Performance evaluation of machine learning based voting classifier system for human activity recognition

Sonika Jindal [1,*] Monika Sachdeva [2], Alok K. S. Kushwaha [3]
[1,2,3] *Dept. of Computer Science and Engineering*
[1,2] *I. K. Gujral Punjab Technical University, Jalandhar, India*
[3] *Guru Ghasidas Vishwavidyalaya , Bilaspur, India*
**Corresponding author: sonikajindal@sbsstc.ac.in*

Abstract

In the last few decades, Human Activity Recognition (HAR) has been a centre of attraction in many research domains, and it is referred to as the potential of interpreting human body gestures through sensors and ascertaining the activity of a human being. The present work has proposed the voting classifier system for human activity recognition. For the voting classifier system, five machine learning classifiers are considered: Logistic Regression (LR), K-Nearest Neighbour (KNN), Random Forest (RF), Naive Bayes (NB), and Support Vector Machine (SVM). These machine learning classifiers are ensembled by analyzing the best performers among them. The ensemble voting classifiers are proposed under two variations, i.e., hard voting and soft voting. The various combinations of voting classifiers are compared and evaluated. For experiments, the benchmark dataset of the UCI-HAR dataset is considered, and all the data files are combined into a single file to avoid bias. The dimensionality of the dataset is reduced by using Principal Component Analysis (PCA) from 561 features to 200 components. The results reveal that Voting Classifier-II (a combination of SVM, KNN, and LR) using soft voting outperformed other machine learning classifiers.

Keywords: Hard voting; human activity recognition; machine learning; pattern recognition; soft voting; voting classifier

1. Introduction

HAR is a process of interpreting human body gestures or movements through sensors that are useful for numerous research domains, including human machine communication, robotics, etc (Yayan *et al.*, 2021). However, recognizing human activities is a very hard job to do because of unresolved challenges such as sensor mobility, sensor deployment, disordered background, and intrinsic unpredictability in the sense of how various human activities are conducted. A HAR system has the ability to simplify or automate many of the routine tasks of humans by recognizing them. The methods employed for HAR can be divided into two types: invasive and non-invasive. The invasive methods make use of wearing sensors to track subjects with the aim of developing a large dataset to learn from models (Sjarif & Shamsuddin, 2015). The non-invasive HAR, on the other hand, reduces the need for any wearable gadgets to monitor human activities. Such systems use Wi-Fi signals that are openly accessible in almost all premises.

HAR involves crucial activities in the classification task, such as sitting, falling, and human non-appearance, etc. All these movements are related to the smart home application, whereas the falling activity is specifically related to the health assistance of the elderly, where it is not possible to install cameras in separate rooms but a requirement to monitor the patients. The fact that the human body has an impact on the signal due to its reflection and several activities lead to the display of diverse attributes motivates the concept of Human Activity Recognition in which Wi-Fi signals are utilized. Furthermore, the concept of HAR using sensors is preferred compared to still pictures because the activity recognition with video frames or still images faces issues such as backdrop clutter, partial occlusion, variations in scale, perspective, lighting, and look. Sensor networks, contact opportunities, and robotic systems for human nature interpretation, to name a few applications, all require a multimodal activity recognition system. In this paper, HAR presents a comprehensive evaluation of current and cutting-edge research developments in the pasture of HAR. A lot of work has been done recently on the HAR model that needs feature engineering and the classic methods. For example, SVMs or decision trees are used to perform the classification portion of the examination (Ijjina & Mohan, 2014). The major contributions in the field are discussed in the next section.

The major contribution of the work is the usage of voting classifier systems for human activity recognition. The voting classifier systems ensemble the machine learning classifiers and evaluate the result outcome based on the combined efforts of the incorporated machine learning classifiers. Here, hard voting and soft voting based classifications are conducted by incorporating the machine learning (Khan *et al.*, 2016) classifiers of Logistic Regression (LR), K-Nearest Neighbour (KNN), Random Forest (RF), Naive Bayes (NB), and Support Vector Machine (SVM). Hard voting determines the output on the basis of the higher number of votes for the classifier, and soft voting evaluates the predictive outcome by considering the mean of the probabilities allotted to the class of voting classifiers. The HAR results are determined by experiments on the UCI-HAR dataset.

The other sections of the paper are structured as follows: Section 2 depicts various studies done in this domain. Section 3 describes the proposed suggested methodology in which the voting classifier has been deployed. Section 4 represents the empirical results of the proposed methodology and a comparative study of the suggested methodology in contrast with others. Finally, in Section 5, a conclusion has been drawn that shows that the Voting Classifier-II (a combination of SVM, KNN, and LR) using soft voting outperformed other machine learning classifiers.

2. Related Work

Voluminous studies have been carried out in the fields of recognition of human activities and other aspects by exploiting artificial intelligence enabled computer-aided procedures. The researchers have used various isolated machine learning (ML) techniques and ensemble approaches wherein several individual models are performed together to get expected prediction results on several types of datasets. The research contributions are discussed in two aspects: firstly for the work related to voting classifiers for different applications; and secondly, for the work related to human activity recognition using different techniques.

