Proceedings of 2020 IEEE 10th International Conference on Intelligent Systems

Comparative Study of Classifiers on Human Activity Recognition by Different Feature Engineering Techniques

Mahbuba Tasmin
Department of Electrical & Computer Engineering
North South University
Dhaka, Bangladesh
mahbuba.tasmin@northsouth.edu

Taoseef Ishtiak
Department of Electrical & Computer Engineering
North South University
Dhaka, Bangladesh
taoseef.ishtiak@northsouth.edu

Sharif Uddin Ruman
Department of Electrical & Computer Engineering
North South University
Dhaka, Bangladesh
sharif.ruman@northsouth.edu

Arif Ur Rahaman Chowdhury Suhan
Department of Electrical & Computer Engineering
North South University
Dhaka, Bangladesh
arif.suhan@northsouth.edu

N.M. Shihab Islam
Department of Electrical & Computer Engineering
North South University
Dhaka, Bangladesh
shihab.islam@northsouth.edu

Sifat Jahan
Department of Electrical & Computer Engineering
North South University
Dhaka, Bangladesh
sifat.jahan@northsouth.edu

Sajid Ahmed
Department of Electrical & Computer Engineering
North South University
Dhaka, Bangladesh
sajid.ahmed1@northsouth.edu

Md. Shahnawaz Zulminan
Department of Electrical & Computer Engineering
North South University
Dhaka, Bangladesh
shahnawaz.zulminan@northsouth.edu

Abdur Raufus Saleheen
Department of Electrical & Computer Engineering
North South University
Dhaka, Bangladesh
abdurraufus.saleheen@northsouth.edu

Rashedur M. Rahman
Department of Electrical & Computer Engineering
North South University
Dhaka, Bangladesh
rashedur.rahman@northsouth.edu

***Abstract*—The paper presents a comparative discussion of classification approaches for human activity recognition tasks based on the feature sets through extensive feature selection techniques. The original dataset on Human Activity Recognition from Continuous Ambient Sensor is collected from UCI machine learning repository. Amongst the 12 activities mentioned in the dataset, five specific activities (Watching TV, Reading, Talking over Phone, Cooking and Eating) have been selected for the purpose of this research. The scraped dataset is analyzed through four feature selection methods for extracting important features upon statistical significance of features and node impurity. From the actual dataset with 37 attributes, the feature selection methodologies give four distinct features sets. Later on, Principal Component Analysis is applied on the five feature sets including the original scraped dataset to reduce feature space and five principal components are selected to cover more than 90% data variance of the feature sets. Based on the 37 features present in the actual dataset and obtained sets of important features, performance of five classifier models (K Nearest Neighbors, Decision Tree, Random Forest, Gaussian Naïve Bayes and MLP classifier using Back propagation) are evaluated. The selection of feature set based on different approaches of feature importance generates difference in outputs for each feature set on each classifier. The result shows that Multi-layer Perceptron using Back propagation algorithm achieves better accuracy on human activity recognition on the five feature sets. The research findings highlight the necessity of data preprocessing and significant feature selection for getting better accuracy score for noisy time-series data of HAR activity.**

Keywords-Human Activity Recognition, Time Series Data, Activity Classification, Feature Engineering, Sensor

I. INTRODUCTION

The advancement of Internet of Things (IoT) has paved the way for smart living in home spaces using intelligent system installed in the framework of indoor spaces. [1]. Smart intelligent appliances have been developed for convenient living style, which is gradually proceeding towards assisted living through interacting system space. Ambient Assisted Living (AAL) [2] emerged with the aim of easing the life of independent elderly citizens in indoor space through the use of smart technologies at home. AAL focuses on health care monitoring and user interaction [3], which requires the necessity of human activity recognition from Activities of Daily Life (ADL). Better performance of an AAL system installed in a home depends on the accuracy of the system to interact with the user and to diagnose the activities to take actions accordingly. AAL infrastructures highly depend on wireless sensor networks [5] placed at home to collect sensor data streams of human

978-1-7281-5456-5/20/$31.00 ©2020 IEEE

activities. In general, activity recognition is a context-aware system [6, 7], aimed to sense the surrounding activities and execute consequent features.

Sensor data for human activity recognition works are collected from wearable devices, smartphones and indoor infrastructure of wireless sensors [8]. The sequential or time-series datasets collected from the above-mentioned systems are complex [9, 10] in terms of interpretation compared to computer vision-based activity recognition. Time-series data for human activity recognition is checked for change detection [11] or activity transition through calculating statistical metrics (e.g. Mean and Covariance). Activity segmentation improves the performance as it provides the information of activity transitions, beginning and ending times [12]. The image/video datasets of human activity are easier to label whereas the sensor collected raw dataset requires intensive feature engineering [13, 14] to achieve an optimum-cost computational algorithm with highest accuracy. The dataset is usually large and requires significant feature selection [15] to discard insignificant attributes-instances and to produce a concentrated important feature set. The feature selection approaches are based on statistical scoring on a threshold and node impurity calculation through Gini index. The common approaches include decision tree implementation for scoring of features and forest-based categorization of features. Feature space reduction concentrates the dataset in execution reducing the dimension of dataset. Feature engineering of time-series data collected from sensors [16] is a necessity to achieve better recognition output through any advanced machine learning model.

