Journal of Theoretical and Applied Information Technology
31st December 2022. Vol.100. No 24

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

A NOVEL APPROACH BASED ON VOTING ENSEMBLE AND PCA DIMENSIONALITY REDUCTION METHOD FOR THE PREDICTION OF HEART DISEASE

SADIYAMOLE P A[1], Dr.S MANJU PRIYA[2]

[1]Research Scholar, Department of Computer Science,Karpagam Academy of Higher Education,Coimbatore 21,India

[2]Professor, Department of CS,Karpagam Academy of Higher Education,Coimbatore 21,India

E-mail: [1]sadiya.pa@gmail.com, [2]smanjupr@gmail.com

ABSTRACT

It is assumed that 32% of all deaths around worldwide are due to different types of CVDs.Advanced noticing and realizing heart diseases can be a boom to patients as they get a chance for switching their habits and life styles to a more healthy way and thus they can save their lives.Scientific researchers all over the world have been working on creating a much more intelligent decision support model for the early prediction of CVD.Healthcare people around the globe collects heart disease datasets .With the help of already available data and application of machine learning algorithms can be a cathartic feeling to the medical area.One of the main reasons of failure of intelligent prediction system is the inaccurate feature set and suffering of high overfitting and low variance.In order to predict the heart health of a patient in a more effective way,some machine learning models can be combined for better performance.Principal Component Analysis(PCA) is used to control the dimensionality of attributes.This paper mainly focuses on voting ensemble method for improving the performance of individual classifiers.

Keywords: *Machine Learning, Ensemble, PCA, Voting, LR, DT*

1. INTRODUCTION

Heart disease affects the normal working condition of human heart.Different types of heart diseases are Coronary heart disease, various heart rhythm problems, arrhythmia,cardio vascular disease,various chest pains,congenital heart defects, etc.Different heart disease symptoms contain chest pressure,discomfort of chest,breathing problems,fatigue, hearbeatabnormality, and various heart defects.The dreadful disease that affects heart chambers is heart failure.Coronary heart disease is the most ordinary form of heart disease.It happens when plaque clogging occurred in the arteries whose function is to supply blood to the heart.It reduces the blood supply and hence the oxygen received by heart lessens .It causes the weakening of heart muscle and thus lead to heart failure.Some persons have a heart problem by birth.This is known as congenital heart defects.Irregular heartbeat is known as arrhythmia.When there is a problem of pumping or relaxing function ,it cause congestive heart failure.All these types of heart failure affects the damage of whole or portion of heart.

There are various reasons for these heart failures.High blood pressure, high cholesterol,s mokingcondition,high intake of alcohol,over weight and obesity, anxiety,unhealthy food habits,diabetes and hereditary.World health organization[10] mentions that various heart diseases are the major cause of human loss in the world.If any type of heart failure occurs, the person immediately needs to be under medical attention.The number of affected people is increasing swiftly every year.Studies shown that about half percentage of patients who suffer heart disease die within the first three years of their illness. As well as the cost of treatment for heart disease is also increasing .

Traditional approach for identifying a person's heart health includes various body invasive tests like blood cholesterol, insulin level, ECG ,angiography,etc.Critical patients must undergo immediate cardiac surgery.Tests like angiography are costly, and it has some side effects too.Substandard tests will lead to calamitous consequences and hence they are inadmissible.Hospitals can also minimize the charge of different clinical tests by employing asuitable computer aided decision support

systems[1].Over the last few decades researchers have been continuously working on various machine learning techniques for the early diagnosis of heart disease[2].Data mining as well as improved deep learning methods can be used for identifying various diseases like cancer[3,4],diabetics [5]etc.

In this paper voting ensemble method along with PCA dimensionality reduction method is applied for the prediction of heart disease. This paper is divided into different sections: Section 2 deals with different literature surveys on this problem, section 3 deals with the methodology implemented, section 4 implies results and section 5 depicts the conclusion.

2. RELATED WORKS

Most of the researchers use Clevland heart disease dataset for study. Dr.S.V.Kogilvani[6] also have used Clevlanddataset.The authors have created some synthetic dataset since the original dataset contains small amount of data.Later they compared Clevland as well as synthetic dataset after preprocessing and the results are compared on the basis of accuracy, precision and recall and synthetic data performed well.

In[9],the authors have suggested an advanced preprocessing based method for the prediction of predict Coronary Heart Diseases (CHD).Random forest, K-nearest neighbor and Decision Tree are used for comparative study. Framingham Heart Study data set is used for the study.Random Forest performed well with an accuracy of 96.8%.

Terrada[11] al. have proposed a medical diagnosis system based on supervised ML technique that is applied on three databases(Clevland,Hungarian,Z-alisadeh) to predict atherosclerosis that is a reason for heart disease.DT,ANN and adaboost are used in three databases for comparison.With 94% accuracy ANN performed well along with Z-alisadeh dataset .Baccouche[12] et.al have suggested an ensemble based method that trains two DNN models on the same data.The research differentiate between a single CNN and ensembling of GRU, LSTM, BiLSTM and BiGRUmethods.Combination of BiGRU and CNN technique has got 96% accuracy.The authors in[13] have developed a web application and compared three ML methods DT,LR,NB and SVM.LR was the best performing algorithm with an accuracy of 82.89% out of all.

