textOsFtextTOsFliningLFliningTLFtextosfliningIftabulartabproportionalprop

superiorSup

su-
pe-
ri-
or-
Sup

fontspechyperref

Comparative Analysis of Machine Learning Approaches for Human Activity Recognition

Machine Learning I - Final Project Report

Burak Cetin
University of Vigo
Spain
brakcetin660@gmail.com

Erik Figueiral Alonso
University of Vigo
Spain
erikfial1@gmail.com

Racim [Surname]
University of Vigo
Spain
racim...@col.udc.es

Roi Cores Cabaleiro
University of Vigo
Spain
roi.cores.cabaleiro@gmail.com

Abstract

This report presents a comparative study of four different machine learning pipelines for Human Activity Recognition (HAR) using the UCI Smartphone dataset. We investigate the effectiveness of dimensionality reduction techniques, including Principal Component Analysis (PCA) and Linear Discriminant Analysis (LDA), against traditional baseline models. Our experimental results demonstrate that the PCA-based Ensemble approach achieves a test accuracy of 93.32%, offering an optimal balance between computational efficiency and performance.

Keywords

Human Activity Recognition, PCA, Ensemble Learning, LDA, Machine Learning

1 Introduction

2 Methodology

In this study, each team member explored a different pipeline to solve the classification problem.

2.1 Approach 1: [Racim's Method]

In this approach, we evaluated classical machine learning models using the full set of 561 engineered features from the UCI HAR dataset. Unlike PCA or LDA, this method applies no dimensionality reduction; instead, it investigates whether the original high-dimensional representation already provides strong discriminative power for activity recognition.

2.1.1 Data Preprocessing. Because the HAR dataset contains features with heterogeneous numerical ranges, we standardized all inputs using Z-score normalization:

$$z = \frac{x - \mu}{\sigma} \tag{1}$$

This step ensures that models sensitive to scale—such as SVM, ANN, and kNN—operate on features with consistent magnitudes. No feature removal or compression was applied, making this a true full-feature baseline.

2.1.2 Model Optimization. We trained four supervised learning models using 5-Fold Stratified Cross-Validation (CV). The following configurations achieved the best performance for each model:

- **Artificial Neural Network (ANN):** A multilayer perceptron with architecture [128, 32] achieved a CV accuracy of 93.04%.
- **Support Vector Machine (SVM):** The RBF kernel with $C = 1.0$ reached a CV accuracy of 92.73%.
- **k-Nearest Neighbors (kNN):** The best performance was observed at $k = 3$ (CV: 91.65%).
- **Decision Tree (DT):** A depth-10 tree achieved 88.47% CV accuracy.

2.1.3 Ensemble Strategy. To improve robustness and reduce single-model biases, we created a **Voting Ensemble** combining ANN, SVM, and kNN. Majority voting improved generalization, particularly for ambiguous activities such as *Sitting* vs. *Standing*.

2.1.4 Performance Analysis. The models were evaluated on the strictly separated Test Set. Figure **??** shows the comparative accuracy between the five models. Quantitative results are summarized in Table **??**.

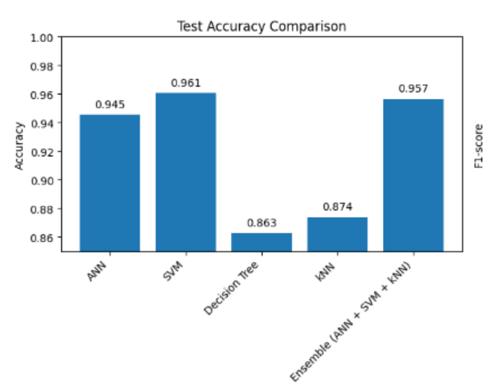

Figure 1: Test accuracy comparison of ANN, SVM, kNN, Decision Tree, and Ensemble models. The Ensemble model achieved the highest accuracy of 95.7%.

