Journal on
Artificial Intelligence

Tech Science Press

https://doi.org/10.32604/jai.2026.080854

ARTICLE

Intra-Modal Feature Translation for Robust Kinematic Modeling in Wearable Health Applications

Ashwin Sankaran*

MSA Department, Central Michigan University, Mount Pleasant, MA, USA

*Corresponding Author: Ashwin Sankaran. Email: sanka1a@cmich.edu

Received: 16 February 2026; Accepted: 02 June 2026; Published: 29 July 2026

ABSTRACT: Human Motion Kinetics Recognition (HMKR) plays a pivotal role in advancing intelligent monitoring systems for public safety, machine interaction, and elderly-care technologies. Despite the increasing availability of affordable Inertial Measurement Units (IMUs), achieving accurate full-body motion analysis remains challenging due to sensor noise and variability across real-world environments. Addressing these limitations requires robust processing strategies capable of handling complex inertial time-series data. This research introduces an optimized HMKR system that leverages a comprehensive pipeline focusing on inertial sensor time-series data. In this research, we introduce an HMK-based recognition system that primarily focuses on six-degree-of-freedom (6-DoF) sensors that capture translational and rotational movements of the human body to improve health and safety applications. The study also performed a simulation study on two public benchmark datasets. The research's main contribution lies in integrating a preprocessing approach, including state-of-the-art denoising and a multi-stream feature set (statistical, frequency-domain, and vector features), that effectively mitigates sensor limitations and high-dimensionality issues via principal component analysis. These results justify the proposed system's reliability and its significance to substantially enhance health and safety monitoring applications, particularly in unconstrained public health contexts.

KEYWORDS: Feature space; human motion analysis; machine learning; classification; movement analysis; categorical boosting

1 Introduction

In recent times, the market has experienced an exponential increase in ubiquitous smart devices, such as smartphones and smart wearables, that operate on the Internet of Things (IoT) principle. A key feature of these devices is their integration of advanced sensors, which have opened new possibilities across various industries. This is particularly true in healthcare, where the widespread adoption of these sensor equipped devices enables the continuous monitoring of body signals. This new capability for data collection has led to rapid advancements in the healthcare sector, primarily by enabling the early detection of potential health risks [1]. For example, human motion kinematics recognition (HMKR) is the comprehensive process of selecting and detecting actions. HMKR involves the analysis of various human activities, which are classified as simple, complex, and postural activities. This technology is employed across diverse fields, including the diagnosis of medical conditions, public safety, the detection of sports injuries [2], and, critically, the care of elderly individuals [3,4]. These individuals often have health issues that demand ongoing, unobtrusive surveillance to identify and distinguish abnormal from normal motion kinetics [5].

The main aim of HMKR systems can be broadly summarized as follows: (a) utilising sensory observation data to find the present actions and activity patterns of individuals. (b) evaluating the ongoing action, such as the player's performance level and errors made during the exercise, given that the action is an indoor squat. (c) identifying specific attributes of the subject(s), including their gender, age, weight, and identity within the designated area.

In this study, intra-modal feature translation refers to the structured transformation of raw inertial sensor signals into multiple complementary feature representations within the same modality. Instead of relying on a single feature space, the proposed system translates the original signal into statistical, frequency-domain, and geometric descriptors, which are subsequently projected into a compact latent space using PCA. This process enhances discriminative capability while preserving essential motion dynamics.

The modality of sensory information implemented can be used to determine the classification of HMK systems, as the categories of features and algorithms are greatly influenced by the specific type of sensory data and architectures [6]. On a general scale, the subsequent research and development streams in HMK systems can be identified: (1) HMKR systems that rely on signals, images, and RGBD data, (2) HMKR systems that depend on inertial motion transmitted from sensors such as IMUs, (3) HMK systems that depend on infrared signals, and (4) HMKR systems that depend on audio streams [6].

To the best of our domain understanding, there is not a significant amount of work on integrating multiple modalities. In this research, we focused specifically on the second modality: inertial motion time-series data. This data type is highly valuable because of its intrinsic ability to monitor motion dynamics without being affected by unpredictable surroundings, such as a lack of light for visible capture. However, this inherent focus on internal motion is also its drawback; it makes the data less effective at capturing external context, such as the historical environment or interactions with other objects and individuals.

On the other hand, numerous machine learning and deep learning models utilize signals from multiple human body parts, including the heart, brain, chest, hands, wrists, and limbs, as documented in the liter-ature. The applications for camera-based systems are limited by their high cost and extensive space requirements compared to other accurate, cost effective sensors, such as wearable inertial measurement units (IMUs) [7].

Additionally, different data acquisition systems and methods rely on distinct sensory readings to identify these actions. The earliest approaches, for instance, collected accelerometer data using heavy, cumbersome devices [8]. However, the field rapidly advanced toward miniaturization: later systems utilized smaller integrated circuits connected to PDAs. These newer, portable systems allowed for the implementation of two primary modalities: camera-based computer vision systems and inertial sensor-based systems [9].

In a similar context, selective features of human body motion can be measured using a variety of sensors, including accelerometers, gyroscopes, compasses, and GPS devices. Sensors of this nature are effectively integrated into wearable devices [10,11]. For example, the use of the Xiaomi Mi Band for sleep monitoring and its impact on the daily lives of elderly individuals staying in nursing homes. Similarly, Apple Watch, Fitbit, and Garmin provide data that outperforms the basic step counts provided by pedometers [12,13]. Additionally, these companies also provide automated and visual feedback that is not present in conventional accelerometers.

The primary motivation and choice of methods used in this research are discussed as follows:

- The research broadly used 6-DoF sensors (including gyroscope and accelerometers) due to their ability to track human motion kinematics with greater accuracy and precision. In addition, these sensors are less affected by environmental factors such as lighting or obstructions, unlike other vision-based systems.

- The proposed system is tested on two benchmark datasets for evaluating system performance across both indoor and outdoor physical activities.
- Additionally, the research incorporates a denoising strategy to improve the quality of the data before feeding it into decision models.
- Furthermore, the study aims to support the deployment of HMKR systems that can effectively improve public health monitoring and elderly care.

This study presents a comparative analysis of HMKR in the context of public health safety using machine learning approaches. The main contributions of this research are as follows:

- We propose a unified multi-stream feature framework that systematically combines statistical, frequency-domain, correlational, and geometric (vector-based) descriptors. Unlike prior studies that rely on a limited subset of features, this design captures complementary motion characteristics across both temporal and spectral domains.
- We emphasize the role of preprocessing by integrating a denoising strategy (median filtering and windowing) as a core component of the pipeline, rather than treating it as a preliminary step. This allows us to explicitly study and demonstrate its impact on downstream classification performance.
- We introduce a consistent evaluation framework across multiple classifiers (logistic regression, k-NN, decision trees, MLP, and CatBoost) using identical feature representations. This enables a fair comparative analysis that highlights how different model classes respond to the same structured feature space.
- The pipeline is validated across two benchmark datasets (UCI-HAR and PAMAP2), covering both controlled and unconstrained activity settings, demonstrating robustness and generalizability.

The structure of this research is as follows: Section 2 reviews recent related work, Section 3 presents the overall system methodology, Section 4 evaluates the benchmark datasets and provides the comparative analysis, and Section 6 concludes the study with future insights into HMKR-based systems.

2 Related Work

Human motion kinetics (HMK) systems have attracted significant interest and development across a variety of application domains. Due to this growing significance, numerous review studies have focused on various aspects of the field [14]. For example, machine learning (ML) is an important part of AI, which is used to develop algorithms by analyzing data trends and historical relationships [15].

Data trends, whether labelled or unlabeled, are used to encode learning from past experiences. The target variable for labelled data may be either numerical or categorical. The tasks that involve machine learning include clustering for unlabeled data, analysis for numerical labels, and categorization for the categorical target output variable. The majority of ML research on wearable devices focuses on classification tasks, while only a small number are suitable for clustering [16]. A small number of them can be addressed as regression problems.

Researchers have focused on several topics, such as the identification of human daily activities, seizure detection, vital sign monitoring and forecasting [17], and fall recognition. Additionally, research has been conducted on the potential of wearable devices to detect stress, heart rate arrhythmia, and therapeutic tasks. Activity recognition enables healthcare professionals to assess a patient's health status by determining whether they can perform activities of daily living (ADLs). In Yen et al. [16], the authors conducted research on human activity recognition using convolutional neural networks. Nevertheless, the precision of machine learning algorithms for activity recognition in human subjects is significantly reduced when they are presented with a context with a distinct data distribution from the training data [17].

