Multimedia Tools and Applications (2023) 82:14815–14835
https://doi.org/10.1007/s11042-022-13716-z

Artificial neural networks for human activity recognition using sensor based dataset

Shahab Geravesh[1] · **Vaibhav Rupapara**[2]

Received: 9 January 2021 / Revised: 12 May 2022 / Accepted: 24 August 2022 /
Published online: 7 October 2022

Abstract

The use of wearable devices, sensors, machine learning, and deep learning for human activity recognition (HAR) applications has increased in recent years. Lots of researchers introduce different methods and techniques for HAR but accuracy and efficiency are still a gap to work in the HAR domain for other researchers. In this study, we design a simple architecture of MLP with a stack of dense layers which can perform accurately and efficiently on small and large features set for HAR. This study used a dataset that contain smartphone sensors (gyroscope and accelerometer) data against six daily human activities. Each sensor has three feature values corresponding to an instance for an activity. First, we train our proposed model individually on gyroscope and accelerometer data and then we combine both sensors data to train the model. We train the model using 70% of the dataset and evaluate the performance of the model on 30% data. MLP outperforms all other stat of the art models in comparison such as random forest, decision tree, logistic regression, and K nearest neighbor. The MLP accuracy scores are 0.74, 0.77, and 0.98 using features of gyroscope, accelerometer, and combination of both respectively. The results show that the proposed approach is also good with single sensor data. In comparison with the proposed MLP, we also deployed state-of-the-art models such as LSTM, CNN, and other machine learning models. Proposed MLP outperforms all used models in terms of all evaluation parameters. We also did a statistical T-test to show the significance of the proposed approach.

Keywords Machine learning · Deep learning · IoT · Human activity recognition · Computer science · Data science

1 Introduction

This era belongs to technology which is serving a human in every field of life such as education, health, security, agriculture, business, etc. especially the evaluations in the health

✉ Vaibhav Rupapara
vaibhav.rupapara.sept@gmail.com

Shahab Geravesh
Shahabgeravesh@gmail.com

[1] University of California Riverside, Riverside, CA, 92521, USA

[2] School of Computing and Information Sciences, Florida International University, 11200 SW 8th St ECS 354, Miami, FL, 33199, USA

field using technology most research-able topic nowadays. This study is also about activity detection to monitor human activity, which can be useful in many domains such as it can be used in the health domain to monitor patient activities, it can be useful in smart cities to monitor the domestic activities using the Internet of Things [32]. Activity detection can also be useful for security purposes as it can be helpful for crowd anomaly detection and object tracking [12].

The use of artificial intelligence to detect human activity is trendy nowadays. In recent years the work on human activity recognition has increased and several computational methods are introduced by different researchers [5, 22]. HAR applications are beneficial in security and surveillance processes, to monitor patient health. In practice, for the case of home care, personnel deployment for such tasks often cannot be financially feasible. The development of the human activity recognition application using shallow algorithms such as decision tree, support vector machine, logistic regression, and random forest which is not too accurate and efficient that's the reason the deep learning approach has some major evaluation in this domain. We deploy an approach that can recognize human activity with accelerometer and gyroscope sensor values individually and also with both sensor combinations. Particularly proposed a simple architecture of MLP which can make a more accurate prediction as compared to state-of-the-art machine learning models.

In this study, we have used the deep learning approach for human activity detection using a sensor dataset. The dataset contains the two smartphone sensors gyroscope and accelerometer data. We combine both sensor data in our proposed approach to achieve high accuracy results. First, extract the equal number of instances from each activity, and then we split the dataset into two sets; the training set and the testing set with a ratio of 70:30. The training set we used to train the designed MLP model. The design model is a stack of dense layers, dropout layers, and ReLU and Softmax activation functions see Section 4.5.1. We also used the machine learning models in comparison with the design MLP model such as random forest, decision tree, logistic regression, and K nearest neighbor. The models also get training on the same dataset as MLP. After training of models, we evaluate the performance of all models in terms of five evaluation parameters which are accuracy, precision, recall, f1 score, and Cohn's kappa. Key-point of the study are:

- The performance comparison of machine learning models on human activity dataset.
- Previous studies work on combined features by accelerometer and gyroscope but this study proposed a model which even can perform better with a single sensor which makes it more effective. Proposed a model which can also perform better on the individual sensor (accelerometer or gyroscope) as compared to state-of-the-art models. This study can be helpful in the medical field for the doctor to monitor their patient activities, and for the employer to detect their employee's activity during on duty with low-cost sensors.
- Designed the architecture of artificial neural networks which can perform better on fewer feature data.
- Analysis on accelerometer and gyroscope feature importance for human activity detection.
- The effects of combined features by accelerometer and gyroscope on the machine learning models.

The rest of the paper has the following section. Section 2 contains related work. Section 3 contains the dataset description and Section 4 describes the used method in detail. The proposed methodology is present in Section 5 while Section 6 Conclusion contains the experimental results and discussion. In the end, Section 7 elaborates on the conclusion of the study.

2 Related work

Human activity recognition (HAR) is a wide-ranging field of study that aims to identify the specific action or movement of a person from sensor data. Sensor-based activity detection requires acute high-level knowledge about the activities performed by humans from massive low-level sensor readings. Data can be recorded directly through smartphones, personal tracking devices, or custom hardware that has gyroscopes or accelerometers. The main objective of HAR frameworks is to observe and analyze human behaviors efficiently, to understand the human behaviors HAR systems retrieve and process the relative data. Activity recognition involves behavior and environment monitoring, activity modeling, data processing, and pattern recognition [13]. There are numerous applications, under the concept of human activity recognition such as healthcare monitoring, tele-immersion (TI), active and assisted living systems for smart homes, and monitoring and surveillance systems for indoor and outdoor activities [18].

Activity recognition is the source for the development of many promising applications in the field of sports, wellness, health, military, lifelogging, or elderly assisted living through health monitoring, discovering patterns, detecting activity, and improving wellbeing [16]. Activity data can be collected with multiple sensors that are further analyzed to identify the complex and simple activities performed like walking, sitting, running, and other activities of daily living [6]. These activities happen to be very important to provide real-time data [14]. Various studies have been previously performed for human activity recognition by using a machine learning approach. A study [36] proposed a systematic feature learning approach for human activity recognition, to automate feature learning in a systematic way from raw data a deep convolutional neural network (CNN) is used. Through its deep architecture, the learned features from low-level time series data are considered the high-level intellectual representation. The learned features are bestowed with more discriminative power by which feature learning and classification are improved respectively. To learn without using prior information and further classifying human activities, a study [29] introduced a deep learning model which is Long Short-Term Memory (LSTM). An annotated dataset that consists of the data collected from three houses was used to apply LSTM. The experimental results of the proposed approach show that it outperformed the existing methods in terms of accuracy. As Human Activity Recognition (HAR) is a research field that mainly focuses on certainly detecting and assessing a particular user activity based on related sensor data for this purpose a novel approach is proposed in [15] consists of architecture based on the use of Deep Recurrent Neural Networks (DRNN). The novel approach is adept to detect the existence of anomalous segments during the execution of the main activity in Human Activity Recognition (HAR) and can also detect areas of a certain secondary activity within the main activity based on three scenarios including intra-dataset with intra-subject validation, intra-dataset with inter-subject validation, and inter-dataset with inter-subject validation.

