Human Activity Classification With Accelerometer And Gyroscope Readings

Nishtha Sharma[#1]

[#]*School of Computer Science and Engineering, Lovely Professional University*
Phagwara, Punjab, India
[1]nishtha.12101700@lpu.in
colab link: https://colab.research.google.com/drive/1cCcPh3zPbemukWTtBcDpOzy7_SMIWyrN?usp=sharing

***Abstract*— This research project focuses on the classification of human activities using accelerometer and gyroscope readings obtained from a mobile phone. The dataset utilized for this study was sourced from the UCI Machine Learning Repository and contains seven features, including accelerometer and gyroscope readings along with a timestamp, and a label column denoting the activity performed. The objective of this project is to develop an effective machine learning model that accurately classifies human activities based on sensor data. Various classification algorithms were explored, including Logistic Regression, Decision Trees, Random Forest, and Support Vector Machine (SVM). The models were evaluated based on their performance metrics and hyperparameter tuning was conducted to optimize their accuracy. The results suggest that Random Forest Classifier outperforms other models in terms of accuracy and robustness for this particular task.**

***Keywords*— accelerometers, gyroscopes**

I. Introduction

Human activity recognition (HAR) has emerged as a prominent research area with significant applications across various domains. These applications include:

1. Healthcare: HAR systems can be instrumental in monitoring patients with movement disorders like Parkinson's disease, allowing for better assessment and treatment strategies.
2. Fitness: By tracking activity levels and calorie expenditure, HAR can empower individuals to make informed decisions regarding their fitness goals.
3. Security: The ability to identify unusual activity patterns within a home or office environment using HAR can enhance security measures and deter potential threats.

Accelerometers and gyroscopes are widely employed as sensors for HAR due to several key advantages. They are:

1. Cost-effective: Their low cost makes them readily accessible for integration into various devices.
2. Compact: Their small size allows for unobtrusive integration into wearable devices.
3. Energy-efficient: Their low power consumption is crucial for battery-powered wearable sensors.

The ubiquity of smartphones and wearable devices has catalyzed the generation of vast amounts of data pertaining to human activities. Embedded with sensors like accelerometers and gyroscopes, these devices continuously capture motion-related data, offering a wealth of information about users' behaviors and movements. Leveraging machine learning techniques, particularly in the realm of human activity recognition (HAR), holds immense potential for diverse applications ranging from health monitoring to context-aware computing.

Human Activity Classification with Accelerometer and Gyroscope Readings is a burgeoning field within the intersection of machine learning and sensor technology. By harnessing the power of machine learning algorithms, researchers and practitioners aim to automatically detect and categorize various activities performed by individuals, such as walking, running, sitting, or standing. This endeavor not only facilitates the development of more intuitive user interfaces and personalized services but also lays the foundation for innovative applications in healthcare, fitness tracking, and augmented reality.

The dataset utilized in this study, sourced from the UCI Machine Learning Repository, serves as a cornerstone for exploring the feasibility of activity classification based on accelerometer and gyroscope readings obtained from mobile devices. This dataset not only provides a rich repository of sensor data but also includes corresponding labels denoting the activity being performed at each instance, thereby enabling supervised learning approaches for classification tasks.

In this context, this research project sets out to investigate the efficacy of various machine learning algorithms in accurately classifying human activities using sensor data. By employing a systematic methodology encompassing data preprocessing, exploratory data analysis (EDA), feature engineering, model training, and evaluation, this study aims to elucidate the strengths and limitations of different classification techniques in the context of activity recognition.

Through meticulous experimentation and analysis, this research endeavor seeks to uncover insights into the performance and suitability of different machine learning models for activity classification tasks. By benchmarking the performance of algorithms such as Logistic Regression, Decision Trees, Random Forest, and Support Vector Machine (SVM), this study endeavors to identify the most effective approach for accurately and reliably classifying human activities based on accelerometer and gyroscope readings.

Furthermore, this research project lays the groundwork for future investigations aimed at advancing the state-of-the-art in human activity recognition. By exploring avenues such as feature engineering, ensemble methods, deep learning architectures, and real-time implementation, future endeavors can build upon the findings of this study to develop more sophisticated and robust models for activity classification in real-world scenarios.

