

Assessment of ML Classifiers in Complex Human Activity Recognition Using Wearable Sensors Data

Saira Hamid*, M. Arshad Awan, Mubashar Saeed

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

*Correspondence: sairahamid84@yahoo.com

Citation | Hamid. S, Awan. M. A, Saeed. M, “Assessment of ML Classifiers in Complex Human Activity Recognition Using Wearable Sensors Data”, IJIST, Special Issue. pp 104-117, Oct 2024

DOI | <https://doi.org/10.33411/IJIST/1088>

Received | Oct 06, 2024 **Revised |** Oct 13, 2024 **Accepted |** Oct 18, 2024 **Published |** Oct 21, 2024.

Human Activity Recognition (HAR) is essential for understanding daily behavior patterns, and wearable sensor data serves as a reliable source for monitoring complex activities. This study uniquely evaluates the performance of nine machine learning classifiers in the context of complex human activity recognition, relying solely on wearable sensors. It offers valuable insights into classifier effectiveness for real-world applications. Data from the PAMAP2 dataset, which was collected using three wearable IMUs placed on the hand, chest, and ankle, along with a heart rate sensor, was utilized to identify six daily complex activities. A 70/30 train-test split methodology was implemented to assess classifier performance. The Random Forest (RF) classifier achieved the highest performance, boasting 93% accuracy, precision, recall, and F1-score, followed closely by the K-Nearest Neighbors (KNN) classifier, which recorded 91% across all metrics. In contrast, the Logistic Regression (LR) classifier underperformed, achieving only 55% accuracy, likely due to its limitations in handling non-linear data. These results demonstrate that RF and KNN classifiers are effective for complex human activity recognition, while linear classifiers like LR are less suitable for such tasks. Overall, the Random Forest and KNN classifiers provide reliable performance for complex human activity recognition using wearable sensors, making them excellent choices for practical applications.

Keywords: Human Activity Recognition, Machine Learning, Classification, Wearable Sensor, Complex Activity.



Introduction:

Wearable sensor-based Human Activity Recognition (HAR) has gained significant attention recently due to its wide-ranging applications in healthcare, sports, and human-computer interaction [1]. HAR systems utilize sensor data collected from wearable devices to recognize human activities, offering the potential to transform various sectors through real-time monitoring and analysis of human behavior [2]. Promising results have been achieved in accurately recognizing and classifying human actions using machine learning (ML) classifiers [3]. These classifiers excel at analyzing complex patterns in sensor data to predict actions being performed. However, the effectiveness of ML classifiers can be influenced by factors such as the types of activities being recognized, the sensor modalities employed, and the classifier algorithms utilized [4].

This study evaluates the performance of different ML classifiers in identifying common daily complex human activities using wearable sensor data (WSD). The specific activities selected for analysis include using a computer, cycling, folding laundry, cleaning the house, ironing, and vacuuming. Data for these activities were collected using inertial measurement units (IMUs) placed on various parts of the body, including the hand, chest, and ankle, to capture a range of postures and movements.

The wearable sensor configuration consists of three wireless Colibri IMUs, each equipped with two accelerometers, one gyroscope, one magnetometer, and a sampling frequency of 100 Hz. Data collection involved nine participants, comprising eight males and one female, with an average age of 27 to 31 years [5]. The study assesses the effectiveness of nine commonly used ML classifiers in HAR applications: Decision Tree (DT), Gaussian Naive Bayes (GNB), Random Forest (RF), AdaBoost (AB), Gradient Boosting (GB), Logistic Regression (LR), Support Vector Machine (SVM), K-Nearest Neighbors (KNN), and Multi-Layer Perceptron (MLP). These classifiers were trained and evaluated using WSD to recognize the activities performed by the participants.

The findings provide valuable insights into the effectiveness of various ML classifiers for identifying complex human activities, indicating that RF and KNN achieved strong performance, while LR performed the weakest. These observations could inform the development of more autonomous and accurate HAR systems for a variety of real-world applications.

Objective of the Study:

The objective of this study is to evaluate and compare the performance of nine distinct machine learning classifiers in recognizing complex human activities solely through wearable sensor data. This research seeks to determine which classifiers are most effective at accurately classifying daily complex activities based on data collected from wearable devices. By doing so, it aims to provide insights into the suitability of various classifiers for real-world human activity recognition tasks.

Novelty Statement:

This study presents a novel evaluation of various machine learning classifiers for complex human activity recognition, utilizing only wearable sensors. It specifically addresses the gap in understanding classifier performance for recognizing intricate activities. By focusing on real-world sensor data and comparing nine classifiers, the research provides unique insights into the most effective algorithms for achieving accurate and resource-efficient human activity recognition—an area that has not been thoroughly explored previously.

Material:

The study utilized data from the publicly available PAMAP2 dataset, which comprises sensor data collected from wearable Colibri wireless IMUs positioned on three body locations: the hand, chest, and ankle, along with a heart rate sensor. This dataset records six common daily complex activities performed by the participants.

