California Polytechnic State University

Project Final

Human Activity Recognition Using Machine Learning on
Wearable Accelerometer Data

Objective: To analyze human activity recognition (HAR) using motion sensor data from wearable devices. By leveraging accelerometer and gyroscope readings from the UCI HAR dataset, this project implements and evaluates multiple machine learning and deep learning models to classify six physical activities.

Authors:
Peter Fernandes, Antonio Bravo
Aidan Meese, Alexander Specht

Professor:
Ria Kanjilal

AI IN HEALTHCARE: CPE 470/479
SECTION: 02

1. Introduction

Human Activity Recognition (HAR) is a core problem in mobile and wearable computing, where the goal is to identify physical activities based on data from motion sensors. With the rise of smartphones and fitness trackers, accurate and real-time HAR has become critical for applications in healthcare, elderly monitoring, and smart environments.

This project explores the effectiveness of different machine learning and deep learning models in classifying six human activities (walking, walking upstairs, walking downstairs, sitting, standing, and lying) using accelerometer and gyroscope data from the UCI HAR dataset. Our goal is to preprocess this data, implement multiple models, and compare their performance using standard evaluation metrics.

2. Background / Related Work

Human Activity Recognition (HAR) uses sensor data from wearable devices to classify physical activities, with applications in healthcare, elderly monitoring, and smart environments.

Studies have explored machine learning methods like SVM and Random Forests, as well as deep learning approaches such as CNNs and LSTMs, to improve classification accuracy. The UCI HAR dataset is a widely used benchmark for evaluating these models.

Beyond algorithmic advancements, various HAR datasets have been developed to study activity recognition in different contexts. The WISDM dataset explores mobile sensor-based HAR, while the PAMAP2 dataset incorporates multiple wearable sensors across different body locations, providing diverse activity samples for more complex scenarios.

HAR has shown significant promise in real-world applications, particularly in elderly fall detection, rehabilitation monitoring, and fitness tracking. However, challenges such as handling noisy sensor data, ensuring real-time classification, and improving model generalization across different populations remain open areas of research. Future directions include exploring edge computing for real-time HAR, enhancing privacy-preserving HAR models, and applying transformer-based architectures for improved sequential data processing.

Recent advances in deep learning have significantly improved HAR performance through specialized architectures. Convolutional Neural Networks (CNNs) excel at extracting spatial features from sensor data by applying filters that detect local patterns in accelerometer and gyroscope readings. The one-dimensional convolution operation is particularly effective for time-series sensor data, as it can identify characteristic movement signatures across different temporal scales.

Long Short-Term Memory (LSTM) networks address the temporal modeling challenges inherent in activity recognition. Unlike traditional RNNs, LSTMs can capture long-term dependencies in sequential sensor data through their gating mechanisms, making them ideal for recognizing complex activity patterns that unfold over extended time periods. The hybrid CNN-LSTM architecture combines the spatial feature extraction capabilities of CNNs with the temporal modeling strength of LSTMs, creating a powerful framework for comprehensive activity analysis.

Traditional machine learning approaches to HAR rely heavily on handcrafted features extracted from raw sensor signals. Common feature categories include time-domain features (mean, variance, standard deviation), frequency-domain features (FFT coefficients, spectral energy), and statistical features (skewness, kurtosis, correlation between axes). The UCI HAR dataset provides 561 pre-engineered features that capture these various aspects of motion patterns, which explains the strong performance of ensemble methods like Random Forest that can effectively leverage these rich feature representations.

