Effective Cardiovascular Disease Prediction using Hybrid Machine Learning Techniques

Madhavi Veeranki , Jayanag Bayana

Abstract: In today's era deaths due to heart disease has become a major issue approximately one person dies per minute due to heart disease. This is considering both male and female category and this ratio may vary according to the region also this ratio is considered for the people of different age groups.Heart disease is one of the most fatal problems in the whole world, which cannot be seen with a naked eye and comes instantly when its limitations are reached.There are various data mining and machine learning techniques and tools available to extract effective knowledge from databases and to use this knowledge for more accurate diagnosis and decision making.In this paper, we propose a novel method that aims at finding significant features by applying machine learning techniques resulting in improving the accuracy in the prediction of cardiovascular disease. We produce an enhanced performance level with an accuracy through the prediction model for heart disease with the proposed method hybrid random forest with a linear model.In this paper commonly used machine learning techniques and their complexities are summarized.

Keywords: Machine Learning, Heart Disease Prediction, Classification Algorithms.

I. INTRODUCTION

Heart disorder was a form of illness that affects cardiac function. The leading cause of mortality is cardiac failure in today's age. WHO-has estimated that Twelve million humans suffer per year from heart problems. Many respiratory conditions, stroke, renal disorder also hit. Knock was a kind of cardiac condition that happens owing to widening, blockage otherwise reducing of platelets arteries that pass across the brain or it may even be caused via elevated platelets pressure [1]. The main problem confronting the energy sector today was infrastructure dominance. The level of care can be determined by accurate evaluation of every condition and delivering appropriate medication to patients. Low treatment has devastating and unaccepted results. Diverse methods were used in data processing but machine learning for assess the extent of human heart failure. Illness frequency was graded according to different approaches such as K-Neighbor Algorithm (KNN), Decision Trees (DT), Genetic Algorithm (GA), and Naive Bayes (NB). Investigators allow use for many open data processing tools to support the clinicians or doctors classify the heart condition. Judgment Line, k-and Naïve Bayes are widely

Madhavi Veeranki*, Studying Master of Technology, Department of computer science and engineering ,velagapudi Ramakrishna Siddhartha engineering college , kanuru, Vijayawada.

Jayanag Bayana, Sr. Asst. Professor in Velagapudi Ramakrishna Siddhartha Engineering college, dept of computer science and engineering, kanuru, Vijayawada.

employed techniques. Knocking algorithms, kernel length, serial minimum computation as well as neural networks, clear kernel self-map and SVM (Support Vector Machine) are certain common designation-strategies utilized.

HEART DISEASE

A primary measure that defies social security organizations (hospitals, therapeutic focus) was the worth management office at fair prices. Performance pleasantries advocate specifically diagnosing patients by administering appropriate medications. Bad health judgments will result with disastrous consequences, as are currently the case. Health hospitals will restrict clinical research expenditures. Such outcomes can be obtained by utilizing appropriate PC-based knowledge with genuinely solid applications decision in addition [4][6]. The core is the vital component of our anatomy. Everything relies on the good functioning of the heart it's own. Throughout the event the heart errand is not true, it may affect certain human body parts, e.g., cerebrum, kidney, etc. Coronary disorder is an illness that has implications only for the heart's operation.
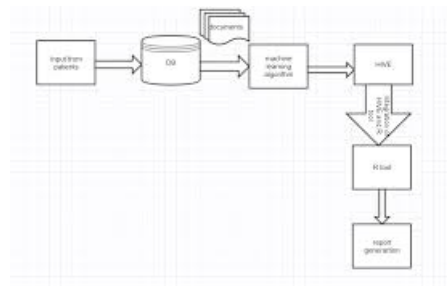

FIGURE.1: Disease prediction using Machine learning technique

II. RELATED WORK

Using PSO Algorithm for Producing Best Rules in Diagnosis of Heart Disease [2]

Heart failure appears to be a growing medical concern total global. Limiting individual skill with competence in physical resolve causes off-research in the context of medical facilities, and information on various illnesses is often inadequate otherwise reliable because we be obtained from specific forms with medical equipment. As the correct understanding of a person's situation is crucial, supplying health science through shrewd tools for detection and management of ailment will minimize the mix-between experts and cash-calamities. Calculating Particle Swarm Optimization (PSO), which is one of the latest amazing revolutionary equations, is being used to build guidelines for cardiovascular disease right now. First the

unusual instructions are encoded and then they are progressed by using PSO measurement based on their precision. Ultimately we equate our findings with the C4.5 estimation.