Method:

A train-test split methodology was employed, utilizing 70% of the data for training and 30% for testing. Nine different machine learning classifiers were evaluated, including Random Forest (RF), K-Nearest Neighbors (KNN), and Logistic Regression (LR), among others. The performance of each classifier was assessed using metrics such as accuracy, precision, recall, and F1-score to determine their effectiveness in recognizing complex human activities from the sensor data.

Literature Review:

Recent years have witnessed significant research focused on Human Activity Recognition (HAR) using wearable sensors, driven by its potential applications in human-computer interaction, sports, and healthcare. Various strategies have been explored to enhance the precision and effectiveness of HAR systems, with a major emphasis on the selection and fusion of sensor modalities. Research indicates that accelerometer data are effective for identifying activities such as jogging, walking, and navigating stairs [6]. However, the integration of gyroscopes and magnetometers has been shown to improve the accuracy of activity recognition, particularly for complex movements involving rotations and changes in orientation [7].

The application of machine learning (ML) techniques has greatly advanced the development of HAR systems. Classifiers like Support Vector Machines (SVM) are popular due to their ability to handle high-dimensional data and nonlinear interactions [8]. Studies have demonstrated that Random Forest (RF) and Gradient Boosting (GB) algorithms also achieve high accuracy in HAR tasks [9]. These research efforts aim to enhance the generalizability and robustness of HAR models. Transfer learning has emerged as a viable strategy to leverage knowledge from related tasks or domains to improve HAR performance [10]. For instance, one study outperformed traditional ML models by employing transfer learning in HAR using wearable sensors [11].

Moreover, deep learning methods have gained traction in HAR research. Convolutional Neural Networks (CNNs) are utilized to automatically extract features from sensor data, eliminating the need for manual feature engineering [12]. Recurrent Neural Networks (RNNs) have shown promise in capturing temporal dependencies within sequential sensor data, further enhancing activity recognition accuracy [13]. Despite these advancements, challenges remain in HAR research, particularly regarding sensor orientation and location variability, which can affect data quality and the effectiveness of HAR models [14]. Ethical considerations concerning the security and privacy of personal data collected by wearable sensors also need to be addressed in the design of HAR systems [15].

This research aims to enhance classifier performance by comparing supervised and ensemble learning classifiers in the context of HAR with mobile devices. It evaluates the effectiveness of walking and sitting movements on two UCI datasets while performing feature selection to reduce dimensionality, thus transforming high-dimensional data into lower-dimensional forms. The findings underscore the significance of HAR in security, fitness, and healthcare [31].

Human Activity Recognition (HAR) is a critical area of research within body area networks and computing, focused on categorizing input data into distinct classes. This study explores both basic and deep learning methods, employing dimensionality reduction and feature extraction via Topological Data Analysis (TDA) to tackle HAR challenges. The WISDM and UCI-HAR datasets serve as public data sources for experimentation. Techniques for data balancing and sampling mechanisms ensure the acquisition of balanced datasets. Consequently, seven machine learning methods are utilized as ensemble classifiers, including 1D-CNN, BiLSTM, and GRU, alongside proposed deep learning approaches. The study presents the best reported results for the proposed methods on the WISDM and UCI-HAR datasets [32].

This paper reviews the application of Inertial Measurement Unit (IMU) sensors for HAR, aiming to enhance the functionality of wearable devices during daily activities. The study focuses on measuring acceleration and angular velocity to personalize gestural control. Four machine learning models are utilized: Artificial Neural Networks, Decision Tree Classifier, Random Forest Classifier, and K-Nearest Neighbors. Notably, the Random Forest Classifier achieves a high accuracy of 97.67%, thereby enabling greater functionality and customization of wearable robotics, ultimately improving the quality of life for individuals with disabilities [33].

In summary, HAR based on wearable sensors is an evolving field with substantial practical applicability. Current research primarily concentrates on sensor modality integration, machine learning and deep learning approaches, as well as challenges like sensor variability and data privacy. Progress in these areas is expected to lead to more precise and reliable HAR systems, with applications across various real-world scenarios.

Material and Methods:

Figure 1 shows the flow of this study.