Table 1: Performance Comparison (Full-Feature Baseline Approach)

Model	CV Accuracy	Test Accuracy	F1-Score
Ensemble	-	**95.7%**	**0.957**
ANN	93.04%	94.5%	0.945
SVM (RBF)	92.73%	96.1%	0.960
kNN ($k = 3$)	91.65%	87.4%	0.873
Decision Tree	88.47%	86.3%	0.863

Figure **??** presents the confusion matrix of the Ensemble model. Consistent with related work, the primary confusions occur between *Sitting* and *Standing*, the two classes with the most similar motion patterns.

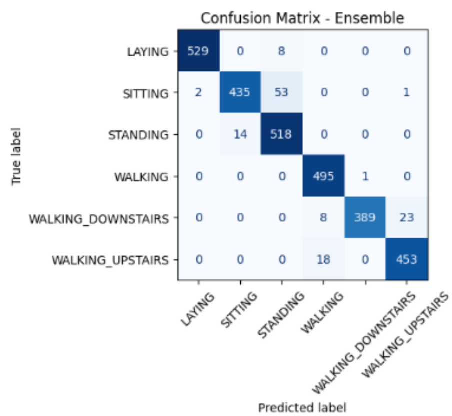

Figure 2: Confusion Matrix of the Full-Feature Ensemble Model. Slight confusion occurs mainly between Sitting and Standing.

2.1.5 Limitations and Future Work. Although the full-feature baseline achieved high performance, its computational cost is significantly higher than PCA or LDA approaches. Future improvements may include:

1. **Feature Selection:** Using mutual information or recursive elimination to reduce redundancy among the 561 features.
2. **Hybrid Dimensionality Reduction:** Combining PCA with a subset of raw features to maintain interpretability.
3. **Weighted Ensembles:** Giving more weight to ANN and SVM predictions to enhance static activity classification.

2.2 Approach 2: PCA & Ensemble Learning

In this approach, we investigated the trade-off between dimensionality reduction and classification accuracy. The primary objective was to design a computationally efficient pipeline suitable for embedded devices by compressing the feature space while maintaining high predictive performance.

2.2.1 Data Preprocessing and PCA. Principal Component Analysis (PCA) is inherently sensitive to the scale of input features. Since the UCI HAR dataset contains features with varying ranges, we first applied Z-Score normalization to standardize all 561 features to zero mean and unit variance:

$$z = \frac{x - \mu}{\sigma} \tag{2}$$

Following normalization, PCA was applied with a cumulative variance threshold of **95%**. This process reduced the original 561 dimensions to **102 principal components**, achieving an **81.8% compression rate**. This significant reduction drastically lowers the computational cost for subsequent training steps.

2.2.2 Model Optimization. We trained four distinct classifiers on the reduced feature space using 10-Fold Stratified Cross-Validation (CV) to ensure robustness. The hyperparameter tuning yielded the following optimal configurations:

- **Artificial Neural Network (ANN):** A multi-layer perceptron with two hidden layers ([200, 100] neurons) achieved the highest individual CV accuracy (97.31%).
- **Support Vector Machine (SVM):** Optimized with a *Linear Kernel* and $C = 1.0$, demonstrating that the PCA-transformed space is linearly separable (CV: 97.23%).
- **k-Nearest Neighbors (kNN):** The best performance was observed at $k = 1$ (CV: 96.54%), though this suggests potential susceptibility to noise (overfitting).
- **Decision Tree (DT):** Achieved only 84.47% CV accuracy. The poor performance is attributed to PCA's feature rotation, which disrupts the axis-aligned splits required by decision trees.

2.2.3 Ensemble Strategy. To mitigate individual model biases—specifically the overfitting of kNN and the decision boundaries of SVM—we constructed a **Voting Ensemble** comprising the top three models (ANN, SVM, and kNN). The ensemble aggregates predictions using majority voting to improve generalization on unseen subjects.

2.2.4 Performance Analysis. The models were evaluated on the strictly separated Test Set. To visualize the comparative performance, we computed accuracy metrics across all models, as illustrated in Figure **??**. The quantitative results are detailed in Table **??**, which compares the cross-validation (CV) and testing phases.

