Heart Attack Prediction in Machine Learning Environment

Dr. P. Senthil Kumari[1], S. Vinitha[2]*

[1] Head In charge, Dept. of Computer Science, Thassim Beevi Abdul Kader College for Women, Kilakarai,
[2] M.Phil. Research Scholar, Tamil Nadu, India.
senthilmathimca@gmail.com

Abstract. Healthcare aims to impart more personal, preventive and predictive clinical treatments to patients nowadays. Since digitized data acquisition is growing more, Artificial Intelligence (AI) applications such as Machine Learning (ML) using Data Mining (DM) are used in the areas which are misunderstood as the province of human expert's fields. AI performs mimicry of human cognitive functions. Patient care requires interaction between clinicians, caretakers, managers and executives to decide on a clinical method (treatment, diagnosis and therapy) and a non-clinical method (technology chosen, budget, etc.). Due to advancements in deep neural networks, researchers and medical practitioners focus on developing an automated system to anticipate the occurrence of heart attacks. In the proposed research work, an extreme learning machine using AI and ML based diagnosis system using R-tool for heart attack prediction using classifiers (Decision Tree (DT) and Random Forest Algorithm (RFA)) is proposed. The objective of this research work is to find the most suitable classification algorithm for predicting the heart disease. The performance of classifiers is evaluated with performance metrics such as classification accuracy and error rate. The performance of the proposed system has been validated on 13 features and a reduced set of features. The feature reduction improved the performance of classifiers in the system performance metrics of accuracy and error rate. The proposed ML-based decision support system will assist the doctors in diagnosing the prediction of heart attack for the patients efficiently.

Keywords - Machine Learning, Data Mining, Decision Tree and Random Forest Algorithm

I. INTRODUCTION

Netflix attracts people by giving people's favorite movies and series. Amazon attracts people by providing people's favorite items. Google attracts people by giving the terms and conditions they search for. Detailed personal profiling can be created with all this data used to understand and predict healthcare trends. Artificial Intelligence can be used in all areas of healthcare, from admitting the patient to discharge [1].

AI is a collection of technology. AI is implemented in diagnostics, therapeutic and population health management. Specific AI technologies relevant to healthcare are Machine Learning (ML), AI based Deep Neural Networks (DNN), Deep Learning (DL) and Natural Language Processing (NLP) [2].

Precision medicine is the existing application of ML in healthcare to predict the treatment protocol. Supervised learning is used for this purpose. Digital healthcare applications, Complex algorithms and genes, proteins and metabolites-based tests are three clinical areas where innovation in precision medicine is used. Reduced healthcare costs, reduced drug response side effects, and improving drug action effectiveness are the advantages of precision medicine [3].

Google collaborates with health delivery networks to provide prediction models from big data to alarm clinicians of heart failure conditions, etc. In addition, ML has complex dialogue management and conversational flexibility to handle translation and dissemination of information to help patients [4].

Healthcare organizations use ANN to provide care delivery at a minimum cost by informing management decisions at the micro, meso and macro levels. Individual patient decisions come under the micro-level. Decisions made at the group (for example. Organization or department) are meso-level. Finally, large groups or public agencies decide the allocation of resources whose decisions reflect society's macro level. Information for decision making (in this paper, thirteen attributes of a patient) are captured and entered into Excel sheet for further analysis of categorization [5].

NLP comprises speech recognition, text analysis, translation, etc. Statistical and semantic NLP are classifications o NLP. Statistical NLP is a combination of DL and ANN. NLP includes the interpretation of clinical documents, reports and conversational AI. Clinical notes in cardiology help derive metadata in the development of downstream ML applications [6].

Empathy and ethical dilemma lack between AI and patients. Streamlining and creative solutions are effective and preferred by humans. During the SARS-COV-2 epidemic, there is a need to promote remote healthcare technologies to save lives and reduce the risk to both healthcare workers and patients [7].

More than 500 people are affected by stroke all over the world. In China, stroke is the first cause of death, and in North America, stroke is the fifth cause of death. US$ 689 billion medical expenses are spent on stroke. To prevent the death rate by stroke and reduce the medical expenses on stroke, this research paper proposes an optimal framework for heart attack prediction in balanced high dimensional big data.

The rest of the paper is structured as follows: Section II includes the existing works related to various big data analytics in healthcare, data mining in healthcare, machine learning in healthcare and artificial neural networks in healthcare. Section III describes the implementation details of the proposed heart attack prediction model in big data environment. Finally, section IV discusses the results and discussions of the proposed work, and section V presents the conclusion of this paper.

