RESEARCH ARTICLE OPEN ACCESS

Comparing Multiple Machine Learning Classification Techniques for Human Activity Recognition (HAR)

Keerthesh Kumar K S*, Ramacharan Avadanam**, Manjunath M***

*(Bangalore, India, email:keerthesh.kumar@gmail.com)

**(Bangalore, India, email :ramcharan1989@gmail.com)

***(Bangalore, India, email:manjug8612@gmail.com)

--************************----------------------------------

Abstract:

Machine learning is subset of artificial intelligence (AI) which aids in training the system to learn by self-using the given data and to predict accurately. Human Activity Recognition (HAR) is highly used technology in many fields namely healthcare, human survey, smart homes and medical research. The data for activity recognition can be captured using different devices such as smart phones, fit bits and various other wearable devices which have built-in sensors such as GPS, accelerometer, barometer, gyroscope etc.; these sensor readings can be taken as inputs then used to predict the human activity using various classification techniques. In this paper, we are using multi-class human activity recognition data to make a comparison between eight different machine learning classification models. The approach we are using to compare these models is to first build the confusion matrix for each model and then use evaluation metrics such as accuracy, precision, recall and F1-Score to measure performance.

Keywords **—Machine Learning, Human Activity Recognition (HAR), Multi-Class Classification,Accuracy, Precision, Recall.**

--************************----------------------------------

I. INTRODUCTION

Smartphone, over the years has become one of the important devices in our life and with its advanced technology, it has become an intelligent tool in our day to day activities [1]. Smartphones these days are not just used for communication but are used for wide variety of interesting applications and this is possible due to the variety of sensors that are installed in the them. Almost all the smartphone these days have tri-axial accelerometer, gyroscope and orientation sensors embedded in them which measure acceleration and rate of rotation around the 3 physical axes i.e. X, Y and Z which can be used to identify different activities done by humans [2]. It is been predicted that smartphones will be able to effectively and discreetly monitor and learn from our actions and help us better decide the future of human behavior [3].

Human activity recognition (HAR) is a significant research area which aims to recognize the actions taken by an individual or a group of subjects by gathering and monitoring the subjects' state in his/her favorable environment [4]. HAR system takes the reading from sensors such as accelerometers, GPS, barometer, gyroscope, light sensors etc. as inputs and predicts individual's motion activity [2].

In this paper, we focus on comparing different classification techniques in predicting the various human activity. The eight different classifiers that were selected and implementedare Logistic Regression, Linear SVC, Kernel SVM (RBF),Decision Tree, Random Forest, Light GBM, XGBoost and K-NN. We have used Accuracy,Precision, Recall and F1-score as evaluation metrics to calculate the efficiency of the classifiers.

II. Related work

Human activity recognition using smartphone is an activeresearch area. Pavel et al [8] have evaluated the performance ofdifferentclassifiers such as KNN, Random Forest, CART, QDA, LDA, NCC with related to three distinct human activity recognition where the fastest was NCC among the algorithms, but recognition accuracy was not enough for the practical purpose. k-NN and RF was most reasonable choice with respect to speed/accuracy ratio.

Ronao et al [9] proposes a two-stage continuous hidden Markov model (CHMM) approach for activity recognition data which was captured from a smartphone. The proposed method consists of first-level CHMMs for coarse classification, second-level CHMMs for fine classification. Random Forests (RF) variable importance measures are exploited to determine the optimal feature subsets for both coarse and fine classification. The CHMM approach showed that overall accuracy of 91.76% can be achieved compared to classification algorithms.

Erhan et al [10] has proposedfew machine learning algorithms namely Support Vector Machines,Decision trees, K-nearest neighbours (KNN). Binary decision tree accuracy was 94.4% when the branching limit was increased to 100. Support Vector machines provided an accuracy of 99.4% using hyper dimensional planes. K-NN provided an efficiency of 97.1% with k value setting to 3.

Akrametal [11] have analyzed the classifiers performance for Multilayer Perceptron, Random forest, LMT, SVM, Simple Logistic and Logit Boost which were available in Weka tool kit. These performances were compared as an individual and combined classifier and validated by K-fold cross validation. With average of probabilitiesthe combined method reached an overall accuracy of 91.15%.

