Journal of Harbin Engineering University
ISSN: 1006-7043
Vol 44 No. 9
September 2023

Hybrid Classification Approach for Heart Disease Prediction

Mohd. Akram[1], Dr. Pooja Sharma[2], Kavita[3]
[1] Research Scholar, Rayat Bahra University, Mohali
ansariakram15031993@gmail.com
[2]Associate Professor, Computer Science Department, Rayat Bahra University, Mohali
pooja.rani@rayatbahrauniversity.edu.in
[3]Uttaranchal University, Dehradun
kavita@ieee.org

Abstract

The term "data mining" describes a technology that aids in extracting significant data from a large amount of amorphous data. On the other side, predictive analysis uses current data to forecast future events. The focus of this initiative is on foreseeing heart organ morbidities. The data is pre-processed, characteristics are extracted, and classification are all done as part of this method of predicting heart disease. The application of LR and RF algorithms results in the building of a hybrid framework. While LR categorizes the features, the RF algorithm seeks to extract them. A range of criteria are used in this study to evaluate how competent the designed framework is. The method suggested has a 95% accuracy rate for predicting heart conditions.

Keywords: Heart Disease, Random Forest, Logistic Regression, Ensemble Classification

1. Introduction

Heart disease is a group of illnesses which impacts on people's hearts and veins. Depending on the precise form of cardiac disease, there are various symptoms. An ongoing task that a skilled practitioner can complete with sufficient expertise and knowledge is identifying and diagnosing cardiovascular disease. Age, diabetes, smoking, being overweight, eating a diet high in junk food, and other factors are only a few [1]. There are a amount of variables/factors that either raise the risk of cardiac disorder or cause it. For monitoring their clinical and/or patient data, the majority of hospitals have management software. These systems produce a tonne of patient data, and it is currently popular. Rarely are these data used to help clinical decision-making. These data are useful and provide information that is largely unused. Converting the accumulated clinical data into information that can help intelligent systems support healthcare practitioners' decision-making is a very challenging issue. It was necessary to develop a quick and effective automated detection method due to the scarcity of experts and the volume of instances that were misdiagnosed. The major goal is to utilise the classifier model to categorise the important aspects of the medical data and use the classifier models to forecast heart illness early on [2].

Several papers, such as Feature Optimization Based Heart Disease Prediction using Machine Learning [3], have proposed research involving machine learning and deep learning. At first, the dataset was run through SVM, KNN, NB, and RF algorithms to make predictions about the heart disorder. According to the findings, the introduced system provided 87.78% accuracy with the help of the NB algorithm.

The implementation of an ensemble classifier for heart disease prediction has inspired my work because it outperforms the current model in terms of accuracy, precision, recall, and F1-score. The presented work focusses on

- Proposes a novel technique for HDP
- Dataset from Cleveland
- The evaluation metrics is used to compare results

2. Literature Review

2.1 Heart Disease Prediction using Machine Learning

D. P. Yadav, et.al (2021) introduced an automated system on the basis of analyzing diverse ML (machine learning) methods [3]. Initially, the

heart disorder was predicted through SVM, KNN, NB and RF algorithms on the dataset. A three-fold validation assisted in avoiding bias efficacy. The results reported that the introduced system offered 87.78% accuracy using NB algorithm and proved applicable. Furthermore, a GA (genetic algorithm) exploited on the given dataset with the objective of optimizing the attributes. The accuracy from NB algorithm was calculated 96% in this process.

M. Chakarverti, et.al (2019) projected KMC and SVM in order to cluster and classify the input information so that the coronary disease was predicted [4]. The information was clustered with the help of BP (back propagation) algorithm with KMC. These algorithms were useful for improving the precision while performing PA (prediction analysis). The projected algorithms were computed on a dataset extracted via UCI repository. There were seventy-six attributes comprised in this dataset. A comparative analysis was carried out on the projected algorithm against the traditional methods with regard to accuracy, ER (error rate) and execution time. The outcomes of experiment indicated that the projected approach was effective.

R. Latha, et.al (2019) suggested POMDP (partially observable markov decision process) algorithm for predicting the coronary disorder [5]. The fog computing helped the patient to send notification to the doctor so that an ambulance was sent at the location where the infection person was present. iFogSim offered the data to the doctors. The CVD (cardiovascular disease) occurred due to the maximization of blood viscosity. The coronary disorder was predicted on the basis of diverse factors such as higher BP, diabetes, higher blood viscosity, etc. The states and timeslots employed to predict the coronary disease using suggested algorithm.

X. Yuan, et.al (2022) established a ML (machine learning) based predictive algorithm for executing the binary and multiple classification to predict the heart disorder [6]. A Fuzzy-GBDT algorithm was put forward in which FL (fuzzy logic) algorithm integrated with GBDT for mitigating the data complexity and boosting the efficacy to generalize the binary classification. Thereafter, the integration of this algorithm was done with bagging for avoiding the issue related to overfitting. This integrated algorithm was effective for predicting the severity of heart disorder. The experimental outcomes confirmed that the integrated algorithm offered superior accuracy and proved stable.

M. Nahiduzzaman, et.al (2019) developed two algorithms namely MLP and SVM in order to foresee the heart disorder [7]. This fundamental focus was on classifying two issues namely 2-class and 5-class of coronary disease. The developed algorithms were evaluated on Cleveland in which303 instances, having five classes and thirteen features, comprised. The results exhibited that the first algorithm offered accuracy of 90.57% and 68.86% both the issues respectively. The second algorithm yielded accuracy of 92.45% for initial issue and 59.01% for second issue.

2.2 Heart Disease Prediction using Deep Learning

S. Bhoyar, et.al (2021) presented a viable and effective mechanism to predict the chances of a patient who suffered from the heart disease optimally [8]. An NN (Neural Network) was designed to predict the heart disorder in which an MLP (Multilayer Perceptron) algorithm was employed. Two datasets: UCI and CVD (Cardiovascular Disease) applied to simulate the presented system. The results of experiment revealed that the designed algorithm attained 85.71% accuracy on first dataset and 87.30% on latter dataset. A simple web application tool was created on the basis of Python to test the designed algorithm. Moreover, a promising tool was created for clinical professional.

