Computer Science and Information Systems 19(1):47–63 https://doi.org/10.2298/CSIS201221043N

Comparative Analysis of HAR Datasets Using Classification Algorithms

Suvra Nayak[1], Chhabi Rani Panigrahi[1], Bibudhendu Pati[1], Sarmistha Nanda[1], and Meng-Yen Hsieh[2]

[1] Department of Computer Science, Rama Devi Women's University, Bhubaneswar, India
{suvra.nayak24, panigrahichhabi, patibibudhendu, sarmisthananda}@gmail.com
[2] Department of Computer Science & Information Engineering, Providence University, Taiwan
mengyen@gm.pu.edu.tw

Abstract. In the current research and development era, Human Activity Recognition (HAR) plays a vital role in analyzing the movements and activities of a human being. The main objective of HAR is to infer the current behaviour by extracting previous information. Now-a-days, the continuous improvement of living condition of human beings changes human society dramatically. To detect the activities of human beings, various devices, such as smartphones and smart watches, use different types of sensors, such as multi modal sensors, non-video based and video-based sensors, and so on. Among the entire machine learning approaches, tasks in different applications adopt extensively classification techniques, in terms of smart homes by active and assisted living, healthcare, security and surveillance, making decisions, tele-immersion, forecasting weather, official tasks, and prediction of risk analysis in society. In this paper, we perform three classification algorithms, Sequential Minimal Optimization (SMO), Random Forest (RF), and Simple Logistic (SL) with the two HAR datasets, UCI HAR and WISDM, downloaded from the UCI repository. The experiment described in this paper uses the WEKA tool to evaluate performance with the matrices, Kappa statistics, relative absolute error, mean absolute error, ROC Area, and PRC Area by 10-fold cross validation technique. We also provide a comparative analysis of the classification algorithms with the two determined datasets by calculating the accuracy with precision, recall, and F-measure metrics. In the experimental results, all the three algorithms with the UCI HAR datasets achieve nearly the same accuracy of 98%.The RF algorithm with the WISDM dataset has the accuracy of 90.69%,better than the others.

Keywords: Machine Learning, Human Activity Recognition, WEKA, Classifier, Classification Algorithms.

1. Introduction

Due to recent scientific research efforts, Machine Learning (ML) [1] as an essential branch of computer science has emerged out of Artificial Intelligence (AI). Based on the importance of the logical and knowledge-based approaches, there is a rift between ML and AI. ML mechanism relying on database technology is the extensive use of data technological items [12]. Various problems such as HAR, air quality prediction[21], key sentence extraction using the comments in the blogs[22], detection of emergency situation[23], face recognition for security purpose [24],[30], film review analysis for maximizing the profit of investors and recommendation for viewers [25], intrusion detection[26], [31] can be solved using the ML techniques[27]. Human Activity Recognition (HAR) plays an essential role in different fields, such as human-computer interaction, health care, and security surveillance [2], as one of the prominent research branches. Due to specific challenges like optimal sensor placement, sensor motion, inherent variability, and cluttered background, HAR remains a very intricate task [3], [4]. HAR systems can replace human operators to intensify the proficiency and fruitfulness of the analysis and observation processes. For example, one of the HAR systems inform users about an emergent situation by tracking their health conditions with the help of specific sensor devices. Disaster management is a research area which attracts researchers of different communities like health care, computer science, business, and disaster management etc. The disaster recovery systems need to be designed using effective fault-tolerant techniques [41]. HAR plays a very important role during any kind of emergency. For collecting information on human behaviour, the steps with raw sensor data are concluded [10], [13], [15]: (a) pre-processing, (b) preliminary classification, (c) select classifier, and (d) performance evaluation. Classification is a widely used way of ML techniques, while the datasets consisting of training data and testing data are required. The training dataset always comprises a set of characteristics, and the primary role of classification divides the dataset to determine the classes [14]. There are several classification algorithms in the literature for resolving multiple challenges such as Multilayer Perceptron (MLP), SMO, decision tree, J48, RF, SL, Artificial Neural Network (ANN), Convolutional Neural Network (CNN), Support Vector Machine (SVM) [28-29], [32], [42] and others. The essential components of classification techniques [15] are shown in Figure 1.

The remaining of the article is organized as follows: Section 2 summarizes the related works proposed in the literature. The methodology adopted in our experiments is presented in Section 3. Section 4 summarizes and analyzes the experimental results. Section 5 presents a detailed comparison of the adopted classification algorithms with the chosen HAR datasets and finally, we provide concluding remarks and future directions of this research in Section 6.

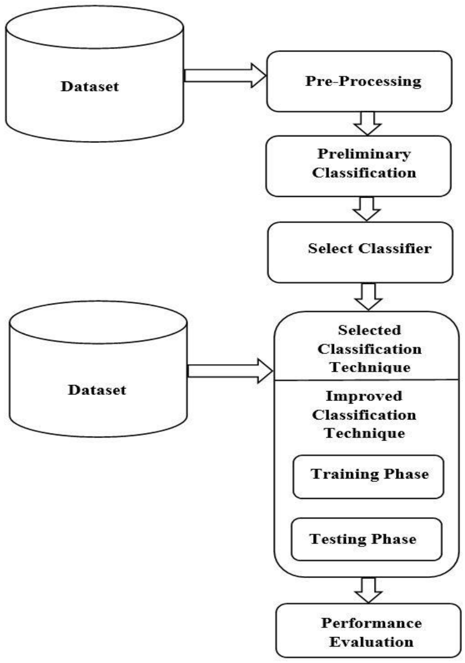

Fig. 1. Components of Classification Techniques

2. Related Work

In this section, a brief study of HAR datasets used for various applications is presented. Various researchers used different datasets for HAR in the literature [40]. The details of the datasets used are presented in Table 1.

