Research Article

On the Utilization of Sensor-Based Machine Learning Approaches for Human Activity Recognition

Saira Hamid, Tahir Javed, Aftab Khan* and Saleem Iqbal

Department of Computer Science, Allama Iqbal Open University (AIOU), Islamabad, 46000, Pakistan.

Abstract: By using smartphone sensors human activity recognition is an active field of research. A fast-growing subject of machine learning in the artificial intelligence field is Human Activity Recognition (HAR) because it makes life easier and more comfortable. Extensive research has been conducted on simple human activity recognition but very small research has been conducted on complex human activity recognition. Humans carry out a variety of activities during the day, the majority of which are complex. For the recognition of complex human activities different sensors are used to achieve high accuracy. This study assesses six classification algorithms in two scenarios: one without accelerometer data and one with it. Naïve Bayes, BayesNet, Logistic Regression, Decision Table, Random Forest, and Lazy IBK were evaluated for accuracy, precision, and recall. In the absence of accelerometer data, all algorithms showed competitive accuracy. Lazy IBK stood out with high precision and recall. With accelerometer data, all algorithms improved accuracy, with Lazy IBK reaching 100%. In the future, the study plans to develop an Android app for data collection and recognize complex human activities using accelerometer data. This research highlights the role of data sources in algorithm performance and Lazy IBK's suitability for activity classification.

Received: November 05, 2023; **Accepted:** January 07, 2025; **Published:** August 22, 2025
***Correspondence:** Aftab Khan, Department of Computer Science, Allama Iqbal Open University (AIOU), Islamabad, 46000, Pakistan; **Email:** dr.aftabkhan87@gmail.com
Citation: Hamid, S., T. Javed, A. Khan and S. Iqbal. 2025. On the utilization of sensor-based machine learning approaches for human activity recognition. *Journal of Engineering and Applied Sciences,* 44: 13-26.
DOI: https://dx.doi.org/10.17582/journal.jeas/44.13.26
Keywords: Human Activity Recognition (HAR), Machine Learning (ML), Classification, Activity, Accelerometer

Introduction

Human Activity Recognition (HAR) strives to discern the actions and motions of either an individual or multiple individuals by analyzing a sequence of observations involving people. HAR represents a highly impactful research domain and finds extensive application in surveillance (Avci *et al*., 2010; Osmani *et al*., 2008; Taha *et al*., 2015). Human activities can be classified in two categories and detail is given in next section. Human activities are categorized into two groups: Simple Human Activities (SHAs) and Complex Human Activities (CHAs) as shown in Figure 1. This classification is contingent upon the body's motion and function.

Simple Human Activities (SHAs) are characterized by repetitive, singular movements. They are also referred to as low-level or elementary activities. Examples of these include sitting, standing, walking, running, and sleeping (Vishakha *et al*., 2017) shown in Figure 2.

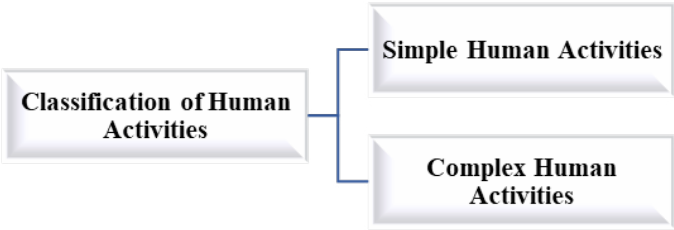


Figure 1: *Classification of human activities.*

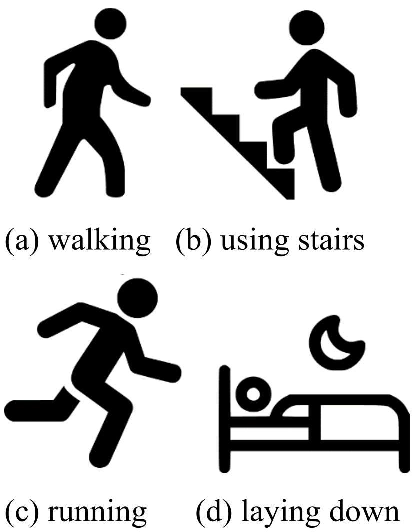


Figure 2: *Examples of simple human activities (Online Web Fonts).*

Complex human activities shown in Figure 3 encompass a series of diverse actions, representing a blend of various simple human activities. They are also referred to as high-level or compound activities. Examples of complex human activities include ironing, dishwashing, baking, watering plants, engaging in physical exercise, and reading the newspaper (Ferrari *et al*., 2023; Vishakha *et al*., 2017). The study focuses on recognizing complex activities, which include cycling, computer work, vacuum cleaning, ironing, folding laundry, and house cleaning.

Following the classification of human activities, the next section provides descriptions of various approaches for human activity recognition.

Limitations of the study

It is proposed that this study has the following limitations. We used the already collected dataset, so this data collection process, while being quite exhaustive, may suffer from set innate biases of the study because of the demographic characteristics of the participants. Moreover, there is the limitation of wearability that may cause noise and, therefore, a poor quality of data collected. As much as the identified algorithms were appropriately chosen, they might not capture all the valuable methods for the recognition of human activity. Future work will extend this work to gather a larger data sample, and consider the use of other algorithms.

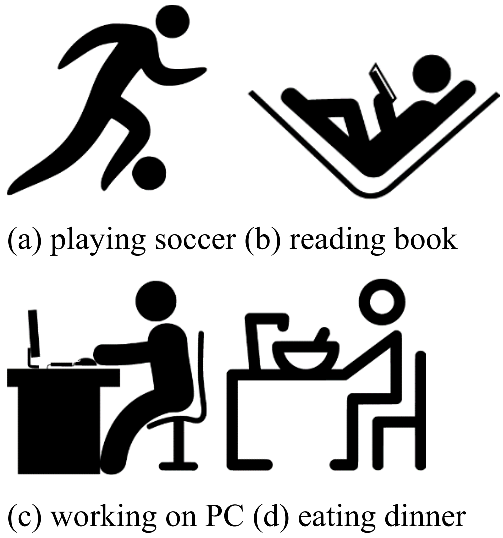


Figure 3: *Examples of complex human activities (online web fonts).*

Literature review

While extensive research has focused on simple human activity recognition, limited investigation has been conducted for complex human activity recognition (Garcia-Gonzalez *et al*., 2023; Geravesh and Rupapara, 2023). This chapter serves as a valuable resource, providing background and literature on techniques for recognizing complex human activities, offering fundamental insights into diverse recognition methods. Motion sensors, typically located in a trouser pocket, play a crucial role in detecting human activities. However, this positioning has limitations for identifying hand gesture activities like smoking, drinking coffee, and talking (Chen *et al*., 2020). To address this, wrist-worn motion sensors are used to detect these activities, and both sensor locations are tested. Three types of motion sensors (accelerometer, gyroscope, and linear acceleration) are employed in this testing (Kuan, 2020).