This research work is motivated to classify five distinct activities (Watching TV, Reading, Taking on Phone, Cooking, and Eating) from the dataset of 12 pre-defined activities including unlabeled activity namely "other activity". The dataset has been acquired from the UCI Machine Learning Repository naming "Human Activity Recognition from Continuous Ambient Sensor Data Dataset" collected by the Washington State University [11]. The motivation is to precisely classify the activities while reducing the number of parameters and selecting important features from original dataset based on statistical approach. This originates from the idea to allow human activity recognition with a simplified model for saving computation power so that real-life applications upon such model will be lighter. The sensors' signals is preprocessed in original dataset, among those, features with statistically significant values have been selected for training, feature set space has been reduced for less computational load and finally five different classifier models have been employed to evaluate activity recognition accuracy and a comparative study of performance is reported towards the end.

The novelty of this proposed approach is to introduce the application of tree-based feature engineering and feature selection methods for improving training accuracy of human activity recognition. Standard baseline classifiers are used to evaluate the performance differences achieved by the preceding feature engineering stage. From the initial training upon the original set of 37 features, the best training accuracy is achieved 76.7%, obtained by Multi-layer perceptron model; whereas, after selecting the important features, the highest accuracy score is achieved by the same model is 78.5%, based on the important features extracted by the Extra tree classifier method. Although the accuracy score is not very high due to limitation of the classifier models, the proposed research aims to indicate the potentiality that features selection can play, in the field of human activity recognition classification with data mining and machine learning algorithms. It is believed that, with more advanced machine learning and data mining algorithms with extensive feature engineering, the accuracy can be improved further. The feature engineering process presented in this paper will lead to executing advanced classifier models faster with less resource requirement and obtain better accuracy. The key finding of this research focuses on the significance of feature engineering for improving human activity recognition accuracy on different feature sets. The results show variance in accuracy depending on the four feature sets through five classifier models.

The major contributions of the proposed research are:

1. Feature Selection: Statistically important feature selection from the five activities- based dataset through four feature selection approaches.

2. Dimensionality Reduction: The acquired feature sets' space reduction through Principal Component Analysis (Five PCs) to prepare for the classifier computation.

3. Classifier Performance: Five classifier models are evaluated on the five sets of features for obtaining best accuracy on activity recognition.

The rest of the paper is organized as follows. Section II presents the related works on human activity recognition. Section III presents data source and Section IV presents methodology where data preprocessing, feature selection approaches, and classifier models are discussed. Section V reports performance evaluation of the five classifier models on the feature sets. Section VI concludes the paper and gives direction of future works.

II. Related Work

The research field of activity recognition considers the combination of embedded sensors, different environmental setups and algorithms to detect activity points. Probabilistic graph-based Markov models, conditional random fields, Bayesian network [16] are some of the state-of-the-art classification models for detecting activity from times-series data.

Distinct activities like Walking, Running, Standing, Sitting, Climbing Stairs and Falling) are classified in [17] using accelerometer placed on the body. Recently smartphones with embedded motion detector and orientation sensors (Accelerometer and Gyroscope) are used as wearable device to recognize gesture and motion patterns [18]. Improvement in performance, increased accuracy and better results can be attained by the Deep Learning based approaches from raw sensor inputs. In indoor HAR system setup, large ranges of activity are observed through embedded sensors at key location of activities. Environment sensors such as motion detector, light sensor, temperature and pressure sensors etc. are used to record stream of sensor data of activities in [11].

In realistic activity recognition tasks, the recognizing activities are performed with interleaved activities, embedded errors and concurrent activities are performed by multiple individuals in the setup [19]. Detecting activities in free movement setup, where the residents perform usual daily routines in a smart home environment [20]. These recorded datasets require manual labeling to segment and analyze the data.

Dedicated HAR architectures recognize sequential and concurrent human activities using multiple sensor data at a time. Two key approaches ar7 followed in HAR: “Data-Driven” and “Knowledge-Driven” technique [22]. Naïve Bayes (NB) classifiers, Decision Trees, Hidden Markov Models, Bayesian Networks and Support Vector Machine (SVM) classifier had been used as the Data-driven method in [22]. Existing works performed with data-driven technique utilize supervised approach using manually labeled data for training. The unsupervised approaches achieve low performance in comparison with the supervised approach in indoor home environment. Activities are classified with the prior knowledge of pre-recorded data of surroundings. Data-driven techniques are useful for detecting basic distinctive activities; on the other hand unsupervised approach is suitable for creating probabilistic models with good accuracy.

Feature selection and data dimensionality reduction have been the goal of many previous research works. As high dimensional data becomes a hurdle in classification and other data mining tasks, the authors in [29] have suggested using gain ration and correlation to reduce the dimensionality. The authors determined the split and feature selection with C4.5 tree along with gain ratio and genetic algorithms helped them to determine the correlation. For faster computational support, the authors chose filter method over wrapper method in the proposed research. They argued that although wrapper methods are better than filter methods in the domain of data reduction, wrapper methods face more computational and time complexity comparing to the filter methods. PSO-based feature selection and tree-based classifiers have been used in [30] in intrusion detection system.

III. Data Source

The primary dataset of the project has been collected from UCI Machine Learning Repository [11], Human Activity Recognition from Continuous Ambient Sensor Dataset, published on September 20, 2019. This dataset recorded multiple sensor data placed at volunteer resident houses where the residents performed their daily activities with no direct contact with the data collector infrastructure. Ambient PIR motion sensors, door/temperature sensors, and Light Switch sensors are used to record activity data as event stream the sensors are located in different corners of resident houses to record event data. Table I describes the key features.