M. A. Khan, et.al (2020) investigated an IoT (Internet of Things) based technique for predicting the coronary disease in more accurate manner for which MDCNN (Modified Deep Convolutional Neural Network) algorithm was implemented [9]. The BP (blood pressure) and ECG (electrocardiogram) of patients monitored with smartwatch and heart monitoring device. The received sensor data was classified into diseased and healthy using the investigated technique. The comparison of investigated technique was done with the traditional techniques for analyzing the efficiency. The results depicted the supremacy of the investigated technique against the existing methods and offered an accuracy of 98.2%.

P. Ramprakash, et.al (2020) recommended a model for understanding the principles to predict the risk profile of patients based on the healthcare data parameters [10]. DNN and χ 2 - statistical framework considered in the generation of this model. The recommended model was capable of removing the issue about under- and over fitting. The results indicated the efficacy of this model for predicting whether the coronary disease was occurred or not over the existing techniques.

S. Hussain, et.al (2021) intended a new DL (deep learning) method called a 1D-CNN (one dimensional convolutional neural network) algorithm to classify the healthy persons from the infected ones with balanced datasets for overcoming the drawbacks of traditional ML (machine learning) technique [11]. The risk contour was computed in the patients on the basis of various medical factors. Consequently, the disorder was diagnosed at initial phases. The issue about overfitting was avoided using several regularization techniques. The intended method attained an accuracy of 97% when the data was trained and 96% when the data was tested. The outcomes reported the effectiveness of the intended method over the conventional models.

H. Tiwari, et.al (2022) formulated a technique to predict the heart disease relied on DL (deep learning) method on UCI (University of California Irvine) dataset [12]. The coronary disease was predicted with regard to diverse factors such as age, blood pressure, and cholesterol. MLP (multilayer perceptron) was deployed with FCD (fully connected dense) layers was put forward. Thirteen attributes and a single output neuron were employed in the primary layer to classify the target class. A comparative analysis was performed on the formulated technique against the existing techniques. The results proved the applicability of the formulated technique for predicting the probability of the heart disorder in patients.

W. Li, et.al (2020) devised a CRMPN model for predicting the grade of coronary disorder through a colored Doppler ECG report, BB indicators of the patient [13]. DRL (deep reinforcement learning) was pre-trained with A3C (asynchronous advantage actor-critic). The RNN (recurrent neural networks) was optimized on training data to parameterize the stochastic policy. Thereafter, module to share the soft and hard metrics, and PDN employed for predicting the coronary disease. The findings indicated that the devised approach had generated efficient results as compared to other methods.

2.3 Heart Disease Prediction using Hybrid Techniques

M. T. Islam, et.al (2020) suggested a method for predicting heart disease in advance [14]. PCA (Principal Component Analysis) and integrated heuristic KMC (k-means clustering) algorithm with metaheuristic GA (Genetic Algorithm) to mitigate the dimensionality of the dataset. The final clustering quality was enhanced using the suggested technique. The experimental outcomes depicted that the suggested presented algorithms were applicable for predicting the coronary disorders in initial phases at an accuracy of 94.06%.

D Deepika, et.al (2021) introduced an optimized method to select the features and new MLP-EBMDA to predict the coronary disorder [15]. This method employed the dataset and the data was pre-processed later on. The optimized algorithm aided in selecting the attributes. these selected attributes were considered to classify the coronary diseases. The results reported that the introduced method offered an accuracy upto 9428%, recall of 96% and F1-score around 96%. Moreover, this technique was adaptable to predict the heart disease in comparison with existing methods.

R. Sonawane, et.al (2021) investigated the hybrid meta-heuristic concept to extract the optimized features relied on J-RDA [16]. The attributes were optimized to construct a hybrid technique in which optimized DBSCAN was integrated with KMC (K-Means Clustering) to tune the important metrics. The objective technique optimized the attributes and the constructed technique predicted the heart disease for tackling the multi-objective function. The results exhibited the efficiency of the presented approach for predicting the heart disorders.

H. Hasanova, et.al (2022) developed a ML based SCA_WKNN model to predict the coronary disorder for learning from the data available in blockchain [17]. This data was employed as authentic source to learn the data and also as a safe storage atmosphere to acquire the patient information. Diverse metrics namely accuracy,

and RMSE were considered to compute the developed algorithm. The results reported that the developed algorithm led to enhance the accuracy around 15.61%.

F. Ali, et.al (2020) projected a smart healthcare model for predicting the coronary disorder on the basis of ensemble DL and FF techniques [18]. At first, the FF technique emphasized on integrating the attributes of sensor data and EMRs for generating valued medical data. Thereafter, IG method assisted in eliminating the inappropriate and dismissed attributes, and selecting the significant data to mitigate the computing burden and enhance the system efficacy. In the end, the projected system computed for predicting the coronary disorder. The experimental results revealed that the accuracy of the projected system was counted 98.5%.

3. Proposed Approach for Heart Disease Prediction (HDP)

HD is often diagnosed by a doctor after reviewing the patient's medical history, the results of their physical exam, and any concerning symptoms. However, the results of this method of diagnostic do not reliably identify HD patients. Additionally, it is costly and computationally challenging to assess. To tackle these problems, it is necessary to create a non-invasive diagnosis system based on machine learning (ML) classifiers [19]. The ratio of deaths falls as a result of the implementation of expert technique used to make decision on the basis of machine learning classifiers and the ability to accurately diagnose HD. A visual representation of a general framework to predict the coronary disease represented in Figure 1.