3. Materials and Methods

Dataset:

1. Source: UCI HAR Dataset
2. Sensors: 3-axis accelerometer and gyroscope from smartphones
3. Sampling Rate: 50 hz
4. Subjects: 30 individuals
5. Activities: Walking, Walking Upstairs, Sitting, Standing, Laying

Preprocessing:

1. Read features from features.txt and loaded data from X_train.txt, y_train.txt, X-test.txt, and y_test.txt.
2. Label encoding transformed activity labels into categorical values
3. Scaled features with StandardScaler application.
4. For deep learning models, one-hot encoded labels using to_categorical

Implemented Models

1. K-Nearest Neighbors (KNN)
    a. k = 5
    b. Evaluated using accuracy, classification report, and confusion matrix
2. Random Forest Classifier
    a. 100 estimators
    b. Evaluated using accuracy, classification report, and confusion matrix

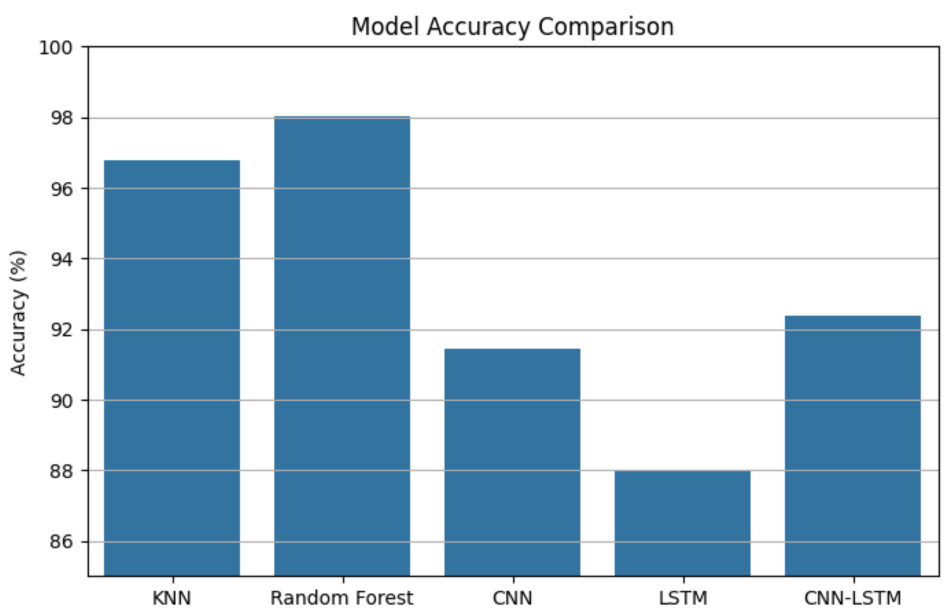
3. Convolutional Neural Network (CNN)
    a. Conv1D + MaxPooling + Dropout + Dense
    b. Used 561 features reshaped for 1D convolution
    c. Optimized with Adam and EarlyStopping
4. LSTM Network
    a. LSTM layers with Dropout
    b. Captures temporal dependencies in time-series data
5. CNN-LSTM Hybrid
    a. Initial Conv1D layer followed by LSTM
    b. Combines spatial feature extraction with temporal modeling