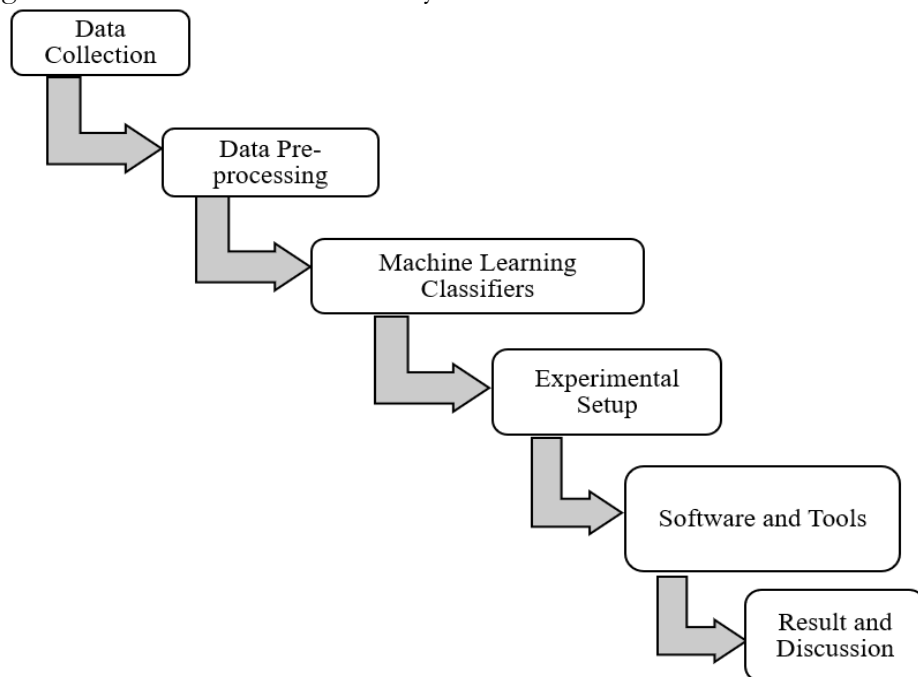


Figure 1. Flow of Study

Data Collection:

The PAMAP2 (Physical Activity Monitoring) [5] dataset is employed in this study. This dataset was gathered using three Inertial Measurement Unit (IMU) sensors and one heart rate sensor, as previously described. Each IMU consists of two accelerometers, one gyroscope, and one magnetometer, all operating at a sampling frequency of 100 Hz. The data collection involved nine volunteers, comprising one female and eight males, with an average age ranging from 27 to 31 years.

Data Preprocessing:

To generate orientation-independent values, data from each IMU's accelerometers, gyroscope, and magnetometer were fused. Time-domain features were collected, including rotation on the x, y, and z axes from the gyroscope; acceleration on the x, y, and z axes from two accelerometers; and magnetic field data on the x, y, and z axes from the magnetometer for each IMU.

For this study, 36 out of 53 potential features were extracted. Table 1 summarizes the sensor data, noting that features associated with simple activities were excluded, as the focus is solely on complex activities. The total number of samples collected was 992,341 (approximately

1 million), with each participant contributing around 20 minutes of data, which includes approximately 3 to 4 minutes of each activity. To evaluate the classifiers, a train-test split methodology was employed, dividing the dataset into 70% (694,638 samples) for training and 30% (297,703 samples) for testing.

Table 1. Extracted features from 3 IMUs

Sensor	Value
Accelerometer (2 at each IMU)	Acceleration at x-, y-, and z-axis
Gyroscope (1 at each IMU)	Rotation at x-, y-, and z-axis
Magnetometer (1 at each IMU)	Magnetic field at x-, y-, and z-axis

Each feature (column) in the CSV file was labeled according to the corresponding sensor data. The six activities (classes) were also labeled appropriately to facilitate the supervised learning approach for the machine learning models. Due to the presence of missing values in the data, the corresponding rows (samples) were removed. Ultimately, 992,341 samples remained for analysis.

Machine Learning Classifiers:

A total of nine machine learning classifiers were selected for evaluation: Gaussian Naive Bayes (GNB), K-Nearest Neighbors (KNN), Decision Tree (DT), Random Forest (RF), AdaBoost (AB), Gradient Boosting (GB), Logistic Regression (LR), Support Vector Machine (SVM), and Multi-Layer Perceptron (MLP). The classifiers were trained using 70% of the dataset, while the remaining 30% was reserved for testing. Each classifier's performance was assessed using key metrics, including F1-score, accuracy, precision, and recall.

Experimental Setup:

The classifiers were evaluated based on their accuracy, precision, recall, and F1-score in classifying the six activities. These metrics provided a comprehensive assessment of each model's performance in recognizing and distinguishing between the various human activities.

Software and Tools:

Preprocessing, feature extraction, and classification were conducted using Microsoft Excel and the Python programming language. The machine learning classifiers were implemented with the Pandas, NumPy, and Scikit-learn libraries. This methodology aimed to evaluate how effectively the classifiers recognized complex human activities based on data collected from wearable sensors. The findings of the study will provide valuable insights into the relative strengths of various classifiers for human activity recognition (HAR) applications, aiding researchers in selecting appropriate sensors and machine learning models. While previous studies primarily utilized inertial sensors, especially accelerometers, to identify simple activities, the integration of additional inertial sensors, such as gyroscopes and magnetometers, enables the recognition of complex human activities with enhanced accuracy.

Result and Discussion:

Experimental Results:

The tests aimed to evaluate the effectiveness of nine machine learning classifiers in identifying complex human activities based on wearable sensor data (WSD). The activities under consideration included computer work, cycling, folding laundry, vacuuming, ironing, and housecleaning. The classifiers were assessed using key performance metrics, including recall, accuracy, precision, and F1-score.