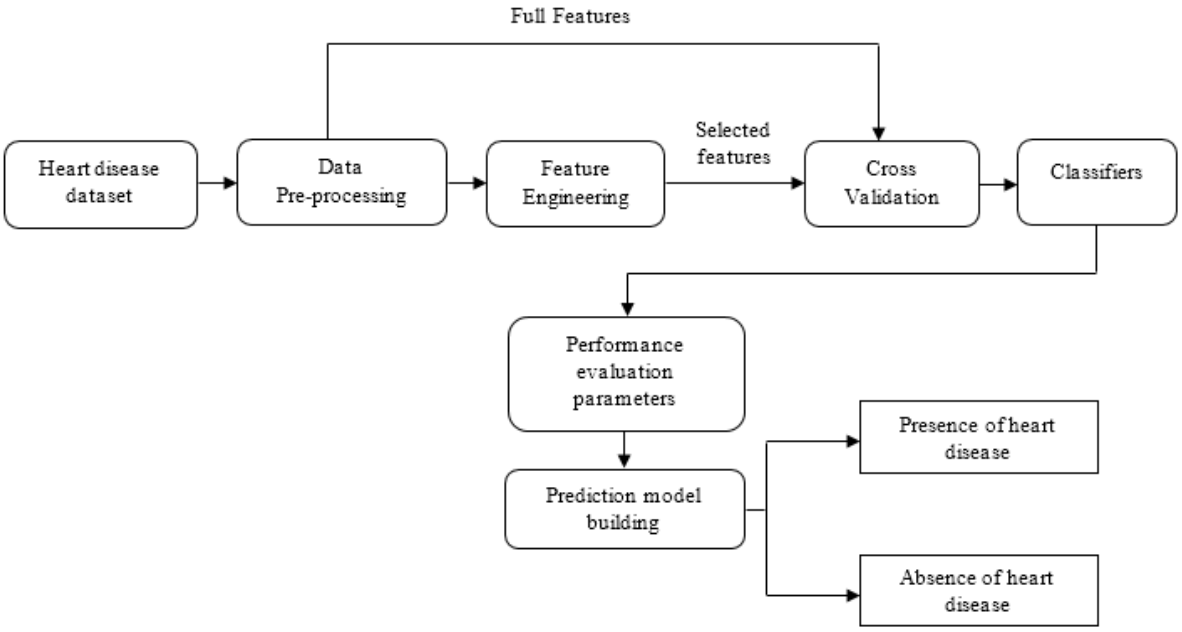

Figure 1: Heart Disease Prediction Model

The first step in HDP is data collecting, which involves obtaining, assessing, and carefully reviewing research visions. Regardless of the topic of the study, gathering data is typically the first and most important step [20]. The next step in the data processing process is to turn raw data into expressive patterns. This refers to a crucial stage in machine learning since the meaningful information that can be deduced from the input and its quality directly affect the model's capacity for learning. After the pre-processing of the data is complete, features are chosen for additional processing. In general, feature engineering is crucial for creating a foundation for classifiers. This process, also known as data transformation, transforms the data points as a whole in addition to choosing the relevant qualities or features. This phase aims to reduce the number of features that are input into a classifier in order to provide effective predictive frameworks with simpler computational requirements [21]. Validating the prediction model is an important stage in

machine learning. Leave-one-subject-out is a well-known cross validation strategy (LOSO). A single instance is used as test data in this validation strategy, and the remaining individuals are divided for the model training. The test subject must be expected to have HD in order for the person to be labelled as normal.

To predict the coronary disease can be seen as a classification or clustering issue from the perspective of automated learning. In the fourth stage of heart disease prediction, the retrieved features are categorized. Classification is vitally significant in a data mining challenge. Data that has been abstracted must be divided into distinct classes and components. The index's size and content determine how the partitioning is classified. The goal of classification is to find and categorise the class in order to increase prediction accuracy for concealed specific objects of data class. Only after separating the data class objects for each class framework is the prediction for each class model complete [22].

In the performance evaluation step, classifier performance is assessed using a variety of performance evaluation metrics. The confusion matrix is employed in the computation of these parameters. To gauge the effectiveness of the applied classification method, a variety of evaluation criteria are available, such as accuracy, sensitivity, specificity, recall, f1-score, Matthew correlation-coefficient (MCC), AUC-score, and ROC curve. The predictive framework is created utilising both conventional and statistical methods and ML algorithms to predict coronary disorder [23]. The framework discrimination and predictive limits of various techniques are also contrasted in this step.

3.1 Machine learning Models for Heart Disease Prediction

To create prediction frameworks, ML community has developed a number of methods based on multi-variables and their plentiful non-linear communications. The widely used ML algorithms can produce beneficial outcomes for heart disease prediction [24]. Machine learning algorithms manage non-linear relationships between existing data with equal computing weight, which is their main advantage. The most popular machine learning (ML) algorithms for predicting cardiac disease are listed below.

i. K-Nearest Neighbors (KNN): It is a robust method which is capable of classifying the data. This algorithm has assumption that same objects are adjacent and are useful for discovering its k-nearest neighbors for similarity. The space vector having multidimensionality elements, having their class name are considered as the training examples. The testing sample feature vectors and class labels comprise to train the K-NN (K-Nearest Neighbors) algorithm. k employs to define a user-defined constant while classifying the data. A mark is assigned to the well-known sample among k samples to recognize an unlabeled vector. The dataset is responsible for selecting the k. In general, the highest k values are capable of alleviating the effect of noise while classifying the data and inserting the boundaries among classes which are least distinctive. The maximization of k indicates the accuracy and reliability of the forecasts because of majority or average voting. This algorithm is implemented easily and found least complex [25]. There is not any necessity of creating a framework, changing various metrics, or generating the supplementary assumptions.

ii. Support Vector Machine (SVM): This model utilizes a hyper-plane for classifying the data points. The data points are differentiated using diverse kinds of hyper-planes. The major focus of this algorithm is on discovering a plan having a maximum range which implies the highest gap among the data points in both groups. The gap is increased from the margins to offer some clarification for recognizing the data points at superior accuracy. Hyper-planes are kind of boundaries implement to make the decision and assist in distinguishing the data points. The allocation of data points available in any side of hyperplane is done to diverse groups. Most frequently, the dimensionality of the hyper-plane is depending upon the amount of features. In case the number of features for the input is found 3, it results in making the hyper-plane to a two dimensional plane. In case the number of attributes are available in more amount, this is

challenging to understand the scenario. This algorithm is cost effective and feasible [26].