R.Indrakumari[14] et.al considered the main risk factors that affect heart disease and un supervised algorithm K-means clustering is applied on Clevland dataset and four types of chest pain are predicted with the help of Tableau visualization tool.Hashi[15] et.al have suggested Grid searchCv based hyper parameter optimization model which is applied on RF, KNN, LR, DT, and SVMclassifiers.The authors have compared normal dataset and dataset after tuning.The later performed well in all evaluation criterias.Sharma[16] et.al bring forward a modified backpropagation method by using CNN model and they claim that the prediction accuracy increases to 9% than the state of the art methods.In heart disease prediction area, many deep learning techniques have been implemented by many researchers[23] for the past couple of years.

3. MATERIALS AND METHODS

Many researchers have been working on different ensemble method [8]for improving the performance of prediction. This research paper mainly focuses on voting ensemble method. This part of the paper describes the dataset taken for study and the methodology applied.

3.1 Datasets

Latest researches [19] have shown that clinical feature based data modality methods perform well than the image based data.In this study, the Clevland data set[17] and KaggleClevland-Statlog-Hungary dataset[18] are taken.TheClevland dataset contains 303 records with 14 features including the output feature.The later dataset is a combination of Clevland(303), Hungarian(294), Switzerland(123), Longbeach VA(200) and Stalog(270) datasets.It contains a total of 1190 records with 12 features.Dataset contains health related data like ECG,sugarlevel,cholesterol,chestpain,ECG, etc.Some researchers create synthetic data as the number of datasets in the Clevland is small[7]. But the data can not be relied upon if artificial data are created much.In this study 70-30 train test split is used.TheClevlanddataset is shown in the Table 1.

Table 1. Clevland Dataset

#	Field	Detail
0	Age	Age of the patient.
1	sex	Patient's gender
2	cp	Chest pain value
3	Trestbps	Blood pressure while resting
4	Chol	Patient's Cholesterol
5	Fbs	Fasting Glucose level
6	Restecg	Resting ECG
7	Thalach	Achieved peak heart rate
8	Exang	exercise induced angina pain.(1 =positive; 0 = negative)
9	Oldpeak	ST depression
10	Slope	the slope of the peak exercise ST segment
11	Ca	number of main vessels coloured by fluoroscopy
12	Thal	3 = normal; 6 = fixed defect; 7 = reversible defect
13	Target	Heart disease or not

3.2 Data Preprocessing

The real data most probably contain some irrelevant information,redundant data and some noise .These data should be properly preprocessed to avoid all these problems.Relationshipof the field sex with target variable in Clevland is shown in Fig 1 and in Statlog Hungary dataset is shown in Fig 2.

The histogram of Clevland dataset is shown in Fig 3 and that of thesecond dataset is shown in Fig 4.

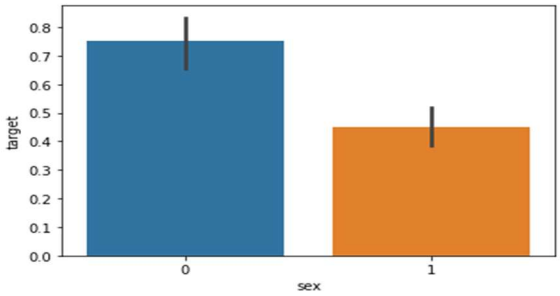


Figure 1.Relationship of sex with target

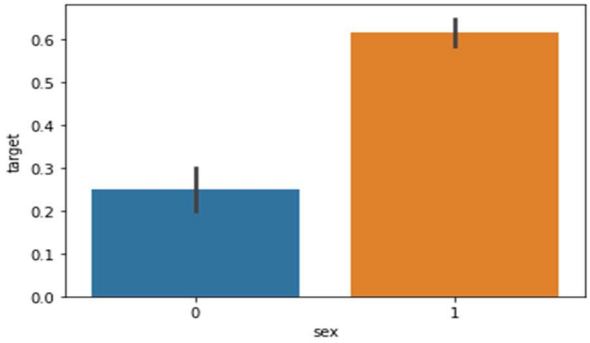


Figure 2.Relationship of age with target in Clevland-Statlog-Hungary--Longbeachdataset

Journal of Theoretical and Applied Information Technology
31st December 2022. Vol.100. No 24