II. RELATED WORKS

Big Data Analytics in Healthcare

Core providers of healthcare services (such as nurses, doctors, lab technicians, etc.), critical healthcare services (such as health insurance policies and research based on medical field) and the recipients of healthcare services (public people and patient) are the three entities of medical healthcare system. Applications of Big Data Analytics (BDA) are five-fold: health awareness among common people, interactions among stakeholders in the healthcare environment, hospital administration practices, treatment of particular diseases and technology used in healthcare [8]. Healthcare defined three levels of big data analytics with enhanced functionality and values: descriptive, predictive and prescriptive. Descriptive level reports current situations and problems. The predictive level simulates and models the AI techniques. Finally, the prescriptive level optimizes clinical, financial and other outcomes [9].

MEDLINE, Web of Science, Embase, Scopus, Cochrane Database of Systematic Reviews via Cochrane Library and Epistemonikos are six databases used for big data analytics. In addition, patients' data are collected from Electronic Health Records (EHR), private hospital's patient databases and general public hospital's information systems. "Probability of dying from cardiovascular disease was the most frequently assessed health indicator and core priority within the WHO (World Health Organization) General Programme of Work 2019/2023 [10].

Hadoop, MapReduce, Google Big Query, Microsoft Windows Azure and Jaql are some of the tools used for analyzing big data. Frequency matching and pattern matching are the diagnostic methods used for heart disease with the help of ECG patterns [11]. Clinical prediction models of supervised learning methods are classified into three types: statistical methods (for example. Linear regression and Bayesian model, etc.), sophisticated methods in ML and Data Mining (DM) (for example. Decision trees and ANN) and survival methods to predict survival outcomes. These prediction models discover the relationship between several attributes and an outcome variable [12].

Data Mining in Healthcare

Data mining is an approach that efficiently mines data from a large database to get necessary data. Predictive and descriptive models are the two types of data mining tasks. Constructive diagnosis and treatment of the disease, detection of misuse and deceit, inpatient records, etc., are examples of data mining [13].

The Data pyramid shows the architecture and illustrates data management in figure 1.a. It demonstrates the data flow from raw data through an analytical engine in the Hadoop environment to get the required output. Data lifecycle shows the different processes in data such as data collection, data pre-processing, data reduction and transformation, data analytics and data output, as in figure 1.b [14].

A multinomial Naïve Bayes algorithm was proposed by Saqlin et al. to diagnose heart failure. Thirty variables data set is used. Various classification algorithms (Neural networks, Decision Tree, Logistic Regression, et.) were used for comparison. An accuracy of 86.7% and a curvature area of 92.4% are achieved [15].

Disease diversity, heterogeneity of treatment and outcome, etc., are the characteristics of medical data. Big data mining I medical field are more difficult to obtain. They are based on several formats and organized data. Therefore, professional knowledge is important to analyze and interpret results. Surveillance Epidemiology and End Results (SEER), Medical Information Mart for Intensive Care (MIMIC), China Health and Nutrition Survey (CHNS), Health and Retirement Research (HRS) and Gene Expression Profiling Interactive Analysis (GEPIA) are some of the public medical database [16].

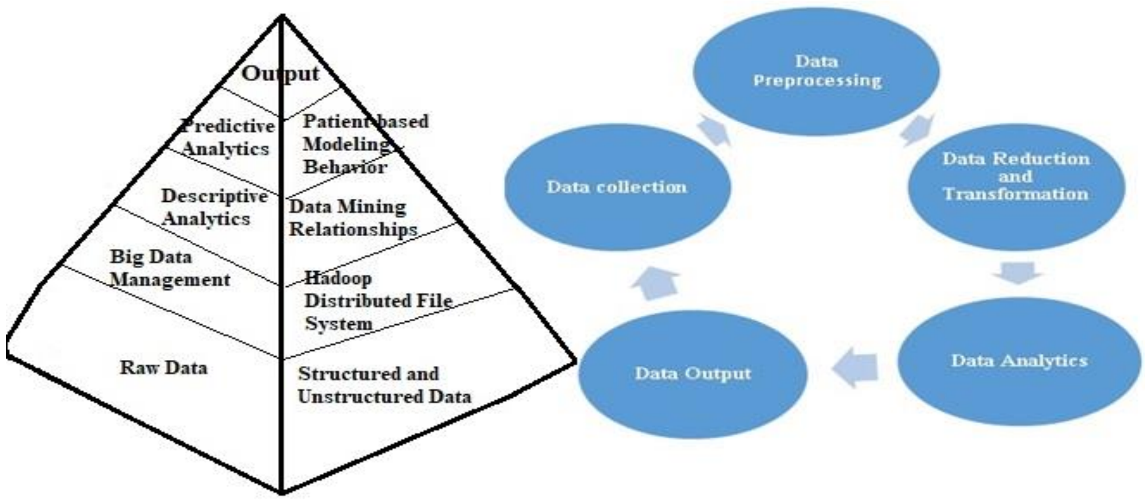


FIGURE 1. a) *Data Pyramid* b) *Data Lifecycle*

WEKA (Waikato Environment for Knowledge Analysis), KEEL (Knowledge Extraction Based on Evolutionary Learning), KNIME (Konstanz Information Miner), R programming language, Rapid miner and Orange are some of the data mining tools used in healthcare. In this proposed work, the R programming language is used for data mining purposes [17].

