Alinteri J. of Agr. Sci. (2021) 36(1): 260-264
e-ISSN: 2587-2249
info@alinteridergisi.com

http://dergipark.gov.tr/alinterizbd
http://www.alinteridergisi.com/
DOI:10.47059/alinteri/V36I1/AJAS21039

RESEARCH ARTICLE

Prediction of Heart Disease Using Machine Learning Algorithms

S. Ravi[1*] • Dr.M. Sambath[2] • Dr.J. Thangakumar[3] • Danam Kumar[4] • Gorantla Naveen[5]• Makka Bramiah[6]

[1*]*Department of Computer Science and Engineering, Hindustan Institute of Technology and Science, Rajiv Gandhi Salai, Padur, Chennai, Tamil Nadu, India.*

[2]*Department of Computer Science and Engineering, Hindustan Institute of Technology and Science, Rajiv Gandhi Salai, Padur, Chennai, Tamil Nadu, India.*

[3]*Department of Computer Science and Engineering, Hindustan Institute of Technology and Science, Rajiv Gandhi Salai, Padur, Chennai, Tamil Nadu, India.*

[4]*Department of Computer Science and Engineering, Hindustan Institute of Technology and Science, Rajiv Gandhi Salai, Padur, Chennai, Tamil Nadu, India.*

[5]*Department of Computer Science and Engineering, Hindustan Institute of Technology and Science, Rajiv Gandhi Salai, Padur, Chennai, Tamil Nadu, India.*

[6]*Department of Computer Science and Engineering, Hindustan Institute of Technology and Science, Rajiv Gandhi Salai, Padur, Chennai, Tamil Nadu, India.*

ARTICLE INFO

Article History:
Received: 04.02.2021
Accepted: 28.03.2021
Available Online: 15.05.2021

Keywords:
Learning Algorithms
Research Process
Knowledge Discovery

ABSTRACT

As big data becomes more prevalent in the healthcare and medical sectors, accurate medical data collection benefits early diagnosis of heart disease, hospital treatment, and government resources. However, where medical data quality is lacking, understanding accuracy suffers. Consequently, some field diseases have unique features in different regions, which can make illness more difficult. It is now more hard to predict outbreaks. We automate machine learning algorithms for efficient epidemic detection in bacterial infection population in this paper. We put the modified forecasts to the test using securely and efficiently datasets. areas of the region to improve the situation of lost data, we use a predictive modeling approach to restore inaccurate value. Focused upon its patient's signs, a heart attack is suspected. Models were built using machine learning techniques. As a consequence, the accuracy is pinpoint accurate. The Flask web interface is used to build the Application. In this research, we shall conduct experiments using machine learning methods.

Please cite this paper as follows:

Ravi, S., Dr. Sambath, M., Dr. Thangakumar, J., Kumar, D., Naveen, G. and Bramiah, M. (2021). Prediction of Heart Disease Using Machine Learning Algorithms. *Alinteri Journal of Agriculture Sciences, 36*(1): 260-264. doi: 10.47059/alinteri/V36I1/AJAS21039

Introduction

Introduction

Diabetes, elevated blood pressure, high cholesterol, slow pulse rhythm, and many other medical conditions make it very hard to diagnose heart failure. Several knowledge discovery and support vector methods were being used to estimate the severity of heart illnesses. Individuals' Different options, including the K-Nearest Neighbor Algorithm (KNN), Decision Trees (DT), and Genetic Algorithms (GA), are used to describe the severity of the disorder, the Optimization Approach (GA) and Naive Bayes (NB) are used. Since the nature of respiratory failure is complex, it would be handled with care. Malfunction to do that would damage blood vessels or prove fatal. Various versions of cardiovascular diseases are dissected using patient management and analytics processing viewpoints. In terms of predicting cardiovascular diseases and Research process, knowledge discovery, and analysis both play essential roles.

