Data Analytics and Artificial Intelligence

Vol: 3(3), 2023

REST Publisher; ISBN: 978-81-948459-4-2

Website: http://restpublisher.com/book-series/daai/

An Automated System for Human Activity Identification and Monitoring for Elderly People Using Data Analytics and Artificial Intelligence Techniques

*** M. Janaki, S. N. Geethalakshmi**

Avinashilingam Institute for Home Science and Higher Education for Women, Coimbatore, Tamil Nadu, India.

*Corresponding Author Email: janakisanjay81@gmail.com

Abstract : *"Human activity recognition" is crucial to the success of a vast number of real-world applications, including the surveillance of the health of elderly people and the identification of abnormal behavior (HAR). A machine learning (ML) model that can accurately identify an activity being done based on raw data collected by wearable technology was developed and tested as part of our present research. This particular study made use of a dataset referred to as the "WISDM Smartphone and Smart watch Activity and Biometrics Database." In addition to the development of AI models, feature selection can help to minimize the overall dimension of a dataset. During the process of determining the values for precision and recall, a confusion matrix was developed for the model. The data that we have chosen to collect is multidimensional, and in order to evaluate human actions, we have employed a number of different machine classification approaches, such as KNN, SVM, DT, and NB. The findings of the experiments demonstrated that the decision tree algorithm that achieved such remarkable outcomes with an exceptionally high degree of precision.*

Keywords: *"Human Activity Recognition (HAR), Machine Learning, Multi-Class Classification, Wearable's Sensor data, Meta-Heuristic Algorithms, Smartphone".*

1. INTRODUCTION

The identification of human activity in computer vision is one of the most significant and demanding topics (HAR). Humans regularly perform a variety of routine and official chores in their daily lives, including driving, cleaning, playing games, and others. It has significant applications in many different fields like robotics, therapy etc. All of these activities have standing, sitting, and walking as their main components. The use of sensor data focuses on automatically identifying the actions done by the adults.[14].

1.1. Human Activity Recognition: Sensor-based physiological monitoring of senior citizens has the potential to significantly enhance their quality of life and avert unfavourable health-related events. Seniors who want to monitor their health may find it helpful to monitor their regular exercise routines. The system for caring an elderly people faces a big challenge with the detection of human activity. People are interested in building systems for the long-term monitoring of human subjects utilising wearable monitoring devices as a result of improvements in sensor technology [1]. Data collection, feature selection, model development, testing, and performance assessment are the four primary stages of the HAR process [9]. The following method is used to conduct an analysis on the HAR system that is depicted in Figure 1.

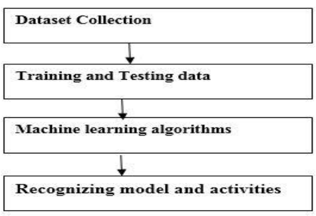

FIGURE 1. Flowchart of HAR

Figure 2 illustrates the use of sensors to identify human activities by sensor signal, feature selection, training and testing.

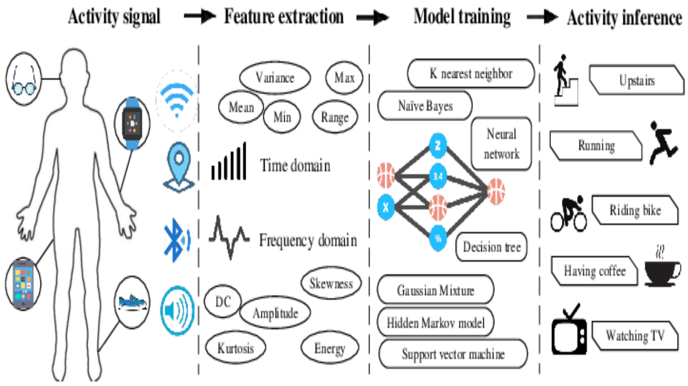

FIGURE 2. Using sensors to recognize Human Activity

2. RELATED WORK

M. M. Hossain Shuvo (2017) researcher demonstrated a 1D CNN-based approach for recognizing elderly movement that makes use of accelerometer data collected from the user's smart phone. In the human activity category, our solution fared better than the industry standard random forest method. The classification accuracy was the highest when longer duration (i.e., Features) accelerometer data was utilized, the neural network must be trained. Author suggest that the performance of activity recognition is impacted by the input vector's dimension. [1] The most current innovations and developing trends in smartphone-based human activity recognition were studied by Anna Ferrari1(2021). For the activity recognition approach, we explicitly went through the processes of data collection, preprocessing, feature selection, and grouping.[2] Zhu et al.(2017) the author has presented a HAR system made up of modules like feature extraction and activity identification, for identifying 33 different physical activities. The proposed system, the author concluded, is a competitive option for applications involving home care monitoring.[3] For the sake of creating and testing the models, the author Jakaria Rabbi (2021) use the HAR dataset that is available in the UCI Machine Learning Repository as a reference dataset. Throughout the entirety of the evaluation, the Support Vector Machine performed noticeably better than the other approaches (average accuracy 96.33%). [4] According to Vasudavan et al. (2021) the majority of elderly individuals live alone, hence it is crucial for their relatives to keep an eye on their daily activities. It is not always possible to be with older folks all the time due to a hectic schedule. Families choose remote monitoring as a result. The ability to monitor an elderly person's activity and get alerts in the event of unusual behaviour, such as a fall, allows older persons to stay safe while requiring little care. Protect the standard of living of older persons who wish to live freely without any manual help.[6].

