a higher-dimensional feature space, where nonlinear patterns can be linearly separated. By leveraging the kernel trick, SVM demonstrates remarkable versatility in handling datasets with intricate and nonlinear relationships. This capability makes SVM a valuable tool for accurately classifying activities and predicting their corresponding labels. The SVM model with default parameters gives the results as both of the accuracy and F1 score as 0.969.

- **K-Nearest Neighbors Classifier**

KNN Classifier, another ML model employed in this analysis, operates by computing the distances to the k nearest neighbors to determine the class of a given instance. The selection of the k value is crucial as it represents a trade-off parameter between overfitting and underfitting. When k is set to 1, the classifier tends to overfit the data, meaning it becomes too specific to the training examples and may not generalize well to new instances. Conversely, if the k value is too high, the classifier may struggle to capture intricate data patterns and become overly generalized. In this study, the initial forecast was performed using the default k value of 5. The results indicate an accuracy and F1 score of 0.95, highlighting the effectiveness of the KNN Classifier in accurately classifying activities in the dataset. The careful selection of the k value allows for a balance between capturing relevant patterns and avoiding overfitting or underfitting.

- **XGBoost**

In addition to the aforementioned models, XGBoost, one of the most renowned and effective gradient boosting algorithms, was employed in this analysis. XGBoost is an ensemble learning method that combines numerous weak predictive models to make accurate predictions. It has gained popularity due to its remarkable performance across various domains. This algorithm excels in both regression and classification tasks, sharing similarities with the Random Forest algorithm. In the context of this dataset, XGBoost demonstrated outstanding performance, outperforming other models. With accuracy and F1 scores of 0.99, XGBoost showcased its ability to provide highly accurate and reliable predictions for the given activities. Its ensemble approach, incorporating multiple weak models, allowed XGBoost to capture intricate patterns and make robust predictions, making it a valuable tool in activity recognition tasks.

- **Neural Network**

Additionally, Neural Network algorithms were utilized in this analysis, capable of learning intricate patterns in large datasets. Inspired by the human brain, Neural Networks consist of weighted nodes in different layers that are interconnected, mimicking the interconnected neurons in the brain. Through the optimization of weights and activation functions, information is propagated through the network. Each node in the network possesses specific weights and activation functions, which contribute to optimizing the score function. Given that this particular problem involves multi-class prediction, the activation function used in the output layer is softmax in the Neural Network model. The output layer comprises six nodes, corresponding to the six different classes in the dataset. With an impressive accuracy, and F1 score of 0.976, the Neural Network algorithm demonstrated its ability to effectively classify activities and make accurate predictions. Its capacity to handle complex patterns and large datasets makes Neural Networks a powerful tool in activity recognition tasks.

4 Experimental Results

	Accuracy	F1 Score
Naive Bayes Classifier	0.566	0.528

SVM	0.969	0.969
KNN	0.951	0.951
XGBoost	0.992	0.992
Neural Network	0.976	0.976

The results obtained from the different machine learning models are presented in the table above. These models include the Naive Bayes Classifier, Support Vector Machine (SVM), K-Nearest Neighbors (KNN), XGBoost, and the Neural Network. Let's analyze and compare these results to draw meaningful conclusions.

Starting with the Naive Bayes Classifier, it achieved an accuracy of 0.566 and an F1 score of 0.528. Although it performed reasonably well, its accuracy and F1 score are comparatively lower than the other models. Although it performed reasonably well, its accuracy and F1 score are comparatively lower than the other models. The high accuracy and F1 score indicate that SVM successfully captured the patterns and boundaries within the dataset. Next, the KNN classifier achieved an accuracy and F1 score of 0.951. The obtained accuracy and F1 score indicate that KNN performed well in understanding the data patterns and making accurate predictions. XGBoost, one of the most famous gradient boosting algorithms, produced outstanding results with an accuracy of 0.992 and an F1 score of 0.922. Lastly, the Neural Network achieved an accuracy and F1 score of 0.976. With a specific focus on multi-class prediction, the Neural Network demonstrated its capability to accurately classify the activities in the dataset.

In conclusion, based on the obtained results, XGBoost emerges as the top-performing model, achieving the highest accuracy and F1 score. However, it is crucial to consider the specific requirements of the task at hand, such as interpretability, computational complexity, and the importance of precision or recall. SVM and the Neural Network also demonstrated exceptional performance, making them suitable alternatives depending on the specific needs of the application. These results provide valuable insights into the strengths and capabilities of each model, enabling informed decision-making for future applications in activity classification and recognition tasks.

4.1 Hyperparameter Tuning

The model that yields the highest accuracy and F1 score values is XGBoost, making it the top-performing model in this analysis. Therefore, the next step is to fine-tune the hyperparameters specifically for XGBoost. The hyperparameters that require tuning include max_depth, learning_rate, n_estimators, min_child_weight, subsample, colsample_bytree, and objective.

Max_depth refers to the maximum depth of an individual tree in the XGBoost ensemble. By adjusting this hyperparameter, we control the complexity and depth of the trees, which can have a significant impact on the model's performance.