Classifier Performance:

The Random Forest classifier achieved an overall accuracy of 93%, with corresponding precision, recall, and F1-score also at 93%, indicating the best performance among the classifiers tested. Table 2 provides a breakdown of precision, recall, and F1-score for each activity. The overall precision, recall, and F1-score are calculated using a weighted average to account for the varying number of samples across activities.

Table 2. Precision, recall & f1-score (Random Forest)

Activity	Precision	Recall	F1-Score
Computer-Work	0.9985	0.9898	0.9941
Cycling	0.9571	0.9681	0.9626
Folding-Laundry	0.9233	0.8432	0.8814
house-cleaning	0.8998	0.8901	0.8949
ironing	0.9052	0.9003	0.8792
vacuum-cleaning	0.9573	0.9305	0.8896

The confusion matrix for Random Forest is given below in the Figure 2.

computer-work	44405	79	3	116	174	87
cycling	3	47439	27	359	339	835
folding-laundry	9	142	25295	1525	2276	753
house-cleaning	16	525	984	49809	2346	2276
ironing	29	179	780	1501	68452	563
vacuum-cleaning	10	1200	307	2048	2038	40774
	computer-work	cycling	folding-laundry	house-cleaning	ironing	vacuum-cleaning

Figure 2. Confusion matrix for Random Forest

The K-Nearest Neighbors (KNN) classifier performed competitively, achieving an overall accuracy of 91%. Table 3 presents the precision, recall, and F1-score for each activity, while Figure 3 illustrates the confusion matrix for a comprehensive view of the classifier's performance.

Table 3. Precision, recall & f1-score (KNN)

Activity	Precision	Recall	F1-Score
Computer-Work	0.9841	0.9935	0.9888
Cycling	0.9672	0.9427	0.9548
Folding-Laundry	0.8033	0.8579	0.8297
House-Cleaning	0.8753	0.8775	0.8764
Ironing	0.8901	0.9307	0.9100
Vacuum-Cleaning	0.9216	0.8296	0.8732

Table 4. Precision, recall & f1-score (Decision Tree)

Activity	Precision	Recall	F1-Score
Computer-Work	0.9816	0.9836	0.9826
Cycling	0.9191	0.9058	0.9124
Folding-Laundry	0.7366	0.7402	0.7384
House-Cleaning	0.7692	0.7757	0.7724
Ironing	0.8544	0.8615	0.8579
Vacuum-Cleaning	0.7770	0.7669	0.7719

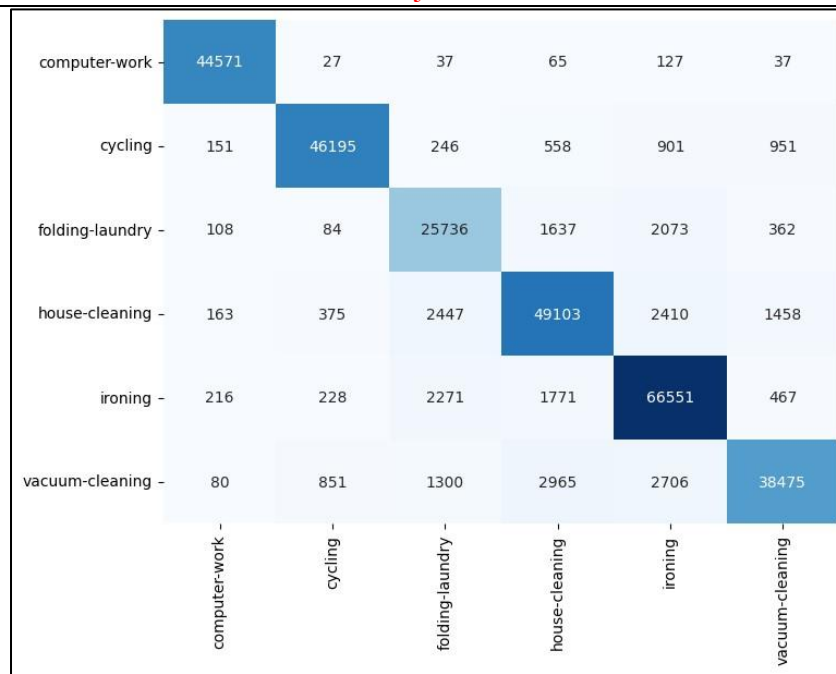


Figure 3. Confusion matrix for KNN

The Decision Tree classifier also demonstrated strong performance, achieving an overall accuracy of 84%. Table 4 provides the precision, recall, and F1-score for each activity, while Figure 4 presents the confusion matrix, offering further insight into the classifier's effectiveness.