Machine Learning in Healthcare

People use natural human language to interact with a chatbot. ELIZA is the premier chatbot. Patient care chatbots improve the communication between clinic-patient and doctor-patient through remote testing, monitoring and telephone consultation. The recent chatbot can answer any query based on the Frequently Asked Questions (FAQ) using latent semantic analysis of NLP and Artificial Intelligence Markup Language (AIML) [4].

Supervised learning has classification and regression as two types. Random Forest Algorithm (RFA) can perform in both these categories. The ground method of RFA is recursion. In this algorithm, decision trees are created. RFA is used for imbalanced datasets since it is insensitive to noise. Moreover, RFA eliminates overfitting [18]. In the proposed work for heart attack prediction, a random forest algorithm is used.

DL finds very complex correlations. ANN with more than ten layers is used for this purpose. Simulation of artificial neurons is triggered. Watson of IBM AI machine and Deep Mind of Google AI machine are used for healthcare applications [1].

Artificial Neural Networks in Healthcare

Trained ANNs behave similarly to the biological neural cluster. Digital model of the human brain and can find the nonlinear relationship between dependent and independent variables. As with first and second-level headings, all words except prepositions. ANNs are used in cardiology, diagnosis, electronic signal analysis and medical image analysis [19]. For example, Senthilkumar et al. proposed a novel method to improve accuracy in predicting a heart

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 06 Issue: 08 | August - 2022 Impact Factor: 7.185 ISSN: 2582-3930

attack. A hybrid of Random Forest and Linear classification techniques is used to get improved accuracy of 88.7% [20].

A real-time heart attack prediction based on apache Spark is used to stream data against ML through in-memory calculations. Streaming processing and data storage visualization are two subparts. Spark MLib with Spark streaming is used for the first part. Apache Cassandra is used for the second part. UCI (University of California, Irvine C. A) is the data set used [21].

DL is an advanced version of NNs. DL explores more complex non-linear patterns in data mining. Arterys – one of the medical imaging clouds AI obtained clearance from the Food and Drug Administration (FDA) of the United States of America to publicize its Arterys Cardio DL package legally. The application implements AI to present editable and automated ventricle segmentation [23].

The computing of the brain-inspired Bioinspired computing. Magnetic Nano-devices display behaviors of neurons and synapses. Fully integrated spintronic Complementary Metal Oxide Semiconductor (CMOS) bioinspired hardware is proposed by Julie Grollier et al. [24].

III. HEART ATTACK PREDICTION IN MACHINE LEARNING ENVIRONMENT

The proposed work is implemented in a Pentium IV processor, 2GB RAM, 1 GB disk space and R 2.0.0 and R tool. Compiled RPMs (Red hat Package Manager) for Linux. Party package plot tree is used to show the plotting of prediction of decision tree classification algorithm. Data collection, feature selection and classification algorithms are the three modules used in this proposed work. Random forest algorithm is implemented with caret package prediction. R tool is used for classification. Misclassification error is analyzed in the decision tree. Error rate and histogram are shown in caret package prediction of random forest.

Conventional clinical practice depends on the manual patients' records. Conventional Decision Support Systems (DSS) use EHR, and other digital medical data use AI-based ML algorithms for decision making. Integrative DSSs use data mining and DNN techniques to make a decision. Fully automated clinical systems behave as e-doctors [22]. Deep Neural networks are the combination of the input layer, output layer and hundreds of hidden layers. Neural networks with hundreds of layers model complex relationships between the input, output and complex mathematical operations for the neurons are designed. The convolutional layer extracts spatial relations. Modern neural networks work with millions of parameters and consume huge amounts of resources to train.

Supervised learning of training the UCI data set is performed systematically. First, data cleaning and preprocessing are performed. Then, missing values (no recorded values for several attributes), outliers (inconsistent data) and scaling transformation of data (range of numeric attributes) are performed. In features selection, numeric variables that are restricted to have continuous values are clustered together. The reduced features set is given for any classification algorithms (in this proposed work, DT and RFA) to predict heart attack. The output of the classification algorithm is given as input to Extreme Learning Machine (ELM) for further process. Feedforward neural networks are used in ELM. Hidden nodes are assigned randomly with nonlinear transformations. ELM, with ML techniques' help, is obtained whether the patient has a heart attack or not. The data mining process is shown in figure 2.