iii. Naive Bayes: This model is planned on the basis of a set of Bayes' Theorem. This algorithm is termed as naïve due to the assumption of independence among attributes which are utilized to classify the data and it is not realistic in real time. This algorithm helps in dividing the data into 2 elements: feature matrix and response vector. The initial is employed to comprise the entire data collected in the form of vectors (rows) at which every row is used to illustrate the relative variable kind, and the second one deploys rows for demonstrating an outcome class. This algorithm classifies the RoI (region of interest) from the data into secure and under-threat regions. Naïve Bayes algorithm predicts the data accessibly on the basis of some linear input metrics in a learning problem. The maximum-likelihood method implements to train the dataset and assess a closed-form formulation, with regard to the linear time. The fundamental intend of this algorithm is to combine the Naïve Bayes probability framework into a DR (decision rule) [27]. The maximum posteriori or MAP-DR is considered as a common rule in order to choose a feasible assumption. A similar classification algorithm called Bayes is defined as a function for allocating a class label $y = C_k, (kv\{1, \dots, k\})$ for some k as:

$$y = armax\ p(C_k) \prod_{i=1}^{n} p\big((x_i|C_k)\big)$$

iv. Random Forest: This algorithm exploits DT (decision trees) due to their applicability on the specific data in training. In DT, after updating the training data, no similar result can be found in DT. These factors consume higher cost and lead to cause a risk of overfitting and focus on exploring the local optima as they are ineffective to go back when the data is portioned. Such kind of drawbacks of decision trees are resolved using RF (Random Forest) algorithm [28]. This algorithm integrates various models for training the DTs so that an output is acquired. To illustrate, for a variable x_i, the permutation technique helps in replacing all x_i with a random value and classifying the permutation as noise. Hence, the original relation amid$x_i$and the result Y is broken. In the meantime, the Gini coefficient is exploited to determine the effective variables for mitigating the purity decrease of a predictor. The data dimensionality is lessened through the RFCY (random forest cross-validation) to select the attributes so that the variable number is verified[29-35].

4. Research Methodology

Heart disease poses a hazard to human life since it is more contagious and has a greater mortality rate. The difficult issue is to identify and treat heart diseases when they are in their early stages. However, due to inadequate access to diagnostic facilities, qualified medical personnel, and other resources, the process of anticipating heart disorders is hampered. Any sort of heart disease affects other bodily parts and causes disturbance[36-50]. To forecast the heart condition, researchers use computer-aided data extraction from massive amounts of records. Many organizations use DM (Data Mining) algorithms and procedures. The DM algorithm is beneficial for disease prediction in the healthcare industry. There are numerous risk factors that can result in heart disease.

Heart conditions are foreseen at several stages, which are as follows:

A. Data Gathering: This process is used to gather data while taking into account various clinical relationships. These data are used to run the tests.

B. Data pre-processing: In this level, the data analysis and wholeness are achieved through ML models. Additionally, the primary goal is to process the data in order to remove redundant features from the dataset. As a result, the training system is improved to clean and de-noise the data that is disseminated during the feature selection phase.

C. Feature selection: A subgroup with distinctive qualities is useful for detecting heart conditions. These attributes encourage us to address the current class of attributes. the major intend of RF algorithm is to choose the features. In this algorithm, the estimator value is given a value of

100. It creates a tree structure that contains all of the relevant features. This algorithm can be modified to choose the pertinent features for identifying coronary diseases.

D. Classification: The chosen attributes are mapped into the training system using this process, which will classify the supplied attributes. As a result, it is now able to accurately forecast the coronary problem. Each distinct class is used to represent a certain type of cardiac condition. The collected characteristics are used as input in the LR (logistic regression) algorithm, which uses them to classify Heart disease. To indicate the likelihood of a person having a heart condition, this research effort divides the data into two categories: heart disease and healthy.

