Human Activity Recognition Using Accelerometer, Gyroscope and Magnetometer Sensors: Deep Neural Network Approaches

Abdul Kadar Muhammad Masum
Dept. of Computer Science and Engineering
International Islamic University Chittagong, Bangladesh
akmmasum@yahoo.com

Erfanul Hoque Bahadur
Dept. of Computer Science and Engineering
International Islamic University Chittagong, Bangladesh
ehb1729@gmail.com

Ahmed Shan-A-Alahi
Dept. of Computer Science and Engineering
International Islamic University Chittagong, Bangladesh
shourabh_iiuc@yahoo.com

Md. Akib Uz Zaman Chowdhury
Dept. of Computer Science and Engineering
International Islamic University Chittagong, Bangladesh
akibzaman0007@gmail.com

Mir Reaz Uddin
Dept. of Computer Science and Engineering
International Islamic University Chittagong
Chittagong, Bangladesh
reaz.ctg2011@gmail.com

Abdullah Al Noman
Dept. of Computer Science and Engineering
International Islamic University Chittagong
Chittagong, Bangladesh
alnomancse143@gmail.com

***Abstract*—Nowadays, smartphone with various extraordinary and notable sensors makes new invigorating open entryways for Data Mining and Machine Learning; other than makes another exploration field for Human Activity Recognition a.k.a. HAR. Such development advanced the manufacture of an assortment of useful datasets, which encourages the investigation of certainties of different research area. The MHEALTH, a multivariate time series dataset is such an optional dataset that was arranged in order to encourage the investigation with respect to HAR. The dataset includes data with respect to twelve human activities oppressing ten volunteers. Sensors were utilized for the information amassing process to be specific Accelerometer, Gyroscope, Magnetometer and Electrocardiogram signal. This paper plays out a relative report on HAR process as far as work of four unique information preprocessing strategies joined by six popular classifiers entitled as Random Forest, Support Vector Machine, Naive Bayes, along with three deep learning approaches, Multilayer Perceptron, Deep Convolutional Neural Network and Long-Short Term Memory. Information preprocessing strategies that were utilized on the referenced dataset are disposal of invalid name occurrences, consistency of unequal classes, low-pass filtering and Principle Component Analysis. The objective of this examination is to break down execution of various classifiers as far as the referenced information preprocessing techniques and furthermore distinguishing the procedure for which the classifiers display prevalent precision.**

Keywords—Human Activity Recognition (HAR); MHEALTH; Accelerometer sensor; Gyroscope sensor; Magnetometer sensor

I. Process and Consumer Researches Prolusion

Physical movement of an individual vows a fathomable origination about one's composite way of life. It can be associated to the components, for example, cardiovascular diseases [1], anomalous exercises, mental health [2] and so forth. In this manner, investigates towards these components request the acknowledgment of an assortment of physical movement performed by human. With the development of this acknowledgment procedure likewise titled as Human Activity Recognition a.k.a. HAR, an intelligible investigation of the variables related with physical movement can be practicable.

As to upgrade of digitalization, an assortment of sensors have been designed. These various sensors were imagined to facilitate the mechanical procedure. In any case, a brilliant side of proposals sensors worries that execution of information gathering process as far as an assortment of activities can be directed by these sensors. Along these lines, to encourage examines about different highlights, an assortment of dataset have been created with the work of various sensors. Such an optional dataset is the MHEALTH dataset that was manufactured to facilitate the investigation of certainties in regards to human movement with the work of Accelerometer, Gyroscope, Magnetometer and Electrocardiogram sensor. Since unique sensors show unalike qualities regarding disparate exercises, thusly with the ramifications of sensors information we can apparently recognize the physical movements performed by an individual. Another factor which ought to be examined with need is that, information gathering procedure can be loud and datasets created by others may should be improved in the event that it might envelops purposeless information in regards to somebody's need. On that account, assorted information dissemination and separating procedure have been acquainted so as with adjust the accessible datasets as indicated by one's order. Separating forms really change the datasets such that, nearness of loud information will in general be invalid and filter the dataset to get a handle on the most extreme sufficiency.