Clinical decision support system: Risk level prediction of heart disease using weighted fuzzy rules [5]

When people are concerned in their well-of late, among the most complex testing domains was the development in medical room use. One example of the medical space application is the PC-coronary disease detection system that helps identify approaches where the knowledge is gathered by a variety of various sources with analyzed depending upon requirements centered on PC. Previous to that, the purpose of PC was to build an evidence-based, socially responsive therapeutic preference network that incorporates input by healthcare experts as well as actively transfers that knowledge through PC measurements. This technique is boring, and is largely focused on the theoretical findings of health experts. To tackle this problem, AI approaches were developed to gather knowledge through models or rough details, obviously. A weighted fuzzy recommendation focused on personal preference emotionally balanced network (CDSS) for detecting coronary disease is presented here, obviously gathering knowledge from personal details of the individual. The conceptual clinically supporting medical preference network for cardiac danger forecasting consists of 2 stages: (1) Computerized solution to weighted flaccid values as well as era (2) creating a fuzzy, socially nurturing system of options focused on guidance. We used the mining method, characteristic commitment, and credit weighting methodology in the primary process to achieve the weighted fluffy criteria. The fluffy structure is designed at a certain stage according to the weighted fluffy criteria and selected characteristics. Ultimately, the research is carried out on the conceptual structure utilizing the data sources collected from the UCI vault and the application view is compared with the design dependent on the neural network utilizing consistency, affect ability with expressivity.
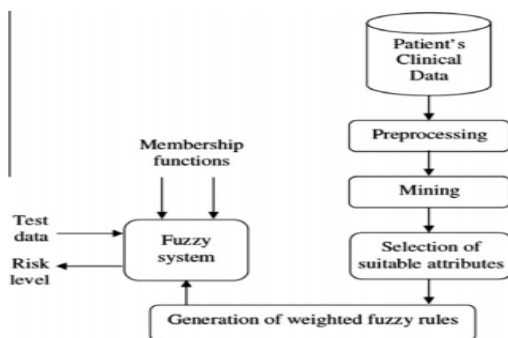

FIGURE.1: Fuzzy-based clinical decision support system.

The Impact of Criterion Weights Techniques in TOPSIS Method of Multi-Criteria Decision Making in Crisp and Intuitionistic Fuzzy Domains [9]

A few parameters (properties) with distinctive significance are used in MCDM issues to determine the best choice from many. Subsequently, there are special weight approaches in writing for determining the value of each basis. Actually. multi-criteria dynamic (MCDM) Strategy is implemented using fresh knowledge index of various weight schemes introduced in text. Accordingly, Intuitionist fluffy TOPSIS is

added to intuitionist fluffy knowledge collecting (IFS) using recorded load techniques. The last two are intuitionist conjecture concerning fresh charges. They plan to apply two weight estimation methods, the Standard Deviation (SD) and the Inclination Selection File (PSI) to practical fuzzy collections. The findings obtained are contrasted with evaluating the impact of weight approaches on the aftermath of complex techniques.

III. FRAMEWORK

In this paper the author tests the efficiency of specific classification / prediction algorithms including SVM, Naïve Bayes and Logistic Regression etc. Both of these algorithms are reliably successful but precision is not strong enough. The author blends two identification algorithms such as Linear Model and Random Forest to construct a new algorithm named Hybrid Machine Learning to increase the prediction precision of the core dataset. Using the Voting Classifier, the hybrid algorithm will be built, the Internally Voting Dataset will construct up to use the Linear Model and the Random Forest and while the Identification Voting Algorithm will determine the predictability of both algorithms and vote for that algorithm that gives greater reliability. But we can still have higher prediction precision algorithm by using hybrid system, which allows to accurately forecast cardiac disease.
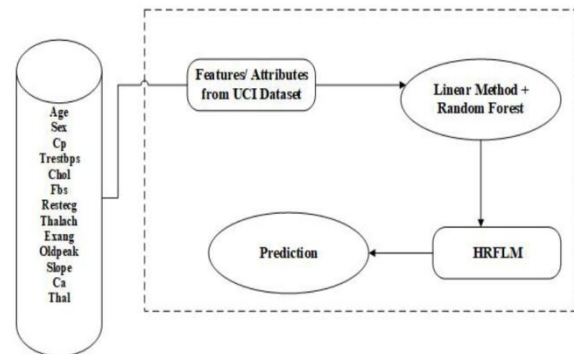