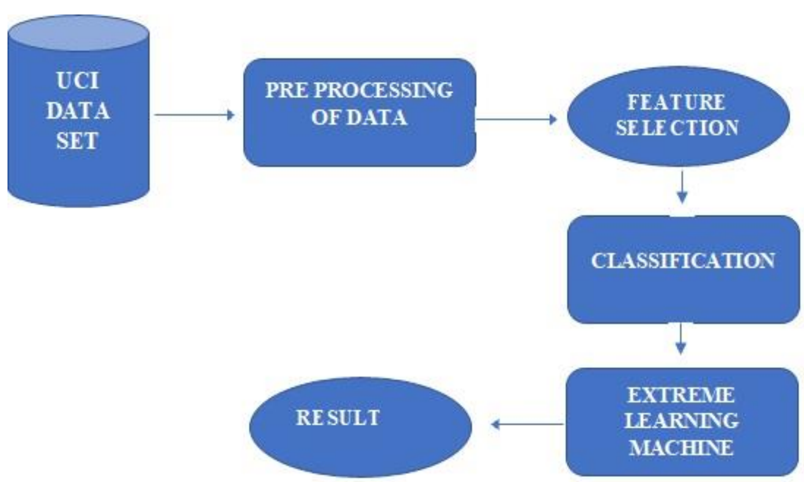


FIGURE 2. *Data Mining in Classification Algorithm*

UCI Cleveland heart diseases data set includes 304 instances with no missing values. The number of major vessels (CA), SLOPe of the peak exercise ST segment (SLOP), EXercise induced rate Achieved (EXCHANGE), RESTing Electro Cardio

Graphic (RESTECG), Cholesterol (CHOL), Resting Blood Pressure (RBP), Age, Sex, Fasting Blood Sugar (FBS), Chest Pain (CP), Maximum heart rate Achieved (THAL), ST depression induced by exercise relative to rest (OLPEAK) and reversible defect are the thirteen attributes taken for consideration to predict the heart attack. All the attributes are numerical. CP determines the pain type, ranging from 1 to 4. RBP ranges from 92 to 100. FBS is a Boolean value true of false with values 1 or 0. RESTECG varies from 0 to 2. THAL ranges from82 to 185. EXANGE is a Boolean value with 1 or 0. Part of UCI Data set is shown in figure 3.

Data cleaning, preprocessing and filtering of data set are performed. Attribute dependency is verified to find the hidden patterns and relationships. After analyzing the data and applying advanced data mining techniques, data is given as classification algorithms. The classification algorithm has two steps: the learning step and the prediction step. In the learning step, the model learns from the trained data. In the prediction step, the model predicts the response from the given input data.

A decision tree creates a training model that can predict the value of the target variable by using decision variables inferred from training data. Categorical variable decision tree and continuous variable decision tree are the two varieties of the decision tree. The decision tree follows disjunctive normal form. Algorithm selection used in the decision tree is based on the target variable. Iterative Dichotomiser 3 (ID3), C4.5 (successor of ID3), Classification and Regression Tree (CART), CHi-square Automatic Interaction Detection (CHAID) and Multivariate Adaptive Regression Splines (MARS) are the algorithms used in the decision tree. The following algorithm 1 shows the steps involved indecision tree based clustering.

Algorithm 1: Decision tree based clustering
Input: D – UCI Data set with thirteen features taken for consideration

1. For all thirteen features do
   a. For each sample do
      Execute the Decision tree algorithm
   b. End for
2. End for
3. Obtain the total number of leaf nodes $l_1, l_2, l_3, \ldots. l_n$.
4. Split the data set D into $d_1, d_2, d_3, \ldots. d_n$ based on leaf nodes.

Output: Partition data sets $d_1, d_2, d_3, \ldots. d_n$

Random forest algorithm creates a forest and makes it random. The larger the number of trees in the forest, the results are more accurate. Random forest uses a bagging algorithm and ensemble learning technique. The random forest can be used for classification and regression tasks. If there are enough trees in the forest, the classifier will not over-fit the model. Random forest creation and prediction from random forest classifier are the two stages in the random forest. It needs a longer training period. The following algorithm 2 shows the random forest based clustering algorithm.

Algorithm 2: Random forest based clustering
Input: D – UCI data set with thirteen features taken for consideration

1. For all thirteen features do
   a. For each sample do
      Execute the random forest algorithm
   b. End for
2. End for
3. Obtain the total number of leaf nodes $l_1, l_2, l_3, \ldots. l_n$.
4. Split the data set D into $d_1, d_2, d_3, \ldots. d_n$ based on leaf nodes.