For classification, three different classifiers are employed: Lazy IBK (KNN), decision tree, and Naive Bayes. Using both wrist and pocket sensor positions together yields superior results with narrower segmentation windows compared to using only the wrist position. Detecting less frequent activities like smoking, dining, conversing, and drinking coffee in shorter segmentation periods poses a challenge. To address this, the study explores the impact of seven window sizes (2–30 seconds) on thirteen different tasks, illustrating how window size affects these activities in various scenarios. The study also suggests optimizations to improve activity detection accuracy. Importantly, the dataset used in this research is publicly available for reproducibility (Shoaib *et al*., 2016).

Zainudin and colleagues introduce a complex model that requires substantial time and memory resources to analyze numerous features. As the volume of less informative elements rises, accuracy decreases. To tackle this, they propose a hybrid feature selection method, combining Relief-f and differential advancement. This technique prioritizes the removal of less significant features before entering the classifier model, enhancing both efficiency and accuracy (Zainudin *et al*., 2017).

The paper in (Dernbach *et al*., 2012) addresses the growing adoption of assistive healthcare technologies. In the realm of assistive healthcare technology, the focus has mainly been on identifying movements, particularly basic ones like walking. However, this paper explores recognizing complex activities such as cooking or washing using smartphone data derived from inertial sensors. It employs these functions as input for supervised machine learning. While basic activities are easier to detect, recognizing complex activities alongside them remains challenging. Another study emphasizes the difficulty of recognizing human activities when they involve multitasking, especially complex activities. Most research centers on basic activity identification, leaving complex activity recognition based on sensors underexplored. Addressing the challenges of recognizing Complex Human Activities (CHA), the research introduces the “DEBONAIR” deep learning model, designed for multimodal CHA recognition. DEBONAIR utilizes specialized sub-networks for various sensor inputs and an LSTM network to learn sequential patterns in CHAs. Experimental results indicate DEBONAIR outperforms other CHA recognition methods, with future plans to enhance its performance further (Kuan, 2020).

This study presents a smartphone-based method for human activity recognition, using data from built-in sensors. It eliminates the need for extra sensors by utilizing smartphone sensors. The method achieved high accuracy in identifying tasks and is implemented as an Android application, ensuring practicality and accessibility (Alruban *et al*., 2019). One study focuses on recognizing human activity networks using sensor data and introduces the “InnoHAR” deep learning model, combining elements from the “Inception Neural Network (INN)” and recurrent neural networks. This innovative approach applies deep neural network expertise to human activity recognition, achieving an impressive 93.5% accuracy rate on the pamap2 dataset. It addresses the limitations of traditional feature engineering methods (Xu *et al*., 2019).

The 2010 survey explores the application of inertial sensors in activity recognition, emphasizing health, home monitoring, and sports. It outlines a structured five-step process for activity recognition: Pre-processing, segmentation, feature extraction, dimensionality reduction, and classification. This approach enhances the analysis of human activities in diverse contexts (Avci *et al*., 2010).

This study emphasizes achieving a profound understanding of human activities via sensor-based recognition. While traditional pattern recognition methods have improved, they often rely on manual feature extraction, limiting performance. Deep learning has transformed this field, enabling automatic high-level feature extraction, making it vital for sensor-based activity recognition. The paper categorizes research into three areas: deep models, sensor types, and applications. It provides thorough analyses, guiding future research. This holistic approach aims to advance sensor-based activity recognition and enhance our comprehension of human activities with cutting-edge technology (Wang *et al*., 2019). This study proposes a sensor-based system to monitor the health of socially isolated elderly individuals. It tracks daily activities like urination, kitchen work, and cleanliness using water flow sensors and microphones on faucets. Regularity and patterns of these activities assess well-being. The system is cost-effective, privacy-conscious, and easily implementable at home. Experiments show its potential for health monitoring and support for the elderly (Tsukiyama, 2015).

In this research, Hassan and colleagues introduce a smartphone sensor-based approach for human activity recognition. They emphasize the role of smartphones in this context and employ a Deep Belief Network (DBN) for training and recognition. A comparison with Support Vector Machine (SVM) reveals that DBN achieves a remarkable 95.89% accuracy in recognizing various human activities, indicating its potential for accurate recognition of complex activities in future applications (Hassan *et al.*, 2018).

Human activity recognition process

Following steps are used for human activity recognition (Vishakha *et al.*, 2017). Firstly, the data is collected then pre-processing will be done on data. After that feature should be extracted that are useful. Then the next step is dimensionality reduction. At the end classification is done in which human activity is recognized. Human Activity Recognition process is shown in Figure 4 and details of these is given in sub-sections.

Data collection: The first step involves collecting data using sensors. These sensors can be placed on the human body or in the environment to capture information related to the activities being monitored. Common sensors include accelerometers, gyroscopes, magnetometers, and more (Reiss and Stricker, 2012).

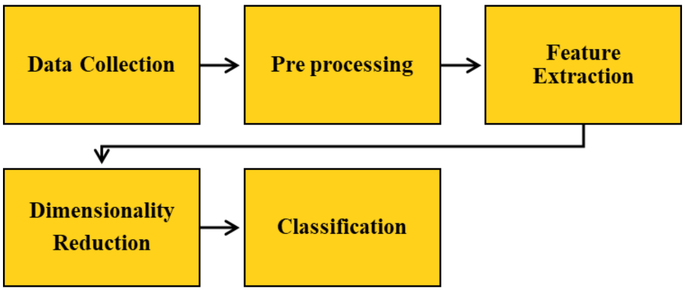


Figure 4: *Human activity recognition process.*

Pre-processing: Once the data is collected, it often needs to be pre-processed. This step includes cleaning the data by removing noise, outliers, or irrelevant information. Additionally, missing data may be handled using imputation techniques (Maharana *et al.*, 2022).

Feature extraction: In this step, relevant features are extracted from the pre-processed sensor data. These features represent characteristics of the activities and are used as input for the classification model. Feature extraction methods depend on the type of data and the specific activities being recognized (Zhang *et al.*, 2022).

Dimensionality reduction: Sensor data can be high-dimensional, which can lead to computational challenges and overfitting. Dimensionality reduction techniques, such as Principal Component Analysis (PCA) or feature selection, may be applied to reduce the number of features while preserving important information (Vishakha *et al.*, 2017).

Classification: The core of HAR is classification, where machine learning or pattern recognition algorithms are used to classify the activities based on the extracted features. Common classifiers include k-Nearest Neighbours (k-NN), Decision Trees, Support Vector Machines (SVM), Neural Networks, and Naive Bayes.

The proposed methodology utilized the model shown in Figure 4 expect it omits the dimensionality reduction stage.

