

HEART DISEASE PREDICTION USING MACHINE LEARNING TECHNIQUES

DR.A.Nagaraju¹, U.Akhila², S.Vidya vani³, A.Chandi priya⁴, D.Bhanu prakash⁵

*¹ Head of Department, ^{2,3,4,5} Students B.Tech -IT, (20S11A1203, 20S11A1244,
20S11A1212,20S11A1209).*

Malla Reddy Institute of Technology and Science., Maisammaguda., Medchal., Ts, India

*¹dranraju.mrits@gmail.com, ²akhilupperla3680@gmail.com, ³vidyavanisingam@gmail.com,
⁴priyareddy.inolla@gmail.com, ⁵dharmavarapubhanuprakash@gmail.com.*

ABSTRACT

As per the recent study by WHO, heart related diseases are increasing. 17.9 million people die every-year due To this. With growing population, it gets further difficult to diagnose and start treatment at early stage. But due to the recent advancement in technology, Machine Learning techniques have accelerated the health sector by multiple researches. Thus, the objective of this paper is to build a ML model for heart disease prediction based on the related parameters. We have used a benchmark dataset of UCI Heart disease prediction for this research work, which consist of 14 different parameters related to heart disease. Machine Learning algorithms such as Random Forest, Support Vector Machine (SVM), Naive Bayes and Decision tree have been used for the development of model. In our research we have also tried to find the correlations between the different attributes available in the dataset with the help of standard Machine Learning methods and then using them efficiently in the prediction of chances of heart disease. Result shows that compared to other ML techniques, Random Forest gives more accuracy in less time for the prediction. This model can be helpful to the medical practitioners at their clinic as decision support system.

1. INTRODUCTION

Healthcare is one of the primary focus for humanity. According to WHO guidelines, good health is the fundamental right for individuals. It is considered that appropriate health care services should be available for regular check up of one's health. Almost 31% of all deaths are due to heart related disease in all over the world. Early detection [1] and treatment of several heart diseases is very complex, especially in developing countries because of the lack of

diagnostic centers and qualified doctors and other resources that affect the accurate prognosis of heart disease. With this concern, in recent times computer technology and machine learning techniques are being used to make medical aid software as a support system for early diagnosis of heart disease.

2. LITERATURE SURVEY

2.1 A.S. Abdullah and R.R. Rajalaxmi

Cardiovascular system conditions area unit a vital pathological state. This study aimed to dissect factors that engender Coronary road criticism exploitation Random forest Classifier. It shows that random forests rule may be accustomed the process and bracket of medical knowledge similar as CAD.

2.2 A.H. Alkeshuosh, M.Z. Moghadam, I. Al Mansoori, and M. Abdar

In this paper, the Particle Swarm Optimization(PSO) rule is employed to induce rules for heart criticism. Rules area unit optimized grounded on their accuracy exploitation PSO rule. militarization medical knowledge with intelligent tools for designation and treating health problem will cut back doctors miscalculations.

2.3 N.AI- milli

Experimenters have latterly planned many software package, tools, and algorithms for developing effective medical call support systems. one in all the foremost vital problems is that the opinion of heart criticism, and diverse experimenters have worked to develop intelligent medical call support systems to assist physicians. A neural network may be a fashionable tool for prognosticating the opinion of heart criticism. A neural network- grounded heart criticism anticipation system is developed during this exploration paper. For heart criticism anticipation, the planned system utilized thirteen medical attributes. The trials distributed during this work incontestible that the planned rule outperformed

analogous state-of-the-art approaches.

2.4 P.K. Anooj The planned clinical call network for threat anticipation in heart cases is split into 2 stages (1) an automatic approach for generating weighted fuzzy rules and call tree rules and (2) the event of a fuzzy rule-grounded call network. To realize weighted fuzzy rules within the initial section, we tend to use the mining fashion, attribute choice, and therefore the attribute weightage system. The fuzzy system is additionally erected exploitation the weighted fuzzy rules and attributes that are named. Eventually, the planned system is tested exploitation datasets earned from the UCI deposit, and its performance is compared to a neural network-grounded system exploitation delicacy, sensibility, and quality.