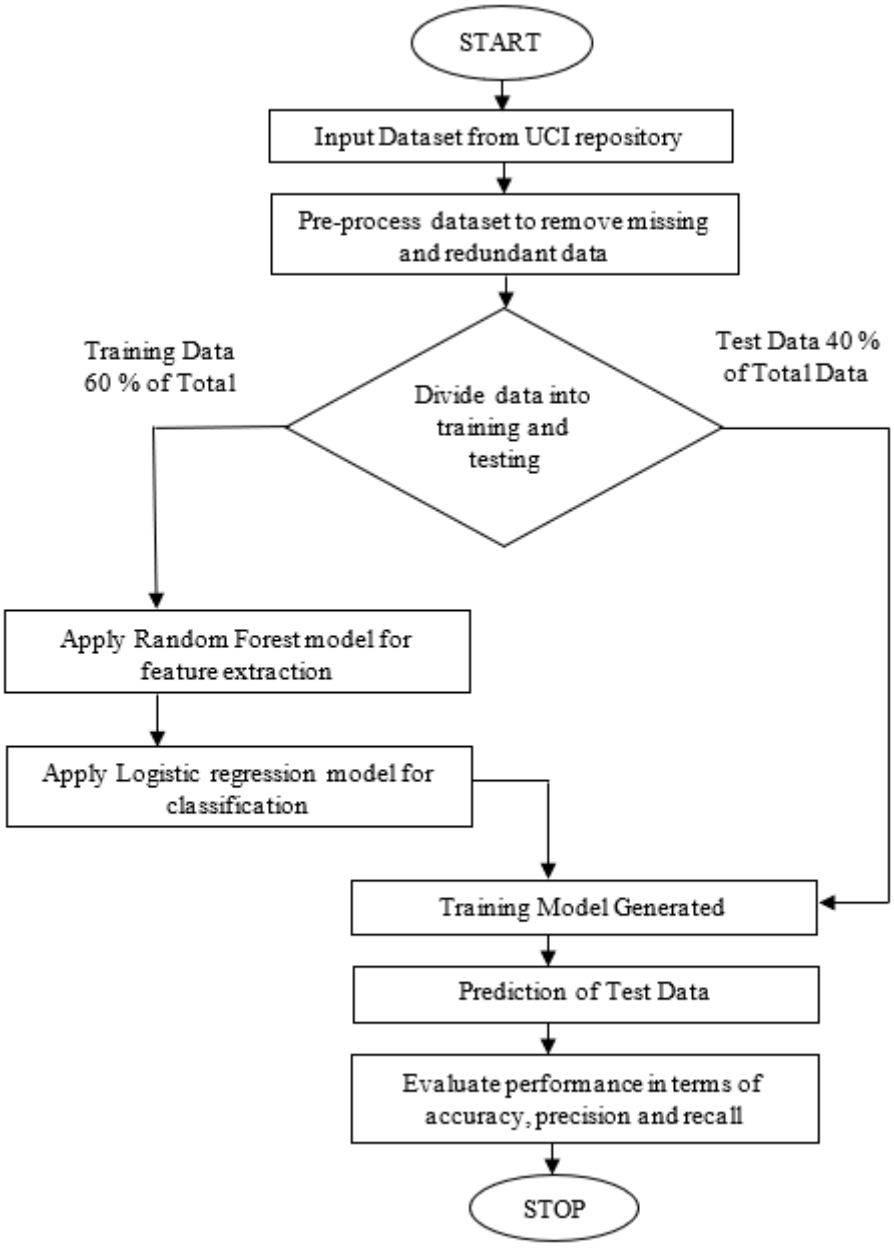

Figure 1: Proposed Methodology

5. Result and Discussion

This work generates the data set from Cleveland. This dataset is employed to predict the CVD (cardiovascular disease). There are 14 attributes comprised in this. This major focus of this project is to implement and compare diverse methods to predict the coronary diseases. A comparison of DT (Decision Tree), NB (Naïve Bayes), MLP (Multi-layer Perceptron) and Ensemble

technique is done with the introduced mechanism concerning some evaluation measures.

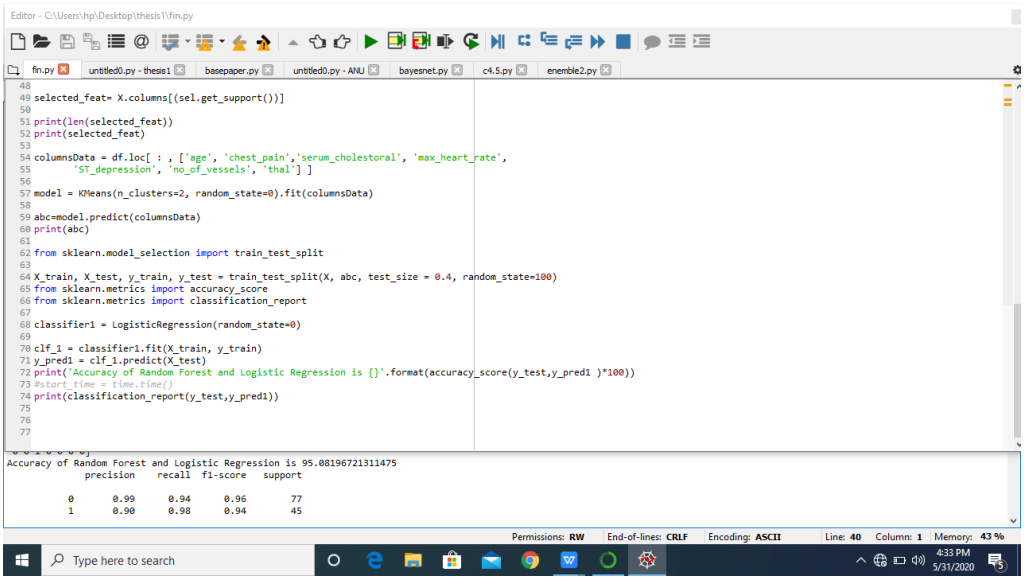

Figure 2: Presented Classifier (Performance of proposed HDPS on logistic regression)

Figure 2 demonstrates the deployment of Cleveland dataset to predict the disorders in heart. The introduced mechanism is implemented for predicting the coronary diseases.

Table 1: Results of Decision Tree

Class	Precision	Recall	F1-Score
0	0.79	0.71	0.75
1	0.66	0.74	0.79

Table 1 displays the results of evaluating the performance of the Decision Tree classifier for the two distinct classes based on precision, recall, and F1-score

Table 2: Results of Naïve Bayes

Class	Precision	Recall	F1-Score
0	0.87	0.90	0.89
1	0.86	0.82	0.84

Table 2 displays the results of evaluating the performance of the Naïve Bayes classifier for the two distinct classes based on precision, recall, and F1-score.

Table 3: Results of Multi-layer

Class	Precision	Recall	F1-Score
0	0.79	0.84	0.84
1	0.82	0.69	0.75

Table 3 displays the results of evaluating the performance of the multi-layer classifier for the two distinct classes based on precision, recall, and F1-score.

Table 4: Ensemble Model

Class	Precision	Recall	F1-Score
0	0.86	0.84	0.84
1	0.84	0.85	0.82

Table 4 displays the results of evaluating the performance of the Ensemble Model classifier for the two distinct classes based on precision, recall, and F1-score.

Table 5: Proposed HDP Model

Class	Precision	Recall	F1-Score
0	0.94	0.93	0.94
1	0.95	0.95	0.95

Table 5 displays the results of evaluating the performance of the Proposed Model for the two distinct classes based on precision, recall, and F1-score.

Table 6: Overall Results

Model	**Accuracy**	**Precision**	**Recall**	**F1-Score**
Decision Tree	75.41 Percent	75 Percent	75 Percent	75 Percent
Naïve Bayes	83.61 Percent	84 Percent	84 Percent	84 Percent

Multilayer Perceptron	83.61 Percent	85 Percent	84 Percent	84 Percent
Ensemble Model	85.25 Percent	86 Percent	85 Percent	86 Percent
Proposed Model	95.08 Percent	95 Percent	95 Percent	95 Percent

Table 6 shows the aggregate performance of all classifiers, and it is clear from the assessment measures that the proposed model fares quite well compared to existing model.