Experimental setup

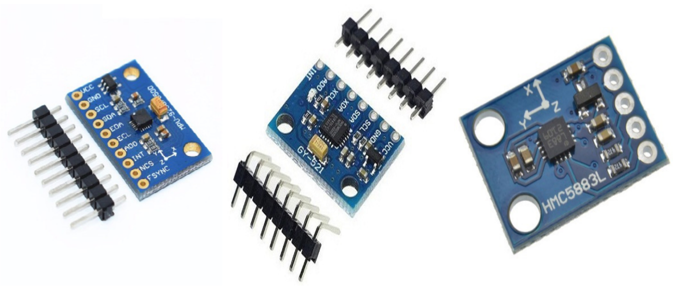
This chapter provides a comprehensive overview of the experimental setup, the experiments conducted, and the corresponding results, along with a detailed discussion of these findings. According to Statista data from 2021, an estimated 6.378 billion people use smartphones worldwide. These devices have become essential tools for various purposes, including making calls, sending emails, engaging with social media, storing and managing data, and performing numerous other tasks. Additionally, smartphones are instrumental in creating smart environments in homes, schools, and various other settings. This information highlights the ubiquity and importance of smartphones in today's society and underscores their role in facilitating not only personal communication but also the development of smart and connected ecosystems in various contexts.

Sensors

Smartphone sensors indeed play a crucial role in recognizing complex human activities. These sensors, which often include gyroscope, accelerometer, and magnetometer sensors, provide valuable data for activity recognition. In the research mentioned, smartphones are utilized due to the accessibility of these sensors, which enable the collection of data necessary for recognizing complex human activities.

The PAMAP2 dataset is specifically employed for training different machine learning algorithms.

This dataset serves as a valuable resource, providing real-world data that researchers can use to develop and test their recognition models. By leveraging smartphone sensors and datasets like PAMAP2, researchers can advance the field of human activity recognition and develop systems capable of identifying complex activities performed by individuals (Zainudin *et al*., 2017). The three sensors' data available in the dataset include the accelerometer, the gyroscope and the magnetometer. These sensor details and functioning is briefly explained as follows with the sensors shown in Figure 5.



(a) (b) (c)

Figure 5: *(a) Accelerometer (WatElectronics, 2021) (b) Gyroscope (WatElectronics, 2019) and (c) Magnetometer (Electronicwings).*

Accelerometers measure acceleration along X, Y, and Z axes, crucial for human activity analysis. Gyroscopes detect orientation changes but require accelerometer data for absolute orientation. Combining both sensors improves precision, ideal for applications like AR/VR. Magnetometers measure Earth's magnetic field along three axes, useful for compass and AR apps, providing comprehensive orientation data.

Datasets

Data collection, often referred to as a dataset, is a fundamental component in research and experimentation. Standard datasets play a crucial role in enabling researchers to conduct experiments, develop models, and compare their results with those of other researchers. This process of using standard datasets fosters collaboration and facilitates the improvement of results and accuracy in various fields. In the ongoing study, the PAMAP2 dataset was utilized. The PAMAP2 dataset, made available for academic research by Reiss and Stricker (Reiss and Stricker, 2012), is a valuable resource. It is provided to the research community without any restrictions, allowing researchers to use it for the purpose of recognizing human activities. Such publicly available datasets are instrumental in advancing the state of the art in human activity recognition, as they provide a common benchmark for testing and comparing different recognition algorithms and approaches. The experiments are done on PAMAP2 dataset by applying different machine learning algorithms to get the result. There are nine participants labelled as subiect 1 to subject 9 in this dataset performing different complex human activities. Eight participants are male and one participant is female.

Classification techniques

Machine learning is considered as a component of artificial intelligence. Machine learning is used as its training is CPU based and it needs less time for training. Some of the machine learning algorithms which are used in research are discussed below.

- Naïve bayes: The Naive Bayesian model is straightforward to construct, making it a preferred choice, especially for handling extensive datasets. Its reputation for surpassing even the most sophisticated classification algorithms is attributed to its simplicity (Zhang and Jiang, 2022).
- Bayes Net: A "BayesNet classifier" is essentially a "Bayesian network" used for predicting the probability P(c|x) of a discrete class variable C, relying on specific attributes X (Dhobale *et al*., 2022).
- Logistic regression: Logistic Regression establishes the connection between independent and dependent variables within the dataset. Its efficiency is notable due to its minimal computational resource requirements. This classification algorithm finds widespread use and is highly accessible for implementation. Additionally, data training in this method is efficient and straightforward (Srimaneekarn *et al*., 2022).
- Lazy IBK (KNN): Lazy IBK essentially implements the K-Nearest Neighbors (KNN) algorithm, where "K" represents the number of neighbors taken into account. One characteristic of this approach is its computational intensity since it necessitates calculating the distances to all neighbors. Consequently, it is computationally expensive. Moreover, it scans the entire dataset for each prediction, leading to high accuracy but

slower processing speeds (Slavova *et al*., 2022).

- Decision table (DT): Decision tables provide a straightforward visual representation of actions to be taken in response to specific situations. These algorithms generate a list of actions as their output based on the input data and conditions (Ariza-Colpas *et al*., 2022).
- Random forest (RF): In the training phase, multiple decision trees are merged to create a collective learning approach for categorization. This collective learning approach either delivers a classification result or predicts the mean of each tree individually (Thakur and Biswass, 2022).

Software tools

In this research, classification is carried out using Weka, a widely used software tool for data analysis known as "Waikato Environment for Knowledge Analysis (WEKA). The experimental model involves a classification process, and WEKA offers a variety of classification algorithms for this purpose (Samawi *et al*., 2022). Some examples of these algorithms include Logistics, Lazy IBK, Decision Table, Random Forest, and Naïve Bayes. To evaluate the performance of the classification model, a "10-fold cross-validation" approach is employed. In this method, the dataset is divided into 10 equal-sized folds. During each iteration, one of the folds serves as the "test set data," while the remaining nine folds act as the "training set." This process is repeated for all 10 folds, with each fold taking a turn as the test set. As a result, the learning algorithm is invoked 10 times in total within the 10-fold cross-validation model. The use of the 10-fold cross-validation model helps assess the validity and robustness of the proposed system by testing it against different subsets of the data, ultimately providing a comprehensive evaluation of the classification performance.

Results and Discussion

The results of the experimentation are divided into two sets. (a) Classification Results without using accelerometer data and (b) Classification Results using accelerometer data. The activities are labelled as follows.

a = cycling, b = vacuum cleaning, c= Ironing, d= folding laundry, e = House cleaning, f = computer work.

Overall accuracy, precision and recall have been analyzed in the two scenarios for the classification results and the respective methods used.

Classification results without using accelerometer data

The paper presents classification outcomes without the utilization of accelerometer data, and various algorithm performance is evaluated. Figures 6 to 11 shows the confusion charts for the following algorithms (a) Naïve Bayes (b) BayesNet (c) Logistic Regression (d) Decision Table (DT) (e) Random Forest (RF) and (f) Lazy IBK.