Statistical methods are most often used to predict the probability of heart condition cases. Jun Wu was the former

* Corresponding author: S. Ravi

director in charge of enforcing the manuscript's review and approving its printing. Different data integration techniques, particularly validated statistical approaches, were also used to forecast heart problems. Many other samples were conducted in this study Investigate how to construct a predictive approach by combining not just to various methods and furthermore matching double or all of those. Many modern techniques that were merged are known as reverse strategies. To integrate computer vision, humans use blood pressure historical data. Block on the left package branch.

AFL, PVC, and Second-Degree Lock (BII) are mostly all checks used to assess the sufferer's exact state in regards to respiratory failure. A schema with both a neural network structure network (RBFN) is often used for detection, with 70% of the evidence use in testing and 30% used during classifying. We also work in the fields of medical advances Introduce the Medical Imaging Visualization Tool as well (CADSS). Subsequent analysis has found in using machine learning methods in the medical field decreases the while doing activities. Prediction of illness with the more meaningful comparison. We are thinking about using the GA to detect heart failure. This method employs efficient associative learning for fitness value, merger, and evolution the suggested performance mechanism is the product of affiliation laws extracted again from GA. For confirmation, humans will use extremely poorly Cleveland model, which again was obtained from with a Data mining technique. We'll like that our results compare to these more particularly poorly classifiers today. Another very sophisticated especially true in developing on the market.

In addition, machine learning is regarded as the most effective tool for detecting illnesses such as diabetes and chronic condition. For forecasting colon cancer, we use a simulation results with texture features. As comparison to existing approaches in experiments such as the results show a higher level of efficiency. Blood Vessel Catheterization (CAS) has also been a widespread medical procedure since current history choice in the health world In deceased heart diseases, the ICC forecasts the likelihood of major cardiovascular conditions (MACE). Their evaluation appears highly important. To generate data, we employ a Genetic Algorithm (ANN). Before it comes to forecasting heart failure, it provides fair outcomes. It implements odds as well as simulated results from many counterpart methods. This project has a precision, reaching up to 89.01 percentile. These findings are remarkable as compared to past research. As historically shown by the Detection and classification of plant datasets is used in all studies using a Back Propagation NN to improve the efficacy of colon cancer.

There are also significant additions in computer science (ML) strategies for the Artificial intelligence (IoT). It has been demonstrated that using machine learning methods on datasets will reliably classify Devices linked to a service. Lead to the contamination and position to acquire and identified moving vehicles from nine separate Internet of Things units, as well as PCs including devices. They prepared a deep learning is used in a low- and mid meta classification task. In the very first point, the classification algorithm will distinguish through traffic generated by the sensors and data analysis shows by non-IoT instruments. In the second level, increasing Access point is allocated to a specific Smart home grade. Computer vision is a promising tool for obtaining accurate results. derived from extraction of information collected from IoT devices embedded in various areas In-depth education Even with its layered nature, computer vision is also appropriate for interface processing. We address the Model Random Forest with Linear Model procedure.

Literature Survey

In the study mentioned above, various machine learning algorithms are often used to classify heart problems. Per the surveys undertaken by Shusaku Tsumoto during the year 2000, since beings are incapable of organizing massive amounts of data, we should use machine learning methods to derive different trends from the critical enabler archive, and can also be used afterwards, this period for disease diagnosis and do a number of moves on it.

