Available online at www.garmian.edu.krd

Passer Journal of Basic and Applied Sciences

Passer 7 (Issue 2) (2025) 1064-1075

Original Article

A Comparative Study of Machine Learning Techniques for Human Activity Recognition Using Smartphone Sensor Data

Ahmed Adil Nafea *[1], Mustafa Muslih Shwaysh [2], Yaqin Taher Al-Ani [3], Mohammed Salah Ibrahim [1], Ahmed Adnan Hadi [4], Aythem Khairi Kareem [5], Khattab M Ali Alheeti [6]

[1] *Department of Artificial Intelligence, College of Computer Science and IT, University of Anbar, Ramadi 31001, Iraq.*
[2] *College of Education for Humanities, University of Anbar, Ramadi 31001, Iraq.*
[3] *Department of Computer Sciences, College of Science, University of Al Maarif, Al Anbar,31001, Iraq.*
[4] *Artificial Intelligence Sciences Department, College of sciences, Al-Mustaqbal University 51001, Babil, Iraq.*
[5] *Department of Heet Education, General Directorate of Education in Anbar, Ministry of Education, Heet 31007 Anbar, Iraq.*
[6] *Department of Computer Networking Systems, College of Computer and Information Technology, University of Anbar, Anbar 31001, Iraq.*

Received 01 November 2024; revised 13 May 2025;
accepted 09 July 2025; available online 29 September 2025

DOI: 10.24271/PSR.2025.44486520.1798

ABSTRACT

Human Activity Recognition (HAR) helps in understanding human behavior through sensor data, particularly from smartphones. Despite advances in smartphone-based HAR, challenges such as high resource consumption and real-time processing remain an issue, often due to reliance on offline learning methods. This paper applied of different machine learning (ML) algorithms Support Vector Machines (SVM), Decision Trees (DT), Naive Bayes (NB), K-Nearest Neighbors (KNN), and Linear Discriminant Analysis (LDA) for human activity classification utilizing two datasets: UCI-HAR and Wild-SHARD. This work utilized k-fold cross-validation to evaluate model robustness and accuracy. The results shows that DT achieved the high accuracy of 98.35% on the UCI-HAR dataset, while KNN led to an accuracy of 96.86% on the Wild-SHARD dataset. Conversely, NB and LDA exhibited lower performance, highlighting their limited effectiveness in these scenarios. The findings underscore the significance of selecting suitable classification techniques to enhance the efficiency and accuracy of HAR systems, offering insights for future research and practical applications in real-time activity monitoring.

Keywords: Human activity recognition, Classification, Smartphone, Convolutional Neural Networks, Machine Learning, Deep Learning.

1 Introduction

HAR is the process of identifying and interpreting human actions through a sequence of observations and analyses of human behavior and the surrounding environment [1-3]. HAR has garnered significant attention due to its wide-ranging applications, including intelligent surveillance systems, anomaly detection, healthcare monitoring, smart homes, virtual reality interactions, and beyond [4-6]. Its ability to provide enhanced individualization and connectivity across various domains has made HAR a pivotal area of research within the computer science community [7,8].

Recent advancements in ML and Deep Learning (DL), supported via the evolution of sophisticated sensor technologies such as NVIDIA GPUs and Microsoft Kinect, have revolutionized HAR [9-11]. These technologies enable rapid identification of human behaviors utilizing different sensors, including motion detectors and video-based systems. Convolutional Neural Networks (CNNs) have emerged as a main tool in HAR due to their exceptional performance, although they often demand substantial computational resources, necessitating GPU utilization for efficient processing [12,13].

Traditional HAR systems have typically depended on specialized motion sensors attached to different body parts, such as the waist, chest, arms, and legs. While effective, these systems often involve complex setups that can be inconvenient and costly, potentially deterring users [14,15]. Technological advancements have begun to streamline these setups, often reducing them to a single smartphone, further enhancing user comfort and system accessibility.

As communication technologies evolve and big data becomes more prevalent, the complexity and heterogeneity of devices, infrastructure, and resources within HAR systems have increased. Managing, optimizing, and maintaining these systems now requires the integration of artificial intelligence, particularly ML

* Corresponding author
E-mail address ahmed.a.n@uoanbar.edu.iq (Instructor).
Peer-reviewed under the responsibility of the University of Garmian.

techniques. ML models, renowned for their predictive analysis and optimization capabilities, have shows effectiveness in addressing real-world challenges across different domains, including email spam filtering, face and object detection, traffic prediction, medical diagnosis, character recognition, and autonomous vehicles [16-21].

This paper proposed ML models for HAR, comparing the performance of the traditional techniques, including LDA, SVM, KNN, DT, and NB. Via evaluating these diverse approaches, the study aims to enhance activity recognition accuracy and provide valuable insights into the effectiveness of each method in improving HAR detection. A key contribution of our study is the integration of K-Fold Cross-Validation with these traditional ML models. This approach ensures a more robust evaluation by partitioning the data into multiple subsets and validating the models across different folds, mitigating issues like overfitting and providing a more reliable assessment of each model's generalizability. The use of K-Fold Cross-Validation strengthens the credibility of the results and demonstrates the potential of traditional ML techniques in achieving competitive performance for HAR tasks.

The work is organized as follows: The related work section reviews ML techniques utilized in HAR. The methodology section details of the proposed models. Finally, the results and discussion section present the findings, present insights and conclusions based on the comparative analysis.

2 Related Works

This part presents an analysis of earlier investigations on HAR and evolution techniques in improving the prediction technique's performance. The literature survey regarding the HAR characteristics and the review results are illustrated below:

This study [22] presented a DNN incorporates convolutional layers with Long Short-Term Memory (LSTM) approach. The presented approach automatically extracts exercise features and classifies them with an occasional instance parameter. LSTM is distinct from the Recurrent Neural Network (RNN), which is more appropriate for temporal sequences. Also, a Global Average Pooling layer (GAP) was involved to substitute the fully connected layer after convolution for decreasing model parameters. The results illustrate that the suggested approach has elevated robustness and more reasonable activity detection ability than some conveyed results.

This paper [2] investigated HAR methods in terms of surviving techniques, different sensing devices, and activity recognition mythologies. This approach classified HAR methods into three classifications: sensor-based, vision-based, and multimodal. The study presented the existent sensing devices for HAR techniques and their advantages and disadvantages in each method. It also discussed various sensor modalities and exercised detected by the sensor devices mentioned above. Then, this study investigated the effectiveness of fusion methods by utilizing ML algorithms.