A research work [26] is motivated by the success of joint learning, to solve multiple self-supervised tasks. Therefore, to identify the transformations applied to the raw input signal the study proposed a temporal convolutional neural network (CNN) and learns features by training. A set of signal transformations are used on each input signal of data and further fed to the CNN with the original data to distinguish between them. It is crucial to use transformations that exploit versatile invariances of the temporal data, for the end task highly-specified features are extracted. Research [33] compared machine learning algorithms such as support vector machine, and some deep learning algorithms including convolutional neural network, multilayer perceptron, and long short-term memory with bidirectional long short-term memory to identify the significant approach for human activity recognition. The experiment has been taken place on two datasets including UCI HAR and Pamap2 where CNN outperformed all the other models with an accuracy score of 92.7%. Similarly, CNN, LSTM, and other neural networks are also effective for other applications of machine learning such as studies [1, 2] for citywide crowd flow prediction in fog computing.

A recognition method and interval-based detection are proposed [37] which can generate candidate intervals to determine the duration of each targeted motion. Therefore, A classification-based periodic matching method is proposed for weak periodic activities. For the end task, various models are utilized for classification purposes such as convolutional neural networks, support vector machine, naive Bayes, random forest, and k-nearest neighbor. Where the accuracy of CNN and SVM has outperformed all the other classifiers with the proposed methodology of periodic matching. An architecture is proposed [9] for unsupervised local feature extraction with shallow CNN and global characteristics of the time series that encoded statistical features. For online human activity classification, they presented a user-dependent deep learning-based approach. present a user-independent deep learning-based approach for online human activity classification. Moreover, the impact of time series length on the accuracy has been investigated and limited up to a specific time interval that makes possible continuous real-time activity classification.

In a paper [5], a systematic study is performed about the data acquisition details and on-body sensor positioning for Human Activity Recognition (HAR) systems. A testbed system is built that consists of eight inertial measurement units (IMU) sensors and a smartphone device for data collection of human activity. A Long Short-Term Memory (LSTM) network is used to support training on gathered human activity data, attained in both controlled environments and the real world. To reflect the characteristic of individual activities, custom weights are elaborated for multimodal sensor fusion, by analyzing the accuracy of each sensor on different types of activity. A deep neural network for complex human activity recognition is proposed [5] named InnoHAR. The proposed model is built on the arrangement of the recurrent neural network and inception neural network. The waveform data of multi-channel sensors are fed to the model end-to-end. Where multi-dimensional features are extracted with inception modules by using several convolutional layers based on kernel. Time series features modeling is recognized to complete classification tasks by making the most use of data characteristics combined with GRU. Their proposed model contributed to the state-of-the-art result when compared with the other traditional method. In the study [21], a transfer learning-based framework is developed using convolutional neural networks that is Transfer Convolutional(TrC), aim to build a personalized model for activity recognition with the least user supervision. The personalized model is developed by retraining the upper layers of the network with a few numbers of illustrations in the target domain and reutilizing the lower layers of the network. therefore, all the training layers of the trained network are freezes except the classification layer when a new user uses the model. A major obstruction in using wearable sensors for personalized Human Activity Recognition is that during the adoption of the system by a new user or behavioral change in user status the performance of the recognition model drops significantly. Therefore, by collecting new labeled data in the new context the model needs to be retrained. Hence, two datasets are used for classification purposes where the developed model outperformed the other traditional models compared.

In research work, [20], a HAR technique based is designed to enable real-time and accurate classification for low-power wearable sensors based on a deep learning methodology. To acquire invariance against changes in sensor acquisition, sensor placement, and sensor orientation rates, a feature generation process is designed to apply to the spectral domain of the inertial data. Precisely, the sum of temporal convolutions of the transformed input is used in the proposed method. The accuracy of the proposed approach is evaluated against the other popular methods using both real-world and laboratory activity datasets. A comparison of activity recognition computation times on smartphone devices and sensor nodes and a systematic analysis of the feature generation parameters are presented. In a recent study, [25], A classification system for Human Activity Recognition is developed to obtain high accuracy by using two low-cost sensors gyroscope and accelerometer with the implementation of an Artificial Neural Network. Precisely, this study proposed a Deep Stacked Multilayer Perceptron where the MLP model is used as a base learner and an ANN model has been used as a meta learner. The proposed model DS-MLP has outperformed traditional machine learning algorithms with the highest accuracy score of 97.3% on the HAR dataset. The study [35] performs human activity recognition experiments using the ensemble model CNN-LSTM. The CNN-LSTM evaluate three datasets by the authors which are UCI, WISDOM, and OPPORTUNITY, and achieved 95%, 95%, and 92% accuracy scores for these public datasets respectively. Another study [38] works on human activity recognition using a supervised machine learning approach. In that study, the author used a hybrid method feature selection process, which includes a filter and wrapper method. The process uses a sequential floating forward search (SFFS) to extract desired features for better activity recognition. Features are then fed to a multiclass support vector machine to create nonlinear classifiers by adopting the kernel trick for training and testing purposes. The author used the UCI dataset and achieved 90% and 96% accuracy scores without and with feature selection techniques respectively. The study [31] used a machine learning approach for human activity detection as they used a support vector classifier in their approach. They also focused on six human activities and achieved a 98% accuracy score. The study [8], proposed a hybrid model named CNN-GRU for human activity detection using wearable sensors. They perform experiments on the UCI-HAR, WISDM, and PAMAP2 datasets are 96.20%, 97.21%, and 95.27% respectively. Similarly, the study [19], used adaptive CNN for human activity detection and used two datasets opportunity and W.HAR. The proposed models achieved a significant 91% and 97% F1 score for both datasets respectively.

Table 1 contains a summary of related work. We analyze that all previous studies work with complex neural network models such as CNN or LSTM which are more costly in terms of computation, So we design a simple architecture for more efficient results. Another thing which we conclude through related work is that all previously done work uses two sensor data. There is no proposed approach according to our good knowledge for human activity detection using single sensor data that's the reason this study also contributes by giving an accurate model as compared to stat-of-the-art models on single sensor data.

3 Dataset

This study is about the classification of human activity using a supervised machine learning approach. Machine learning models are used to get training on human activity datasets.