Authors in [41] used different classifiers such as RF, IBK, J48, Bagging, and MLP for HAR. From the experimental results of the authors, it is indicated that RF performs well as compared to other considered classifiers and they achieve the 87.19 % accuracy. In [33], different classifiers like PEF, FNN, PTN, and PDF were used by the authors on various HAR datasets i.e. WISDM, MHEALTH, and SPAR. From the experimental results obtained, FNN was found to be the best classifier by the authors. The authors used categorical cross-entropy loss, the embedding and triplet loss for training. Also the authors observed that the training can be improved by using subject triplet selection. In [37], Yang *et al.* used DPCRCN for classification which uses end-to-end learning. During experimentation, they used Adam optimizer with the ReLU activation function. In [38], authors used HMDB, UCF101 and Kinetic datasets in their work and used a supervised approach where the weights of the branch were learned with standard back-propagation. The relation schema was integrated with an appearance branch and a Smart

Block was created to capture the spatiotemporal information. Then multiple Smart Blocks were stacked up to construct ARTNet.

Table 1. Details of HAR Datasets used for Various Applications

Author [Year]	**Dataset Used**	**Tools/ Framework Used**	**Classifiers Used**	**Best Classifier**	**Accuracy in %**
Nanda *et al.* [2021] [41]	WISDM Smartphone and Smartwatch Activity and Biometric Dataset	WEKA	RF, IBK, Bagging, J48 and MLP	RF	87.19
Burns *et al.* [2020] [33]	WISDM, MHEALTH, and SPAR	Seglearn, Keras, and Python Scikit-learn library	FCN, PEF, PDF, PTN	PTN	WISDM: 91. 3 MHEALTH: 99.9 SPAR: 99.0
Yang *et al.* [2019] [37]	AReM	LSTM, Fully connected Layer, and Softmax	LR, RF, SVM, DPCN, LSTM, XgBoost, LISEN, IDNNs, Dual Path Convolutional Neural Network (DPCRCN)	DPCRCN	99.97
Wang *et al.* [2018] [38]	Kinetics, HMDB51, UCF101	SMART Blocks, Two stream CNN, 3D CNNs, and ARTNets	C3D and ARTNet	ARTNet	Kinetics: 78. 7 HMDB51: 70.9 UCF101: 94.3
Min-Cheol Kwon *et al.* [2018] [39]	HAR dataset	Python Scikit-learn library	DT, RF, and SVM, ANN	ANN	95
Jain *et al.* [2017] [34]	Physical Activity Sensor data, UCI HAR	Score level fusion and Feature level fusion	k-NN and Multiclass SVM	Multiclass SVM	Physical Activity Sensor data: 96.83, UCI HAR: 97.12
Walse *et al.* [2016] [35]	WISDM	WEKA and Adaboost.M1	Decision Stump, Random Tree, RF, Hoeffding Tree, REP Tree, J48	J48	97.83
Kutlay *et al.* [2015] [36]	MHEALTH	WEKA	SVM and MLP	MLP	91.7

In [39], authors developed a HAR system where data from an off-the-shelf smartwatch was collected and ANN was used for human activity classification. The proposed system was improved by the authors by considering location information. From the experimental results of authors it was observed that an accuracy of 95% was achieved using ANN. Jain *et al.* [34] used UCI-HAR dataset and Physical Activity Sensor data for activity recognition and are publicly available sensor-based datasets. Here SVM and kNN classifiers were used by the authors. The simulation results indicate that SVM performs best for both of the datasets. Authors used a histogram of centroid and for feature extraction, gradient signature-based Fourier descriptor was utilized and then for information fusion, score and feature level fusion were combined. In [35], Walse *et al.* used different classifiers like Random Tree, J48, Hoeffding tree, RF, Decision Stump, and REP tree a log with MetaAdaboost.M1 for classification. All considered classification models were experimented with 10-fold cross-validation technique and J48 with MetaAdaboost.M1 was found to give improved results as compared to other algorithms. In [36], Kutlay *et al.* applied SVM and MLP classifiers on the MHEALTH dataset for classification. From the experimental results by the authors, it was found that MLP with 10 fold cross-validations gave 91.70% accuracy.

3. Methodology

In this section, we choose and discuss two HAR datasets downloaded from the UCI ML repository [16] and three classification algorithms. In this paper, we use the WEKA software as an open-source tool to demonstrate the classification algorithms.

Table 2. Description of UCI HAR and WISDM Datasets

Description/Dataset	**UCI-HAR**	**WISDM**
Instances	10299	15630426
Attributes	561	6
No. of Subjects	30	51
Activities	6	18
Characteristics	Multivariate/ Time-Series	Multivariate/ Time-Series
Associated Tasks	Classification/ Clustering	Classification
Sampling Rate	50 Hz	20 Hz
Device used	Smartphone: Samsung Galaxy 2	Smartphone: Google Nexus 5/5x or Samsung Galaxy S5 Smart watch: LG G Watch
Type	Filtered (Butterworth low pass filter)	Raw data as collected
Sensors used	Accelerometer, Gyroscope	Accelerometer, Gyroscope
File Type	Text	CSV

3.1. Datasets

We detail each of the datasets in Table 2. The UCI-HAR dataset is one of the chosen HAR datasets constructed from the recordings of 30 subjects performing activities of daily living while carrying a waist-mounted smartphone with embedded inertial sensors [5]. The dataset is characterized by multivariate and time series. The other dataset, WISDM, contains time-series sensor data of accelerometer and gyroscope, collected from 51 test subjects with 18 activities using smartphones and smart watches [6]. The characteristic of this dataset is multivariate on actual time-series and attribute characteristics.