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

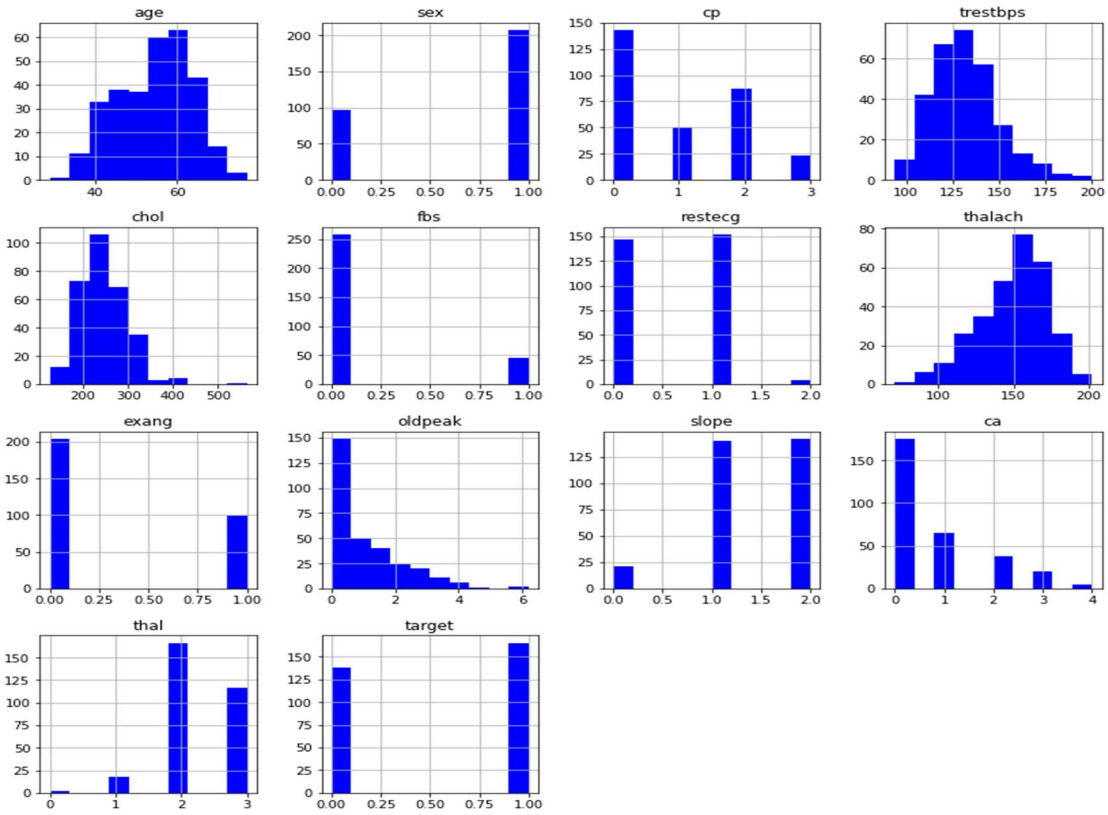


Figure 3.Histogram of Clevland dataset

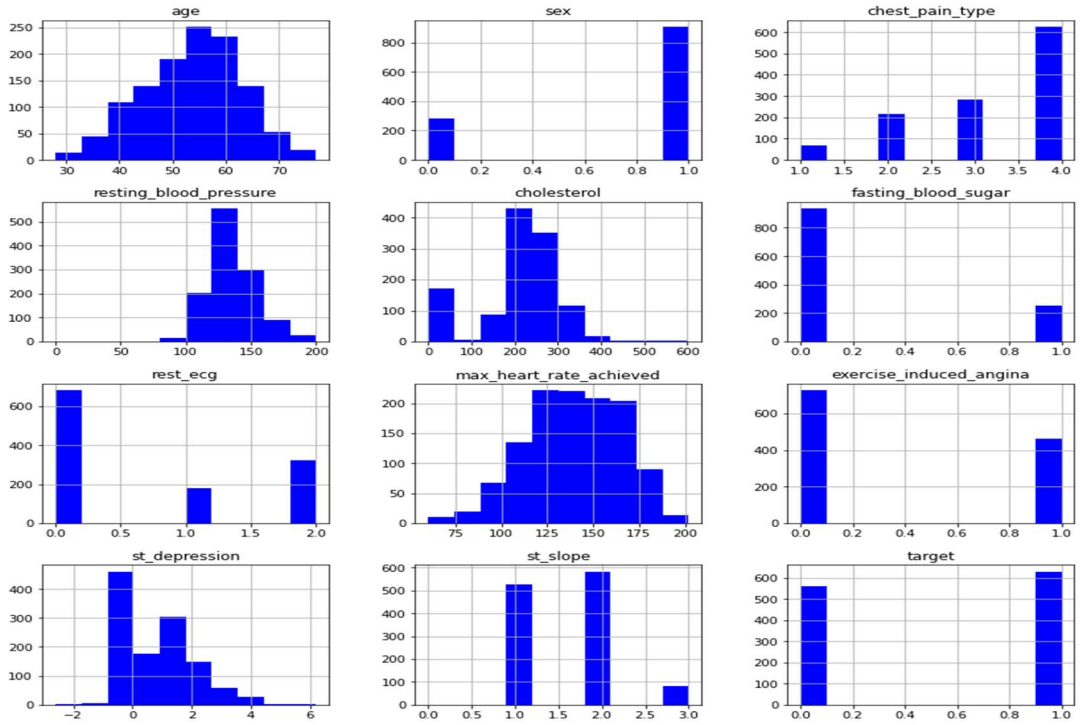
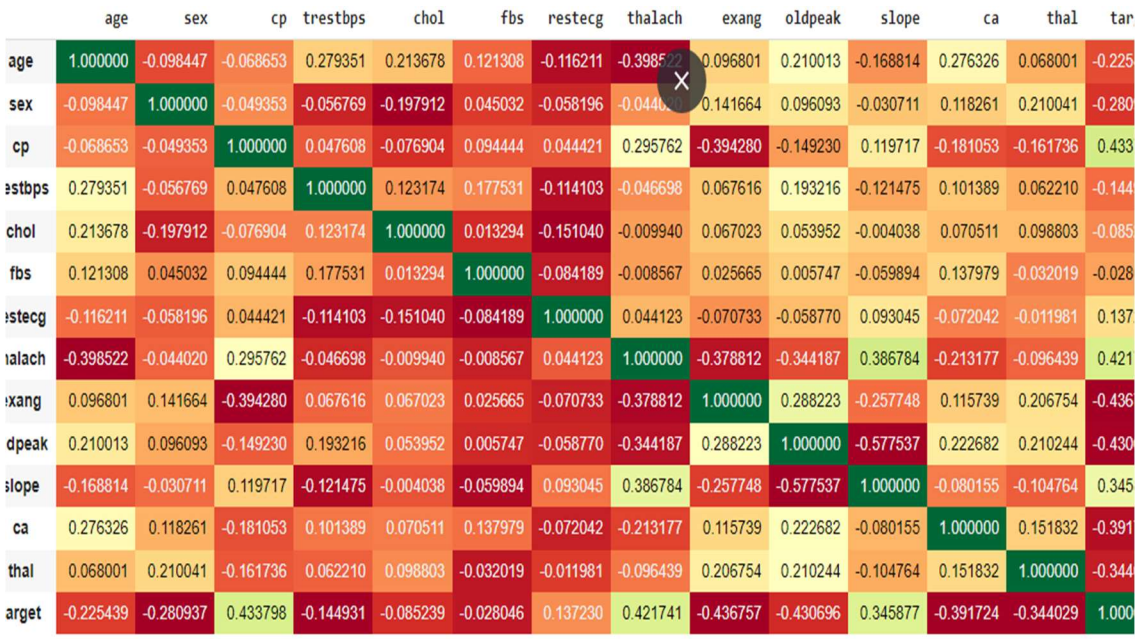