Table 1 Summary of Related work

Ref	Year	Dataset	Model
[36]	2015	Body movement and hand movement datasets	CNN
[29]	2017	Annotated sensor dataset	LSTM
[15]	2019	RealWorld (HAR) dataset	DRNN
[26]	2019	Six datasets including: HHAR, UniMiB, UCI HAR, MobiAct, WISDM, MotionSense	CNN
[33]	2020	UCI HAR and Pamap2 dataset	CNN
[37]	2019	Badminton and swimming dataset	CNN
[9]	2018	WISDM and UCI HAR	CNN
[5]	2019	Daily living activity	LSTM
[5]	2019	Three datasets including: Opportunity activity recognition, Pamap2, and Smart phone dataset	InnoHAR
[21]	2018	SDA, WISDM datasets	TrC
[20]	2019	ActiveMiles, WISDM, Skoda, and DaphnetFoG datasets	CNN
[35]	2020	UCI, WISDOM, and OPPORTUNITY	CNN-LSTM
[38]	2020	UCI	SFFS with SVM
[31]	2021	UCI HAR Dataset	SVM
[8]	2021	UCI-HAR, WISDM, and PAMAP2	CNN-GRU
[19]	2021	Opportunity and W.HAR	Adaptive CNN
Our Approach	2020	HHAR dataset	MLP

This study for experiments acquired and dataset ‘Heterogeneity Dataset for Human Activity Recognition from the UCI data repository contains human activity features http://archive.ics.uci.edu/ml/datasets/heterogeneity+activity+recognition. This dataset is collected using the smartphone and watch sensor. This dataset contains the features for the six activities which (sit, stand, walk, bike, walk upstairs, and walk downstairs). This activity dataset is collected using the gyroscope and accelerometer sensor of smartphones and the smartwatch. Each collected data CSV file contain ‘Index’, ‘Arrival_Time’, ‘Creation_Time’, ‘x’, ‘y’, ‘z’, ‘User’, ‘Model’, ‘Device’, ‘gt’ attribute. The attributes ‘x’, ‘y’, and ‘z’ contain the 3- axial value from the sensor. The data count for each activity is shown in Table 2. The size of the data is too large we extract an equal number of instance from each activity of both sensors and used 12000 instances as shown in Table 3.

The used dataset features are extracted by carrying the smart devices by the order in no specific order. The sensors of the smartphone device capture the 3-axial linear acceleration and 3- axial velocity sampled at the highest frequency the respective device allows. In this data collection, 9 subjects participate. Each person performs all six activities carrying the smart device. For the preprocessing of the signal, a noise filter is applied to the signals collected through the gyroscope and accelerometer. The characteristic of the used dataset is shown in Table 4.

The sample of the dataset is shown in Table 5 and the used data for the experiment is shown in Table 6. For the experiment, we extract the 3 axial attributes x, y, and z from both

Table 2 Instance count for each activity according to sensors

Activity	Accelerometer	Gyroscope
Sit	1991919	2218501
Stand	1851492	2024206
Walking	2192401	2350429
Bike	1845557	1911730
Stairs up	1782010	1884306
Stairs down	1615896	1673833
Total	13062475	13932632

gyroscope and accelerometer files as shown in Table 6. As we make an experimental dataset by combining the gyroscope and accelerometer features and the instance for each activity in the experimental dataset are equal so there are very fewer chances of the model over-fitting as compared to the original dataset.

4 Supervised machine learning algorithms

Supervised machine learning is an approach for classification and regression tasks. In the classification task, the target value is categorical and in the regression task, the target value continues. Supervised machine learning is where the input variable is (x) and the output variable is (y) and an algorithm is used to map the input (x) to output (y) [11].

$$y = f(x) \tag{1}$$

We used a supervised machine learning approach for human activity detection using the smartphone sensor. We used different machine learning algorithms such as decision tree, random forest, logistic regression, k nearest neighbor, and ensemble to achieve accurate result. We used these models with the optimal hyper-parameters setting to select the best performance on the used dataset. The hyper-parameters are shown in Table 7.

4.1 Decision tree

The decision tree is a rule-based model that can be used for both classification and regression tasks [4]. It generates a tree by decision rules extract from the input data. Decision trees learn from decision rules and predict the target value as the base of learning rules. Decision tree accuracy highly depends on the splitting rules of nodes in the tree. The decision tree

Table 3 Instance count for each activity in dataset sued for experiments

Activity	Accelerometer	Gyroscope
Sit	2000	2000
Stand	2000	2000
Walking	2000	2000
Bike	2000	2000
Stairs up	2000	2000
Stairs down	2000	2000
Total	12000	12000

Table 4 Characteristic of used dataset

Attribute	Count	Description
smartwatch	4	2 Samsung Galaxy Gears, 2 LG watches
Smart Phones	8	2 Samsung Galaxy S3 mini, 2 Samsung Galaxy S+, 2 LG Nexus 4, 2 Samsung Galaxy S3
User	9	Perform 6 activities
Sensors	2	Collect the activity data by subject performance
CSV file attributes	10	Contain the data-related information

model uses multiple algorithms to split a node into multiple sub-node. The purity of the node can increase the accuracy of the decision tree. In this study, for experiments, we used 'Entropy' algorithms to find the purity of nudes. We used the 'max_depth' hyper-parameter with the decision tree to reduce the complexity of the decision tree. This hyper-parameter is used with the value of 300 which will restrict a decision tree to max 300 level depths. This value for max depth we find using the hit and trial method. This is the best parameter setting for the decision tree to get higher results on the used dataset.

4.2 Random forest

Random forest is an ensemble model that can be used for both classification and regression tasks [3]. It is a combination of multiple decision trees under the majority voting criteria. It generates the number of decision trees to learn from the input data and then combines the output of these generated trees to make a final prediction. Its performance is good as compared to the individual decision tree because multiple algorithms in combination can perform better as compared to an individual model. we can define a random forest as:

$$Prediction = mode\{dt_1, dt_2, dt_3, \ldots, dt_n\} \tag{2}$$

$$Prediction = mode\{\sum_{i=0}^{N} dt_i\} \tag{3}$$

In (3) N is the number of decision trees used to obtain a final prediction using the majority voting criteria. The 'dt' is the decision tree in a random forest.

We used a random forest with the number of hyper-parameters as shown in Table 7, the n_estimator parameters define the number of decision trees will generate in the random forest and we used this hyper-parameter with a 300 value which means that the random

Table 5 Sample dataset from gyroscope file

Index	Arrival_Time	Creation_Time	x	y	z	User	Model	Device	gt
0	1.42E+12	1.42E+18	0.013748	-0.00063	-0.02338	a	nexus4	nexus4_1	stand
1	1.42E+12	1.42E+18	0.014816	-0.00169	-0.02231	a	nexus4	nexus4_1	stand
2	1.42E+12	1.42E+18	0.015884	-0.00169	-0.02124	a	nexus4	nexus4_1	stand
3	1.42E+12	1.42E+18	0.016953	-0.00383	-0.02017	a	nexus4	nexus4_1	stand

Table 6 Sample dataset used in experiments

Gyroscope			Accelerometer			Target
X	Y	Z	X	Y	Z	
0.013748	-0.00063	-0.02338	-5.95819	0.688065	8.135345	stand
0.014816	-0.00169	-0.02231	-5.95224	0.670212	8.136536	stand
0.015884	-0.00169	-0.02124	-5.99509	0.653549	8.204376	stand
0.016953	-0.00383	-0.02017	-5.94272	0.676163	8.128204	stand

forest will generate 300 decision trees and these decision trees will make their prediction and then the random forest will combine all prediction against single instance and will make a final prediction on the base of majority voting. For the next hyper-parameter, we use max_depth with the same value as in the decision tree. This hyper-parameter will restrict each decision tree to a maximum of 300-level depths. This hyper-parameter will reduce complexity in a random forest.