III. Data collection

All the data for the study was gathered from UCI Machine learning repository, were the activities of 30 participants monitored and captured for further processing. The participants belonged to the age group of 19-48 years and were asked to perform six different activities which are listed below:

- Walking
- Walking_Upstairs
- Walking_Downstairs
- Sitting
- Standing
- Laying

The triaxial linear acceleration along with angular velocity was captured by leveraging the accelerometer and gyroscope sensors embedded in the smartphone at 50Hz of sampling rate. The acquired signals were sampled into 2.56 sec fixed width sliding windows with 50% overlap between consecutive windows [1], [6]. The accelerationsignal, which contains both gravitational and body motion components were split by employing the Butterworth low-pass channel into gravity and body acceleration [6]. The gravitational power which had only low frequency elements and a frequency of 0.3 Hz was employed for constant gravity signal [1], [6].

From the above described sampled window standard measures such as mean, correlation, signal magnitude etc. were employed to derive new features to increase the learning performance of the model and a total of 561 features were extracted to explain activity window [7].The list of all those measures that were applied to derive new features are provide in Table I. The dataset was randomly split for evaluating the performance of the models, 70% of the data was used to training while 30% were used for testing the models.

TABLE 1
LIST OF MEASURE FOR PROCESSING FEATURE VECTORS

SL No	Function	Description
1	mean	Mean value
2	mad	Median absolute value
3	min	Smallest value in array
4	energy	Average sum of the squares
5	entropy	Signal Entropy
6	correlation	Correlation coefficient
7	meanFreq	Frequency signal weighted average
8	kurtosis	Frequency signal Kurtosis
9	angle	Angle between two vectors
10	std	Standard deviation
11	max	Largest values in array
12	sma	Signal magnitude area
13	iqr	Interquartile range
14	arCoeff	Autorregresion coefficients
15	maxFreqInd	Largest frequency component
16	skewness	Frequency signal Skewness
17	energyBand	Energy of a frequency interval

IV. MODEL SELECTION

In this study we use Logistic regression, decision tree, random forest, SVM, Linear SVC, Light GBM, XGBoost and KNN were used on the training data set for evaluation. We train the models using the training set and we choose the best hyper parameters using the 3-fold cross validation. Once we select the best hyperparameters we apply the same to test dataset to perform prediction.

A. Decision Tree

Decision tree is one of the important and one of the most popular models used for multi class classification. Decision Tree is very beneficial when we want to classify data into a class of homogenous nodes. It is a supervised learning technique where the features are spilt into internal nodes further split into leaves. Trees with low cost function are considered as the best split, in some cases pruning will be done to avoid overfitting by removing features with low importance[13].

B. Random Forest

Random Forest is a type of ensemble model, it constitutes of multiple decision tree models which are spitted into multiple internodes and leaves. These decision trees are uncorrelated to each other. A voting is carried out to identify the best model to do the prediction across these decision trees and to check if the predicted value is same from all the derived decision trees. HAR data will utilize the random forest approach to classify the activities into right set of classes [1].

C. Logistic Regression

Logistic Regression is one of most popular supervised classification algorithms.Logistic Regression is derived from the logit function which uses the S-shaped curve that takes the real value number and maps it to 0 or 1. A threshold must be setup to decide whether the new value falls intobinary category. In case of the multiple classes, the model uses one versus rest principle to make prediction. Mean squared error is the cost function considered in Logistic Regression [14]

D. K-NN regression

K-NN regression assume that all the similar things exist in the proximity of each other. KNN algorithm hinges on this assumption being true enough for algorithm to be useful. It calculates the distance between points, Euclidean distance is one of the most used method to calculate the distance between points. The best K value is obtained after multiple iterative with different K values [2].

E. XGboost

XGboost is a decision tree-based ensemble learning algorithm that uses gradient boosting framework. Gradient boosting employs gradient algorithm to minimize errors in sequential models. XGboost is also an open source library with high performing decision tree-based models. In boosting methods, a single model is trained on the different on the dataset but with different parameters. Whereas, XGboost considers iterative approach, it trains the model in succession, the new model is trained to correct the errors of the previous model. This method continues until no further residuals are left from previous model [15].