Figure 4.Histogram of Clevland-Statlog-Hungary-Longbeach.

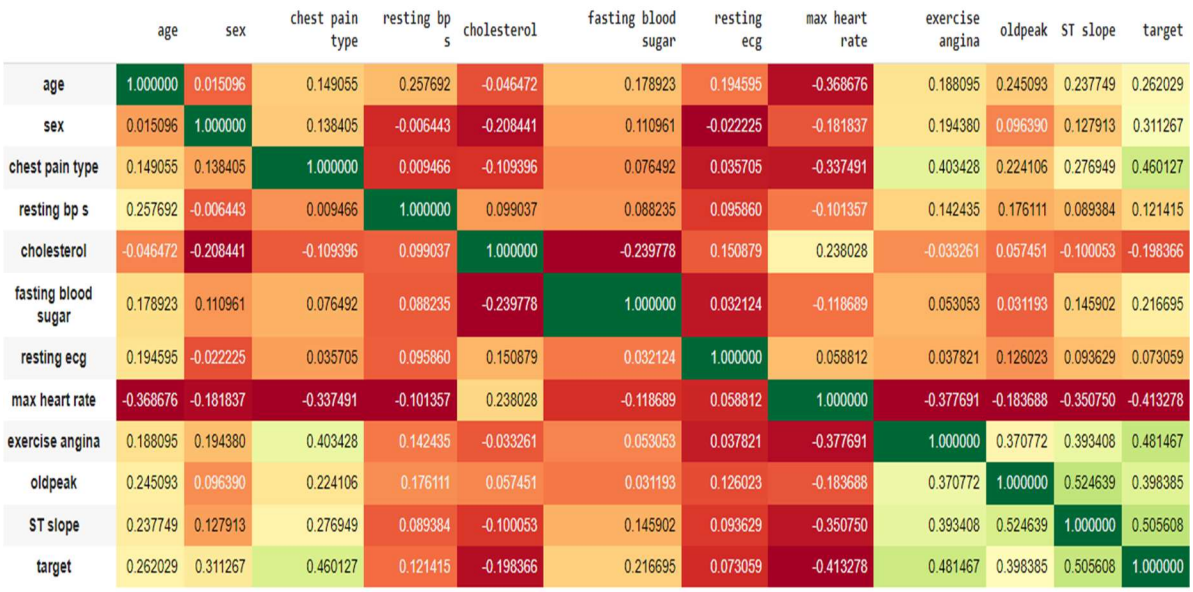
Journal of Theoretical and Applied Information Technology
31st December 2022. Vol.100. No 24

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195



	age	sex	cp	trestbps	chol	fbs	restecg	thalach	exang	oldpeak	slope	ca	thal	tar
age	1.000000	-0.098447	-0.068653	0.279351	0.213678	0.121308	-0.116211	-0.398	0.096801	0.210013	-0.168814	0.276326	0.068001	-0.225
sex	-0.098447	1.000000	-0.049353	-0.056769	-0.197912	0.045032	-0.058196	-0.044	0.141664	0.096093	-0.030711	0.118261	0.210041	-0.280
cp	-0.068653	-0.049353	1.000000	0.047608	-0.076904	0.094444	0.044421	0.295762	-0.394280	-0.149230	0.119717	-0.181053	-0.161736	0.433
estbps	0.279351	-0.056769	0.047608	1.000000	0.123174	0.177531	-0.114103	-0.046698	0.067616	0.193216	-0.121475	0.101389	0.062210	-0.144
chol	0.213678	-0.197912	-0.076904	0.123174	1.000000	0.013294	-0.151040	-0.009940	0.067023	0.053952	-0.004038	0.070511	0.098803	-0.085
fbs	0.121308	0.045032	0.094444	0.177531	0.013294	1.000000	-0.084189	-0.008567	0.025665	0.005747	-0.059894	0.137979	-0.032019	-0.028
estecg	-0.116211	-0.058196	0.044421	-0.114103	-0.151040	-0.084189	1.000000	0.044123	-0.070733	-0.058770	0.093045	-0.072042	-0.011981	0.137
alach	-0.398522	-0.044020	0.295762	-0.046698	-0.009940	-0.008567	0.044123	1.000000	-0.378812	-0.344187	0.386784	-0.213177	-0.096439	0.421
exang	0.096801	0.141664	-0.394280	0.067616	0.067023	0.025665	-0.070733	-0.378812	1.000000	0.288223	-0.257748	0.115739	0.206754	-0.436
dpeak	0.210013	0.096093	-0.149230	0.193216	0.053952	0.005747	-0.058770	-0.344187	0.288223	1.000000	-0.577537	0.222682	0.210244	-0.430
slope	-0.168814	-0.030711	0.119717	-0.121475	-0.004038	-0.059894	0.093045	0.386784	-0.257748	-0.577537	1.000000	-0.080155	-0.104764	0.345
ca	0.276326	0.118261	-0.181053	0.101389	0.070511	0.137979	-0.072042	-0.213177	0.115739	0.222682	-0.080155	1.000000	0.151832	-0.391
thal	0.068001	0.210041	-0.161736	0.062210	0.098803	-0.032019	-0.011981	-0.096439	0.206754	0.210244	-0.104764	0.151832	1.000000	-0.344
arget	-0.225439	-0.280937	0.433798	-0.144931	-0.085239	-0.028046	0.137230	0.421741	-0.436757	-0.430696	0.345877	-0.391724	-0.344029	1.000