4.3 Multinomial logistic regression

Logistic regression is a linear model used for classification and can be used for binary and multi-class classification [10]. Its name logistic regression is because of the mathematical logistic function. This function is also known as the sigmoid function. It is an s-shaped function that can map all real values between 0 and 1. Logistic regression finds the relationship between the dependent and independent variables to get a good fit. This is about the classification of six activities in the used dataset that's the reason we used the multinomial logistic regression. Logistic regression can be defined as:

$$\frac{1}{(1+e^{-value})} \tag{4}$$

$$v = \frac{e^{b0+b1*x}}{(1+e^{b0+b1*x})} \tag{5}$$

In (5) v is the prediction by the model and b0 is the intercept by the model and b1 is the coefficient for the single input value x. We used multinomial logistic regression with the 3 hyper-parameters as shown in Table 7. Solver ‘liblinear’ which is used for the fast computation and the second parameter multi_class with value ‘multinomial’ which is used for multi-class classification and third is C which is Inverse of regularization strength; must be a positive float. We use it with a value of 3.0.

Table 7 Hyper-parameters Setting

Model	Hyper-parameters
Random forest	n_estimators = 300, max_depth = 10
Decision tree	max_depth = 10
Logistic regression	solver = ‘saga’, multi_class = ‘multinomial’, C = 3.0
K nearest neighbor	Neighbour = 5
MLP	Parameters and values are in MLP section

4.4 K nearest neighbor

K nearest neighbor is a simple supervised machine learning algorithm to solve the classification and regression algorithms [30]. This model is also known as the lazy learner This model is significant when the dataset size is small that's the reason we used the k nearest neighbor model in this study because the used dataset has a very small feature set. We used the single parameter with the model which is a neighbor.

4.5 Artificial neural networks

Artificial neural networks (ANN) are used to perform the classification, clustering, and pattern recognition task [7, 34]. ANN is a complex system containing many neurons. The use of the neural network to perform the classification task is increased in recent years. This study is also used the ANN to perform the activity recognition task because the use of ANN in the human activity recognition domain is increased in recent years. ANN consists of many layers that contain the number of neurons to perform computation on input data. Each neuron is interconnected and takes input for computation and processes the input and sends information. Lots of models are available based on ANN and we design a multilayer perceptron (MLP) to solve the human activity recognition problem. ANN with one or more hidden layers is known as MLP. MLP model consists of three parts: an input layer, hidden layers, and an output layer. ANN is a fully connected model which means that each neuron is interconnected with the other.

In recent years, lots of work is done using the MLP model such as the study [5]. They apply the model in the five activity dataset collected wearing the device in the trouser front pocket and trouser back pocket. Their method achieved 92% accuracy on the used dataset. The study [21] used the ANN model for the activity recognition using the UCI HAR dataset and performs a comparison between ANN and the state-of-the-art model and achieved 94% accuracy. Another study [20] used the accelerometer and gyroscope 3- axial sensor to collect the dataset and ANN model to achieve the highest accuracy. Similarly, this study also used the MLP model to fit the activity dataset (Table 8).

4.5.1 The architecture of used MLP

In this study, we designed an MLP model with one input layer, two hidden layers, one output layer, two dropout layers, and two activation functions ('ReLU' and 'sigmoid'). The proposed model outperforms all other used models because of its architecture. As the dataset contains a small feature set so in the proposed model we make a stack of dense layers and dropout layers with the number of neurons. The simple architecture of the proposed model used 64 neurons at each dense layer with ReLUs activation function which makes it less complicated and better on non-linear data. The input layer also contains the 'ReLU' (Rectified linear unit) which overcomes the vanishing gradient problem, allowing models to learn faster and perform better [27]. The 'ReLU' function can be defined mathematically shown in (6):

$$y = \max(0, x) \tag{6}$$

This input layer follows the dropout layer with a 0.2 dropout rate. The dropout rate is helpful to reduce the over-fitting of the model, it drops the randomly selected neurons to reduce the over-fitting of the model. Hidden layers contain the 64 neurons and 'ReLU' activation function. Each hidden layer is also followed by the dropout layer with a 0.2 dropout rate. In

Table 8 Proposed MLP layers and parameters values

Layers	Parameters & Values
Input/Dense Layer	Neurons=64
	Input Dim. =3 (Gyroscope Features)
	Input Dim. =3 (Accelerometer Features)
	Input Dim. =6 (Combine Features)
	Activation Function= ReLUs
Dropout Layer	Dropout Rate=0.2
Dense Layer	Neurons=64
	Activation Function= ReLUs
Dropout Layer	Dropout Rate=0.2
Dense Layer	Neurons=64
	Activation Function= ReLUs
Dropout Layer	Dropout Rate=0.2
Output/Dense Layer	Neurons=6
	Activation Function= SoftMax
Compile	Loss= categorical_crossentropy
	Optimizer=Adam
	Epochs= 500

the end out we used the output layer with the softmax activation function which is used to predict the multinomial probability distribution. Softmax can define as shown in (7):

$$\sigma(z)_i = \frac{e^{z_i}}{\left(\sum_{j=1}^{k} e^{z_j}\right)} \tag{7}$$

We compile the model with the ‘adam’ optimizer and the ‘categorical_crossentropy’ loss function. The ‘adam’ optimizer is an optimization algorithm that can be used instead of the classical stochastic gradient descent procedure to update network weights iterative based on training data. The ‘categorical_crossentropy’ loss was used because of the multi-class classification problem and we set epochs value to 500. The architecture of MLP is shown in Fig. 1.

5 Experimental flow

This study is about the recognition of human activity using a supervised machine learning approach. For this, we acquired a dataset from the UCI dataset repository containing six activity features (see section). The dataset contains both smartphone and smartwatch data files. In this study, we used only a smartphone dataset with both gyroscope and accelerometer sensors. We used 3- axial data from gyroscope and accelerometer sensors and combine both sensor data values to train learning models. Figure 2 show how we combine the feature sets from both sensors.

After combining both sensor’s features we extract the equal number of the instance from each activity from both gyroscope and accelerometer files as shown in Table 9. After obtaining the experimental dataset split the dataset into training and testing sets. We split the data

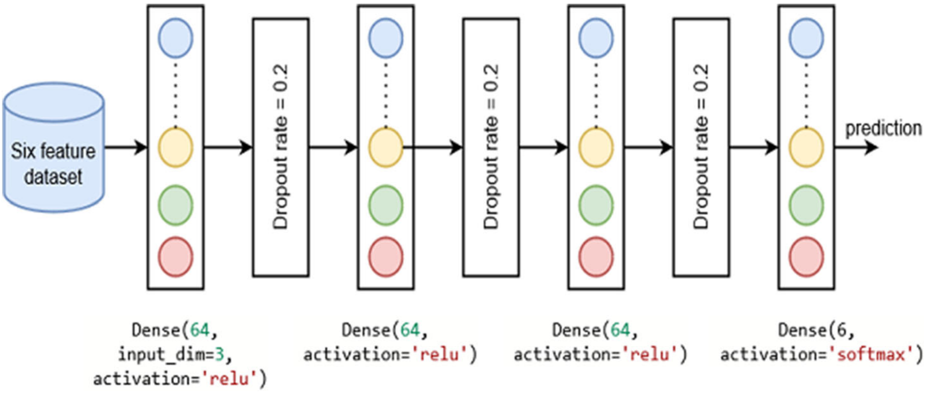


Fig. 1 Architecture diagram for design MLP

into training and testing sets with a ratio of 70:30. The 70% of the dataset we used for the training of models and then we used the 30% test data to evaluate the performance of the model. The training and testing data size is shown in Table 9.