Liang et al. [18] employed a cross-subject transfer learning method to link source and target signals by developing feature-level manifolds, as personalized exercises may not be directly suitable as training data for another subject. An alternative approach to addressing this issue was explored by Compagnon et al. [19], who developed a personalized model for each subject. They noted that individuals engage in activities in various ways and that general models may average out specific important individual characteristics.

An activity recognition task may include fitness monitoring as an additional application. The research mainly estimated dynamic activities such as jogging, running, and walking using accelerometers and found that there is no tangible advantage to using accelerometers at both the hip and ankle locations compared to using only an accelerometer [20].

Kulsoom et al. [21] conducted a detailed analysis of HMKR, identifying the most effective ML algorithms, methods, and sensors for specific HMKR applications. They conclude that DL methods demonstrate superior accuracy and efficacy compared to conventional ML methods and emphasize the potential, limitations, and future directions in HAR. Gupta et al. [22] provided a thorough examination of HAR, with a particular emphasis on applications, AI, and acquisition devices. They suggest that the AI framework is synchronized with the development of HAR devices. Additionally, they recommend that researchers broaden the scope of HAR to encompass a range of domains and enhance human health and well-being.

Bian et al. [23] conducted a comprehensive study on the sensing modalities employed in HAR tasks, including the classification of human activities and sensing techniques. In another study, Dang et al. [24] provided a more in-depth discussion of the role of machine learning in HAR. The research performed a comprehensive HAR examination. The ongoing discussion regarding HAR technologies is enriched by their conclusion, which highlights current challenges and proposes directions for future research.

Table 1 summarizes the methods and contributions in literature.

Table 1: The key research focus and methods.

Reference	Research Focus	Methods Used	Contributions
Dunn et al. [15]	Wearable devices for health monitoring (e.g., fall detection, seizure detection, vital sign tracking)	Various wearable sensors	Discusses the potential of wearables in healthcare applications such as stress detection and arrhythmia monitoring.
Yen et al. [16]	Human Activity Recognition (HAR)	Convolutional Neural Networks (CNNs)	Identifies limitations of ML algorithms in new contexts with different data distributions.
Saeedi et al. [17]	Activity recognition in contexts with distinct data distributions	Machine Learning	Shows precision reduction when ML algorithms are presented with different context data.

(Continued)

Table 1 (continued)

Reference	Research Focus	Methods Used	Contributions
Liang et al. [18]	Cross-subject transfer learning in HAR	Transfer learning, Manifold development at the feature level	Develops a method for cross-subject HAR by linking source and target signals. Suggests personalized models can improve performance.
Compagnon et al. [19]	Personalized models for HAR	Personalized modeling	Argues that personalized models offer better performance due to individual differences and data privacy.
Kulsoom et al. [21]	Comprehensive review of HAR techniques	Analysis of various ML and DL methods	Concludes that DL methods outperform traditional ML in HAR tasks, highlighting future directions.
Gupta et al. [22]	Examination of HAR applications, AI, and acquisition devices	AI frameworks, Wearable devices	Recommends synchronizing AI development with HAR devices and broadening HAR applications.
Dang et al. [24]	Review of HAR technologies for IoT and healthcare applications	Classification of sensors, Feature Engineering, Data Collection, Training methods	Highlights challenges in HMKR and suggests possibilities for future research.

3 Materials and Methods

3.1 Datasets

We selected two benchmark datasets, UCI-HAR [25,26] and PAMAP2 [27], which are publicly available for training, validation, and prediction of human motion kinetics. Additionally, these datasets were chosen due to their diverse representation of human activity kinetics. The data were collected using various sensor modalities, including smartphones and inertial sensors, in both controlled and unconstrained environments.

3.1.1 UCI-HAR

The UCI-HAR dataset is generated by documenting 30 individuals aged 19–48 who were wearing a smartphone with integrated inertial sensors around their waists. Each activity protocol was followed by all subjects during the recording. Sitting, standing, lying, walking, walking downstairs, and walking upstairs are the six activities that must be identified in this work. These activities are the most frequently conducted in daily life. Tri-axial acceleration and tri-axial angular speed were used to capture the activity signals at a sampling rate of 50 Hz. Fig. 1 shows the rest and moving activities of UCI-HAR dataset.

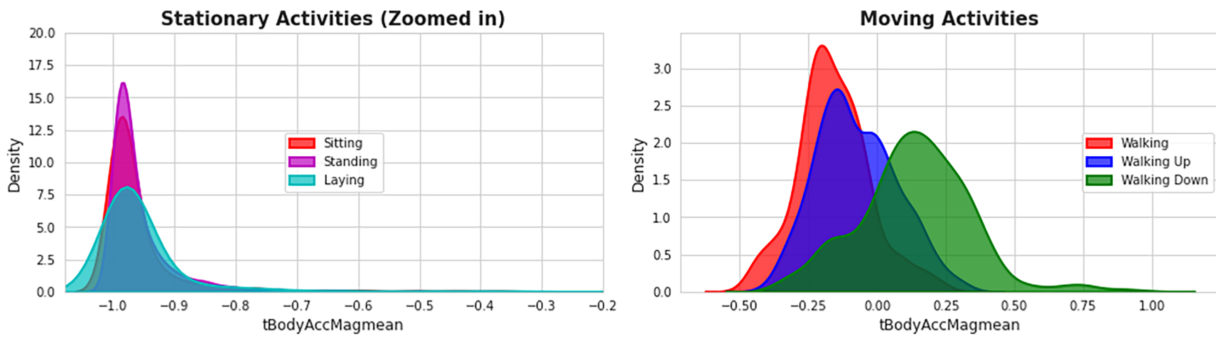


Figure 1: The stationary and moving activities of UCI-HAR database.

3.1.2 PAMAP2

The final dataset we used for testing was PAMAP2. Nine subjects participated in 18 physical activities. In addition, IMU sensors accelerometers, gyroscopes, and magnetometers were affixed to various body regions, including the wrist, chest, and ankle. However, we utilized our approach for 12 dynamic activities (see Fig. 2).

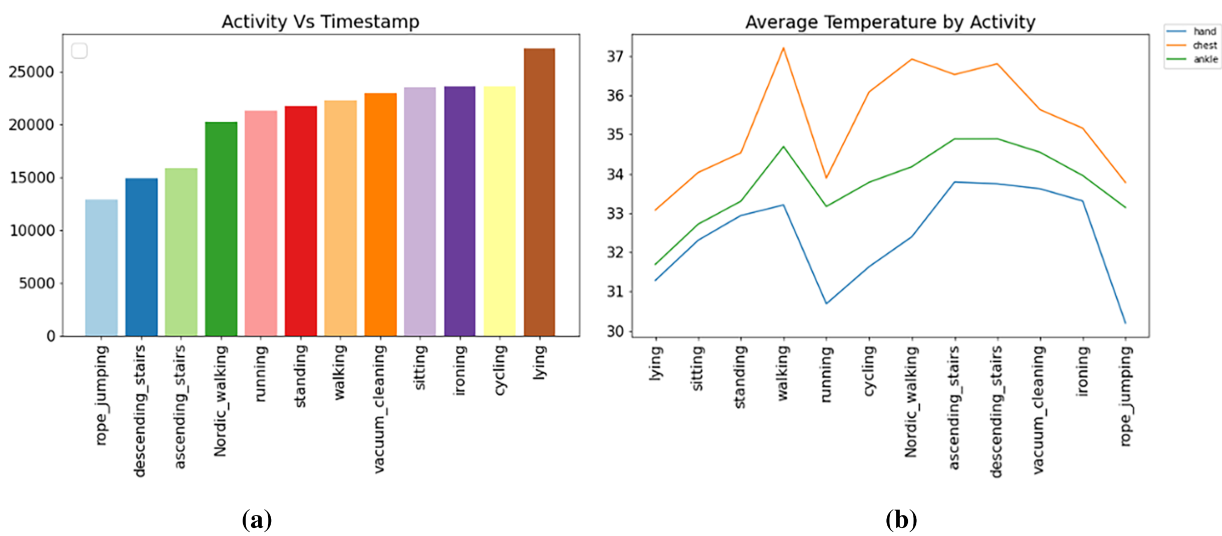


Figure 2: (**a**) Class distribution of PAMAP2 database and (**b**) Average temperature by activity.

3.2 Statistical Analysis

This section involves the overall strategical flow of the proposed system. The Fig. 3 shows the architecture of the proposed system.

3.2.1 Data Preprocessing

This section explains the denoising method, which include median filtering and windowing. Tri-axial angular velocity and linear acceleration signals were obtained from the triaxial inertial sensors in the PAMAP2 and UCI-HAR datasets. Signal processing is negatively affected by noise generated during data recording or acquisition. To obtain more precise results, it is necessary to remove disturbances from the data using filters. Denoising is one of the preprocessing stages used to enhance the outcomes of subsequent stages. The median filter's output is determined by calculating the mean of the input data within the window centered on the designated point (see Eq. (1)).

$$y[a] = \mathrm{med}\{\, z[i] \mid i \in w \,\} \tag{1}$$

where w represents a window centered at position a in the signal, and $y[a]$ denotes the filter output.