3.2. Classification algorithms

The three classification algorithms are as follows: SMO, RF, and SL. We used WEKA tool to evaluate the algorithms with the two HAR datasets, provide a comparison of these algorithms, and suggest which algorithms may perform best. The specifics of the classification algorithms are described in this section below.

- **Sequential Minimal Optimization Algorithm (SMO):** This algorithm is developed by John C. Platt, Microsoft researcher in the year 1998 [18]. During the training of SVM, SMO is proposed to solve various quadratic problems. The worst-case time complexity of SMO as one of supervised classification algorithms is $O(n^3)$. Generally, SMO breaks down a significant problem into the sub-problems using the divide and conquer method, and solves them through analysis. SMO performs the functions of polynomial or RBF kernels to solve the classification problems, implemented in the popular LIBSVM tool [9] and widely used for training SVM. Using Support Vector Machine with sparse datasets, SMO is found to be the fastest algorithm.
- **Random Forest Algorithm (RF):**RF is developed by Breiman [19].This algorithm is designed for regression, classification, and other tasks by building a multitude of decision trees during training time and output the class, while the class represents the mode of classification or mean prediction of the trees [7],[8]. RF is one of the supervised learning algorithms, and mostly used for classification. By originating decision trees on the training data, the algorithm can make the prediction from each of the data samples. Generally, RF selects the best solution using the voting technique, and reduces the over fitting by averaging the results of calculated classification. Consequently, RF consisting of multiple single trees is an ensemble method, better than a single decision tree.
- **Simple Logistic Algorithm (SL):** This algorithm proposed by Sumner *et al.* in the year 2005 [20] is used for modelling the possibility of a particular event or certain class such as fail/pass, or lose/win. The algorithm for supervised learning tasks applies the simple logistic function on the data to predict the probability of a target variable. Besides, the algorithm is used to model a binary dependent variable, while every event would be assigned a probability value between 0 and 1. In logistic regression, estimating the parameters of a logistic model is applied in various fields such as ML, medical fields, and social sciences. Logistic regression is a statistical model used to model a binary dependent variable using a logistic function, although

many variations exist. In regression analysis, logistic regression estimates a logistic model with given parameters, as a form of binary regression [11].

The different features of these three classification algorithms are given in Table 3.

Table 3. Features of SMO, RF, and SL Classification Algorithms

Algorithm	SMO	RF	SL
Primary Problem	Classification	Classification and Regression	Classification can be done but good for Regression
Class Type	Binary and Multiclass	Binary and Multiclass	Good for Binary but Multiclass is also possible
Solution Approach	Quadratic Programming	Uncorrelated forest of trees	Statistical Learning
Dataset Type	Large	Large	Small
Time Complexity	$O(n^3)$	O(v * n (log (n))	O(n)
Data Normalization	Required	Not Required	Not Required
Raw Implementation	Difficult	Difficult	Easy
Predictors	Categorical or Numeric	Categorical or Numeric	Numeric

4. Results and Analysis

We have presented the results obtained from our experimentation, along with the accuracy analysis of the classifiers in this section. We considered two HAR datasets, UCI HAR and WISDM, to evaluate the quality of each of the classifiers. Three particular classifiers corresponding to SMO, RF, and SL algorithms were generated with the two datasets. We measured a number of parameters such as the accuracy percentage; the number of correctly classified instances, and the error percentage, while as the three metrics of the accuracy are adopted, corresponding to the values of precision, recall, and F-measure. Using the confusion matrix and weighted average computation, we calculated all these measures for each of the classifiers. The sum of diagonals in the matrix denotes the number of correctly classified instances. True Positive (TP) and False Positive (FP) denote the true positive and false-positive rates. Suppose that TP_A, TP_B, and TP_C denote the TP number of class A, class B, and class C, individually, the accuracy value is computed in Equation (1) as follows:

$$\text{Accuracy} = \frac{TP_A+TP_B+TP_C}{Total\ number\ of\ classification} \quad (1)$$

The other metrics related to accuracy are used for evaluating the performance results of each of the classifiers, as follows:

Precision: This metric is called positive predictive value as the fraction of appropriate instances among the retrieved instances. The formula of Precision is defined in Equation (2), as follows:

$$\text{Precision}= \frac{TP}{TP+FP} \quad (2)$$

Recall: Recall, also called sensitivity, is the fraction of the number of appropriate instances retrieved [17], defined in Equation (3):

$$\text{Recall}= \frac{TP}{TP+FN} \quad (3)$$

F-Measure: Precision and accuracy are combined into the calculation of F-Measure. Keeping in view the weighted average of both values, F-measure in Equation (4) is calculated as follows:

$$\text{F}= \frac{2*precision*recall}{precision+recall} \quad (4)$$

MCC: The Matthews Correlation Coefficient (MCC) correlates with the actual and predictive series. The formula of MCC always returns a value between -1 and +1 [17], defined as Equation (5).

$$\text{MCC}=\sqrt{\frac{X^2}{n}}, \text{ where } n \text{ is the total number of observations.} \quad (5)$$

Kappa statistic: The kappa statistic, κ measures the inter-related reliability for the qualitative items. Since κ considers the agreement possibility that occur by chance [17], the formula of κ is defined as follows.

$$\kappa \equiv \frac{P_0 - P_e}{1-P_e}=1-\frac{1-P_0}{1-P_e} \quad (6)$$

Where, P_0 is the relative observed agreement among raters and is identical to accuracy. P_e represents the hypothetical probability of chance agreement. The observed data is used to calculate the probabilities of each observer randomly by considering each category [17].