We evaluate the performance of models in terms of accuracy precision, recall, and f1 score. The experimental flow diagram is shown in Fig. 3. These experiments are performed on the Cori7, 7th generation machine, Jupytor Notebook with python 3.6, and Google Colab.

5.1 Evaluation parameters

In this study about the multi-class classification using the machine learning algorithms so to evaluate the performance of the machine learning model, we have used several evaluation parameters such as accuracy, precision, recall f1 score, and Cohn's kappa [24].

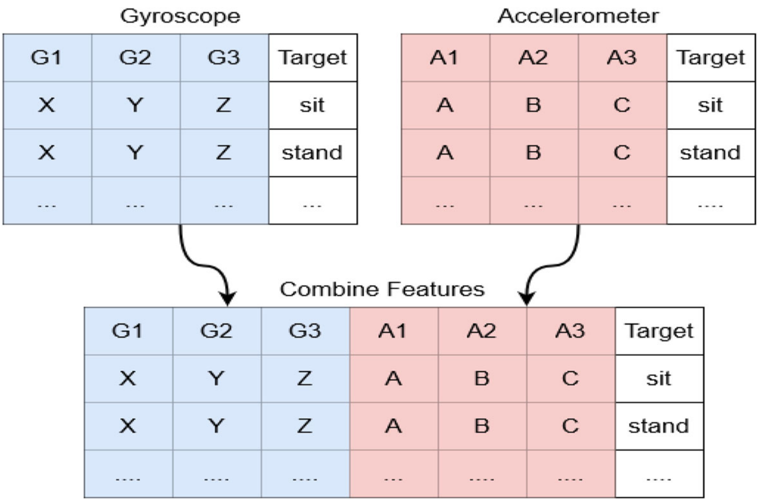


Fig. 2 Visualization for combining of features from both sensors

Table 9 Training and testing sets count

Activity	Total	Training Set	Testing Set
Sit	2000	1383	617
Stand	2000	1384	616
Walking	2000	1425	588
Bike	2000	1392	608
Stairs up	2000	1425	575
Stairs down	2000	1404	596
Total	12000	8400	3600

Accuracy parameters give a score about model correct perdition. The accuracy score is the number of correct predictions divided by the number of total predictions. The maximum accuracy score for a model is 1 and the minimum accuracy score is 0.

$$Accuracy = (Number\ of\ correct\ predictions)/(Total\ predictions) \tag{8}$$

It can also be calculated as:

$$Accuracy = (TP + TN)/(TN + TP + FN + FP) \tag{9}$$

In (9):

- TP is the True positive when the model label is yes and the actual label is also yes.
- TN is True Negative when the model predicts the label as No and the actual label was also no.
- FP is False positive when the model predicts the label as no but the actual label was yes.
- FN is False Negative when modeling the label as yes but the actual label was no.

Similarly, precision can be found as

$$Precision = TN/(TN + FP) \tag{10}$$

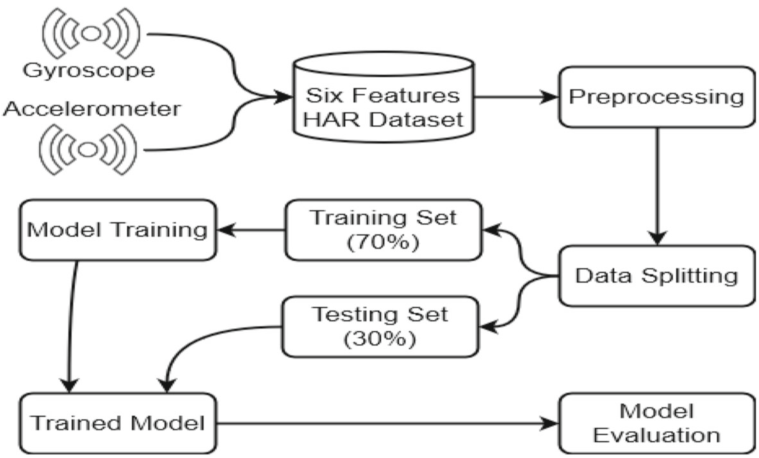


Fig. 3 Experimental Flow Diagram

Precision maximum score can be and minimum 0. Recall can be found as:

$$Recall = TN/(TN + FN) \tag{11}$$

Recall the maximum score can be and the minimum 0. F1 score is also known as F measure score and it's a harmonic mean between precision and recall. It's also the maximum score can be and the minimum can be 0.

$$F1score = 2 * (Precision * Recall)/(Precision + Recall) \tag{12}$$

6 Results

This section contains the experimental results of the classification of human activities using the smartphone data and machine learning approach. The used dataset is collected for the sensor as described in the dataset section. The performance of the machine learning model highly depends on the importance of the features. The sensor which will generate more correlated features corresponding to the target class will be more helpful to train a good fit learning model. We train over design MLP and other states of the art model first on the accelerometer and gyroscope data and then we combine both sensor features to make a prediction.

6.1 Results on gyroscope features

The dataset contains a gyroscope and accelerometer 3- axial features which can be useful to train machine learning models. First, we train machine learning models on the gyroscope features and perform a comparison between their performances. MLP model outperforms all other models in terms of all evaluation parameters and achieved 0.74 accuracies and the highest precision of 0.75. The performance of all models is not prominent but overall the MLP is the top performer. LR performs worst in this scenario because training data have only 3 features which are not enough for a good fit for LR, that's the reason LR is the worst performer. RF, DT, and KNN also perform well because they didn't require more features for training. The results of all models are shown in Table 10 while training and testing accuracy, and loss are shown in Table 11. Figure 4a & b show that the training and testing accuracies are too parallel which shows the good fit of MLP.

The confusion matrix in Fig. 5 shows the performance of MLP in terms of an accurate and wrong prediction. According to the confusion matrix, the MLP model makes 2679 correct predictions and 921 wrong predictions out of 3600 predictions. MLP gives more accurate predictions on stand activity which are 543 out of 616.