The original dataset is built under the lead of Diane J. Cook from School of Electrical Engineering and Computer Science at Washington State University, and the other creators are Aaron S. Crandall, and Brian L. Thomas. [10]. Figure 1 shows the layout of sensor placement in the indoor environment for data collection in their proposed system.

TABLE I. Key Features of the Scraped Dataset

Data Set Characteristics	Multivariate, Sequential, Time Series	**Number of Instances**	4475631
Attribute Characteristics	Integer, Real	**Number of Attributes**	37
Associated Tasks	Feature Engineering, Classification	**Missing Values**	Yes

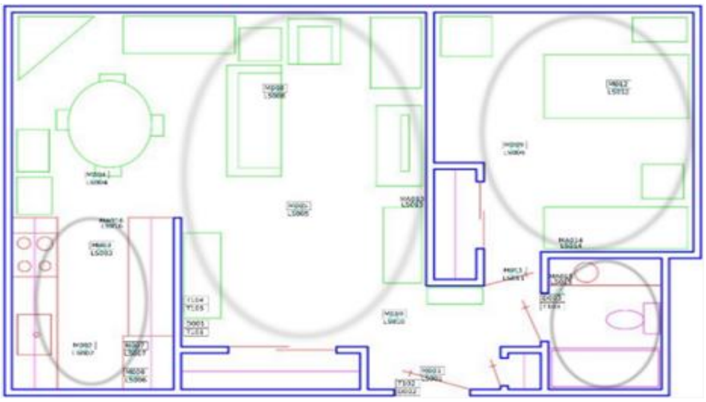


Figure 1. Sensor Layout of One of the Volunteer Resident House

IV. Methodology

This research work focuses on the comparative performance evaluation of the five classifier models on the acquired feature sets through extensive feature engineering of five activities from “Human Activity Recognition from Continuous Ambient Sensor Data Dataset” [11]. Four variations of tree-based algorithms have been chosen to perform the feature engineering. Genetic algorithms and decision trees have been applied successfully to identify network intrusion detection [31]. The aim of the research team is to detect activity based on the 37 attributes that have been enlisted in the aforementioned dataset. Based on the experimental results which are discussed in section V of this paper, training 37 attributes do not give satisfactory results upon five basic classifiers.

The novelty of the proposed approach stands on such a ground that, when such large numbers of features are hard to train well, it is possible to reduce the number of features with feature engineering and improve the performance of training through any classifier. The basic classifiers are thus experimented to defend this proposition. Decision tree-based algorithms and other wrapper function methods are well established approaches for many data mining projects [32], bioinformatics [33]. But decision tree-based algorithms have not been explicitly used to reduce number of insignificant features and feature engineering to improve model’s performance. The research team aims to improve the explored models’ performance with the proposed tree-based algorithms. Two variants of feature engineering approach have been introduced in this paper, first technique focuses on feature engineering through machine learning models based on tree-based classifier and random forest classifier. In this way, the models do the necessary mathematical calculations on the background which are described in the following subsection and generates own set of important features based on the score. Another approach is-feature rank generation for all the attributes. In this approach, the Gini index and Entropy calculation creates a feature rank for all the attributes. Extra tree classifier and Random forest classifier have been applied in this case to rank all the features. This is a more deterministic approach, because the research team will choose from different set of attributes to obtain the highest accuracy.

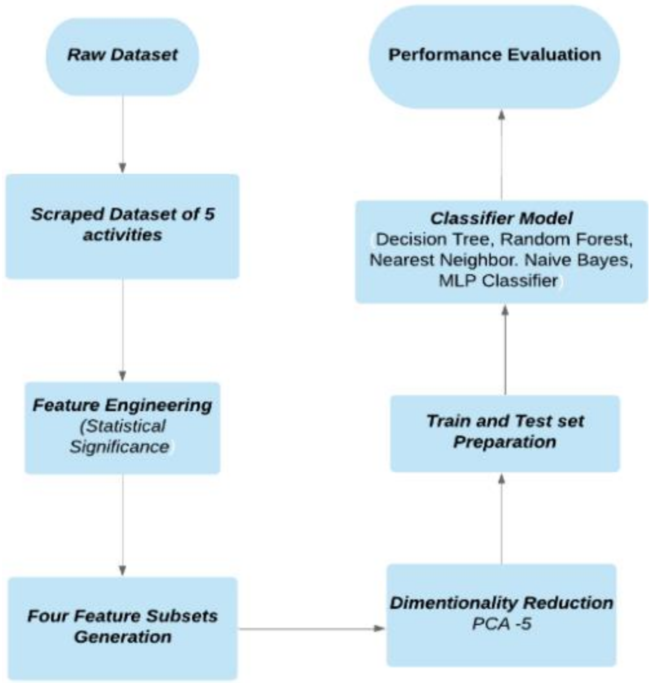


Figure 2. Workflow of the proposed Approach

The proposed research methodology has explored both of these approaches and summarized the training results. The research team has scraped the dataset for this research from the UCI dataset, for the five selected activities (Watching TV, Reading, Talking over Phone, Cooking, and Eating). The overview of the selected dataset are presented in Table I. The proposed work is primarily divided into three major segments- Feature Engineering, Classifier Model Training, and Performance Evaluation of Classifier models. Figure 2 presents the basic workflow of the proposed research work.