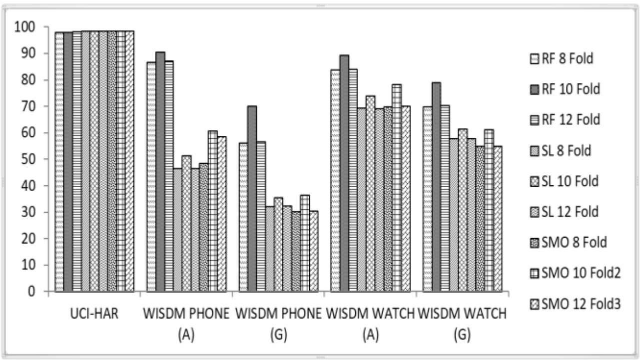

Fig. 2. Accuracy % Graph for 8, 10, and 12 Fold Cross Validation

In this work, the experimental results of the classifiers were operated on the considered datasets, while the classifiers were implemented by the SMO, RF, and Simple Logistics algorithms. We have used 8-fold, 10-fold, and 12-fold cross-validation to estimate the performance of all of the three considered algorithms during our experimentation. Based on the results, it was found that the cases of 10-fold cross-validation have a good accuracy performance on all the three algorithms. The experimental results of all the considered algorithms with different cross validations are presented in Figure 2.

The results obtained for each of the algorithms in terms of the different parameters such as TP rate, FP rate, Precision, Recall, F-Measure, MCC, ROC area, and PRC area are given in Tables 4-12 based on the datasets. The WISDM dataset contains the sensor data, collected using mobile devices such as phone and watch. Accelerometer and Gyroscope as sensors are embedded to the considered devices. The classification algorithms identify their activities with the data received from the phone and watch devices.

Table 4. Weighted Average of Accuracy of the SMO Algorithm with 8 Fold Cross Validation

Datasets	Device	TP rate	FP rate	Precisi on	Recall	F-Measure	MC C	ROC Area	PRC Area
UCI HAR	Phone	**0.99**	0.00	0.99	0.99	0.99	0.98	0.99	0.98
WISDM	Phone A	0.49	0.03	0.48	0.49	0.48	0.45	0.89	0.39
	Phone G	0.30	0.04	0.29	0.30	0.29	0.25	0.81	0.23
	Watch A	**0.70**	0.02	0.70	0.70	0.70	0.68	0.95	0.61
	Watch G	0.55	0.03	0.55	0.55	0.55	0.52	0.90	0.43

Table 5. Weighted Average of Accuracy of the RF Algorithm with 8 Fold Cross Validation

Datasets	Device	TP rate	FP rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area
UCI HAR	Phone	**0.98**	0.00	0.98	0.98	0.98	0.78	1.00	1.00
ISDM	Phone A	**0.87**	0.01	0.87	0.87	0.87	0.86	0.99	0.93
	Phone G	0.56	0.03	0.56	0.56	0.56	0.53	0.92	0.60
	Watch A	0.84	0.01	0.84	0.84	0.84	0.83	0.99	0.90
	Watch G	0.70	0.02	0.70	0.70	0.69	0.68	0.96	0.75

Table 6. Weighted Average of Accuracy of the SL Algorithm with 8 Fold Cross Validation

Datasets	Device	TP rate	FP rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area
UCI HAR	Phone	**0.99**	0.00	0.99	0.99	0.99	0.98	0.99	0.99
WISDM	Phone A	0.47	0.03	0.45	0.47	0.45	0.43	0.89	0.45
	Phone G	0.32	0.04	0.30	0.32	0.30	0.27	0.84	0.30
	Watch A	**0.69**	0.02	0.69	0.69	0.69	0.67	0.97	0.75
	Watch G	0.58	0.03	0.57	0.58	0.57	0.55	0.93	0.59

Tables 4-6 present the classification results in terms of accuracy measures such as TP, FP, Recall, F-Measure, Precision, MCC, PRC area, and ROC area for the considered datasets with 8 fold cross validation. The dataset is partitioned into 8 different sets randomly. Among them one set behaves as validation set whereas remaining sets act as training set. This was repeated for 8 times by considering each partition as validation set and the results are averaged to get the prediction. In UCI-HAR dataset, SMO, RF, and SL algorithms produce nearly the same results. Also, in UCI HAR dataset, SMO, RF as well as SL algorithm results in higher accuracy percentage whereas in WISDOM dataset the accuracy percentage was found to be less in all types of sensors data and Phone with

Accelerometer sensor performs better among them in RF algorithm with an accuracy of 87%.