Output: Partition data sets $d_1, d_2, d_3, \ldots. d_n$

"Read CSV" operator of rapid miner tool is first loaded for secondary data retained in Comma Separated Values (CSV) file. Then, only a subset of data is selected from the loaded data. "The operator uses select Attributes" for selecting the subset. Finally, it is given to the "X-Validation" operator. The AdaBoost technique is implemented for classification by the R tool. Results of the decision tree and random forest are compared. The accuracy of a random forest is greater than a decision tree. The error rate of the random forest classifier is lesser than the decision tree. The proposed system architecture is shown in figure 4.

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 06 Issue: 08 | August - 2022 Impact Factor: 7.185 ISSN: 2582-3930

Age	Sex	chesp pain type	resting bp	cholestrol	fasting blood sugar	electrocardiographic	max heart rate	exercise induced angina	oldpeak	slope of peak exercise	ca	thal	nsp
63	1	1	145	233	1	2	150	0	2.3	3	0	6	1
67	1	4	160	286	0	2	108	1	1.5	2	3	3	2
67	1	4	120	229	0	2	129	1	2.6	2	2	7	1
37	1	3	130	250	0	0	187	0	3.5	3	0	3	3
41	0	2	130	204	0	2	172	0	1.4	1	0	3	1
56	1	2	120	236	0	0	178	0	0.8	1	0	3	2
62	0	4	140	268	0	2	160	0	3.6	3	2	3	3
57	0	4	120	354	0	0	163	1	0.6	1	0	3	2
63	1	4	130	254	0	2	147	0	1.4	2	1	7	2
53	1	4	140	203	1	2	155	1	3.1	3	0	7	1
57	1	4	140	192	0	0	148	0	0.4	2	0	6	3
56	0	2	140	294	0	2	153	0	1.3	2	0	3	1
56	1	3	130	256	1	2	142	1	0.6	2	1	6	2
44	1	2	120	263	0	0	173	0	0	1	0	7	2
52	1	3	172	199	1	0	162	0	0.5	1	0	7	1
57	1	3	150	168	0	0	174	0	1.6	1	0	3	