Table 10 Model performance on gyroscope features

Model	Accuracy	Precision	Recall	F1 score	Cohen's Kappa
MLP	0.74	0.75	0.74	0.74	0.70
KNN	0.71	0.71	0.70	0.70	0.65
RF	0.73	0.73	0.73	0.73	0.67
DT	0.72	0.72	0.72	0.72	0.67
LR	0.37	0.36	0.37	0.36	0.24

Table 11 Accuracy and loss for the MLP model on gyroscope features

	Accuracy	Loss
Training	0.821	0.444
Testing	0.744	0.588

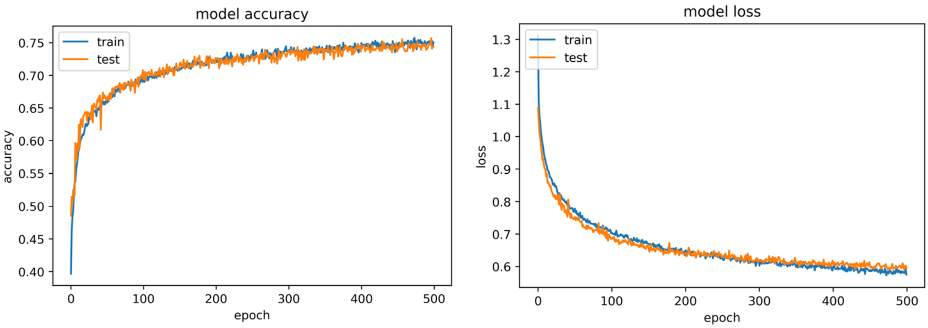
6.2 Results on accelerometer features

We experiment on accelerometer 3- axial features using the machine learning approach. The performance of the machine learning model on accelerometer features is good compared to the performance on gyroscope features. The MLP model outperforms all other models in terms of all evaluation parameters. The accuracy of MLP is 0.77 which is the highest on accelerometer features. The results of all models used in this study are shown in Table 12. Table 13 shows the training and testing accuracy and loss. The training and testing accuracy are in parallel as shown in Fig. 6a, b, which show the good fitting of the model.

The confusion matrix in Fig. 7 shows that the model makes a more correct prediction on the accelerometer features as compared to gyroscope features. MLP gives the highest correct prediction ratio as compared to other models. It gives 2766 correct predictions out of 3600 and 834 wrong predictions. This correct prediction ratio is high on accelerometer features as compared to gyroscope features.

6.2.1 Results on both gyroscope and accelerometer features

In this section, we will disuse the results of machine learning models on the combination of gyroscope and accelerometer datasets. The all used model increase their performances on the large features dataset as compared to individual sensor data. MLP our design model also outperform here and achieved a 0.95 accuracy score. The results of models show that the combination of gyroscope and accelerometer features make more correlation with the target class as compared to individually and a combination of features also makes a large feature set with is good model training. All models improve their performance and MLP is the top performer in Table 14. The LR is again the worst performer with 0.56 accuracy because of the small feature set which is not enough for the LR. Table 14 shows the results of all used



(a) (b)

Fig. 4 MLP training and testing scores for each using gyroscope features (a) Accuracy, (b) Loss

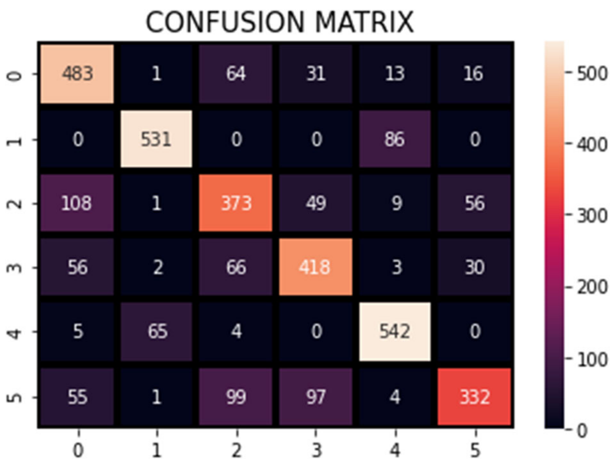


Fig. 5 Confusion matrix for the best performer MLP on gyroscope features

Table 12 Model performance on Acceleoromter features

Model	Accuracy	Precision	Recall	F1 score	Cohen's Kappa
MLP	0.77	0.76	0.76	0.76	0.72
KNN	0.75	0.74	0.75	0.75	0.71
RF	0.75	0.75	0.75	0.75	0.70
DT	0.73	0.73	0.73	0.73	0.67
LR	0.48	0.44	0.48	0.45	0.37

Table 13 Accuracy and loss for the MLP model on accelerometer features

	Accuracy	Loss
Training	0.791	0.501
Testing	0.770	0.551

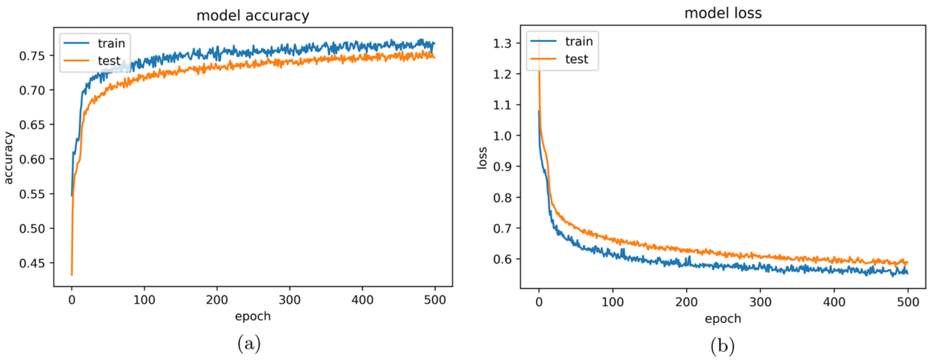


Fig. 6 MLP training and testing scores for each using accelerometer features (a) Accuracy, (b) Loss

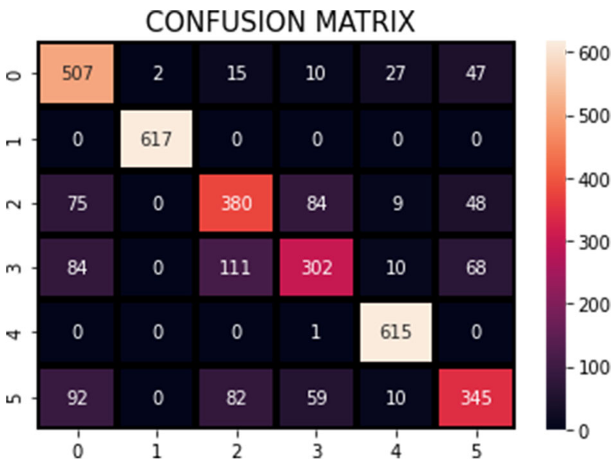


Fig. 7 Confusion matrix for the best performer MLP using accelerometer features

Table 14 Model performance on gyroscope and accelerometer combine features

Model	Accuracy	Precision	Recall	F1 score	Cohen's Kappa
MLP	0.98	0.98	0.98	0.98	0.97
KNN	0.94	0.93	0.93	0.93	0.92
RF	0.93	0.93	0.93	0.93	0.91
DT	0.91	0.91	0.91	0.91	0.89
LR	0.56	0.53	0.56	0.54	0.46

Table 15 Accuracy and loss for the MLP model on gyroscope and accelerometer combine features