Y. Alp Aslandogan et al. (2004) combined three different classification models termed K-nearest Neighbor (KNN), Decision Tree, and Nave Bayesian to create a single conclusion judgement using Dempster's' law. Focused on the merged idea, the precision of this grouping has increased. Franck Le Duff (2004) investigated how to quickly and potential to help a decision tree using scientific proof from even a practitioner or facility He suggested a few machine learning techniques to help neurosurgeons estimate therapeutic efficacy. The main drawback of the design was that the customer had to be aware of the things, and we had to collect appropriate data to construct a simple plan. Kiyong Noh et al. (2006) used a classifier to delete low- and mid-complexes by obtaining HRV and ECG indicators. Kiyong engaged the FP growth framework in this recursive approach. A level of rule continuity was reached, enabling for a good push on thinning concepts in the future. design production method (nine) Heon Gyu Lee et al. (2007) working for the extension leader of the intra function by direct and recursive mechanisms of Heart Rate Variability, also categorizing for the programs of Arithmetical and categorizing for the programs of Arithmetical and archiving for the programs of Arithmetical and archiving for the programs of Arithmetical and categorizing for the programs of Arithmetical and categorizing for the programs comparison of both the multi-parametric function by similar and recursive Heart Rate Variability features (HRV). The analysis of the differential equation's aspects of the HRV columns is analyzed utilizing Decision Tree (C4.5), Multiple Association Rules (CMAR), or Bayesian classifiers, and also Support Vector Machine (SVM).

Overview of the System

- ***Existing System***

The results of this approach are analyzed using the LIBSVM and WEKA data mining tools. This experiment uses five data sets from the Irvine UC machine learning repository (Iris, diabetes disease, breast cancer disease, heart disease, and hepatitis).

Both LIBSVM and Weka machine learning applications are used to interpret the effects of this method. This observed change use of five data sources from deep learning framework at the University of California, Irvine (Iris, diabetes disease, breast cancer disease, heart disease, and hepatitis). Otoom et al. offered an outline as well as a framework for following up. The unit under consideration detects and tracks cardiovascular events. Cleveland Heart received the data from the UCI. This dataset includes 303 cases with 76 characteristics. Thirteen of the 76 components have been used. Two tests are conducted for recognition using proposed algorithm: Bayes Naive, Support Vector Machine, and Functional Trees FT. The WEKA tool will be used for recognition. Since passing the Stronghold inspection, the SVM technique yielded an accuracy of 88.3 percent.

Disadvantages of Existing System

Either of these methods has same target: to use hybrid learning techniques to identify cardiovascular disease. Even so, they only used one identification and optimizing method.

- ***Proposed System***

Machine learning techniques are available for forecasting the incidence of cardiovascular disease outbreaks in large environments. We take a cardiovascular disease database from its Kaggle website, prepare and test it with the random forest model, and then forecast based around each patient's user feedback.

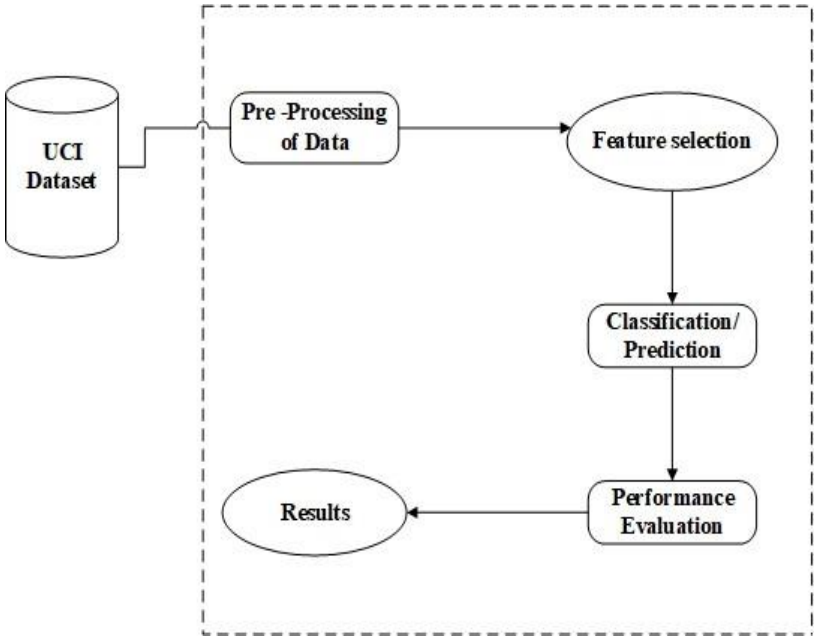
Advantages of Proposed System

By integrating numerous technical optimizations and machine learning techniques, the proposed strategy offered a comprehensive solution to achieving the desired goals.