FIGURE 3. *Part of UCI Data Set*

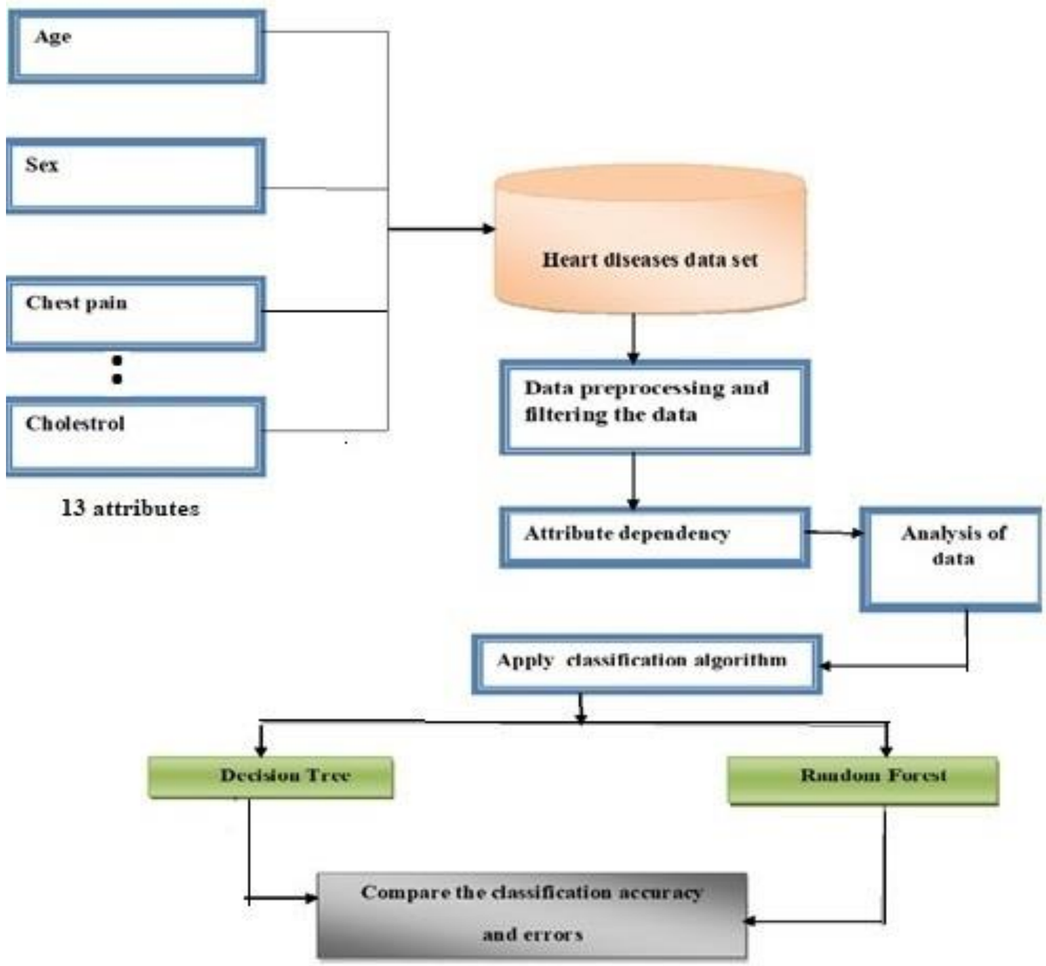


FIGURE 4. *Proposed Framework for Heart Attack Prediction*

IV. RESULTS AND DISCUSSION

R tool displays the predictions such as n-normal, s-suspect and p-pathological to depict the results. The performance of the classification algorithm is analyzed with thirteen parameters. If a classification algorithm is trained, the predictive model will present the best accuracy with the help of the following classifications:

- TP (True Positive) – Number of patients who are correctly classified as having the disease
- TN (True Negative) – Number of patients who are correctly classified as healthy
- FP (False Positive) – Number of patients who are misclassified as healthy
- FN (False Negative) – Number of patients who are misclassified as having the disease

The data set is prepared and partitioned. A decision tree with a party package of the R tool is created. It is shown in figure 5. Prediction is applied to the trained data. Partitioning of the predicted data set is performed. A decision tree is created with rpart package. Data of tree are validated. Misclassification errors are found and validated.

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 06 Issue: 08 | August - 2022 Impact Factor: 7.185 ISSN: 2582-3930

