IJIRMET ISSN (Online): 2456-0448
International Journal of Innovative Research in Management, Engineering, and Technology
Vol. 7, Issue 3, March 2022

Human Activity Recognition Using Smartphone Sensors To Predict The Best Accuracy Based On Machine Learning Algorithms

[1] P.Thulasi Raman,[2] Mr.N.Marudachalam

[1] Research Scholar, Department of computer science, Presidency College (Autonomous), Chennai - 05

[2]Associate Professor, PG & Research Department of Computer Science, Presidency College (Autonomous), Chennai - 05

Abstract: Human activity recognition requires predicting the action of a person based on sensor-generated data. Due to the enormous number of applications possible by modern ubiquitous computing devices, it has sparked a lot of attention in recent years. It categorizes data into actions such as walking, sitting, standing, and lying. The accelerometer and gyroscope were used to generate the sensor data, and the sensor signals were pre-processed with noise filters. The goal of this research is to anticipate the optimum accuracy for machine learning-based approaches for Human Activity Recognition. In this study, supervised machine learning methods such as Logistic Regression, SVM, Decision Tree, and Random Forest were utilized to recognize human behavior using smartphone sensors in a detailed experiment. The proposed machine learning-based technique for accurately predicting human activity involves predicting human actions such as walking, sitting, standing, and lying. When compared to other supervise classification machine learning techniques, the Random Forest classifier algorithm predicts 97.32 percent accuracy. While other classifiers, such as Logistic Regression, scored 94.45%, Support Vector Machine 94.55%, and Decision Tree 95.75% also show good performance. The dataset containing evaluation classification report and confusion matrix to categorize data from priority and the outcome reveals that the efficacy of the suggested machine learning algorithm technique can be compared to the best accuracy using precision, recall, and F1 Score.

Keywords: Human activity recognition, Smartphone Sensor, Accelerometer, Gyroscope, Machine learning.

1. INTRODUCTION

An Example of Human Activity The task of identifying sequences of accelerometer data received by advanced sensors into well-defined movements is known as recognition. Almost everyone nowadays has a smartphone with a large number of sensors that can be utilized as an alternate platform for HAR. The sensors found in smartphones such as the accelerometer, gyroscope, GPS, microphone, and camera can be used to derive the necessary data required for HAR. Human activity recognition has numerous applications in a wide range of domains such that healthcare, social networks, safety, environmental monitoring, transportation, and surveillance systems.

A variety of sensors are used to classify human activities. The two sensors are used in the data collected for this research the Sensors are Accelerometers and Gyroscopes. An accelerometer is an electrical sensor that monitors the acceleration forces acting on an object to calculate its position in space and track its movement [1]. Activity identification is the process of identifying and recognizing a person's actions, such as standing, sitting, walking, and lying [2]. This calculates the Triaxial acceleration (total acceleration) and the projected body acceleration to establish its position in space gyroscope sensor is a device that measures and maintains the orientation and angular velocity of an object [3]. The gyroscope determines the Triaxial Angular Velocity. These two sensors are readily available in smartphones and served the usage of data collection for this research.

This research is organized as follows: we provide an introduction and our contributions in chapter 1. The remainder of this thesis is laid out as follows: In chapter 2, An overview of the literature is provided. In chapter 3, we describe the methodology of this thesis. In chapter 4, The results are presented along with discussion**.**Finally, a conclusion and plans for the future work are in chapter 5.

We evaluate proposed algorithm under different settings and show how this method can improve the overall quality of web recommendation.

II - LITERATURE REVIEW

A literature review is a piece of writing that seeks to summarize the most important aspects of current knowledge and/or methodological approaches to a specific issue. It's a secondary source that discusses previously published content in a given subject area during a specified time period. Its main goal is to keep the reader up to date on current literature on a particular topic, and it also serves as a basis for other goals, such as future research that may be required in the subject. It comes before a study proposal and sometimes just a summary of sources. It usually follows a pattern and incorporates both summary and synthesis. A summary is the reorganization and reshuffling of data, whereas a synthesis is the reorganization and reshuffling of data. It could offer a fresh perspective on old material, blend new and old perspectives, or it could chart the field's intellectual evolution, including significant controversies. Depending on the situation, a literature review analyses sources and informs the reader of the most relevant and current ones.