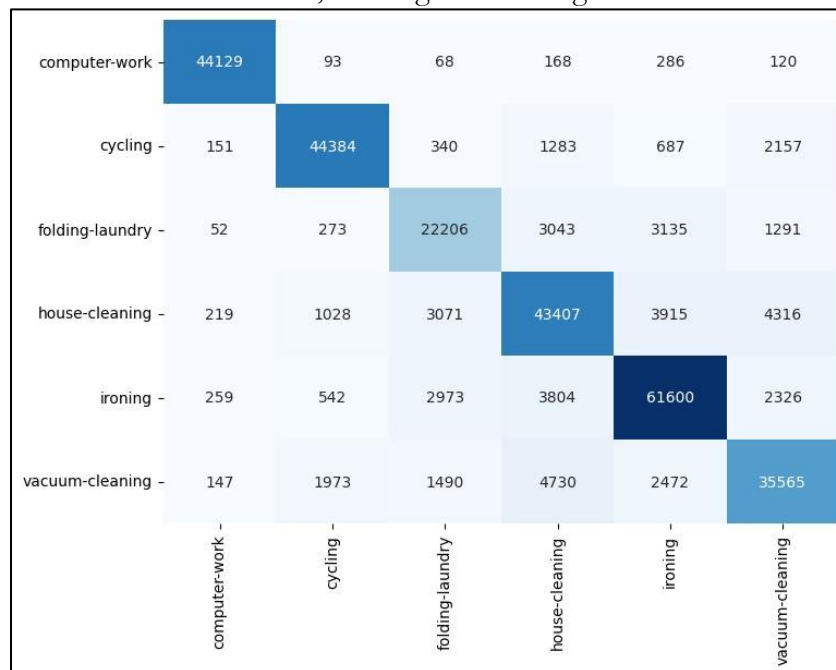


Figure 4. Confusion matrix for Decision Tree

The MLP classifier achieved an overall accuracy of 80%. Table 5 details the accuracy, precision, recall, and F1-score for the MLP classifier, while Figure 5 displays the corresponding confusion matrix. Although the Support Vector Machine (SVM) classifier produced reasonable results, it required an excessive amount of time—over 10 hours—to build the model on a Google Colab setup. This long training time makes the use of SVM a critical consideration, particularly for real-time applications, where speed is essential. The SVM classifier achieved an overall accuracy of 77%. Table 6 presents the precision, recall, and F1-score for each activity,

and Figure 6 illustrates the confusion matrix, highlighting the correctly classified instances for each class.

Table 5. Precision, recall & f1-score (MLP)

Activity	Precision	Recall	F1-Score
Computer-Work	0.9823	0.9747	0.9784
Cycling	0.9400	0.9033	0.9213
Folding-Laundry	0.6404	0.6160	0.6280
House-Cleaning	0.7274	0.6450	0.6838
Ironing	0.7470	0.8738	0.8054
Vacuum-Cleaning	0.7615	0.7220	0.7412

computer-work	43727	87	79	276	509	186
cycling	67	44264	234	891	1686	1860
folding-laundry	104	151	18479	3654	6389	1223
house-cleaning	245	819	5306	36094	7455	6037
ironing	315	364	3518	3642	62483	1182
vacuum-cleaning	59	1406	1239	5063	5128	33482
	computer-work	cycling	folding-laundry	house-cleaning	ironing	vacuum-cleaning

Figure 5. Confusion matrix for MLP

computer-work	43138	80	34	559	737	316
cycling	209	44449	216	974	1903	1251
folding-laundry	413	163	13638	5324	9221	1241
house-cleaning	677	1102	3758	35665	8865	5889
ironing	1548	912	2508	5069	60134	1333
vacuum-cleaning	104	2030	881	5375	6027	31960
	computer-work	cycling	folding-laundry	house-cleaning	ironing	vacuum-cleaning

Figure 6. Confusion matrix for SVM

Another ensemble classifier, Gradient Boosting (GB), was evaluated and achieved an overall accuracy of 76%. Details regarding precision, recall, and F1-score for the GB classifier can be found in Table 7, while Figure 7 presents the corresponding confusion matrix.

Table 6. Precision, recall & f1-score (SVM)

Activity	Precision	Recall	F1-Score
Computer-Work	0.9360	0.9615	0.9486
Cycling	0.9120	0.9071	0.9096
Folding-Laundry	0.6483	0.4546	0.5345
House-Cleaning	0.6734	0.6374	0.6549
Ironing	0.6921	0.8410	0.7593
Vacuum-Cleaning	0.7611	0.6891	0.7233



Figure 7. Confusion matrix for Gradient Boosting

Table 7. Precision, recall & f1-score (Gradient Boosting)

Activity	Precision	Recall	F1-Score
Computer-Work	0.9756	0.9614	0.9684
Cycling	0.8971	0.9028	0.9000
Folding-Laundry	0.6558	0.4680	0.5462
House-Cleaning	0.6492	0.6148	0.6315
Ironing	0.6978	0.8322	0.7591
Vacuum-Cleaning	0.7155	0.6865	0.7007

Table 8. Precision, recall & f1-score (AdaBoost)