A. *Feature Engineering*

The attributes of the selected five activities are analyzed to compute sstatistical significance scores with view to reducing and eliminating the number of features that do not contribute to training. Four approaches have been considered in feature selection and feature importance paradigm. All the approaches are run on scikit-learn [21].

Feature selection approach not only reduces the training time and computational cost but also reduces the variance of the model to avoid over fitting. This section describes the feature selection techniques applied in the proposed research. Table II lists the set of significant column attributes found by the four feature engineering approaches. The selected features through different approaches will be used to train a corresponding model.

1) Tree-based Feature Selection: The tree-based estimators are used to calculate the statistical significance of features and to discard the irrelevant features [24]. Node impurities measure the importance of features in decision trees.

2) Feature Selection with Random Forest: Random forest classifier uses the tree-based strategies [25] to rank the features for improving purity of the node. The interpretability of this approach is very efficient to derive the importance of each attribute in the dataset based on the tree decision. Feature selection is done by embedded methods in this approach. Such embedded methods are scalable across any dataset for their high accuracy, better generalization and efficient interpretability; including built-in functions for feature selection. Random forest executes a random number of feature selection against each tree. To make the approach less likely to over fit, the trees' chances of correlation gets decreased as every tree does not observe every variable of the whole dataset. Impurity measure is executed through information gain or entropy in this approach. Across each tree the average impurity decrease determines the final importance of the variable. Feature importance calculates the score for each feature in a dataset through the implementation of forests of tree-based approach. The Extra Tree classifier and Random Forest classifier have extracted the top 21 significant features out of 37 original attributes of the research dataset. Figure 3 and 4 present the significant-feature ranks with bar charts. These significant features along with the features stated in Table II *have also been considered for the final training of the classifiers. The four different approaches of feature selection and feature importance implemented by the research team have given four feature sets against each technique.* Table II *shows the feature sets against each tree-based feature selection and feature importance approach.*

3) Extra Tree Classifier: The ensemble learning approach of Extremely Randomized Trees Classifiers [26] performs the aggregation of de-correlated decision trees' results in a forest for classification. The decision trees construction differs in this aspect comparing to the construction of Random Forest Classifier. Each tree in the forest gets a random K sample of feature-set. Each decision tree selects the best feature to split the data on mathemtical basis calculated by the Gini Index [28] and entropy calculation. The multiple sampling of features aggreagates the multiple de-correlated decision trees. Extra tree classifier performs feature selection in descending order based on the Gini importance and entropy calculation with respect to each feature. The research team in this approch has selected the top 21 features based on the obtained score in the range [0.153283, 0.020804] as demonstrated in Figure 3. Gini importance is calculated by the equation given by:

$$\boldsymbol{Gini(S) = 1 - \sum_{i=1}^{c} p_i^2}$$

The Gini Importance, also known as the Mean Decrease in Impuritiy (MDI) computes the sum over the quanitity of splits for finding the importance of each feature. It includes the feature with proportionally the sample splits. Entropy calculation measures the rate of impurity in the recursively produced set of features by decision tree. The formula for entropy calculation is given by:

$$\boldsymbol{Entropy(S) = \sum_{i=1}^{c} -p_i log_2(p_i)}$$

4) Random Forest Classifier: The "random forest" concept brought in this classifier works with several decision trees [27]. Every node in the tree splits the dataset in to sub-set conditioning a single feature. It ensures the similar response values come to the same set. Impurity, here chooses the locally optimal condition by Gini impurity [28] or by the information gain or entropy. While training, thus, the computation of decrease in weighted impurity is measured. The average impurity decrease is measured with this approach and the features' importance is ranked thereby. The importance scores demonstrated in Figure 4, are obtained in the range [0.166521, 0.014055] with this approach.

B. Final Feature Set Generation

After the feature engineering completion, at this step, unsupervised dimensionality reduction is introduced to reduce the feature space. After this step, the feature set of each feature selection approaches are feed into the classifier models. This step of the methodology used Principal Component Analysis and five principal components (PCA = 5) are selected after calculation of expected variance ratio (99%) to project the reduced linear subspace. The dataset is split into train and test set, and standardized. The evaluation is performed by a K (K=3) nearest neighbor (KNN) classifier on the 5-dimensional projected points. Figure 5 presents the PCA-5 projected subspace on the original scraped dataset. The test accuracy 91% is evaluated against 3-nearest neighbors classifier. While feature engineering selects the significant attributes, feature set generation through PCA-reduction afterwards contracts the instances of the dataset. Since the classifier models in the following sections are standard classifiers without advanced neural network architecture, the reduced feature set works faster with less resource requirement.

C. Classifier Models

Five classifier models have been evaluated against the five feature sets (Four feature sets through feature engineering and one original feature set). The variance of the classifiers against the variant feature sets shows the significance of feature engineering and a good classification model for achieving higher accuracy on human activity recognition. The models are –K Nearest Neighbors, Decision Tree, Random Forest, Gaussian Naïve Bayes and MLP classifier using Back propagation. The feed forward neural network of MLP Classifier achieves the highest accuracy in all the feature sets. All of the models are based on scikit-learn [23]. A brief description of the applied algorithms are introduced below.