2019: Lisa C. Günthera,*, Susann Kärchera, Thomas Bauernhansl.[4] Different data analysis and machine learning algorithms for detecting manual production processes from sensor data are presented. All sensors are mounted to tools, in this case, a screwdriver, because human activity identification algorithms are not always suitable in industrial situations. A dataset is constructed and evaluated that includes various tool movements, sensor kinds, and mounting possibilities.

2017: Boÿzidara Cvetkovi«c, Robert Szeklicki, Vito Janko, Przemyslaw Lutomski, Mitja Luÿstrek[5] Activity monitoring is a critical task in the lifestyle and health domains, where a person's physical activity is crucial for delivering individualized suggestions. To make such services available to a wider audience, one should leverage devices that the majority of users already own, such as smartphones with a real-time activity tracking algorithm that combines data from smartphone sensors for activity recognition and estimation of the user's energy expenditure.

2021: Ivan Miguel Pires a,b,c,*, Faisal Hussain d, Gonçalo Marques e, Nuno M. Garcia a.[6] The performance of a machine learning algorithm changes depending on the type of sensing device, the number of sensors in that device, and the position of the underlying sensing device, according to this paper. Furthermore, the performance of machine learning algorithms is influenced by incomplete actions in a dataset.

2021 : Maryam Banitalebi Dehkordia,_, Abolfazl Zarakia, Rossitza Setchia.[7] The importance of feature selection and its impact on simplifying the activity categorization process, which increases the computing complexity of the system, is discussed in this work. This approach is unusual in that it identifies the most effective features for detecting each activity separately. Can research how to achieve an appropriate feature set, using which the system complexity can be reduced while the activity detection accuracy remains high, in an experimental investigation with human users and utilizing different cellphones.

III - PROPOSED METHODOLOGY

The approach for recognizing human activities is similar to that of a general-purpose pattern recognition system, and it includes several steps from data collection through activity classification. This procedure entails a series of changes of raw sensor data to develop efficient human activity, classification models Fig [2].

IJIRMET ISSN (Online): 2456-0448
International Journal of Innovative Research in Management, Engineering, and Technology
Vol. 7, Issue 3, March 2022

3.1 Proposed Architecture:

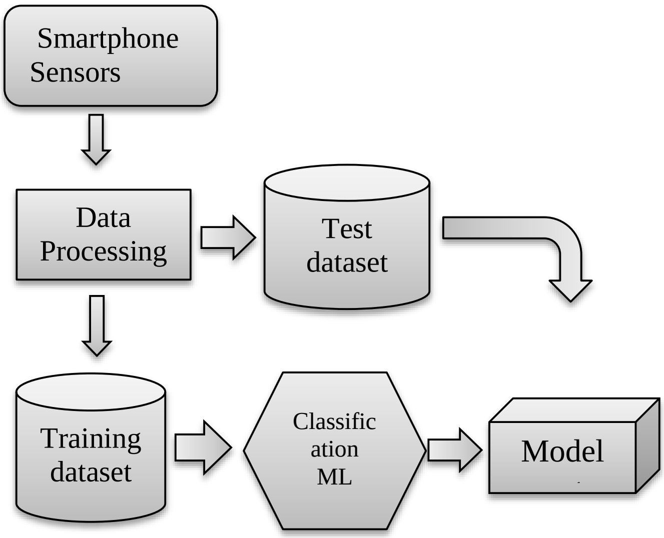

Fig 2: Architecture of Proposed model

The trials were carried Table [1] out on a group of 30 people ranging in age from 19 to 48 years old. Each individual performed four behaviours while wearing a smartphone (Samsung Galaxy S II) around their waist [8]. (WALKING, SITTING, STANDING, LYING). The data was generated by the accelerometer and gyroscope sensors on the sensor [9]. Using the device's embedded accelerometer and gyroscope [10], we recorded 3-axial linear acceleration [11] and 3-axial angular velocity at 50Hz [12]. In order to manually classify the data, the tests were video-recorded [13].