Our targets point to the work of the MHEALTH dataset to play out the acknowledgment of twelve unique exercises, entitled as, standing still, frontal elevation of arms, lying down, climbing stairs, waist bends forward, walking, knees bending (crouching), cycling, jogging, sitting and relaxing, running, jump front and back by separating the dataset with four information preprocessing forms to be specific disposal of invalid mark cases, consistency of uneven classes, low-pass filtering and Principle Component Analysis. For every one of the separating procedure, six classifiers specifically Random Forest a.k.a. RF, Support Vector Machine a.k.a. SVM, Naive Bayes a.k.a. NB along with three deep learning approach, Multilayer Perceptron a.k.a. MLP, Deep Convolutional Neural Network a.k.a. CNN and Long-Short Term Memory a.k.a. LSTM were practiced and a near report was directed to finish up for which sifting process the classifiers display the unrivaled result.

As to examines, HAR got a handle on a quintessential position in specialists' perspectives. HAR was led utilizing distinctive techniques [3, 4] and disparate recommendations. Ronao and Cho [5] practiced the profound CNN for the execution of HAR alongside investigating the issues of convnets layers with the acknowledgment procedure and accomplished a precision of 94.97% on test set with grungy sensor information and 95.75% with auxiliary learning of first Fourier change. Lu, et al. [6] practiced an unsupervised learning approach qualified as MCODE for assess the viability of their strategy on three genuine world datasets collected by accelerometer sensor. Another exploration directed by Chamroukhi, et al. [7] encompasses the work of wearable sensors put on chest, thigh and lower leg and practicing an unsupervised learning model entitled as desire amplification calculation. They obtained a normal grouping precision of 90.3%. Moreover, looks into, for example, [8-10] were performed HAR by different wearable sensors at divergent anatomical positions and closed normal exactness of 91.50%, 89.08% and 98% individually. A decentquantity of investigates have been led on HAR process with the work of camera perceptions. Liu, et al. [11] played out the human movement acknowledgment practicing comparable strategy. Alongside the wearable sensors, consideration of RGB-camera information was additionally presented for the acknowledgment procedure of 20 different human movements by practicing the SVM calculation. Another use of camera information was presented by Liciotti, et al. [12] for the achievement of HAR process so as to facilitate the way toward identifying irregular human conduct. An assortment of examines presented HAR with enhancement of encompassing elements, for example, vitality, cost and so forth. Ding, et al. [13] led the acknowledgment procedure of 10 exercises utilizing Arbitrary Woods calculation with an exactness of 93.01% yet declining vitality utilization by 74.9%. An extra research achieved by Ghasemzadeh, et al. [14] presented an engineering for action acknowledgment process which diminish the power exhaustion of the handling by 70% though the location affectability continue at least 80%.

As to achieved explores, it tends to be seen that the execution of HAR has been performed utilizing unique models alongside assorted components enhancement. A bunch of correlations among dissimilar models have additionally been shown. Be that as it may, our investigation aches for dealing with the examination on premise of order models as well as based on information preparing models. We examined the impacts of information dissemination as far as HAR and closed an unmistakable examination between two different information preprocessing procedures in regards to action acknowledgment.

The later fragments of this paper portray the modus operandi in segment II, execution assessment in segment III, discussion in segment IV and conclusion in segment V.

II. Modus Operandi

Starting move towards the instatement of our disclosures joins the thought of setting up the data for our arranged action. We preprocessed our data utilizing four approaches in specific transfer of invalid values, consistency of uneven classes, low-pass filter and Principle Component Analysis a.k.a. PCA. Taking after, we associated six different classifiers on the dataset after utilization of each preprocessing strategy and evaluated how the classifiers show as distant as each preprocessing proposition.