The proposal techniques are described to optimization learning based on current approximate sets, and their performance is measured using classifier performance tests.

- ***Over View of Method and Result***

We choose a mathematical approach only with three mining association concepts, including a priori, predictive, and Tertius, to explore the reason of cardiovascular disease on the UCI.



Cleveland information gathering. Females get a lower risk of cardiovascular than men, according to the data. The significance of early detection of cardiovascular disease cannot be overestimated. Traditional approaches, on the other hand, are inadequate for accurate estimation and assessment. HRFLM employs ANN with deep network as the input, along with 13 clinical functions. The results are successfully detected using standard methods. The risk levels escalate to toxic levels, and the accuracy of the diagnosis is determined by a number of factors. the disease. A successful care plan is required due to the severity and complications of heart disease. Data mining methods are helpful in addressing challenges in the medical field. Methods such as DT, NN, SVM, and KNN are also used in data mining. Amongst these different approaches employed, the results of SVM have proven to be beneficial in terms of developing precision in the region. Detection of disease. A nonlinear technique with a module for monitoring cardiac function is used to diagnose irregular heartbeats such as bradycardia, tachycardia, atrial, atrial-ventricular flutters, and others. The precision of the ECG-based outcome results could be used to assess the method's performance effectiveness. The ANN instructional tool is used in clients to diagnose illness and forecast possible abnormalities. KNN, LR, SVM, NN, and Vote are a few of the data mining techniques and statistical techniques that also have gained significant popularity for detecting and forecasting cardiovascular disease. The Probabilistic Principal Component Analysis (PPCA) methodology is developed for assessment using three UCI data sets: Cleveland, Switzerland,

and Hungarian. The approach removes increased vectors and reduces the function component using vector projections.

In order to more effectively forecast heart disease, we compare the A priori algorithm with SVM to 9 other existing approaches. In contrast to other current methods, the classification's production has shown a higher degree of precision and consistency in the estimation of cardiovascular disease. The choice of features is critical in the prediction of diseases. An ANN with back propagation is recommended for improved predictive modeling.

The effects of using ANN are highly reliable and detailed. The Recurrent Fuzzy Neural Network (RFNN), an optimization algorithm with fuzzy NN, is used to diagnose cardiovascular disease. The UCI data set considers 297 patient information in all, with 252 details used for preparation and the remainder for the research.

The results were considered to be acceptable based on the assessment. It is planned to use SVM and ANN to forecast cardiovascular disease. Two methods have been used in this method with the assumption of accuracy and processing time.

As seen in the suggested framework splits the existing data into two classes for further review using SVM and ANN. The Back Propagation Neural Network (BPNN) with grading system is used to create and classify the hypertension gene series.

A. Data Processing

Data on cardiac event is pre-processed after several records have been compiled. The dataset contains 303 medical reports, with 6 of them containing certain null values. Since those 6 documents were removed from the database, the current 297 medical information are included in well before. The micro vector and classification techniques are used for the characteristics of the listed dataset. The micro predictor is used to assess whether a person has cardiovascular disease or not. The advantage is intensified because the client has cardiovascular disease.