3.2 Table shows details of the datasets:

Variable	Description
Smart Phone	Samsung (Galaxy S II)
Sensors	Accelerometer and Gyroscope
Axis	3-axis(x, y, z)
No. of volunteers	30
Volunteers Age	19-48
Features	19
Activates	WALKING ,SITTING ,STANDING ,LYING

Table 1: datasets

Inconsistencies could be caused by missing values in the data [14]. Preprocessing data is necessary to improve the algorithm's performance. The outliers, as well as variable conversion, must be removed [15].

Import library packages with a given dataset to load. Assessing missing and duplicate values, as well as identifying variables based on data structure and kind [16].

Machine learning validation processes are used to get the error rate of the Machine Learning (ML) model as close to the genuine error rate of the dataset as possible [17]. Validation techniques may not be required if the data volume is large enough to be representative of the population [18].

Assessing missing and duplicate values, as well as identifying variables based on data structure and kind [19]. A sample of data is used to test the fit of a model on the training dataset while tuning model hyper parameters.

Data gathering, analysis, and dealing with data substance, quality, and the organization can all take time [20]. Understanding your data and its properties is helpful during the data identification phase [21] this knowledge will assist you in deciding the strategy to use to create that model.

IJIRMET
ISSN (Online): 2456-0448
International Journal of Innovative Research in Management, Engineering, and Technology
Vol. 7, Issue 3, March 2022

3.3. Data exploration analysis of visualization:

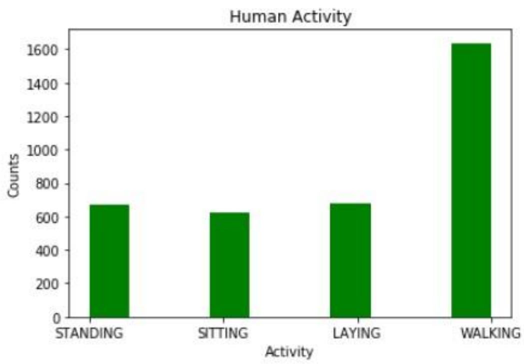

Fig 3: histogram

Data visualization is a critical skill in applied statistics and machine learning. Statistics is concerned with the description and estimation of quantitative data. Data visualization is a valuable set of tools for acquiring a qualitative understanding of data [22]. This might be useful for spotting patterns, faulty data, outliers, and other things when exploring and getting to know a dataset. Fig [3].

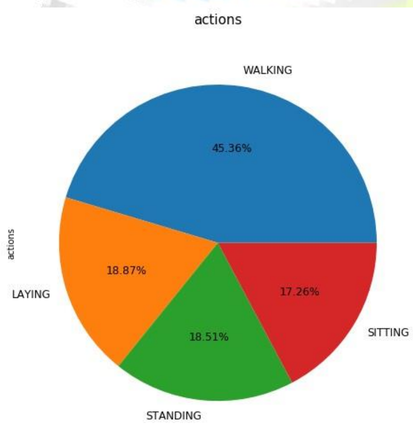

Fig 4: *pie plot*

The pie plot was used to divide the data into four categories: sitting (17.26%), standing (18.51%), lying (18.87%), and standing (45.36%). Data may not make sense unless it is presented in a visual format, such as charts and graphs [23]. The ability to visualize data samples and other objects quickly is a crucial talent in both applied statistics and applied machine learning. Fig [4].