Figure 5.Heatmap of Clevland dataset



	age	sex	chest pain type	resting bp s	cholesterol	fasting blood sugar	resting ecg	max heart rate	exercise angina	oldpeak	ST slope	target
age	1.000000	0.015096	0.149055	0.257692	-0.046472	0.178923	0.194595	-0.368676	0.188095	0.245093	0.237749	0.262029
sex	0.015096	1.000000	0.138405	-0.006443	-0.208441	0.110961	-0.022225	-0.181837	0.194380	0.096390	0.127913	0.311267
chest pain type	0.149055	0.138405	1.000000	0.009466	-0.109396	0.076492	0.035705	-0.337491	0.403428	0.224106	0.276949	0.460127
resting bp s	0.257692	-0.006443	0.009466	1.000000	0.099037	0.088235	0.095860	-0.101357	0.142435	0.176111	0.089384	0.121415
cholesterol	-0.046472	-0.208441	-0.109396	0.099037	1.000000	-0.239778	0.150879	0.238028	-0.033261	0.057451	-0.100053	-0.198366
fasting blood sugar	0.178923	0.110961	0.076492	0.088235	-0.239778	1.000000	0.032124	-0.118689	0.053053	0.031193	0.145902	0.216695
resting ecg	0.194595	-0.022225	0.035705	0.095860	0.150879	0.032124	1.000000	0.058812	0.037821	0.126023	0.093629	0.073059
max heart rate	-0.368676	-0.181837	-0.337491	-0.101357	0.238028	-0.118689	0.058812	1.000000	-0.377691	-0.183688	-0.350750	-0.413278
exercise angina	0.188095	0.194380	0.403428	0.142435	-0.033261	0.053053	0.037821	-0.377691	1.000000	0.370772	0.393408	0.481467
oldpeak	0.245093	0.096390	0.224106	0.176111	0.057451	0.031193	0.126023	-0.183688	0.370772	1.000000	0.524639	0.398385
ST slope	0.237749	0.127913	0.276949	0.089384	-0.100053	0.145902	0.093629	-0.350750	0.393408	0.524639	1.000000	0.505608
target	0.262029	0.311267	0.460127	0.121415	-0.198366	0.216695	0.073059	-0.413278	0.481467	0.398385	0.505608	1.000000

Figure 6.Heatmap of Clevland-Statlog-Hungary-Longbeach Dataset

All null values in the both datasets are eliminated. Most of the researchers use min max scalar for standardization[20]In order to make all features to acommon scale standardization is applied as feature scaling method.The equation of standardization is

$$Z=X-/X-\frac{\mu}{g}$$

where u is the mean and z is the standard deviation.

3.3 PCA

In ML technology, too much fields are not a good sign.In some datasets large number of features will lead to poor accuracy.In order to reduce the curse of dimensionality Principal Component Analysis is used as dimensionality reduction method[21].The main aim of PCA is to figure out the patterns and correlations among different features in the dataset.So all redundant data and inconsistent data should be removed.PCA is done through a series of steps.Let A is an n*m matrix and B is the resultant matrix after PCA .

Step1:Standardization of the data
Step 2:Compute mean values of all columns.M=mean(A)
Step 3: To center the values in each colum,subtract the mean value column ,C=A-M
Step 4: Compute covariance matrix of C,V=cov(C)
Step 5: Calculate eigen values and vectors of V,eig(V)
Step 6:Select q eigen vectors that have the q largest eigen values, known as principal components.
From the initial set of variables, the new set of variables obtained are Principal components.

3.4 Proposed Voting Ensemble Method

Ensemble methods are widely use in ML techniques[22].It combines various base models for producing one optimum predictive model.Different ensemble techniques are bagging,boosting , stacking and majority voting.In this research majority voting technique is applied for combining predictions from different other models.By using majority voting better performance can be achieved than using a single model.Voting ensemble method is used when all models have equal good performance.Here
Logistic Regression, Decision Tree, KNN and SVM are
used as base models.Voting ensemble method can be
shown in the Fig 7.