Recent studies related to the usability of voting classifiers for different applications are discussed here. Kour *et al.* (2022) used the hybrid voting classifier system for the COVID-9 prediction for the dataset collected from the COVID-19 cases in Mexico. The methods of SVM, Bernoulli Naive Bayes, random forest, and Naive Bayes were adapted for the hybridization of voting classifiers. Nisar & Chhabra (2021) have projected a voting ensemble categorization method for the revealing of undesirable emails. In this research, the first experiment was performed for the individual classifiers and the second was voting, wherein the results of each individual classifier were summed up to form the outcome class using common votes. Kasubi & Huchaiah (2021) have concentrated on distinguishing human physical actions from day-to-day routines using the ARAS dataset. Adaboost collective and bagging techniques of classification were used for experimentation. Kumari *et al.* (2021) have intensively studied the machine learning algorithms that aid in finding cases of diabetes mellitus with greater accuracy. The authors have put forward an ensemble soft voting classifier that gives binary classification and uses a combination of three machine learning algorithms, i.e., Logistic Regression, Random Forest, and Gaussian Naive Bayes for the classification. Mahabub (2020) has proposed an ensemble voting classifier based on a sharp detection system to categorise real and fake news. The authors have tried and categorically analyzed at least eleven notable ML classification algorithms.

Furthermore, the work related to human activity recognition using different techniques is described. The recent trends, datasets, and issues in the fields are discussed by Nguyen *et al.* (2021). The authors majorly focused on the techniques of machine learning and deep learning for HAR. Tan *et al.* (2022)

presented an ensemble learning method in which deep neural network (DNN) is ensembled with a convolutional neural network (CNN) stacked on the GRU (a gated recurrent unit). Muhammad *et al.* (2021) adapted the methods of bi-directional long short-term memory (BiLSTM) and dilated CNN for the specific feature selection to recognize human activities. The experiments were conducted for the different datasets of J-HMD, UCF sports, and UCF11. Zaki *et al.* (2020) have studied and acknowledged the best performing classifiers by experiments on two publicly available datasets of HAR. The inference result shows that the logistic regression gave more accurate predictions as compared to other evaluated classifiers in this study for HAR. Tarafdar & Bose (2021) proposed boosting-based ML (Machine learning) methods and assessed their ability to predict natural human behaviour. Balli *et al.* (2019) have presented a hybrid method of Principal Component Analysis (PCA) and Random Forest (RF) which combines an efficient feature extraction technique with a classification algorithm. Deshmukh *et al.* (2018) used the data obtained from the accelerometer sensor, which is embedded in smartphones named as MobiAct dataset. The author used the KNN classifier as a classification algorithm to evaluate the performance of the model. Sukor *et al.* (2018) focused on the recognition of activities of daily living using an embedded accelerometer sensor in smartphones. The authors also used the feature selection method of Principal Component Analysis (PCA), and a comparison has been made on the performances between PCA-based features and the original raw data. For the classification of activities, machine learning classifiers such as Decision Tree (DT), Support Vector Machine (SVM), and Multi-Layer Perceptron Neural Network (MLP-NN) were used. The experiment results revealed that the PCA-based features have a better recognition rate than frequency-domain features.

Li *et al.* (2018) used hybrid deep-learning procedures such as convolutional and long-short-term memory (LSTM) for the Opportunity and UniMiB-SHAR datasets to obtain features. For classification, machine learning classifiers were adapted. Fu *et al.* (2018) analyzed that humans have started focusing on wearable gadgets that facilitate human beings to monitor the level and graph of their fitness on a regular basis. The authors explored the use of a speaker and microphones that are available on smartphones to identify the exercises performed nearer to them. Several classification methods were tested, ranging from SVMs, Naive Bayes, Random Forest, and AdaBoost to CNNs. Milenkoski *et al.* (2018) used LSTM, which is able to attain features from raw accelerometer data. The authors evaluated their algorithms on data collected in a laboratory-enabled framework, as well as on data collected in open and natural settings, and showed that their algorithm is strongly constructed on all logical parameters and outperforms almost equally well for both scenarios. Inoue *et al.* (2018) proposed a new mode of HAR with high throughput from unprocessed accelerometer sensor data by applying a deep recurrent neural network (DRNN) and identifying several structures and their integration to discover the optimum parameter values. Vavoulas *et al.* (2017) introduced a benchmark dataset, the MobiAct dataset, for smartphone-based HAR. This dataset is comprised of data for 50 subjects who performed 9 various activities of Daily Living (ADLs), recorded through the Orientation Sensors, Gyroscope, and Accelerometer embedded in smartphones, and 54 subjects who performed 4 different activities under types of falls. Further, feature selection was performed on this dataset to obtain optimized features, and classification techniques were also applied for the recognition of ADLs using the accelerometer recorded data. Ronao & Cho (2016) proposed a deep convolutional neural network (convnet) to perform as an explicit model for human activity recognition using the embedded sensors of smartphone by taking the advantages of inherent attributes of activities and one dimensional time - series signals. Machado *et al.* (2015) described a HAR methodology that relies on fundamental principles of feature extraction and feature selection techniques. A feature selection method is deployed in order to achieve improved clustering accuracy and decrease computational complexity. The various clustering techniques named Spectral Clustering, Affinity Propagation, K-Means, and Mean Shift were experimented with and evaluated.