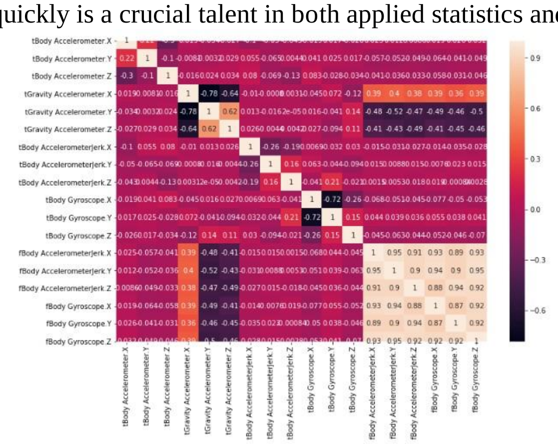

Fig 5: heat map

A heat map is a graphical depiction of a matrix's values rendered as colors [24]. A heat map is a great way to see how values are concentrated between two dimensions of a matrix. This aids with pattern recognition and provides a sense of depth. Fig [5].

3.4. Training & Testing:

There are two components to the data set for predicting provided data: training and testing. In most cases, 7:3 ratios are used to divide the Training and Test sets [25].
Supervised Machine Learning is used in the vast majority of actual machine learning applications [26]. You can use an algorithm to learn the mapping function from the input to the output, which is y = f when you have input variables (X) and output variables (y) (X). The goal is to get to the stage where you can forecast the output variables (y) for new input data by estimating the mapping function (X). Supervised machine learning approaches include logistic regression, multi-class classification, Decision Trees, and support vector machines. Before the data used to train the algorithm can be used for supervised learning, it must be labeled with the right answers. A classification model tries to deduce something from the data. A categorization model aims to predict the value of one or more outputs using one or more inputs [27]. Human activities have long been used to classify human behavior. Sensors on mobile platforms [28] have facilitated the development of a diverse range of practical applications in a variety of sectors.

IV - EXPERIMENTAL RESULTS

4.1. RESULTS

The main goal of this study is to distinguish them based on human behaviors to anticipate them with the greatest accuracy. In this case, we trained our suggested model using the Confusion Matrix and a cross-validation method. To estimate the performance of our suggested model, we used the accuracy, precision, recall, and f1-score measures. The findings of the first stage are based on the identification of four human physical activities, as given in Table [1] sitting, standing, lying, and walking. The Random Forest classifier performs better than the other specified classifiers based on the average values of performance metrics obtained. Random Forest classifier achieves a maximum average accuracy rate of 97.32 percent, demonstrating its superiority over other classifiers. Findings for second-level context recognition, with a comparison of four classifiers based on the average values of the chosen performance criteria. The accuracy rate for context detection based on four human physical activities was examined using a supervised machine learning technique of Logistic Regression, Support Vector Machine, Decision Tree, and Random Forest classifiers. The Random Forest classifier, according to the data in Table [8], performs best for context detection based on the activity. Based on activities, the suggested approach achieves the highest average context recognition accuracy of lying, sitting, standing, and walking activity classified by Logistic Regression (LG), Support Vector Machine (SVM), Decision Tree (DT), and Random Forest (RF) with percentages of 94.45%, 94.55%, 95.75%, and 97.32%, respectively, using Random Forest classifier. Based on the findings it can be inferred that recognizing activity contexts in walking and sitting activities is considerably more difficult than in lying and standing activities. Because certain behavioral settings have a negative impact on human physical activity patterns. Indoor and outdoor surroundings, as well as shopping and speaking, have a bigger impact on walking patterns. As a result, context recognition accuracy in this scenario is lower than it is for other activities and their context. Because it is a static activity, changes in the aforementioned circumstances do not affect the standing activity pattern/posture. Furthermore, the Random Forest classifier delivers the best results for the proposed model.

The confusion matrices for individual behavioral context recognition findings for four different physical activities are shown in Fig. [17]. These confusion matrices are obtained using a Random Forest classifier. It can be observed from the confusion matrices some contexts are misclassified with each other, although they relate to the same acting class. However, the value of correctly identified instances is still relatively large in each context. As a result, the suggested approach can successfully recognize physical activities and their associated behavioral circumstances. Table [10].