3. EXISTING SYSTEM

TABLE 1. Existing works in the HAR system: Summary

Sl. No	Author(year)	Title	Methods	Accuracy
1	Jakaria Rabbi et al.,(2021) [4]	Human Activity Analysis and Recognition from Smartphone using Machine Learning Techniques	SVM	96.33%
2	Z Chen et al.,(2017) [5]	Robust human activity recognition using Smartphone sensors via CT-PCA and online SVM	CT-PCA ,SVM	95%
3	K. Butchi Raju ,Suresh Dara (2022) [18]	Smart Heart Disease Prediction System with IOT and Fog Computing Sectors Enabled by Cascaded Deep Learning Model	GSO-CCNN	94%
4	Kun Xia et al., (2020) [7]	LSTM-CNN Architecture for Human Activity Recognition	LSTM-CNN	95.85%
5	Federico Cruciani et al., (2019) [19]	Feature learning for Human Activity Recognition using Convolution Neural Networks	3-CNN	92.30%
6	Syed K. Bashar et al., 2020 [9]	Smartphone Based Human Activity Recognition with Feature Selection and Dense Neural Network	Neighborhood component Analysis(NCA) -SVM	95.79%

4. PROPOSED SYSTEM

The framework that was used in this paper, which was based on an optimization model for human activity analysis using sensor data (the Benchmark dataset), was used to build a strategy that visually illustrates and improves the classification performance by utilising optimization techniques and machine learning. The framework was used to build this strategy. The following outline constitutes the framework for this investigation: The second part of this chapter addresses the topic of the literature review. The third part of the chapter offers a comprehensive illustration of the system architecture. The fourth part of the chapter discusses the dataset, the evaluation methods that were used in this study, and the experimental results that were obtained by utilising the suggested system. The sixth section provides an overview of the primary findings and conclusions. In Figure: 3 Proposed System Architecture is shown

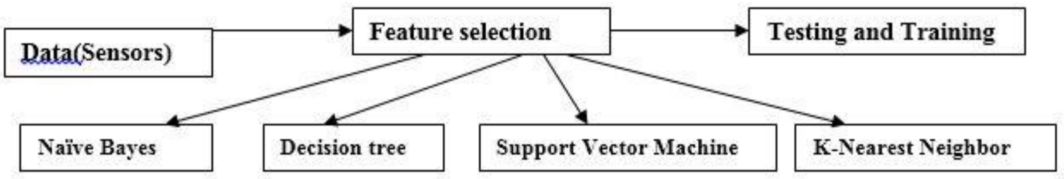

FIGURE 3. Architecture Diagram of the Proposed System

5. MATERIALS AND METHOD

The WISDM (Wireless Sensor Data Mining) Smartphone, Smart Watch Activity, and Biometrics dataset was utilised in the current investigation to make predictions regarding human activity. Human activity detection is the process of predicting a person's activities by using a sensor-based mobility record as the data source. For the "WISDM (Wireless Sensor Data Mining) dataset," information was collected from 7,352 different individuals. Every individual subject spent three minutes completing six different tasks. In addition to wearing a smart watch on their dominant wrist, each participant also carried a smart phone in their pocket. The procedure of data collection was directed and supervised by a specially developed programme that could be executed on a smart watch as well as a smart phone. For the purpose of recording the sensor data, the accelerometer and gyroscope of both the smart watch and the phone were used. This results in a total of four sensors. Based on the data for the subjects being examined, the data can be partitioned into 80% for training and 20% for testing. Participants (referred to as subjects) give accelerometer and gyroscope measurements while undertaking the following 6 Activities. [10].

i. Walking
ii. Walking Upstairs
iii. Walking Downstairs
iv. Standing
v. Sitting
vi. Lying.

A span of 3 minutes with 50% overlap separates the readings. The two types of accelerometer readings are readings of body acceleration and readings of gravity acceleration, both of which contain three axes (x, y, and z). This axis of angular velocities is measured by a gyroscope. Using jerk signals, body acceleration measurements are computed. To acquire frequency readings, Fourier Transforms are applied to the aforementioned time readings. On readings for each base signal for each window, the following metrics are calculated: mean maximum, minimum, standard deviation, coefficient, entropy, etc. The dataset contains 561 features that make up the feature vector that we receive. Each reading window represents a data point with 561 characteristics.