Furthermore, the contributions specific to human activity recognition are briefed. Table 1 shows a comparison of related studies in terms of results obtained on various datasets and algorithms performed on them.

Table 1. Existing studies related to human activity recognition.

Authors and Year	Algorithm	Dataset Used	Best Classifier
Tan *et al.* (2022)	DNN ensembled with CNN stacked on the GRU	UCI-Opportunity, UCI-HAR, and UCI-WIDSM	DNN with CNN stacked on the GRU
Muhammad *et al.* (2021)	BiLSTM and dilated CNN	UCF sports, J-HMD, and UCF11	BiLSTM and dilated CNN
Zaki *et al.* (2020)	NB, KNN, LR, RF, GB	UCI-HAR and HAPT	LR
Tarafdar & Bose (2021)	XgBoost, AdaBoost, Boosted C5.0	Public Dataset	Adaboost
Balli *et al.* (2019)	RF, SVM, C4.5 and KNN	Sensor data of Moto 360 smart watch	PCA + RF combination
Deshmukh *et al.* (2018)	KNN	MobiAct	KNN
Sukor *et al.* (2018)	DT, SVM, MPL-NN (Multilayer Perceptrons-Neural Networks)	Accelerometer	MLP-NN + PCA
Li *et al.* (2018)	SVM, CNN, and LSTM	OPPORTUNITY UniMiB-SHAR	LSTM layers
Fu *et al.* (2018)	NB, SVM, RF, AdaBoost and CNN (Convolutional Neural Network)	Google Nexus 5X.	SVM
Milenkoski *et al.* (2018)	LSTM	Laboratory Data and Field Collections	LSTM
Inoue *et al.* (2018)	DRNN (Dilated Residual Networks)	HASC (Human Activity Sensing Consortium)	DRNN
Vavoulas *et al.* (2017)	IBk (Instance Based Learner)	MobiAct	IBk
Ronao & Cho (2016)	Convnet (Convolutional Neural Network)	Benchmark dataset	convnet
Machado *et al.* (2015)	K-Means, Affinity Propagation, Mean Shift and Spectral Clustering	Accelerometer	K-means

3. Research Methodology

In this research, we have presented the voting classifier system for improving the significant accuracy of activity classification under its superset, Human Activity Recognition. As we have suggested, a voting classifier for HAR, i.e., a group of individual classifiers that perform together as a voting type soft to get the desired results. We have done data preprocessing before feeding the training dataset to the suggested model, which is followed by Principal Component Analysis (PCA) for feature reduction. Figure 1 depicts the workflow of the proposed architecture.

The sub-sections of the proposed work for recognizing human activities based on data obtained from the embedded sensors of mobile phones are discussed as follows.

3.1 Data Collection

The input benchmark UCI-HAR dataset was composed by a team of thirty volunteers (age group 19–48 years) who performed the 6 activities of their normal daily routine while wearing a smartphone on their waist (Anguita *et al.*, 2013). These activities are labelled as standing, sitting, laying, walking, walking upstairs, and walking downstairs. With the availability of gyroscope and accelerometer embedded sensors in smartphones, the participant's tri-axial angular velocity at a constant rate of 50Hz and tri-axial linear acceleration are recorded in digital form. To obtain feature vectors, variables from the time and