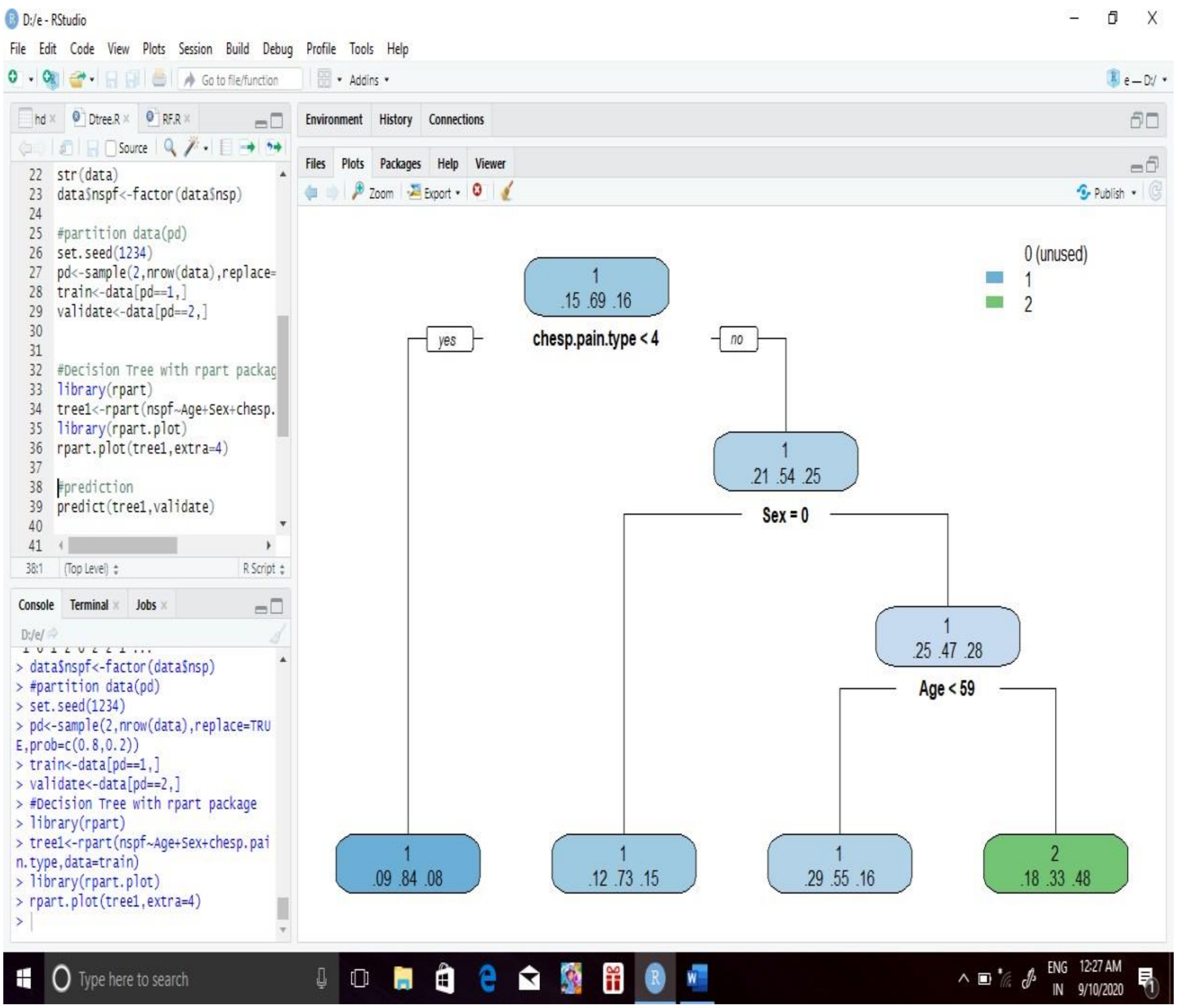


FIGURE 5. *Decision Tree Created using Party Package*

The data set is prepared and partitioned for the random forest. Random forest is created using a caret package. To train the data, prediction is performed, and a confusion matrix is created. The confusion matrix created in the random forest classifier is shown in Figure 6. The error rate of random forest is plotted in the graph. Misclassification errors are validated. The accuracy of the two classifier algorithms is shown in table 1 and figure 7.a. Error rate of two classifier algorithms is shown in table 2 and figure 7.b.

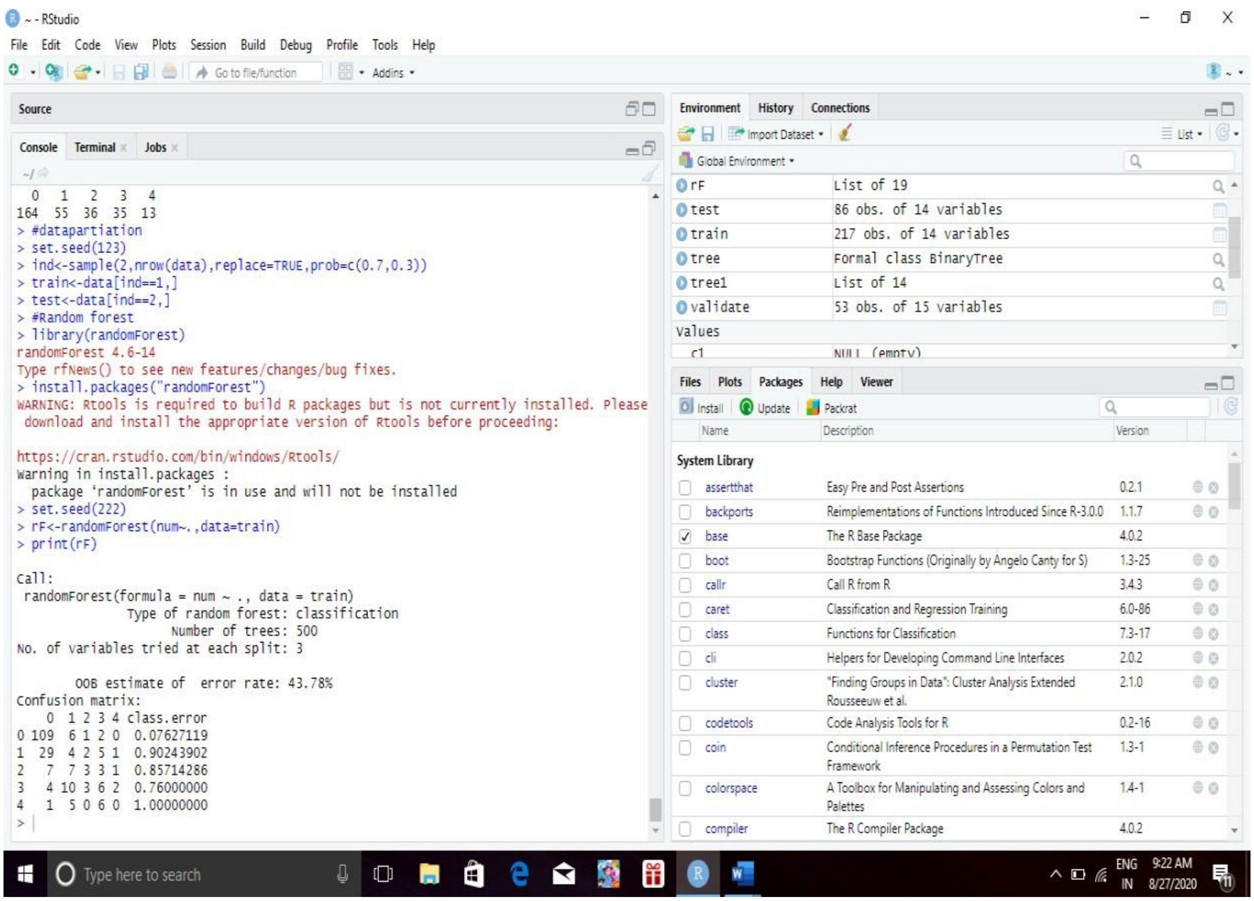


FIGURE 6. *Random Forest Confusion Matrix*

Accuracy is the fraction of many correct predictions over the total number of predictions.