6. METHODOLOGY

When using the feature selection approach, in order to reach a classification accuracy that was satisfactory, the number of features in the recovered features was reduced by getting rid of irrelevant and unnecessary features [6]. This was done so that the recovered features could be used. In classification problems, a common and essential strategy for improving accuracy and accelerating convergence is to pick relevant features. Feature selection is to be converted to a more reliable and suitable form of input for the classifier to classify the different features category. This section covers optimizers for selecting the finest features from an obtained dataset. PSO, Grey wolf and whale optimization feature selection methods are exposed here to compare the suggested optimizer's performance. The optimizers for successful feature selection are discussed in this section. The preliminary backdrop of the PSO, whale, and Grey wolf the working mechanism of the suggested optimizer is discussed using GWO approaches. In figure 4 particle swarm optimization –Decision tree the number of selected features is 266 out of 570 features using matlab

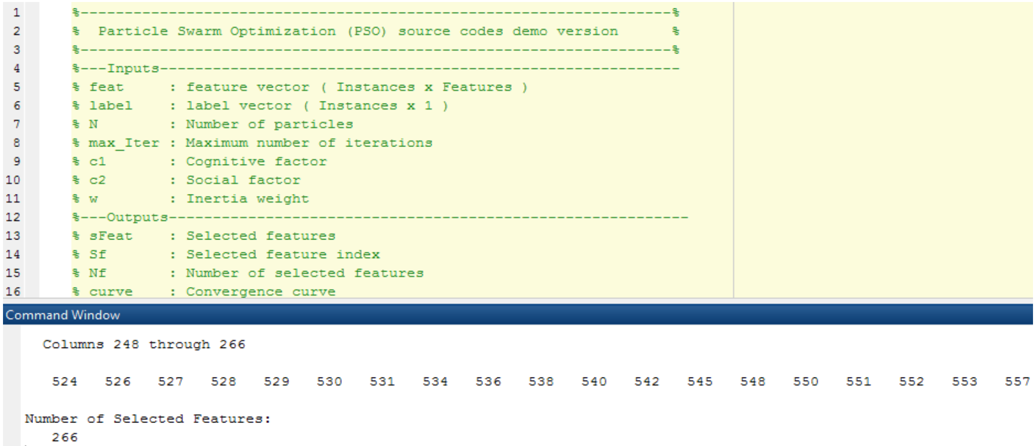

FIGURE 4. Particle Swarm Optimization-Decision Tree

In figure 5 whale optimization –Decision tree the number of selected features is 292 Out of 570 features.

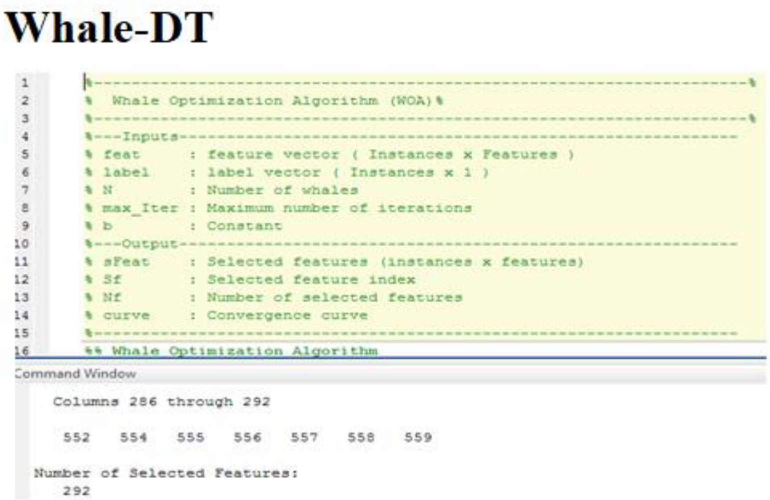

FIGURE 5. Whale –Decision tree

In figure 6 proposed chaotic GWO –Decision tree the number of selected features is 252 out of 570