This work [23] proposed a new DL model called Multi-Modal HAR Ensemble Network (MMHAR-EnsemNet), which constructs four distinct modalities to perpetrate sensor-based method. Two particular CNN are constructed for structure data. This approach is trained on LSTM and one CNN for RGB pictures. Gyroscope and raw accelerometer data are transformed to signal; formerly extra CNN model is applied. Ultimately, the outputs of the expressed pattern have been employed to construct a train to enhance the HAR approach's performance. The suggested approach has been investigated on two available datasets, Berkeley-MHAD and UTD-MHAD, including four various input data modalities. Experimental results demonstrate that this model has surpassed some recently suggested models regarded for comparison.

This paper [24] proposed a novel lower-expense learning method named NOvelty discrete data stream for HAR (NOHAR), concentrated on the continual data analysis flow. NOHAR is an online internet classification technique established on symbolic data developed by a discretization strategy. The benefits of this method are its ability to decrease and compress the data dimensionality. This investigation presents a new framework named DIscrete STream learning for Activity Recognition (DISTAR. The paper's objective includes normalizing the evolution of online techniques for symbolic data. The proposed model showed good performance that overcome state-of-the-art approaches.

This study [25] proposed a double-channel approach that includes LSTM, observed by a concentration instrument for the temporal integration of data equipment from inertial device coincident with a convolutional residue network for the spatial integration of sensor's data. This study also introduces an adaptive channel-squeezing process to fine-tune CNN feature extraction ability using multichannel dependence. Two commonly public and available datasets of HAR are employed in investigations to estimate the implementation of this approach. The results show that the presented strategy can overcome state-of-the-art approaches.

This work [26] analyzed sensorial Heterogeneous HAR (HHAR) system. The system utilized data into three-channel image illustrations, regaling the HHAR assignment as a classification of image issue. Since the present, CNN models have been computationally heavily used in Internet of Thing (IoT) applications. Therefore, this study suggests a lightweight prototype image encoded HHAR, named Multi-Scale image encoded HHAR (MS-IE-HHAR). The suggested study utilizes a Hierarchical Multi-Scale Extraction (HME) system observed by an Improved Spatial-wise and Channel-wise Attention (ISCA) module to construct the prominent architecture of model. The HME module has been created by a set of residually interconnected Shuffle Group Convolutions (SG-Conv) to learn and drag image illustrations from various sensor regions while decreasing the number of parameters. The presented study utilized two commonly available HHAR datasets: HHAR UCI and MHEALTH. The results show that the presented strategy overcome state-of-the-art approaches.

This work [27] proposed a generic HAR model for smartphone device sensors. This model used LSTM networks. Four model LSTM networks have been investigated to analyze the influence of utilizing various types of phone devices. In addition, a hybrid LSTM network named 4-layer CNN-LSTM is presented to enhance recognition HAR performance. The HAR approach is estimated on a public smartphone-based dataset of UCI-HAR via different varieties of model generation strategies. The testing results demonstrate that the performance in activity recognition improved accuracy.

This work [28] developed a compact wireless wearable detector node, which blends an air pressure detector and Inertial Measurement Unit (IMU) to supply multi-modal data for the HAR prototype exercise. This study suggests a new transfer learning model, a standard probability discipline adaptive form with enhanced pseudo-labels. This study adds the improved pseudo-label approach to avoid accumulative errors due to incorrect initial pseudo-labels. In order to confirm this approach and equipment, this study utilizes the newly developed sensor node to equip seven daily activity lives (ADL) of seven subjects. The experimental results demonstrate that the multi-modal method enhances the accuracy of the HAR system.

This paper [29] developed an inertial of smartphone accelerometer HAR architecture. When the person performs routine ADL, the smartphone acquires the sensorial raw data, extracts the appropriate features from the raw data, and conveys the person's physical behavior data via multiple x,y,z-axis accelerometers. Then the data are done by performing preprocessing operation, such as clearing the noise, normalization, and segmentation to produce useful features. In addition, a HAR technique established on a CNN is approaching. As shown in Table 1 a comparison based on the provided related works:

Table 1: A comparison of related work studies.

Ref	Proposed	App-Areas	Dataset Used	Computational Complexities	Advantages	Disadvantages
[22]	CNN and LSTM	HAR	UCI, WISDM, and OPPORTUNITY	High computational complexity due to multiple layers and parameters	Improved robustness and activity detection ability.	High model complexity; requires significant computation resources.
[2]	Classification of HAR methods based on sensor, vision, and multimodal fusion	HAR	Data Fusion	Low computational cost for classification	Comprehensive classification of HAR methods, sensor-based, vision-based, and multimodal.	Lack of specific model implementation; general discussion.
[23]	MMHAR-EnsemNet	HAR with sensor fusion	Berkeley-MHAD, UTD-MHAD	High computational complexity with multiple CNN and LSTM networks	Improved performance by combining multiple sensor modalities (accelerometer, gyroscope, RGB)	High computational load with multiple networks; complex model structure.
[24]	Novel discrete data stream method for HAR	HAR	UCI-HAR, and WISDM	Online processing with reduced dimensionality	Reduced data dimensionality; effective for continuous, real-time data.	Limited scope for real-world applications, lacks detailed dataset usage.
[25]	Double-channel LSTM and CNN for temporal and spatial integration	HAR	Public HAR datasets	Moderate computational load with dual-channel data integration	Improved performance in activity recognition through adaptive channel integration.	Limited to inertial data sensors; may not generalize to all activity types.

[26]	MS-IE-HHAR with CNN	Sensor fusion in HAR	HHAR UCI, MHEALTH	Lightweight model with reduced parameter complexity	Efficient, lightweight, and effective for heterogeneous sensor data in IoT applications.	Requires specialized hardware for efficient deployment; computationally intensive in real-time applications.
[27]	Generic HAR model with LSTM for smartphone sensors	Smartphone-based HAR	UCI-HAR	Computational complexity based on LSTM network with hybrid CNN	High accuracy using LSTM network; easy deployment on smartphones.	Limited by smartphone capabilities; LSTM model requires significant computation.
[28]	Transfer learning	HAR with wearable sensors	Custom developed multi-modal sensor data	Low computational cost; focuses on improving pseudo-label accuracy	Effective transfer learning method with improved accuracy in multi-modal data.	Requires specialized sensors; pseudo-label approach can lead to error accumulation if not properly managed.
[29]	Smartphone accelerometer-based HAR using CNN	Smartphone-based HAR	Real-life HAR dataset	Computational complexity with CNN for feature extraction	Effective HAR system with smartphone accelerometer data; CNN-based feature extraction.	Sensitive to noise and segmentation errors in accelerometer data.

Existing methods in HAR offer several advantages, such as high accuracy through DL models like CNN-LSTM, and the ability to process multimodal data for richer insights. Techniques like multi-modal fusion and online learning enable real-time processing and improved robustness. However, they also come with disadvantages, including high computational requirements, sensor dependency, and challenges in data fusion. Models like DL can be prone to overfitting, especially in limited data scenarios, and require substantial resources, which may limit their applicability in resource-constrained environments. The trade-off between accuracy and efficiency remains a key challenge for real-time HAR systems.