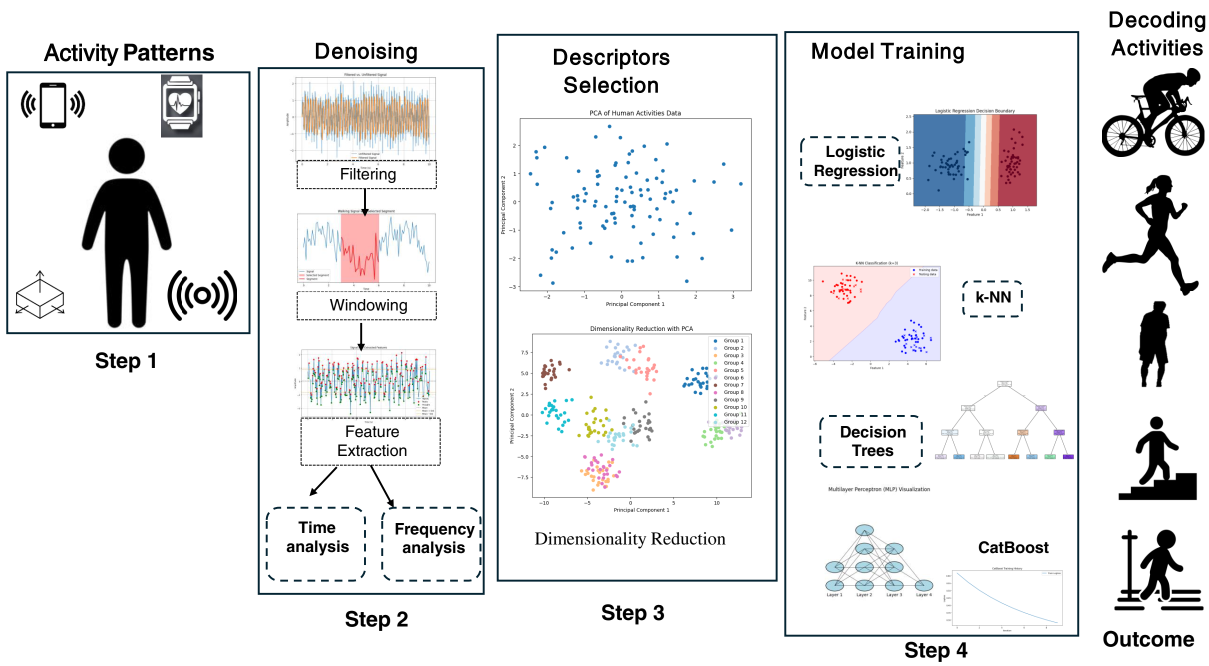


Figure 3: The flow of our proposed system.

3.2.2 Time Frequency Analysis

Signals with time-varying frequency components are analyzed in the time-frequency domain. It offers a method for comprehending the temporal evolution of a signal's spectral characteristics, a critical aspect of non-stationary signals. In the time-frequency plane, a scalogram is a visual representation of this analysis, generated by plotting the squared magnitudes of wavelet coefficients. This representation facilitates the detailed examination of intricate signals in a variety of disciplines, including finance, engineering, medicine, and HAR, by emphasizing the signal's frequency components as they develop. Using time-frequency-domain values for HAR rather than unprocessed data provides substantial benefits for feature extraction and classification. Raw sensor data, which is composed of time series data, commonly lacks to obtain frequency related information that is crucial for distinguishing activities.

The research adopted several features for exploration. We split all features into five streams: statistical features (mean, min, max, and standard deviation), signal-specific measures (energy and entropy), correlational descriptors, frequency-domain features (band energy and mean frequency), and vector features (geometric-based angle features). All the features collaboratively propose an effective set of cues for acquiring both temporal and spectral features of the inertial signal.

3.2.3 Principal Component Analysis (PCA)

Dimensionality reduction approaches are statistical methods that transform a data set into subspaces of reduced dimension derived from the original space [28]. These subspaces enable a more cost-effective description of the data. These techniques are important because the performance of numerous algorithms from a variety of disciplines, including numerical analysis, machine learning, and data mining, often deteriorates when used with high dimensional data. The curse of dimension is a term that describes the diverse phenomena that occur when data is analyzed and organized from multi-dimensional spaces. We can emphasize the significance of the following algorithms.

In this research, the data are evaluated using principal component analysis to identify previously unknown patterns and insights, as well as to eliminate excessive value correlation. PCA was applied to reduce the dimensionality of the features and to capture information about human activities in the frequency and time domains for classification [29]. Fig. 4 shows how PCA reduces the complexity of high dimensional data to two principal components. The results showed that PCA robustly captures the data's variance. For example, walking and walking upstairs are closer to each other, which shows similar movement patterns. Fig. 4 shows the PCA result on a feature vector.

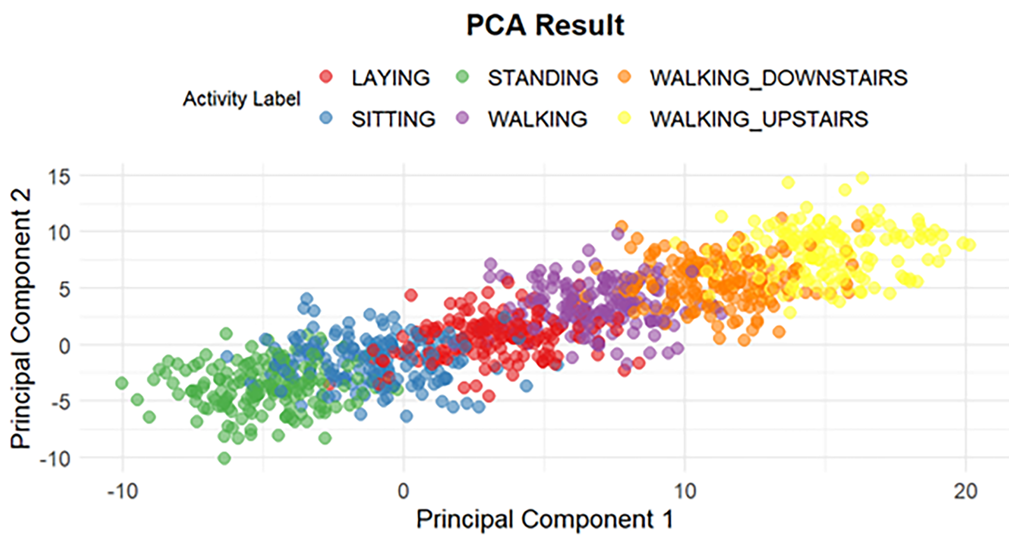


Figure 4: The PCA result over dynamic activities.

3.2.4 Pattern Classification

The last step involves classifying the HMKR in indoor and outdoor environments to improve overall public health safety. For this reason, we have used multiple classifiers to improve recognition performance on two benchmark datasets. The proposed system comprises five state-of-the-art classifiers: logistic regression [30], k-nearest neighbors (KNN) [31], decision trees [32], CatBoost [33], and a multi-layer perceptron (MLP) [34].

- **Logistic Regression**
  Logistic regression is a predictive modelling technique used to analyse the relationship between a binary dependent variable and one or more independent variables. It is commonly applied to describe data patterns and estimate the probability of class membership. When at least ten events per explanatory variable are available, the model provides stable and reliable parameter estimates.
- **Decision Trees**
  Decision trees represent decision processes through branching structures. Internal nodes test an attribute, branches indicate test outcomes, and leaf nodes correspond to class labels. Starting from a root node, the model applies the test function to the input and follows the appropriate branch until reaching a leaf node. The goal is to construct the smallest tree that accurately represents the decision boundaries.

- **CatBoost**
  CatBoost is a supervised machine learning algorithm designed for both classification and regression tasks. Originating from "categorical boosting," it is particularly effective for structured data containing categorical features and missing values. Its predictive strength comes from combining multiple weak learners into a robust ensemble. CatBoost incorporates advanced optimisation strategies, including ordered boosting, which reduces overfitting and enhances model accuracy—making it well suited for activity recognition in public health applications.
- **Multilayer Perceptron (MLP)**
  An MLP is a feedforward neural network composed of an input layer, one or more hidden layers, and an output layer. Information flows in a single direction without feedback loops. In fully connected networks, each neuron in one layer connects to every neuron in the next. With nonlinear activation functions and multiple hidden layers, MLPs can model complex relationships and outperform single-layer perceptrons.
- **k-Nearest Neighbours (kNN)**
  kNN is a non-parametric, instance-based learning method used for both classification and regression. Unlike model-based approaches, kNN does not build an explicit model during training. Instead, it makes predictions by analysing the structure of the training data when a new instance arrives. Because of this lazy-learning behaviour, kNN determines the output based on the majority class (or average value) among the k closest neighbours.