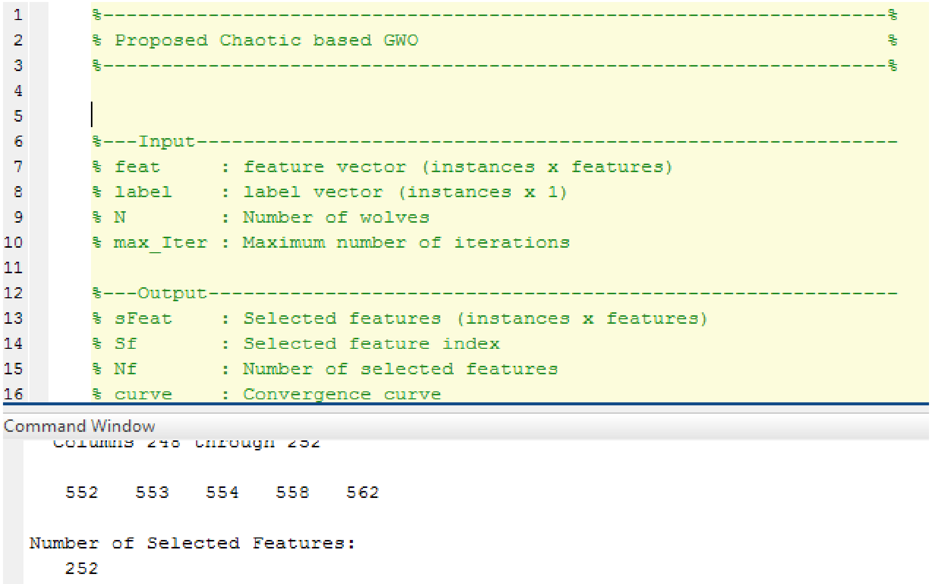

FIGURE 6. Proposed Chaotic GWO – Decision tree

6.1. Chaotic Maps for Feature Selection: The grey wolf optimization (GWO), the genetic algorithm (GA), the particle swarm optimization (PSO), and the whale optimization algorithm (WOA) are all examples of optimization algorithms [17]. We investigate the idea that chaotic systems are difficult to optimise because they are extremely sensitive to the initial conditions at the outset; that is, even minute shifts in the initial conditions can result in significantly different outcomes, and the techniques that are currently in use for optimization have slow convergence speeds. The chaotic technique is utilised in this study to increase GWO performance by minimising stuck situations at local optima and boosting the speed of convergence. In addition to that, it is employed to deal with the random factor values of GWO. It has been demonstrated that the Chaos technique can boost the value of looking for a global optimum in a variety of optimization applications. This is accomplished by avoiding challenges associated with becoming caught in local optima. As a direct consequence of this, the searching range [0, 1] may be utilised in the development of a feature selection system that bases its optimization on chaos theory. [8].

We have three types of chaotic map like
1. Tent map
2. Circle map
3. Logistic map
4. Sine map
5. Gaussian map

6.2. Machine learning algorithms**:** Machine learning (ML) is a highly interdisciplinary field that draws inspiration from many different fields of science and mathematics, including cognitive psychology, computer science, statistics, and optimization.The supervised learning technique of classification is used in machine learning to analyse a piece of data and create a model that divides it into the right number of distinct classes. Machine learning is a vital step in understanding human behaviour. A range of machine learning algorithm types have been deployed by HAR. In Figure 7 we can see the overview of the methodology where datasets are taken from benchmark datasets for research purposes. Here we are using a feature selection strategy based on a metaheuristic optimization algorithm. It is an advanced technique for finding a worthwhile solution to a complex problem. Here we are using optimization algorithms like PSO, GWO, and WO. smearing existing machine algorithms like DT, SVM, NB, and KNN to see how our feature selection algorithm will work in classifying the activities and to check the best result. This is compared using accuracy, precision, and recall.

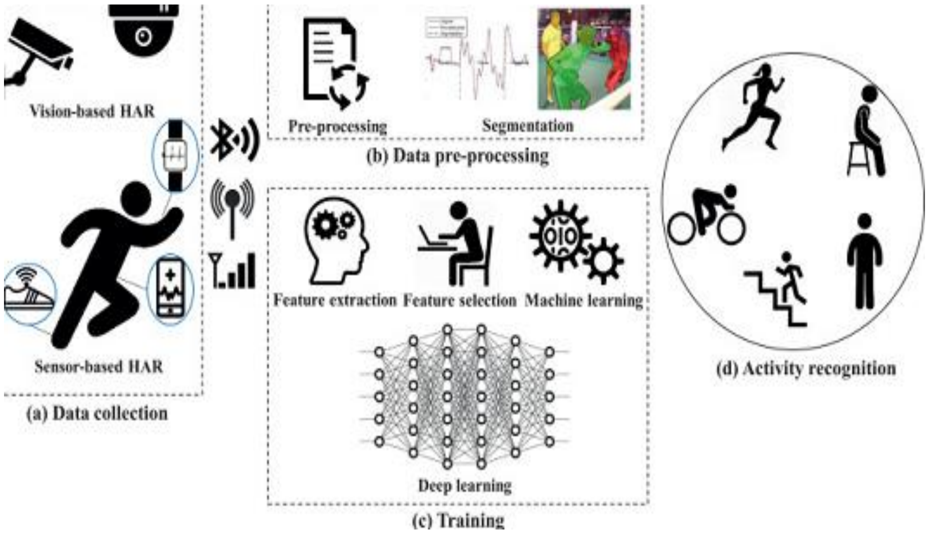

FIGURE 7. Overview of Methodology

6.3. Models Used for Activity Recognition

6.3.1 Naïve Bayes: In order to derive a list of probabilities from a dataset, a fundamental probabilistic analyzer known as "Naive Bayes (NB)" adds up all of the frequencies and values contained in the dataset. This method, which is based on the Bayes theorem, takes into consideration all of the independent and non-dependent characteristics that are revealed by the value of the class variable. The effectiveness of Naive Bayes has been demonstrated in a wide variety of real-world applications, including text classification, medical diagnosis, and the monitoring of system performance, amongst others. [11]. The "Naive Bayes" method offers quite a few distinct advantages. A condensed amount of training data is required in order to calculate the approximations of the values of the classification-related parameters.

6.3.2 Support Vector Machine: In many sectors of application, SVM are primarily reliable and efficient classification and regression techniques. The study of pattern recognition, which is a particularly active and well-liked field of research, has greatly benefited from the use of the SVM.Support Vector Machine has been combined with other cutting-edge methodologies, like evolve algorithms, to enhance classification and parameter optimization. The support vector machine method's main objective is to divide the training set into various classes using a surface that maximizes their distance from one another. In other words, SVM allows for the most generalizability of a model [12].