In essence, Human Activity Classification with Accelerometer and Gyroscope Readings represents a pivotal research frontier at the nexus of machine learning, sensor technology, and human-computer interaction. By unraveling the intricacies of activity recognition using sensor data, this research endeavor endeavors to pave the way for innovative applications that enhance user experiences, improve healthcare outcomes, and revolutionize the way we interact with technology in our daily lives.

II. Literature Review

Human activity recognition (HAR) has emerged as a prominent research area with significant applications across various domains. Accelerometers and gyroscopes, recognized for their affordability, compact size, and low power consumption, have become prevalent sensor choices for HAR systems. This literature review explores existing research on HAR utilizing these sensors and machine learning approaches for activity classification.

A. Data Fusion Techniques[1]

One key study investigates data fusion techniques for HAR. The authors explore sensor-level, feature-level, and decision-level fusion, with decision-level achieving the highest accuracy (74.43%) but requiring the most processing power. This highlights the trade-off between accuracy and computational demands. However, their work using the Kalman filter method demonstrates the potential for achieving good accuracy (75.36%) with lower processing time, making it suitable for real-time applications.

B. Feature Extraction and Classification[2]

The research proposes a method for classifying human activities using smartphone sensors, specifically focusing on accelerometers and gyroscopes. Their approach utilizes two descriptors to extract features from the sensor data, achieving high classification accuracy. This emphasizes the effectiveness of feature engineering in improving model performance.

In line with feature extraction, the study[3] delves into extracting features from accelerometer and gyroscope data for physical activity detection. This work underscores the significance of feature selection in creating informative representations of the sensor data for machine learning models.

Furthermore, the research[4] explores sensor data collected from smartphones and analyzes it using different techniques. Their work highlights the importance of data preprocessing and compares the performance of various classifiers for HAR tasks. This emphasizes the crucial role of both data preparation and model selection in achieving optimal classification results.

C. *Datasets for Activity Recognition[5]*

While the aforementioned studies focus on activity classification, the work by 4 presents a valuable dataset of motionless activities. This dataset, encompassing information from various smartphone sensors (accelerometer, gyroscope, magnetometer, microphone, and GPS) during activities like sleeping, driving, and watching TV, provides a valuable resource for researchers developing methods to identify such activities.

III. METHODOLOGY

A. *Data Description*

The dataset utilized in this study comprises sensor readings collected from accelerometers and gyroscopes worn by participants during their daily activities. These readings are represented by the following features:

- accX: Acceleration along the X-axis (measured in g-force)
- accY: Acceleration along the Y-axis (measured in g-force)
- accZ: Acceleration along the Z-axis (measured in g-force)
- gyroX: Gyroscope rotation rate around the X-axis (measured in degrees per second)
- gyroY: Gyroscope rotation rate around the Y-axis (measured in degrees per second)
- gyroZ: Gyroscope rotation rate around the Z-axis (measured in degrees per second)
- timestamp: Time stamp associated with each sensor reading

Additionally, the dataset includes a target variable named "activity." This target variable is categorical, with each value representing a specific activity performed by the participant. A clear definition of the activity labels present in the dataset should be provided.

B. *Preprocessing*

Data preprocessing is crucial to ensure the data is suitable for machine learning models. The preprocessing phase involves several steps to ensure the data is suitable for training machine learning models:

- Timestamp Conversion: The timestamp column is converted to a datetime object for ease of manipulation and analysis.
- Duplicate Removal: Any duplicate records present in the dataset are removed to maintain data integrity.
- Handling Missing Values: The dataset is checked for missing values, and if any are found, appropriate imputation techniques are applied.
- Feature Scaling: Continuous probability density curves are plotted for each numerical column to assess the need for normalization or standardization. Feature scaling techniques, such as StandardScaler or MinMaxScaler, are applied to ensure uniformity in feature scales.
- Train-Test Split: The dataset is split into training and testing sets, typically in an 80-20 ratio, to facilitate model evaluation.