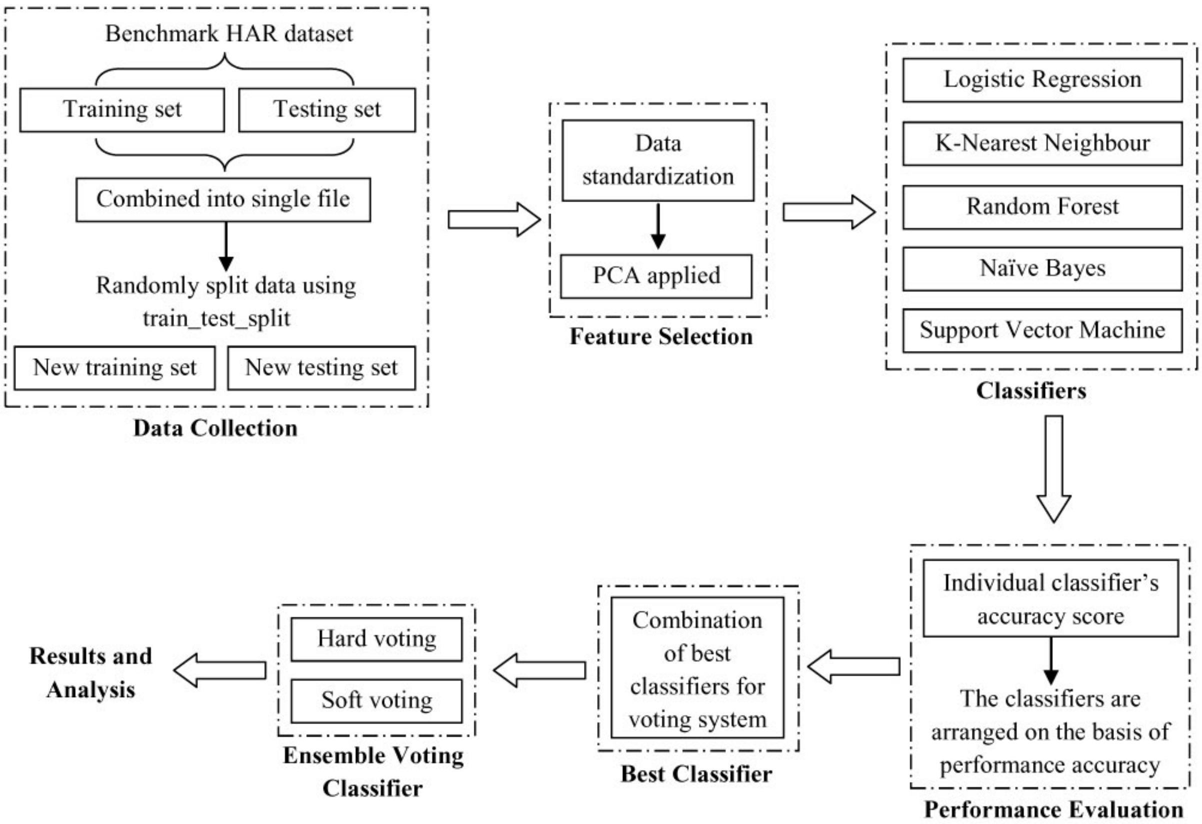

Fig. 1. Workflow of proposed technique.

frequency domains were calculated and listed in Table 2.

3.2 Feature Selection

This stage focuses on processing the entire dataset that is utilized as input so that the attributes of the dataset are mitigated. To seek a dimensionality lessening procedure, we found that Principal Component Analysis (PCA) is considered the most prominent preprocessing method that performs transforms of m-dimensional X data to n-dimensional X' data with the perseverance of relevant information in the dataset (Ray *et al.*, 2020). Therefore, PCA is an effective method for dimensionality reduction of the larger data set. It is a technique of linear dimensionality reduction that can be used by projecting it into a lower-dimensional sub-space to extract information from a high-dimensional field. Here, Principal Component Analysis is used in our implementation using python and the Scikit-learn library for dimensionality reduction. Moreover, by using PCA, we can speed up a machine learning algorithm. In some cases, decreasing the number of vectors or features causes a loss of accuracy, thus making the large data set smaller, easier to use, and visualize. To implement Principal Component Analysis practically using Python, the following automatic steps have been followed to obtain the required optimum set of features.

Step I: Standardized the values available in the dataset CSV file's cells. (With a mean of zero and a variance of one).

Step II: A matrix that shows the covariance of dimensions has been computed.

Step III: The eigenvalues and eigenvectors were acquired from the matrix referenced in step II.

Step IV: The projection matrix W has been constructed from the chosen k Eigenvectors.

Step V: The original data set X via W has been transformed to get the new k-dimensional feature subspace Y.

Summarily, the core principle of this technique is to reduce the dimensionality of a dataset consisting of additional variables interrelated with each other, either lightly or heavily, without affecting the variance that existed in the dataset.

Table 2. List of measures for computing feature vectors.

Sr. No.	Function	Description
1	mean	Mean value
2	std	Standard deviation
3	mad	Mean Absolute deviation
4	max	Maximum Value
5	min	Minimum Value
6	sma	Signal magnitude area
7	energy	Energy measures
8	iqr	Inter quartile range
9	entropy	Signal Entropy
10	arCoeff	Auto regression coefficients
11	correlation	Correlation coefficient between signals
12	maxFreqInd	With Largest Magnitude, Index of Frequency Components
13	meanFreq	Weighted Average of Frequency Component For Obtaining Mean Frequency
14	skewness	Skewness of frequency domain
15	kurtosis	Frequency signal Kurtosis
16	energyBand	Energy of a frequency interval
17	angle	Angle between the vectors

In this study, we have a total of 561 features in the dataset. After the implementation of PCA, 200 features are selected at the cost of minimum loss of information and also in view of increasing accuracy through the proposed ensemble voting method.

The features are reduced by using the PCA algorithm, in which the initial image data is normalized. Then the covariance matrix from the image data is calculated. The matrix data is processed for Single Value Decomposition (SVD) and finally, the projection of image data to the new basis with reduced is determined. The reduced features are the selected features that are used for further processing.

3.3 Classifiers

After the reduction of features in the dataset, the five different machine-learning classifiers for the classification of human activities are applied. These classifiers are LR, KNN, NB, RF, and SVM.

3.3.1 Logistic Regression (LR)

This classification algorithm is used to obtain a binary prediction as an output. The required binary outcome is determined by analyzing the independent variables with results exactly falling into one of two categories. Here, the dependent variable is always categorical, but the independent variables can be categorical or numeric.

3.3.2 K-Nearest Neighbour (KNN)

This algorithm classifies objects based on the most frequent class among the k-nearest neighbors. The k is the number of neighbours chosen by considering the minimal distance. To classify the input classes, it evaluates the distance from each object, and the object with the minimal distance is selected. The most frequent class among the k-nearest neighbours is assigned to the object.

3.3.3 Random Forest (RF)

This algorithm uses various decision trees on subsets of datasets to obtain better prediction accuracy. A resampling bootstrap technique is implemented on each decision tree from the set of training data. During the classification, each tree votes individually for the given input. This algorithm selects the class that gets the most votes.