6.3.3 K- Nearest Neighbor: The K-Nearest Neighbour (KNN) method is straightforward but effective. It works well for both classification and regression techniques. However, it is most typically used in classification prediction. By organising the data into coherent clusters or subsets, KNN categorises newly entered data according to how much it resembles previously trained data. The class with which the input has the greatest number of close neighbours is given the assignment. Unlabelled observations are categorised by the KNN classifier by placing them in the category of the most comparable labelled samples. The main drawbacks of KNN are (1) its low efficiency, which prohibits it from being applied in many applications because it is a lazy learning technique, such as dynamic web mining for a large repository. A case-based learning approach called KNN retains all of the training data for categorization.[13]

6.3.4 Decision Tree Classifier: A normal tree has leaves, branches, and a root system. The format is the same for the "Decision Tree" or "DT" for short. It is composed of a base node, branch nodes, and leaf nodes. Every internal node performs a test on a property, the results of which are displayed on the branch, and the class label is displayed on the leaf node [15]. A root node is the node that is at the very top of a tree and is considered to be the parent of all of the other nodes in the tree. A decision tree node represents qualities, links represent rules, and leaves reflect outcomes (a category or continuing value). Because decision trees simulate the degree of reasoning that a human would employ, it is quite simple to collect the data and come to some fascinating conclusions. The objective is to get everything in order.

Pseudocode: The DT classification algorithm

Contribution: enhanced features

Output: Validity and Accuracy

y - Target Feature; T - Input Feature Set; tree - Training Set

Step 1: Begin

Step 2: Set up a new tree T (feature space) only with one root node

Step 3: Set the training instances as a tree.

Step 4. If the instances of the tree (y->tree) are A, Return leaf node "Class A" in step 5.

Step 6: Alternatively, if the tree instance (y->tree) is B,

Step 7: Return leaf node "Class B" in

Step 8: if **End**

Step 9: Finish

Confusion Matrix for Decision tree **When applying the decision tree plain in the dataset, we get 88.4%. is shown in Figure:8**

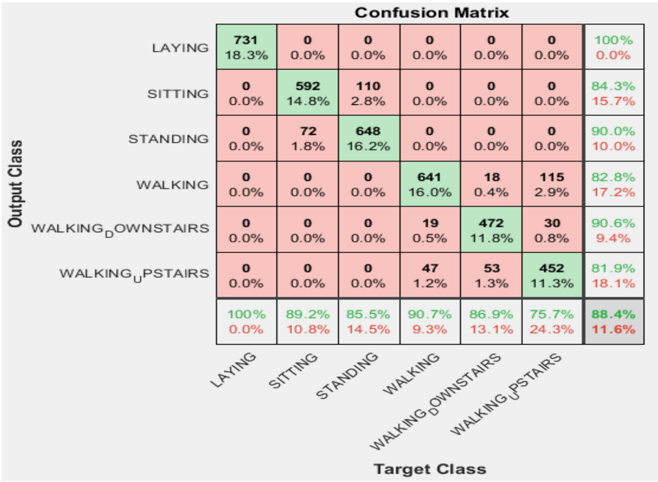

FIGURE 8. Confusion Matrix for Decision tree

When applying the PSO-DT (The PSO algorithm is a hybrid of evolutionary algorithms and swarm system. The particles (Birds) of the swarm are free to fly across the multidimensional search space. Each particle acquires its own velocity and location throughout the voyage. The entire population is updated by updating each particle. The particles gradually concentrate around the upper goal function value as the swarm arrangement forces itself toward it.) in the dataset ,we get 89.9% in Figure :9

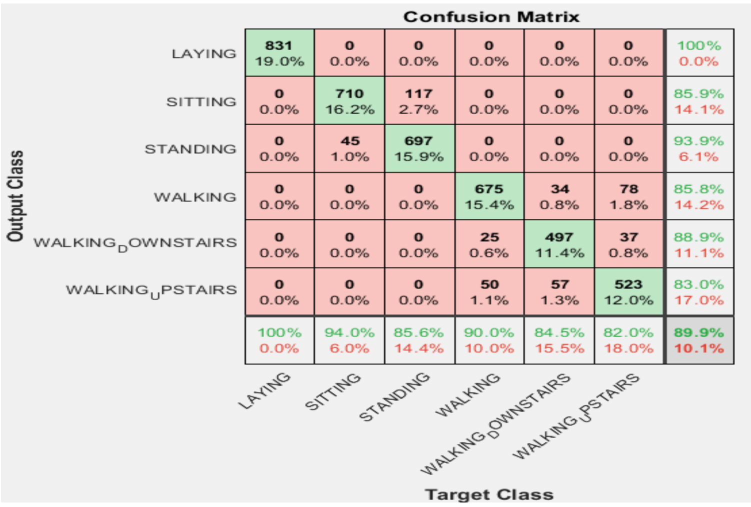

FIGURE 9. Confusion matrix: PSO-DT

When applying the Whale optimization –DT(An innovative method for handling optimization issues is the Whale Optimization Algorithm (WOA). Three operators are used in this algorithm to replicate how humpback whales hunt by searching for prey, circling prey, and using bubble nets.) in the dataset ,we get 88.4% in Figure :10