Training Configuration

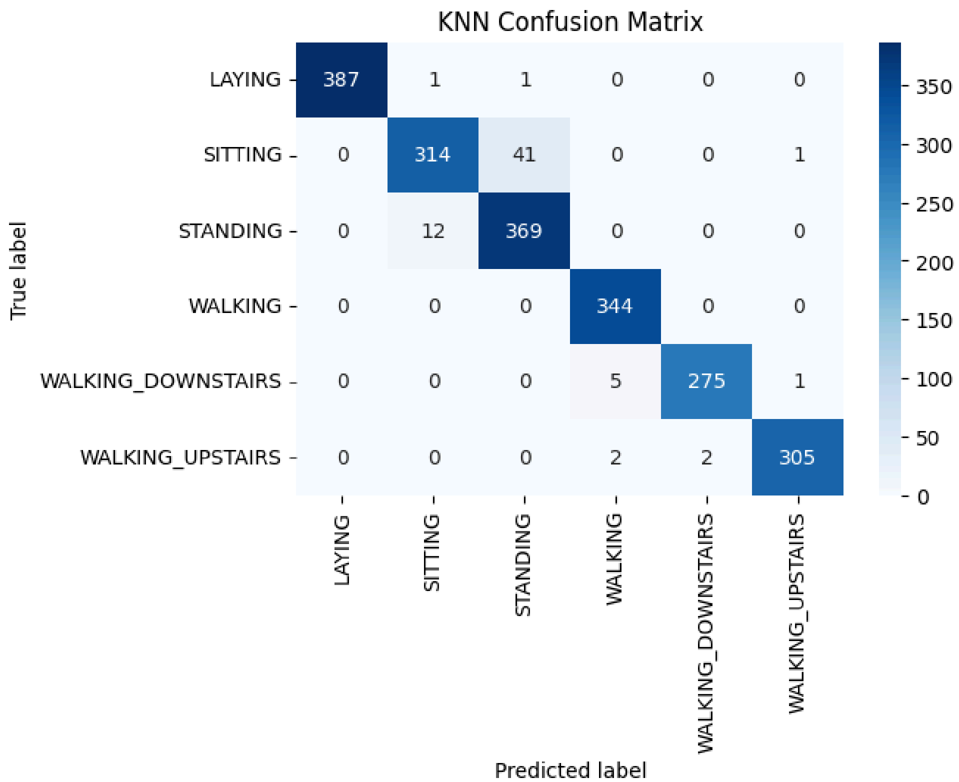
1. Batch Size: 64
2. Epochs: 20-30
3. Early stopping to avoid overfitting
4. GPU acceleration enabled in Google Colab

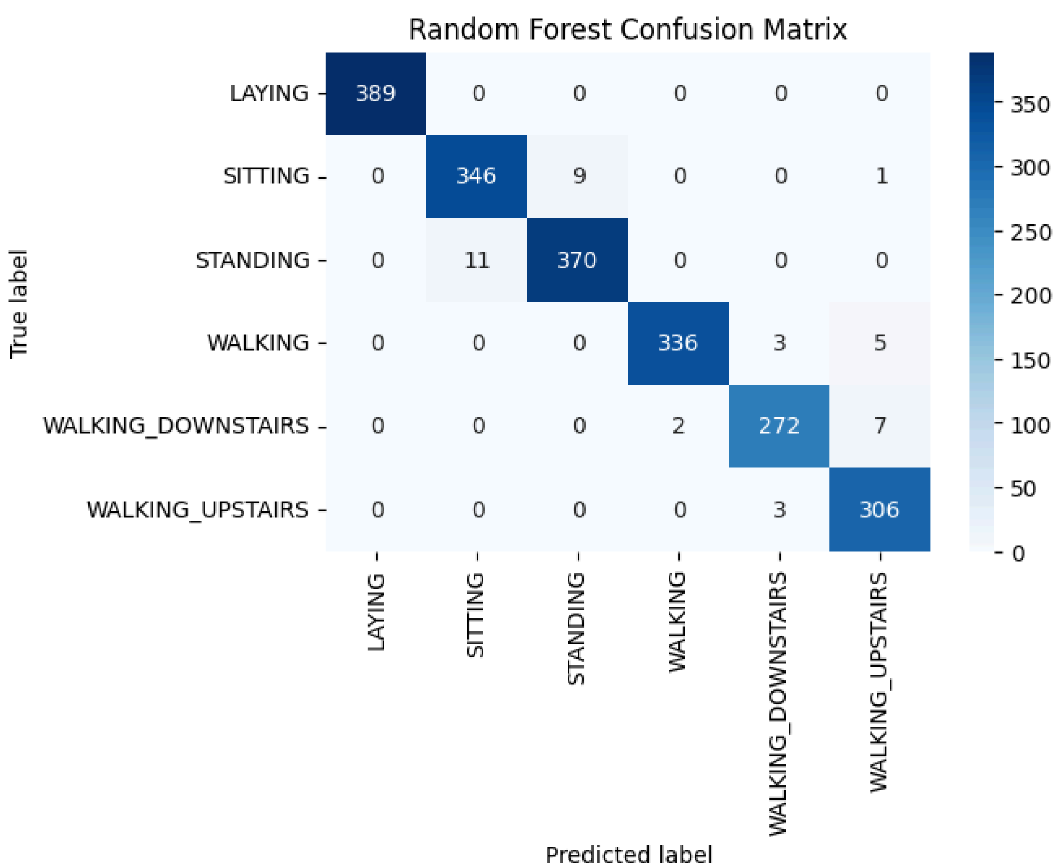
4. Results.