A. MHEALTH Dataset

Since we directed our exploration on the MHEALTH dataset, thusly, a concise portrayal about the qualities of dataset ought to be raised. The MHEALTH dataset includes the body developments and imperative sign discoveries for ten volunteers of unique representations at the season of performing twelve different physical exercises. Work of wearable sensors was led for the store of information. Sensors were joined at differing bit of body to be specific right wrist, chest and left lower leg with the assistance of versatile ties. Information was amassed with fuse of a recurrence about 50Hz, which has been demonstrated to be adequate for typify esteems in regards to the body developments. The exercises for which information store was led involves standing still for 1 min, sitting and relaxing for 1 min, lying down for 1 min, walking for 1 min, climbing stairs for 1 min, waist bends forward repeated for 20 times, frontal elevation of arms repeated for 20 times, knees bending (crouching) repeated for 20 times, cycling for 1 min, jogging for 1 min, running for 1 min, jump front and back repeated for 20 times. Here Accelerometer sensors data were scaled in m/s^2 unit, Gyroscope sensors data were scaled in radian/s unit, Magnetometer sensors data were scaled in local unit and Electrocardiogram sensors data were scaled in mV unit.

B. Data Cleansing and Preprocessing

Despite the overwhelming evidence that in every instance the MHEALTH dataset was gathered with such an enormous arrangement, there has been some adjustment that we required to accomplish our exploring goal. For a while there, we didn't actually believe about the Electrocardiogram signal sensor information with lead 1 and lead 2 axes for our exploration. We consequently confirmed these two portions from the dataset.

A various number of invalid mark examples can be seen from the MHEALTH dataset. This invalid mark example actually indicates the information related with the change starting with one action then onto the next. Be that as it may, a matter of critical concern is, presence of various invalid name occurrences may cause lopsidedness in the dataset and an imbalanced dataset holds extensive measure of impacts to cause difficult issue in assessment of effectiveness of research. In this manner, we organized four preprocessing approach so as to decrease the imbalance in dataset.

On a very basic level we reduced all the invalid mark occurrences. The invalid mark occasions truly don't speak to any physical movement. Along these lines, evacuation of invalid name occasions ought not to have any kind of repercussions in the acknowledgment procedure of physical action. A pictorial perspective on the dataset with and without invalid mark occurrences are portrayed underneath to pass on the imbalance brought about by invalid name examples:

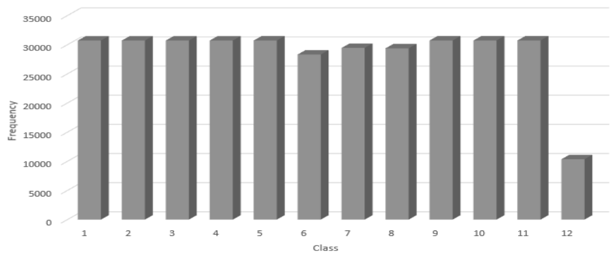

Figure 1: Class histogram before removing null label class

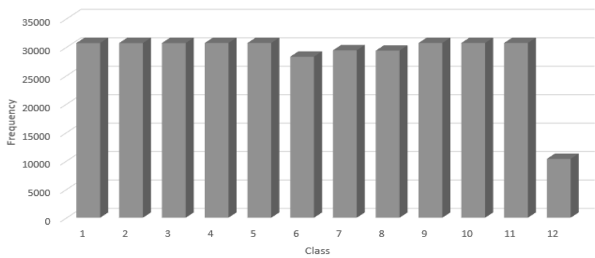

Figure 2: Class Histogram after removing null label class

The diversity of the unequaled category is another preprocessing process that we conducted to lessen the disproportion in the data set. The technique of the ClassBalancer filter from the WEKA framework obtained uniformity in all classes.