  Table 2 summarises the workflow of the five algorithms.

Table 2: Algorithmic workflow comparison across classifiers used for activity recognition.

Step	Logistic Regression	MLP	CatBoost	Decision Tree	k-NN
Initialization	Weight matrix W and bias b initialized.	Layers with weights and biases initialized (hidden + output).	Boosting model initialized with parameters (trees, depth, learning rate).	Root node created.	Training samples with labels stored.
Training Function	Gradient descent minimizes log-loss.	Backpropagation with Adam/SGD updates weights to reduce classification error.	Ordered gradient boosting: each tree fits residuals from the previous iteration.	Recursive splitting on features, maximizing information gain (Gini, entropy).	No explicit training; stores data for neighbour search.
Feature Handling	Uses PCA-reduced features directly.	Learns feature transformations via hidden layers.	Handles numeric/categorical features internally and reduces prediction shift.	Splits directly on raw features without transformation.	Computes distances in PCA-reduced feature space.
Prediction Function	Softmax: $\hat{y} = \arg\max \sigma(W\tilde{x} + b)$.	Forward pass through layers produces class probabilities.	Ensemble of boosted trees outputs class probabilities.	Follow decision path to leaf node → majority class.	Find k nearest neighbours; majority or weighted vote gives $\hat{y}$.
Evaluation	Accuracy, Precision, Recall, F1.	Same metrics.	Same metrics.	Same metrics.	Same metrics.

3.3 Assessment Metrics

In this study, we integrated the test and train datasets to combine a unified dataset for k-fold cross-validation. Instead of a fixed 70:30 split, the dataset was split into five equal folds. For training, we implemented five distinct classifiers on two benchmark datasets. Each of these five classifiers was trained using K-fold cross-validation on normalized data. After training, the models were evaluated using test folds, and predictions were made for each iteration. The key performance metrics considered were accuracy, precision, recall, and F1-score. To mitigate bias and ensure stable results, the experiment was repeated across multiple folds, and the mean accuracy was used as the final performance measure.

4 Comparative Analysis and Results

4.1 System Configuration

The whole experiment was employed on the machine HP 840 G6 Intel Core i7 vPro-8665U having 32 GB of RAM. In addition, Windows 11 pro was employed for the deployment and Jupyter Notebook (Anaconda 3) is utilized for simulations.

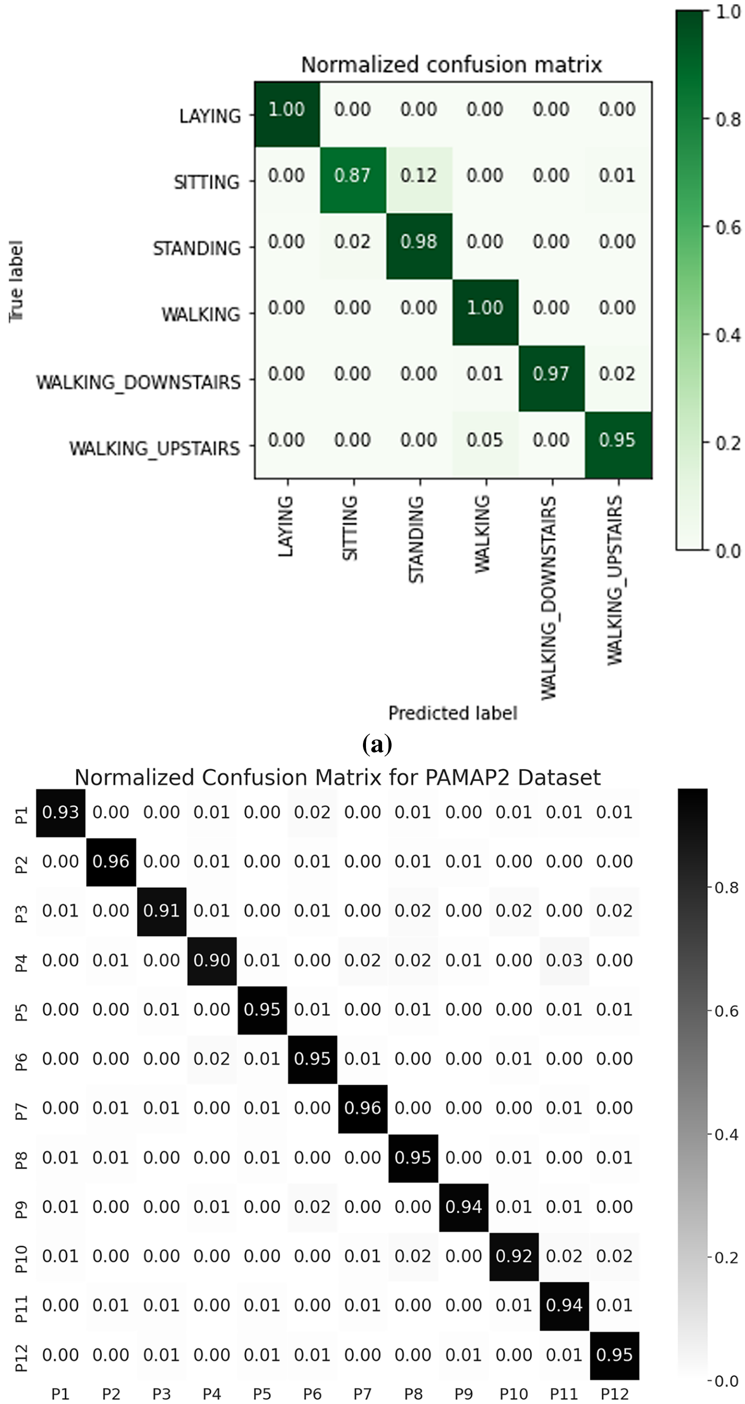
4.2 Experimental Analysis

This study's analytical framework offers a comprehensive approach to the primary domains of HAR in public health safety. However, it is essential to examine these domains thoroughly to ensure that the development of smart living environments addresses potential challenges and concerns. For evaluation metrics, the actual test labels are stored in y_{test}. The confusion matrix is computed using the function $\text{confusion\_matrix}(y_{\text{test}}, \text{all\_classifier\_pred})$, which compares the true labels y_{test} with the predicted labels all_classifier_pred. The confusion matrix function returns the total count of instances for each actual–predicted class pair. Fig. 5 illustrates the confusion matrix for the logistic regression classifier.

The Fig. 5a,b shows that most activities are classified accurately, however a few pairs exhibit notable confusion due to their similar motion patterns. In the UCI-HAR dataset, sitting is most frequently misclassified as standing, likely because both are low-movement postures that produce nearly static sensor readings. In similar context, minor confusion also appears between walking and stair-related activities, reflecting the shared rhythmic gait patterns that make these motions difficult to distinguish. On the other hand, PAMAP2 misclassifications are generally lower but occur between activity pairs with similar intensity or biomechanics. The results suggest that overlapping sensor signatures and natural variation in how individuals perform tasks contribute to these errors. The logistic regression model is fitted with tuned parameters based on cross-validation. The model was validated via 3 cross-validation sets. The regression model achieved accuracies of 94.6% and 93.83% on the UCI-HAR and PAMAP2 datasets, respectively.

Fig. 6 presents the confusion matrix of k-nearest neighbors.

In Fig. 6a,b, the k-NN classifier utilized 20 neighbors for the best performing configuration based on the cross-validation results. The model achieved 89.39% and 88.66% of recognition rate over two benchmark datasets. In addition, both confusion matrices show that similar misclassification patterns persist across both datasets. In the UCI-HAR dataset, sitting and standing remain the most frequently confused activities, with the model mislabeling 18% of sitting samples and 7% of standing samples, reflecting the difficulty of distinguishing low movement postures using inertial data. Likewise, stair related activities show higher confusion, particularly between walking downstairs and walking upstairs, as well as with regular walking, due to their overlapping gait dynamics. In the PAMAP2 dataset, diagonal accuracy decreases for several classes, indicating broader confusion among activities with similar motion intensity or biomechanics. These results highlight that activities involving subtle postural differences or comparable movement patterns remain the most challenging for classification.



(b)

Figure 5: (**a**) Confusion matrix for 6 and 12 classes over UCI-HAR (**b**) and PAMAP-2 database using logistic regression.