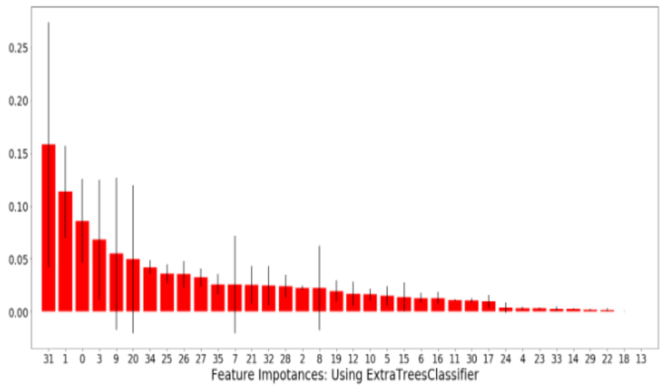


Figure 3. Extra Tree Classifier based Feature Selection Ranking: based on Importance Score

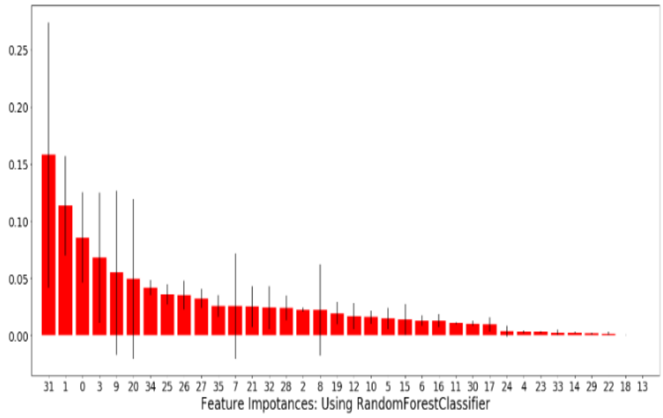


Figure 4. Random Forest Classifier based Feature Selection Ranking: based on Importance Score

1) K Nearest Neighbors: KNN works on the principle of least distance between similar objects, whereby it assumes that similar classes/features stay nearer. The algorithm predicts on the majority support of nearest neighbors of each point.

2) Decision Tree: Decision tree is a supervised decision support structure consisting of nodes and leaf. The root to leaf path represents the classification rules. The leaves present the labels, internal nodes are features and branches present the outcome of the test.

3) Random Forest: Random forest approach is an ensemble method of decision trees. It fits decision trees on random sub-sampled dataset and uses averaging from prediction of all the trees to improve the prediction score.

4) Gaussian Naïve Bayes: This classifier is the Gaussian distribution implementation of Naïve Bayes approach. The supervised approach applies Bayes' theorem with the "naïve" assumption of pairwise independence of features. It calculates the probability of an instance belonging to a certain class through mean and standard deviation calculation.

5) MLP Classifier using Backpropagation: Multi-Layer Perceptron or Feed Forward Neural Network is one of the basic deep learning models. This supervised learning model follows repeated execution of forward pass and backward pass. In the forward pass, the signal traverses through the input and hidden *layers to the output layer. Output findings are evaluated against the ground-truth labels to calculate the loss function. In the backward pass, the error term is back propagated and weights are updated through gradient calculation until convergence state is reached.*

TABLE II. SELECTED FEATURE SETS THROUGH DIFFERENT FEATURE SELECTION APPROACH

Original Set of Features	**Model Obtained Selected Features with Tree-based Feature Selection**	**Model Obtained Selected Features with Random Forest Classifier**	**Top 21 Extracted Features by Extra Tree Classifiers**	**Top 21 Extracted Features by Random Forest Classifier**
lastSensorEventHours	√	√	√	√
lastSensorEventSeconds	√	√	√	√
lastSensorDayOfWeek	√		√	√
windowDuration		√	√	√
timeSinceLastSensorEvent				
prevDominantSensor1	√		√	
prevDominantSensor2			√	
lastSensorID			√	√
lastSensorLocation	√		√	√
lastMotionLocation	√	√	√	√
complexity			√	√
activityChange				
areaTransitions				√
numDistinctSensors				
sensorCount-Bathroom				
sensorCount-Bedroom			√	√
sensorCount-Chair				
sensorCount-DiningRoom				
sensorCount-Hall				
sensorCount-Ignore			√	√
sensorCount-Kitchen	√	√	√	√
sensorCount-LivingRoom	√		√	√
sensorCount-Office		√		
sensorCount-OutsideDoor				
sensorCount-WorkArea				
sensorElTime-Bathroom			√	√
sensorElTime-Bedroom			√	√
sensorElTime-Chair			√	√
sensorElTime-DiningRoom			√	√
sensorElTime-Hall				
sensorElTime-Ignore				
sensorElTime-Kitchen			√	√
sensorElTime-LivingRoom				√
sensorElTime-Office				
sensorElTime-OutsideDoor			√	√
sensorElTime-WorkArea			√	√

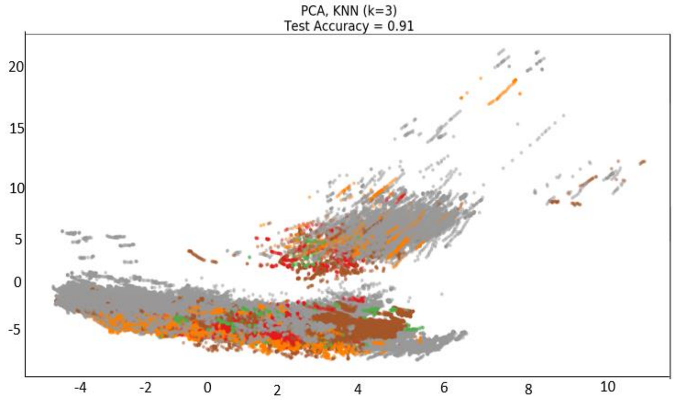


Figure 5. 91% Variance Test Accuracy by PCA -5 in feature 3et

V. PERFORMANCE EVALUATION

In this section, the accuracy of the five classifier models on the five feature sets are evaluated and compared against each other. The best model output is presented with evaluation metric scores.