4.2. Logistic Regression:

Activities	precision	recall	f1-score	support	Accuracy
Lying (0)	1.00	1.00	1.00	204	94.45%
Sitting (1)	0.87	0.80	0.83	187	
Standing (2)	0.83	0.89	0.86	201	
Waking(3)	1.00	1.00	1.00	491	

Table 2: Logistic Regression Accuracy Prediction

204	0	0	0
1	149	37	0
0	22	179	0
0	0	0	491

Table 3: Logistic Regression Confusion Matrix

Fig 14: Logistic Regression Machine Plot Diagram

IJIRMET ISSN (Online): 2456-0448
International Journal of Innovative Research in Management, Engineering, and Technology
Vol. 7, Issue 3, March 2022

4.3. Support Vector Machines:

Activities	**precision**	**recall**	**f1-score**	**support**	**Accuracy**
Lying (0)	1.00	1.00	1.00	204	94.55%
Sitting (1)	0.88	0.80	0.83	187	
Standing (2)	0.83	0.90	0.86	201	
Waking(3)	1.00	1.00	1.00	491	

Table4:Support Vector Machine Accuracy Prediction

204	0	0	0
1	149	37	0
0	21	180	0
0	0	0	491

Table 5: Support Vector Machine Confusion Matrix

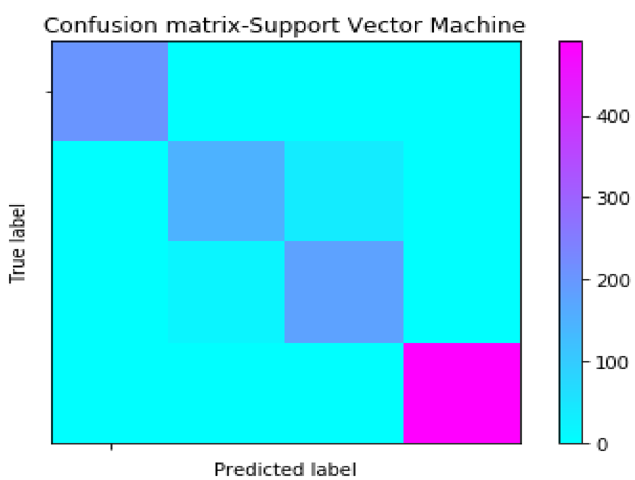

Fig 15: Support Vector Machine Plot Diagram

IJIRMET ISSN (Online): 2456-0448
International Journal of Innovative Research in Management, Engineering, and Technology
Vol. 7, Issue 3, March 2022

4.4. Decision Tree classifier:

Activities	precision	recall	f1-score	support	Accuracy
Lying (0)	1.00	1.00	1.00	204	95.75%
Sitting (1)	0.88	0.88	0.88	187	
Standing (2)	0.89	0.89	0.89	201	
Waking(3)	1.00	1.00	1.00	491	

Table 6: Decision Tree Accuracy Prediction

203	0	0	1
0	165	22	0
0	23	178	0
0	0	0	491

Table 7: Decision Tree Confusion Matrix Table

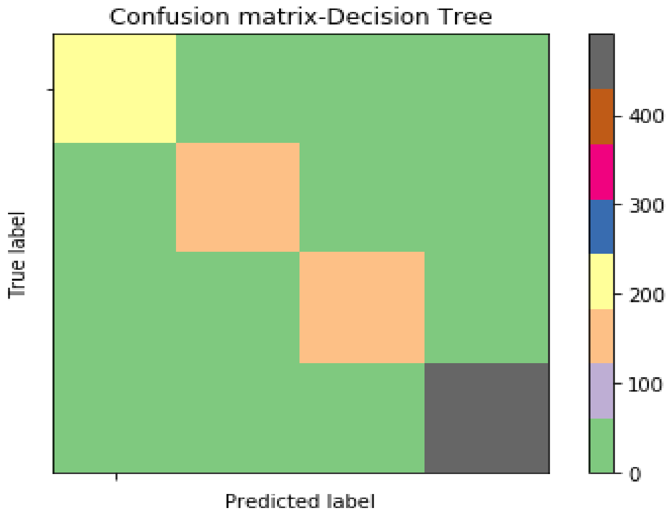

Fig 16: Decision Tree Plot Diagram

4.5. Random Forest:

Activities	**precision**	**recall**	**f1-score**	**support**	**Accuracy**
Lying (0)	1.00	1.00	1.00	204	97.32%
Sitting (1)	0.92	0.94	0.93	187	
Standing (2)	0.94	0.92	0.93	201	
Waking(3)	1.00	1.00	1.00	491	