3. DATASET

The University of California Irvine Heart Disease Dataset :

A quite condensed and widely used dataset, the UCI dataset uses 14 features such as gender, age, cp (chest pain type), cholesterol level, fasting blood sugar level, exercise-induced angina, maximum achieved heart rate (thalach), maximum achieved heart rate (thal), and the number of major vessels. Such biological parameters often represent the heart's health, and using them in a data-oriented fashion for prediction will certainly yield favorable patterns. It has around 300 samples.

4. METHODOLOGY

Existing System

Identification of any heart-related illness at primary stage can reduce the death risk. Various ML techniques are used in medical data to understand the pattern of data and making prediction from them. Healthcare data are generally massive in volumes and complex in structure. ML algorithms are capable to handle the big data and mine them to find the meaningful information. Machine Learning algorithms learn from past data and do prediction on real-time data. This sort of ML framework for coronary illness expectation can encourage cardiologists in taking quicker actions so more patients can get medicines within a shorter timeframe, thus saving a large number of lives.

Proposed System

Machine Learning is a branch of AI research [2] and has become a very popular aspect of data science. The Machine Learning algorithms are designed to perform a large number of tasks such as prediction, classification, decision making etc. To learn the ML algorithms,

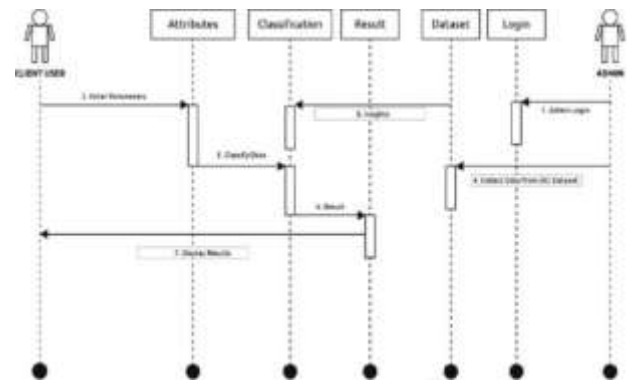
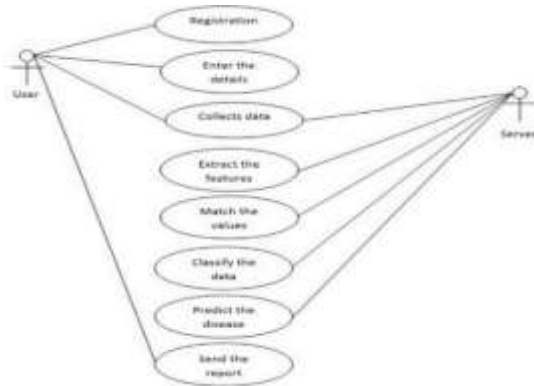
training data is required. After the learning phase, a model is produced which is considered as an output of ML algorithm. This model is then tested and validated on a set of unseen real-time test dataset. The final accuracy of the model is then compared with the actual value, which justifies the overall correctness of predicted result.

UML DIAGRAMS

UML stands for Unified Modeling Language. UML is a standardized general-purpose modelling language in the field of object-oriented software engineering. The standard is managed and was created by the Object Management Group. The goal is for UML to become a common language for creating models of object-oriented computer software. In its current form, UML is comprised of two major components: a Meta-model and a notation. In the future, some form of method or process may also be added to; or associated with, UML. The Unified Modeling Language is a standard language for specifying, Visualization, Constructing and documenting the artefacts of software systems, as well as for business modelling and other non-software systems. The UML represents a collection of best engineering practices that have proven successful in the modelling of large and complex systems. The UML is a very important part of developing object-oriented software and the software development process. The UML uses mostly graphical notations to express the design of software projects.