	Accuracy	Loss
Training	0.997	0.014
Testing	0.982	0.056

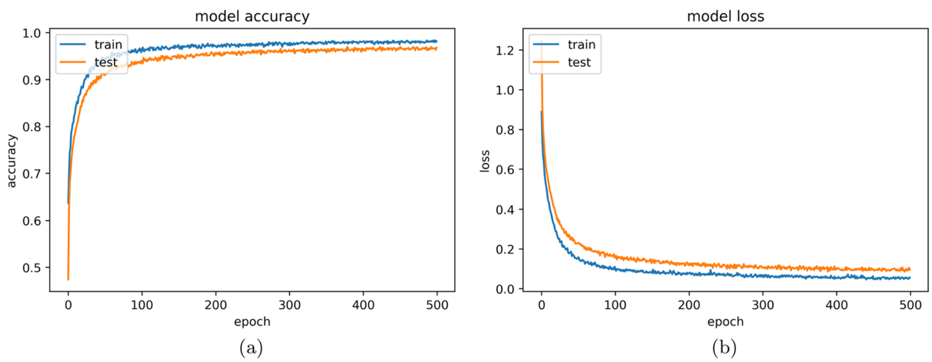


Fig. 8 MLP training and testing scores for each epoch using gyroscope and accelerometer combine features (a) Accuracy, (b) Loss

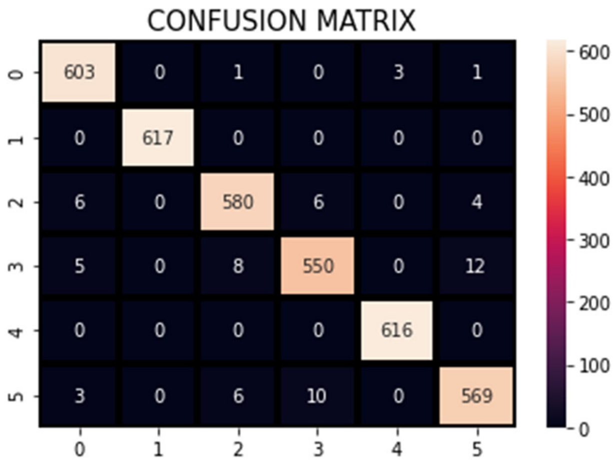


Fig. 9 Confusion matrix for the best performer MLP on Gyroscope and accelerometer combine features

models on the combined features and Table 15 shows the accuracy and loss for the MLP model. The training and testing accuracy are in parallel as shown in Fig. 8a and b, which show the good fitting of the model.

MLP confusion matrix in Fig. 9 shows the correct and wrong prediction for each activity using the combined features of the gyroscope and accelerometer. Overall the MLP gives 3535 correct predictions out of 3600 and 65 wrong predictions which very the lowest wrong predictions as compared to previous results. For sit activity, the MLP model gives 100% results by giving all 617 correct predictions. Figure 10 shows the comparison between the model's performance on Gyroscope, accelerometer, and combined features.

We also did a comparison of the proposed approach with other state-of-the-art deep learning models such as CNN and LSTM. The models are deployed with the state-of-the-art architecture used by the studies [23, 28]. Both CNN and LSTM are not good on the used

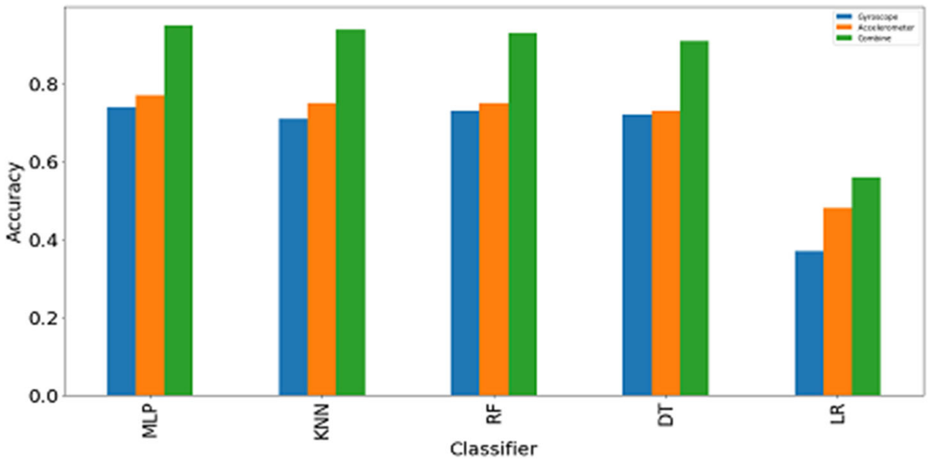


Fig. 10 Model performance comparison on Gyroscope and accelerometer combine features

Table 16 Deep learning models results

Model	Accuracy	Precision	Recall	F1 Score
CNN	0.66	0.64	0.65	0.64
LSTM	0.71	0.70	0.70	0.69

dataset because of the small feature set. Especially CNN, required a large feature set for a good fit as compared to machine learning models. Table 16 shows the results of CNN and LSTM and Fig. 11 show the accuracy and loss for both.

6.3 T-test to show statistical significance of proposed MLP

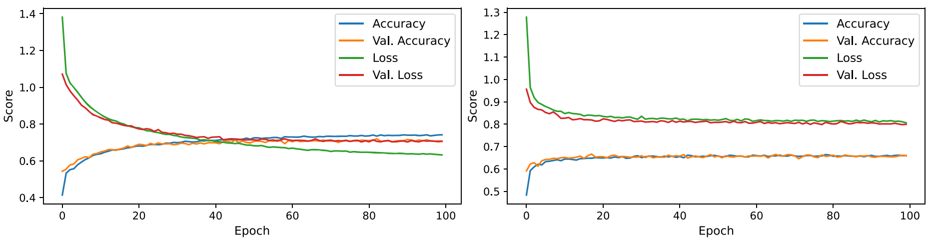
This study performs a statistical T-test to show the significance of the proposed MLP model as compared to all other used models [17]. The T-test is mostly used to show the statistical significance of one approach on other approaches. This study also used a T-test to check the statistical significance of the proposed model MLP on other models. To support the T-test we consider two hypotheses:

- Null Hypothesis (Ho): Proposed MLP is statistically significant as compared to other used models.
- Alternative Hypothesis (Ha): Proposed MLP is not statistically significant as compared to other used models.

Our T-test accepts Ho which means that the proposed model is statistically significant when we used gyroscope features, Accelerometer features, and a combination of both feature sets for model training. Statistical T-test shows that the proposed MLP is statistically significant as compared to all other used models on the human activity dataset.

7 Conclusion and future work

In this study, we have experimented with the human activity dataset to perform the classification task. We design the MLP model for this and perform its comparison with the state-of-the-art algorithms. To train the MLP model we split the dataset into training and testing sets with a ratio of 70:30. We perform the comparison on gyroscope and accelerometer features individually and also by combining them. The deigned model outperforms all



(a) (b)

Fig. 11 Accuracy and loss graph (a) LSTM, (b) CNN other models with the highest accuracy of 0.98 with combination gyroscope and accelerometer features. We conclude that gyroscope features are not too much correlated with the target class as compared to accelerometer features that are the reason gyroscope features are not so prominent individually but the combination of accelerometer and gyroscope features can increase the accuracy of the machine learning model. The combination of accelerometer and gyroscope increases the size of the feature set which is helpful in a good fit of the model. We also conclude that LR performance is depending on the feature set size. Small feature sets are not good for LR training if we will increase the size of features the performance of LR will be increased. Another thing we concluded is that neural networks are stronger than traditional machine learning algorithms because neural networks learn the pattern more deeply. In future work, we will generate a real-time dataset to test the trained model and will evaluate our designed model on multiple datasets. Especially, more activities will be under consideration in future work to evaluate the proposed model.