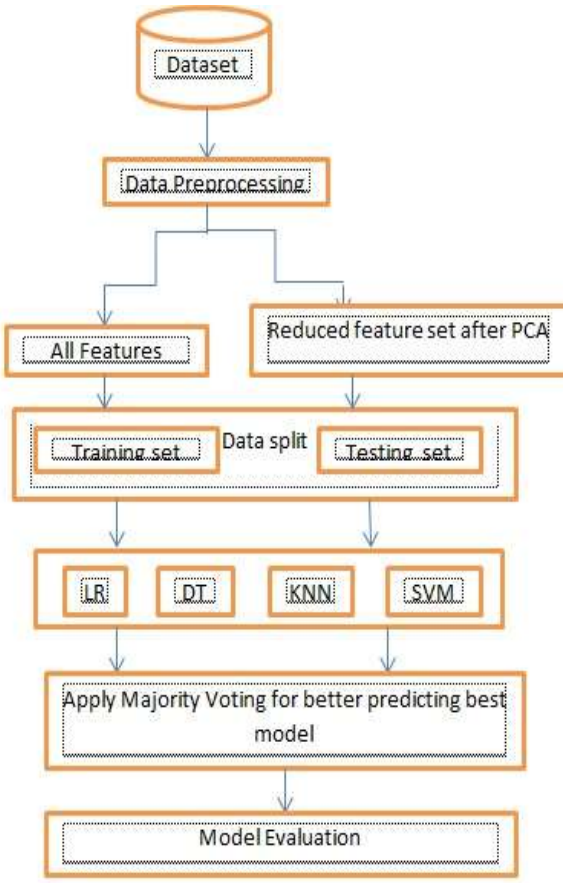


Figure 7.Proposed Majority Voting method

In majority voting , different classifiers combine through majority voting technique. The final prediction has been decided by the highest probability of selected class.Voting ensemble method is shown in figure 8.
Let C1,C2…Cn are the individual classifier predictions,the majority voting classifier
Ý=mode{C1,C2,….Cn}.

In the proposed methodology, the individual learners Logistic Regression, Decision Tree, KNN,SVM and majority voting are implemented on both Clevland as well as Clevland-Statlog-Hungary dataset.Then the above classifiers are applied on both these datasets after using the dimensionality reduction method PCA.Then the
performance of both the method are compared using the metrics, accuracy, precision and recall.

Journal of Theoretical and Applied Information Technology
31st December 2022. Vol.100. No 24

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

4. DISCUSSION AND PERFORMANCE COMPARISON

To assess the benefit of the proposed method, the experiments are conducted on Clevland as well as Statlog-Hungary dataset.In Table 2, the results are presented by comparing performance of Clevland dataset with and without using PCA dimensionality reduction method.Table 3 describes the performance measures of the proposed method on Clevland-Statlog-Hungary dataset.

Table 2 .Performance measures on Clevland dataset

Models	Precision	Recall	f1-score	Accuracy_ score
LR	0.8159	0.8109	0.8109	0.8109
LR(PCA)	0.8159	0.8109	0.8109	0.8298
SVM	0.8135	0.8135	0.8110	0.8109
SVM(PCA)	0.8135	0.8135	0.8111	0.8239
DT	0.8736	0.8697	0.8692	0.8697
DT(PCA)	0.8921	0.8908	0.8908	0.8908
KNN	0.8412	0.8361	0.8361	0.8211
KNN(PCA)	0.88412	0.8361	0.8361	0.8361
Hard Voting	0.8944	0.8861	0.8861	**0.9167**

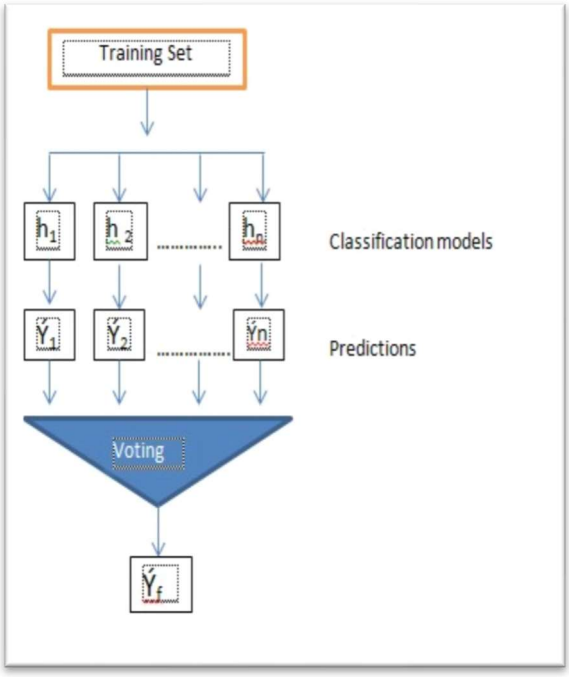


Figure 8.Majority Voting Classifier Model

Table 3 Performance measures of Clevland-Statlog-Hungary-Longbeach dataset

Models	Precision	Recall	f1-score	Accuracy_score
LR	0.7878	0.78689	0.787115	0.786885
LR(PCA)	0.8183	0.8113	0.87115	0.8885
SVM	0.8214	0.8184	0.8184	0.81967
SVM(PCA)	0.8207	0.8196	0.8184	0.8727
DT	0.7866	0.7868	0.7865	0.7869
DT(PCA)	0.7721	0.7778	0.7377	0.7877
KNN	0.820502	0.8196	0.8199	0.8197
KNN(PCA)	0.8205	0.7573	0.7872	0.8571
Hard Voting	0.9432	0.9123	0.9248	**0.9557**

In Clevland dataset, the majority voting with PCA has the highest accuracy of 87%.In the case of the second dataset Decision Tree with PCA shows better accuracy of 89%.The performance of different classifiers and voting ensemble method of Clevland dataset and the
second dataset both with and without PCA is shown in the Figure 9,Figure 10,Figure 11 and Figure 12.A comparison of the above results show that PCA method improves the accuracy than using without PCA.In both the dataset Majority voting performs well than in many recent studies[24].