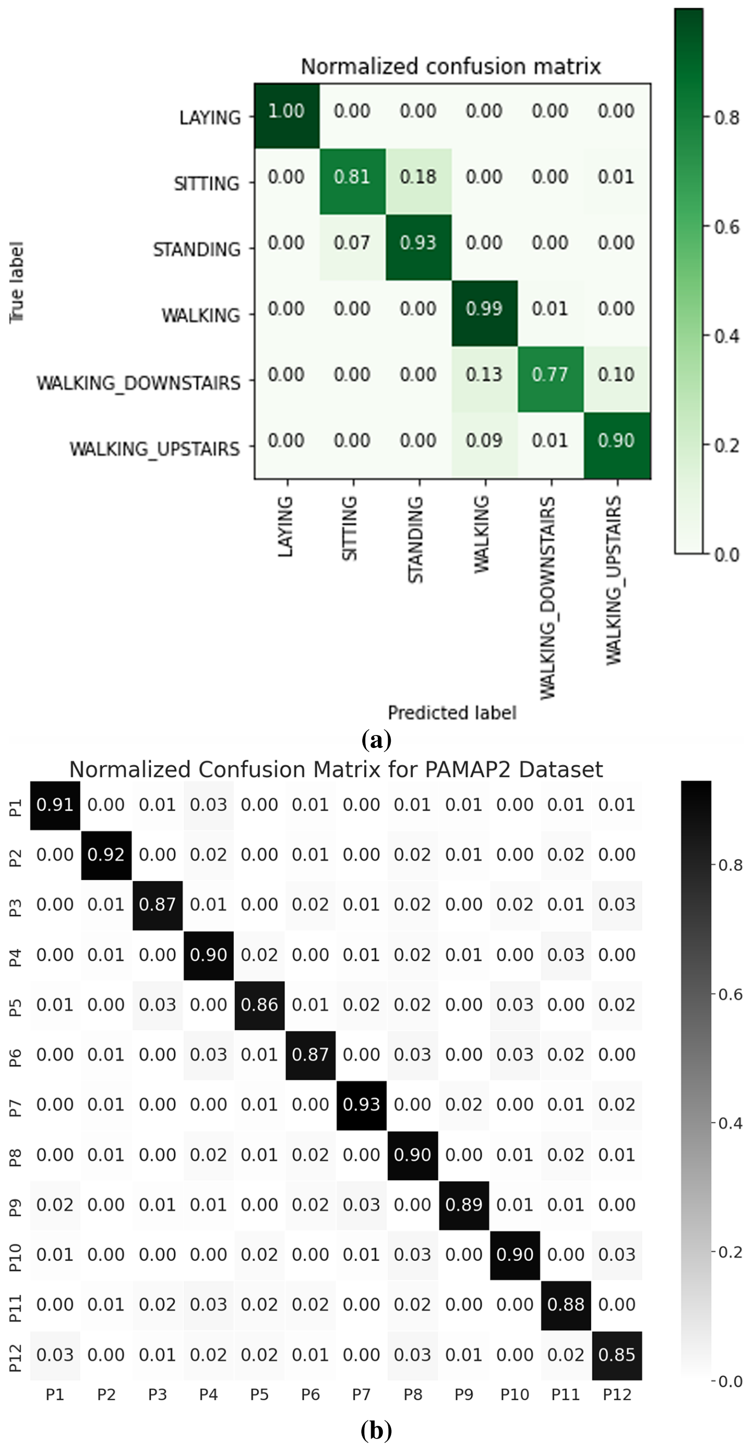


Figure 6: (**a**) Confusion matrix for 6 and 12 classes over UCI-HAR (**b**) and PAMAP-2 database using k-NN.

Fig. 7 presents the confusion matrix of decision trees.

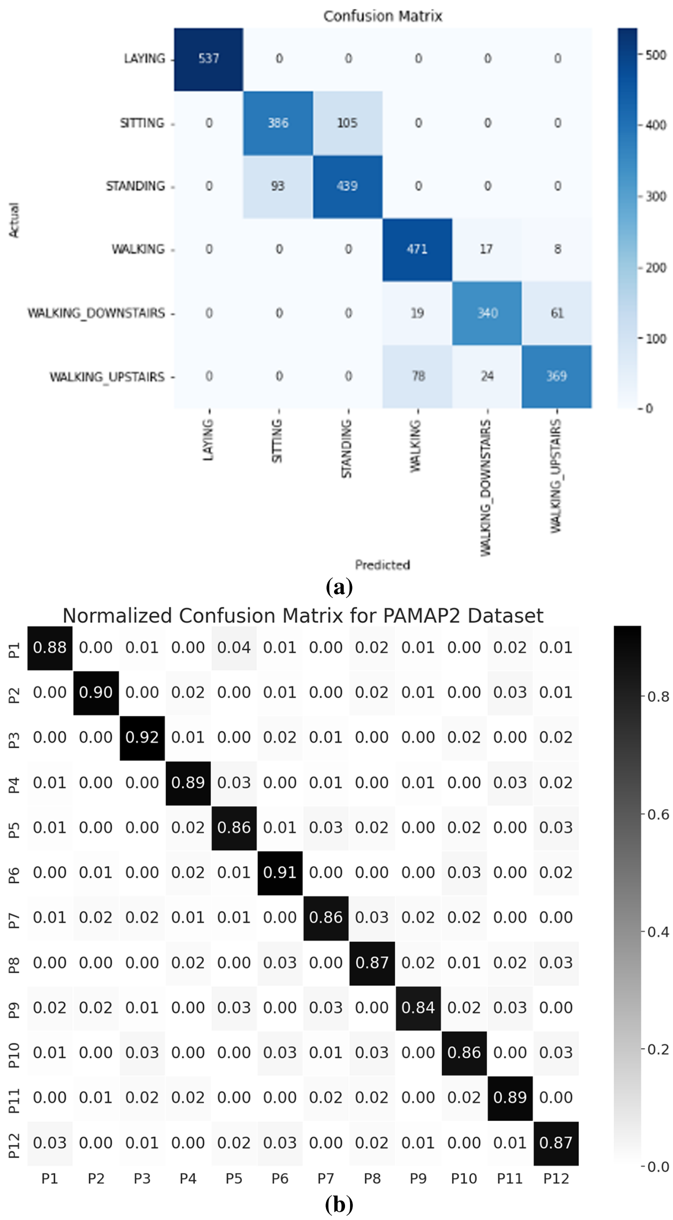


Figure 7: (**a**) Confusion matrix for 6 and 12 classes over UCI-HAR and (**b**) PAMAP-2 database using decision trees.

The decision tree model achieved an accuracy of 86.26% and 87.91% on the test set over two datasets. The model is performing better with a reasonable training time and better performance. However, the confusion matrices again show that misclassifications mainly occurs between activities with similar motion characteristics. In the HAR matrix, the majority of errors arise between sitting and standing, where many samples are incorrectly exchanged due to their low-movement nature and highly overlapping sensor patterns. Significant confusion is also observed between walking upstairs and walking downstairs, as well as between these stair-related activities and regular walking, reflecting the similarity in gait dynamics across these movements. In the PAMAP2 dataset, classification remains strong overall but shows scattered confusion among several activity pairs, especially those with comparable intensity. These results consistently indicate that activities producing similar inertial signatures remain the most challenging for the model to distinguish.

Fig. 8 presents the confusion matrix of Multilayer Perceptron.