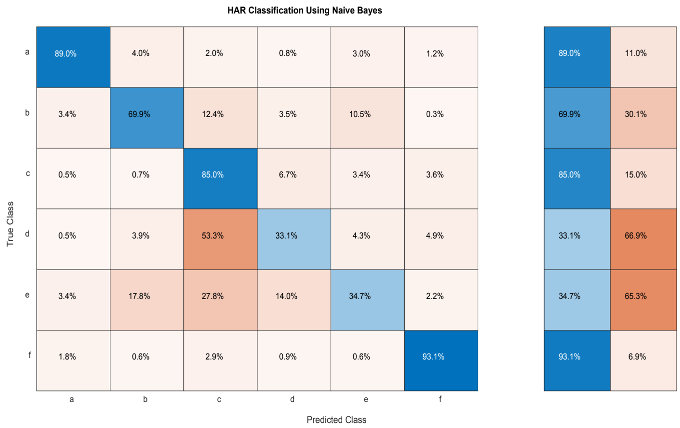


Figure 6: *Confusion matrix using Naïve bayes.*

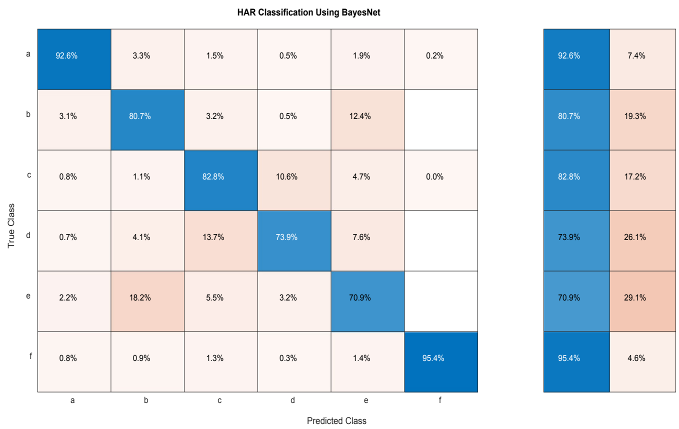


Figure 7: *Confusion matrix using bayesnet.*

Figure 8: *Confusion matrix using logistic regression.*

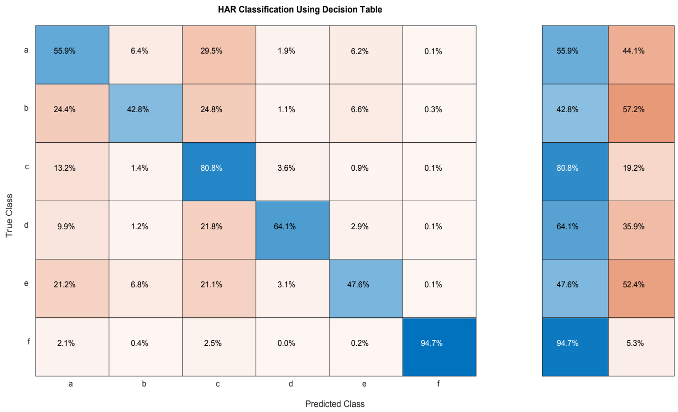


Figure 9: *Confusion matrix using decision table (DT).*

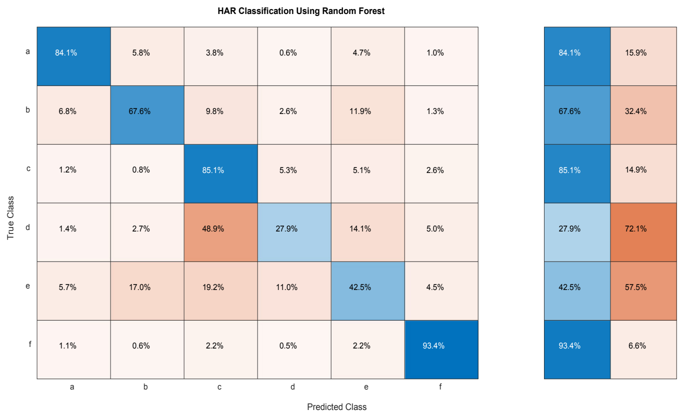


Figure 10: *Confusion matrix using random forest (RF).*

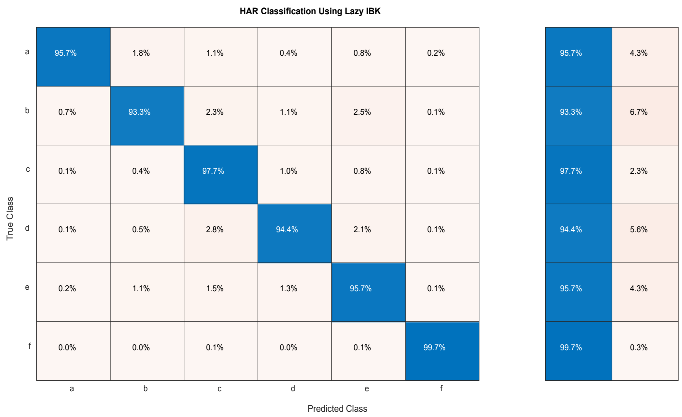


Figure 11: *Confusion matrix using lazy IBK.*

In assessing the performance of six different algorithms on the same classification task involving six distinct classes (Cycling, Vacuum Cleaning, Ironing, Folding Laundry, House Cleaning, and Computer Work), it's crucial to begin with an examination of overall accuracy. While overall accuracy serves as a valuable starting point, it must be complemented by a deeper analysis of other performance metrics:
C1 (Naïve Bayes): The confusion matrix for Naïve Bayes (C1) exhibits high overall accuracy, indicating that this algorithm performs well in correctly classifying instances across all classes.
C2 (BayesNet): Similar to C1, BayesNet (C2) demonstrates a high level of overall accuracy, suggesting strong performance in this classification task.
C3 (Logistic Regression): Logistic Regression (C3) maintains consistently high overall accuracy, signifying effective classification across the given classes.
C4 (Decision Table): The confusion matrix for Decision Table (C4) also indicates high overall accuracy, affirming the model's reliability in classification.
C5 (Random Forest): C5, representing the Random Forest algorithm, shows high overall accuracy, underscoring the algorithm's predictive power.
C6 (Lazy IBK - KNN): C6 demonstrates high overall accuracy as well, suggesting strong performance by the Lazy IBK (KNN) algorithm.

Moving beyond overall accuracy, it is essential to delve into precision and recall metrics, which provide valuable insights into the trade-offs between false positives and false negatives for each class.

Tables 1 and 2 shows the precision and recall values for the classification schemes. In Table 1, precision and recall values for six distinct classification algorithms (Naïve Bayes, BayesNet, Logistic Regression, Decision Table, Random Forest, and Lazy IBK - KNN) are presented across six diverse classes.

Naïve Bayes emerges as a strong contender, showcasing high precision and recall across most classes, notably excelling in Vacuum Cleaning, Ironing, and Computer Work. While it demonstrates overall robust performance, it exhibits comparatively lower recall in Folding Laundry. BayesNet, much like Naïve Bayes, performs admirably in Vacuum Cleaning, Ironing, and Computer Work, although its performance in Folding Laundry lags slightly. Logistic Regression demonstrates consistent and well-balanced precision and recall, signifying dependable and uniform performance across all classes. Decision Table delivers noteworthy performance in Cycling and Ironing but grapples with Folding Laundry, where both precision and recall dip. Random Forest shines with remarkable precision and recall in Cycling and House Cleaning, sustaining its competitive edge in other classes. Lazy IBK (KNN) stands out as a consistent performer, achieving high precision and recall values across all classes, with exceptional proficiency in house cleaning and computer work.