3.3.4 Naive Bayes (NB)

This classification algorithm is based on the Bayes' theorem and classesifies datasets based on past results. The working effectiveness of Bayesian algorithms is subject to the correctness of their strong assumptions. The benefit of Naive Bayes is that it is fast to design and does not require a large training set.

3.3.5 Support Vector Machine (SVM)

This algorithm is based on the concept of the hyperplane in an N-dimensional space for the classification of data points. Here, N is the number of features. Data points are attributed to various classes based on their placement to the different sides of the hyperplane. The accuracy of the results is directly dependent on the selection of the hyperplane. A plane with the maximum distance between data points of both classes is selected.

3.4 Best Classifier

The above-mentioned classifiers are evaluated on the basis of their confusion matrix and classification reports. The F1 score and overall accuracy score are considered to give them a ranking as per their performance.

3.5 Voting Classifier

A Voting Classifier is a suggested machine learning procedure that works on an ensemble of various models and gives an output or class based on the result of their combined efforts, i.e., the highest probability predicted for the chosen class as the output. Hard voting and soft voting classifiers are used in this paper.

3.5.1 Hard Voting

In this type of voting, the projected output class is a class with the highest number of votes received by each individual classifier, i.e. the class that has obtained the highest probability of being chosen by each of the classifiers. For example, 3 classifiers predicted the output classes (P, P, Q), so here class P has the highest majority of votes. Hence, P will be the final predicted class.

3.5.2 Soft Voting

In this type of voting, the resulted class is the predictive output based on the average of the probabilities allotted to that class by the individual classifiers who participated in the voting process. For example, if there are three models, the prediction probability for class P = (0.27, 0.42, 0.32) and Q = (0.32, 0.28, 0.25). So the average for class P is 0.3367 and the average for class Q is 0.2833, indicating that class P is clearly the winner because it received the highest probability averaged by each classifier.

After experimentation, the performance results of the selected ML classifiers used for study and the combinations of best performer classifiers have been chosen for the proposed ensemble voting classifier under two variations, i.e. Hard Voting and Soft Voting. Apart from this, to increase the accuracy, the optimum tuning of hyperparameters and weighted voting have also been performed with considered ML classifiers and proposed ensemble voting classifiers, respectively, used for experimentation.

4. Results and Analysis

Each model is analyzed in terms of recognition accuracy, precision, recall, and F1 Score. The formulations of these measures are illustrated in Equations (1)-(4). The prediction outcome is evaluated and analyzed by calculating the confusion matrix and classification report. The parameters for the confusion matrix are depicted in Table 3.

$$\text{Precision} = \frac{\text{TP}}{\text{TP+FP}} \tag{1}$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP+FN}} \tag{2}$$

$$\text{F-Measure} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision+Recall}} \tag{3}$$

$$\text{Recognition Accuracy} = \frac{\text{Correctly Classified Instances}}{\text{Total Number of Instances}} \tag{4}$$

Where, TP, FN, TN, and FP are the True Positive, False Negative, True Negative, and False Positive instances, respectively.

Table 3. Parameters of confusion matrix.

		Predicted Class	
		Yes	No
Actual Class	Yes	True Positives	False Negatives
	No	False Positives	True Negatives

This research work proposes a hybrid methodology of ensemble voting classifiers for recognizing human activities, wherein, on the one hand, the performance of five ML classifiers has been evaluated and ranked one to five on the basis of their performance in terms of overall accuracy on the Benchmark dataset. On the other hand, the proposed ensemble voting classifiers are ML algorithms that are trained on an ensemble of combinations of top-performing classification models using both hard and soft voting.

For experiments, the overall data is initially combined to remove the bias, which is further divided into the training and testing ratio of 70:30. The dimensions of the training set are 7209 rows and 561 columns, while the dimensions of the testing set are 3090 rows and 561 columns, and the number of numeric features are 561. The PCA value is tuned to 200 components, which is kept common for all types of experiments. The results of the individual ML classifiers are depicted in Table 4.

The results outcomes (Table 4) of individual classifiers are represented by their ranks. The rank of the classifier is defined on the basis of the value of the recognition accuracy. Rank 1 is assigned to the classifier possessing the highest recognition accuracy, then rank 2 to the classifier having the second highest recognition accuracy, and similarly, rank 5 to the classifier determined with the least value of recognition accuracy. As per Table 4, the ranks of the classifiers are assigned from 1 to 5 to the classifiers SVM, KNN, LR, RF, and NB as per their recognition accuracy of 89.02%, 86.56%, 83.92%, 80.18%, and 72.05%, respectively. It is clearly demonstrated that SVM performs better among the classifiers chosen for this study.

By analyzing the performance of different ML classifiers, the combinations of best performer classifiers have been chosen for the proposed ensemble voting classifier under two variations, i.e. Hard Voting and Soft Voting. Apart from this, to increase the accuracy, the optimum tuning of hyperparameters, and weighted voting have also been performed with the considered ML algorithms and proposed ensemble voting classifier, respectively, used for experimentation.

Tables 5 and 6 depict the results of soft voting and hard voting, respectively, for all voting classifiers. The Voting Classifier-II (a combination of SVM, KNN, and LR) performs best among other voting classifiers, with a recognition accuracy of 92.78% with soft voting and 91.64% with hard voting.