Table 7. Weighted Average of Accuracy of the SMO Algorithm with 10 Fold Cross Validation

Datasets	Device	TP rate	FP rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area
UCI HAR	Phone	**0.98**	0.00	0.98	0.98	0.98	0.98	0.99	0.97
WISDM	Phone A	0.60	0.02	0.60	0.60	0.60	0.58	0.93	0.53
	Phone G	0.36	0.03	0.35	0.36	0.35	0.31	0.84	0.29
	Watch A	**0.78**	0.01	0.78	0.78	0.78	0.77	0.97	0.72
	Watch G	0.61	0.02	0.60	0.61	0.60	0.58	0.93	0.51

Table 8. Weighted Average of Accuracy of the RF Algorithm with 10 Fold Cross Validation

Datasets	Device	TP rate	FP rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area
UCI HAR	Phone	**0.98**	0.00	0.98	0.98	0.98	0.97	0.99	0.99
WISDM	Phone A	**0.90**	0.00	0.90	0.90	0.90	0.90	0.99	0.95
	Phone G	0.70	0.01	0.69	0.70	0.69	0.68	0.95	0.76
	Watch A	0.89	0.00	0.89	0.89	0.89	0.88	0.99	0.94
	Watch G	0.79	0.01	0.78	0.79	0.78	0.77	0.97	0.84

Table 9. Weighted Average of Accuracy of the SL Algorithm with 10 Fold Cross Validation

Datasets	Device	TP rate	FP rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area
UCI HAR	Phone	**0.98**	0.00	0.98	0.98	0.98	0.98	0.99	0.99
ISDM	Phone A	0.51	0.02	0.50	0.51	0.50	0.48	0.91	0.52
	Phone G	0.35	0.03	0.34	0.35	0.34	0.30	0.85	0.34
	Watch A	**0.74**	0.01	0.73	0.74	0.73	0.72	0.97	0.78
	Watch G	0.61	0.02	0.60	0.61	0.60	0.58	0.94	0.63

Table 10. Weighted Average of Accuracy of the SMO Algorithm with 12 Fold Cross Validation

Datasets	Device	TP rate	FP rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area
UCI HAR	Phone	**0.99**	0.00	0.99	0.99	0.99	0.98	0.99	0.98
WISDM	Phone A	0.49	0.03	0.48	0.49	0.48	0.45	0.89	0.39
	Phone G	0.31	0.04	0.29	0.30	0.29	0.26	0.81	0.23
	Watch A	**0.70**	0.02	0.70	0.70	0.70	0.68	0.95	0.61
	Watch G	0.55	0.03	0.55	0.55	0.55	0.52	0.90	0.44

Like 8 fold cross validation, the result of 10 fold cross validation was also estimated and are provided in the Tables 7-9 using all the considered algorithms and datasets. From the simulation results, it is found that the prediction accuracy is improved in 10 fold cross validation as compared to the 8 fold cross validation. Using the UCI-HAR dataset, SMO, RF, and SL algorithms produce the same results and the accuracy percentage is about 98%. Using the WISDM dataset, we computed the accuracy of the classifiers and is found to give better accuracy as compared to 8 fold cross validation in all types of sensors data.

Table 11. Weighted Average of Accuracy of the RF Algorithm with 12 Fold Cross Validation

Datasets	Device	TP rate	FP rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area
UCI HAR	Phone	**0.98**	0.00	0.98	0.98	0.98	0.78	0.99	0.99
WISDM	Phone A	**0.87**	0.01	0.88	0.87	0.87	0.87	0.99	0.94
	Phone G	0.57	0.03	0.56	0.57	0.56	0.54	0.92	0.60
	Watch A	0.84	0.01	0.84	0.84	0.84	0.83	0.99	0.90
	Watch G	0.70	0.02	0.70	0.70	0.70	0.68	0.96	0.75

Table 12. Weighted Average of Accuracy of the SL Algorithm with 12 Fold Cross Validation

Datasets	Device	TP rate	FP rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area
UCI HAR	Phone	**0.99**	0.00	0.99	0.99	0.99	0.98	0.99	0.99
WISDM	Phone A	0.47	0.03	0.45	0.47	0.46	0.42	0.89	0.46
	Phone G	0.33	0.04	0.31	0.33	0.31	0.28	0.84	0.31
	Watch A	**0.69**	0.02	0.69	0.69	0.69	0.67	0.97	0.75
	Watch G	0.58	0.03	0.57	0.58	0.57	0.55	0.93	0.59

Table 13. Comparative Analysis Results of SMO, RF, and SL on UCI-HAR and WISDM Datasets

Datasets	Device	Classifiers	Accuracy in %	Kappa statistics	Mean Absolute Error	Root Mean Squared Error	Relative Absolute Error in %	Root Relative Squared Error in %
UCI-HAR	Phone	SMO	**98.52**	0.98	0.22	0.31	80.32	83.47
		RF	98	0.97	0.04	0.10	17.03	28.47
		SL	98.47	0.98	0.00	0.06	2.94	16.62
WISDM	Phone A	SMO	60.94	0.58	0.09	0.21	94.99	95.91
		RF	**90.69**	0.90	0.02	0.09	26.11	42.55
		SL	51.48	0.48	0.07	0.18	67.56	80.69
	Phone G	SMO	36.65	0.32	0.10	0.22	95.88	96.88
		RF	**70.26**	0.68	0.68	0.16	59.87	70.91
		SL	35.65	0.31	0.08	0.20	80.80	89.19
	Watch A	SMO	78.47	0.77	0.09	0.21	94.48	95.38
		RF	**89.51**	0.88	0.03	0.10	31.43	46.62
		SL	74.10	0.72	0.04	0.14	41.54	62.55
	Watch G	SMO	61.27	0.58	0.09	0.21	94.95	95.90
		RF	**79.01**	0.77	0.05	0.14	49.35	62.24
		SL	61.44	0.59	0.06	0.17	58.46	74.57

Table 10-12 present the results with 12 fold cross validation. The obtained results indicate that the accuracy decreases as compared to those with 10 fold cross validation. So, from this experimental study we can conclude that for both the datasets 10-fold cross validation works well as compared 8 fold and 12 fold cross validation.

