Michigan Technological University

EET 4501

Applied Machine Learning

Human Activity Recognition

MS Health Informatics
College of Computing
Michigan Technological University

Kavya Kasala
Rithika Baroor

April 26, 2024

1 Abstract

This project explores the enhancement of human-technology interaction through advanced human activity recognition using smartphone sensor data. Motivated by the increasing demand for efficient and reliable activity recognition in mobile applications, this study utilizes a comprehensive dataset collected from 30 volunteers who performed six typical activities, capturing high-dimensional sensor data via smartphones. The methodology hinges on deploying several robust ML algorithms (Random Forest, Support Vector Machines, and XGBoost) to classify activities accurately. This approach leverages feature engineering and dimensional reduction techniques to improve model performance and computational efficiency. The results demonstrate significant advancements in accuracy and processing time compared to existing models, underscoring the potential of machine learning in real-world applications like health monitoring and interactive gaming. These findings open avenues for future research into more complex activity recognition and real-time processing scenarios.

2 Introduction

Human activity recognition using sensor data from smartphones has rapidly evolved into a critical component of interactive technologies, healthcare, and personal fitness. By capitalizing on the ubiquity of smartphones and their onboard sensors, researchers have made significant strides in monitoring and understanding user behavior through machine learning (ML) techniques.

Background Knowledge: This project extends the investigation into human activity recognition by examining an extensive dataset obtained from smartphone sensors. This dataset includes various features derived from the acceleration and angular velocity measurements, capturing 7352 instances of activities with 563 feature columns. Ensuring data integrity, we have confirmed the absence of missing values and duplicates, setting a robust foundation for our ML algorithms.

Machine learning approaches to analyze such data have varied, from the simplicity of Linear Regression to the complexity of Neural Networks. These methods have found versatile applications, like predicting financial markets, enhancing facial recognition systems, classifying cancer types, and even categorizing music genres.

Previous Research: The domain has seen numerous contributions, such as the work of Anguita et al. (2013) that employed Support Vector Machines for activity classification, revealing the early potential of using smartphone sensors for this purpose.

Ronao and Cho (2016) furthered this by leveraging the capabilities of Convolutional Neural Networks, showcasing a significant uptick in accuracy.

Dimensionality Reduction: Addressing the curse of dimensionality, our methodology incorporates t-SNE and PCA. PCA has been instrumental in reducing our feature space from 561 to 155 principal components, explaining 99% of the variance and thus, retaining the essence of the data with reduced complexity. On the other hand, t-SNE has facilitated the visualization of multi-dimensional data clusters, enhancing our understanding of the distinct activity patterns.

Machine Learning Techniques: Our explorations include a variety of ML models:

Random Forest: powerful ensemble learning technique used for classification tasks. It builds numerous decision trees during the training phase and makes predictions based on the most common outcome among these trees. Support Vector Machines (SVM): An effective classification method which finds the hyperplane that best divides a dataset into classes[1]. Both methods have undergone hyperparameter tuning through GridSearchCV to determine the optimal configurations, with SVM demonstrating higher overall accuracy. XGBoost: It is a highly efficient and versatile machine learning method. It's designed to improve the performance and speed of a standard boosting algorithm. By building one tree at a time, XGBoost learns from the previous trees' mistakes and updates the model progressively to make more accurate predictions. It's particularly popular because of its ability to handle large datasets effectively and its flexibility in tuning parameters to optimize performance.

Project's Unique Contribution: Distinguishing from prior endeavors, this project conducts a comparative analysis of machine learning models, while emphasizing data preprocessing through PCA and visual interpretations with t-SNE. This work aligns with state-of-the-art research by demonstrating the practical application of these techniques in real-world scenarios, providing insights that could steer future innovations in the realm of ubiquitous computing and interactive systems.