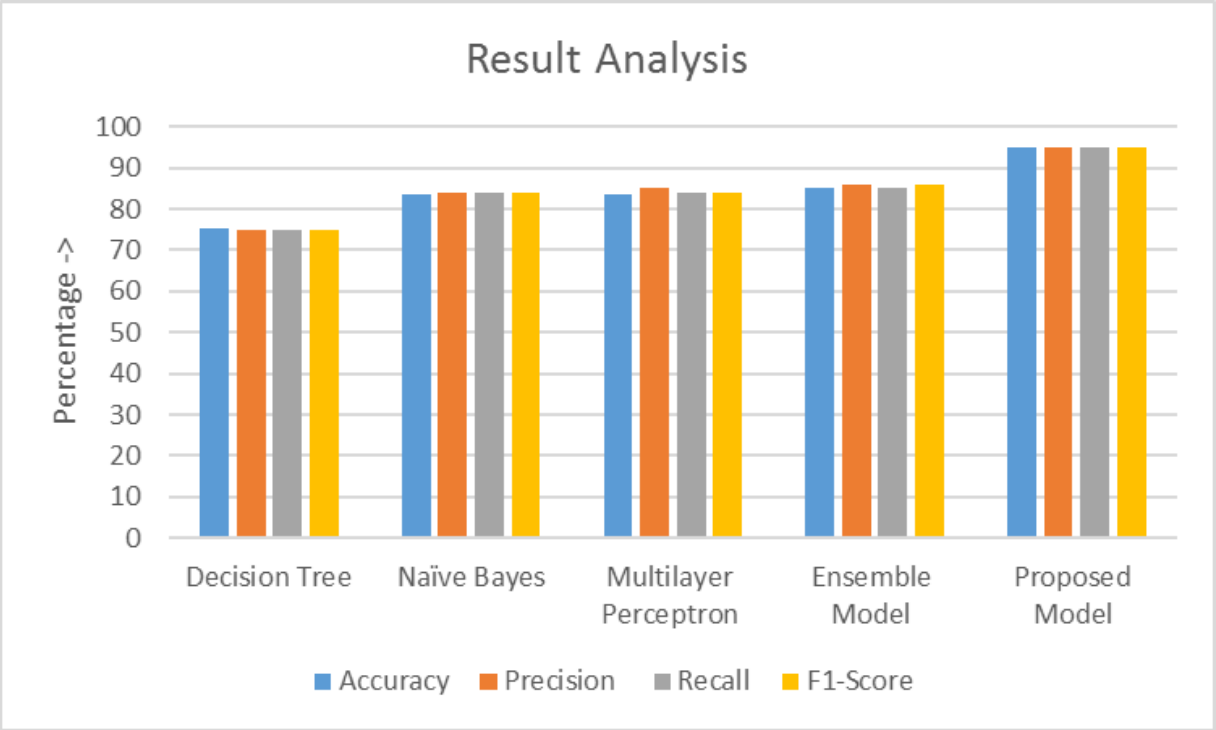

Figure 3: Overall Result Analysis

Figure 3 illustrates the comparison of overall results of the suggested model against the existing models like DT, NB, Multilayer Perceptron and ensemble model. The suggested approach achieves maximum accuracy of 95% for predicting the coronary disorder which is approx. 5 percent higher than the existing heart disease prediction models.

6. Conclusion

The coronary diseases become the major cause of severe health condition and various individuals are affected due to this disorder. The breathing issues, body paleness and inflammation in the legs etc. are some major indicators of coronary disorder. Researchers emphasizes on discovering an effective methodology to recognize the heart ailments, as existent diagnostic techniques for heart illness. However, these methods are ineffective to detect this disease at the primary phase because of various reasons concerning accuracy and execution time. Detecting and curing the coronary infections is a complex task in case of absence of advanced technology and medical professionals. This work indicates that the fundamental objective is of predicting the coronary disorder which becomes complicated in the presence of enormous amount of properties. A testing is performed on diverse methods such as DT (Decision Tree), NB, MLP and Ensemble technique for predicting the coronary disease. A novel mechanism is introduced that employed RF and LR algorithms for predicting CVD (cardiovascular diseases). The disease is classified to extract the attributes from RF and Logistic Regression algorithms. The

results indicate that the introduced mechanism offers the accuracy, recall and precision of 95%.

References

[1] K. Gomathi, Dr. Shanmugapriyaa, "Heart Disease Prediction Using Data Mining Classification", 2016, International Journal for Research in Applied Science & Engineering Technology (IJRASET), Volume 4 Issue II

[2] H. Benjamin Fredrick David and S. Antony Belcy, "Heart Disease Prediction using Data Mining Techniques", 2018, ICTACT Journal on Soft Computing, Volume: 09, Issue: 01

[3] D. P. Yadav, P. Saini and P. Mittal, "Feature Optimization Based Heart Disease Prediction using Machine Learning," 2021 5th International Conference on Information Systems and Computer Networks (ISCON), 2021, pp. 1-5

[4] M. Chakarverti, S. Yadav and R. Rajan, "Classification Technique for Heart Disease Prediction in Data Mining," 2019 2nd International Conference on Intelligent Computing, Instrumentation and Control Technologies (ICICICT), 2019, pp. 1578-1582

[5] R. Latha and P. Vetrivelan, "Blood Viscosity based Heart Disease Risk Prediction Model in Edge/Fog Computing," 2019 11th International Conference on Communication Systems & Networks (COMSNETS), 2019, pp. 833-837

[6] X. Yuan, J. Chen, K. Zhang, Y. Wu and T. Yang, "A Stable AI-Based Binary and Multiple Class Heart Disease Prediction Model for IoMT," in IEEE Transactions on Industrial Informatics, vol. 18, no. 3, pp. 2032-2040, March 2022

[7] M. Nahiduzzaman, M. J. Nayeem, M. T. Ahmed and M. S. U. Zaman, "Prediction of Heart Disease Using Multi-Layer Perceptron Neural Network and Support Vector Machine," 2019 4th International Conference on Electrical Information and Communication Technology (EICT), 2019, pp. 1-6