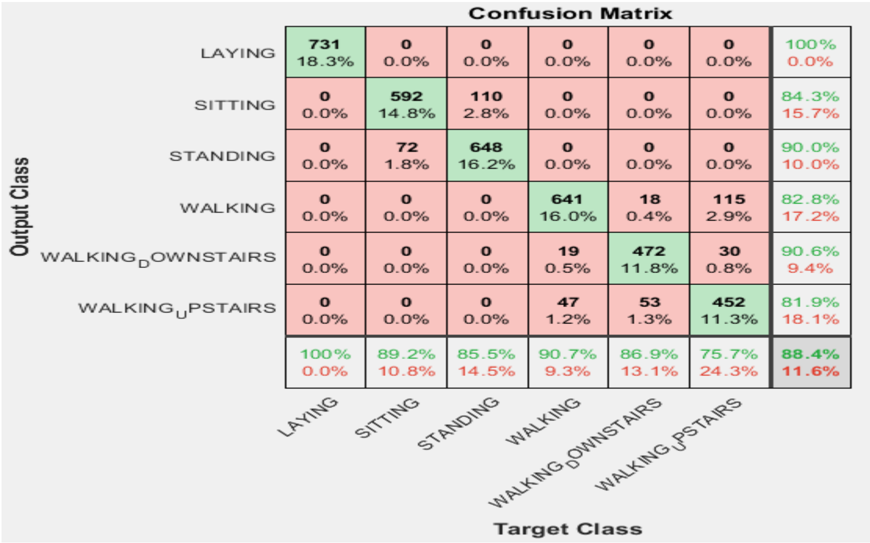

FIGURE 10. Confusion matrix: Whale-DT

When applying the GWO -DT (The leadership structure and hunting strategy of grey wolves in nature are modelled by the GWO algorithm. For the purpose of mimicking the leadership hierarchy, four different varieties of grey wolves, including alpha, beta, delta, and omega, are used.)in the dataset ,we get 89.3% in Figure :11

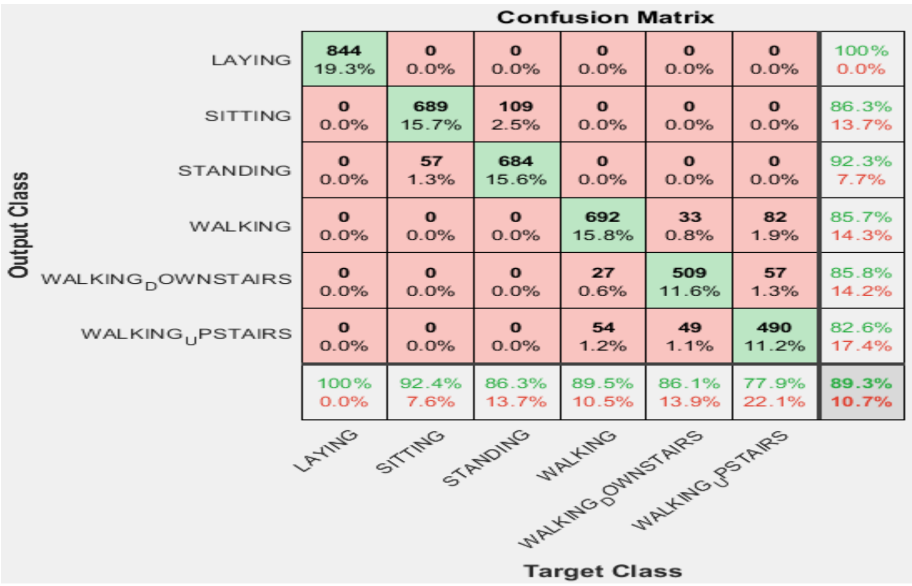

FIGURE 11. Confusion matrix: GWO-DT

Confusion matrix: CLM-GWO: When applying Chaotic logistic map **(**The Chaos technique is proved to increase the value of searching for global optimum in a range of optimization applications by avoiding stuck difficulties in local optima. As a result, the searching range [0, 1] can be used to create a feature selection system that uses chaos theory to optimize features.) **decision tree** plain in the dataset we getting 92.3% in figure :12

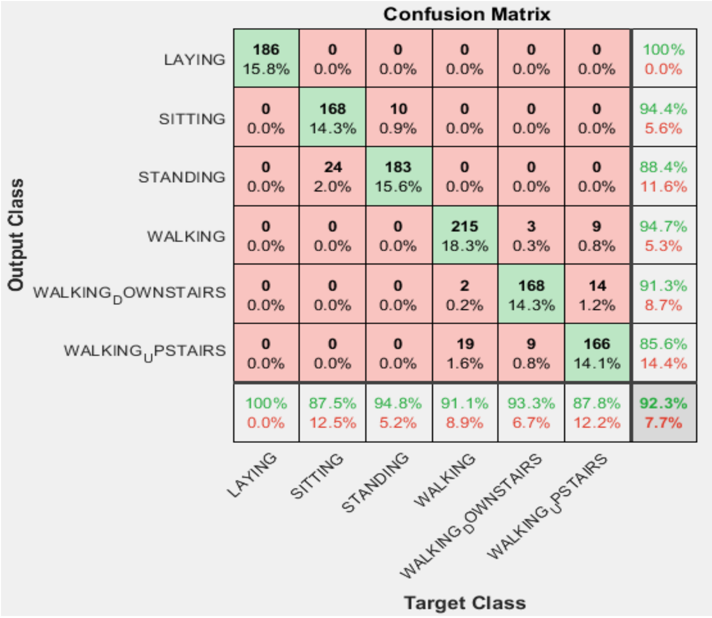

FIGURE 12. Confusion matrix: CLM-GWO

In table 2 we can observe the Different human activities with their count with the Proposed Model

