



Performance Analysis of Classification Algorithms on a Novel Unified Clinical Decision Support Model for Predicting Coronary Heart Disease Risks

Thendral Puyalnithi^{1*}

Madhuviswanatham Vankadara¹

¹*School of Computing Science and Engineering, VIT University, Vellore-632014, Tamilnadu, India*

* Corresponding author's Email: thendral.p@vit.ac.in

Abstract: Most of the existing Clinical Decision Support Systems (CDSS) for predicting Coronary Heart Disease (CHD) risks rely solely on the analysis of single type of diagnostic data or a certain type of classification algorithm, thus making the predicted component unreliable. Since the Risk value very much influences the further course of the treatment, the correctness of the risk value is eminent. So the novelty of the proposed method lies in considering many classification algorithms and many diagnostics data sets to predict the risk value. Moreover in the proposed method we have given flowchart for finding the risk value, which is a novelty. The results shows that the selection of classification algorithms cannot be universal as proposed by many recent researches and each type of data set will be having a certain type of classifier which works best for the data set. The proposed system has four classifiers namely Mixed cardio data classifier, Single Photon Emission Computed Tomography (SPECT) data classifier, Electrocardiogram (ECG) data classifier and 1-year survival Electrocardiogram data classifier, each trained with benchmark datasets of University of California, Irvine (UCI) Machine learning repository. The performance of algorithms such as Iterative Dichotomiser3 (ID3), Naïve Bayes (NB), Support Vector Machine (SVM), k-Nearest Neighbors (kNN) and Ensemble classifiers such as Bagging and Boosting are evaluated for every data set and best classifier is chosen by its accuracy measure. Ensemble method is the most efficient for 1-year survival ECG data classifier with 89.5% accuracy, ID3 is the most efficient algorithm for Mixed cardiac data classifier with 96.9% accuracy and SVM is the most efficient algorithm for SPECT data classifier and ECG data Classifier with accuracy of 98.3% and 100% respectively. The proposed work also finds the optimal number of samples needed for Bagging and Boosting type of ensemble classifiers.

Keywords: Data mining, Classification algorithms, Naïve bayes, Support vector machine, Decision tree, k-Nearest Neighbour, Ensemble, Bagging, Boosting, Coronary Heart Disease, Clinical Decision Support system, Disease risk prediction.

1. Introduction

The Coronary Heart Disease (CHD) is one of the major fatal diseases in the world. It is now prevalent in many countries. The coronary heart condition starts the lifecycle in the coronary arteries. The coronary arteries are the ones which supply oxygenated blood to the heart muscles. In the inner walls of the coronary artery starts the deposition of wax like substance called plaque due to various reasons. The plaque like substance build up happens over the years. The plaque deposit goes on increasing and it may harden. If the plaque size

reaches a certain threshold, then it will block the blood supply to the heart muscles. The hardened plaque can also dislodge itself from the inner walls of coronary artery and this kind of rupture of plaque mass wounds the inner walls of the coronary artery and this leads to the formation of clot in the inner walls of the coronary artery. This blood clot can also block the supply of oxygenated blood to the heart muscles. The ruptured plaque or the dislodged blood clot can also wander to different parts of the coronary artery and can block the blood flow to the heart muscles. Thus this plaque deposition starts the starvation of oxygen in the heart muscles. The heart muscle which is being deprived of oxygen starts to

die and it leads to the formation of scar tissue or dead tissue. The scar tissue in the heart leads to the heart attack or angina. The heart attack or angina is a symptom of coronary heart disease which is characterised with a severe pressurising pain the chest and the shoulder. The heart attack can be so massive that it may prove fatal to the patient. The occurrence heart attack is a not a guaranteed one, meaning if the patient with coronary heart condition need not get heart attack, but the patient might be having a blocks which are so risky, that may lead to heart attack. So it is highly advised to alleviate the possibility of getting heart attack by going through a regular health check-up. The person with the symptom of coronary heart disease like pain the chest area has to undergo various diagnostics test. To summarise we can say coronary heart disease is characterised by the blocks in the artery. Thus the number of blocks, position of the blocks and the size of blocks will affect the functioning of the heart in different in different ways. (<https://www.nhlbi.nih.gov/>)