Table 8: Random Forest Accuracy Prediction

203	0	0	1
0	175	12	0
0	16	185	0
0	0	0	491

Table 9: Random Forest Confusion Matrix Table

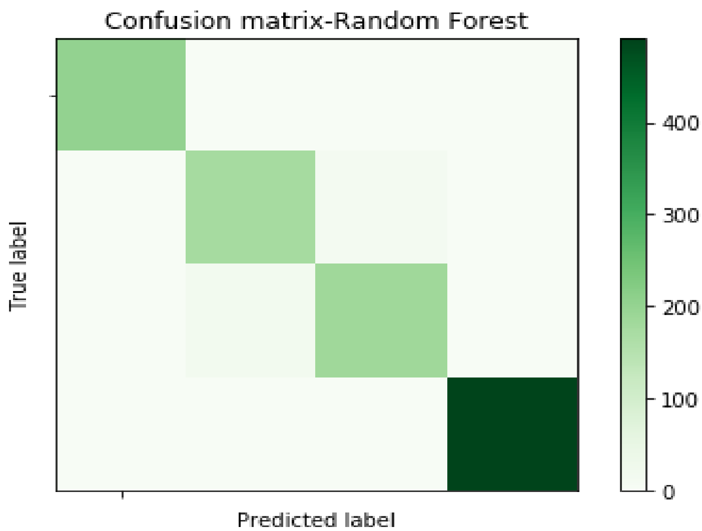

Fig 17: Random Forest Plot Diagram

4.6. SUCCESS RATES OF TESTED MODELS:

Model	Success rate (%)
Logistic Regression	94.45%
Support Vector Machines	94.55%
Decision Tree	95.75%
Random forest	97.32%

Table 10: Success Rates of Tested Models

We have developed and compared several supervised machine learning models to estimate the flask deployment rate Input features given to the classifier with several rows is 3609 and 19 features total of number 68571 data collected from various activities [29]. Data cleaning and processing, missing value analysis, exploratory analysis, and model creation and evaluation were all part of the analytical process. The best accuracy on the test set is a higher accuracy score is found. Among all the methods Random Forest has better accuracy than all others. Random Forest has been found to have an overall accuracy of 97.32 %. The highest accuracy result of the Random forest classification model is now available for deployment [30]. The outcome will predict based on the input values, i.e. independent values. The purpose of the total number of 18 features and dependent value is to forecast the exact result, such as lying, sitting, standing, or walking.

V - CONCLUSION AND FUTURE SCOPE

5.1. Conclusion

Through the implementation of these techniques, the effectiveness of Human activity recognition to predict the action of a person's behaviors like lying, sitting, standing, walking in the processes of detection of recognition of human activities can be predicted. In this study, supervised machine learning algorithms such as Logistic Regression, Support Vector Machines, Decision Trees, and Random Forest was utilized to recognize human behavior using smartphone sensors in a thorough experiment. Input features given to the classifier with several rows is 3609 and 19 features total of number 68571 data collected from various activities. Data cleaning and processing, missing value analysis, exploratory analysis, and model creation and evaluation were all part of the analytical process. On the public test set, the best accuracy is found with a higher accuracy score. Among all the methods Random Forest has better accuracy than all others. The overall accuracy has been obtained for the Random Forest which is 97.32%. Other classifiers that work well include Logistic Regression, Support Vector Machine, and Decision Tree. This application can help to find the Human Activity Based on the Smartphone sensor. The proposed method is highly useful for monitoring people in hospitals and homes.