This paper is different from the other proposed studies in that it employs different data preprocessing methods such as data encoding, balancing, and standardization. In addition to that, we employed the k-fold cross-validation to assess model robustness and accuracy. This concept is different than other proposed methods in that it helps the ML model to learn on different instances of the data with less computation time and high accuracy results. The next section will explain all the steps implemented in our research.

3 Proposed Methodology

The main scope of this project is to implement a HAR System. The purpose of this system is to identify the action which the user would be doing, solely based on the changes in the motion of the user's body during the performance of specific actions. The project aims to develop a simple HAR prototype that only uses the built-in sensors found in an average smartphone and eliminates the use of a belt mount, allowing the user to carry their phone in their pockets. While this may result in less accurate predictions, it allows users to retain their habits, keeping their phones in their pockets. After the data is collected, it is pre-processed. Finally, the data is modelled and classified using five different supervised ML algorithms: SVM, DT, NB, KNN, and LDA as shows in Figure 1.

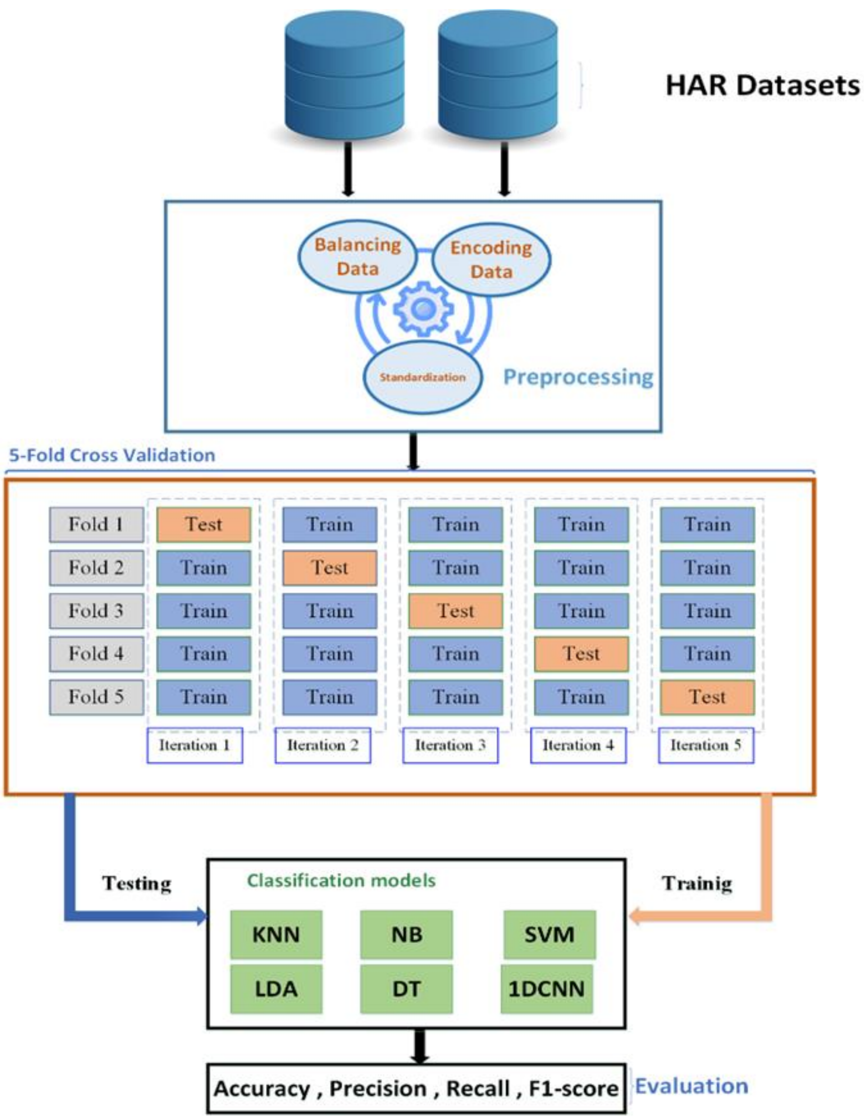


Figure 1: Proposed HAR system methodology.

3. 1 Dataset Description

3. 1. 1 Systems UCI-HAR Dataset Description

The proposed approach utilized publicly available datasets. This dataset is collected using the free smartphone application 'AndroSensor'[30]. This application allowed for data collection using the four main inertia sensors: gyroscope, gravity, accelerometer and linear acceleration. The activities are Cycling, Football, Swimming, Tennis, Jump Rope, walking, sitting, Walking Upstairs, Walking Downstairs, and Push-ups. The data collected involves roughly one hour worth of data for each of the 10 activities, allowing the model to be developed with an even distribution of data across all the categories.

The available dataset includes 60,070 samples, split into two testing and training sets. The dataset is imbalanced, as depicted in Figure (2).

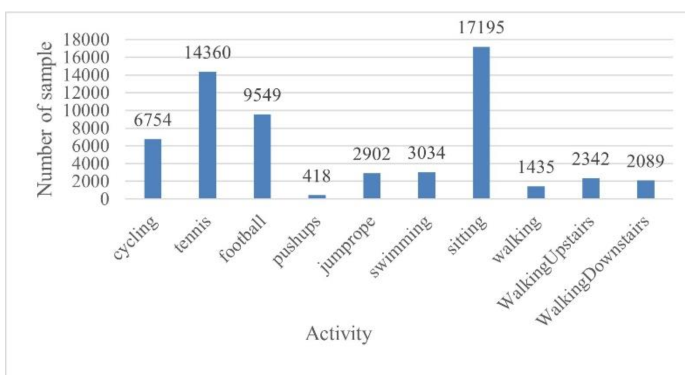


Figure 2: HAR dataset label distribution.

Figure 2 shows the imbalanced data where the sample number of sitting activity is 17195 while the sample number of pushups activity is 418.

The 'AndroSensor' application has more than one sensor, comprising an accelerometer, gravity, linear accelerometer, and gyrscope, each sensor employed to induce a time sequence of data for the experimentation. Each sensor generates a three-axial data sequence depending on the orientation of the human.

3. 1. 2 Wild-Shard Dataset Description

The proposed approach utilizes the 'Wild-SHARD' dataset. It introduces a practical HAR dataset acquired in a real-world (wild) environment to handle the constraints of simulated data [31,32]. It includes a time series of Activities of Daily Living (ADLs) captured using smartphone devices such as One Plus 9 Pro, Poco X2, Galaxy A30s, and Samsung Galaxy F62. These devices improve data robustness and variability with their various sensor manufacturers. The sensor module consists of gyroscopes and accelerometers of 40 volunteer subjects of various genders, ages, heights, and weights. Subjects performed activities naturally to capture authentic ADL data, including running, standing, walking, sitting, and navigating stairs outdoors and indoors. The dataset was collected at a selection frequency of 100 Hz, covering six ADLs: running, standing, walking, sitting, going downstairs and going upstairs. The characteristics documented in the dataset include linear acceleration, gravity, acceleration due to gravity, rotational vector, rotational rate, and the cosine of the rotational vector. The utilizing dataset includes 483,895 instances, each with sixteen features.