Declarations

Conflict of Interests The authors declare that they have no conflict of interest.

References

1. Ali A, Zhu Y, Zakarya M (2021) A data aggregation based approach to exploit dynamic spatio-temporal correlations for citywide crowd flows prediction in fog computing. Multimed Tools Appl 80(20):31401–31433
2. Ali A, Zhu Y, Zakarya M (2021) Exploiting dynamic spatio-temporal correlations for citywide traffic flow prediction using attention based neural networks. Inf Sci 577:852–870
3. Biau G, Scornet E (2016) A random forest guided tour. Test 25(2):197–227
4. Brijain M, Patel R, Kushik M, Rana K (2014) A survey on decision tree algorithm for classification
5. Chung S, Lim J, Noh KJ, Kim G, Jeong H (2019) Sensor data acquisition and multimodal sensor fusion for human activity recognition using deep learning. Sensors 19(7):1716
6. Cornacchia M, Ozcan K, Zheng Y, Velipasalar S (2016) A survey on activity detection and classification using wearable sensors. IEEE Sensors J 17(2):386–403
7. Dongare AD, Kharde RR, Kachare AD (2012) Introduction to artificial neural network. Int J Eng Innov Technol (IJEIT) 2(1):189–194
8. Dua N, Singh SN, Semwal VB (2021) Multi-input CNN-GRU based human activity recognition using wearable sensors. Computing 103(7):1461–1478
9. Ignatov A (2018) Real-time human activity recognition from accelerometer data using convolutional neural networks. Appl Soft Comput 62:915–922
10. Kleinbaum DG, Dietz K, Gail M, Klein M, Klein M (2002) Logistic Regression. Springer, New York
11. Kotsiantis SB, Zaharakis I, Pintelas P (2007) Supervised machine learning: a review of classification techniques. Emerg Artif Intell Appl Comput Eng 160(1):3–24
12. Lara OD, Labrador MA (2012) A survey on human activity recognition using wearable sensors. IEEE Commun Surv Tutor 15(3):1192–1209
13. Lockhart JW, Pulickal T, Weiss GM (2012) Applications of mobile activity recognition. In: Proceedings of the 2012 ACM conference on ubiquitous computing, pp 1054–1058
14. Mendes S, Queiroz J, Leitão P (2017) Data driven multi-agent m-health system to characterize the daily activities of elderly people. In: 2017 12th Iberian Conference on Information Systems and Technologies (CISTI). IEEE, (pp 1–6)
15. Munoz-Organero M (2019) Outlier detection in wearable sensor data for human activity recognition (HAR) based on DRNNs. IEEE Access 7:74422–74436
16. Nweke HF, Teh YW, Mujtaba G, Al-Garadi MA (2019) Data fusion and multiple classifier systems for human activity detection and health monitoring: Review and open research directions. Inf Fusion 46:147–170

17. Omar B, Rustam F, Mehmood A, Choi GS (2021) Minimizing the overlapping degree to improve class-imbalanced learning under sparse feature selection: application to fraud detection. IEEE Access 9:28101–28110
18. Ranasinghe S, Al Machot F, Mayr HC (2016) A review on applications of activity recognition systems with regard to performance and evaluation. Int J Distrib Sens Netw 12(8):1550147716665520
19. Rashid 3- N, Utku Demirel B, Abdullah Al Faruque M (2021) AHAR: adaptive CNN for energy-efficient human activity recognition in lowpower edge devices, arXiv:2102.01875
20. Ravi D, Wong C, Lo B, Yang GZ (2016) Deep learning for human activity recognition: A resource efficient implementation on low-power devices. In: 2016 IEEE 13th international conference on wearable and implantable body sensor networks (BSN). IEEE, (pp 71–76)
21. Rokni SA, Nourollahi M, Ghasemzadeh H (2018) Personalized human activity recognition using convolutional neural networks. arXiv:1801.08252
22. Ronao CA, Cho SB (2016) Human activity recognition with smartphone sensors using deep learning neural networks. Expert Syst Appl 59:235–244
23. Rustam F, Imtiaz Z, Mehmood A, Rupapara V, Choi GS, Din S, Ashraf I (2022) Automated disease diagnosis and precaution recommender system using supervised machine learning. Multimed Tools Appl, pp 1–24
24. Rustam F, Mehmood A, Ahmad M, Ullah S, Khan DM, Choi GS (2020) Classification of shopify app user reviews using novel multi text features. IEEE Access 8:30234–30244
25. Rustam F et al (2020) Sensor-based Human Activity Recognition using Deep Stacked Multilayered Perceptron Model. In: IEEE Access, https://doi.org/10.1109/ACCESS.2020.3041822
26. Saeed A, Ozcelebi T, Lukkien J (2019) Multi-task self-supervised learning for human activity detection. Proc ACM Interact Mob Wearable Ubiquitous Technol 3(2):1–30
27. Schmidt-Hieber J (2020) Nonparametric regression using deep neural networks with reLU activation function. Ann Stat 48(4):1875–1897
28. Siddiqui HUR, Shahzad HF, Saleem AA, Khan Khakwani AB, Rustam F, Lee E, Ashraf I, Dudley S (2021) Respiration based Non-Invasive approach for emotion recognition using impulse radio ultra wide band radar and machine learning. Sensors 21(24):8336
29. Singh D, Merdivan E, Psychoula I, Kropf J, Hanke S, Geist M, Holzinger A (2017) Human activity recognition using recurrent neural networks. In: International cross-domain conference for machine learning and knowledge extraction). Springer, Cham, pp 267–274
30. Soucy P, Mineau GW (2001) A simple KNN algorithm for text categorization. In: Proceedings 2001 IEEE international conference on data mining. IEEE, (pp 647–648)
31. Swarnakar SK, Agrawal H, Goel A (2021) Smartphone inertial sensors-based human activity detection using support vector machine. In: Soft computing: theories and applications. Springer, Singapore, (pp 231–241)
32. Vrigkas M, Nikou C, Kakadiaris IA (2015) A review of human activity recognition methods. Frontiesr Robotics AI 2:28
33. Wan S, Qi L, Xu X, Tong C, Gu Z (2020) Deep learning models for real-time human activity recognition with smartphones. Mob Netw Appl 25(2):743–755
34. Webber WRS, Lesser RP, Richardson RT, Wilson K (1996) An approach to seizure detection using an artificial neural network (ANN). Electroencephalogr Clin Neurophysiol 98(4):250–272
35. Xia K, Huang J, Wang H (2020) LSTM-CNN architecture for human activity recognitions. In: IEEE Access, vol 8, pp 56855–56866. https://doi.org/10.1109/ACCESS.2020.2982225
36. Yang J, Nguyen MN, San PP, Li X, Krishnaswamy S (2015) Deep convolutional neural networks on multichannel time series for human activity recognition. In: Ijcai, vol 15, pp 3995–4001
37. Zhuang Z, Xue Y (2019) Sport-Related Human activity detection and recognition using a smartwatch. Sensors 19(22):5001
38. hmed N, Rafiq JI, Islam MR (2020) Enhanced human activity recognition based on smartphone sensor data using hybrid feature selection model. Sensors 20:317. https://doi.org/10.3390/s20010317