Table 4. Ordering of classifiers based on performance analysis.

Classifier	Accuracy (%)	Precision (%)	Recall (%)	F1 Score (%)	Rank
LR	83.92	84.79	83.92	84.35	3
KNN	86.56	87.98	86.56	87.26	2
RF	80.18	81.37	80.18	80.77	4
NB	72.05	74.42	72.05	73.22	5
SVM	89.02	89.91	89.02	89.46	1

Table 5. Performance analysis of Soft Voting Classifiers.

Classifier	Algorithm Combinations	Accuracy (%)	Precision (%)	Recall (%)	F1 Score (%)
I	SVM, KNN	91.03	92.16	91.03	91.59
II	SVM, KNN, LR	92.78	93.06	92.78	92.92
III	SVM, KNN, LR, RF	91.38	92.47	91.38	91.92
IV	SVM, KNN, LR, RF, NB	90.11	91.54	90.11	90.82

Table 6. Performance analysis of Hard Voting Classifiers.

Classifier	Algorithm Combinations	Accuracy (%)	Precision (%)	Recall (%)	F1 Score (%)
I	SVM, KNN	90.18	90.41	90.18	90.29
II	SVM, KNN, LR	91.64	92.33	91.64	91.98
III	SVM, KNN, LR, RF	91.11	91.47	91.11	91.29
IV	SVM, KNN, LR, RF, NB	90.73	91.08	90.73	90.90

The comparative analysis of Table 5 and Table 6 reveals that the performance of the soft voting ensemble classifiers performed better as compared to the hard voting ensemble classifiers. The Voting Classifier – II obtained an overall accuracy of 92.78% in soft voting and 91.64% in hard voting ensemble classifiers. Further, the comparison of Voting Classifier – II (using soft voting type) has been presented with all the individual classifiers in Table 7.

Table 7 has compared the performance of the Voting Classifier-II (a combination of SVM, KNN, and LR) with the performance of individual classifiers. Voting Classifier – II (a combination of SVM, KNN, and LR) performs better than other classifiers. The graphical representation of this comparison is illustrated in Figure 2.

Figure 2 depicts the accuracy achieved by the proposed method, i.e., Voting Classifier – II soft voting is 1.14% better than hard voting. Among the individual classifiers, Voting Classifier – II (soft voting) is 8.86% better than LR, 6.22% than KNN, 12.6% than RF, 20.73% than NB, and 3.76% than SVM, respectively. This proves that Voting Classifier – II (soft voting) is superior to others.

Furthermore, the comparison of the voting classifiers is conducted with state-of-the-art techniques. The incorporated research studies are Ma *et al.* (2021), Xu *et al.* (2020), Chen *et al.* (2019), and Seto *et al.* (2015). Ma *et al.* (2021) presented the weighted support tensor machine (WSTM) for activity recognition. Xu *et al.* (2020) proposed an integrated concept of Gramian angular field (GAF) with multi-dilated kernel residual (Mdk-Res), and presented the novel approach of GAF+Fusion-Mdk-ResNet. The authors have also evaluated the results using the methods of GAF+GoogLeNet, GAF+ResNet, GAF+Conv_2D, LSTM, Conv_1D, and MLP. Chen *et al.* (2019) adapted the methods of the Semisupervised Recurrent Convolutional Attention Model (SRCAM) for activity recognition. Seto *et al.* (2015) used the approach of Dynamic Time Warping (DTW) Barycenter Averaging (DBA), which is termed as DTW-DBA ap-

Table 7. Performance Analysis of Voting Classifier – II with Individual Classifiers.

Classifier	Accuracy (%)	Precision (%)	Recall (%)	F1 Score (%)
Voting Classifier-II (Soft voting classifier)	92.78	93.06	92.78	92.92
Voting Classifier-II (Hard voting classifier)	91.64	92.33	91.64	91.98
LR	83.92	84.79	83.92	84.35
KNN	86.56	87.98	86.56	87.26
RF	80.18	81.37	80.18	80.77
NB	72.05	74.42	72.05	73.22
SVM	89.02	89.91	89.02	89.46