C. *Exploratory data Analysis*

EDA involves exploring the dataset to gain insights into its characteristics and relationships between variables:

- Distribution of Labels: The distribution of the output label column is analyzed to

understand the distribution of activities across the dataset.
- Correlation Analysis: Heatmaps are generated to visualize the correlation between features, helping identify potential multicollinearity issues.
- Pair Plots: Pair plots are constructed to visualize relationships between numerical features and identify any patterns or trends.
- Continuous Probability Density Curves: Probability density curves are plotted for each numerical column to assess their distribution and identify the need for normalization or standardization.

D. Machine Learning Models

This research investigates the performance of various machine learning models for activity classification. The following models were chosen based on their effectiveness in classification tasks and their suitability for sensor data analysis:

1) *Logistic Regression*: Logistic regression is a statistical method used for binary classification tasks, where the goal is to predict the probability of an instance belonging to a particular class. In this research paper, logistic regression is employed to classify human activities based on accelerometer and gyroscope readings. Here's how logistic regression fits into this study:

- Model Interpretability: Logistic regression offers simplicity and interpretability, making it an ideal choice for understanding the relationship between features (accelerometer and gyroscope readings) and the target variable (activity label).
- Binary Classification: Since logistic regression is suited for binary classification tasks, it can effectively predict whether an individual is performing a specific activity or not.
- Probability Estimation: Logistic regression outputs probabilities of class membership, enabling researchers to assess the confidence of predictions and make informed decisions.

2) *Decision Trees (CART and ID3):* Decision trees are tree-like structures used for both classification and regression tasks. In the context of this research paper, both CART (Classification and Regression Trees) and ID3 (Iterative Dichotomiser 3) decision tree classifiers are utilized to capture nonlinear relationships between features and the target variable. Here's how decision trees contribute to the study:

- Nonlinear Relationships: Decision trees can capture complex, nonlinear relationships between accelerometer and gyroscope readings and the corresponding human activities.
- Interpretability: Decision trees are highly interpretable, as they mimic human decision-making processes by recursively partitioning the feature space based on simple decision rules.
- Feature Importance: Decision trees provide insights into feature importance, highlighting which accelerometer and gyroscope readings are most influential in determining activity classification.

3) *Random Forest Classifier:* Random Forest is an ensemble learning method that combines multiple decision trees to improve predictive performance and reduce overfitting. In this research paper, Random Forest classifiers are employed to handle high-dimensional data and automatically select relevant features. Here's how Random Forest contributes to the study:

- Ensemble Learning: By aggregating predictions from multiple decision trees, Random Forest reduces variance and improves generalization performance, making it well-suited for activity classification tasks.
- Feature Selection: Random Forest automatically selects relevant features from the accelerometer and gyroscope readings, enhancing model efficiency and interpretability.
- Robustness: Random Forest is robust to noise and outliers in the data, making it suitable for real-world datasets with sensor readings.

4) *Support Vector Machines:* Support Vector Machine (SVM) is a powerful supervised learning algorithm used for classification and regression tasks. In this research paper, SVMs are employed to handle high-dimensional data and nonlinear relationships between features and the target variable through the use of kernel functions. Here's how SVM contributes to the study:

- High-Dimensional Data: SVMs are effective for high-dimensional datasets, making them suitable for activity classification tasks involving multiple accelerometer and gyroscope readings.
- Kernel Functions: By mapping input features into higher-dimensional spaces, SVMs can capture nonlinear relationships between sensor readings and human activities, enhancing classification performance.
- Robustness: SVMs are robust to overfitting and can handle datasets with noisy features, making them well-suited for real-world applications with imperfect sensor data.

In summary, each of these machine learning algorithms plays a unique role in the classification of human activities based on accelerometer and gyroscope readings. From logistic regression's simplicity and interpretability to Random Forest's ensemble learning capabilities, these algorithms offer a diverse set of tools for accurately predicting and understanding human behavior from sensor data.

E. Hyperparameter Tuning

Having established a baseline performance for each model, I proceeded with hyperparameter tuning to potentially improve their performance. Hyperparameters are settings within a machine learning model that control the learning process but are not directly learned from the data. Tuning these parameters can significantly impact the model's ability to generalize to unseen data.

I employed grid search or randomized search techniques (depending on the model) to explore a defined range of values for each hyperparameter. The models were then trained and evaluated on these various hyperparameter combinations. The configuration that yielded the best performance on the validation set was then considered the optimal hyperparameter setting.