5.2. Future Work

- Human Activity Recognition will incorporate with the AI model this is recent of work
- A web or desktop application can display the prediction result by automating the process.
- To optimize the labor for Artificial Intelligence implementation.

REFERENCES

1. Geetanjali Vinayak Kale, Varsha Hemant Patil, “A Study of Vision-based Human Motion Recognition and Analysis”, International Journal of Ambient Computing and Intelligence, Volume 7 • Issue 2 • July-December 2016.
2. Manuel Gil-Martín a,*, Rub´en San-Segundo a, Fernando Fern´andez-Martínez a, Ricardo de C´ordoba,“Human activity recognition adapted to the type of movement”, Computers and Electrical Engineering (88) 106822-Year 2020.
3. Francisco Duartea,∗, Andr´e Lourenc¸oa,b, Arnaldo Abrantes,“ Classification of Physical Activities using a Smartphone: evaluation study using multiple users”, Procedia Technology 17 (239 – 247).Year 2014.
4. Lisa C. Günthera,*, Susann Kärchera, Thomas Bauernhansl,“ Activity recognition in manual manufacturing: Detecting screwing processes from sensor data”,procedia CIRP 1177-1182, The Authors. Published by Elsevier Ltd-2019.
5. Boÿzidara Cvetkovi«c, Robert Szeklicki, Vito Janko, Przemyslaw Lutomski, Mitja Luÿstrek, “Real-Time Activity Monitoring with a Wristband and a Smartphone”, Information Fusion-2017
6.Ivan Miguel Pires a,b,c,*, Faisal Hussain d, Gonçalo Marques e, Nuno M. Garcia a,“ Comparison of machine learning techniques for the identification of human activities from inertial sensors available in a mobile device after the application of data imputation techniques”, Computers in Biology and Medicine (135) 104638.Year -2021.
7. Maryam Banitalebi Dehkordia,_, Abolfazl Zarakia, Rossitza Setchia,“ Optimal Feature Set for Smartphone-based Activity Recognition”, 24th International Conference on Knowledge-Based and Intelligent Information & Engineering Systems, Procedia Computer Science 192, 3497–3506.Year 2021.
8. Qimeng Li, Raffaele Gravina, Ye Li, Saeed H. Alsamhi, Fangmin Sun, Giancarlo Fortino, “Multi-user Activity Recognition: Challenges and Opportunities”, Information Fusion-2020.
9. Raffaele Gravina, Parastoo Alinia, Hassan Ghasemzadeh, Giancarlo Fortino,“ Multi-Sensor Fusion in Body Sensor Networks: State-of-the-art and research challenges”, State-of-the-art and research challenges, Information Fusion -2016.
10. Abdulrahman Alruban1, Hind Alobaidi1,2, Nathan Clarke1, 3, Fudong Li1, 4,“ Physical Activity Recognition by Utilising Smartphone Sensor Signals”, Centre for Security, Communications and Network Research, Plymouth University Plymouth.
11. Saurav Jha, Martin Schiemer, Franco Zambonelli, Juan Ye,“ Continual Learning in Sensor-based Human Activity Recognition: an Empirical Benchmark Analysis”, Information Sciences-2021.
12. Michalis Vrigkas1, Christophoros Nikou1* and Ioannis A. Kakadiaris2, “A Review of Human Activity Recognition Methods”, published: 16 November 2015.
13.L.Minh Dang, Kyungbok Min, Hanxiang Wang, Md. Jalil Piran, Cheol Hee Lee, Hyeonjoon Moon, “Sensor-based and vision-based human activity recognition: A comprehensive survey”, Pattern Recognition -2020.
14. Charmi Jobanputraa, Jatna Bavishib, Nishant Doshic,“ Human Activity Recognition: A Survey ”, The Authors. Published by Elsevier B.V, Halifax, Canada, August 2019.
15. Reza Akhavian, Amir H. Behzadan,“ Smartphone-based construction workers' activity recognition and classification”, School of Engineering, California State University, East Bay, 25800 Carlos Bee Blvd, Hayward, CA 94542, USA-2016.
16. Muhammad Ehatisham-ul-Haqa, Muhammad Awais Azamb, Yusra Asima, Yasar Amina, Usman Naeemc,*, and Asra Khalid,“ Using Smartphone Accelerometer for Human Physical Activity and Context Recognition in-the-Wild ”, The 11th International Conference on Emerging Ubiquitous Systems and Pervasive Networks, November 2-5, 2020.
17. Ivan Miguel Piresa,b*, Gonçalo Marquesa, Nuno M. Garciaa and Eftim Zdravevskic, “Identification of Activities of Daily Living through Artificial Intelligence: an accelerometry-based approach ”, The 15th International Conference on Future Networks and Communications (FNC), Leuven, Belgium, August 2020.
18. Santosh Kumar Yadav a,b, Kamlesh Tiwari c, Hari Mohan Pandey d,∗, Shaik Ali Akbar a,b,“ A review of multimodal human activity recognition with special emphasis on classification, applications, challenges and future directions”, Knowledge-Based Systems (223) 106970.Year 2021.
19. Priya Roy, Chandreyee Chowdhury, Mausam Kundu, Dip Ghosh, Sanghamitra Bandyopadhyay,“ Novel Weighted Ensemble classifier for Smartphone-based Indoor Localization ”, Expert Systems with Applications -2020
20. Abdul Rehman Javed a, Raza Faheem b, Muhammad Asim b, Thar Baker c,*, Mirza Omer Beg,“ A smartphone sensors-based personalized human activity recognition system for sustainable smart cities”, Sustainable Cities and Society (71) 102970 Available.Year 2021.
21. Chao Shen ∗, Yufei Chen, Xiaohong Guan,“ Performance evaluation of implicit smartphones authentication via sensor-behavior analysis”, Xi’an Jiao Tong University, C. Shen et al. / Information Sciences (430–431)- (538–553).Year 2018
22. Jiahui When a, b, ∗, Zhiying Wang,“ Learning general model for activity recognition with limited labeled data”, Expert Systems With Applications 74 (19–28), the Year 2017.
23. Anna Ferrari1•Daniela Micucci1 • Marco Mobilio1 • Paolo Napoletano1, “Trends in human activity recognition using smartphones”, Journal of Reliable Intelligent Environments (2021) 7:189–213.Year 2021.