Some of the major risk factors of CHD are smoking, being diabetic, devoid of regular exercise, being obese, high blood cholesterol, hypertension, poor diet, excessive alcohol consumption and depression. The CHD can be diagnosed through various ways such as clinical tests (body fluid analysis), imaging of heart and surrounding arteries, cardiac stress testing and monitoring electrical activity over heart muscles. The CHD is characterized by the blocks' count, position and size and these parameters affect the heart's functioning. So it is necessary to find the details about blocks that are formed in the arteries. The blocks position and size can be found out through imaging and how far the block is affecting the heart can be found through ECG test and cardiac stress test. The body fluid tests try to alleviate the CHD risks in the initial stages of plaque formation and not that useful in the final stages of CHD or after getting the heart attack symptom. Thus body fluid tests data and the person's behavioural data can be used to find the reason for the formation of plaque or he risk of getting the blocks. Various imaging techniques are available to access the health of the heart by detecting the blocks, such as Single Photon Emission Computed Tomography (SPECT) and Coronary angiogram. SPECT scan uses radioactive materials for detecting the blocks. The radioactive material is pumped into the heart. Then the scanning shows the blocks by different colouring. SPECT scan is more accurate compared to cardio angiogram. That is why in the proposed method we have used the SPECT, ECG and mixed cardio analysis data of

the patients to train the classifier. (<https://www.nhlbi.nih.gov/>)

Recently there has been an enormous development in the application of Computational Intelligence (CI) in various spheres of human life. Data mining and soft computing techniques play a vital role in health care diagnostics. Many automatic diagnostics system arrives in the market which uses machine learning techniques. These CDSSs (Clinical Decision Support Systems) assists medical practitioners to diagnose the disease effectively. The test data obtained from various sources are said to be static and the Diagnostic System analyses the data with the help of Machine learning techniques. Data mining and soft computing methods like probabilistic methods (naïve bayes), support vector machine, k-nearest neighbor, decision tree, neural network, fuzzy systems, evolutionary algorithm such as particle swarm optimization, firefly, bee colony, ant colony, cuckoo search algorithms based systems are applied to predict the coronary heart disease. The classification system uses clinical data, non-clinical data to train the model. The clinical data includes body fluid test results, imaging and Electro cardiogram data. Non clinical data includes demographic details and behavioral data. The survey covers recent papers in the area of coronary heart disease prediction.

In the research work [1], the severity level of coronary heart disease is found out using data mining algorithms such as linear regression and c4.5, even though the authors analyzed many algorithms to arrive at the risk value for coronary heart disease, they have relied solely on a single heart disease data set (Cleveland data set of UCI machine learning) to arrive at the risk value, which is considered as a drawback, whereas in our proposed method four heart disease related bench mark data sets are analyzed to arrive at the risk value. Since the proposed work is considering various non-invasive types of diagnosis parameters to predict the risk level, predicted value will be more reliable. Moreover the work [1] analyzes the accuracy values of only two classification algorithms, whereas the proposed method compares many data mining algorithms' accuracy including ensemble techniques.

In some of the recent research works [2,11, 12,14] the importance is given to single algorithm only. The dependence of heart disease prediction relies solely on a single algorithm turns the decision making system prone to error. For example in [2], Fuzzy rule base is created out of the rules extracted from the decision tree, thus making Decision tree algorithm as the only contributor of rules for the rule base, whereas the proposed method using many

algorithms to come to the final decision. In the similar fashion, the research works [11,12,14] uses rule based classifiers as the only decision making method to predict heart diseases. In [4,9,13] SVM is the only algorithm tried for diagnosing cardiovascular problem. In some other research recent research works, prediction of heart diseases is done just by analyzing a single type of data sets like ECG alone. In [3,7] only ECG signals are processed to predict the heart diseases, thus making the decision making system prone to error, that is why in our proposed method the risk prediction uses many types of diagnostics data.