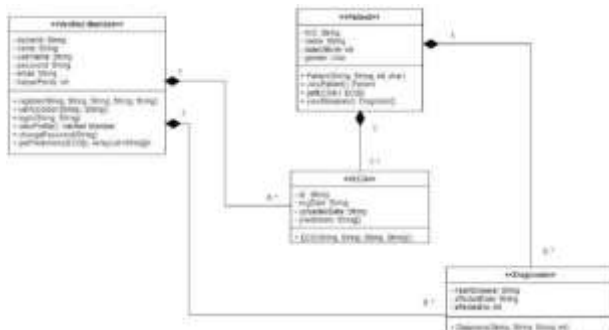
USE CASE DIAGRAM

A use case diagram in the Unified Modeling Language (UML) is a type of behavioral diagram defined by and created from a Use-case analysis. Its purpose is to present a graphical overview of the functionality provided by a system in terms of actors, their goals (represented as use cases), and any dependencies between those use cases. The main purpose of a use case diagram is to show what system functions are performed for which actor. Roles of the actors in the system can be depicted.



CLASS DIAGRAM

In software engineering, a class diagram in the Unified Modeling Language (UML) is a type of static structure diagram that describes the structure of a system by showing the system's classes, their attributes, operations (or methods), and the relationships among the classes. It explains which class contains information.



SEQUENCE DIAGRAM

A sequence diagram in Unified Modeling Language (UML) is a kind of interaction diagram that shows processes operate with one another and in what order. It is a construct of a Message Sequence Chart. Sequence diagrams are sometimes called event diagrams, event scenarios, and timing diagrams.

SYSTEM STUDY

FEASIBILITY STUDY

This section discusses the feasibility of the task and provides business thinking with a well-known plan for the study and some rate estimates. While evaluating the device, a viability test of the proposed device will carry out. It is essential to have a little information about the predominant needs of the machine to assess feasibility. The feasibility analysis takes three key factors into account

- ECONOMICAL FEASIBILITY
- TECHNICAL FEASIBILITY
- SOCIAL FEASIBILITY

ECONOMICAL FEASIBILITY

This work is performed to test the financial results a machine can have on the business. The amount of money that a business owner can invest in assignments and system development is limited. Therefore, the device has also developed in finance, which happened because most of the technology used is freely available. It was only necessary to purchase specially designed goods.

TECHNICAL FEASIBILITY

This study carried out to test the technical feasibility, that is, the device's technical requirements. There should not be much demand on any high-end device for the available technological resources. It will place excessive orders on the available technical resources. It will lead to high demands placed on the user. An advanced machine should have modest requirements as minimal modifications are practical or no changes are required to enforce this machine.

SOCIAL FEASIBILITY

The factor to consider is testing the degree of user recognition of the system. It consists of a consumer education technology to use the system effectively. The

consumer should no longer feel threatened with the system's help, and as an alternative, he needs to obtain it as a necessity. Through the users' use, the acceptance stage depends entirely on the strategies contracted to educate the user and familiarize him with the system. His confidence level should be raised to file a constructive complaint, which is welcome because you are the last consumer of the device.

SYSTEM STUDY

The reason for the test is the detection of errors. A test is a machine trying to identify all possible faults or weaknesses in a product. It provides a method for testing additives, sub- assemblies, assemblies and / or stopping goods for capacity. It is a tutorial way to ensure that this device meets your needs and expectations from the person and now does not fail unacceptably. There are special types of tests. All of them analyse the specific needs of the checks.

TYPES OF TESTS

Unit Testing

The unit test consists of the creation of test cases that demonstrate that the internal software's logic is functioning properly and that programme inputs produce valid outputs. All decision branches and internal instructions must be validated. This is trying out person software devices from the software. After final touch of person unit is finished before integration. The unit assessments perform primary tests on the component stage and examine a specific enterprise strategy for utility and device configuration. The unit tests ensure that each specific path of the enterprise method leads to substantiated requirements and includes clearly described suggestions and predicted outcomes.

Integration Testing

Integration assessments are intended to test the components of the included software to decide whether they truly paintings as one software or no longer. The take a look at is completed primarily based on the occasion, and it is extra associated with the underlying output of the screens or fields. The integration assessments display that despite the fact that the components had been in my opinion nice, as tested within the unit test efficaciously, the set of additives is accurate and regular. The integration testing particularly targets to discover problems arising from the combination of additives.