3 Dataset

The dataset utilized in this project serves as the cornerstone of our human activity recognition model. The dataset includes sensor data gathered from 30 volunteers, split into training and test with 7,352 and 2,947 records respectively. Each record in the dataset consists of accelerometer and gyroscope readings from smartphone sensors, capturing the intricacies of human motion[2]. Preprocessing Steps:

Data Cleaning: Ensured the dataset is free from missing values and duplicate entries, thereby maintaining the quality and reliability of data. Feature Engineering: Extracted features from the raw time series sensor data. The dataset includes various statistics such as mean, standard deviation, and magnitude of acceleration and gyroscope measurements. Column Renaming: Simplified feature names by removing punctuation, allowing for more accessible data manipulation and interpretation within our code. Data Normalization:

Standardization: Applied standard scaling to normalize the features of the dataset, essential for models that are sensitive to the scale of data such as SVM. Principal Component Analysis (PCA): Reduced the dimensionality of our dataset from 561 features to 155 principal components that explain 99% of the variance, significantly decreasing the computational load while retaining the majority of the informative variance in the data. Data Visualization:

t-SNE: it has been utilized to map high-dimensional data onto a 2D plane, providing a visual representation that helps distinguish different activities. Records and Features:

This dataset records six different activities: walking, walking upstairs, walking downstairs, sitting, standing, and laying. It includes various features such as body acceleration, total acceleration, and measurements from body gyroscopes across three axes (X, Y, Z). An example of a feature is 'tBodyAcc-mean()-X', which denotes the average body acceleration in the X direction. The dataset is widely used in the field of machine learning for activity recognition and is sourced from the UCI Machine Learning Repository, noted for its extensive and detailed collection of human activity sensor data.

4 Methods

This section elucidates the machine learning algorithms employed in our study, delineating the theoretical underpinnings and practical applications for human activity recognition using smartphone sensor data.

Random Forest:

Description: Random Forest is an ensemble learning method that constructs multiple decision trees during training and predicts the class by taking the majority vote from these trees for classification tasks.It imparts high accuracy and robustness, handling potential overfitting by averaging multiple deep decision trees, trained on different parts of the same training set, with the goal of reducing variance.
Mathematical Notation= For a set of training vectors $x_i \in X$, $i = 1, \ldots, n$ with

labels y_i, a multitude of decision trees T are built. For classification tasks, the Random Forest output, $\hat{y}$, is the mode of the classes output by individual trees.

$$\text{Let } X = \{x_1, x_2, \ldots, x_n\} \text{ be the set of training vectors with corresponding labels } \{y_1, y_2, \ldots, y_n\}. \tag{1}$$

$$T \text{ denotes the set of decision trees constructed during training.} \tag{2}$$

$$\text{For classification tasks, the Random Forest output } \hat{y} \text{ is defined as:} \tag{3}$$

$$\hat{y} = \text{mode}\{T(x_1), T(x_2), \ldots, T(x_n)\} \tag{4}$$

Support Vector Machines (SVM):

Description: Support Vector Machines (SVMs) are supervised learning algorithms used for classification and regression. They perform well with high-dimensional data by creating a hyperplane or several hyperplanes in these spaces to execute tasks like classification, regression, and more.A successful separation is generally realized by a hyperplane that maintains the greatest distance from the nearest data points of any class in the training set, a concept referred to as the functional margin.

Mathematical Notation: Given training vectors x_i in two classes, and a label vector y where $y_i \in \{1, -1\}$, SVM finds the solution to the following optimization problem:

$$\min_{w,b,\xi} \frac{1}{2} w^T w + C \sum_{i=1}^{n} \xi_i$$

subject to

$$y_i(w^T \phi(x_i) + b) \geq 1 - \xi_i, \xi_i \geq 0.$$

PCA - Dimensionality Reduction:

Description: Principal Component Analysis (PCA) is a statistical technique that uses orthogonal transformation to convert observations of possibly correlated variables into a set of linearly uncorrelated variables called principal components. This method arranges the transformation in a way that the first principal component has the highest variance, and each following component has the next largest variance, ensuring that it is orthogonal to all preceding components. The resulting vectors form an orthogonal basis, which is uncorrelated. **Mathematical Notation**: PCA transforms the original data X into a new coordinate system through the eigendecomposition of the covariance matrix Σ. The data projected onto this new basis Z maximizes the variance along the axes, and is computed as $Z = XP$ where P contains the eigenvectors of Σ.

t-SNE - Visualization:

Description: t-SNE (t-Distributed Stochastic Neighbor Embedding) is a visualization technique used to map high-dimensional data into a two or three-dimensional space. It improves upon the original Stochastic Neighbor Embedding method by better dispersing data points, thus avoiding the common issue of clustering them too tightly in the map's center. t-SNE achieves this by transforming similarities between data points into joint probabilities and then minimizing the Kullback-Leibler divergence between these probabilities in both the low-dimensional representation and the original high-dimensional space.