There are some recent research works which does not consider a single classifier's decision to predict the heart disease, like in [6], for predicting heart arrhythmia random forest classifier is used, which is a collection of classification trees. In our method too ensemble technique is proposed which is similar to random forest, since both are collection of classifiers. But we have tried boosting along with bagging, which will try to over fit the data during the training phase and most of the times over fitting of data during the training produce more accurate results. In [5] they are going with neural network for predicting risk in congenital heart surgery and they have gone for comparison of many algorithms for finding the risk. They have compared multi-layer perceptron model, radial basis and classification tree classifiers for predicting the risks. They have found that multi-layer perceptron as a winner classifier among the classifiers used. The work has given much importance to neural networks and they have used only classification tree from the data mining family of algorithms and that too a basic algorithm. The main drawback in using neural network is the time to converge and getting into local optimum value, so the neural networks are not used for analysis in our work and we have tried only with the rich set of data mining algorithms. Research works [8, 10] uses Bagging ensemble classifier in medical decision support system. In our proposed methods boosting ensemble method too is tried along with bagging. Since every data set will be having one efficient classifier, it is better to try many algorithms on a single data set and the prediction of heart disease or any disease should also be not relied upon on a single data set. That is the reason for the proposed method to go with many types of data mining algorithms on various types of heart disease data sets to arrive at a reliable risk value. The proposed method emphasizes algorithms of data mining family instead of neural network and evolutionary algorithms which were followed in recent research works. The issues of neural networks

along with evolutionary and hill climbing algorithms lies in annealing, that is they suffer local optima problem. The proposed work has an edge over the other methods of recent research since all the algorithms considered here are very reliable and proven data mining algorithms which do not suffer local optima problem.

2. Proposed method

The proposed method is a unified approach, which takes many diagnostics test parameters into account before deciding the risk level for a patient. The proposed diagnostic model is shown in the Figure 1. The diagnostic model uses four different classifiers such as Mixed cardiac data classifier, SPECT data classifier, ECG data classifier, ECG data 1 year survival classifier. The performance of various data classification methods are analysed and the best method is found out for every classifier. Figure 1 shows the flow diagram of proposed CHD risk prediction model.

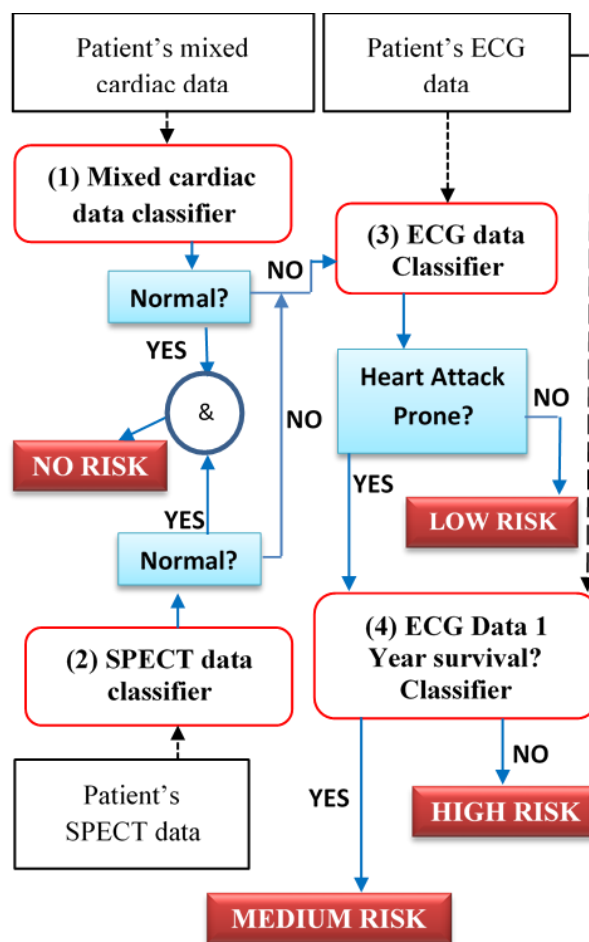


Figure.1 Flow diagram of proposed unified diagnostic model

2.1 Proposed diagnostic model

2.1.1. Mixed cardiac data classifier

The patient suffering from chest pain undergoes various diagnostics tests such as Angiogram, ECG (Exercise + rest) and cardio exercise test. The results from these tests will have many attributes. Out of the attributes of all the tests taken, 14 prominent attributes are selected and it has been compiled in processed Hungarian dataset given in UCI machine learning repository. The 14 attributes are age, sex, chest pain type (typical angina, atypical angina, non-anginal pain , asymptomatic) , resting blood pressure (clinical data), serum cholesterol, fasting blood sugar (clinical data),resting ECG results (ECG data),Maximum heart rate achieved during exercise (exercise data),Exercise induced angina(1=yes and 0 no - exercise data),old peak (ST depression in ECG signal induced by exercise relative to rest - ECG during exercise), slope of the peak exercise ST segment (Value 1: up-sloping, flat, down-sloping - ECG during exercise),Number of major vessels coloured by fluoroscopy, Thalassemia(normal, fixed and reversible defect) and the final attribute, the diagnosis of heart disease which is the class attribute for this dataset and it is binary, value 0 refers to less than 50% diameter narrowing in coronary arteries which is considered normal and value 1 which refers to more than 50% diameter narrowing is considered abnormal. The 14 attributes are used for training the system. The data set has 294 patients' records. The data set is created by Hungarian Institute of Cardiology, Budapest. Then the trained system will find whether an unknown sample data belongs to normal or abnormal category.