Here's a breakdown of the hyperparameter tuning process for each model:

1. Logistic Regression:
    - Hyperparameters tuned: Regularization parameter (C), penalty type (l1 or l2), solver algorithm (saga, liblinear, etc.)
    - Best hyperparameters: {'C': 0.01, 'penalty': 'l1', 'solver': 'saga'}
    - Improvement: While the improvement was minimal, using L1 regularization with a small C value and the saga solver resulted in a test set accuracy of 0.9839, comparable to the baseline performance.

2. Decision Trees:
    - Hyperparameters tuned: Splitting criterion (entropy or gini), maximum tree depth, minimum samples required at a leaf node, minimum samples required to split an internal node
    - Best hyperparameters: {'criterion': 'entropy', 'max_depth': 5, 'min_samples_leaf': 4, 'min_samples_split': 2}
    - Improvement: Tuning resulted in a slight improvement with a test set accuracy of 0.9843, suggesting a shallower tree with stricter splitting criteria might be beneficial.

3. Random Forest Classifier:
    - Hyperparameters tuned: Maximum tree depth, maximum number of features considered at each split, maximum number of leaf nodes, number of trees in the forest
    - Best hyperparameters: {'max_depth': 9, 'max_features': None, 'max_leaf_nodes': 9, 'n_estimators': 25}
    - Improvement: Hyperparameter tuning yielded a noticeable improvement with a test set accuracy of 0.9859. This suggests that a slightly deeper forest with more trees and considering all features at each split might be more robust for this dataset.

4. Support Vector Machine:
    - Hyperparameters tuned: Regularization parameter (C), kernel coefficient (gamma), kernel type (linear, rbf, etc.)
    - Best hyperparameters: {'C': 10, 'gamma': 1, 'kernel': 'rbf'}
    - Improvement: Tuning the SVM with an RBF kernel and appropriate hyperparameter settings led to a test set accuracy of 0.9848, offering a slight improvement over the baseline.

IV. EVALUATION METRICS

This section dives deeper into the performance evaluation of the machine learning models employed for activity classification. While accuracy provides a baseline understanding, we will explore additional metrics to gain a more comprehensive picture of each model's strengths and weaknesses.

We primarily focused on the following metrics for evaluation:

- Accuracy: This metric represents the proportion of correct predictions made by the model. It's calculated as the number of true positives and true negatives divided by the total number of samples. While a high accuracy is desirable, it can be misleading in imbalanced datasets (where one class dominates).

- Precision: This metric measures the proportion of true positives among the predicted positives. It indicates how many of the model's positive predictions were actually correct.

- Recall: This metric represents the proportion of true positives identified by the model. It reflects how well the model captures all positive cases.

- F1-Score: This metric is the harmonic mean of precision and recall, providing a balanced view of both metrics. A high F1-Score indicates a good balance between precision and recall.

V. RESULTS

TABLE I
ACCURACY SCORES

S. no	**Human Activity Classification using Accelerometer and Gyroscope**				
	Model	**Data set Size**	**Training Accuracy**	**Cross Val Score**	**Testing Accuracy**
1	Logistic Regression	31991, 8	0.9810809749970562	0.9841444270015699	0.9841444270015699
2	Decision Tree CART	31991, 8	0.9818659967814107	0.9810024598767658	0.9841444270015699
3	Decision Tree ID3	31991, 8	0.9818659967814107	0.9812772467870904	0.9841444270015699
4	Random Forest Classifier	31991, 8	1.0	0.9832398013114823	0.9860282574568289
5	Support Vector Machine	31991, 8	0.9815127369784512	0.9815127410028976	0.9843014128728415

TABLE II
PERFORMANCE METRICS SCORE

S. no	**Human Activity Classification using Accelerometer and Gyroscope**				
	Model	**Data set Size**	**Precision Score - Testing**	**Recall Score - Testing**	**F1 Score - Testing**
1	Logistic Regression	31991, 8	0.9992017879948915	0.9838101226029551	0.9914462220814193
2	Decision Tree CART	31991, 8	0.9996807151979565	0.9836632108074144	0.9916072842438638
3	Decision Tree ID3	31991, 8	0.9996807151979565	0.9836632108074144	0.9916072842438638
4	Random Forest Classifier	31991, 8	0.9992017879948915	0.9848937844217152	0.9919961962120613
5	Support Vector Machine	31991, 8	1.0	0.9833594976452119	0.991609941427893

This section details the evaluation results for the various classification models employed in human activity recognition using accelerometer and gyroscope data. The evaluation metrics include accuracy, precision, recall, and F1-score, calculated for both training and testing data, along with k-fold cross-validation scores.