AGE	NUMERIC[20 TO 77];UNIQUE=41;MEAN54;MEDIUM:56
SEX	NUMERIC[0 TO 1];UNIQUE=2;MEAN 0;MEDIUM:1
CP	NUMERIC[1 TO 4];UNIQUE=4;MEAN 3.1;MEDIUM:3
TESTBPS	NUMERIC[94 TO 200];UNIQUE=50;MEAN 131.1;MEDIUM:130
CHOL	NUMERIC[126TO 564];UNIQUE=152;MEAN:246;MEDIUM:241
FBS	NUMERIC[0 TO 1];UNIQUE=2;MEAN:0.3;MEDIUM:0
RESTECG	NUMERIC[0 TO 2];UNIQUE=3;MEAN:0.99;MEDIUM:1
THALACH	NUMERIC[71 TO 201];UNIQUE=91;MEAN:146MEDIUM:156
EXANG	NUMERIC[0 TO 1];UNIQUE=2;MEAN:0.3;MEDIUM:0
OLPEAK	NUMERIC[0 TO 6.20];UNIQUE=41;MEAN:1.3;MEDIUM:0.8
SLOPE	NUMERIC[1 TO 3];UNIQUE=3;MEAN:1;MEDIUM:2
CA	CATAGORIAL [5]LEVELS
THAL	CATEGORIAL [4]LEVELS
TARGET	NUMERIC[0TO 4];UNIQUE=5;MEAN:0.7;MEDIUM:0

Instead that, the score is set to zero, indicating that the individual does not have cardiovascular disease. As part of the pre-processing of outcomes, medical files are translated into diagnosis values. According to the condition characterized findings for 297 patient reports, 137 get a level of 1, indicating the presence of cardiovascular disease, whereas 160 have a level of 0, shows the absence of cardiovascular disease.

B. Feature Selection and Reduction

To separate the individual data from a data collection's 13 attributes, two components related to gender and age are being used.

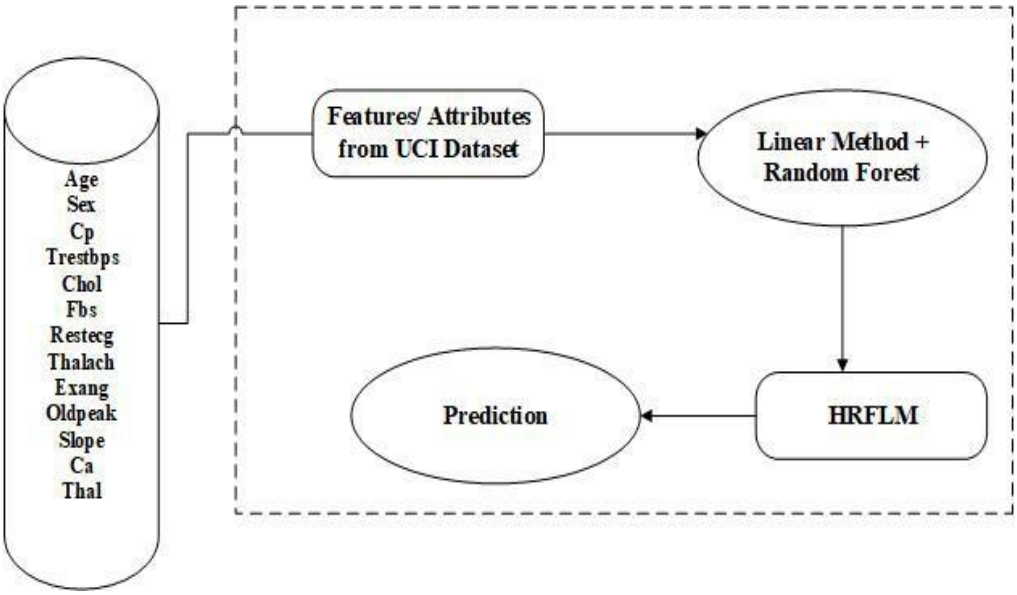


Figure 2. Prediction of heart disease with HRFLM

The personal health history the remaining 11 characteristics are significant as they include vital clinical evidence. Field documentation are important for assessing the incidence and diagnosis of cardiovascular disease. This study employs many (ML) methods, namely NB, GLM, and LR, as mentioned earlier. A few of the words shown are DL, DT, RF, GBT, and SVM. The process was conducted for all selected features and all ML techniques. Figure 2 depicts the HRFLM measured process.