2.1.2. SPECT data classifier

The patient undergoes cardiac Single Proton Emission Computed Tomography (SPECT) test. Based on the patient's data, each patient will be categorized into normal or abnormal. The SPECT data classifier is trained with 267 instances each one having 23 attributes. The 267 SPECT image sets of patients are obtained and each image set is processed to arrive at 23 prominent binary values and it is made available in UCI machine learning repository. Out of the 23 attributes, first attribute is the class attribute for the data set, which is a binary attribute and it conveys whether that sample is normal or abnormal.

2.1.3. ECG data classifier

The ECG data classifier predicts whether the patient's ECG data is heart attack prone or not. The data set used for training this classifier is partly generated. The part which is not generated is obtained from UCI machine learning repository, which is a dataset about survival of patient after 1 year from getting the heart attack. The generated part of the data set will have normal and abnormal ECG data values. The abnormal refers to the data which may be prone to heart attack. Thus the entire data set has data which has two classes, the samples which are heart attack prone and not heart attack prone. The generated part of the ECG data set is generated based on the benchmark information from the website 'echopedia.org', which has details about various attributes of ECG values and its range information for normal and abnormal heart conditions. The data set has 450 instances and 7 attributes with two classes. The class attribute is a binary one which refers to normal or abnormal ECG data.

2.1.4. ECG data 1 year survival classifier

This classifier evaluates the patients ECG data and it predicts whether a patient can survive beyond one year of getting the heart attack. The classifier is trained with the data set containing 132 samples with 12 attributes. The data set contains the records of patients who have experienced heart attack. The class attribute of this data set is a binary attribute which specifies whether a patient is alive or not after 1 year of getting the heart attack. In the proposed model, the ECG data of patient who has experienced angina or angina like symptom will be subjected to this classification and it will be used for predicting the risk levels. The dataset to train the classifier is obtained from UCI machine learning repository.

2.1.5. Flow chart of the proposed diagnostic model

The system as in Figure 1 starts by checking the normality of the heart by using the SPECT data classifier and the Mixed cardiac data classifier. The patient would go through SPECT test and ECG test and other needed tests, so that the data needed for the classifiers are available. The patient's SPECT data and Mixed cardiac data will be fed through the SPECT classifier and Mixed cardiac data classifier respectively. If both of the classifiers classifies the data as into normal means, the patient will be put up under SAFE category otherwise he patient data will be examined further. The third classifier, the ECG Classifier will be given with the ECG values of the

patient. The system will predict the whether the data is heart attack prone or not. If the data is not attack prone, then the patient is put under Low Risk category of coronary heart disease. If the data is heart attack prone, the ECG data will be fed through the fourth classifier, the 1-year survival classifier and if the classifier does not affirm the survival after 1 year, then the patient is put under High Risk category, otherwise will be put up under Medium Risk category.

3. Results and analysis

There are four classifiers in this decision support system. Each classifier should use the best classification algorithm for its prediction. So the performance of various classification algorithms for each classifier is found out. Using Orange tool, we are able to compare and analyze the efficiencies of the different classification algorithms separately that are used to test and score the data set. The scoring methods that we have considered to determine the efficiency of each individual method of Classification are accuracy and precision which is the ratio of number of true positives to sum of number of true positives and false positives. The algorithms under consideration are ID3 Decision tree, k-Nearest Neighbor, Support Vector Machine, Naïve Bayes and Ensemble classifier (Bagging and Boosting) which comprises decision trees. Number of models present in the group is varied for bagging and boosting classifiers. Bootstrapping is used for finding the accuracy values in bagging and boosting type classifiers. Thus the best number of models needed for bagging and boosting classifiers is also found out.