A. Evaluation Metric Calculation

The performance of the five-classifier model on 5 datasets is evaluated through four key metrics of accuracy: precision, recall, f1-score and support. The confusion matrix and evaluation metrics are presented here on the basis of the different datasets and the result of classification models on those datasets. The activities are coded into numerical values in here ('Cook':0,'Eat':1,'Phone':2,'Read':3,'Watch_TV':4).

Accuracy gives the sum of correct classifications to the total number of instances. Recall presents the proportion of actual positive classes those were identified in proportion to all samples in the actual class. F1-score is a performance metric that measures the weighted harmonic mean of precision and recall. It is used to evaluate the classification accuracy of an algorithm. TP i.e. true positive is the category of positive attributes correctly classified as positive attributes, TN i.e. true negative is the set of negative samples identified as negative samples. FP i.e. false positive is the category of negative attributes classified as positive attributes. And FN i.e. false negative are the positive samples being classified as negative samples. This performance metric presents the proportion of positive attributes those were classified correctly.

Feed forward neural network showed highest 78.5% accuracy on the Extra-tree classifier-based feature set. Figure 6 presents the evaluation metric scores of the MLP classifier on

Extra Tree Classifier-based feature set. The activities in the confusion matrix have been enlisted as: ('Cook':0,'Eat':1,'Phone':2,'Read':3,'Watch_TV':4). The obtained result from f1-score activities reached higher accuracy for the following activities: 'Cook':0,'Read':3,'Watch_TV':4. The precision and recall metrics also follow the same pattern of hierarchy in accuracy maintaining the formulae mentioned in Table III.

TABLE III. EVALUATION METRICS FORMULA

Accuracy	Precision
$\frac{TP+TN}{TP+TN+FP+TN}$	$\frac{TP}{TP+FP}$
Recall	**F1-Score**
$\frac{TP}{TP+FN}$	$2*\frac{Recall*Precision}{Recall+Precision}$

B. *Accuracy Comparison of Classifier Models*

The prediction score of correctly classifying the selective five activities denotes performance state of the classifier models and the necessity of proper feature engineering to achieve higher performance. There have been multiple works on the original dataset from UCI repository, and the results from the previous research works align with this research results for correctly labeling activities, while the difference of the works lies in the detection approach taken by different research works. The accuracy score of the activities “Eat” and “Phone” is the lowest in one of the first researches executed under the lead of Diane J.Cook [34].

The scores of activities labeling were respectively 0.17 and 0.01 for the two above mentioned activities, which turns in this case to 0.00 and 0.11 respectively. The scarcity of enough training instance is one of the concerns of such low accuracy, which is reflected in the support score of the activities in the classification report in Fig.7(A).

Figure 6(a) represents the evaluation metric scores obtained by the evaluation parameters (Accuracy, Precision, Recall and F1-score). The evaluation scores of each activity are strongly correlated with the support value of respective activity. The overall accuracy reaches to 78.5% upon application of feed forward neural network on the Extra-tree classifier-based feature set. The activities ('Cook':0,'Read':3,'Watch_TV':4) have dominated the scores of evaluation parameters in this case with high correct prediction scores. Figure 6(b) presents the normalized confusion matrix on the output of MLP classifier on Extra-tree classifier-based feature set. The diagonal values represent the percentage of correct predictions, which is higher for activities ('Cook':0,'Read':3,'Watch_TV':4).

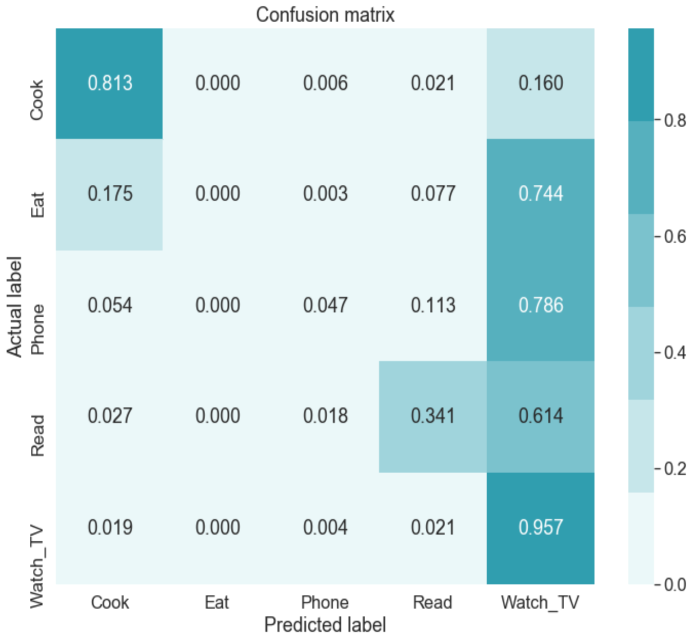
The off-diagonal elements also present noticing percentage value, representing mislabeled values by the classifier. The baseline classifier used in this case has achieved accuracy of 78.5%.