Table 2: Performance Comparison (PCA Approach)

Model	CV Accuracy	Test Accuracy	F1-Score
Ensemble	-	**93.32%**	**0.9329**
ANN	97.31%	93.04%	0.9302
SVM	97.23%	92.20%	0.9216
kNN ($k = 1$)	96.54%	84.39%	0.8437
Decision Tree	84.47%	78.38%	0.7836

While the individual kNN model suffered a significant drop in testing (indicating overfitting due to $k = 1$), the **Ensemble model** successfully compensated for this, achieving the highest overall accuracy of **93.32%**.

To analyze the specific error patterns, we examined the Confusion Matrix shown in Figure **??**. The matrix reveals that the model

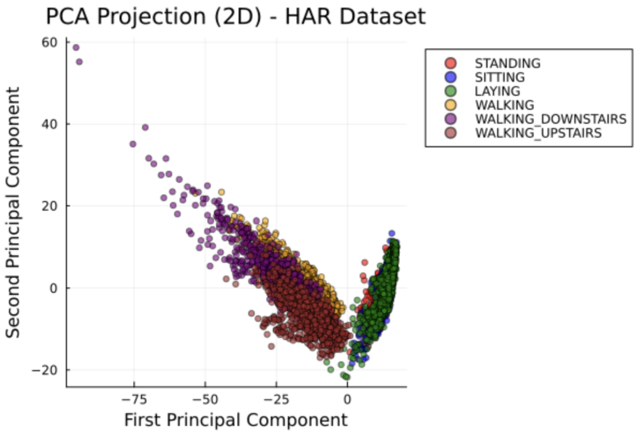

Figure 3: Test Accuracy Comparison. The Ensemble model outperforms individual classifiers.

is highly robust, yet slight confusions exist between static activities. To further validate this, Table **??** presents the detailed precision and recall values for each class.

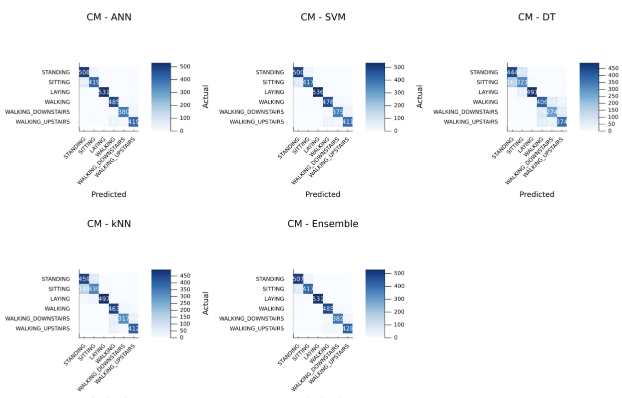

Figure 4: Confusion Matrix of the Ensemble Model. Note the slight confusion between Sitting and Standing.

Table 3: Per-Class Performance of the Best Model (Ensemble)

Activity	Precision	Recall	F1-Score	Support
Standing	0.8696	0.9530	0.9094	532
Sitting	0.9350	0.8493	0.8901	491
Laying	0.9981	0.9888	0.9935	537
Walking	0.9363	0.9778	0.9566	496
Walk Down	0.9409	0.9095	0.9249	420
Walk Up	0.9264	0.9087	0.9175	471
Weighted Avg			**93.32%**	**2947**

2.2.5 Limitations and Future Work. As confirmed by the confusion matrix, misclassifications primarily occurred between *Sitting* and *Standing*. Since PCA transforms the feature space, the explicit vertical orientation signal (gravity-z) may have been diluted. Future improvements include:

(1) **Explicit Gravity Injection:** Concatenating the raw ‘tGravityAcc-mean()-Z‘ feature with PCA components to restore orientation data.
(2) **k-Value Optimization:** Increasing k (e.g., to 7) in kNN to improve generalization on unseen subjects.

2.3 Approach 3: [Erik's Method]

2.4 Approach 4: [Roi's Method]

3 Experimental Results

4 Conclusion

Temporary page!

LaTeX was unable to guess the total number of pages correctly. As there was some unprocessed data that should have been added to the final page this extra page has been added to receive it.

If you rerun the document (without altering it) this surplus page will go away, because LaTeX now knows how many pages to expect for this document.