FIGURE.3: Proposed Model

This project consist of 9 modules

Upload Module: Using this feature to import samples with former patients with heart failure

Pre-process Module: We can delete all of those records that include missing properties using this device. Dataset will be divided into two sections named training and testing, both classifiers must construct train model using training data and then check train model by adding test data to that train model to ensure accuracy of identification.

SVM Module: Using this device, we can use the SVM algorithm to construct a train framework and then add testing data to the SVM model for consistency in the identification.

Naïve Bayes: Using this device, we will construct train prototype using the Naïve Bayes methodology and apply testing data to obtain accuracy of Naïve Bayes identification.

Logistic Regression: Accuracy of the train prototype will be verified here with the Logistic Progression method

ANN Module: Deep Learning Artificial Neural Network predict prototype will be developed, and performance data will be used to measure its reliability.

HRFLM: Propose Hybrid Algorithm mixing linear design with algorithm Random Wood. The hybrid model is created using both algorithms, and the voting perception is being used to choose the better executing algorithms.

Extension Extreme Machine Learning Module: This is an additional module that is designed for extension purposes and this module is based on an innovative Severe Machine Learning algorithm that can be more predictive than other algorithms.

Graph: This module displays precision of all algorithms in diagram design as a compare.

IV. EXPERIMENTAL RESULTS

Information on heart attack is pre-handled after different documents have been numerous. The file includes a total of 303 accounts of patients, where 6 reports are with certain attributes lacking. Such 6 documents were removed from the registry and the 297 medical records left are used for pre-preparation. For the characteristics of the specified dataset the multiclass vector and the simultaneous classification are provided. The multi-class component is used to test if coronary disease is close or non-appearing. In the case of a patient with cardiovascular disease the interest is set. The pre-handlingof knowledge is achieved by moving the judgment esteems regarding health data. The after-effects of pre-preparing details for 297 medical reports indicate that 137 documents show the estimation of 1 indicating the presence of coronary disease while the other 160 showed the estimation of 0 suggesting that coronary disease did not exist.

FIGURE.4: Home Screen

FIGURE.5: Proposed HRFLM Screen

Extreme Learning Machine (ELM) Is a novel approach for detection of patterns and the approximation of functions. This system is basically a single neural feed forwarding network; its configuration consists of a single layer of hidden nodes, where the weights between inputs and visible nodes are distributed arbitrarily and stay constant through training and prediction phases. The weights which bind secret nodes to outputs, on the contrary, can be trained very quickly. Experimental literature studies have shown that ELMs can

deliver reasonable predictive results and their computational costs are considerably lower than those educated by the back-propagation methodology.
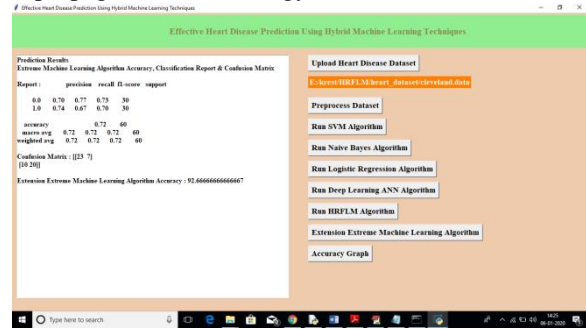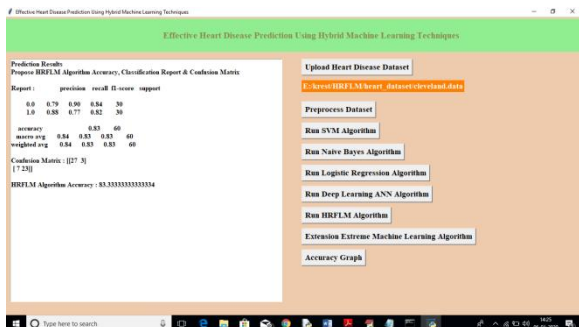

FIGURE.6: Extension Extreme Machine Learning Algorithm Screen

FIGURE.7: Accuracy graph Screen

V. CONCLUSION

Identifying the transmission of raw cardiac details safety data can continue to preserve individual life in the long run, including early diagnosis of heart disease anomalies. In this study, machine learning techniques were used to process raw data and provide a modern and novel discernment of heart disease. Predicting cardiac failure is difficult in the medical profession and is very important. However, once the illness is diagnosed in the early stages, the death rate may be significantly reduced and prevention steps are implemented as quickly as possible.