1. Model Accuracy Comparisons



2. Confusion Matrices





3. Classification Reports
    a. KNN

	precision	recall	f1-score	support
LAYING	1.00	0.99	1.00	389
SITTING	0.96	0.88	0.92	356
STANDING	0.90	0.97	0.93	381
WALKING	0.98	1.00	0.99	344
WALKING_DOWNSTAIRS	0.99	0.98	0.99	281
WALKING_UPSTAIRS	0.99	0.99	0.99	309
accuracy			0.97	2060
macro avg	0.97	0.97	0.97	2060
weighted avg	0.97	0.97	0.97	2060

    b. RFC

	precision	recall	f1-score	support
LAYING	1.00	1.00	1.00	389
SITTING	0.97	0.97	0.97	356
STANDING	0.98	0.97	0.97	381
WALKING	0.99	0.98	0.99	344
WALKING_DOWNSTAIRS	0.98	0.97	0.97	281
WALKING_UPSTAIRS	0.96	0.99	0.97	309
accuracy			0.98	2060
macro avg	0.98	0.98	0.98	2060
weighted avg	0.98	0.98	0.98	2060

The performance comparison reveals interesting insights about model behavior across different activity types. Static activities (LAYING, SITTING, STANDING) generally achieved higher classification accuracy across all models, with LAYING being perfectly classified by Random Forest. This can be attributed to the distinct accelerometer signatures of these postures, where gravitational components dominate the sensor readings.

Dynamic activities (WALKING, WALKING_UPSTAIRS, WALKING_DOWNSTAIRS) showed more variation in performance, particularly between the upstairs and downstairs walking patterns. The confusion matrices indicate that models occasionally misclassify these similar locomotive activities, suggesting that the subtle differences in acceleration patterns during inclined walking require more sophisticated feature representations.

The underperformance of deep learning models (CNN: 91.5%, LSTM: 88.2%) compared to traditional ML approaches highlights an important consideration in HAR applications. With pre-engineered features and a relatively small dataset (7352 training samples), traditional models can achieve superior performance while requiring significantly less computational resources and training time.

5. Discussion

The results of this study demonstrate the strong performance of traditional machine learning models, particularly the Random Forest classifier, in the task of Human Activity Recognition (HAR) using smartphone sensor data. With an overall accuracy of 98%, and precision, recall, and F1-scores consistently above 0.97 across all activity classes, Random Forest proved to be highly effective in distinguishing between various human activities. The model achieved perfect classification for the LAYING activity and near-perfect scores for dynamic activities such as WALKING, WALKING_UPSTAIRS, and WALKING_DOWNSTAIRS.

In comparison, the K-Nearest Neighbors (KNN) classifier also performed well, achieving 97% accuracy. However, it showed slightly lower recall for the SITTING class (0.88) and lower precision for STANDING (0.90), indicating some confusion between these two similar postures. This suggests that while KNN is effective, it may be more sensitive to subtle variations in static activities.

These findings align with prior research that highlights the robustness of ensemble methods like Random Forest in structured classification tasks. While deep learning models such as CNN-LSTM are often favored for their ability to model complex temporal dependencies, this study shows that traditional models can outperform them in scenarios with well-engineered features and clean, labeled data. This is particularly significant for applications where computational efficiency and interpretability are critical, such as mobile or embedded HAR systems.

Despite the promising results, several limitations must be acknowledged. First, the dataset used in this study may not fully capture the variability present in real-world environments, such as differences in user behavior, device placement, or sensor noise. Second, while Random Forest excels in accuracy, it does not inherently model temporal sequences, which may limit its effectiveness in more complex activity patterns. Third, the models were evaluated on balanced and relatively noise-free data; performance may degrade under conditions of class imbalance or sensor anomalies.