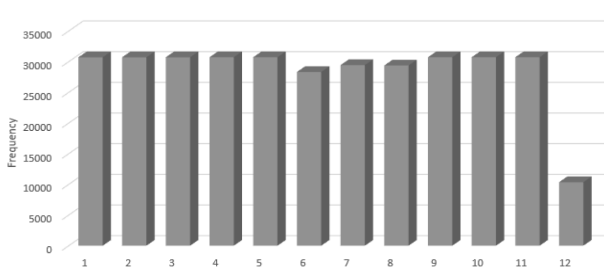

Figure 3: Class histogram before balancing

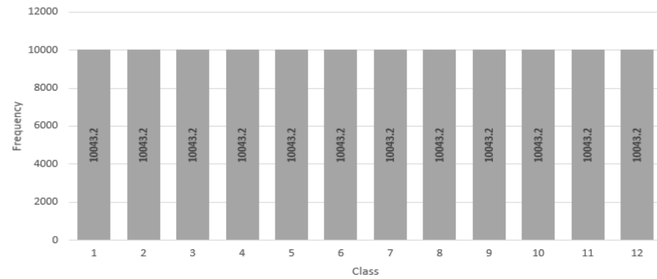

Figure 4: Class histogram after balancing

A sort of signal processing filter can be described as the low queue filter which is intended to turn the frequency produced into soft frequency. Butterworth filter takes a limit number throughout sampling and extracts every single signal value above it. Data abnormalities caused by the ground wave gravitational influence lose its smoothness and there is therefore a divergence. And that the Butterworth low-pass filter mechanism was used to eliminate anomalies related to the gravity of the earth. An individual may use a distinct kind of mathematical formulation when using the Butterworth low pass filter. Such a formula is the Fourier Transform that can divide any signals and display them as a sine vector. This implies slowing down a fluctuated signal, which makes it as soft as a sinusoidal signal. Here is the instance of filtering approach on Accelerometer data:

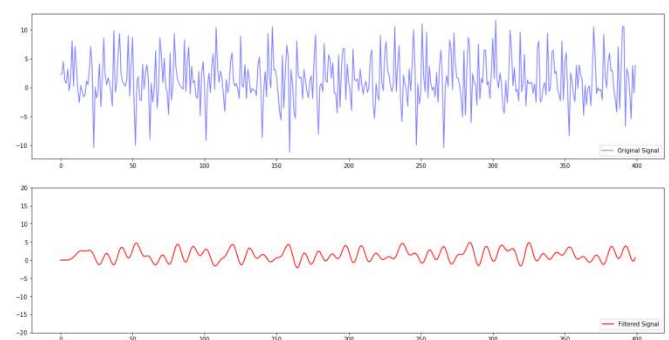

Figure 5: ButterWorth Low Pass Filter on Accelerometer sensor data

The goal of the mill of a Principle Component Analysis a.k.a. PCA is to decrease its dimensionality by anticipating it in a small subspace, in which the prescient vectors frame the tomahawks. In any event, the individual vectors characterize the coils of the new pivot, because they are all unit length 1 similar. We must explore the corresponding proper values in order to choose which prophetic vectors can be removed without leaving any residual information for low-dimensional subspace growth. The vectors with the smallest value of their own have minimal data on the communication of the information; these can be discarded. In order to do this, the fundamental methodology is to select the base k vectors from the highest to the less so. Following on from the arrangement of the Eigen pairs, the question below is "how many important sections will we select for our fresh subspace?" The alleged "clear fluctuation," that can be determined by our own values, is a useful measure [15]. The clear shift shows us the amount of information fluctuation in each of the main sections can be attributed. In the key components' inquiry of the PCA model, there are essentially no proven avenues of identifying, in some manner, an optimal or any other appropriate amount of vital section PCs, especially given a volume restricted adaption information matrix. Therefore, the PCs those are having co-variance is greater than 0.1, Eigen value is greater than 1.0 and co-variance among the Eigen values is greater than 0.1are chosen. In this criterion we selected first 17 PCs only.