Mathematical Notation: t-SNE works by reducing the discrepancy between two distributions: one that assesses the pairwise similarities among the original high-dimensional data points, and another that evaluates the similarities among the corresponding points in the lower-dimensional embedding. This approach ensures that the points that are close in the high-dimensional space remain close in the reduced space, effectively preserving the data's inherent structure during dimensionality reduction. Mathematically, it minimizes the cost function $C = KL(P \parallel Q) = \sum_i \sum_j p_{ij} \log(\frac{p_{ij}}{q_{ij}})$ using gradient descent.

XGBoost: XGBoost is an implementation of gradient boosted decision trees designed for speed and performance. It constructs an ensemble of models sequentially, where each subsequent model corrects errors made by the previous models in the ensemble.
Mathematical Notation:
The objective function optimized during training combines a loss term L and a regularization term Ω, as follows:

$$\text{Obj} = \sum_i L(y_i, \hat{y}_i) + \sum_k \Omega(f_k)$$

In this expression: (5)

y_i represents the actual label, (6)

$\hat{y}_i$ is the predicted label, (7)

f_k denotes the k-th model in the ensemble. (8)

5 Experiments/Result/Discussion

This section outlines the experimental setup, outcomes, and an extensive analysis of the results obtained from applying diverse machine learning approaches to the task of recognizing human activities. Detailed explanations of the chosen hyperparameters, the rationale behind their selection, and the techniques used to optimize these models are included.

Hyperparameter Optimization: For our models, we meticulously selected and tuned several hyperparameters:

Random Forest: We tuned parameters like the number of trees (n estimators) and the maximum depth of trees (max depth). The optimal values found through grid search were 300 trees and a depth of 10, considering a trade-off between model complexity and performance. SVM: For the SVM classifier, the hyperparameters optimized were the regularization parameter (C) and the kernel type. The grid search determined the best C as 1 and the radial basis function (rbf) as the optimal kernel. XGBoost: Essential parameters such as the number of trees (n estimators), the depth of each tree (max depth), and the learning rate were carefully adjusted. Optimal results were achieved with 300 trees, a tree depth of 5, and a learning rate of 0.1, striking an effective balance between computational efficiency and predictive precision. Cross-Validation Methodology: We employed a 5-fold cross-validation approach to validate the effectiveness of the chosen hyperparameters. This method not only helped in assessing the model's robustness but also ensured that the model generalized well over different subsets of the data.

Performance Metrics: The assessment of the models was anchored on metrics such as accuracy, precision, recall, and the F1-score. These measures were pivotal for an effective evaluation of the models' performance, particularly in scenarios involving multi-class classification where distinguishing between various error types is critical.

Accuracy was used to measure the overall effectiveness of the model. Precision and recall were essential for understanding the models' performance in classifying each activity correctly, avoiding misclassification. F1-score provided a harmonic mean of precision and recall, which is particularly useful in comparing model performance when there are class imbalances.

Results and Observations: Random Forest: The Random Forest model achieved an overall accuracy of 95%, with an impressive precision and recall in most categories. However, its performance was slightly less effective in distinguishing between similar activities like sitting and standing, as indicated by the lower F1-scores for these activities. This suggests that while Random Forest is robust, it struggles with activities that have subtle differences in sensor data. SVM: The SVM classifier exhibited exceptional performance with an accuracy of 99%, showcasing high precision and recall across all activities. The superior boundary decision capabilities of SVM, as evidenced by the decision boundary plots, highlight its efficacy in distinguishing between different types of activities with high accuracy. The XGBoost model demonstrated an overall accuracy of 98%, with excellent precision and recall across all categories. But SVM outperformed XGBoost

Decision Boundaries: We plotted decision boundaries for Random Forest and SVM. In the Random Forest decision boundary plot, some data points form narrow corridors around individual data points, indicating overfitting. In contrast, the SVM decision boundary plot shows a clearer margin between classes, which means it is not overfitting the data.