FUTURE SCOPE

For improved predictive strategies, the future direction of this work can be done with complex mixtures of machine learning strategies. In addition, modern app selection approaches may be introduced to achieve a broader understanding of the significant features to improve heart disease prediction output.

REFERENCES

1 A. S. Abdullah and R. R. Rajalaxmi, ''A data mining model for predicting the coronary heart disease using random forest classifier,'' in Proc. Int. Conf. Recent Trends Comput. Methods, Commun. Controls, Apr. 2012, pp. 22–25.
2 A. H. Alkeshuosh, M. Z. Moghadam, I. Al Mansoori, and M. Abdar, ''Using PSO algorithm for producing best rules in diagnosis of heart disease,'' in Proc. Int. Conf. Comput. Appl. (ICCA), Sep. 2017, pp. 306– 311.

3 N. Al-milli, ''Backpropogation neural network for prediction of heart disease,'' J. Theor. Appl.Inf. Technol., vol. 56, no. 1, pp. 131–135, 2013.

4 C. A. Devi, S. P. Rajamhoana, K. Umamaheswari, R. Kiruba, K. Karunya, and R. Deepika, ''Analysis of neural networks based heart disease prediction system,'' in Proc. 11th Int. Conf. Hum. Syst. Interact. (HSI), Gdansk, Poland, Jul. 2018, pp. 233–239.

5 P. K. Anooj, ''Clinical decision support system: Risk level prediction of heart disease using weighted fuzzy rules,'' J. King Saud Univ.-Comput. Inf. Sci., vol. 24, no. 1, pp. 27–40, Jan. 2012. doi: 10.1016/j.jksuci.2011.09.002.

6 L. Baccour, ''Amended fused TOPSIS-VIKOR for classification (ATOVIC) applied to some UCI data sets,'' Expert Syst. Appl., vol. 99, pp. 115–125, Jun. 2018. doi: 10.1016/j.eswa.2018.01.025.

7 C.-A. Cheng and H.-W. Chiu, ''An artificial neural network model for the evaluation of carotid artery stenting prognosis using a national-widedatabase,'' in Proc. 39th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC), Jul. 2017, pp. 2566–2569.

8 H. A. Esfahani and M. Ghazanfari, ''Cardiovascular disease detection using a new ensemble classifier,'' in Proc. IEEE 4th Int. Conf. Knowl.-Based Eng. Innov. (KBEI), Dec. 2017, pp. 1011–1014.

9 F. Dammak, L. Baccour, and A. M. Alimi, ''The impact of criterion weights techniques in TOPSIS method of multi-criteria decision making in crisp and intuitionistic fuzzy domains,'' in Proc. IEEE Int. Conf. Fuzzy Syst. (FUZZ-IEEE), vol. 9, Aug. 2015, pp. 1–8.

10 R. Das, I. Turkoglu, and A. Sengur, ''Effective diagnosis of heart disease through neural networks ensembles,'' Expert Syst. Appl., vol. 36, no. 4, pp. 7675–7680, May 2009. doi: 10.1016/j.eswa.2008.09.013.

AUTHORS PROFILE:

Ms. Madhavi Veeranki , Studying Master of Technology, Department of computer science and engineering ,velagapudi Ramakrishna Siddhartha engineering college , kanuru, Vijayawada. Her special fields of interest included machine learning and datascience.

Mr B. Jayanag is currently working as a Sr. Asst. Professor in Velagapudi Ramakrishna Siddhartha Engineering college, dept of computer science and engineering, kanuru, vijayawada. So far he is having 15 Years of Teaching Experience in various reputed engineering colleges. His special fields of interest included Artificial intelligence, machine learning, data science.