C. Exercised Models

We exercised six models in terms of each preprocessing procedure for the acquisition of clear collation between the four mentioned earlier preprocessing procedures. We implemented three base level classifiers: Naïve Bayes, Random Forest and Support Vector Machine and three deep classifiers: Multi-Layer Perceptron, Convolutional Neural Network and Long-Short Term Memory.

Naïve Bayes Classifier is a narrow yet robust data mining algorithm constructed on the Bayes theorem. Bayes theorem demonstrates the conditional probability, i.e. the probability of an case appearing even though another event occurred. The Bayes theorem can be summarized as follows:

$$P(X|Y) = \frac{P(Y|X)P(X)}{P(Y)} \tag{1}$$

Where P(X) reflects the subsequent frequency of incidence of case X, P(Y) indicates the subsequent frequency of incidence of case Y, P(Y) provides the later likelihood of incidence of case Y provided that case X is true and P(X) illustrates the expected likelihood of case X provided Y is true. But in order of categorization issue we can narrate Y as the uncertain class information test, X as the Y test refers to category C and P(X|Y) as the probability that for given data sample Y the hypothesis holds.

Narrating the Random Forest classifier, we can classify it as an instant training strategy that amplifies the interaction of the decision tree. Random Forest initialization presents a set of choice forests where preceding nodes show the predictor components. Subsequently, the density of decision trees produced helps to achieve the final product of results. Despite the fact that the beginning of high variances can be initialized by generated subtrees, but low overfitting can be achieved with the amalgamation of the learning rate. In our organized classifier, we chose to produce 100 decision trees to achieve the optimal result.

Another well-liked classifier, Support Vector Machine, establishes the emergence of hyper-plane variation to simplify the classification problem. In terms of classification, an effective hyper-plane is manufactured for a specified amount of classifications that helps to categorize future samples. Choosing the correct kernels as well as the regularization parameter and gamma is evident for tuning the SVM model. The kernels are integrated here to identify the resemblance metric between different characteristics and the gamma value helps to effectively adjust the SVM model. Generally speaking, kernels effectively improve the simplification of non-linear categorization as SVM basically presents binary classification. We practiced the "RBF" kernel and a gamma value of 0.01 in our SVM model architecture for which we obtained the highest results result.

Briefing on the Multi-Layer Perceptron a.k.a. MLP, it initializes its three significant parts, namely input layer, one or more hidden layers, and output layer, by incorporating it. Inheritance of input vectors (predictor characteristics) is pulled out on the hidden layer; on the contrary, the output layer finalizes some specific judgment regarding the predictor characteristic values. We can demonstrate the entry vector by O_x in order to formulate an MLP model, where x presents the entry element. To calculate the hidden layer y entry, an adder feature is practiced and can be mathematically stamped by,

$$I_y = \sum w_{xy} O_x + b_y \tag{2}$$

Here, w_{xy} stands for the weighted linkage between entry layer x neurons and hidden layer y and specifies the preference significance. The calculated concealed layer entry I_y is then fed into an activation function for calculating the result and can be mathematically explained by,

$$O_y = Act(I_y) \tag{3}$$

Where the activation function is indicated by Act(). Subsequently, the output layer is incorporated and the defect fraction can be calculated to decrease the incidence of the misstatement by means of the back-propagation algorithm. We initiated 17 hidden layers with 0.5 momentum and decreased mistake proportion through the introduction of back-propagation algorithm for the implementation intent in MLP.