The accuracy percentages of all of the algorithms for both the datasets are presented in Table 13. From the comparative analysis results, it showed that the RF algorithm in all of the cases performs well in predicting human activity as compared to the SMO and SL algorithms in WISDM dataset and SMO showed higher accuracy of 98.52% in UCI-HAR dataset as compared to RF and SL.

5. Comparative Analysis

A comparative analysis results in terms of accuracy is shown in Figure 3 for the SMO, RF, and SL algorithms with both UCI and WISDM datasets. SL algorithm is a good binary classifier which performs better than the others while adopting small datasets. In the current work, we have dealt with the multiclass data to the algorithms, and the experimental results showed that the RF and SMO performed better than the SL. In this paper, WISDM denotes a large dataset and UCI HAR denotes a small dataset. All of the three considered algorithms give good results in UCI HAR. On the contrary, the results obtained with the WISDM dataset were found to have less accuracy in the algorithms. Another observation from our experiment is that compared with using the gyroscope sensor, no matter what kind of device's accelerator sensor is used, it can provide better results.

From the obtained results, it is also found that the weighted average of TP rate is high in SMO and SL classifiers with the UCI HAR dataset. The weighted average of TP rate is high in RF with WISDM by the phone accelerometer. Consequently, SMO and SL with the UCI dataset provide better performance in terms of precision, recall, and F-measure values. However, RF with WISDM gives the best performance in the same conditions.

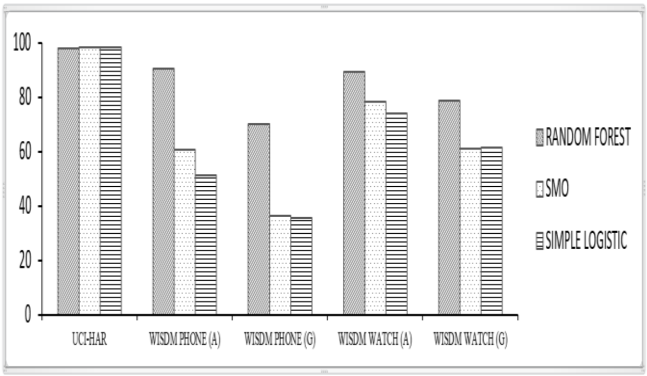

Fig. 3. A comparative analysis in terms of accuracy for SMO, RF, and Simple Logistic algorithms with UCI and WISDM datasets

Using the UCI HAR dataset, the kappa values of the three classifiers are almost same. Nevertheless, the kappa values for the classifiers with the WISDM dataset are different widely. By performing the RF with the WISDM dataset in Phone A, the kappa value is 0.90, higher than the others. For UCI-HAR, although SMO results in a slightly higher accuracy as compared to RF and SL algorithms but the different error percentages such as Mean Absolute Error, Relative Absolute Error, Root Mean Squared Error, and Root Relative Squared Error was found to be less for SL than RF and SMO and is presented in Table 13. However, for WISDM dataset, in all types of sensors RF was found to give better results along with less error percentages as compared to SMO and SL.

A MCC value of 0.98 was obtained in SMO and SL classifiers with the UCI dataset. While the three classifiers were applied in different devices, RF with the WISDM dataset obtained the largest value of MCC in Phone A, but a low value of MCC in the other devices. In Table 5, the three algorithms with the UCI dataset returned the ROC area value of 0.99as the maximum value from the range between 0 and 1 when being compared with these with the WISDM dataset.

In addition, Watch A returns the ROC value of 0.97 as the maximum value in all the cases of WISDM dataset in Table 5. RF and SL with the UCI dataset have the maximum value of the PRC area, 0.99.In addition, RF with the WISDM dataset in Phone A, returned the PRC value of 0.95 in Table 4. According to the experimental results, it is evident that when the number of activities is increasing, the accuracy is decreasing, and when the number of activities is decreasing, the values of accuracy and the other metrics are increasing.

Based on the above reason, the accuracy value is inversely proportional to the number of activities. That is why the classifiers with the UCI dataset consisting of six activities achieve better results than them with the WISDM dataset consisting of eighteen activities.

6. Conclusion

In this work, we have evaluated the performance of the three classification algorithms, namely SMO, RF, and SL in terms of the metrics related to accuracy such as TP rate, FP rate, F-measure, Precision, Recall, ROC area, and PRC area. On the UCI HAR dataset, SMO is a better algorithm than the others; however all of the algorithms provide nearly equal results. On the WISDM dataset, we found that RF is the best algorithm. According to the experimental results, we infer that the adopted classification algorithms are suitable to classify human activities in the domain of HAR. Recognizing human activities is very important and is useful for various applications like elderly care service, healthcare, assisted living, smart home, etc. This study can provide a reference to help researchers make decisions on applying classification algorithms into human activity recognition.

Acknowledgment. This work is funded by CURIE Grant, Department of Science and Technology, Govt. of India under Grant no. DST/CURIE/01/2019(G) dt. 20.05.2019.