3.2 Preprocessing

3. 2. 1 Standardization

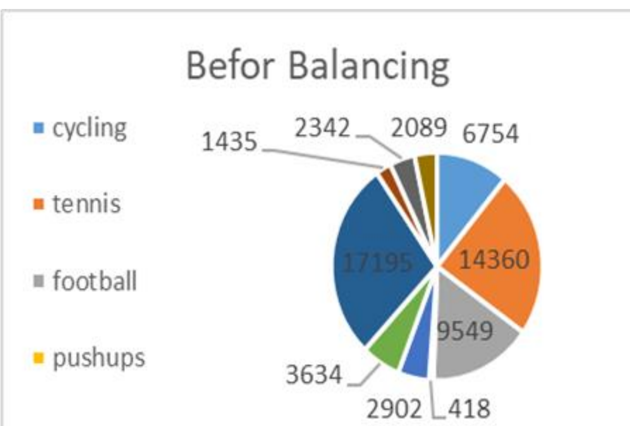
It's often implemented using tools such as StandardScaler in Python's scikit-learn library. It is a preprocessing step used to normalize the features of a dataset. This technique transforms the data to have a mean of zero and a standard deviation of one [21].

3. 2. 2 Encoding data

Encoding is a crucial preprocessing step for converting categorical data into a format that can be used by ML algorithms, which typically require numerical input. There are several encoding methods, each suited to different types of categorical data. This study utilized Label Encoding that Converts each category into a unique integer.

3. 2. 3 Balancing data

Imbalanced datasets are a relevant issue typically encountered in real-world applications that can notably influence the classification performance of ML techniques. The fundamental concept is to resample the original data. Two methods can be employed to perform this operation: Over-Sampling (OS) and Under-Sampling (US). Since the US may discard some meaningful data and consequently worsen the performance of the classifiers, OS tends to be preferred. Random oversampling, a simple process, expands the minority class by constructing randomly duplicated copies. Nevertheless, this approach can significantly increase the computational burden in large datasets. Also, as it generates exact duplicates of the juvenile class, it can potentially conduct overfitting. SMOTE, a synthetic juvenility oversampling method, offers an advantageous explanation. It doesn't just duplicate the juvenility class but synthetically produces new samples, thereby improving the dataset's diversity. By interpolating between several juvenility class observations that are 'close' to each other, SMOTE possesses new juvenility examples, potentially improving the model's performance. The proposed approach employs SMOTE method to perform balancing data. Figure (3) provides the segments for the presented model preprocessing result before and after data balancing for UCI-HAR while Figure (4) for Wild-SHARD.



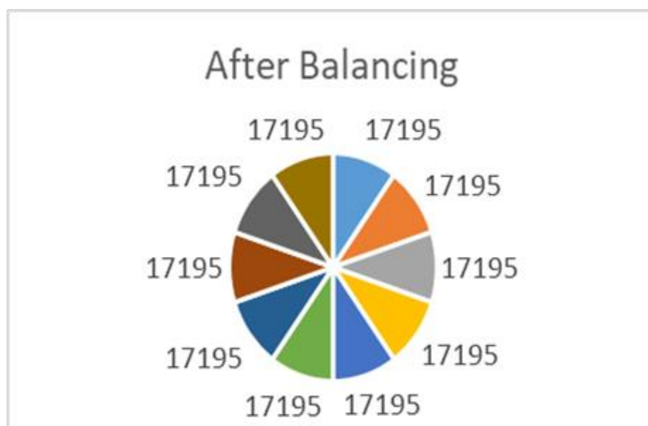


Figure 3: Before and after data balancing of UCI-HAR dataset.

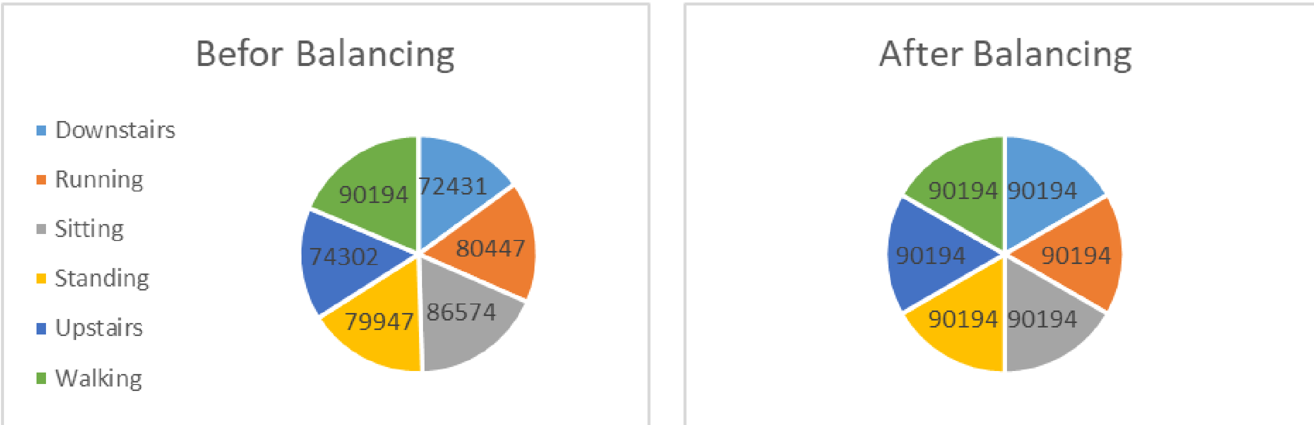


Figure 4: Before and after data balancing of Wild-Shard Dataset.

3. 2. 4 K-Fold Cross Validation

The traditional Train-Test Split process splits the dataset into just two parts: one for testing and the other for training. It is quick and simple to implement. However, the performance estimation can be highly dependent on the explicit split, leading to high variance in the results. K-Fold Cross-Validation (K-FCV) is a robust method for evaluating the performance of ML techniques. It is beneficial to confirm that the model generalizes unrecognized data well by utilizing various portions of the dataset for testing and training in multiple iterations. Figure (5) shows an overview of K-FCV. In k-FCV.

The dataset splits into K folds. A fold is utilized once in each iteration for testing, while the remaining folds are used for training. Therefore, the method is repetitive until all datasets are evaluated. KF-CV results are typically re-iterated with the average score. In this analysis, the value of k=5.