A Convolutional Neural Network a.k.a. CNN model is closely related to the MLP but differentiates by bringing the layer of convolution and the layer of pooling. CNN architecture can generally be segmented into two similarly important parts, namely the removal part function that includes convolutionary layer and pooling layer, and the other part is the identification part that includes fully linked layer and execution of softmax.CNN input generally reflects the shape of the matrix. It uses the sliding window to supply its required entry framework, to be accurate. The convolution coating subsequently manipulates the segmented screen outputs by using kernels to obtain low-volume functionality. In CNN, the function extractors can define the kernels as it is initially liable for input processing. Afterwards, when the modified characteristics are acquired, we can implement the use of the pooling strands to decrease the temporal magnitude of the extracted characteristics which can promote the reduction of computing burdens. For example, by returning the maximum value of a manipulated feature segment, a max pooling can decrease the volume of the feature. Subsequently, we can transfer the production of the convolution part into a fully linked load forward neural network for evaluation and the implications. In our CNN model, we employed a 3*3 kernel for the convolution portion and as an activation function, we elected the Relu activation. For the fully connected layer, softmax function was introduced in case of activation.

From the architectural point of view, the various number of memory cell blocks make up aLong-Short Term Memory a.k.a. LSTM network. Twostates, called cell state and hidden state, move data and create tiny changes to the data through multiplications and comparison respectively. Through the significant mechanics, called gates, sections and manipulations in memory blocks are performed. There is a gate called forget gate to optimize the performance of the LSTM network, which removes information that is less important or the information is no longer needed for the LSTM by multiplication a filter.The logistics function, which contains the hidden state of the preceding unit (h_{t-1}) and the entry as a domain at that specific moment phase (x_t), is accountable for deciding which series to hold and which to discard. Thus, after leaving the state, the tissue condition becomes:

$$C_t = f_t * C_{t-1} \tag{4}$$

Where f_t is the forget state and $C_{t-1}$is the previous cell state. Forget state is stated as:

$$f_t = \frac{1}{(1+e^{-(W_f*[h_{t-1},x_t]+b_f)})} \tag{5}$$

Where W_f is the matrix of weight and where b_f is the vector of prejudice. The output vector is increased to the cell state from the sigmoid function. Then an input gate adds sequences to the cell state. Another sigmoid function acts as a filter, controlling the variables to be introduced to the cell state and creating a vector with a hyperbolic tangent function that contains all feasible variables to be inserted. The combination of the sigmoid layer and the hyperbolic tangent feature applicant scores are subsequently introduced to the cell state. Thus, only significant and non-redundant components are chosen first and then introduced to the state of the cell. Thus, cell state after the input state becomes:

$$C_t = f_t * C_{t-1} + i_t * C'_t \tag{6}$$

Where i_t is the input state, holding sigmoid layer and C'_t is holding candidate values. Both functions stated below:

$$i_t = \frac{1}{(1+e^{-(W_i*[h_{t-1},x_t]+b_i)})} \tag{7}$$

$$C'_t = \frac{(1-e^{-2(W_c*[h_{t-1},x_t]+b_c)})}{(1+ e^{-2(W_c*[h_{t-1},x_t]+b_c)})} \tag{8}$$

Where W_i and W_c are weight matrices and b_i and b_c are bias vectors of initial cell and medium cell state. Another filter is made to decide which sections of the cell state are supposed to perform via sigmoid function and a vector is produced after placing the cell state through hyperbolic tangent function, thereby scaling the scores in the spectrum from -1 to +1. Finally, the hyperbolic tangent vector item and the legislative filter transmit input to the next hidden state and this item ensures that only the significant parts are generated as inputs. The output state therefore generates the status of the unit as:

$$C_t = o_t * h_t \tag{9}$$

Where o_t is the regulatory filter and ht is the vector of hyperbolic tangent values. Both functions are stated below:

$$o_t = \frac{1}{(1+e^{-(W_c*[h_{t-1},x_t]+b_o)})} \tag{10}$$

$$h_t = \frac{o_t*(1-e^{-2(C_t)})}{(1+e^{-2(C_t)})} \tag{11}$$

Where W_o and b_o is the destination gate's weight matrix and bias vector.