F. Light GBM

Light GBM is another decision tree-based algorithm with Gradient boosting framework, on the other hand in Light GBM the trees grow vertically meaning growing by leaf wise. This model uses histogram-based method to train the model as a result improves the efficiency of the model.

G. Support Vector Machine

Support Vector Machine is used to find a hyperplane that classifies the data points into their respective categories. Linear SVC is one of the effective algorithms, data points are mapped from data space to high dimensional space using a Gaussian Kernel[16]. As the width parameter of Gaussian kernel decreases the number of clusters increases. Radial based function is a kernel utilized method in SVMs which is features the data points into high space dimensional using Kernels.

V. MODEL EVALUATION

Evaluation of a model is a crucial part of any model development procedure. The model evaluation process not just helps us in choosing the best model for our data but also helps us to get a good understanding of the chosen model works with unseen or future data [12]. In this paper we have used Accuracy, Precision, Recall and F1-Score as the evaluation metrics to compare the performance of each models, these metrics are calculated using True Positives (TP), True Negative (TN), False Positive (FP) and False Negative (FN) given in Table II.

True Positive: These are values which were predicted as “Yes” while the actual values were “Yes”.
True Negative: These are values which were predicted as “No” while the actual values were “No”.
False Positive: These are values which were predicted as “Yes” while actual values were “No”.
False Negative: These are values which were predicted as “No” while actual values were “Yes”.

TABLE II
TWO CLASS CLASSIFICATION PROBLEM CONFUSION MATRIX

	Predicted Yes	**Predicted No**
Actual Yes	TP	FN
Actual No	FP	TN

Accuracy:In statistics, accuracy means a ratio of total number of correctly predicted values to total observations [12]. We might think that higher the accuracy of the model the better is its prediction capacity, however this is true when the data is symmetric.

$$Accuracy = \frac{(TP + TN)}{(TP + TN + FP + FN)} \qquad (1)$$

Precision:The ratio of total true positive i.e. correctly predicted positive values to the total predicted positive values. Basically, what it means is that of the total predicted values how many of them are relevant.

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

Recall: The ratio of total true positive to all the observations in a specific class i.e. it is a ratio of the positive cases that are correctly classified as positive.

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

F1-score: F1-Score is nothing but a weighted average of Precision and Recall. It is slightly difficult to understand as compared to accuracy, however it is generally more beneficial than accuracy particularly when class is imbalanced.

$$F1 - Score = 2 * \frac{Precision * Recall}{(Precision + Recall)} \quad (4)$$

VI. RESULTS

A comparison of the evaluation metrics such as Accuracy, Precision, Recall and F1-Score for the various classification models on the test data set is provided in Table III. When we look at these evaluation metrics, we see that Logistic Regression, Linear SVC, and Kernel SVM models seems to perform significantly better than other models with an accuracy of 96.0%, 96.6% and 96.3% respectively while Decision Tree models is the worst performing model with accuracy of 87.3%.

TABLE III
COMPARISON OF MODEL EVALUATION METRICS

Models	**Best Estimator**	**Accuracy**	**Precision**	**Recall**	**F1-Score**
Logistic Regression	C : 20 Penalty : L2	96.0%	96.3%	96.0%	96.0%
Linear SVC	C :1 Penalty : L2	96.6%	97.0%	96.5%	96.7%
Kernel SVM (RBF)	C : 16 Gamma : 0.00781	96.3%	96.5%	96.3%	96.2%
Decision Tree	Max_depth : 9	87.3%	87.5%	87.0%	86.8%
Random Forest	Max_depth : 9 N_Estimators : 110	92.1%	92.2%	91.8%	92.2%
Light GBM	Max_depth : 1 N_Estimators : 110	89.8%	90.2%	89.7%	89.5%
XG Boost	Max_depth : 5 N_Estimators : 110	93.6%	93.7%	93.5%	93.5%
KNN	N_Neighbors : 15	90.5%	90.8%	89.8%	90.2%

When we look at accuracy as a standalone evaluation metrics the difference between Logistic Regression, Linear SVC, and Kernel SVM seems to be not that significant. However, Linear SVC seems to do marginally better than the other models when we compare Precision, Recall and F1-score with a score of 97.0%, 96.5% and 96.7% respectively.