Learning_rate determines the step size used during the search for optimized trees. It serves as a regularization parameter to prevent overfitting by controlling the contribution of each tree in the ensemble.

N_estimators corresponds to the number of runs XGBoost will attempt to learn. It represents the number of boosting rounds or iterations the algorithm will perform, with a higher value potentially leading to better performance but also increasing computation time.

Min_child_weight represents the minimum sum of instance weight required in a child. It affects the tree's partitioning process by controlling the minimum weight required to create a new tree node.

Subsample indicates the subsample ratio of the training instances. It controls the percentage of data that will be randomly sampled and used in each boosting iteration, allowing for better regularization and reducing the risk of overfitting.

Colsample_bytree is a family of parameters that determine the subsampling of columns (features) for each tree. It specifies the fraction of columns to be randomly sampled at each tree construction, adding another level of regularization.

Objective refers to the loss function used in XGBoost. Different objectives are available, depending on the nature of the problem being solved. Selecting the appropriate objective is crucial for achieving optimal performance.

After performing hyperparameter tuning, it was found that the default parameters provided by XGBoost outperformed the candidate hyperparameters. The best model, with the default parameters, achieved an accuracy and F1 score of 0.963, which is lower than the accuracy and F1 score of 0.992 obtained by the top-performing model. This indicates that the default hyperparameters provided satisfactory results and did not require further adjustment for this particular dataset.

4.2 Explainable AI

Explainable Artificial Intelligence (XAI) is gaining significant traction as a powerful tool to enhance the understandability of black-box machine learning algorithms. It serves the dual purpose of aiding in feature selection and providing insights into the relative importance assigned to different features by the models. In our research, we leveraged XAI techniques to shed light on the pivotal role played by specific features in the recognition of human activity.

By employing state-of-the-art XAI methods such as SHAP (SHapley Additive exPlanations) and ELI5 (Explain Like I'm 5), we were able to extract importance values for various attributes across the six activity classes under consideration. The obtained results highlight the discriminative features for each class, offering valuable insights into the underlying patterns captured by the model.

- **Class 0 “STANDING”**: tGravityAcc-max()-Y, tGravityAcc-mean()-Z, tGravityAcc-sma()
- **Class 1 “SITTING”:** tGravityAcc-max()-Y, tGravityAcc-mean()-Z, angle(Y.gravityMean)
- **Class 2 “LAYING”:** tGravityAcc-max()-Y, tBodyAcc-min()-Y, tBodyGyro-min()-X,
- **Class 3 “WALKING”:** fBodyAccJerk-std()-Y, fBodyAccMag-std(), tBodyAccJerk-entropy()-X
- **Class 4 “WALKING_DOWNSTAIRS”:** fBodyAccMag-mad(), tBodyAcc-entropy()-X, tBodyAccJerk-entropy()-X
- **Class 5 “WALKING_UPSTAIRS”:** fBodyGyro-maxInds-Z, fBodyGyro-max()-Z, tGravityAcc-arCoeff()-Z,4

By leveraging XAI techniques and uncovering the relative importance of different features for each activity class, our research provides a comprehensive understanding of the underlying dynamics of human activity recognition. These findings have significant implications for the design and development of robust activity recognition systems, enabling informed decision-making in real-world applications.

In conclusion, our research showcases the effectiveness of XAI in unraveling the importance of features in human activity recognition. The insights gained from this study contribute to the growing body of knowledge in XAI and pave the way for further advancements in interpretable machine learning for activity recognition tasks.

5 Conclusions

In conclusion, this report contributes to the field of smartphone-based human activity recognition by investigating a combination of classifiers and introducing the XGBoost algorithm. The findings indicate that XGBoost achieves the highest accuracy and F1 score compared to the other tested models. However, it is important to consider specific task requirements, such as interpretability, computational complexity, and the significance of precision or recall, when selecting the most appropriate model. Both SVM and the Neural Network also exhibited outstanding performance, providing viable alternatives for different application scenarios. These results provide valuable insights into the strengths and capabilities of each model, enabling informed decision-making for future activity recognition tasks.

References

1. Davide Anguita, Alessandro Ghio, Luca Oneto, Xavier Parra and Jorge L. Reyes-Ortiz. Human Activity Recognition on Smartphones using a Multiclass Hardware-Friendly Support Vector Machine. International Workshop of Ambient Assisted Living (IWAAL 2012). Vitoria-Gasteiz, Spain. Dec 2012

2. Mannini, A., Sabatini, A.M.: Machine learning methods for classifying human physical activity from on-body accelerometers. Sensors 10(2), 1154–1175 (2010)

3. Ravi, N., Dandekar, N., Mysore, P., Littman, M.L.: Activity recognition from accelerometer data. In: Proceedings of the Seventeenth Conference on Innovative Applications of Artificial Intelligence, IAAI, pp. 1541–1546. AAAI Press (2005)

4. Davide Anguita, Alessandro Ghio, Luca Oneto, Xavier Parra, Jorge L. Reyes-Ortiz. Energy Efficient Smartphone-Based Activity Recognition using Fixed-Point Arithmetic. Journal of Universal Computer Science. Special Issue in Ambient Assisted Living: Home Care. Volume 19, Issue 9. May 2013