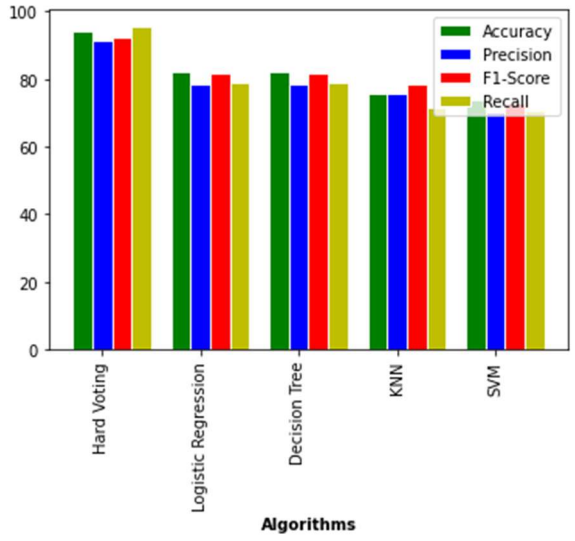


Figure 9.Clevlandwithout PCA

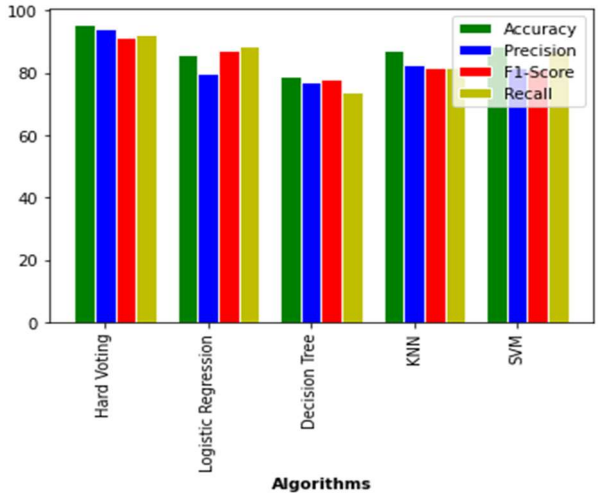


Figure 10.Clevland with PCA

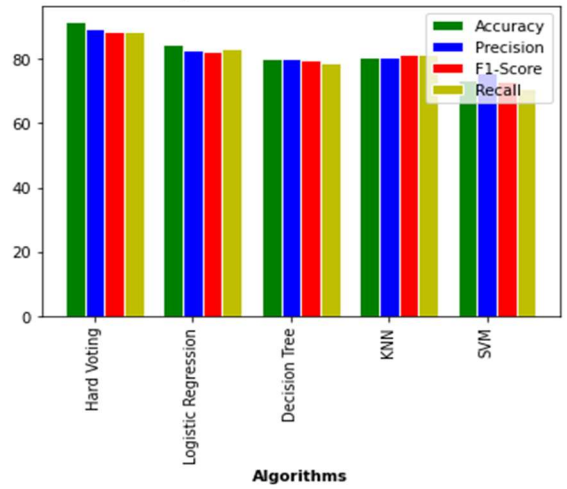


Figure 11.Clevland-Statlog_hungary_Longbeach without PCA

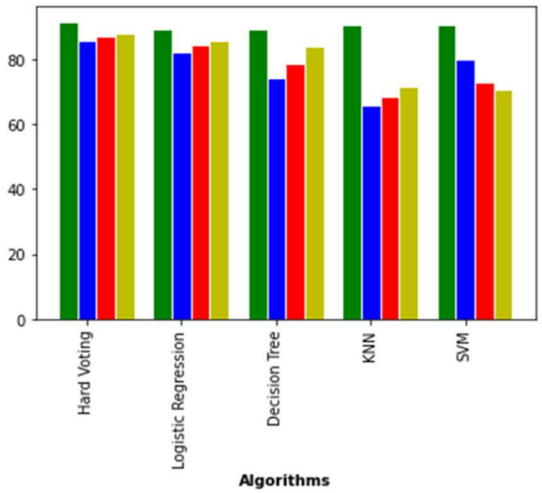


Figure 12.Clevland-Statlog_hungary_Longbeach with PCA

5. CONCLUSION

In this paper, a model based on ensemble voting technique along with PCA is used for improving the performance of heart disease prediction.It is not feasible ,when all the features of a dataset are used for training.The aim of the classifier is to predict whether a person has heart disease or not.In this research,two datasets Clevland and Clevland-Statlog-Hungary-Longbeach are used for comparison.Different classifiers Logistic regression,Decision Tree Classifier,KNN and SVM are applied on these two datasets with and without using PCA.Then majority voting is applied on these classifiers to predict the best accuracy.In both datasets, classifiers with PCA method performs well than using the classifiers alone.From the analysis, majority voting has the highest accuracy of 95.57% in Clevland dataset and 91.67% in Clevland-Statlog-Hungary-Longbeachdataset.The future scope of this study aims to focus on a comparatively large dataset with different feature selection technique and ensemble deep learning.