III. Performance evaluation

In order to assess the efficiency of all the above-mentioned models, we implemented 10-fold cross validation. We slice the whole surveillance in 10 slides with a 10 consecutive cross-validation and choose one of them as the validation range, the rest 9 leaves function as the workout range. This is done 10 trials, and the validation matrix for each iteration is chosen for each loop.

As already indicated, we have preprocessed information with four methods and therefore have conducted the validation method by using both information sets to discover a better preprocessing method in HAR. We initially carried out the data set validation method that was developed by deleting all the cases of a void tag. We have conducted the validation method for all classifiers and registered the respective results assessment. The second dataset ready by separating all the classifications was carried out by the same procedure. We can see that, as shown in the previous two tables, all the classifiers displayed greater data collecting accuracy, accurateness, and recall by extraction of NULL tag cases.

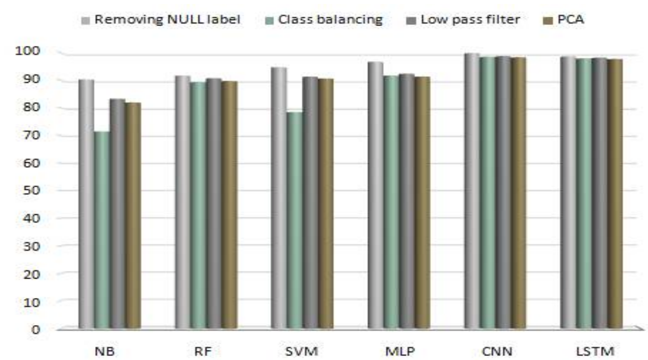

Figure 6: Performance evaluation pre-processing methods on practiced models

The preprocessing technique; the removal of null mark instances is gaining a margin of 90.26%-71.46% precision in the basic classifier, 91.63%-89.28% recall, 94.69%-78.46% accuracy respectively in the case of Naïve Bayes, Random Forest and SVM. The context is the same with regard to low-level classification where the ratio is 96.66%-91.68%-, 99.8%-98.48% for MLP, CNN and LSTM. Moreover, since we are able to gain a straightforward understanding of the finest method for the pre-processing of our work, we can assess which types of classifications are available on figure 6. Figure 6 show that deep-level classifiers are performing higher classification than the base-level classifiers. Another remark also says that inferior efficiency is measured by the deep CNN classification among three deep classifiers.

IV. Discussion

The MHEALTH data set is a popular dataset, and many scholars have carried out a number of study projects using this dataset. We have also tried, in our presentation, to improve the results of two pre-processing techniques that used the MHEALTH dataset. We were allowed to assume the superior by observing the classifiers' efficiency estimates for both pre-processing techniques. In spite of our contestant preprocessing technique, we also decided which sort of category has advanced. In addition, a greater precision procurement can also be regarded a major input. With CNN, we have reached 99.8 percent precision which is more accurate than many other research findings. We also think that the four preprocessing methods efficiency estimates encourage other researchers to convey our results to enhance the assessment of their study findings.

Null tag cases effectively constitute the temporary event significance in MHEALTH dataset; therefore, the extraction of null cases can influence if someone tries to identify the scenario in motion. But because our aim was not only the temporary contexts but the specific activity, we were therefore prepared to prevent that issue. The stock balance pre-processing technique believes the temporary attributes to be a spare class and thus the other classes might be faced with an problem of a misclassification owing to the existence of a spare group. A further fact should be discussed that the temporary points are considered as a spare category in the category balancing technique, which means other classes can experience a problem of misclassification because of the existence of a spare category. However, the suppression of null tag cases obviously removes the problem of error because of the presence of null cases. Therefore, we obtained a great deal of efficacy for pre-processing null tag cases rather than category filtration processing.