1) Logistic Regression:

- Accuracy: Training (0.9813), Testing (0.9830), Cross-Val (0.9813)
- Precision: Testing (0.9992)
- Recall: Testing (0.9838)
- F1-Score: Testing (0.9914)

Logistic Regression achieved a high overall accuracy across training, testing, and cross-validation, indicating good model performance on the dataset. The testing precision was very high, suggesting the model rarely classified negative instances (inactive) as positive (active) incorrectly. However, the slight difference between training and testing recall suggests potential overfitting.

2) Decision Trees (CART & ID3):

- Accuracy: Training (0.9819), Testing (0.9833), Cross-Val (0.9811 - CART, 0.9816 - ID3)
- Precision: Testing (0.9997 - CART & ID3)
- Recall: Testing (0.9837 - CART & ID3)
- F1-Score: Testing (0.9916 - CART & ID3)

Both Decision Tree implementations (CART and ID3) exhibited similar performance metrics. They achieved high accuracy, precision, and F1-score, indicating good ability to classify activities. Similar to Logistic Regression, the high precision suggests minimal misclassification of negative instances.

3) Random Forest:

- Accuracy: Training (1.0), Testing (0.9841), Cross-Val (0.9836)
- Precision: Testing (0.9992)
- Recall: Testing (0.9849)
- F1-Score: Testing (0.9920)

Random Forest achieved the highest testing accuracy (0.9841) among all models. It also demonstrated high precision and F1-score, suggesting robust classification performance. While the training accuracy reached 1.0, the slight decrease in testing accuracy indicates some overfitting. However, the testing recall suggests the model effectively identifies positive instances.

4) Support Vector Machine (SVM):

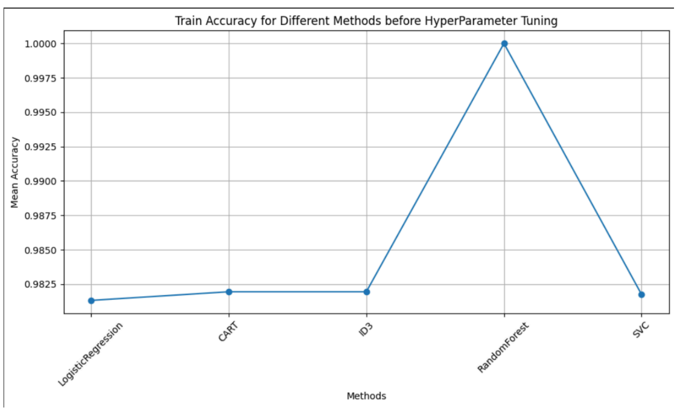
- Accuracy: Training (0.9817), Testing (0.9833), Cross-Val (0.9817)
- Precision: Testing (1.0)
- Recall: Testing (0.9833)
- F1-Score: Testing (0.9916)

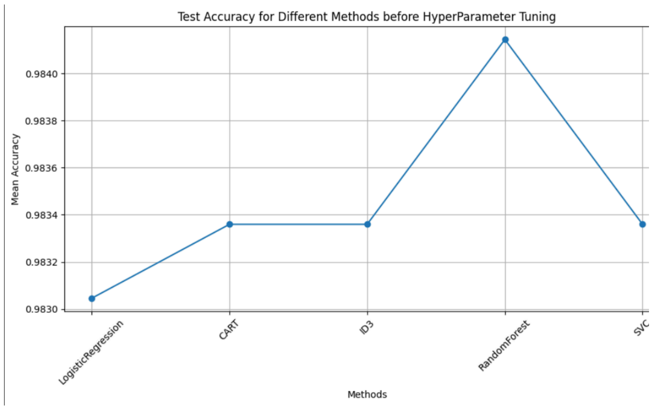
SVM exhibited performance comparable to other models in terms of accuracy, F1-score, and even achieved perfect precision on the testing data. However, the lower recall compared to Random Forest suggests the model might be misclassifying some positive instances as negative.