TABLE 2. Different human activities with their count with the Proposed Model

S.No	Human Activities	Number of persons
1	Standing	183
2	Sitting	168
3	Laying	186
4	Walking	215
5	Walking_downstairs	168
6	Walking_upstairs	166
Total		1086

TABLE 3. Performance Comparison with Accuracy

Description	Algorithms used	Accuracy
ML Algorithm	DT	88.4 %
Feature selection Techniques	PSO-DT	89.9 %
	Whale-DT	88.4 %
	GWO-DT	89.3 %
	Proposed	92.3 %

In table 3, When applying Decision tree along with Particle swarm optimization, Grey wolf optimization, Whale optimization and proposed Chaotic logistic map. We get high accuracy with CLM Decision tree 92.3% Performance Comparison of CLM-G WO-Machine Learning with Accuracy.

TABLE 4. Performance Comparison of CLM-GWO-Machine Learning with Accuracy

Algorithm Details	Performance Metrics				
	Accuracy	Sensitivity	Specificity	Precision	F1-Score
CLM-GWO-NB	81.5	80.4	81.0	79.6	80.4
CLM-GWO-SVM	87.56	85.67	85.89	84.37	85.8
CLM-GWO-KNN	88.6	87.45	86.74	85.46	85.78
CLM-GWO-DT	**92.3**	**90.55**	**92.0**	**91.6**	**91.78**

In table 4 when applying CLM-GWO with machine learning algorithms like Naïve bayes, support vector machine, kNN and Decision tree. we get high accuracy with CLM Decision tree 92.3% In figure 13 the performance analysis with machine learning algorithms is shown in chart method

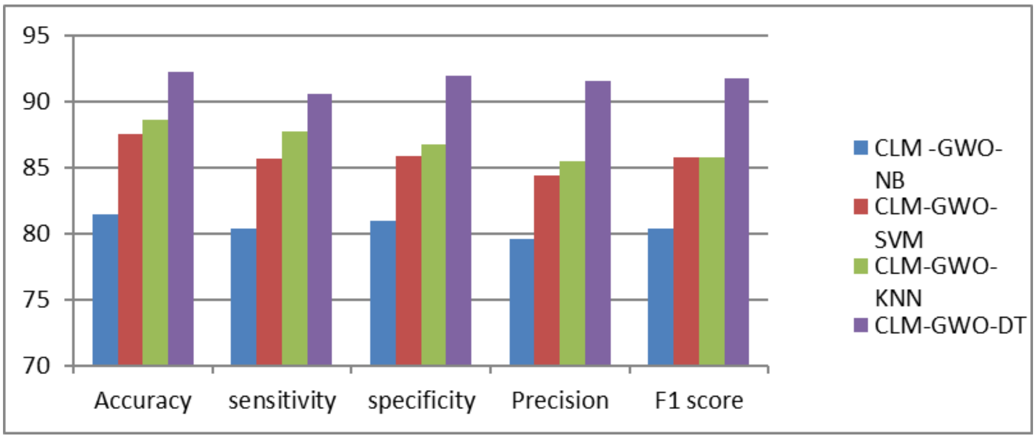

FIGURE 13. Performance Analysis with Machine Learning Algorithms

The collection of observed data from the sensors can be used to examine activities. Wearable sensors can be used for data collection. For instance, wearing an accelerometer on the body can capture data while getting around some limitations like flexibility and privacy.

Justification for Proposed Methodology: Feature optimization is used here to establish recognition models to capture best features from input sequence data to achieve more accurate recognition. This research work utilized the chaotic based GWO with machine learning scheme that can recognize specific activities in short duration of time.

The experimental results show that the proposed approach can help to improve the recognition accuracy rate up to 92.3%.

Performance Metrics and evaluation: To better identify human behaviors, the architecture employs feature selection and machine learning algorithms. The partitioned datasets that were used to train and test the technique are shown in Table 5

TABLE 5. Total number of data sets that were used in the training and testing procedure

Sl. No	Total number of data	Training data	Testing data
01	7352	5882	1470

TABLE 6. Measures of performance

SL.NO	Performance Metrics	Mathematical Expression
01	Accuracy	$\frac{TP+TN}{TP+TN+FP+FN}$
02	Sensitivity or recall	$\frac{TP}{TP+FN}$ x100
03	Specificity	$\frac{TN}{TN+FP}$
04	Precision	$\frac{TN}{TP+FP}$
05	F1-Score	$2.\frac{Precison * Recall}{Precision + Recall}$

True Positive values are denoted by TP, True Negative values by TN, False Positive values by FP, and False Negative values by FN.

7. RESULT AND DISCUSSION

"Decision Tree" abbreviation Methods of machine learning are utilised, as well as the data sets that are chosen by these algorithms for the purposes of training and testing. In order to do an evaluation, we tested using 20% of the data, while we trained with 80% of the data. During the evaluation, the accuracy parameter will be utilised. The entire experiment is powered by a computer with a 3.0 GHz Intel I7 processor, a 2GB NVIDIA Geoforce K+10 GPU, 16GB of RAM, and a 2TB hard drive.