References

1. Sarle, W. S. Neural networks and statistical models, *Proceedings of the Nineteenth Annual SAS Users Group International Conference*, Cary, NC: SAS Institute, USA, pp. 1538-1550. (1994).
2. Ann, O. C., and Theng, L. B. Human activity recognition: a review. In *2014 IEEE international conference on the control system, computing, and engineering (ICCSCE 2014)* (pp. 389-393). IEEE. (2014).
3. Oikonomopoulos, A., and Pantic, M. Human activity recognition using hierarchically-mined feature constellations. In *International Symposium on Visual Computing*, Springer, Berlin, Heidelberg, pp. 150-159. (2013).
4. Roshtkhari, M. J., and Levine, M. D. Human activity recognition in videos using a single example. *Image and Vision Computing*, *31*(11), pp. 864-876. (2013).
5. Anguita, D., Ghio, A., Oneto, L., Parra, X., and Reyes-Ortiz, J. L. A public domain dataset for human activity recognition using smartphones. *In ESANN proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning,* Bruges (Belgium), 24-26 April 2013, ISBN 978-2-87419-081-0. (2013).
6. Weiss, G. M., Yoneda, K., and Hayajneh, T. Smartphone and Smart watch-Based Biometrics Using Activities of Daily Living. *IEEE Access*, *7*, pp. 133190-133202. (2019).
7. Ho, T. K. Random decision forests. In *Proceedings of 3rd international conference on document analysis and recognition*, Vol. 1, pp. 278-282. IEEE. (1995).
8. Ho, T. K. The random subspace method for constructing decision forests. *IEEE transactions on pattern analysis and machine intelligence*, *20*(8), 832-844. (1998).
9. Chang, C. C., and Lin, C. J. LIBSVM: A library for support vector machines. *ACM transactions on intelligent systems and technology (TIST)*, *2*(3), pp. 1-27. (2011).
10. Zanni, L., Serafini, T., and Zanghirati, G. Parallel software for training large scale support vector machines on multiprocessor systems. *Journal of Machine Learning Research*, *7*(Jul), pp. 1467-1492. (2006).
11. Tolles, J.,andMeurer, W. J. Logistic regression: relating patient characteristics to outcomes. *Jama*, *316*(5), pp. 533-534. (2016).
12. Singaravelan, S., Arun, R., Arun Shunmugam, D., Ruba Soundar, K., Mayakrishnan, R., and Murugan, D. Analysis of classification algorithms on different datasets. *Review of Innovation and Competitiveness: A Journal of Economic and Social Research*, *4*(2), pp. 41-54. (2018).
13. Ranasinghe, S., Al Machot, F., and Mayr, H. C. A review of applications of activity recognition systems with regard to performance and evaluation. *International Journal of Distributed Sensor Networks*, *12*(8), 1550147716665520. (2016).
14. Krishnan, N. C., Juillard, C., Colbry, D., and Panchanathan, S. Recognition of hand movements using wearable accelerometers. *Journal of Ambient Intelligence and Smart Environments*, *1*(2), pp. 143-155. (2009).

15. Florence, A. M., and Savithri, R. Talent knowledge acquisition using the C4.5 classification algorithm. *International Journal of Emerging Technologies in Computational and Applied Sciences*, pp. 406-410. (2013).
16. Dua, D., and Graff, C. UCI Machine Learning Repository, http://archive.ics.uci.edu/ml, Accessed: 2020-05-01 (2019).
17. Alghobiri, M. A Comparative Analysis of Classification Algorithms on Diverse Datasets, Engineering, Technology & Applied Science Research, 8(2), pp. 2790-2795. (2018).
18. Platt, J. Sequential minimal optimization: A fast algorithm for training support vector machines, Technical Report MSR-TR-98-14(1998).
19. Breiman, L. Random forests. Machine learning, 45(1), 5-32. (2001).
20. Sumner, M., Frank, E., & Hall, M. Speeding up logistic model tree induction. In European conference on principles of data mining and knowledge discovery (pp. 675-683). Springer, Berlin, Heidelberg. (2005).
21. Chang, F., Ge, L., Li, S., Wu, K., & Wang, Y. Self-adaptive spatial-temporal network based on heterogeneous data for air quality prediction. Connection Science, 1-20. (2020).
22. Zhang, S., Hu, Z., Zhu, G., Jin, M., & Li, K. C. Sentiment classification model for Chinese micro-blog comments based on key sentences extraction. Soft Computing. (2020).
23. Nanda, S., Panigrahi, C. R., Pati, B., & Mishra, A. A Novel Approach to Detect Emergency Using Machine Learning. In Progress in Advanced Computing and Intelligent Engineering (pp. 185-192). Springer, Singapore.
24. Verma, R. K., Singh, P., Panigrahi, C. R., & Pati, B. ISS: Intelligent Security System Using Facial Recognition. In Progress in Advanced Computing and Intelligent Engineering (pp. 96-101). Springer, Singapore. (2019).
25. Wang, Q., Zhu, G., Zhang, S., Li, K. C., Chen, X., & Xu, H. Extending emotional lexicon for improving the classification accuracy of Chinese film reviews. Connection Science, 1-20. (2020).
26. Tian, Q., Han, D., Li, K. C., Liu, X., Duan, L., & Castiglione, A. An intrusion detection approach based on improved deep belief network. Applied Intelligence. (2020).
27. Ramasamy, V., Gomathy, B., Obulesu, O., Sarkar, J. L., Panigrahi, C. R., Pati, B., & Majumder, A. Machine Learning Techniques and Tools: Merits and Demerits. New Age Analytics, 23. (2020).
28. Qin, P., Chen, J., Zhang, K., & Chai, R. Convolutional neural networks and hash learning for feature extraction and of fast retrieval of pulmonary nodules. Computer Science and Information Systems, 15(3), 517-531. (2018).
29. Zhang, C., Pan, X., Li, H., Gardiner, A., Sargent, I., Hare, J., & Atkinson, P. M. A hybrid MLP-CNN classifier for very fine resolution remotely sensed image classification. ISPRS Journal of Photogrammetry and Remote Sensing, 140, 133-144. (2018).
30. Cao, L., &Shen, H. Imbalanced data classification based on hybrid resampling and twin support vector machine. Computer Science and Information Systems, 14(3), 579-595. (2017).