[8] S. Bhoyar, N. Wagholikar, K. Bakshi and S. Chaudhari, "Real-time Heart Disease Prediction System using Multilayer Perceptron," 2021 2nd International Conference for Emerging Technology (INCET), 2021, pp. 1-4

[9] M. A. Khan, "An IoT Framework for Heart Disease Prediction Based on MDCNN Classifier," in IEEE Access, vol. 8, pp. 34717-34727, 2020

[10] P. Ramprakash, R. Sarumathi, R. Mowriya and S. Nithyavishnupriya, "Heart Disease Prediction Using Deep Neural Network," 2020 International Conference on Inventive Computation Technologies (ICICT), 2020, pp. 666-670

[11] S. Hussain, S. K. Nanda, S. Barigidad, S. Akhtar, M. Suaib and N. K. Ray, "Novel Deep Learning Architecture for Predicting Heart Disease using CNN," 2021 19th OITS International Conference on Information Technology (OCIT), 2021, pp. 353-357

[12] H. Tiwari, "Early prediction of heart disease using deep learning approach", Deep Learning for Medical Applications with Unique Data, vol. 7, pp. 120475-120499, 18 February 2022

[13] W. Li, D. Chen and J. Le, "Coronary Heart Disease Prediction Based on Combined Reinforcement Multitask Progressive Networks," 2020 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), 2020, pp. 311-318

[14] M. T. Islam, S. R. Rafa and M. G. Kibria, "Early Prediction of Heart Disease Using PCA and Hybrid Genetic Algorithm with k-Means," 2020 23rd International Conference on Computer and Information Technology (ICCIT), 2020, pp. 1-6

[15] D Deepika and N. Balaji, "Effective heart disease prediction using novel MLP-EBMDA approach", Biomedical Signal Processing and Control, vol. 10, no. 1, pp. 41720-41739, 17 November 2021

[16] R. Sonawane and H. Patil, "Automated heart disease prediction model by hybrid heuristic-based feature optimization and enhanced clustering", Biomedical Signal Processing and

Control, vol. 19, no. 2, pp. 1165-1172, 9 November 2021

[17] H. Hasanova, M. Tufail and M-S. Kim, "A novel blockchain-enabled heart disease prediction mechanism using machine learning", Computers and Electrical Engineering, vol. 10, pp. 9340-9356, 16 May 2022

[18] F. Ali, S. El-Sappagh and K-Sup. Kwak, "A smart healthcare monitoring system for heart disease prediction based on ensemble deep learning and feature fusion", Information Fusion, vol. 9, no. 4, pp. 75801-75825, 26 June 2020

[19] J. Vijayashree and N.Ch. Sriman Narayana Iyengar, "Heart Disease Prediction System Using Data Mining and Hybrid Intelligent Techniques: A Review", 2016, International Journal of Bio-Science and Bio-Technology, Vol.8, No.4, pp. 139-148

[20] D. Ratnam, P. Hima Bindu, V. Mallik Sai, S. P. Rama Devi and P. Raghavendra Rao, "Computer-Based Clinical Decision Support System for Prediction of Heart Diseases Using Naïve Bayes Algorithm", 2014, International Journal of Computer Science and Information Technologies, vol. 5, no. 2, pp.2384- 2388

[21] T. Santhanam and E. P. Ephzibah, "Heart Disease Prediction Using Hybrid Genetic Fuzzy Model", 2015, Indian Journal of Science and Technology, vol.8, no. 9, pp.797–803

[22] G. Purusothaman and P. Krishnakumari, "A Survey of Data Mining Techniques on Risk Prediction: Heart Disease", 2015, Indian Journal of Science and Technology, vol. 8, no. 12

[23] K. Srinivas, G. Raghavendra Rao and A. Govardhan, "Analysis of Coronary Heart Disease and Prediction of Heart Attack in Coal Mining Regions Using Data Mining Techniques", 2010, Proceedings of 5th International Conference on Computer Science & Education, China, pp. 24–27

[22] K. Thenmozhi and P. Deepika, "Heart Disease Prediction Using Classification with Different Decision Tree Techniques", 2014, International Journal of Engineering Research and General Science, vol. 2, no. 6, pp.6-11

[23] S. Chaitrali, Dangare and S. Apte, "Improved Study of Heart Disease Prediction System using Data Mining Classification Techniques", 2012, International Journal of Computer Applications, vol. 47, no. 10, pp. 0975 –888

[24] J Peter and K. Somasundaram, "An Empirical Study on Prediction of Heart Disease Using Classification Data Mining Techniques", 2012, Proceedings of IEEE International Conference on Advances in Engineering, Science and Management (ICAESM), pp. 514-518.

[25] H. D. Masethe and M. A. Masethe, "Prediction of Heart Disease using Classification Algorithms", 2014, Proceedings of the World Congress on Engineering and Computer Science (WCECS), San Francisco, USA

[26] Jyoti Soni, Ujma Ansari, Dipesh Sharma, Sunita Soni, "Predictive Data Mining for Medical Diagnosis: An Overview of Heart Disease Prediction", 2011, International Journal of Computer Applications (0975 – 8887) Volume 17, No.8

[27] Asha Rajkumar, G. Sophia Reena, "Diagnosis of Heart Disease Using Datamining Algorithm", 2010, Global Journal of Computer Science and Technology, Vol. 10 Issue 10

[28] Sunita Soni, O.P.Vyas, "Using Associative Classifiers for Predictive Analysis in Health Care Data Mining", 2010, International Journal of Computer Application, Volume 4, No.5, pages 33-34

[29] Kaur, M.; Singh, A.; Verma, S.; Jhanjhi, N.Z.; Talib, M.N. FANET: Efficient routing in flying ad hoc networks (FANETs) using firefly algorithm. In Intelligent Computing and Innovation on Data Science Lecture Notes in Networks and Systems; Springer: Singapore, 2021; Volume 248, pp. 483–490

[30] Arora, M.; Verma, S.; Kavita; Wozniak, M.; Shafi, J.; Ijaz, M.F. An efficient ANFIS-EEBAT

approach to estimate effort of Scrum projects. Sci. Rep. 2022, 12, 7974.