Tables 1, 2, 3, and 4 shows accuracy and precision values of individual classification algorithms for every classifier. Tables 5, 6, 7 and 8 shows precision values for bagging ensemble classifiers for mixed cardiac data classifier, SPECT data classifier, ECG data classifier and 1-year survival ECG data classifier respectively. Tables 9, 10, 11 and 12 shows precision values for boosting ensemble classifiers for mixed cardiac data classifier, SPECT data classifier, ECG data classifier and 1-year survival ECG data classifier respectively. Using bootstrapping method of data set selection for training, 50,100,150 and 200 number of classifiers are created for bagging and boosting type of classifiers and the classifiers are evaluated with 50%,60%,70% and 80% of total training samples that also selected based on bootstrapping mechanism.

Table 1. Performance of Individual classification algorithms on Mixed Cardiac Data Classifier

	Accuracy	Precision
ID3	0.935	0.969
kNN	0.732	0.744
SVM	0.783	0.678
NB	0.824	0.786

Table 2. Evaluation of Individual classification algorithms on ECG data classifier

	Accuracy	Precision
ID3	0.912	0.893
kNN	0.821	0.815
SVM	1.000	1.000
NB	0.921	0.922

Table 3. Evaluation of Individual classification algorithms on SPECT data classifier

	Accuracy	Precision
ID3	0.724	0.969
kNN	0.711	0.963
SVM	0.695	0.966
NB	0.768	0.983

Table 4. Evaluation of Individual classification algorithms on 1-year survival data classifier

	Accuracy	Precision
ID3	0.769	0.368
kNN	0.686	0.321
SVM	0.861	0.857
NB	0.686	0.357

Table 5. Precision values of Bagging ensemble classifier for mixed cardiac data classifier

Mixed cardiac data Classifier		Number of models			
		50	100	150	200
Test data set size	50%	0.802	0.809	0.802	0.802
	60%	0.813	0.830	0.836	0.813
	70%	0.825	0.815	0.820	0.820
	80%	0.817	0.817	0.813	0.817

Table 7. Precision values of Bagging ensemble classifier for ECG data classifier

ECG data Classifier Accuracy		Number of models			
		50	100	150	200
Test data set size	50%	0.754	0.745	0.764	0.754
	60%	0.715	0.723	0.731	0.739
	70%	0.692	0.685	0.706	0.692
	80%	0.652	0.664	0.658	0.664

Table 8. Precision values of Bagging ensemble classifier for 1-year survival ECG data classifier

1-year survival ECG data Classifier Accuracy		Number of models			
		50	100	150	200
Test data set size	50%	0.878	0.878	0.878	0.878
	60%	0.873	0.873	0.873	0.873
	70%	0.869	0.869	0.880	0.880
	80%	0.866	0.895	0.895	0.895

Table 9. Precision values of Boost ensemble classifier for mixed cardiac data classifier

Mixed cardiac data Classifier		Number of models			
		50	100	150	200
Test data set size	50%	0.802	0.789	0.761	0.782
	60%	0.768	0.723	0.745	0.745
	70%	0.791	0.796	0.796	0.776
	80%	0.758	0.775	0.775	0.754

Table 10. Precision values of Boost ensemble classifier for SPECT data classifier

SPECT Data Classifier		Number of Models			
		50	100	150	200
Test data set size	50%	0.625	0.625	0.625	0.625
	60%	0.729	0.729	0.729	0.729
	70%	0.625	0.625	0.625	0.625
	80%	0.671	0.671	0.671	0.671

Table 11. Precision values of Boost ensemble classifier for ECG data classifier

ECG Data Classifier		Number of models			
		50	100	150	200
Test data set size	50%	0.696	0.696	0.686	0.696
	60%	0.642	0.642	0.642	0.642
	70%	0.748	0.741	0.748	0.741
	80%	0.658	0.652	0.658	0.658

Table 12. Precision values of Boost ensemble classifier for 1 year survival ECG data classifier

1 year survival ECG data Classifier		Number of models			
		50	100	150	200
Test data set size	50%	0.878	0.848	0.863	0.863
	60%	0.873	0.873	0.873	0.860
	70%	0.880	0.880	0.880	0.880
	80%	0.895	0.885	0.885	0.885