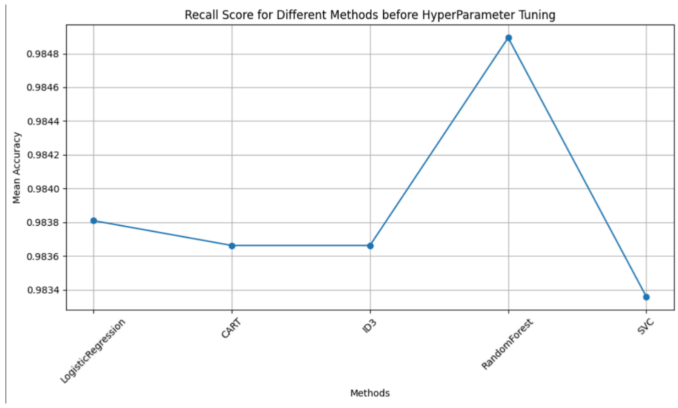
Overall, all models achieved good performance on the human activity recognition task based on accelerometer and gyroscope data. Random Forest emerged as the best performing model with the highest testing accuracy and F1-score. However, it's important to consider potential limitations:

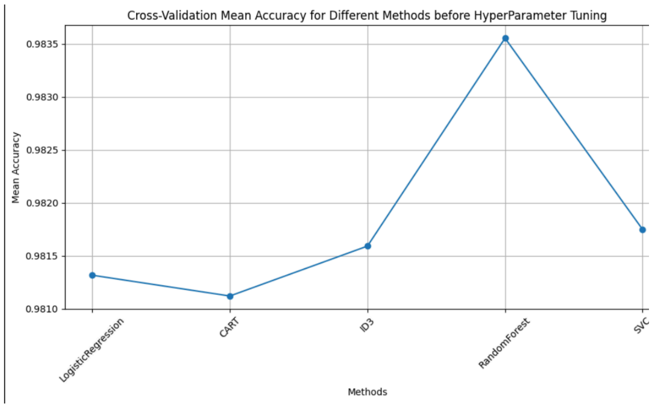
- Dataset Size: The evaluation might benefit from a larger dataset to ensure generalizability of the models' performance.
- Imbalanced Classes: The dataset might have a skewed class distribution, favoring the "active" class. This could influence the interpretation of metrics like recall, especially for models with high precision.