Conclusion

The incidence of cardiovascular disorders will exceed the target line and reach a climax. Cardiovascular disease is a complicated disease that affects a large number of people per year. One of the main drawbacks of these writings is that they mainly focus on the use of labeling different algorithms for labeling. Assumption of heart attack Trying to investigate various data cleaning and mining strategies for planning and constructing a data mining-ready collection So how I can use binary logistic equations to decide whether an individual has cardiovascular disease or not. Any non-medical person can use this software to predict cardiac arrest and save persons time.

References

Kaelber, D.C., Jha, A.K., Johnston, D., Middleton, B., and Bates, D.W. (2008). A research agenda for personal health records (PHRs). *Journal of the American Medical Informatics Association*, 15(6), 729-736.

Yu, S., Wang, C., Ren, K., and Lou, W. (2010). Achieving secure, scalable, and fine-grained data access control in cloud computing. *In Proceedings IEEE INFOCOM*, 1-9.

Kamara, S., and Lauter, K. (2010). Cryptographic cloud storage. *In International Conference on Financial Cryptography and Data Security*, *Springer, Berlin, Heidelberg*, 136-149.

Chen, T.S., Liu, C.H., Chen, T.L., Chen, C.S., Bau, J.G., and Lin, T.C. (2012). Secure dynamic access control scheme of PHR in cloud computing. *Journal of medical systems*, 36(6), 4005-4020.

Xiao, Z., and Xiao, Y. (2012). Security and privacy in cloud computing. *IEEE Communications Surveys and Tutorials*, 15(2), 843-859.

Li, J. (2013). Electronic personal health records and the question of privacy. *Computer*, 48(2, 2015), 20-27.

Li, M., Yu, S., Zheng, Y., Ren, K., and Lou, W. (2012). Scalable and secure sharing of personal health records in cloud computing using attribute-based encryption. *IEEE transactions on parallel and distributed systems*, 24(1), 131-143.

Abbas, A., and Khan, S.U. (2014). A review on the state-of-the-art privacy-preserving approaches in the e-health clouds. *IEEE Journal of Biomedical and Health Informatics*, 18(4), 1431-1441.

"Health Insurance Portability and Accountability," http://www.hhs.gov/ocr/privacy/hipaa/administrative/privacyrule/

Khan, A.N., Kiah, M.M., Madani, S.A., Ali, M., and Shamshir band, S. (2014). Incremental proxy re-encryption scheme for mobile cloud computing environment. *The Journal of Supercomputing*, 68(2), 624-651.

Abbas, A., Bilal, K., Zhang, L., and Khan, S.U. (2015). A cloud based health insurance plan recommendation system: A user centered approach. *Future Generation Computer Systems*, 43, 99-109.

Gai, K., Qiu, M., and Sun, X. (2018). A survey on FinTech. *Journal of Network and Computer Applications*, 103, 262-273.

Wu, R., Ahn, G.J., and Hu, H. (2012). Secure sharing of electronic health records in clouds. *In 8th international conference on collaborative computing: networking, applications and work sharing (CollaborateCom)*, 711-718.

Au, M.H., Yuen, T.H., Liu, J.K., Susilo, W., Huang, X., Xiang, Y., and Jiang, Z.L. (2017). A general framework for secure sharing of personal health records in cloud system. *Journal of computer and system sciences*, 90, 46-62.

Gai, K., and Qiu, M. (2017). Blend arithmetic operations on tensor-based fully homomorphic encryption over real numbers. *IEEE Transactions on Industrial Informatics*, 14(8), 3590-3598.

https://doi.org/10.1109/TII.2017.2780885.

Gai, K., Qiu, M., Xiong, Z., and Liu, M. (2018). Privacy-preserving multi-channel communication in edge-of-things. *Future Generation Computer Systems*, 85, 190-200.