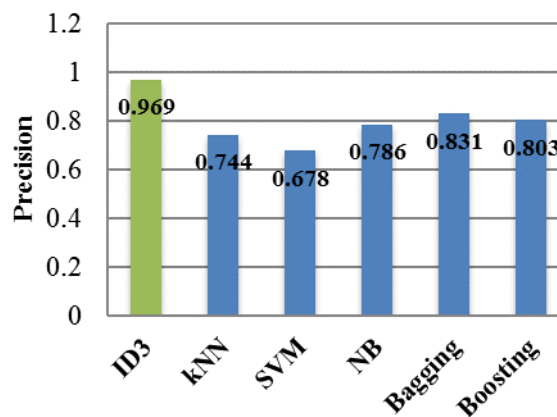


Figure.2 Mixed cardiac data classifier evaluation

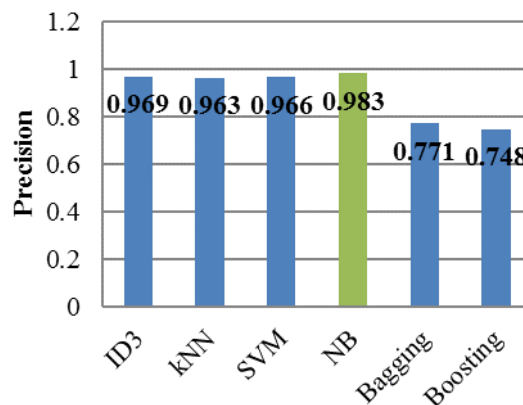


Figure.3 SPECT data classifier evaluation

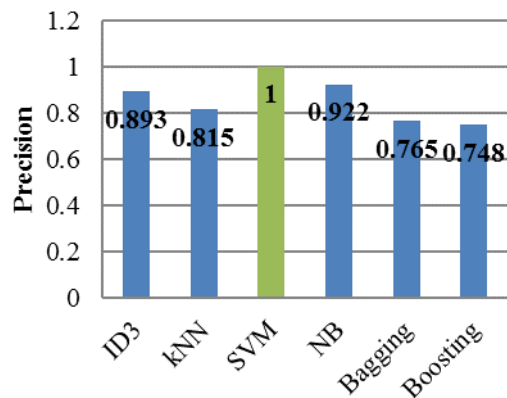


Figure.4 ECG data classifier evaluation

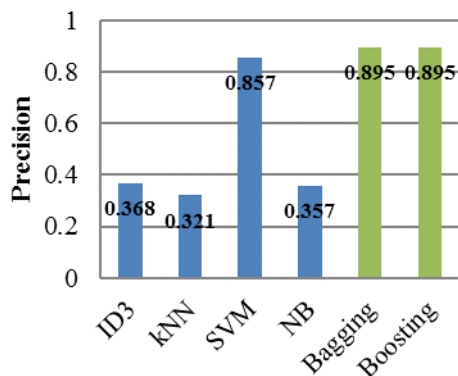


Figure.5 1-Year survival ECG data classifier evaluation

Table 13. Winner classifiers

Classifier	Winner classifier	Precision
Mixed	Decision Tree(ID3)	96.9%
SPECT	Naïve Bayes	98.3%
ECG	SVM	100%
1-year survival	Ensemble	89.5%

Thus based on the precision value, which is the evaluation parameter, the best algorithm for Mixed cardiac data classifier, SPECT data classifier, ECG data classifier and 1-year survival ECG data classifier are given in the Fig. 2, 3, 4 and 5 respectively. The best classifier along with its precision value is given in Table 13.

The results shows clearly that each type of data set that are related to heart diseases, have different type of classifiers as the best performing classifiers for them. This is based on the principle that each classifiers performance depends on number of instances, number of attributes and type of attributes. Thus the Unified model of Decision making system considers many algorithms on different type of data sets to get a reliable risk level, which is the main advantage of this proposed system compared to the existing recent works. The another advantage of the proposed work compared to other recent research works which considers ensemble approach, is that it gives importance to the number models to be present for bagging and boosting method of ensemble classifiers. Moreover the annealing effects of evolutionary algorithms and hill climbing algorithms are nullified by going for data mining algorithms.

The data set will be growing eventually. So the algorithms that are found to be the best might not be the best forever, the algorithm selection procedure must be done whenever there is a substantial increase in the size of the data sets and the increase in size of data sets can occur due to the inclusion of new attributes or instances.