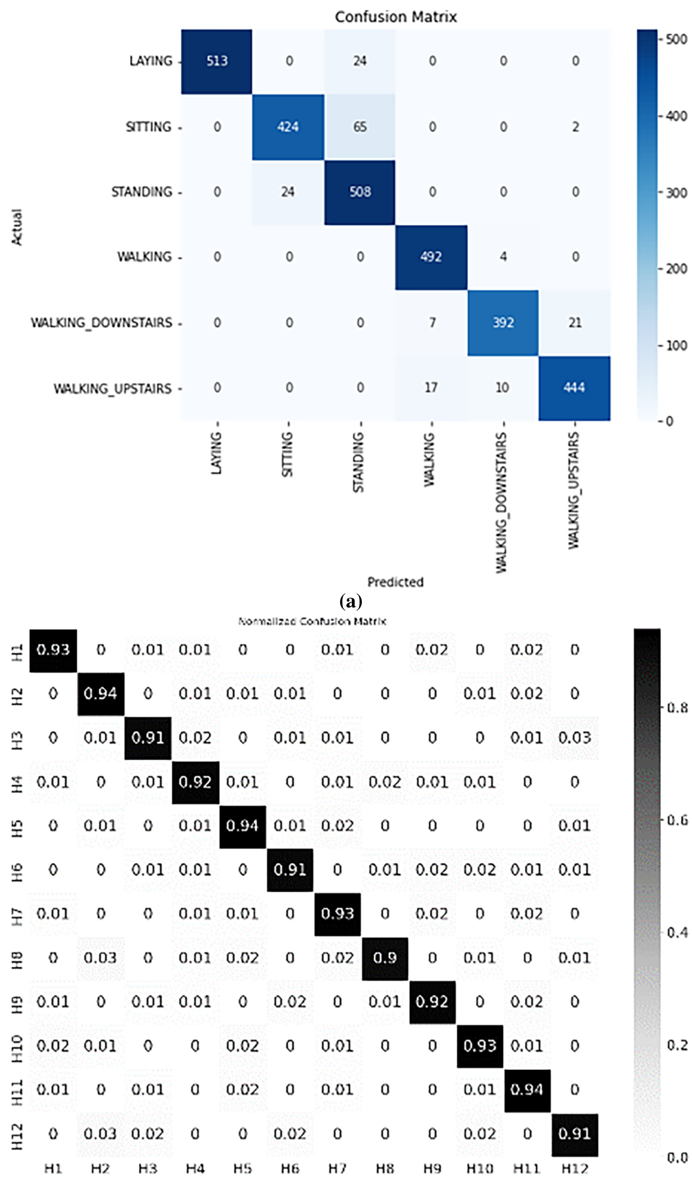
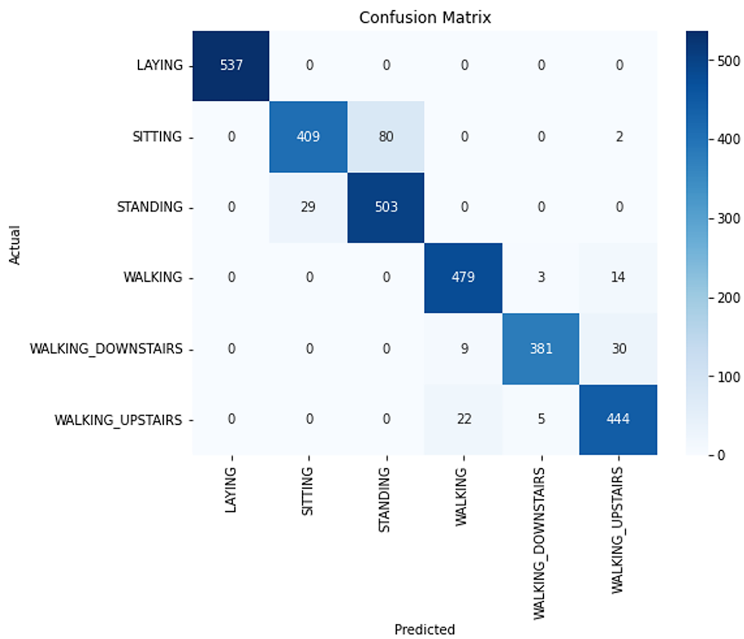


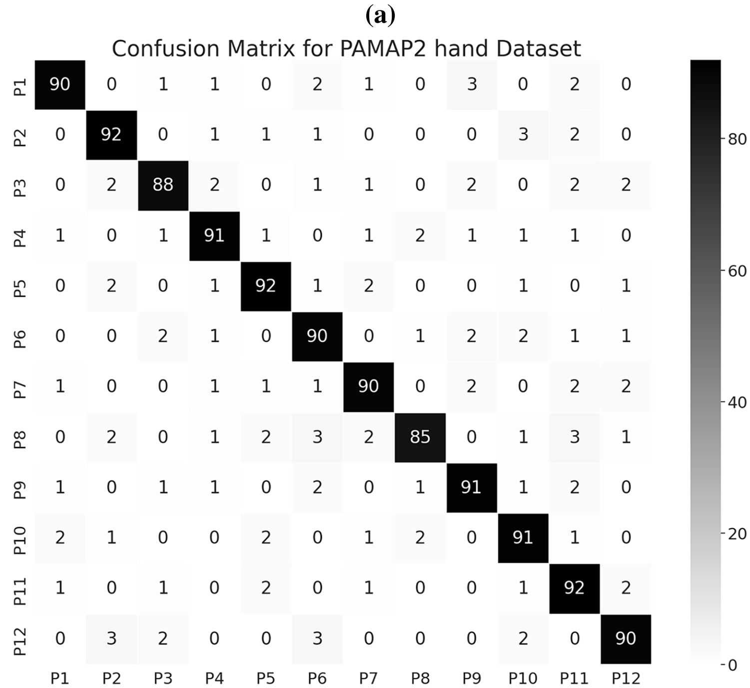
Figure 8: (**a**) Confusion matrix for 6 and 12 classes over UCI-HAR (**b**) and PAMAP-2 database using MLP.

The model achieved 94.10% and 92.30% accuracy on both datasets. The model shows strong clas-sification capability overall, although a few activity pairs still lead to noticeable errors. Most instances of laying, walking, and walking upstairs are recognized correctly, indicating that their motion signatures remain distinct across sensor channels. However, the model occasionally confuses sitting with standing, suggesting that this method also struggles to separate posture based activities that produce similar low intensity signals. A smaller degree of overlap is also evident between stair related movements, particularly where walking downstairs is misassigned to walking upstairs or level walking, reflecting the close resemblance in their dynamic patterns. In the PAMAP2 dataset results, diagonal values remain high, yet modest dispersion of predictions across neighboring classes suggests that activities with comparable movement style or intensity remain difficult to fully differentiate for this algorithm. Overall, the errors highlight that subtle biomechanical differences continue to challenge the classifier, even though general performance remains robust.

Finally, we applied the CatBoost classifier to evaluate performance across different levels of variance. Fig. 9 presents the CatBoost confusion matrix.



(a)



(b)

Figure 9: (**a**) Confusion matrix for 6 and 12 classes over UCI-HAR (**b**) and PAMAP-2 database using CatBoost.

The model attained accuracy of 93.42% and 90.16% on both benchmark datasets. Table 3 summarizes the key confusion trends across all algorithms, highlighting which activity pairs were most often mis-classified and which classes were consistently recognized accurately.

Table 3: Analysis of confusion patterns across all algorithms.

Class	Trend across Algorithms	Models	Reason
Sitting ↔ Standing	Most frequent confusion	Alg 1, Alg 2, Alg 3	Similar static postures with minimal movement.
Walking ↔ Walking Upstairs	Moderate confusion	Alg 2, Alg 3	Overlapping gait rhythm.
Walking ↔ Walking Downstairs	Moderate confusion	Alg 1, Alg 3	Similar vertical acceleration patterns.
Walking Upstairs ↔ Walking Downstairs	Occasional confusion	Alg 1	Similar motion style with slight intensity variation.
Laying	Best classified class	All algorithms	Distinct body orientation and stillness.
Walking	High accuracy	Alg 2, Alg 3	Strong rhythmic signature.

Fig. 10 shows the overall model performance for all classifiers.

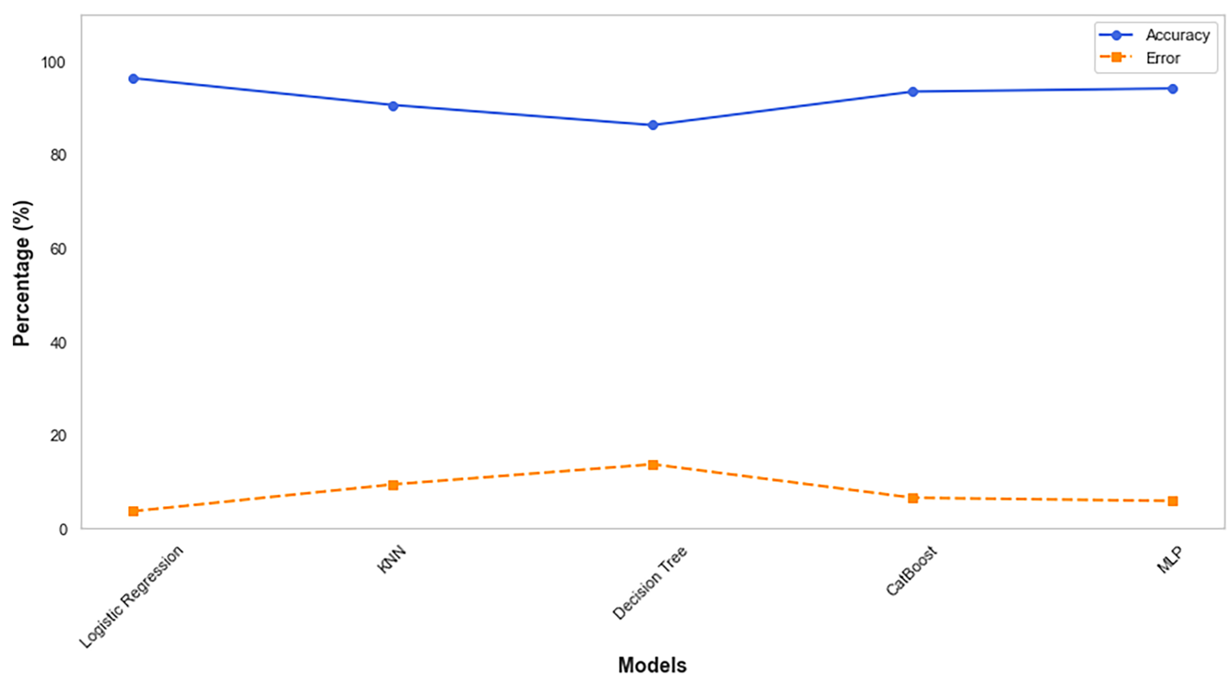


Figure 10: Model performance with error of margins.