Accuracy Score : 0.7852490335843407

Classification Report

	precision	recall	f1-score	support
0	0.80	0.82	0.81	21335
1	0.00	0.00	0.00	3721
2	0.36	0.07	0.11	12117
3	0.61	0.35	0.45	21336
4	0.81	0.95	0.87	120473
Micro avg	0.79	0.79	**0.79**	179012
Macro avg	0.51	0.44	0.45	179012
Weighted avg	0.73	0.79	0.74	179012

a)



b)

Figure 6. Evaluation Metric of Neural Network on Extra Tree Classifier Feature set. (a) Metric Scores (b) Normalized Confusion Matrix.

In the Table IV, the classification accuracy scores upon the five selected activities on the five different features-set based datasets trained on the five-classifier models is presented. All the feature sets are run through PCA-5 analysis before the classifiers are applied. Application of tree-based feature engineering enhance the accuracy score for all the classifiers as shown in Table IV. Here, the Extra Tree Classifier feature set shows the highest accuracy on average for all the classifier models. Among the five classifiers, Feed forward Neural Network performs the best.

TABLE IV. COMPARISON OF PERCENTAGE IN ACCURACY SCORES OF CLASSIFIERS ON FEATURE SETS

Classifier	**Original Set of 37 Features**	**Model Obtained Selected Features with Tree-based Feature Selection**	**Model Obtained Selected Features with Random Forest Classifier**	**Top 21 Extracted Features by Extra Tree Classifier**	**Top 21 Extracted Features by Random Forest Classifier**
Nearest Neighbor	74.2	75.4	72.9	75.7	75.3
Decision Tree	75.3	76.9	75.9	76.3	75.5
Random Forest	74.4	76.4	75.7	75.9	74.5
Naive Bayes	74.9	76.5	76.4	76.9	76.3
Neural Net	76.7	78.3	77.3	78.5	78.1

VI. CONCLUSION AND FUTURE WORK

In this paper, we have presented the use of feature selection in improving the performance of classifiers. The classifier models show significant changes after application of precise data preprocessing and feature selection approaches. Feed forward Neural Network has persistently detected all five activities with varying metric score in the five datasets of this research. The highest accuracy score achieved on these feature sets is 78.5%, provided that the feature engineering followed baseline approaches along with baseline classifiers. Hence, this research presents that for human activity recognition systems, data preprocessing and feature selection greatly affects the classification performance and consequently the AAL and AML structures on the basis of HAR. State-of-the-art classifier models have presented varying accuracy score on the basis of how well the dataset has been preprocessed for running machine learning model on the dataset. To suggest more amicable work based on such data, this research could be explored in variety of fields in health, administration and security issues where such dataset generation and model implementation will be useful for activity recognition.

REFERENCES

[1] Q. Ni, A. B. García Hernando, and I. P. de la Cruz, "The Elderly's Independent Living in Smart Homes: A Characterization of Activities and Sensing Infrastructure Survey to Facilitate Services Development." Sensors. vol. 15, no. 5, pp. 11312-62, May 2015.

[2] Banos, O., Garcia, R., Holgado-Terriza, J. A., Damas, M., Pomares, H., Rojas, I., Saez, A., & Villalonga, C. (2014). mHealthDroid: a novel framework for agile development of mobile health applications. In International Workshop on Ambient Assisted Living (pp. 91-98): Springer.

[3] K. Davis, E. Owusu, C. Regazzoni, L. Marcenaro, L. Feijs, and J. Hu,"Perception of human activities a means to support connectedness between the elderly and their caregivers," in Proceedings of the 1st International Conference on Information and Communication Technologies for Ageing Well and e-Health. SCITEPRESS, 2015, pp. 194-199.

[4] A. Alberdi, A. Weakley, A. Goenaga, M. Schmitter-Edgecombe, and D. Cook. Automatic assessment of functional health decline in older adults based on smart home data. Journal of Biomedical Informatics, 18:119-130, 2018.

[5] Y. Chen, L. Yu, K. Ota, M. Dong, "Robust activity recognition for aging society", IEEE J. Biomed. Health Inform., vol. 22, no. 6, pp. 1754-1764, Nov. 2018.

[6] A. Alberdi, A. Weakley, A. Goenaga, M. Schmitter-Edgecombe, and D. Cook. Automatic assessment of functional health decline in older adults based on smart home data. Journal of Biomedical Informatics, 18:119-130, 2018.

[7] N. A. Capela, E. D. Lemaire, and N. Baddour, "Feature selection for wearable smartphone-based human activity recognition with able bodied, elderly, and stroke patients," PloS one, vol. 10, no. 4, p. e0124414, 2015.

[8] Chung, Seungeun et al. "Sensor Data Acquisition and Multimodal Sensor Fusion for Human Activity Recognition Using Deep Learning." Sensors (Basel, Switzerland) vol. 19,7 1716. 10 Apr. 2019, doi:10.3390/s19071716.

[9] Cornacchia, M., Ozcan, K., Zheng, Y., & Velipasalar, S. (2017). A Survey on Activity Detection and Classification Using Wearable Sensors. IEEE Sensors Journal, 17, 386-403.

[10] S. Aminikhanghahi and D. Cook. Enhancing activity recognition using CPD-based activity segmentation. Pervasive and Mobile Computing, 53:75-89, 2019.