VII. CONCLUSIONS

In this study, we used the HAR dataset to compare the performance of classification various models in order to provide evidence on the general inclination of the data set as well as help our fellow researchers select best model for solving classification model. Based on our finding we suggest that the researcher should further explore the data and include new features which might assist the model in learning better. In future, researchers should avoid drawing conclusions from only one model, instead explore multiple models before putting forth their findings.

REFERENCES

[1] S. Roobiniand.J. Fenila Naomi, "Smartphone Sensor Based Human Activity Recognition using Deep Learning Models", International Journal of Recent Technology and Engineering (IJRTE), vol. 8, no. 1, pp 2740-2248, May. 2019.

[2] Usharani J and Dr. Usha Sakthivel, "Human Activity Recognition using Android Smartphone", 1st International Conference on Innovations in Computing & Networking (ICICN16), May. 2016, pp 191-197.

[3] Cook, D. J., Das, S. K., "How smart are our environments? An updated look at the state of the art", Pervasive and Mobile Computing, Mar 2007, pp 53-73.

[4] Anguita, D, Ghio, A, Oneto, L, Parra, X.; Reyes-Ortiz, J.L. "Energy Efficient Smartphone-Based Activity Recognition using Fixed-Point Arithmetic". Journal of Universal Computer Science, vol. 19, no. 9, pp 1295–1314, May 203.

[5] Mevlut T, Imran K, Turhan K, Kazim O, "Comparing classification techniques for predicting essential hypertension", Elsevier Expert Systems with Applications, vol. 20, pp 583-588, April 2005.

[6] Davide A, Alessandro G, Luca O, Xavier P and Jorge L. R. O., "A Public Domain Dataset for Human Activity Recognition Using Smartphones", Computational Intelligence and Machine Learning, Bruges, Belgium, ESANN 2013.

[7] J.Y. Yang, J.S. Wang, and Y.P. Chen, "Using acceleration measurements for activity recognition: An effective learning algorithm for constructing neural classifiers", Pattern recognition letters, vol. 29, no. 16, pp 2213-2220, Dec 2008.

[8] Pavel Dohnalek, Petr Gajdoš and TomášPeterek, "Human activity recognition: classifier performance evaluation on multiple datasets", Journal of Vibroengineering, Vol. 16, Issue 3, 2014.

[9] C. A. Ronao, S. B. Cho, "Human activity recognition using smart-phone sensors with two-stage continuous hidden markov models", 2014 10th International Conference on Natural Computation (ICNC), pp. 681-686, 2014.

[10] Erhan BÜLBÜL, AydınÇetin and İbrahim Alper DOĞRU, "Human Activity Recognition Using Smartphones", IEEE, 978-1-5386-4184, 2018.

[11] AkramBayat∗, Marc Pomplun, Duc A. Tran, "A Study on Human Activity Recognition Using Accelerometer Data from Smartphones", Department of Computer Science, University of Massachusetts, Boston, 100 Morrissey Blvd Boston, MA 02125, USA, Elsevier Procedia Computer Science, Vol 34, 450 -457 ,2014.

[12] Yunwei Z, "Binary Classification for Teacher Donor's Project", International Conference on Education, Economics and Social Science, vol. 223, pp 157-160, 2018

[13] Veralia Gabriela Sánchez, Nils-Olav Skeie, Decision Trees for Human Activity Recognition Modelling in Smart House Environments, 2018, SNE 28(4), 2018, 177-184, DOI: 10.11128/sne.28.tn.10447.

[14] [14] A Lawi, F Aziz and S L Wungo, Increasing accuracy of classification physical activity based on smartphone using ensemble logistic regression with boosting method, Journal of Physics: Conference Series 1341 (2019) 042002, IOP Publishing, doi:10.1088/1742-6596/1341/4/042002

[15] Ole M. Brastein, Roland Olsson, et.all.,Human Activity Recognition by machine learning methods,2017: NorskInformatikkonferanse, 2017-11-26

[16] Xizhe Yin, Weiming Shen, et.all. Human Activity Detection Based on Multiple Smart Phone Sensors and Machine Learning Algorithms, Conference Paper · May 2015 DOI: 10.1109/CSCWD.2015.7231023.

[17] Trevor Hastie, Robert Tibshirani, Jerome Friedman, The Elements of Statistical Learning Data Mining, InferenceandPrediction, Springer, 2nded.