Activity	Precision	Recall	F1-Score
Computer-Work	0.8877	0.9018	0.8947
Cycling	0.7310	0.8605	0.7905
Folding-Laundry	0.5465	0.1714	0.2610
House-Cleaning	0.4831	0.3945	0.4343
Ironing	0.5989	0.7789	0.6771
Vacuum-Cleaning	0.5781	0.5779	0.5780

The AdaBoost (AB) classifier achieved an overall accuracy of 65%. Additional details regarding its performance are provided in Table 8 and illustrated in Figure 8.

computer-work	40460	938	48	1179	2062	177
cycling	625	42164	21	1788	2278	2126
folding-laundry	702	487	5143	7166	13450	3052
house-cleaning	1514	5404	2322	22073	12507	12136
ironing	1739	2707	1081	8214	55691	2072
vacuum-cleaning	539	5979	795	5266	6995	26803
	computer-work	cycling	folding-laundry	house-cleaning	ironing	vacuum-cleaning

Figure 8. Confusion matrix for Ada Boost

Table 9. Precision, recall & f1-score (GNB)

Activity	Precision	Recall	F1-Score
Computer-Work	0.7340	0.8617	0.7927
Cycling	0.7989	0.8322	0.8152
Folding-Laundry	0.2371	0.1799	0.2046
House-Cleaning	0.5595	0.1828	0.2756
Ironing	0.5319	0.8357	0.6501
Vacuum-Cleaning	0.6249	0.5470	0.5833

computer-work	38660	2101	488	153	3350	112
cycling	1270	40779	437	2322	1915	2279
folding-laundry	4430	244	5398	834	15844	3250
house-cleaning	3449	4217	9001	10229	20493	8567
ironing	4328	1091	3924	1388	59756	1017
vacuum-cleaning	534	2614	3522	3357	10984	25366
	computer-work	cycling	folding-laundry	house-cleaning	ironing	vacuum-cleaning

Figure 9. Confusion matrix for GNB

The Gaussian Naive Bayes (GNB) classifier took the least amount of time to build the model, completing the process in just a few minutes. GNB achieved an overall accuracy of 61%. The precision, recall, and F1-score for each activity are detailed in Table 9, while Figure 9 presents the corresponding confusion matrix.

Table 10. Precision, recall & f1-score (Logistic Regression)

Activity	Precision	Recall	F1-Score
Computer-Work	0.6222	0.8014	0.7005
Cycling	0.6204	0.6825	0.6500
Folding-Laundry	0.3009	0.0212	0.0397
House-Cleaning	0.4078	0.3367	0.3688
Ironing	0.5461	0.7302	0.6248
Vacuum-Cleaning	0.5720	0.5188	0.5441

Table 11. Overall accuracy, precision, recall & f1-score

Classifier	Accuracy	Precision	Recall	F1-Score
RF	0.9277	0.9278	0.9277	0.9274
KNN	0.9091	0.9103	0.9091	0.9091
DT	0.8441	0.8443	0.8441	0.8442
MLP	0.8012	0.8020	0.8012	0.7998
SVM	0.7692	0.7679	0.7692	0.7647
GB	0.7630	0.7619	0.7630	0.7593
AB	0.6461	0.6339	0.6461	0.6256
GNB	0.6053	0.5963	0.6053	0.5731
LR	0.5547	0.5231	0.5547	0.5207

computer-work	35956	4998	4	432	3184	290
cycling	2186	33445	795	2741	6175	3660
folding-laundry	4883	887	637	9514	10756	3323
house-cleaning	8654	4804	450	18840	14184	9024
ironing	4842	3698	30	9015	52210	1709
vacuum-cleaning	1272	6081	201	5658	9103	24062
	computer-work	cycling	folding-laundry	house-cleaning	ironing	vacuum-cleaning

Figure 10. Confusion matrix for Logistic Regression

In contrast, the Logistic Regression (LR) model demonstrated the weakest performance, with an overall accuracy of 55%. Figure 10 shows the confusion matrix for activity recognition, and Table 10 displays the precision, recall, and F1-score for each activity.

Comparative Analysis:

Table 11 provides a comprehensive overview of the performance of all classifiers, summarizing accuracy, precision, recall, and F1-score. As noted, the overall precision, recall, and F1-score for these classifiers are calculated using a weighted average. The results indicate that the Random Forest (RF) classifier achieved the highest performance, followed closely by K-Nearest Neighbors (KNN). In contrast, Logistic Regression (LR) exhibited the poorest performance among the nine classifiers assessed for recognizing complex activities using inertial sensors. Additionally, it is worth mentioning that Support Vector Machine (SVM) required the longest time to build the model, exceeding 10 hours, while Gaussian Naive Bayes (GNB) completed the process in just a few minutes. Previous studies have primarily focused on comparing simple daily activities using data from wearable sensors.

Discussion:

The results of the tests illustrate the effectiveness of using wearable sensor data with machine learning classifiers to recognize complex human activities. This study emphasizes the critical role of inertial sensors in identifying such activities, particularly in contexts where privacy and security are significant concerns. Furthermore, it provides a comparative analysis of various machine learning algorithms, aiding researchers in selecting the most suitable methods for similar studies. Ultimately, the primary contribution of this research lies in the performance comparison of different machine learning models.