Fig. 10 shows the performance of all the models, including their margin errors. The results show that logistic regression, MLP, and CatBoost performed well on the benchmark dataset and achieved better recognition than other methods, including decision trees and the k-NN classifier. In contrast, KNN and the Decision Tree model exhibit noticeably weaker performance, with low accuracy and elevated error values. The results reveal that both models have greater sensitivity to noise.

To evaluate the effectiveness of the proposed approach, we compare its performance with existing methods reported in the literature on the same benchmark datasets. For the PAMAP2 dataset, prior work using ensemble learning achieved 90.11% accuracy, while deep learning approaches such as attention-based CNN models reported 77.78% and BiLSTM-based methods achieved 64.10%. In contrast, the proposed method achieves 93.83% accuracy, outperforming several existing approaches.

Similarly, on the UCI-HAR dataset, SVM-based ensemble models and sequence learning approaches report accuracies of 88.20% and 88.50%, respectively, while personalized models achieve 81.00%. The proposed method achieves 94.60% accuracy, demonstrating competitive performance compared to existing methods (see Table 4).

Table 4: Comparison with existing HAR methods on UCI-HAR and PAMAP2.

Study	Method	Dataset	Accuracy (%)
Heroy et al. [35]	Bidirectional LSTM	PAMAP2	64.10
Murahari & Plötz [36]	Deep CNN (Attention-based)	PAMAP2	77.78
Priyadharshini et al. [37]	Ensemble Learning	PAMAP2	90.11
Batool et al. [38]	SVM + PSO (Actionlet Ensemble)	UCI-HAR	88.20
Compagnon et al. [39]	Sequence Metric Learning (RNN-based)	UCI-HAR	88.50
Compagnon et al. [19]	Personalized Model (SSMN)	UCI-HAR	81.00
Proposed Method	**Hierarchical Model**	**UCI-HAR**	**94.60**
–	–	**PAMAP2**	**93.83**

5 Ablation Study

5.1 Ablation Analysis of Preprocessing and PCA

To assess the contribution of individual components in the proposed pipeline, we conducted an ablation-style analysis focusing on preprocessing and dimensionality reduction. When denoising was not applied, model performance decreased due to the presence of sensor noise, which introduced variability in extracted features. The inclusion of median filtering improved signal stability and resulted in more consistent classification performance.

Similarly, removing PCA led to reduced performance and increased computational complexity due to high-dimensional feature space and redundant information (see Fig. 4). PCA improved performance by retaining the most informative components while reducing noise and correlation among features.

5.2 Feature Importance Analysis

The results indicate that frequency-domain features and energy-based measures contributed significantly to classification performance, particularly for dynamic activities such as walking and stair-related movements (see Table 5).

Table 5: Feature importance analysis.

Feature Group	Relative Importance (%)
Frequency-domain	32
Statistical	24
Energy/Entropy	18
Correlation features	14
Geometric features	12

Statistical features (e.g., mean and standard deviation) were effective in distinguishing static activities such as sitting and standing, while geometric features provided additional discriminative power by capturing motion orientation.

5.3 Classifier Performance Analysis

The comparative evaluation of classifiers reveals distinct performance characteristics. Logistic Regression and MLP achieved the highest accuracy, indicating that the feature space is well-structured and linearly separable to a certain extent, while also benefiting from non-linear modeling in the case of MLP.

CatBoost also demonstrated strong performance due to its ability to handle complex feature interactions and reduce overfitting through ordered boosting.

In contrast, k-NN exhibited lower performance, likely due to its sensitivity to noise and high-dimensional feature space, even after PCA. Similarly, Decision Trees showed comparatively lower accuracy due to their tendency to overfit and limited generalization capability.

These findings suggest that models capable of capturing both linear and non-linear relationships are better suited for the proposed feature representation.

6 Conclusion

This study effectively improved the efficiency of categorizing human activities using inertial data by optimizing sample sizes and window segments in logistic regression, MLP, CatBoost, decision trees, and kNN classifiers. The study found that logistic regression, MLP, and CatBoost classifiers outperformed decision trees and kNN in terms of accuracy. Additionally, the research revealed that PCA reduced the dimensionality of the raw data features and extracted the most important features for classifying human activities.

The experimental evaluation further demonstrates that the proposed approach achieves strong classification performance, with accuracies of 94.6% on the UCI-HAR dataset and 93.83% on PAMAP2. The ablation analysis confirms that preprocessing and PCA significantly enhance performance by reducing noise and feature redundancy. Moreover, feature importance analysis indicates that frequency-domain and statistical features contribute most to distinguishing between different activity classes.

However, this study had a few limitations. For example, the information content that IMUs provide is comparatively limited compared to that of vision-based systems from cameras, particularly in the context of HAR. In addition, certain activities that involve nuanced or complex movements, such as running or jogging, may not be accurately captured. Furthermore, sensor-based systems may be significantly more prone to errors and inaccuracies due to calibration issues or other technical factors.

Acknowledgement: Not applicable.

Funding Statement: The authors received no specific funding for this study.

Availability of Data and Materials: Data are openly available in public repositories. The datasets that support the findings of this study can be accessed from the UCI-HAR repository [25,26] and the PAMAP2 dataset [27].

Ethics Approval: Not applicable.

Conflicts of Interest: The authors declare no conflicts of interest.

Nomenclature

The key symbols, variables, and abbreviations used throughout the manuscript are given below

Symbol/Term	Description
$y[a]$	Output of the median filter at position a
$z[i]$	Input signal value at index i
w	Sliding window centered at position a
$X[n]$	Feature vector for sample n
β	Coefficient vector in the model
k	Number of neighbors in k-NN
PCA	Principal Component Analysis
IMU	Inertial Measurement Unit
HAR	Human Activity Recognition
HMK	Human Motion Kinetics
HMKR	Human Motion Kinetics Recognition
MLP	Multilayer Perceptron
k-NN	k-Nearest Neighbors
LR	Logistic Regression
DT	Decision Tree
CatBoost	Categorical Boosting algorithm

References

1. Akilandeswari K, Sivakumar NR, Alkahtani HK, Basheer S, Ghorashi SA. Smart healthcare activity recognition using statistical regression and intelligent learning. Comput Mater Contin. 2024;78(1):1189–205. doi:10.32604/cmc.2023.034815.
2. Liu C, Cao L. Automatic detection of sports injuries based on multimedia intelligent 3D images. Adv Multimed. 2023;2023(3):4180887. doi:10.1155/2023/4180887.
3. Demrozi F, Pravadelli G, Bihorac A, Rashidi P. Human activity recognition using inertial, physiological and environmental sensors: a comprehensive survey. IEEE Access. 2020;8(3):210816–36. doi:10.1109/access.2020.3037715.
4. Hosseinzadeh M, Koohpayehzadeh J, Ghafour MY, Ahmed AM, Asghari P, Souri A, et al. An elderly health monitoring system based on biological and behavioral indicators in Internet of Things. J Ambient Intell Humaniz Comput. 2023;14(5):5085–95. doi:10.1007/s12652-020-02579-7.
5. Rana SP, Dey M, Ghavami M, Dudley S. Markerless gait classification employing 3D IR-UWB physiological motion sensing. IEEE Sens J. 2022;22(7):6931–41. doi:10.1109/jsen.2022.3154092.
6. Ramanujam E, Perumal T, Padmavathi S. Human activity recognition with smartphone and wearable sensors using deep learning techniques: a review. IEEE Sens J. 2021;21(12):13029–40. doi:10.1109/jsen.2021.3069927.
7. Serpush F, Menhaj MB, Masoumi B, Karasfi B. Wearable sensor-based human activity recognition in the smart healthcare system. Comput Intell Neurosci. 2022;2022(5):1391906. doi:10.1155/2022/1391906.