[31] R. Dogra, S. Rani, H. Babbar, S. Verma, K. Verma and J. J. P. C. Rodrigues, "DCGCR: Dynamic Clustering Green Communication Routing for Intelligent Transportation Systems," in IEEE Transactions on Intelligent Transportation Systems, vol. 23, no. 9, pp. 16197-16205, Sept. 2022, doi: 10.1109/TITS.2022.3148471.

[32] Kumar, M.; Mukherjee, P.; Verma, S.; Shafi, J.; Wozniak, M.; Ijaz, M.F. A smart privacy preserving framework for industrial IoT using hybrid meta-heuristic algorithm. Sci. Rep. 2023, 13, 5372.

[33] Kumar, Ashwani, Mohit Kumar, Rajendra Prasad Mahapatra, Pronaya Bhattacharya, Thi-Thu-Huong Le, Sahil Verma, Kavita, and Khalid Mohiuddin. 2023. "Flamingo-Optimization-Based Deep Convolutional Neural Network for IoT-Based Arrhythmia Classification" Sensors 23, no. 9: 4353. https://doi.org/10.3390/s23094353

[34] Kaur, M.; Verma, S.; Kavita. Flying Ad-Hoc Network (FANET): Challenges and Routing Protocols. J. Comput. Theor. Nanosci. 2020, 17, 2575–2581

[35] M. Kumar, Kavita, S. Verma, A. Kumar, M. F. Ijaz and D. B. Rawat, "ANAF-IoMT: A Novel Architectural Framework for IoMT-Enabled Smart Healthcare System by Enhancing Security Based on RECC-VC," in IEEE Transactions on Industrial Informatics, vol. 18, no. 12, pp. 8936-8943, Dec. 2022, doi: 10.1109/TII.2022.3181614.

[36] Hossam M. Moftah , Taha M. Mohamed, A Novel Fuzzy Bat Based Ambulance Detection and Traffic Counting Approach, Journal of Cybersecurity and Information Management, Vol. 1 , No. 2 , (2020) : 41-54 (Doi : https://doi.org/10.54216/JCIM.010203)

[37] Ashish Sharma , Yogesh Sharma , Radhika Bansal , Sushant Verma, Implementation of Crowd Sale using ERC-20 Tokens, Journal of Cybersecurity and Information Management, Vol. 2 , No. 1 , (2020) : 05-12 (Doi : https://doi.org/10.54216/JCIM.020101)

[38] Ashish Sharma, Sandeep Tayal, Radhika Bansal, and Sushant Verma, Energy Efficiency Techniques in Heterogeneous Networks, Journal of Cybersecurity and Information Management, Vol. 2 , No. 1 , (2020) : 13-19 (Doi : https://doi.org/10.54216/JCIM.020102)

[39] Ashish Sharma, Yogesh Sharma, Radhika Bansal, and Sushant Verma, Empirical Study of Function Point Analysis during Software Development Phase, Journal of Cybersecurity and Information Management, Vol. 2 , No. 1 , (2020) : 20-24 (Doi : https://doi.org/10.54216/JCIM.020103)

[40] Nada M. Alhakkak, A Validation Model for ERP systems, Journal of Cybersecurity and Information Management, Vol. 2 , No. 1 , (2020) : 25-34 (Doi : https://doi.org/10.54216/JCIM.020104)

[41] Hussain, I., Tahir, S., Humayun, M., Almufareh, M. F., Jhanjhi, N. Z., & Qamar, F. (2022, November). Health monitoring system using internet of things (iot) sensing for elderly people. In 2022 14th International Conference on Mathematics, Actuarial Science, Computer Science and Statistics (MACS) (pp. 1-5). IEEE.

[42] Almuayqil, S. N., Humayun, M., Jhanjhi, N. Z., Almufareh, M. F., & Javed, D. (2022). Framework for improved sentiment analysis via random minority oversampling for user tweet review classification. Electronics, 11(19), 3058.

[43] Humayun, M., Niazi, M., Almufareh, M. F., Jhanjhi, N. Z., Mahmood, S., & Alshayeb, M. (2022). Software-as-a-service security challenges and best practices: A multivocal literature review. Applied Sciences, 12(8), 3953.

[44] Humayun, M., Jhanjhi, N. Z., & Almotilag, A. (2022). Real-time security health and privacy monitoring for Saudi highways using cutting-edge technologies. Applied Sciences, 12(4), 2177.

[45] Jindal, U., Dalal, S., Rajesh, G., Sama, N. U., Jhanjhi, N. Z., & Humayun, M. (2022). An

integrated approach on verification of signatures using multiple classifiers (SVM and Decision Tree): A multi-classification approach. Int. J. Adv. Appl. Sci., 9(1), 99-109.

[46] Jindal, U., Dalal, S., Rajesh, G., Sama, N. U., Jhanjhi, N. Z., & Humayun, M. (2022). An integrated approach on verification of signatures using multiple classifiers (SVM and Decision Tree): A multi-classification approach. Int. J. Adv. Appl. Sci., 9(1), 99-109.

[47] Nawaz, A. (2021). Feature engineering based on hybrid features for malware detection over Android framework. Turkish Journal of Computer and Mathematics Education (TURCOMAT), 12(10), 2856-2864.

[48] Shafiq, D. A., Jhanjhi, N. Z., & Abdullah, A. (2021, March). Machine learning approaches for load balancing in cloud computing services. In 2021 National Computing Colleges Conference (NCCC) (pp. 1-8). IEEE.

[49] Sangkaran, T., Abdullah, A., & JhanJhi, N. Z. (2020). Criminal network community detection using graphical analytic methods: A survey. EAI Endorsed Transactions on Energy Web, 7(26), e5-e5.

[50] A. Almusaylim, Z., Jhanjhi, N. Z., & Alhumam, A. (2020). Detection and mitigation of RPL rank and version number attacks in the internet of things: SRPL-RP. *Sensors*, *20*(21), 5997.