Classifier Performance:

The strong performance of the Random Forest (RF) classifier highlights the effectiveness of ensemble methods for human activity recognition (HAR) tasks. The K-Nearest Neighbors (KNN) classifier, a non-parametric lazy learner, also showed commendable results, with the number of neighbors set to 5 for these experiments. Both the Decision Tree and Multi-Layer Perceptron (MLP) classifiers achieved reasonable accuracies of 84% and 80%, respectively. In contrast, the poorer performance of Logistic Regression underscores the limitations of basic, non-ensemble classifiers in handling complex activity recognition tasks.

Practical Implications:

The results are significant for the development of human activity recognition (HAR) systems across various applications, including human-computer interaction, sports performance monitoring, and healthcare. For real-world HAR applications where high accuracy and robustness are essential, classifiers such as Random Forest (RF) and K-Nearest Neighbors (KNN) should be considered.

Comparison with Previous Studies:

The outcomes align with previous studies [16-30] that demonstrated the effectiveness of classifiers such as Random Forest (RF), K-Nearest Neighbors (KNN), Decision Tree (DT), Multi-Layer Perceptron (MLP), and Support Vector Machine (SVM) in human activity recognition (HAR) tasks. However, this investigation expands on earlier research by evaluating a broader range of classifiers and comparing their capabilities in recognizing a diverse array of activities.

Limitations and Future Directions:

This investigation focuses on the activities, sensor combinations, and machine learning classifiers. Future studies could explore the performance of ML classifiers in recognizing additional activities and utilizing different sensor modalities. Moreover, this study did not address the impact of ambient factors, such as noise and sensor orientation, on activity recognition. Incorporating these elements could enhance the robustness of HAR systems.

Conclusion:

This study evaluated the efficacy of nine machine learning classifiers in identifying complex human activities using data from wearable sensors. The activities examined included computer work, cycling, folding laundry, vacuuming, ironing, and housecleaning. Our findings

indicate that the Random Forest (RF) classifier outperformed the others, achieving the highest F1-score, accuracy, precision, and recall. The K-Nearest Neighbors (KNN) classifier also demonstrated competitive performance, making it a suitable choice for activity recognition tasks. In contrast, the Logistic Regression (LR) model had the lowest performance. These results highlight the effectiveness of ensemble and instance-based classifiers like RF and KNN in complex human activity recognition from wearable sensor data. The implications of this research are significant for developing accurate and reliable human activity identification systems across various sectors, including sports, healthcare, and human-computer interaction. Further investigation is needed to explore the effectiveness of these classifiers in different activity recognition scenarios and sensor configurations.

References:

- [1] Y. Wang et al., "Ensemble Classification of Accelerometry Data for Continuous Activity Detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 38, no. 9, pp. 1868-1880, Sep. 2016.
- [2] M. Shoaib et al., "A Survey of Deep Learning Techniques for Social Signal Processing," *IEEE Access*, vol. 7, pp. 118634-118678, 2019.
- [3] O. D. Luis et al., "Deep Learning for Human Activity Recognition using Multiple Sensors," *IEEE Sensors Journal*, vol. 16, no. 12, pp. 5366-5375, Dec. 2016.
- [4] A. M. Khan et al., "Towards Wearable Systems for Gait Analysis: A Survey," *IEEE Reviews in Biomedical Engineering*, vol. 7, pp. 1-13, 2014.
- [5] Reiss, A. (2012). PAMAP2 Physical Activity Monitoring. UCI Machine Learning Repository. [Online]. Available: <https://doi.org/10.24432/C5NW2H>.
- [6] Kwapisz, J. R., Weiss, G. M., & Moore, S. A. (2011). Activity recognition using cell phone accelerometers. *ACM SIGKDD Explorations Newsletter*, 12(2), 74-82.
- [7] Anguita, D., Ghio, A., Oneto, L., Parra, X., & Reyes-Ortiz, J. L. (2013). A public domain dataset for human activity recognition using smartphones. 21st European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning (ESANN).
- [8] Bao, L., & Intille, S. S. (2004). Activity recognition from user-annotated acceleration data. *Pervasive computing*, 3001, 1-17.
- [9] Ronao, C. A., & Cho, S. B. (2016). Human activity recognition with smartphone sensors using deep learning neural networks. *Expert Systems with Applications*, 59, 235-244.
- [10] Ordóñez, F. J., & Roggen, D. (2016). Deep convolutional and LSTM recurrent neural networks for multimodal wearable activity recognition. *Sensors*, 16(1), 115.
- [11] Chen, L., Duan, Y., Zhang, J., & Wang, J. (2020). Transfer learning for human activity recognition: A survey. *IEEE Access*, 8, 57909-57928.
- [12] Hamid, R., Tufail, A., Alam, M. M., & Saba, T. (2018). Deep convolutional neural networks for human activity recognition using mobile sensors. *Wireless Personal Communications*, 103(2), 889-901.
- [13] Banerjee, S., Bhardwaj, A., Misra, A., & Saha, A. (2020). Human activity recognition using wearable sensors and deep learning. *Procedia Computer Science*, 167, 279-288.
- [14] Banos, O., Galvez, J. M., Damas, M., Pomares, H., & Rojas, I. (2014). Window size impact in human activity recognition. *Sensors*, 14(4), 6474-6499.
- [15] A. Bulling, U. Blanke, and G. Schiele, "A Tutorial on Human Activity Recognition Using Body-Worn Inertial Sensors," *ACM Computing Surveys (CSUR)*, vol. 46, no. 2, pp. 1-33, Oct. 2014.
- [16] Figo, D., Diniz, P. C., Ferreira, D. R., & Cardoso, J. M. (2010). Preprocessing techniques for context recognition from accelerometer data. *Personal and Ubiquitous Computing*, 14(7), 645662.
- [17] Ellis, K., Godbole, S., Kerr, J., Lanckriet, G., & Wing, D. (2016). Multi-sensor physical activity recognition in free-living. *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp'16)*, 506-517.