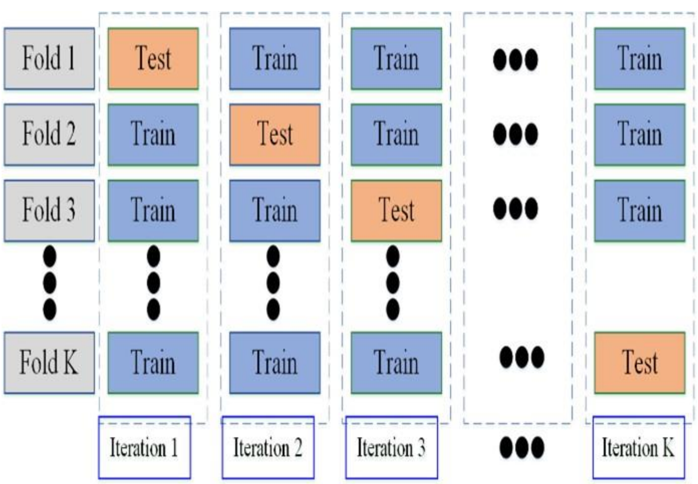


Figure 5: K-fold cross validation for our proposed HAR system.

3. 5 Machine Learning Models

In this study, we employed several ML models to analyze and predict outcomes based on the datasets. The following algorithms were implemented:

3. 5. 1 Support Vector Machine

It is a supervised learning model used for classification and regression tasks [33]. It works by finding the optimal hyperplane that maximizes the margin between different classes. In this research, we utilized the radial basis function (RBF) kernel to handle non-linear data, optimizing hyperparameters through grid search and cross-validation to enhance model performance.

3. 5. 2 Decision Trees

DT are intuitive models that split the dataset into subsets based on feature values, creating a tree-like structure of decisions [34]. We employed the CART (Classification and Regression Trees) algorithm for our decision tree implementation. The tree's depth and minimum samples per leaf were fine-tuned using cross-validation to prevent overfitting while ensuring that the model captured the underlying patterns in the data.

3. 5. 3 Linear Discriminant Analysis

LDA is a statistical method used for classification, which seeks to find a linear combination of features that best separates two or more classes [35]. This method assumes normal distribution of the data and equal covariance among classes. We implemented LDA to evaluate its effectiveness in distinguishing between the target categories, particularly in situations where class separation is essential.

3. 5. 4 K-Nearest Neighbors

KNN is a non-parametric algorithm used for classification and regression [36]. It classifies instances based on the majority label of their KNN in the feature space. We optimized the value of 'k' through cross-validation, experimenting with different distance metrics (Euclidean and Manhattan) to determine the optimal configuration for our dataset.

3. 5. 5 Naive Bayes

NB are based on Bayes' theorem and assume independence between features [37]. This model is particularly useful for text classification and situations with high dimensionality. We implemented the Gaussian NB variant due to the continuous nature of our data, evaluating model performance against other classifiers to assess its effectiveness.

3. 6 Evaluation Metrics

This section comprises the method, the metrics, and the deductions employed to evaluate the performance of the presented approach. The evaluation of the classification technique depends on the accuracy, F score, recall and precision the questions as a follow [38-40].

Accuracy: measures the overall proportion of correctly classified cases.

$$Accuracy = \frac{TP + TN}{TP + FP + FN + TN} \quad (1)$$

Precision: calculates the proportion of TP predictions relative to all positive predictions (both true and false).

$$Precision = \frac{TP}{TP + FP} \quad (2)$$

Recall: measures the proportion of TP identified by the model relative to all actual positives.

$$Recall = \frac{TP}{TP + FN} \quad (3)$$

F-measure: is the harmonic mean of precision and recall, providing a balanced measure when there is an uneven class distribution.

$$F - measure = \frac{2 \times Precision \times Recall}{Precision + Recall} \quad (4)$$

where the FP refers to False Positives, TP refers to True Positives, FN refers to False Negatives and TN refers to True Negatives.

4 Results

The evaluation of various ML techniques on two datasets provides a comprehensive view of their performance capabilities. For UCI-HAR dataset, DT achieved the highest accuracy at 98.35%, significantly outperforming other techniques. SVM also demonstrated strong performance, with accuracy of 95.88% and 95.95%, respectively. The NB got lowest accuracy at 86.72%, and LDA shows the least accuracy of 71.18%.

In Wild-Shard Dataset KNN emerged as the top performer, reaching an accuracy of 96.86%. DT also performed well but shows a decrease in accuracy to 92.33%. NB again underperformed with an accuracy of 64.18%, and LDA performance was notably poor, achieving only 41.92% accuracy. The variation in accuracy between the datasets highlights the influence of dataset characteristics on the effectiveness of different techniques.

The big drop in performance for NB and LDA across both datasets refer that these methods may not be suitable for the specific challenges presented by the data. NB, with its simplistic probabilistic approach, struggled particularly with Wild-Shard Dataset, where KNN outperformed other techniques. Similarly, LDA low accuracy in both datasets suggests it may not effectively capture the underlying patterns required for accurate classification.

Overall, DT and KNN demonstrated superior adaptability and performance across the datasets, with DT excelling in UCI-HAR and KNN in Wild-Shard Dataset. These findings suggest that while some techniques show remarkable effectiveness in specific contexts, others may require further optimization or alternative approaches to improve their utility in diverse data scenarios.

Table 2: UCI-HAR dataset results.

Technique	Accuracy	Precession	Recall	F1-score
SVM	95.95	90.77	89.85	90.11
DT	**98.35**	95.45	95.49	95.46
NB	86.72	79.23	82.28	78.69
KNN	93.8	87.8	87.73	87.82
LDA	71.18	50.35	57.85	51.46

Table 3: Wild-Shard Dataset results.

Technique	Accuracy	Precession	Recall	F1-score
SVM	89.13	88.73	88.62	88.64
DT	92.33	92.04	99.04	92.04
NB	64.18	64.77	63.76	63.49
KNN	**96.86**	96.77	96.74	96.75
LDA	41.92	41.24	41.6	41.01

5 Discussions

Table 4 summarizes key state-of-the-art HAR studies and their respective methodologies along with their accuracy on different datasets. Our proposed models, DT and KNN, achieved 98.35% accuracy on the UCI-HAR dataset and 96.86% on the Wild-SHARD dataset. When compared to other approaches, our DT model outperforms methods. This study contribution of integrating K-Fold Cross-Validation with traditional ML models enhances their reliability by ensuring robust performance evaluation between multiple subsets of the data, reducing the risk of overfitting and providing a more generalizable solution for HAR tasks.