[11] S. Aminikhanghahi, T. Wang, and D. Cook. Real-time change point detection with application to smart home time series data. IEEE Transactions on Knowledge and Data Engineering, to appear.

[12] B. Chen, Z. Fan, and F. Cao, "Activity Recognition Based on Streaming Sensor Data for Assisted Living in Smart Homes," in 2015 International Conference on Intelligent Environments, 2015, pp. 124-127.

[13] D. Cook, N. Krishnan, and P. Rashidi. Activity discovery and activity recognition: A new partnership. IEEE Transactions on Systems, Man, and Cybernetics, Part B, 43(3):820-828, 2013.

[14] K. Robertson, C. Rosasco, K. Feuz, M. Schmitter-Edgecombe, and D. Cook, "Prompting technologies: A comparison of time-based and context-aware transition-based prompting.," Technol. Health Care, vol. 23, no. 6,pp. 745-56, Jan. 2015.

[15] Nhan Duc Nguyen, Duong Trong Bui, Phuc Huu Truong, and Gu-Min Jeong, "Position-Based Feature Selection for Body Sensors regarding Daily Living Activity Recognition," Journal of Sensors, vol. 2018, Article ID 9762098, 13 pages, 2018. https://doi.org/10.1155/2018/9762098.

[16] FangH,etal. Human activity recognition based on feature selection in smart home using back-propagation algorithm. ISATransactions(2014).

[17] J.-L. Reyes-Ortiz, L. Oneto, A. Samà, X. Parra, and D. Anguita,"Transition-Aware Human Activity Recognition Using Smartphones," Neurocomputing, vol. 171, pp. 754-767, Jan. 2016.

[18] Chung, Seungeun et al. "Sensor Data Acquisition and Multimodal Sensor Fusion for Human Activity Recognition Using Deep Learning." Sensors (Basel, Switzerland) vol. 19,7 1716. 10 Apr. 2019, doi:10.3390/s19071716.

[19] Gu T, Wu Z, Tao X, Pung HK, Lu J. epSICAR: an emerging patterns based approach to sequential, interleaved and concurrent activity recognition. Proceedings of the IEEE International Conference on Pervasive Computing and Communications; 2009. pp. 1-9.

[20] Nweke, Henry & Wah, Teh & al-garadi, Mohammed & Alo, Uzoma. (2018). Deep Learning Algorithms for Human Activity Recognition using Mobile and Wearable Sensor Networks: State of the Art and Research Challenges. Expert Systems with Applications. 105. 10.1016/j.eswa.2018.03.056.

[21] A. Jordao, A. C. Nazare, J. Sena, W. R. Schwartz, "Human activity recognition based on wearable sensor data: A standardization of the state-of-the-art" in arXiv:1806.05226, 2018.

[22] G. A. Oguntala et al., "SmartWall: Novel RFID-Enabled Ambient Human Activity Recognition Using Machine Learning for Unobtrusive Health Monitoring," in IEEE Access, vol. 7, pp. 68022-68033, 2019. doi: 10.1109/ACCESS.2019.2917125.

[23] Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, et al. (2011) Scikit-learn: Machine Learning in Python. J Mach Learn Res 12: 2825-2830.

[24] Grabczewski, Krzysztof & Jankowski, Norbert. (2005). Feature selection with decision tree criterion. 6 pp.. 10.1109/ICHIS.2005.43.

[25] Kursa, Miron & Rudnicki, Witold. (2011). The All Relevant Feature Selection using Random Forest.

[26] Geurts, Pierre & Ernst, Damien & Wehenkel, Louis. (2006). Extremely Randomized Trees. Machine Learning. 63. 3-42. 10.1007/s10994-006-6226-1.

[27] Ali, Jehad & Khan, Rehanullah & Ahmad, Nasir & Maqsood, Imran. (2012). Random Forests and Decision Trees. International Journal of Computer Science Issues(IJCSI). 9.

[28] Raileanu, Laura & Stoffel, Kilian. (2004). Theoretical Comparison between the Gini Index and Information Gain Criteria. Annals of Mathematics and Artificial Intelligence. 41. 77-93. 10.1023/B:AMAI.0000018580.96245.c6.

[29] A. G. Karegowda, A. S. Manjunath, & M. A. Jayaram, (2010). Comparative study of attribute selection using gain ratio and correlation based feature selection. International Journal of Information Technology and Knowledge Management, 2(2), 271-277.

[30] B. A. Tama, & K. H. Rhee, (2015). A combination of PSO-based feature selection and tree-based classifiers ensemble for intrusion detection systems. In Advances in Computer Science and Ubiquitous Computing (pp. 489-495). Springer, Singapore.

[31] G. Stein, B. Chen, A. S. Wu, & K. A. Hua, (2005, March). Decision tree classifier for network intrusion detection with GA-based feature selection. In Proceedings of the 43rd annual Southeast regional conference-Volume 2 (pp. 136-141).

[32] S. Beniwal, & J. Arora, (2012). Classification and feature selection techniques in data mining. International journal of engineering research & technology (ijert), 1(6), 1-6.

[33] Y. Saeys, I. Inza, & P. Larrañaga, (2007). A review of feature selection techniques in bioinformatics. bioinformatics, 23(19), 2507-2517.

[34] Cook, Diane & Crandall, Aaron & Thomas, Brian & Krishnan, Narayanan. (2013). CASAS: A smart home in a box. Computer. 46. 10.1109/MC.2012.