Table 1: *Precision value for the classification schemes.*

Technique/ Activity	Naïve Bayes	BayesNet	Logistic regression	Decision table (DT)	Random forest (RF)	Lazy IBK (KNN)
Cycling	83.31	89.54	69.97	57.40	76.64	95.73
Vacuum cleaning	54.58	66.96	52.13	37.88	53.36	93.32
Ironing	83.53	71.82	73.23	70.77	78.38	97.66
Folding laundry	17.86	61.40	5.05	29.59	11.46	94.39
House cleaning	18.16	40.44	35.91	29.78	27.03	95.70
Computer work	86.47	93.68	81.91	88.76	84.19	99.73
Average	57.32	70.64	53.03	52.36	55.18	96.09

Table 2: Recall value for the classification schemes.

Technique/ Activity	Naïve Bayes	BayesNet	Logistic regression	Decision table (DT)	Random forest (RF)	Lazy IBK (KNN)
Cycling	90.16	92.22	77.07	42.89	83.48	98.77
Vacuum cleaning	70.00	72.45	68.36	70.05	69.20	95.62
Ironing	60.55	85.58	70.09	55.94	64.97	95.40
Folding laundry	39.54	68.82	45.21	76.58	41.67	92.92
House cleaning	65.40	75.38	54.92	77.55	58.77	95.15
Computer work	87.40	99.74	84.24	99.09	85.78	99.43
Average	68.84	82.37	66.65	70.35	67.31	96.22

Table 2 offers further insights into algorithmic performance. Naïve Bayes, while maintaining its strength in Vacuum Cleaning, Ironing, and Computer Work, experiences a slight decline in precision and recall for Cycling compared to Table 1. BayesNet maintains relatively stable precision and recall values across classes, retaining its effectiveness in Vacuum Cleaning, Ironing, and Computer Work. Logistic Regression, akin to Table 1, demonstrates steady and harmonized precision and recall metrics for all classes. Decision Table remains consistent in its performance, with Folding Laundry still posing a challenge. Random Forest continues to impress with its robust precision and recall, particularly excelling in Cycling and House Cleaning. Lazy IBK (KNN) maintains its high precision and recall values across all classes, consistently excelling in House Cleaning and Computer Work.

Across both tables, Naïve Bayes, BayesNet, Logistic Regression, and Lazy IBK (KNN) consistently exhibit competitive precision and recall values. However, variations in performance emerge for specific classes, suggesting that algorithm selection should align with the specific objectives and priorities of the classification task. Notably, Decision Table encounters difficulties with Folding Laundry in both tables, indicating a potential area for improvement. Random Forest consistently performs well, particularly excelling in Cycling and House Cleaning. Fine-tuning and optimization options exist to further enhance algorithmic performance, tailoring them to specific classes or overall accuracy requirements. Ultimately, the choice of the most suitable algorithm hinges on the precise demands of the classification task, balancing precision and recall considerations across diverse classes.

Results

The research presents classification results utilizing accelerometer data to assess the effectiveness of various algorithms in human activity recognition. Figures 12 to 17 shows the confusion charts for the classification algorithms when accelerometer data is also utilized.

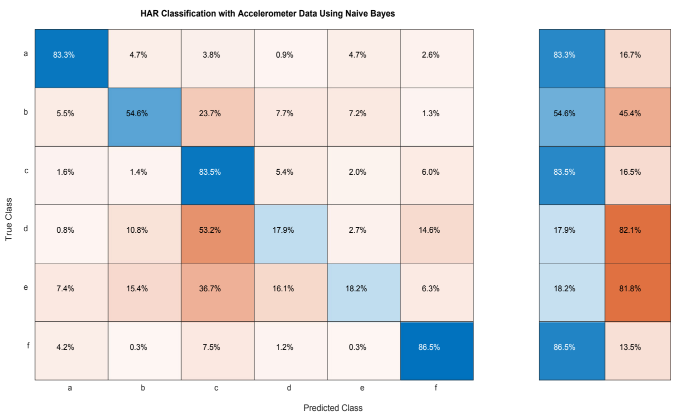


Figure 12: *Confusion matrix using Naïve Bayes*

In this comprehensive analysis of six confusion matrices (C1 to C6), generated for various classification algorithms (Naïve Bayes, BayesNet, Logistic Regression, Decision Table, Random Forest, and Lazy IBK - KNN), we delve into their performance across six distinct classes (Cycling, Vacuum Cleaning, Ironing, Folding Laundry, House Cleaning, and Computer Work). In the initial confusion matrix, Naïve Bayes exhibits strong overall classification performance. It excels in correctly classifying instances related to Vacuum Cleaning, Ironing, and Computer Work, boasting high true positive counts in these categories. However, it encounters challenges in accurately classifying instances of Folding Laundry, with a comparatively lower count of true positives. BayesNet demonstrates competitive classification performance across most classes. It particularly shines in accurately classifying Cycling and House Cleaning instances, achieving substantial true positive counts in these categories. Similar to Naïve Bayes, it faces difficulties when classifying Folding Laundry accurately.

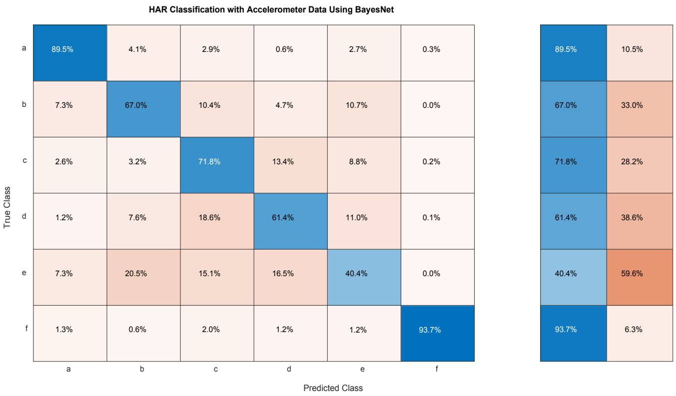


Figure 13: *Confusion matrix using BayesNet.*

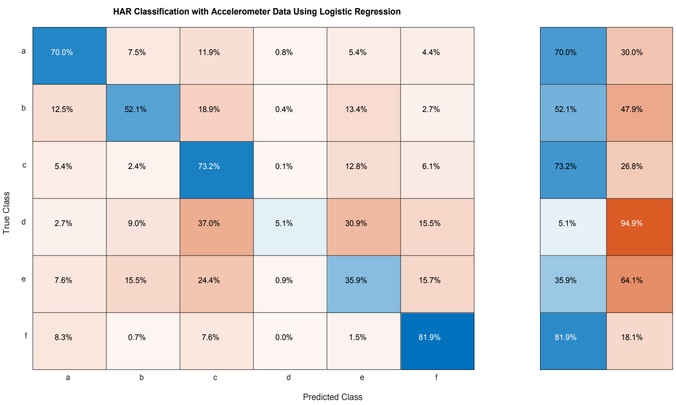


Figure 14: *Confusion matrix logistic regression.*

Logistic regression maintains consistent and balanced classification performance across all classes. It sustains relatively high true positive counts in each category, reflecting its stable and dependable performance. The strength of logistic regression lies in its uniformity, as it doesn't exhibit significant variations in performance among different classes. Decision Table presents mixed performance, displaying strengths in some categories and weaknesses in others. It excels in classifying cycling and ironing instances but struggles in accurately classifying folding laundry, where the true positive count is lower. This variation indicates that decision table might not be the optimal choice for all categories. Random forest consistently demonstrates robust classification performance across all classes. It particularly excels in accurately classifying cycling and house cleaning instances, attaining high true positive counts in these categories. Overall, random forest showcases impressive and stable performance.