8. CONCLUSION AND FUTURE WORK

According to the findings of the studies, the approach that was recommended has the potential to improve recognition accuracy by as much as 92.3%. This research presents an example of how a decision-making algorithm and wearable sensors might be used to monitor the activities of elderly individuals and maintain track of their whereabouts. The existing method has a few flaws, one of which is that it does not provide precise tracking. Future research will make use of additional sensor data collected from a wide variety of physical activities in order to examine the usefulness of different alternative models. The data that we chose to work with is multidimensional, and in order to evaluate human behaviour, we have employed a number of different machine classification methods, such as k-Nearest Neighbour (kNN), Support Vector Machines (SVM), Decision Trees (DT), and Naive Bayes (NB).

REFERENCES

[1]. Lee, S. M., Yoon, S. M., & Cho, H. (2017, February). Human activity recognition from accelerometer data using Convolutional Neural Network. In *2017 ieee international conference on big data and smart computing (bigcomp)* (pp. 131-134). IEEE.

[2]. Ferrari, A., Micucci, D., Mobilio, M., &Napoletano, P. (2021). Trends in human activity recognition using smartphones. *Journal of Reliable Intelligent Environments*, *7*(3), 189-213.

[3]. Zhu, J., San-Segundo, R., &Pardo, J. M. (2017). Feature extraction for robust physical activity recognition. *Human-centric Computing and Information Sciences*, *7*(1), 1-16.

[4]. Rabbi, J., Fuad, M., Hasan, T., &Awal, M. (2021).Human Activity Analysis and Recognition from Smartphones using Machine Learning Techniques. *arXiv preprint arXiv:2103.16490*.

[5]. Chen, Z., Zhu, Q., Soh, Y. C., & Zhang, L. (2017). Robust human activity recognition using smartphone sensors via CT-PCA and online SVM. *IEEE transactions on industrial informatics*, *13*(6), 3070-3080.

[6]. Yu, L., & Liu, H. (2003). Feature selection for high-dimensional data: A fast correlation-based filter solution. In *Proceedings of the 20th international conference on machine learning (ICML-03)* (pp. 856-863).

[7]. Xia, K., Huang, J., & Wang, H. (2020). LSTM-CNN architecture for human activity recognition. *IEEE Access*, *8*, 56855-56866.

[8]. Duan, L., Yang, S., & Zhang, D. (2020). The Optimization of Feature Selection Based on Chaos Clustering Strategy and Niche Particle Swarm Optimization. *Mathematical Problems in Engineering*, *2020*.

[9]. Bashar, S. K., Al Fahim, A., & Chon, K. H. (2020, July). Smartphone based human activity recognition with feature selection and dense neural network. In *2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)* (pp. 5888-5891). IEEE.

[10]. Rani, S., Babbar, H., Coleman, S., Singh, A., &Aljahdali, H. M. (2021). An efficient and lightweight deep learning model for human activity recognition using smartphones. *Sensors*, *21*(11), 3845.

[11]. Harahap, F., Harahap, A. Y. N., Ekadiansyah, E., Sari, R. N., Adawiyah, R., &Harahap, C. B. (2018, August). Implementation of Naïve Bayes classification method for predicting purchase.In *2018 6th International Conference on Cyber and IT Service Management (CITSM)* (pp. 1-5). IEEE.

[12]. J. Cervantes, J., Garcia-Lamont, F., Rodríguez-Mazahua, L., & Lopez, A. (2020). A comprehensive survey on support vector machine classification: Applications, challenges and trends. *Neurocomputing*, *408*, 189-215.

[13]. Taunk, K., De, S., &Verma, S. MACHINE LEARNING CLASSIFICATION WITH K-NEAREST NEIGHBOURS.

[14]. Subasi, A., Khateeb, K., Brahimi, T., &Sarirete, A. (2020). Human activity recognition using machine learning methods in a smart healthcare environment.In *Innovation in health informatics* (pp. 123-144). Academic Press.

[15]. Patel, H. H., &Prajapati, P. (2018). Study and analysis of decision tree-based classification algorithms. *International Journal of Computer Sciences and Engineering*, *6*(10), 74-78.

[16]. Vasudavan et al: "AStudy on Activity Recognition Technology for Elderly Care",2021 scopus Index.

[17]. Chakraborty, A., &Kar, A. K. (2017). Swarm intelligence: A review of algorithms. *Nature-inspired computing and optimization*, 475-494.

[18]. Raju, K. B., Dara, S., Vidyarthi, A., Gupta, V. M., & Khan, B. (2022). Smart Heart Disease Prediction System with IoT and Fog Computing Sectors Enabled by Cascaded Deep Learning Model. *Computational Intelligence and Neuroscience*, *2022*.

[19]. Cruciani, F., Vafeiadis, A., Nugent, C., Cleland, I., McCullagh, P., Votis, K., ...&Hamzaoui, R. (2020). Feature learning for human activity recognition using convolutional neural networks. *CCF Transactions on Pervasive Computing and Interaction*, *2*(1), 18-32.