In this part, we present analysis of the results got from our proposed method and compare them with state-of-the-art approaches, as shown in Table 3. Our proposed method, using DT and KNN, achieved an accuracy of 98.35% on the UCI-HAR dataset and 96.86% on the Wild-SHARD dataset. These results are highly competitive compared to existing methods in the literature. For instance, the LSTM model [25] achieved 97.70%, while CNN-LSTM [41] achieved 92.00%, indicating that our DT and KNN-based approach provides a significant improvement in terms of accuracy on the UCI-HAR dataset. Additionally, our approach outperformed models such as Bi-LSTM + CNN [42]

(97.96%, 97.15%) and XGBoost/LightGBM [43] (97.23%, 96.67%), demonstrating that traditional ML classifiers like DT and KNN can rival more complex DL models in specific scenarios.

However, there are notable limitations in the current state-of-the-art approaches. While many DL models like Bi-LSTM + CNN [42] and 1D-CNN [44] perform well across multiple datasets (e.g., UCI-HAR, WISDM, PAMAP2), they tend to require high computational resources, which can be a disadvantage for real-time applications or on devices with limited processing power. Moreover, models such as SDFL [45] and EDL [46], despite achieving good results (97.00% and 97.35%, respectively), still do not match the performance of our proposed method, which indicates that simpler models may offer better trade-offs between accuracy and efficiency.

Moving forward, future research could explore hybrid models that combine the strengths of both DL and traditional ML methods to enhance performance while reducing computational overhead. Additionally, addressing the generalization of models across different sensor types, datasets, and real-world scenarios will be crucial for improving the robustness and scalability of HAR systems. Another potential avenue for future work is the optimization of these models for real-time, low-latency activity recognition, which remains a significant challenge for existing methods.

Table 4: A comparison of results of proposed with state-of-the-art approaches.

Ref	Year	Proposed	Dataset used	Accuracy
[25]	2020	LSTM	UCI-HAR	97.70%
[41]	2020	CNN-LSTM	UCI-HAR	92.00%
[45]	2021	SDFL	UCI-HAR	97.00%
[46]	2022	EDL	UCI-HAR	97.35%
[42]	2023	Bi-LSTM + CNN	WISDM, UCI-HAR	97.96%, 97.15%
[44]	2024	1D-CNN	UCI-HAPT, WISDM, PAMAP2	98%, 97.8%, 90.00%
[47]	2024	RNN	UCI-HAR	97.00%
[43]	2024	XGBoost, LightGBM	UCI HAR and MotionSense datasets	97.23%, 96.67%
[48]	2024	FusionActNet	UCI HAR	97.35%
Proposed	2025	DT, KNN	UCI-HAR, WILD-SHARD	**98.35**%,96.86%

As shown in Figure (6) a comparation accuracy of the proposed system with state-of-the-art methods.

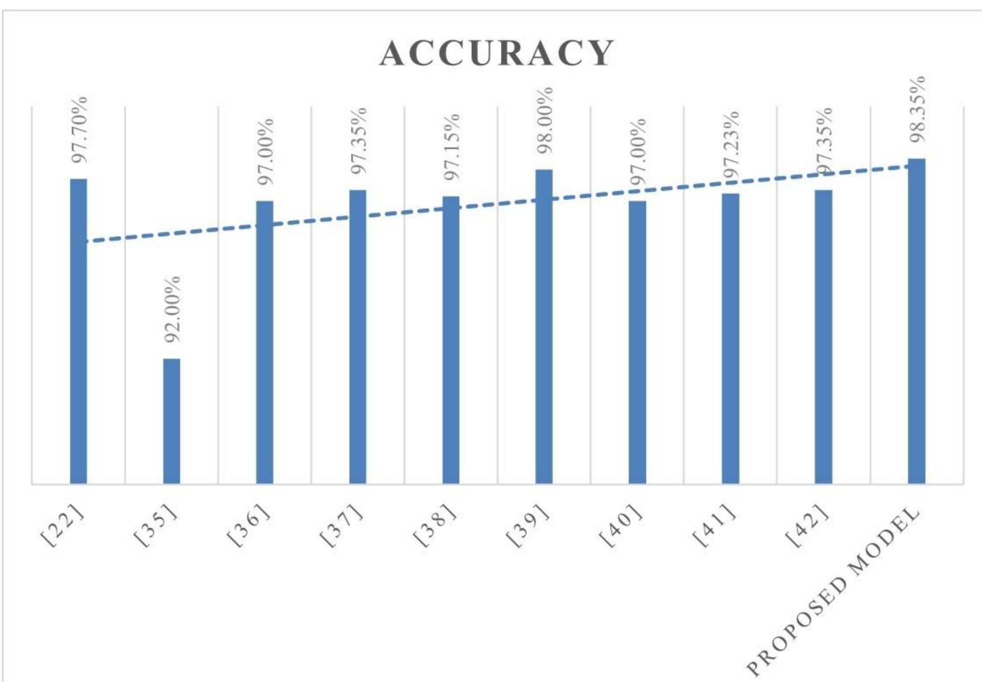


Figure 6: A comparison of accuracies of the proposed system with state-of-the-art approaches.

This proposed approach leverages real-world datasets, minimal smartphone sensor requirements, and SMOTE for data balancing to improve model performance. It evaluates multiple models across different datasets, offering a practical, efficient, and adaptable solution for HAR.
This study has several limitations one of them the dataset remains imbalanced despite SMOTE, models were trained with default parameters which may not be optimal, and they weren't tested on entirely unseen data, limiting real-world performance insights. Computational complexity wasn't analyzed, and larger models may require more resources. Additionally, feature selection or engineering wasn't applied, which could improve performance. These limitations will be addressed in future versions to guide improvements in the project.

6 Conclusion

The evaluation of different ML techniques on two distinct datasets offers valuable insights into their performance for HAR tasks. DT shows high accuracy in UCI-HAR dataset, achieving 98.35%, while KNN performed best in WILD-SHARD, with an accuracy of 96.86%. These results highlight that not all classification techniques are equally suited to the specific characteristics of the datasets, with simpler models like NB and LDA struggling to capture the complex patterns required for accurate classification. The significant variations in performance across datasets emphasize the need to choose ML methods carefully, considering the nature of the data. DT and KNN emerged as the most adaptable and effective approaches, with DT excelling in UCI-HAR dataset and KNN outperforming other techniques in Wild-Shard Dataset. This suggests that while certain methods can deliver outstanding results in specific contexts, others may require refinement or alternative strategies to improve their effectiveness. For future work, real-time implementation will focus on optimizing models for mobile devices, ensuring efficient data streaming and preprocessing, and balancing latency with accuracy. Additionally, continuous learning and adaptation to new data will be explored, alongside pilot studies and user feedback to refine and improve the real-time system.

Authors contribution

A.A.N, M.M.S., Y.T.A, M.S.I., A.A.H, A.K.K., and K.M.A. contributed to the development of ideas and the writing of this manuscript, and all authors have approved the final version.

Conflict of intrestrst

None.

References

[1] F. Ayaz *et al.*, "Radar signal processing and its impact on deep learning-driven human activity recognition," *Sensors* **25**(3), 724, doi: 10.3390/s25030724 (2025).