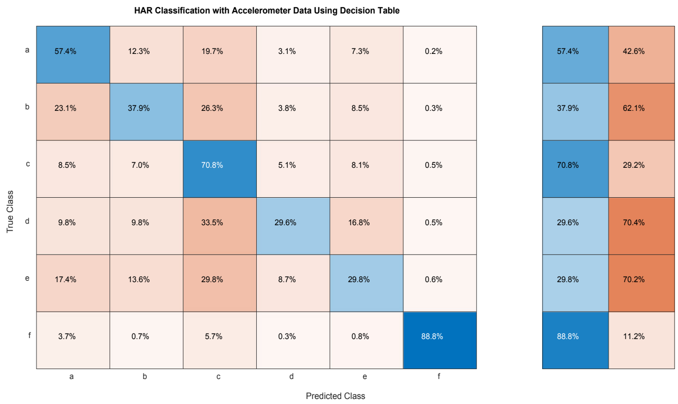


Figure 15: *Confusion matrix using decision table (DT).*

Figure 16: *Confusion matrix using random forest (RF).*

Lazy IBK, represented as KNN, delivers strong classification performance across all categories. It performs exceptionally well in House Cleaning and Computer Work, achieving high true positive counts in these categories. Similar to other algorithms, it faces challenges when classifying Folding Laundry instances accurately but maintains strong overall performance.

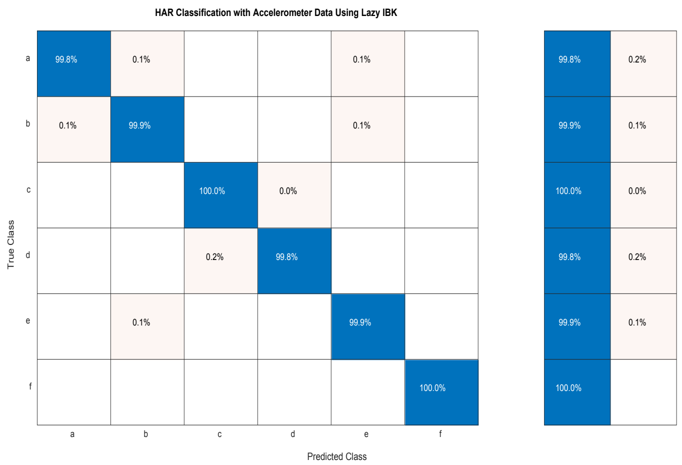


Figure 17: *Confusion matrix lazy IBK (KNN).*

In comparing these categories against each other: Lazy IBK (KNN) emerges as a consistent performer, achieving high true positive counts across multiple categories, particularly excelling in house cleaning and computer work. Naïve Bayes and BayesNet perform well in categories like vacuum cleaning, ironing, and computer work but may face challenges with Folding Laundry. Logistic regression demonstrates balanced and uniform performance across all categories, making it a reliable choice. Decision table performs well in categories like cycling and ironing but encounters difficulties with folding laundry. Random forest consistently excels, particularly in categories like cycling and house cleaning, showcasing its versatility and reliability.

Tables 3 and 4 shows the precision and recall values for the classification. In Table 3, Naïve Bayes shows high precision, excelling in classes a (cycling) and f (computer work). However, it has lower recall in classes d (folding laundry) and e (house cleaning). BayesNet maintains competitive precision and uniform recall across classes. Logistic Regression maintains balanced precision and recall, reflecting its reliability. Decision Table 3 exhibits varying precision with challenges in class d. Random Forest consistently excels in both precision and recall. Lazy IBK (KNN) achieves high precision and recall across all classes, with exceptional performance in classes a (cycling) and f (computer work).

Table 4 continues to show Naïve Bayes excelling in precision, especially in classes b (vacuum cleaning), c (ironing), and f (computer work). BayesNet maintains stable precision and recall. Logistic regression's precision remains consistent, with uniform recall. Decision table's precision varies across classes, excelling in a and c. Random forest consistently performs well. Lazy IBK (KNN) maintains high precision and recall across all classes, excelling in e and f.

Table 3: *Precision value for the classification schemes using accelerometer data.*

Technique/ Activity	Naïve Bayes	BayesNet	Logistic regression	Decision table (DT)	Random forest (RF)	Lazy IBK (KNN)
Cycling	88.98	92.58	79.23	55.89	84.10	100.00
Vacuum cleaning	69.92	80.71	65.25	42.83	67.58	99.99
Ironing	85.04	82.79	85.07	80.77	85.05	100.00
Folding laundry	33.10	73.86	22.71	64.14	27.90	100.00
House cleaning	34.75	70.85	50.29	47.56	42.52	99.99
Computer work	93.10	95.37	93.66	94.72	93.38	100.00
Average	67.48	82.69	66.04	64.32	66.76	100.00

Table 4: *Precision value for the classification schemes using accelerometer data.*

Technique/ Activity	Naïve Bayes	BayesNet	Logistic regression	Decision Table (DT)	Random Forest (RF)	Lazy IBK (KNN)
Cycling	80.31	80.99	65.00	47.45	72.51	100.00
Vacuum cleaning	62.42	62.88	58.68	44.55	60.53	99.99
Ironing	53.09	70.79	55.02	49.69	53.97	100.00
Folding laundry	23.33	45.61	57.87	42.38	26.87	100.00
House cleaning	56.04	58.07	41.89	47.12	45.77	99.99
Computer work	73.54	99.27	63.56	97.37	68.32	100.00
Average	58.12	69.60	57.00	54.76	54.66	100.00

Lazy IBK (KNN) consistently demonstrates strong precision and recall across all classes. Naïve Bayes, BayesNet, and Logistic Regression offer competitive performance with slight variations. Decision Table exhibits mixed performance, and Random Forest consistently excels, particularly in classes a and e. The choice of the most suitable algorithm depends on specific objectives and trade-offs between precision and recall for different classes, with potential for further fine-tuning and optimization.