- [18] Khan, A. M., Siddiqi, M. H., & Lee, S. W. (2010). Human activity recognition using hierarchical hidden Markov models with automatic feature selection. *Journal of Ambient Intelligence and Humanized Computing*, 1(1), 43-52.
- [19] Ravi, D., Wong, C., Lo, B., & Yang, G. Z. (2016). Deep learning for human activity recognition: A resource-efficient implementation on low-power devices. *Proceedings of the ACM Conference on Embedded Networked Sensor Systems (SenSys'16)*, 71-84.
- [20] illelsen, P., Schermer, T., & Coninx, K. (2013). Recognizing activities with a Kinect sensor. *Proceedings of the 15th International Conference on Multimodal Interaction (ICMI'13)*, 533-540.
- [21] Lee, Y. S., & Mase, K. (2002). Activity and location recognition using wearable sensors. *IEEE Pervasive Computing*, 1(3), 24-32.
- [22] He, Y., Li, Y., Duan, L., & Chen, W. (2015). Activity recognition with smartphone sensors. *Tsinghua Science and Technology*, 20(3), 235-249.
- [23] Gjoreski, M., Ciliberto, M., D'Andrea, V., Di Castro, D., Ermes, M., Gjoreski, H., & Lombriser, C. (2016). Activity recognition on smartphones and smartwatches. *Pervasive and Mobile Computing*, 36, 1-17.
- [24] Wu, M., & Hu, Y. (2016). Real-time activity recognition on smartphones using deep learning. *Journal of Ambient Intelligence and Humanized Computing*, 7(6), 783-798.
- [25] Alsheikh, M. A., Selim, A., Niyato, D., Doyle, L., Lin, S., & Tan, H. P. (2016). Deep activity recognition models with triaxial accelerometers. *Proceedings of the 25th International Joint Conference on Artificial Intelligence (IJCAI'16)*, 2670-2676.
- [26] Wang, H., Li, Z., & Huang, K. (2014). Human activity recognition based on smartphone sensors using hybrid feature selection. *Journal of Sensors*, 2014, 1-10.
- [27] Liu, H., & Lee, W. S. (2017). Recognizing daily activities from accelerometer data on smartphones. *Proceedings of the 2017 IEEE International Conference on Pervasive Computing and Communications (PerCom'17)*, 240-248.
- [28] Cleland, I., Kikhia, B., Nugent, C., Boytsov, A., Hallberg, J., Synnes, K., ... & McClean, S. (2013). Optimal placement of accelerometers for the detection of everyday activities. *Sensors*, 13(7), 9183-9200.
- [29] Ignatov, A. D., & Samuylov, K. E. (2016). Human activity recognition using quaternions and machine learning algorithms. *Journal of Sensors*, 2016, 1-10.
- [30] Bulling, A., Roggen, D., & Tröster, G. (2014). Wearable ECG for human activity recognition: A review. *Sensors*, 14(4), 6107-6128.
- [31] Kaur, K., & Batth, J. S. (2024, June). Human Activity Recognition using Machine Learning. In 2024 OPJU International Technology Conference (OTCON) on Smart Computing for Innovation and Advancement in Industry 4.0 (pp. 1-5). IEEE.
- [32] Balaha, H. M., & Hassan, A. E. S. (2023). Comprehensive machine and deep learning analysis of sensor-based human activity recognition. *Neural Computing and Applications*, 35(17), 12793-12831.
- [33] Nia, N. G., Kaplanoglu, E., Nasab, A., & Qin, H. (2023, June). Human activity recognition using machine learning algorithms based on IMU data. In 2023 5th International Conference on Bio-engineering for Smart Technologies (BioSMART) (pp. 1-8). IEEE.



Copyright © by authors and 50Sea. This work is licensed under Creative Commons Attribution 4.0 International License.