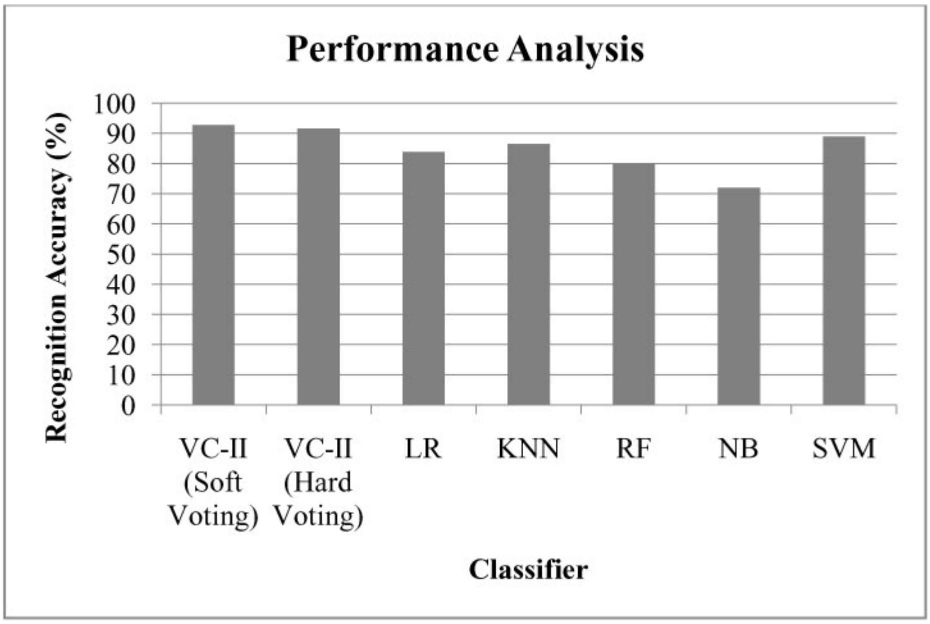

Fig. 2. Graphical representation of Performance Analysis of Voting Classifier-II with Individual Classifier.

Table 8. Comparison of Voting Classifier – II with state-of-the-art Techniques.

Classifier	Recognition Accuracy
Voting Classifier-II (Soft voting classifier)	92.78%
Voting Classifier-II (Hard voting classifier)	91.64%
WSTM (Ma *et al.* (2021))	89.4%
GAF+Fusion-Mdk-ResNet (Xu *et al.* (2020))	89.48%
GAF+GoogLeNet (Xu *et al.* (2020))	87.61%
GAF+ResNet (Xu *et al.* (2020))	87.75%
GAF+Conv_2D (Xu *et al.* (2020))	88.19%
LSTM (Xu *et al.* (2020))	80.90%
Conv_1D (Xu *et al.* (2020))	85.41%
MLP (Xu *et al.* (2020))	80.79%
SRCAM (Chen *et al.* (2019))	81.32%
DTW-DBA (Seto *et al.* (2015))	86%

proach. The comparison of the proposed voting classifiers with these techniques is illustrated in Table 8.

The comparison of the proposed voting classifiers with the state-of-the-art techniques indicates that the proposed voting classifiers are superior to other techniques for activity recognition.

5. Conclusion

The paper has proposed the voting classifier systems using machine learning algorithms for human activity recognition. Initially, five notable problem-solving machine procedures (LR, KNN, RF, NB, and SVM) were empirically evaluated in terms of recognition accuracy. The superiority order of performance is achieved as SVM, KNN, LR, RF, and NB for the benchmark UCI – HAR dataset. Further, this experimentation has nominated voting classifier models with hard and soft voting. Among the different combinations of classifiers, a combination of three machine learning algorithms (SVM, KNN, and LR) has performed best of all. This study reveals that the Voting Classifier-II (a combination of SVM, KNN, and LR) using soft voting outperformed the machine learning classifiers and stat-of-the-art techniques by achieving an accuracy of 92.78%.

The future possibilities are testing this proposed model with video-based HAR datasets such as the UCI-101 dataset can also be possible. This model can also be optimized with deep neural networks and optimization techniques such as swarm optimization techniques.

References

Anguita, D., Ghio, A., Oneto, L., Parra, X., & Reyes-Ortiz, J. L. (2013). A public domain dataset for human activity recognition using smartphones. In *European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning* (pp. 437-442). Bruges, Belgium.

Balli, S., Sağbaş, E. A., & Peker, M. (2019). Human activity recognition from smart watch sensor data using a hybrid of principal component analysis and random forest algorithm. *Measurement and Control*, **52**(1-2), 37-45.

Chen, K., Yao, L., Zhang, D., Wang, X., Chang, X., & Nie, F. (2019). A semisupervised recurrent convolutional attention model for human activity recognition. *IEEE transactions on neural networks and learning systems*, **31**(5), 1747-1756.

Deshmukh, R., Aware, S., Picha, A., Agrawal, A., & Wable, S. D. (2018). Human Activity Recognition using Embedded Smartphone Sensors. *International Research Journal of Engineering and Technology*, **5**(4), 4217-4220.

Fu, B., Kirchbuchner, F., Kuijper, A., Braun, A., & Gangatharan, D. V. (2018). Fitness activity recognition on smartphones using Doppler measurements. *Informatics*, **5**(2), 24(1-14).

Ijjina, E. P., & Mohan, C. K. (2014). Human action recognition based on recognition of linear patterns in action bank features using convolutional neural networks. In *2014 13th International Conference on Machine Learning and Applications* (pp. 178-182). Detroit, MI, USA: IEEE.

Inoue, M., Inoue, S., & Nishida, T. (2018). Deep recurrent neural network for mobile human activity recognition with high throughput. *Artificial Life and Robotics*, **23**(2), 173-185.

Kasubi, J.W., & Huchaiah, M. D. (2021). Human Activity Recognition for Multi-label Classification in Smart Homes Using Ensemble Methods. In *International Conference on Artificial Intelligence and Sustainable Computing* (pp. 282-294). Springer, Cham.

Khan, W., Daud, A., Nasir, J.A., & Amjad, T. (2016). A survey on the state-of-the-art machine learning models in the context of NLP. *Kuwait journal of science*, **43**(4), 95-113.