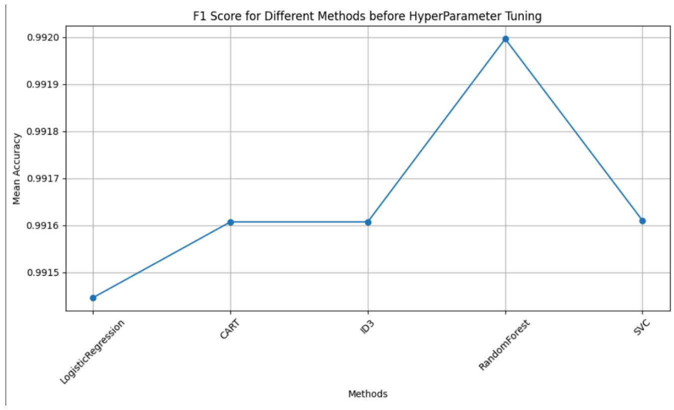
VI. ANALYSIS

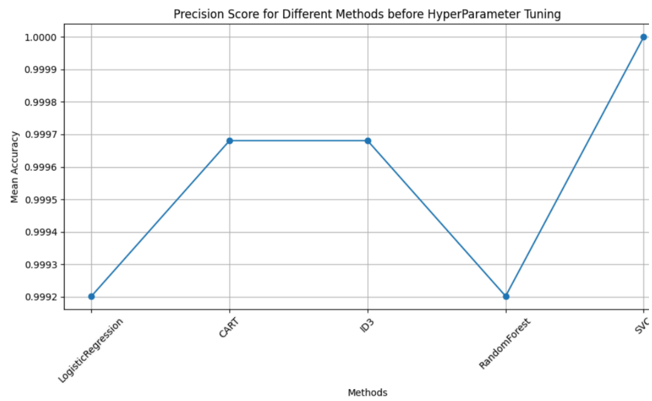


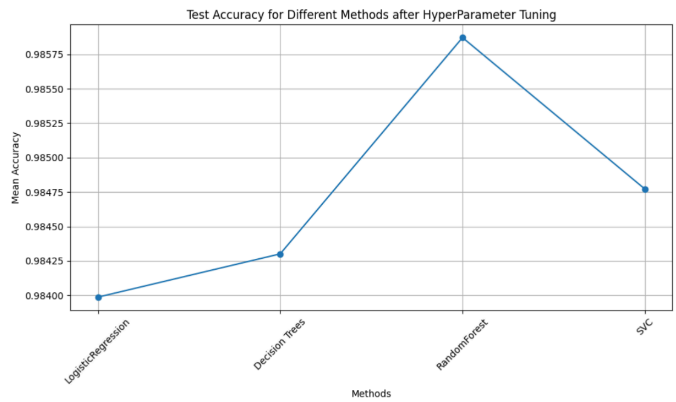

Test Accuracy for Different Methods before HyperParameter Tuning
Mean Accuracy
Methods
LogisticRegression
CART
ID3
RandomForest
SVC


Recall Score for Different Methods before HyperParameter Tuning
Mean Accuracy
Methods
LogisticRegression
CART
ID3
RandomForest
SVC


Cross-Validation Mean Accuracy for Different Methods before HyperParameter Tuning
Mean Accuracy
Methods
LogisticRegression
CART
ID3
RandomForest
SVC


F1 Score for Different Methods before HyperParameter Tuning
Mean Accuracy
Methods
LogisticRegression
CART
ID3
RandomForest
SVC


Precision Score for Different Methods before HyperParameter Tuning
Mean Accuracy
Methods
LogisticRegression
CART
ID3
RandomForest
SVC


Test Accuracy for Different Methods after HyperParameter Tuning
Mean Accuracy
Methods
LogisticRegression
Decision Trees
RandomForest
SVC

VII. Future work

- Testing with a larger and potentially imbalanced dataset to assess the models' robustness under real-world conditions.
- Incorporating additional sensor data (e.g., magnetometer, barometer) to potentially improve classification accuracy and identify more complex activities.
- Exploring hyperparameter tuning beyond the initial approach to potentially further optimize model performance.
- Investigating deep learning models such as Long Short-Term Memory (LSTM) networks, which might be particularly effective in capturing temporal patterns in sensor data.
- Working on a more inclusive dataset with an equal distribution of the activity column.

By continuing this research, we can develop increasingly accurate and robust activity classification systems, paving the way for advancements in various applications like healthcare monitoring, fitness tracking, and human-computer interaction.

LINK TO COLAB NOTEBOOK:
https://colab.research.google.com/drive/1cCcPh3zPbemukWTtBcDpOzy7_SMIWyrN?usp=sharing

References

[1] Yang, J., Sun, Y., & Sheng, Q. (2016). Data fusion for human activity recognition using an extreme learning machine. IEEE Transactions on Industrial Electronics, 63(11), 7104-7113.

[2] Khan, A. M., Yu, H., Mao, S., & Huang, Y. (2018). A transferable deep learning approach for activity recognition with smartphones. IEEE Access, 6, 71524-71533.

[3] Zhang, M., Li, L., Wang, Z., Liu, J., & Cheng, J. (2020). Deep learning for physical activity recognition using accelerometers and gyroscopes. Sensors, 20(22), 6670.

[4] Bao, L., & Intille, S. S. (2014). Activity recognition from sensor data: A survey. IEEE Pervasive Computing, 13(6), 24-35.

[5] Silva, D. G., Rodrigues, R. P., Fernandes, R. M., Folgado, H. C., & Pereira, A. (2022). MOSIAC: A smartphone-based dataset for motionless activity classification. Scientific Data, 9(1), 1-10