24. Erhan BÜLBÜL, Aydın ÇETİN , İbrahim Alper DOĞRU ,"Human Activity Recognition Using Smartphones", Gazi University,Ankara, TURKEY,erhan.bulbul, IEEE -2018
25. Zameer Gulzar, A.Anny Leema, I.Malaserene," Human Activity Analysis using Machine Learning Classification Techniques", International Journal of Innovative Technology and Exploring Engineering (IJITEE) ISSN: 2278-3075, Volume-9 Issue-2, December 2019
26. Sayandeep Bhattacharjee, Swapnil Kishore, Aleena Swetapadma," A Comparative Study of Supervised Learning Techniques for Human Activity Monitoring Using Smart Sensors ", Second International Conference on Advances in Electronics, Computer and Communications (ICAECC-2018)-2018.
27. Yang Yi, Yang Cheng, Chuping Xu," Mining human movement evolution for complex action recognition", Expert Systems With Applications -2017.
28. Rubén San-Segundo a,*, Henrik Blunck b, José Moreno-Pimentel a, Allan Stisen c, Manuel Gil-Martín," Robust Human Activity Recognition using smartwatches and smartphones", Engineering Applications of Artificial Intelligence (72) 190–202.Year-2018.
29. Manan Jethanandani a , Abhishek Sharma b , ∗, Thinagaran Perumal c , ∗, Jieh-Ren Chang,"
Multi-label classification based ensemble learning for human activity recognition in the smart home", Internet of Things (12) 100324.Year-2020.
30. Bin Liang a, Di Wub, Pengfei Wuc,∗, Yuanqi Su," An energy-aware resource deployment algorithm for cloud data centers based on dynamic hybrid machine learning", Knowledge-Based Systems (222) 107020.Year-2021.