Kour, G., Abrolb, P., Kalrupiac, N., & Kaur, J. (2022). Hybrid voting classifier model for COVID-19 prediction by embedding machine learning techniques. *Turkish Journal of Computer and Mathematics Education (TURCOMAT)*, **13** (2), 171-183.

Kumari, S., Kumar, D., & Mittal, M. (2021). An ensemble approach for classification and prediction of diabetes mellitus using soft voting classifier. *International Journal of Cognitive Computing in Engineering*, **2**, 40-46.

Li, F., Shirahama, K., Nisar, M.A., Köping, L., & Grzegorzek, M. (2018). Comparison of feature learning methods for human activity recognition using wearable sensors. *Sensors*, **18**(2), 679(1-22).

Ma, Z., Yang, L.T., Lin, M., Zhang, Q., & Dai, C. (2021). Weighted support tensor machines for human activity recognition with smartphone sensors. *IEEE Transactions on Industrial Informatics*, DOI: 10.1109/TII.2021.3061559.

Machado, I.P., Gomes, A.L., Gamboa, H., Paixão, V., & Costa, R. M. (2015). Human activity data discovery from triaxial accelerometer sensor: Non-supervised learning sensitivity to feature extraction parametrization. *Information Processing & Management*, **51**(2), 204-214.

Mahabub, A. (2020). A robust technique of fake news detection using Ensemble Voting Classifier and comparison with other classifiers. *SN Applied Sciences*, **2**(4), 1-9.

Milenkoski, M., Trivodaliev, K., Kalajdziski, S., Jovanov, M., & Stojkoska, B. R. (2018). Real time human activity recognition on smartphones using LSTM networks. In *2018 41st International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO)* (pp. 1126-1131). Opatija, Croatia: IEEE.

Muhammad, K., Ullah, A., Imran, A. S., Sajjad, M., Kiran, M.S., Sannino, G., & de Albuquerque, V.H.C. (2021). Human action recognition using attention based LSTM network with dilated CNN features. *Future Generation Computer Systems*, **125**, 820-830.

Nguyen, B., Coelho, Y., Bastos, T., & Krishnan, S. (2021). Trends in human activity recognition with focus on machine learning and power requirements. *Machine Learning with Applications*, **5**, 100072(1-17).

Nisar, N., Rakesh, N., & Chhabra, M. (2021). Voting-Ensemble Classification for Email Spam Detection. In *2021 International Conference on Communication information and Computing Technology (ICCICT)* (pp. 1-6). Mumbai, India: IEEE.

Ray, S., AlGhamdi, A., Alshouiliy, K., & Agrawal, D. P. (2020). Selecting Features for Breast Cancer Analysis and Prediction. In *2020 International Conference on Advances in Computing and Communication Engineering (ICACCE)* (pp. 1-6). Las Vegas, USA: IEEE.

Ronao, C. A., & Cho, S. B. (2016). Human activity recognition with smartphone sensors using deep learning neural networks. *Expert systems with applications*, **59**, 235-244.

Seto, S., Zhang, W., & Zhou, Y. (2015). Multivariate time series classification using dynamic time warping template selection for human activity recognition. In *2015 IEEE symposium series on computational intelligence* (pp. 1399-1406). Cape Town, South Africa: IEEE.

Sjarif, N. N. A., & Shamsuddin, S. M. (2015). Human action invarianceness for human action recognition. In *2015 9th International Conference on Software, Knowledge, Information Management and Applications (SKIMA)* (pp. 1-7). Kathmandu, Nepal: IEEE.

Sukor, A. A., Zakaria, A., & Rahim, N. A. (2018). Activity recognition using accelerometer sensor and machine learning classifiers. In *2018 IEEE 14th International Colloquium on Signal Processing & Its Applications (CSPA)* (pp. 233-238). Penang, Malaysia: IEEE.

Tan, T.H., Wu, J.Y., Liu, S.H., & Gochoo, M. (2022). Human activity recognition using an ensemble learning algorithm with smartphone sensor data. *Electronics*, **11** (3), 322(1-17).

Tarafdar, P., & Bose, I. (2021). Recognition of human activities for wellness management using a smartphone and a smartwatch: a boosting approach. *Decision Support Systems*, **140**, 113426(1-40).

Vavoulas, G., Chatzaki, C., Malliotakis, T., Pediaditis, M., & Tsiknakis, M. (2017). The mobiact dataset: Recognition of activities of daily living using smartphones. In *International Conference on Information and Communication Technologies for Ageing Well and e-Health* (pp. 143-151). SCITEPRESS.

Xu, H., Li, J., Yuan, H., Liu, Q., Fan, S., Li, T., & Sun, X. (2020). Human activity recognition based on Gramian angular field and deep convolutional neural network. *IEEE Access*, **8**, 199393-199405.

Yayan, U., Yazici, A., & SARICICEK, I. (2021). Prognostics-aware multi-robot route planning to extend the lifetime. *Kuwait journal of science*, DOI: 10.48129/kjs.13237. 1-14.

Zaki, Z., Shah, M.A., Wakil, K., & Sher, F. (2020). Logistic regression based human activities recognition. *Journal of Mechanics of Continua and Mathematical Sciences*, **15** (4), 228-246.

Submitted:	28/02/2022
Revised:	08/06/2022
Accepted:	15/06/2022
DOI:	10.48129/kjs.splml.19189