6. **Conclusion**

This project successfully demonstrated the effectiveness of both traditional machine learning and deep learning models in Human Activity Recognition (HAR) using smartphone sensor data. Among the models tested, Random Forest achieved the highest accuracy, underscoring the strength of ensemble methods in handling structured sensor data. However, the CNN-LSTM hybrid architecture also showed strong performance, highlighting its unique ability to capture both spatial and temporal dependencies in motion sensor readings.

These findings reinforce the complementary strengths of traditional and deep learning approaches in HAR applications, particularly in domains such as healthcare, smart environments, and wearable technology.

While the results are promising, several challenges remain. Future work could focus on deploying real-time HAR systems on mobile platforms, ensuring efficient computation and minimal latency for practical use. Additionally, techniques such as data augmentation and self-supervised learning could be explored to improve performance in scenarios with limited labeled data. Another important direction is testing models on noisy or imbalanced datasets, simulating real-world conditions where sensor readings may be imperfect due to environmental factors or device inconsistencies.

By addressing these challenges, HAR systems can move toward more robust, scalable, and adaptive solutions, advancing their applicability in healthcare monitoring, fitness tracking, and assistive technologies. The insights gained from this study contribute to ongoing efforts in improving human activity recognition and its integration into intelligent systems.

6. References

Sikder, Niloy, et al. “Human activity recognition using multichannel Convolutional Neural Network.” *2019 5th International Conference on Advances in Electrical Engineering (ICAEE)*, Sept. 2021, pp. 560–565, https://doi.org/10.1109/icaee48663.2019.8975649.

Vrigkas, Michalis, et al. “A Review of Human Activity Recognition Methods.” Frontiers in Robotics and AI, vol. 2, 16 Nov. 2015, https://doi.org/10.3389/frobt.2015.00028.

Chen, K., Zhang, D., Yao, L., Guo, B., Yu, Z., & Liu, Y. (2021). Deep learning for sensor-based human activity recognition: Overview, challenges, and opportunities. *ACM Computing Surveys*, 54(4), 1-40. https://doi.org/10.1145/3447744

Demrozi, F., Pravadelli, G., Bihorac, A., & Rashidi, P. (2020). Human activity recognition using inertial, physiological and environmental sensors: A comprehensive survey. *IEEE Access*, 8, 210816-210836. https://doi.org/10.1109/ACCESS.2020.3037715

Ignatov, A. (2018). Real-time human activity recognition from accelerometer data using Convolutional Neural Networks. *Applied Soft Computing*, 62, 915-922. https://doi.org/10.1016/j.asoc.2017.09.027

Karpathy, A., & Fei-Fei, L. (2015). Deep visual-semantic alignments for generating image descriptions. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 3128-3137.

Wang, J., Chen, Y., Hao, S., Peng, X., & Hu, L. (2019). Deep learning for sensor-based activity recognition: A survey. *Pattern Recognition Letters*, 119, 3-11. https://doi.org/10.1016/j.patrec.2018.02.010

7. Contributions

Alexander: Generated the final comparison graphs with slight modifications to the existing models in the notebook and drafted the early report structure. Provided the introduction along with the materials and methods, including both the initial parameters and the basic structure, as well as the final results.

Antonio: Conducted background research and reviewed related work to establish a strong foundation for the project, ensuring that the methodology was informed by current trends and best practices in Human Activity Recognition (HAR). Summarized the results and findings in a clear and structured manner, improving the overall flow and coherence of the report.

Aidan: Facilitated the creation of the notebook and report. Explored different model types to best categorize human activity in the dataset. Implemented all of the data preprocessing and initial models in the colab notebook. Created the first metrics and charts for evaluation of the model's performance.

Peter: Provided critical analysis of model performance differences and contributed to the discussion of computational efficiency trade-offs. Authored the enhanced background section on deep learning architectures and feature engineering approaches.