31. Albawi, S., Mohammed, T. A., & Al-Zawi, S. Understanding of a convolutional neural network. In 2017 International Conference on Engineering and Technology (ICET) (pp. 1-6). IEEE. (2017).
32 Bhargava, N., Sharma, G., Bhargava, R., & Mathuria, M. Decision tree analysis on j48 algorithm for data mining. Proceedings of International Journal of Advanced Research in Computer Science and Software Engineering, 3(6). (2013).
33 Burns, D. M., and Cari M. W. Personalized Activity Recognition with Deep Triplet Embeddings. arXiv preprint arXiv: 2001.05517 (2020).
34 Ain, A., and Vivek K. Human activity classification in smartphones using accelerometer and gyroscope sensors. IEEE Sensors Journal 18, no. 3 (2017): 1169-1177.
35 Walse, K. H., Rajiv, V. D., and Vilas, M. T. A study of human activity recognition using Ada Boost classifiers on WISDM dataset. The Institute of Integrative Omics and Applied Biotechnology Journal 7, no. 2: 68-76. (2016)
36 Kutlay, M. A., and Sadina, G. Application of machine learning in healthcare: Analysis on mhealth dataset. Southeast Europe Journal of Soft Computing 4, no. 2. (2016)
37 Yang, C., Wenxiang J., and Zhongwen G. Time Series Data Classification Based on Dual Path CNN-RNN Cascade Network. IEEE Access 7: 155304-155312. (2019)
38 Wang, L., Wei, L., Wen, L., and Luc Van Gool. Appearance-and-relation networks for video classification. In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 1430-1439. (2018)
39 Kwon, M. and Choi, S. Recognition of Daily Human Activity Using an Artificial Neural Network and Smartwatch. Wireless Communications and Mobile Computing, Hindawi, https://doi.org/10.1155/2018/2618045. (2018)
40 Biswal, A., Nanda, S., Panigrahi C. R., Cowlessur, S. K., Pati, B. Human Activity Recognition Using Machine Learning: A Review. Progress in Advanced Computing and Intelligent Engineering, Advances in Intelligent Systems and Computing 1299, pp. 323-333. (2021)
41 Mendonça, J., Andrade, E., Endo, P.T., Lima, R. Disaster recovery solutions for IT systems: A Systematic mapping study, Journal of Systems and Software, Volume 149, pp. 511-530. (2019)
42 Feng, F., Li, K-C, Shen, J., Zhou, Q., Yang, X. Using cost-sensitive learning and feature selection algorithms to improve the performance of imbalanced classification. IEEE Access, vol. 8, pp. 69979-69996. (2020)

Suvra Nayak received her M.Sc and M. Phil degree in Computer Science from Rama Devi Women’s University, Bhubaneswar, India in 2021. Her research interests include Machine Learning and Computer Networks.

Chhabi Rani Panigrahi received her Ph.D. in Computer Science and Engineering from IIT Kharagpur, India. She is currently an Assistant Professor in the Department of Computer Science at Rama Devi Women's University, Bhubaneswar, India. Prior to this,

she was working as Assistant Professor in Central University of Rajasthan, India. Her research interests include Software Testing, Mobile Cloud Computing, and Machine Learning. She holds 20 years of teaching and research experience. She has published several international journals, conference papers, and books. She served as chairs and technical program committee member in several conferences of international repute.

Bibudhendu Pati completed his Ph.D. degree from IIT Kharagpur, India. He is currently working as Associate Professor in the Department of Computer Science at Rama Devi Women's University, Bhubaneswar, India. He has around 23 years of experience in teaching and research. His areas of research interests include Wireless Sensor Networks, Cloud Computing, Big Data, Internet of Things, and Advanced Network Technologies. He has got several papers published in reputed journals, conference proceedings, and books of international repute. He has been involved in many professional and editorial activities. He is a Life Member of Indian Society for Technical Education, Computer Society of India and Senior Member of IEEE.

Sarmistha Nanda is a Ph.D. research scholar in the Department of Computer Science at Rama Devi Women's University, Odisha, India. She received her M.Tech. degree in Computer Science from Sambalpur University, Odisha, India. She worked as a JRF in the National Institute of Science Education and Research, Bhubaneswar, India, and Research Associate at Central Rice Research Institute, Cuttack. She also served as a Senior Software Developer in a software firm. Her area of interest is Machine Learning, Cloud Computing, and Programming.

Meng-Yen Hsieh received his MS degree in Computer Science & Information Engineering from National Central University, Taiwan, R.O.C. in 2001, and PhD degree in Engineering Science from National Cheng Kung University, Taiwan, R.O.C. in 2007. He is currently a professor of the Department of Computer Science & Information Engineering, Providence University, Taiwan, R.O.C. His research interests include computer security, machine learning, block chain application, and software engineering. Dr. Hsieh has served on symposium chairs and technical program committees for several international conferences. **http://www1.pu.edu.tw/~mengyen**

Received: December 21, 2020; Accepted: June 17, 2021.