[2] F. Fereidoonian, F. Firouzi, and B. Farahani, "Human activity recognition: From sensors to applications," in *2020 International Conference on Omni-layer Intelligent Systems (COINS)*, 1–8. doi: 10.1109/COINS49042.2020.9191417 (2020).

[3] P. Kumar, S. Chauhan, and L. K. Awasthi, "Human activity recognition (har) using deep learning: Review, methodologies, progress and future research directions," *Arch. Comput. Methods Eng.*, **31**(1), 179–219, doi 10.1007/s11831-023-09986-x (2024).

[4] N. Manakitsa, G. S. Maraslidis, L. Moysis, and G. F. Fragulis, "A review of machine learning and deep learning for object detection, semantic segmentation, and human action recognition in machine and robotic vision," *Technologies* **12**(2), 15, doi: 10.3390/technologies12020015 (2024).

[5] N. Ahmed, J. I. Rafiq, and M. R. Islam, "Enhanced human activity recognition based on smartphone sensor data using hybrid feature selection model," *Sensors* **20**(1), 317, doi: 10.3390/s20010317 (2020).

[6] H. Dingeto and J. Kim, "Exploring Cutout and Mixup for Robust Human Activity Recognition on Sensor and Skeleton Data," *Appl. Sci.* **14**(22), 10286,. doi: 10.3390/app142210286 (2024).

[7] V. Dentamaro, V. Gattulli, D. Impedovo, and F. Manca, "Human activity recognition with smartphone-integrated sensors: A survey," *Expert Syst. Appl.* **246**, 123143,. doi: 10.1016/j.eswa.2024.123143 (2024).

[8] S. Abbaspour, F. Fotouhi, A. Sedaghatbaf, H. Fotouhi, M. Vahabi, and M. Linden, "A comparative analysis of hybrid deep learning models for human activity recognition," *Sensors* **20**(19), 5707, doi: 10.3390/s20195707 (2020).

[9] D. Khan *et al.*, "Advanced IoT-based human activity recognition and localization using Deep Polynomial neural network," *Ieee Access*, doi: 10.1109/ACCESS.2024.3420752 (2024).

[10] J. Xie *et al.*, "A survey of machine learning techniques applied to software defined networking (SDN): Research issues and challenges," *IEEE Commun. Surv. Tutorials* **21**(1), 393–430, doi:10.1109/COMST.2018.2866942 (2018).

[11] L. El Haddad, M. Hanoune, and A. Ettaoufik, "Computer Vision with Deep Learning for Human Activity Recognition: Features Representation," *Eng. Appl. Artif. Intell.*, 41–66, doi: 10.1007/978-3-031-50300-9_3 (2024).

[12] S. Zhang *et al.*, "Deep learning in human activity recognition with wearable sensors: A review on advances," *Sensors***22**(4), 1476, Doi: 10.3390/s22041476 (2022).

[13] S. K. Yadav, K. Tiwari, H. M. Pandey, and S. A. Akbar, "A review of multimodal human activity recognition with special emphasis on classification, applications, challenges and future directions," *Knowledge-Based Syst.* **223**, 106970, doi: 10.1016/j.knosys.2021.106970 (2021).

[14] F. Demrozi, G. Pravadelli, A. Bihorac, and P. Rashidi, "Human activity recognition using inertial, physiological and environmental sensors: A comprehensive survey," *IEEE access* **8**, 210816–210836, doi: 10.1109/ACCESS.2020.3037715 (2020).

[15] Z. Hussain, "Ubiquitous human activity recognition for healthcare applications: A study." Macquarie University, doi: 10.25949/24493705.v1 (2022).

[16] H. J. Mohammed, A. A. Nafea, H. K. Almulla, S. A. S. Aliesawi, and M. M. Al-Ani, "An Effective Hybrid Model for Skin Cancer Detection Using Transfer Learning," in *2023 16th International Conference on Developments in eSystems Engineering (DeSE)*, 840–845. doi: 10.1109/DeSE60595.2023.10468994 (2023).

[17] E. G. Dada, J. S. Bassi, H. Chiroma, A. O. Adetunmbi, and O. E. Ajibuwa, "Machine learning for email spam filtering: review, approaches and open research problems," *Heliyon* **5**(6)., doi: 10.1016/j.heliyon.2019.e01802 (2019).

[18] O. J. Kadhim, A. A. Nafea, S. A. S. Aliesawi, and M. M. Al-Ani, "Ensemble Model for Prostate Cancer Detection Using MRI Images," in *2023 16th International Conference on Developments in eSystems Engineering (DeSE)*, 492–497. doi: 10.1109/DeSE60595.2023.10468866 (2023).

[19] Z.-Q. Zhao, P. Zheng, S. Xu, and X. Wu, "Object detection with deep learning: A review," *IEEE Trans. neural networks Learn. Syst.* **30**(11), 3212–3232, doi: 10.1109/TNNLS.2018.2876865 (2019).

[20] J. Rzeszótko and S. H. Nguyen, "Machine learning for traffic prediction," *Fundam. Informaticae* **119**(3–4), 407–420, doi: 10.3233/FI-2012-745 (2012).

[21] A.-M. Manar, A. K. Kareem, A. A. Nafea, A. M. Shaban, S. A. S. Aliesawi, and M. M. AL-Ani, "An Effective Deep Learning Approach for the Estimation of Proton Energy by Using Artificial Neural Network," in *2024 21st International Multi-Conference on Systems, Signals & Devices (SSD)*, 257–262, doi: 10.1109/SSD61670.2024.10549706 (2024).

[22] K. Xia, J. Huang, and H. Wang, "LSTM-CNN architecture for human activity recognition," *IEEE Access* **8**, 56855–56866, doi: 10.1109/ACCESS.2020.2982225 (2020).

[23] A. Das, P. Sil, P. K. Singh, V. Bhateja, and R. Sarkar, "MMHAR-EnsemNet: A multi-modal human activity recognition model," *IEEE Sens. J.* **21**(10), 11569–11576, doi: 10.1109/JSEN.2020.3034614 (2020).

[24] W. S. Lima, H. L. S. Bragança, and E. J. P. Souto, "NOHAR-NOvelty discrete data stream for Human Activity Recognition based on smartphones with inertial sensors," *Expert Syst. Appl.* **166**, 114093, doi: 10.1016/j.eswa.2020.114093 (2021).

[25] M. Abdel-Basset, H. Hawash, R. K. Chakrabortty, M. Ryan, M. Elhoseny, and H. Song, "ST-DeepHAR: Deep learning model for human activity recognition in IoHT applications," *IEEE Internet Things J.* **8**(6), 4969–4979, doi: 10.1109/JIOT.2020.3033430 (2020).