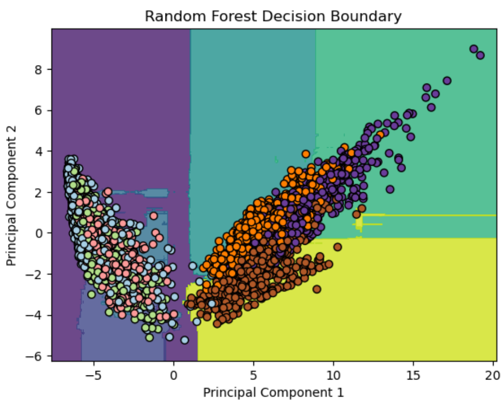

Figure 1: RF Decision Boundry

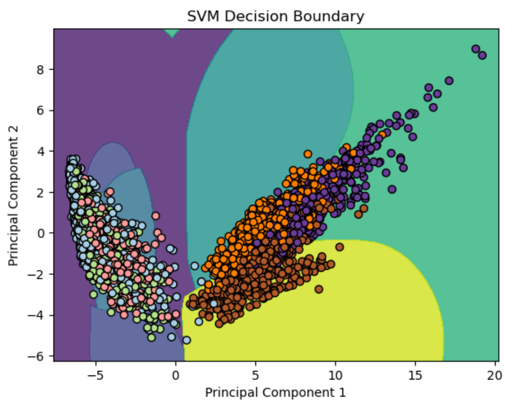

Figure 2: SVM Decision Boundry

6 Conclusion & Future Work

To conclude, our study examined the performance of various machine learning techniques, such as Random Forest, SVM, and XGBoost, within the scope of recognizing human activities.. Through rigorous preprocessing, we ensured data quality and employed dimensionality reduction via PCA to enhance computational efficiency and model performance. Our findings revealed that while Random Forest classifiers demonstrated high accuracy, they experienced difficulties distinguishing between certain activities, like sitting and standing, as evidenced by lower F1-scores in these categories. Conversely, the SVM showed exceptional classification prowess across all activity classes, underlined by near-perfect precision and recall statistics.

Our insights suggest that the integration of PCA with these classifiers yields a significant impact on performance, balancing between precision and computational complexity. Notably, the visualization of decision boundaries via PCA underscored the distinct separation capabilities of SVM in a reduced feature space, emphasizing its robustness as a classifier.

Looking forward, there are several exciting avenues for future research. We intend to explore advanced deep learning techniques, including Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs)[3]. These models are well-suited for better understanding the temporal and spatial relationships present in sensor data.. Additionally, investigating the impact of feature engineering, such as the creation of composite features or the use of signal processing techniques, could further enhance model performance. Furthermore, the deployment of unsupervised or semi-supervised learning methods could capitalize on unlabeled data, which is abundant in real-world scenarios. Lastly, we plan to examine the real-time deployment of these models in wearable devices, emphasizing the balance between accuracy and power consumption, crucial for the practical application of HAR systems.

7 Contributions

KAVYA KASALA was responsible for data collection, data preprocessing, and model training. She also covered these areas in the presentation and report writing. RITHIKA BAROOR handled the introduction, experiments, results, and conclusion sections. She addressed these same sections in both the presentation and the report writing.

8 References

1. Anguita, D. et al. "Human Activity Recognition on Smartphones Using a Multiclass Hardware-Friendly Support Vector Machine." International Workshop on Ambient Assisted Living and Home Care (2012).
2. Davide Anguita, Alessandro Ghio, Luca Oneto, Xavier Parra and Jorge L. Reyes-Ortiz. A Public Domain Dataset for Human Activity Recognition Using Smartphones. 21st European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning, ESANN 2013. Bruges, Belgium 24-26 April 2013.
3. Ronao, C. A., & Cho, S. B. (2016). Human Activity Recognition with Smartphone Sensors Using Deep Learning Neural Networks. ScienceDirect.