Accuracy comparison of the classification models utilised

Figure 18 displays accuracy scores for various classification algorithms under two different conditions: without accelerometer data and with accelerometer data. Without accelerometer data, the Naïve Bayes algorithm achieved an accuracy of 57.32%. BayesNet outperformed Naïve Bayes with an accuracy of 70.64%, demonstrating its effectiveness in this context. Logistic Regression achieved an accuracy of 53.03%, while Decision Table (DT) and Random Forest (RF) yielded accuracies of 52.36% and 55.18%, respectively. Lazy IBK (KNN) exhibited the highest accuracy without accelerometer data, with an impressive score of 96.09%.

When accelerometer data was incorporated into the classification, there was a notable improvement in accuracy across all algorithms. Naïve Bayes saw a significant boost, reaching an accuracy of 67.48%. BayesN*et al*so benefited from accelerometer data, achieving an accuracy of 82.69%. Logistic Regression improved to 66.04%, DT to 64.32%, and RF to 66.76%. Remarkably, Lazy IBK (KNN) reached a perfect accuracy of 100%. In summary, the inclusion of accelerometer data substantially enhanced the accuracy of all classification algorithms. Lazy IBK (KNN) achieved the highest accuracy in both scenarios, while BayesNet showcased its effectiveness in the absence of accelerometer data. The choice of algorithm should consider the specific context and the availability of sensor data to maximize classification accuracy.

Comparison with benchmark schemes

Table 5 summarizes benchmark schemes and Table 6 shows the classification results of the methods used in research studies, focusing on the PAMAP2 dataset and complex activities. Notable achievements include Convolutional Neural Networks (CNN) reaching 93.7% accuracy and Lazy IBK (KNN) achieving a perfect 100% accuracy when accelerometer data is incorporated. These results underscore the effectiveness of machine learning techniques in activity classification, with Lazy IBK (KNN) emerging as a standout performer. Additionally, the Relief-f method demonstrated impressive accuracy of 98.8%, highlighting its utility in the context of the PAMAP2 dataset.

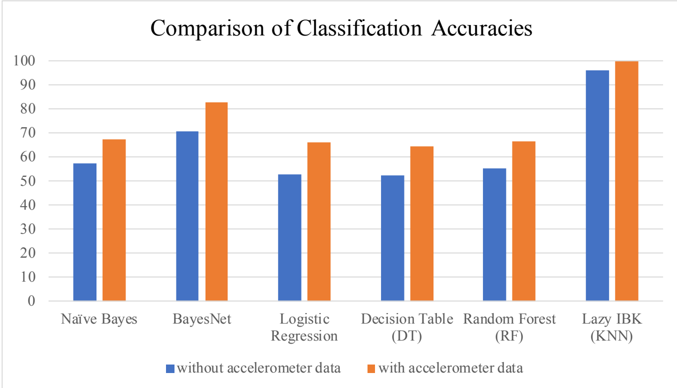


Figure 18: *Accuracy comparison of different classification algorithms.*

Table 5: *Comparison of classification accuracies with benchmark schemes.*

S. No	Benchmark Scheme	Classification method	Dataset/Activity	Accuracy
1	Shoaib *et al.* (2016)	Naïve Bayes	Complex human activities	85.00%
2	Hammerla *et al.* (2016)	CNN	PAMAP2	93.70%
3	Guan *et al.* (2017)	Ensembles LSTM	PAMAP2	85.40%
4	Murad and Pyun (2017)	Deep RNN	Opportunity/ complex	92.00%
5	Zainudin *et al.* (2017)	Relief-f	PAMAP2	98.80%

Table 6: *Classification Results using PAMAP2 dataset.*

S. No	Classification method	Accuracy without accelerometer	Accuracy with accelerometer
6	Naïve Bayes	57.32	67.48
7	BayesNet	70.64	82.69
8	Logistic Regression	53.03	66.04
9	Decision Table (DT)	52.36	64.32
10	Random Forest (RF)	55.18	66.76
11	Lazy IBK (KNN)	96.09	100

Conclusions and Recommendations

The study assessed six classification algorithms in two scenarios: One without accelerometer data and one with it. Naïve bayes, bayesnet, logistic regression, decision table, random forest, and Lazy IBK were analyzed for overall accuracy, precision, and recall. Without accelerometer data, all algorithms demonstrated competitive accuracy. Naïve Bayes and bayesnet excelled in vacuum cleaning, ironing, and computer work but faced challenges in folding laundry. Logistic regression showed consistent performance, and decision table performed well in cycling and ironing but struggled with folding laundry. Random Forest consistently performed well, especially in cycling and house cleaning. Lazy IBK stood out with high precision and recall across all classes.

With accelerometer data, all algorithms improved accuracy, with Lazy IBK achieving a perfect 100%. It consistently displayed strong precision and recall. Naïve Bayes, BayesNet, and logistic regression offered competitive performance. Decision table had mixed results, and random forest excelled in cycling and house cleaning. In an accuracy comparison, accelerometer data significantly enhanced results, with Lazy IBK (KNN) leading in overall accuracy. This highlights the importance of data sources in algorithm performance and underscores Lazy IBK's suitability for activity classification tasks. In the future, it is planned to develop an android application for data collection and to recognize more complex human activities using accelerometer information.

Acknowledgements

The authors would like to thank Allama Iqbal Open University for providing the necessary re-sources and facilities. Special thanks to Prof. Dr. Abdul Majid from the Pakistan Institute of Ap-plied Sciences for his valuable feedback. We also sincerely appreciate the constructive comments from the reviewers, which significantly improved the quality of this manuscript.

Novelty Statement

This study reveals that Lazy IBK excels in complex human activity recogni-tion, achieving perfect accuracy with accelerometer data, highlighting the critical role of data sources in machine learning performance.

Author's Contribution

Saira Hamid: Conceptualization, design, manuscript draft-ing, reviewing, analysis and figure design.
Tahir Javed: Experimentation, data processing and analysis.
Aftab Khan: Conceptualization, design, manuscript draft-ing, reviewing, analysis and figure design.
Saleem Iqbal: Data processing, analysis, manuscript draft-ing, reviewing and figure design.

Conflict of interest
The authors have declared no conflict of interest.

References

Alruban, A., Alobaidi, H., Clarke, N. and Li, F., 2019. Physical activity recognition by utilising smartphone sensor signals proceedings of the 8th international conference on pattern recognition applications and methods. https://doi.org/10.5220/0007271903420351

Ariza-Colpas, P.P., Vicario, E., Oviedo-Carrascal, A.I., Butt, A.S., Pineres-Melo, M.A., Quintero-Linero, A. and Patara, F., 2022. Human activity recognition data analysis: History, evolutions, and new trends. Sensors, 22(9): 3401. https://doi.org/10.3390/s22093401

Avci, A., Bosch, S., Marin-Perianu, M., Marin-Perianu, R. and Havinga, P., 2010. Activity recognition using inertial sensing for healthcare, wellbeing and sports applications: A survey. 23th International conference on architecture of computing systems 2010.