8. Jalal A, Quaid MAK, Tahir SBUD, Kim K. A study of accelerometer and gyroscope measurements in physical life-log activities detection systems. Sensors. 2020;20(22):6670. doi:10.3390/s20226670.
9. Kumrai T, Korpela J, Maekawa T, Yu Y, Kanai R. Human activity recognition with deep reinforcement learning using the camera of a mobile robot. In: Proceedings of the 2020 IEEE International Conference on Pervasive Computing and Communications (PerCom); 2020 Mar 23–27; Austin, TX, USA. p. 1–10. doi:10.1109/percom45495.2020.9127376.
10. Capuano V, Xu L, Estrada Benavides J. Smartphone MEMS accelerometer and gyroscope measurement errors: laboratory testing and analysis of the effects on positioning performance. Sensors. 2023;23(17):7609. doi:10.3390/s23177609.
11. Mekruksavanich S, Jitpattanakul A. Sensor-based complex human activity recognition from smartwatch data using hybrid deep learning network. In: Proceedings of the 2021 36th International Technical Conference on Circuits/Systems, Computers and Communications (ITC-CSCC); 2021 Jun 27–30; Jeju, Republic of Korea. p. 1–4. doi:10.1109/itc-cscc52171.2021.9501477.
12. Bibbò L, Carotenuto R, Della Corte F. An overview of indoor localization system for human activity recognition (HAR) in healthcare. Sensors. 2022;22(21):8119. doi:10.3390/s22218119.
13. Nguyen B, Coelho Y, Bastos T, Krishnan S. Trends in human activity recognition with focus on machine learning and power requirements. Mach Learn Appl. 2021;5(10):100072. doi:10.1016/j.mlwa.2021.100072.
14. Logacjov A, Bach K, Kongsvold A, Bårdstu HB, Mork PJ. HARTH: a human activity recognition dataset for machine learning. Sensors. 2021;21(23):7853. doi:10.3390/s21237853.
15. Dunn J, Kidzinski L, Runge R, Witt D, Hicks JL, Schüssler-Fiorenza Rose SM, et al. Wearable sensors enable personalized predictions of clinical laboratory measurements. Nat Med. 2021;27(6):1105–12. doi:10.1038/s41591-021-01339-0.
16. Yen CT, Liao JX, Huang YK. Human daily activity recognition performed using wearable inertial sensors combined with deep learning algorithms. IEEE Access. 2020;8:174105–14. doi:10.1109/access.2020.3025938.
17. Saeedi R, Sasani K, Norgaard S, Gebremedhin AH. Personalized human activity recognition using wearables: a manifold learning-based knowledge transfer. In: Proceedings of the 2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC); 2018 Jul 18–21; Honolulu, HI, USA. p. 1193–6. doi:10.1109/embc.2018.8512533.
18. Liang Y, Feng K, Ren Z. Human activity recognition based on transformer via smart-phone sensors. In: Proceedings of the 2023 IEEE 3rd International Conference on Computer Communication and Artificial Intelligence (CCAI); 2023 May 26–28; Taiyuan, China. p. 267–71. doi:10.1109/ccai57533.2023.10201297.
19. Compagnon P, Lefebvre G, Duffner S, Garcia C. Learning personalized ADL recognition models from few raw data. Artif Intell Med. 2020;107(12):101916. doi:10.1016/j.artmed.2020.101916.
20. Abhishek K, Badar ud din Tahir S. Human verification over activity analysis via deep data mining. Comput Mater Contin. 2023;75(1):1391–409. doi:10.32604/cmc.2023.035894.
21. Kulsoom F, Narejo S, Mehmood Z, Chaudhry HN, Butt A, Bashir AK. A review of machine learning-based human activity recognition for diverse applications. Neural Comput Appl. 2022;34(21):18289–324. doi:10.1007/s00521-022-07665-9.
22. Gupta D, Singh AK, Gupta N, Vishwakarma DK. SDL-net: a combined CNN & RNN human activity recognition model. In: Proceedings of the 2023 International Conference in Advances in Power, Signal, and Information Technology (APSIT); 2023 Jun 9–11; Bhubaneswar, India. p. 1–5. doi:10.1109/apsit58554.2023.10201657.
23. Bian S, Liu M, Zhou B, Lukowicz P. The state-of-the-art sensing techniques in human activity recognition: a survey. Sensors. 2022;22(12):4596. doi:10.3390/s22124596.
24. Dang LM, Min K, Wang H, Piran MJ, Hee Lee C, Moon H. Sensor-based and vision-based human activity recognition: a comprehensive survey. Pattern Recognit. 2020;108(6):107561. doi:10.1016/j.patcog.2020.107561.
25. Anguita D, Ghio A, Oneto L, Parra X, Reyes-Ortiz JL. A public domain dataset for human activity recognition using smartphones. In: Proceedings of the ESANN 2013; 2013 Apr 24–26; Bruges, Belgium. p. 437–42.

26. Reyes-Ortiz J, Anguita D, Ghio A, Oneto L, Parra X. Human activity recognition using smartphones [Dataset]. UCI Machine Learning Repository. 2013 [cited 2026 Jan 1]. Available from: https://archive.ics.uci.edu/dataset/240/human+activity+recognition+using+smartphones.
27. Reiss A, Stricker D. Introducing a new benchmarked dataset for activity monitoring. In: Proceedings of the 2012 16th International Symposium on Wearable Computers; 2012 Jun 18–22; Newcastle, UK. p. 108–9. doi:10.1109/iswc.2012.13.
28. Zhao Y, Zhou H, Lu S, Liu Y, An X, Liu Q. Human activity recognition based on non-contact radar data and improved PCA method. Appl Sci. 2022;12(14):7124. doi:10.3390/app12147124.
29. Balli S, Sağbaş EA, Peker M. Human activity recognition from smart watch sensor data using a hybrid of principal component analysis and random forest algorithm. Meas Control. 2019;52(1–2):37–45. doi:10.1177/0020294018813692.
30. Hafeez S, Alotaibi SS, Alazeb A, Al Mudawi N, Kim W. Multi-sensor-based action monitoring and recognition via hybrid descriptors and logistic regression. IEEE Access. 2023;11:48145–57. doi:10.1109/access.2023.3275733.
31. Patro SGK, Mishra BK, Panda SK, Kumar R, Long HV, Taniar D, et al. A hybrid action-related K-nearest neighbour (HAR-KNN) approach for recommendation systems. IEEE Access. 2020;8:90978–91. doi:10.1109/access.2020.2994056.
32. Maswadi K, Ghani NA, Hamid S, Rasheed MB. Human activity classification using Decision Tree and Naïve Bayes classifiers. Multimed Tools Appl. 2021;80(14):21709–26. doi:10.1007/s11042-020-10447-x.
33. Chadha J, Jain A, Kumar Y, Modi N. Hybrid deep learning approaches for human activity recognition and postural transitions using mobile device sensors. SN Comput Sci. 2024;5(7):925. doi:10.1007/s42979-024-03300-7.
34. Wang S, Zhang L, Wang X, Huang W, Wu H, Song A. PatchHAR: a MLP-like architecture for efficient activity recognition using wearables. IEEE Trans Biom Behav Identity Sci. 2024;6(2):169–81. doi:10.1109/tbiom.2024.3354261.
35. Heroy AM, Gill Z, Sprague S. Stationary exercise classification using IMUs and deep learning. SMU Data Sci Rev. 2020;3(1):1. doi:10.21203/rs.3.rs-574901/v1.
36. Murahari VS, Plötz T. On attention models for human activity recognition. In: Proceedings of the 2018 ACM International Symposium on Wearable Computers; 2018 Oct 8–12; Singapore. p. 100–3. doi:10.1145/3267242.3267287.
37. Priyadharshini JMH, Kavitha S, Bharathi B. Classification and analysis of human activities. In: Proceedings of the 2017 International Conference on Communication and Signal Processing (ICCSP); 2017 Apr 6–8; Chennai, India. p. 1207–11. doi:10.1109/iccsp.2017.8286571.
38. Batool M, Jalal A, Kim K. Sensors technologies for human activity analysis based on SVM optimized by PSO algorithm. In: Proceedings of the 2019 International Conference on Applied and Engineering Mathematics (ICAEM); 2019 Aug 27–29; Taxila, Pakistan. p. 145–50. doi:10.1109/icaem.2019.8853770.
39. Compagnon P, Lefebvre G, Duffner S, Garcia C. Sequence metric learning as synchronization of recurrent neural networks. In: Proceedings of the 2021 International Joint Conference on Neural Networks (IJCNN); 2021 Jul 18–22; Shenzhen, China. p. 1–8. doi:10.1109/ijcnn52387.2021.9533627.