V. Conclusion

Recognition of human activity has a major effect on different apps, including diagnostic disease and identification of abnormal operations. We have researched the efficiency of numerous well-known classificators and numerous techniques of preprocessing used in the acceptance of natural behavior. The results also include the choice of finest preprocessing methods between four preprocessing processes that are quickly relevant. We also evaluated for the chosen pre-processing technique the finest sort of classifier. In recognition of natural activities with a precision of 99.8 percent, the Deep CNN classification showed exceptional efficiency. We believe that the study findings will make further studies into classification and data pre-processing techniques possible in this field.

References

[1] N. C. Barengo, R. Antikainen, K. Borodulin, K. Harald, and P. Jousilahti, "Leisure□time physical activity reduces total and cardiovascular mortality and cardiovascular disease incidence in older adults," *Journal of the American Geriatrics Society,* vol. 65, pp. 504-510, 2017.

[2] D. Lubans, J. Richards, C. Hillman, G. Faulkner, M. Beauchamp, M. Nilsson*, et al.*, "Physical activity for cognitive and mental health in youth: a systematic review of mechanisms," *Pediatrics,* vol. 138, p. e20161642, 2016.

[3] A. Ignatov, "Real-time human activity recognition from accelerometer data using Convolutional Neural Networks," *Applied Soft Computing,* vol. 62, pp. 915-922, 2018.

[4] A. Murad and J.-Y. Pyun, "Deep recurrent neural networks for human activity recognition," *Sensors,* vol. 17, p. 2556, 2017.

[5] C. A. Ronao and S.-B. Cho, "Human activity recognition with smartphone sensors using deep learning neural networks," *Expert systems with applications,* vol. 59, pp. 235-244, 2016.

[6] Y. Lu, Y. Wei, L. Liu, J. Zhong, L. Sun, and Y. Liu, "Towards unsupervised physical activity recognition using smartphone accelerometers," *Multimedia Tools and Applications,* vol. 76, pp. 10701-10719, 2017.

[7] F. Chamroukhi, S. Mohammed, D. Trabelsi, L. Oukhellou, and Y. Amirat, "Joint segmentation of multivariate time series with hidden process regression for human activity recognition," *Neurocomputing,* vol. 120, pp. 633-644, 2013.

[8] A. Bayat, M. Pomplun, and D. A. Tran, "A study on human activity recognition using accelerometer data from smartphones," *Procedia Computer Science,* vol. 34, pp. 450-457, 2014.

[9] A. Moncada-Torres, K. Leuenberger, R. Gonzenbach, A. Luft, and R. Gassert, "Activity classification based on inertial and barometric pressure sensors at different anatomical locations," *Physiological measurement,* vol. 35, p. 1245, 2014.

[10] P. Gupta and T. Dallas, "Feature selection and activity recognition system using a single triaxial accelerometer," *IEEE Transactions on Biomedical Engineering,* vol. 61, pp. 1780-1786, 2014.

[11] X. Liu, L. Liu, S. J. Simske, and J. Liu, "Human daily activity recognition for healthcare using wearable and visual sensing data," in *2016 IEEE International Conference on Healthcare Informatics (ICHI)*, 2016, pp. 24-31.

[12] D. Liciotti, T. Duckett, N. Bellotto, E. Frontoni, and P. Zingaretti, "Hmm-based activity recognition with a ceiling rgb-d camera," 2017.

[13] G. Ding, J. Tian, J. Wu, Q. Zhao, and L. Xie, "Energy efficient human activity recognition using wearable sensors," in *2018 IEEE Wireless Communications and Networking Conference Workshops (WCNCW)*, 2018, pp. 379-383.

[14] S. Valle, W. Li, and S. J. Qin, "Selection of the number of principal components: the variance of the reconstruction error criterion with a comparison to other methods," *Industrial & Engineering Chemistry Research,* vol. 38, pp. 4389-4401, 1999.

[15] H. Ghasemzadeh, R. Fallahzadeh, and R. Jafari, "A hardware-assisted energy-efficient processing model for activity recognition using wearables," *ACM Transactions on Design Automation of Electronic Systems (TODAES),* vol. 21, p. 58, 2016.