[26] M. Abdel-Basset, H. Hawash, V. Chang, R. K. Chakrabortty, and M. Ryan, "Deep learning for heterogeneous human activity recognition in complex iot applications," *IEEE Internet Things J.* **9**(8), 5653–5665, doi: 10.1109/JIOT.2020.3038416 (2020).

[27] S. Mekruksavanich and A. Jitpattanakul, "Lstm networks using smartphone data for sensor-based human activity recognition in smart homes," *Sensors* **21**(5), 1636, doi: 10.3390/s21051636 (2021).

[28] Z. Fu, X. He, E. Wang, J. Huo, J. Huang, and D. Wu, "Personalized human activity recognition based on integrated wearable sensor and transfer learning," *Sensors* **21**(3), 885, doi: 10.3390/s21030885 (2021).

[29] S. Wan, L. Qi, X. Xu, C. Tong, and Z. Gu, "Deep learning models for real-time human activity recognition with smartphones," *Mob. networks Appl.* **25**(2), 743–755, doi: .1007/s11036-019-01445-x (2020).

[30] H. Nematallah, S. Rajan, and A.-M. Cretu, "Logistic model tree for human activity recognition using smartphone-based inertial sensors," in *2019 IEEE SENSORS*, 1–4. doi: 10.1109/SENSORS43011.2019.8956951 (2019).

[31] N. A. Choudhury and B. Soni, "An adaptive batch size-based-CNN-LSTM framework for human activity recognition in uncontrolled environment," *IEEE Trans. Ind. Informatics* **19**(10), 10379–10387, doi: 10.1109/TII.2022.3229522 (2023).

[32] N. A. Choudhury and B. Soni, "An efficient and lightweight deep learning model for human activity recognition on raw sensor data in uncontrolled environment," *IEEE Sens. J.*, doi: 10.1109/JSEN.2023.3312478 (2023).

[33] D. A. Salh and R. M. Nabi, "Kurdish fake news detection based on machine learning approaches," *Passer J. basic Appl. Sci.* **5**(2), 262–271, doi 10.24271/psr.2023.380132.1226, (2023).

[34] A. A. Nafea, M. S. Ibrahim, A. A. Mukhlif, M. M. AL-Ani, and N. Omar, "An Ensemble Model for Detection of Adverse Drug Reactions," *ARO-THE Sci. J. KOYA Univ.* **12**(1), 41–47, doi: 10.14500/aro.11403 (2024).

[35] P. Xanthopoulos, P. M. Pardalos, T. B. Trafalis, P. Xanthopoulos, P. M. Pardalos, and T. B. Trafalis, "Linear discriminant analysis," *Robust data Min.*, 27–33, doi: 10.1007/978-1-4419-9878-1_4 (2013).

[36] Z. H. Abdaljabar, O. N. Ucan, and K. M. A. Alheeti, "An intrusion detection system for IoT using KNN and decision-tree based classification," in *2021 International conference of modern trends in information and communication technology industry (MTICTI)*, 1–5, doi: 10.1109/MTICTI53925.2021.9664772 (2021).

[37] N. N. Jamil and A. K. Kareem, "Comparative Analysis on Machine Learning and One-Dimensional Convolutional Neural Network to Predict Surface Enhanced Raman Spectroscopy," in *2023 3rd International Conference on Computing and Information Technology (ICCIT)*, 216–221. doi: 10.1109/ICCIT58132.2023.10273940 (2023).

[38] S. A. Rafa, Z. M. Al-qfail, A. A. Nafea, S. F. Abd-hood, M. M. Al-Ani, and S. A. Alameri, "A Birds Species Detection Utilizing an Effective Hybrid Model," in *2024 21st International Multi-Conference on Systems, Signals & Devices (SSD)*, 705–710. doi: 10.1109/SSD61670.2024.10549480 (2024).

[39] Al-Rami, B. ., Alheeti, K. M. A. ., Aldosari, W. M., Alshahrani, S. M. ., & Al-Abrez, S. M. A New Classification Method for Drone-Based Crops in Smart Farming. International Journal of Interactive Mobile Technologies (iJIM) **16**(09), 164–174. doi: 10.3991/ijim.v16i09.30037 ((2022)).

[40] A. T. Guron, M. A. Anwer, S. K. Sulaiman, and S. J. AbdulSamad, "Classification of the cause of eye impairment using different kinds of machine learning algorithms," *Passer J. Basic Appl. Sci.* **5**(2), 410–416, doi: 10.24271/psr.2023.397078.1328 (2023).

[41] R. Mutegeki and D. S. Han, "A CNN-LSTM approach to human activity recognition," in *2020 international conference on artificial intelligence in information and communication (ICAIIC)*, 362–366. doi: 10.1109/ICAIIC48513.2020.9065078 (2020).

[42] V. Soni, H. Yadav, V. B. Semwal, B. Roy, D. K. Choubey, and D. K. Mallick, "A novel smartphone-based human activity recognition using deep learning in health care," in *Machine Learning, Image Processing, Network Security and Data Sciences: Select Proceedings of 3rd International Conference on MIND 2021*, 493–503, doi: 10.1007/978-981-19-5868-7_36 (2023).

[43] G. Türkmen and A. Sezen, "XGBoost and LightGBM A Comparative Analysis of XGBoost and LightGBM Approaches for Human Activity Recognition: Speed and Accuracy Evaluation," *Int. J. Comput. Exp. Sci. Eng.* **10**(2), doi: 10.1007/978-3-031-98284-2_6 (2024).

[44] Y. Kaya and E. K. Topuz, "Human activity recognition from multiple sensors data using deep CNNs," *Multimed. Tools Appl.* **83**(4), 10815–10838, doi: 10.1007/s11042-023-15830-y (2024).

[45] Y. H. Pang, L. Y. Ping, G. F. Ling, O. S. Yin, and K. W. How, "Stacked deep analytic model for human activity recognition on a

UCI HAR database," *F1000Research* **10**, doi: 10.12688/f1000research.73174.3 (2021).

[46] D. Bhattacharya, D. Sharma, W. Kim, M. F. Ijaz, and P. K. Singh, "Ensem-HAR: An ensemble deep learning model for smartphone sensor-based human activity recognition for measurement of elderly health monitoring," *Biosensors* **12**(6), 393, doi: 10.3390/bios12060393 (2022).

[47] M. Khaliluzzaman, M. Furquan, M. S. Zaman Khan, and M. J. Hoque, "STA-HAR: A Spatiotemporal Attention-Based Framework for Human Activity Recognition," *Appl. Comput. Intell. Soft Comput.* **2024**(1), 1832298, doi: 10.1155/2024/1832298 (2024.).

[48] U. Saha, S. Saha, M. T. Kabir, S. A. Fattah, and M. Saquib, "Decoding human activities: Analyzing wearable accelerometer and gyroscope data for activity recognition," *IEEE Sensors Lett.*, doi: 10.1109/LSENS.2024.3423340 (2024).