4. Conclusion

Thus various classification algorithms are used over various heart disease related data sets to arrive at the risk value. Since the proposed method is considering best classifiers for every type of data set and multitude of data sets for decision making the result comes out to be very effective. In 1-year survival classifier, Bagging and Boosting methods were top performers with 89.5% accuracy. In the Classifier which is trained with Mixed cardiac data set, decision tree algorithm, ID3 is the top performer with the accuracy of 96.9%. In SPECT and ECG data classifiers, SVM is the top performer with 98.3% and 100% accuracy. Another inference got from the results is the number of classifier models required to get an efficient performance is 150 with Bagging and 50 with Boosting approach. In Boosting approach, since we are over fitting the training data, the number of models required is less to get an efficient performance. As a future work this Clinical Decision Support System Model for heart disease risk level prediction can be integrated with cloud databases of various hospitals and clinical research institute, so that the training can be done over an enormous data set thereby the prediction accuracy can also be increased.

References

- [1] L. Verma , S. Srivastava, and P.C. Negi, “A hybrid data mining model to predict coronary artery disease cases using non-invasive clinical data”, *Journal of Medical Systems*, Vol. 40, No. 7, pp.1-7, 2016.
- [2] M.G. Tsipouras, T.P. Exarchos, D.I. Fotiadis, A.P. Kotsia, K.V. Vakalis, K.K. Naka, and L.K. Michalis, “Automated diagnosis of coronary artery disease based on data mining and fuzzy modelling”, *IEEE Transactions on Information Technology in Biomedicine*, Vol.12, No. 4, pp. 447-558, 2008.
- [3] D. Giri, R. Acharya, R.J. Martis, S. V. Sree, T.C. Lim, T.A. VI, and J.S. Suri, “Automated diagnosis of coronary artery disease affected patients using LDA,PCA,ICA and Discrete Wavelet Transform”, *Knowledge-Based Systems*, Vol.37, pp.274-282, 2013.
- [4] M. Peker, “A decision support system to improve medical diagnosis using a combination of k-medoids clustering based attribute weighting and SVM”, *Journal of Medical Systems*, Vol. 40, No. 5, p.116, 2016.
- [5] D. Ruiz-Fernández, A.M. Torra, A. Soriano-Payá, O. Marín-Alonso, and E. Triana Palencia, “Aid decision algorithms to estimate the risk in

- congenital heart surgery”, *Computer Methods and Programs in Biomedicine*, Vol. 126, Issue C, pp.118-127, 2016.
- [6] E. Alickovic, and A. Subasi, “Medical decision support system for diagnosis of heart arrhythmia using DWT and random forests classifier”, *Journal of Medical Systems*, Vol. 40, No. 4, pp.1-12, 2016.
- [7] U.R. Acharya, M. Sankaranarayanan, J. Nayak, C. Xiang, and T. Tamura, “Automatic identification of cardiac health using modelling techniques: A comparative study”, *Information Sciences*, Vol. 178, No. 23, pp. 4571-4582, 2008.
- [8] S. Bashir, U. Qamar, and F.H. Khan, “Intellihealth: a medical decision support application using a novel weighted multi-layer classifier ensemble framework”, *Journal of Biomedical Informatics*, Vol. 59, Issue C, pp.185-200, 2016.
- [9] R. Sali, H. Shavandi, and M. Sadeghi, “A clinical decision support system based on support vector machine and binary particle swarm optimisation for cardiovascular disease diagnosis”, *International Journal of Data Mining and Bioinformatics*, Vol. 15, No. 4, pp.312-327, 2016.
- [10] S. Kang, S. Cho, and P. Kang, “Multi-class classification via heterogeneous ensemble of one-class classifiers”, *Engineering Applications of Artificial Intelligence*, Vol. 43, Issue C, pp. 35-43, 2015.
- [11] J.J. Christopher, H.K. Nahemiah, and A. Kannan, “A swarm optimization approach for clinical knowledge mining”, *Computer Methods and Programs in Biomedicine*, Vol.121, No. 3, pp.137-148, 2015.
- [12] N.C. Long, P. Meesad, and H. Unger, “A highly accurate firefly based algorithm for heart disease prediction”, *Expert Systems with Applications*, Vol. 42, No. 21, pp.8221-8231, 2015.
- [13] E. Çomak, A. Arslan, and İ. Türkoğlu, “A decision support system based on support vector machines for diagnosis of the heart valve diseases”, *Computers in Biology and Medicine*, Vol. 37, No.1, pp. 21-27, 2007.
- [14] A. Sengur and I. Turkoglu, “A hybrid method based on artificial immune system and fuzzy k-NN algorithm for diagnosis of heart valve diseases”, *Expert Systems with Applications*, Vol. 35, No. 3, pp.1011-1020, 2008.