Functional testing

Job exams provide ordinary demonstrations of

examined capability as determined by way of commercial and technical necessities, machine files and person manuals. Functional exams are targeted on the following factors: Correct enter: Valid categories of legitimate inputs should be typical. Invalid enter: Invalid categories of invalid enter should be rejected. Functions: Must work out specific features. Output: Specific training of software output need to be practiced. Systems / Procedures: Communication structures or approaches need to be referred to as. The organization and training of purposeful tests makes a specialty of requirements, simple capabilities. Before the useful check is finished, extra assessments are selected and the actual fee of the cutting-edge exams is decided.

System Testing

System testing ensures that the final included computer program requirements meet. It validates the layout to help insure recognized and equivalent to roughly. A system test would be checking the integrity of a router's configuration device.. The system take a look at is based on method descriptions and flows, with emphasis on previous operation links and integrity points.

a) White Box testing

The White Box test is atake a look at inwhich the program exams the inner information, shape, and language ofthe program, or at least its motive. It is a cause. It is used to check regions that are not accessible from the black field stage.

b) Black box testing

The black box test evaluates this system without understanding the unit's internal flaps, structure or language under test. Like most other types of checks, black system evaluations must be written from the final supply record, along with a specification or needs register.

Acceptance Testing

User acceptance testing is a critical stage in any business that necessitates lead from the individual leaving. It also guarantees that the device fulfils beneficial requirements.

5. ACKNOWLEDGEMENT

We wish to acknowledge our sincere thanks to our project guide Dr. A. NAGARAJU, Assistant Professor, for formulation of the problem, analysis, guidance and his continuous supervision during the course of work. We acknowledge our sincere thanks to Dr. Vaka Murali Mohan, Principal and Dr. A. NAGARAJU, Head of the Department and Coordinator, faculty members of

INFORMATION TECHNOLOGY Department for their kind cooperation in making this Project work a success.

6. CONCLUSION AND FUTURE SCOPE

In this article, we have a natural inclination to planned a method for prognosticating heart complaints mistreatment machine skill ways, and therefore the results incontestable a high delicacy customary for manufacturing a far better estimation result. we have a tendency to solve the matter of prevision rate while not outfit by introducing a brand new planned Random forest algorithm and proposing an approach to estimate the center rate and condition. Heart grievance prevision is arduous and crucial within the medical field. The process of raw aid knowledge of heart info can aid within the semipermanent saving of mortal lives. These methods were used here to employ data and supply a brand new and distinctive perspective on heart condition. This analysis ought to be dilated to incorporate real-world datasets simply alternative application. Moreover, unless the illness is found early and precautionary controls are taken as soon as possible, the death rate could be drastically reduced. The proposed hybrid HRFLM method combines the benefits of Random Forest (RF) and Linear methodology (LM).

7. 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, "Backpropagation 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. [11] M. Durairaj and V. Revathi, "Prediction of heart disease using back propagation MLP algorithm," Int. J. Sci. Technol. Res., vol. 4, no. 8, pp. 235–239, 2015. [12] M. Gandhi and S. N. Singh, "Predictions in heart disease using techniques of data mining," in Proc. Int. Conf. Futuristic Trends Comput. Anal. Knowl. Manage. (ABLAZE), Feb. 2015, pp. 520–525. [13] A. Gavhane, G. Kokkula, I. Pandya, and K. Devadkar, "Prediction of heart disease using machine learning," in Proc. 2nd Int. Conf. Electron., Commun. Aerosp. Technol. (ICECA), Mar. 2018, pp. 1275–1278. [14] B. S. S. Rathnayake and G. U. Ganegoda, "Heart diseases prediction with data mining and neural network techniques," in Proc. 3rd Int. Conf. Converg. Technol. (I2CT), Apr. 2018, pp. 1–6. [15] N. K. S. Banu and S. Swamy, "Prediction of heart disease at early stage using data mining and big data analytics: A survey," in Proc. Int. Conf. Elect., Electron., Commun., Comput. Optim. Techn. (ICECCOT), Dec. 2016'