Chen, L., Liu, X., Peng, L. and Wu, M., 2020. Deep learning based multimodal complex human activity recognition using wearable devices. Appl. Intell., 51(6): 4029-4042. https://doi.org/10.1007/s10489-020-02005-7

Dernbach, S., Das, B., Krishnan, N.C., Thomas, B.L. and Cook, D.J., 2012. Simple and complex activity recognition through smart phones 2012 Eighth international conference on intelligent environments. https://doi.org/10.1109/IE.2012.39

Dhobale, N., Mulik, S.S. and Deshmukh, S.P., 2022. Naïve bayes and bayes net classifier for fault diagnosis of end mill tool using wavelet analysis: A comparative study. J. Vibration Eng. Technol., 10(5): 1721-1735. https://doi.org/10.1007/s42417-022-00478-z

Electronicwings, Magnetometer. https://www.electronicwings.com/sensors-modules/hmc5883l-magnetometer-module

Ferrari, A., Micucci, D., Mobilio, M. and Napoletano, P., 2023. Deep learning and model personalization in sensor-based human activity recognition. J. Reliable Intell. Environ., 9(1): 27-39. https://doi.org/10.1007/s40860-021-00167-w

Garcia-Gonzalez, D., Rivero, D., Fernandez-Blanco, E. and Luaces, M.R., 2023. New machine learning approaches for real-life human activity recognition using smartphone sensor-based data. Knowl. Based Syst., 262: 110260. https://doi.org/10.1016/j.knosys.2023.110260

Geravesh, S. and Rupapara, V., 2023. Artificial neural networks for human activity recognition using sensor based dataset. Multimed. Tools Appl., 82(10): 14815-14835. https://doi.org/10.1007/s11042-022-13716-z

Guan, Y. and Plötz, T., 2017. Ensembles of deep LSTM learners for activity recognition using wearables. Proc. ACM Interact. Mobile, Wearable Ubiquitous Technol., 1(2): 1-28. https://doi.org/10.1145/3090076

Hassan, M.M., Uddin, M.Z., Mohamed, A. and Almogren, A., 2018. A robust human activity recognition system using smartphone sensors and deep learning. Future Generat. Comp. Syst., 81: 307-313. https://doi.org/10.1016/j.future.2017.11.029

Khan, A., Hammerla, N., Mellor, S. and Plötz, T., 2016. Optimising sampling rates for accelerometer-based human activity recognition. Pattern Recogn. Lett., 73: 33-40. https://doi.org/10.1016/j.patrec.2016.01.001

Kuan, C.B.W.H.C.C.W., 2020. Self-attention networks for human activity recognition using wearable devices.

Maharana, K., Mondal, S. and Nemade, B., 2022. A review: Data pre-processing and data augmentation techniques. Glob. Transit. Proc., 3(1): 91-99. https://doi.org/10.1016/j.gltp.2022.04.020

Murad, A. and Pyun, J.Y., 2017. Deep recurrent neural networks for human activity recognition. Sensors (Basel, Switzerland), 17(11): 2556. https://doi.org/10.3390/s17112556

Online Web Fonts. https://www.onlinewebfonts.com/icon

Osmani, V., Balasubramaniam, S. and Botvich, D., 2008. Human activity recognition in pervasive health-care: Supporting efficient remote collaboration. J. Netw. Comp. Appl., 31(4): 628-655. https://doi.org/10.1016/j.jnca.2007.11.002

Reiss, A. and Stricker, D., 2012. Introducing a new benchmarked dataset for activity monitoring 2012 16th international symposium on wearable computers. https://doi.org/10.1109/ISWC.2012.13

Samawi, V.W., Yousif, S.A. and Al-Saidi, N.M., 2022. Intrusion detection system: An automatic machine learning algorithms using auto-WEKA. 2022 IEEE 13th Control and System Graduate Research Colloquium (ICSGRC). https://doi.org/10.1109/ICSGRC55096.2022.9845166

Shoaib, M., Bosch, S., Incel, O.D., Scholten, H. and Havinga, P.J.M., 2016. Complex human activity recognition using smartphone and wrist-worn motion sensors. Sensors (Basel, Switzerland), 16(4): 426-426. https://doi.org/10.3390/s16040426

Slavova, V., Ropelewska, E., Sabanci, K., Aslan, M.F., Nacheva, E.J.E.F.R. and Technology, 2022. A comparative evaluation of Bayes, functions, trees, meta, rules and lazy machine learning algorithms for the discrimination of different breeding lines and varieties of potato based on spectroscopic data. Eur. Fd. Res. Technol., 248(7): 1765-1775. https://doi.org/10.1007/s00217-022-04003-0

Srimaneekarn, N., Hayter, A., Liu, W. and Tantipoj, C., 2022. Binary response analysis using logistic regression in dentistry. Int. J. Dentist., 2022(1): 1-7. https://doi.org/10.1155/2022/5358602

Taha, A., Zayed, H.H., Khalifa, M.E. and El-Horbaty, E.S.M., 2015. Human activity recognition for surveillance applications the 7th International Conference on Information Technology. https://doi.org/10.15849/icit.2015.0103

Thakur, D. and Biswas, S., 2022. An integration of feature extraction and guided regularized random forest feature selection for smartphone based human activity recognition. J. Netw. Comp. Appl., 204: 103417. https://doi.org/10.1016/j.jnca.2022.103417

The revised version of the above-mentioned paper, 'Manuscript ID MH20231105175404-R3,' is attached herewith for your reference.

Tsukiyama, T., 2015. In-home health monitoring system for solitary elderly. Proc. Comp. Sci., 63: 229-235. https://doi.org/10.1016/j.procs.2015.08.338

Vishakha V.V., Walse, K.H. and Dharaskar, R.V., 2017. A survey on human activity recognition using smartphone. Int. J. Adv. Res. Comp. Sci. Manage. Stud., 5(3): 118-124.

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

WatElectronics, 2019. Gyroscope (Electronics, 2021). Retrieved May 16, 2020 from https://www.watelectronics.com/what-is-a-gyroscope-sensor-working-its-applications/

WatElectronics, 2021. Accelerometer. Retrieved October 29, 2021 from https://www.watelectronics.com/accelerometer/

Xu, C., Chai, D., He, J., Zhang, X. and Duan, S., 2019. InnoHAR: A deep neural network for complex human activity recognition. IEEE Access, 7: 9893-9902. https://doi.org/10.1109/ACCESS.2018.2890675

Zainudin, M.S., Sulaiman, M.N., Mustapha, N., Perumal, T. and Mohamed, R., 2017. Recognizing complex human activities using hybrid feature selections based on an accelerometer sensor. Int. J. Technol., 8(5): 968-978. https://doi.org/10.14716/ijtech.v8i5.879

Zhang, C., Mousavi, A.A., Masri, S.F., Gholipour, G., Yan, K. and Li, X., 2022. Vibration feature extraction using signal processing techniques for structural health monitoring: A review. Mech. Syst. Signal Proc., 177: 109175. https://doi.org/10.1016/j.ymssp.2022.109175

Zhang, H. and Jiang, L., 2022. Fine tuning attribute weighted naive Bayes. Neurocomputing, 488: 402-411 https://doi.org/10.1016/j.neucom.2022.03.020.