REFERENCES:

[1] G. EasonHaq, Amin Ul,Li, Jian Ping,Memon, Muhammad HammadNazir, ShahSun, Ruinan-A hybrid intelligent system framework for the prediction of heart disease using machine learning algorithms

[2] M.Swathy, K.Saruladha2022,A Comparative Study of Classification and Prediction of Cardio-Vascular Diseases (CVD) using Machine Learning and Deep Learning Techniques, ICT Express,V 8, I 1, P109-116

[3] Abed Mohammed, Mazin Khanapi Abd Ghani, Mohd Mostafa, Salama Taha Al-Dhief, FahadIbrahim Obaid, Omar Mostafa, Salama A Taha AL-
Dhief, Fahad-"Evaluating the Performance of Machine Learning Techniques in the Classification of Wisconsin Breast Cancer" - Article in International Journal of Engineering and Technology-2018:V7-P 160-166.

[4] V. Kirubha1 , S. Manju Priya-Comparison of Classification Algorithms in Lung Cancer Risk Factor Analysis-International Journal of Science and Research (IJSR) ISSN (Online): 2319-7064

[5] Aishwarya Mujumdara, Dr. Vaidehi V(2019)-Diabetes Prediction using Machine Learning Algorithms-International Conference on Recent Trends in Advanced Computing-ICRTAC 2019

[6] Dr.S.V.Kogilavani, K.Harsitha, P.Jayapratha and S.G.Mirtthula- Heart Disease Prediction using Machine Learning Techniques-International Journal of Advanced Science and Technology Vol. 29, No. 3s, (2020), pp. 78-87.

[7] Sharma, VijetaYadav, Shrinkhala Gupta, Manjari(2020),Heart Disease Prediction using Machine Learning Techniques-Proceedings - IEEE 2020 2nd International Conference on Advances in Computing, Communication Control and Networking,,V 29,I 3 pp-177-181

[8] Latha, C. Beulah ChristalinJeeva, S. Carolin(2019), Improving theaccuracy of prediction of heart disease risk based on ensemble classification techniques-Informatics in Medicine Unlocked-V16,P 100203

[9] Krishnani, Divya,Kumari, Anjali,Dewangan, Akash,Singh, Aditya,Naik,NenavathSrinivas(2019)- Prediction of Coronary Heart Disease using Supervised Machine Learning Algorithms-IEEE Region 10 Annual International Conference, Proceedings/TENCONP: 367-372.

[10] https://www.who.int/news-room/fact-eets/detail/cardiovascular- diseases-(cvds)

[11] Terrada, Oumaima, Hamida, Soufiane ,Cherradi, Bouchaib, Raihani, Abdelhadi, Bouattane, Omar(2020)- Supervised machine learning based medical diagnosis support system for prediction of patients with heart disease- Advances in Science, Technology and Engineering Systems-Volume5-Issue5-P 269-277.

[12] AsmaBaccouche, Begonya Garcia-Zapirain ,Cristian Castillo Olea Adel ,Elmaghraby (2020).Ensemble Deep Learning Models for Heart Disease Classification: A Case Study from Mexico,Information 207; doi:10.3390/info11040207; www.mdpi.com/journal/information

[13] SaiShekhar, M.Mani Chand, Y.MaryGladence, L(2021). Heart Disease Prediction Using Machine Learning, Lecture Notes in Electrical Engineering,V708,p 603-609

[14] R. Indrakumari, T. Poongodi, SoumyaRanjanJena(2020)- Heart Disease Prediction using Exploratory Data Analysis-Procedia Computer Science-V 173,I C,P130-139.

[15] Hashi, EmranaKabir Md. ShahidUzZaman(2020), Developing a Hyperparameter Tuning Based Machine Learning Approach of Heart Disease Prediction, Journal of Applied Science & Process Engineering,V-7,P631-647

[16] Sharma, VineetRasool, AkhtarHajela, Gaurav(2020),Prediction of Heart disease using DNN, Proceedings of the 2nd International Conference on Inventive Research in Computing Applications, ICIRCA , P554-562

[17] UCI machine learning repository: heart disease data set. http://archive.ics.uci.edu /ml/datasets/Heart+Disease.

[18] https://www.kaggle.com/datasets/sid321axn /heart-statlog-cleveland-hungary-final

[19] Javeed, Ashir Khan, ShafqatUllah Ali, LiaqatAli, SardarImrana, YakubuRahman, Atiqur (2022),Machine Learning-Based Automated Diagnostic Systems Developed for Heart Failure Prediction Using Different Types of Data Modalities: A Systematic Review and Future Directions, Computational and Mathematical Methods in Medicine

[20] Alim, Muhammad AffanHabib, ShamsheelaFarooq, YumnaRafay, Abdul(2020)-Robust Heart Disease Prediction: A Novel Approach based on Significant Feature and Ensemble learning Model, 3rd International Conference on Computing, Mathematics and Engineering Technologies: Idea to Innovation for Building the Knowledge Economy, iCoMET 2020.

[21] Gárate-Escamila, Anna Karen Hajjam El Hassani, Amir Andrès, Emmanuel(2020)-Classification models for heart disease prediction using feature selection and PCA, Informatics in Medicine Unlocked,V19.

[22] Yekkala, Indu Dixit, SunandaJabbar, M. A(2018), Prediction of heart disease using ensemble learning and Particle Swarm Optimization, Proceedings of the 2017 International Conference On Smart Technology for Smart Nation, V-November,P691-698

[23] P A Sadiyamole,SManjupriya(2021)-Optimized Feature Driven Decision Support System for Heart Disease Prediction Using Deep Learning,Design Engineering ISSN: 0011-9342,Issue 9,P4716-4729.

[24] Mienye, IbomoiyeDomorSun, Yanxia,Wang, Zenghui(2020), An improved ensemble learning approach for the prediction of heart disease risk, Informatics in Medicine Unlocked,p100402.