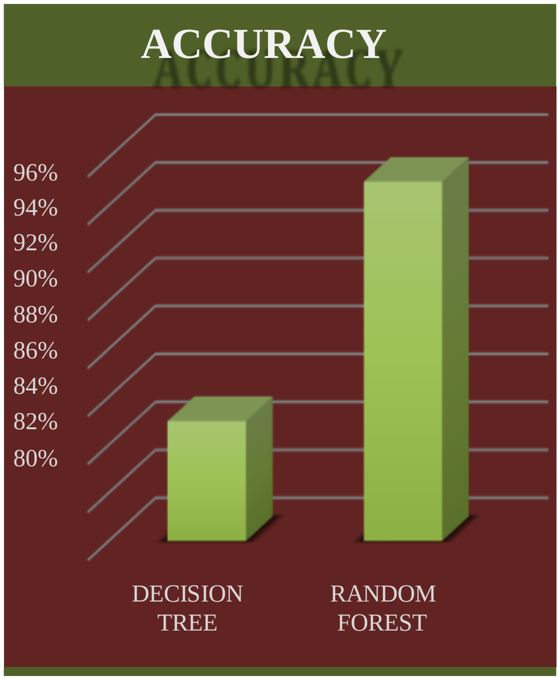
TABLE 1. *Accuracy of two classifiers*

Algorithm	Accuracy
Decision Tree	85%
Random Forest	95%

The error rate is based on the wrong predictions in the testing process. Therefore, error Rate is calculated as (1 – (Number of wrong predictions / Total number of predictions)).

TABLE 2. *Error Rate of two classifiers*

Algorithm	Error Rate
Decision Tree	47%
Random Forest	43%



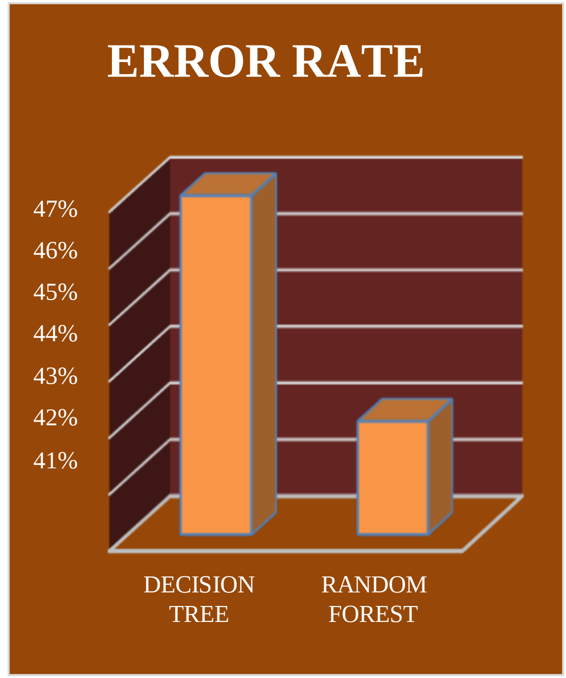


FIGURE 7. a) *Accuracy* of Classifiers b) *Error Rate of Classifiers*

Results are compared with the existing work [20]. Senthilkumar Mohan et al. proved that their research work provides 88.4% with the hybrid random forest with a linear model (HRFLM) when compared with Naïve Bayes, generalized linear model, logistic regression, deep learning, etc. The proposed research work provides 95% accuracy rate which is better than exiting work.

V. CONCLUSION

Machine Learning is a flourishing field of medicine to fuse statistics and computer science into the medical field, and the process is semi-automated. ML is the future of biomedical research, computer-aided diagnosis and personalized medicine to enhance global healthcare. Public illiteracy about AI increases the burden of trust in AI. According to a South Korean survey, only 5.9% of physicians know AI techniques from a professional perspective. Some people will not trust the government to safeguard personal EHR in the research field. There are many ethical values and hazards to be overridden to implement AI techniques in the more sensitive medical field. The predictive model will predict the accuracy of heart attack based on four classifications: True Positive (TP), True Negative, False Positive and False Negative. Sensitivity and specificity denote the persons' ability to identify TP rate and TN rate. To conclude the work, the Random forest algorithm has the highest classification accuracy performance with the lowest error rate compared with the decision tree classification algorithm to predict the heart attack.

REFERENCES

1. Adam Bohr and Kaveh Memarzadeh,"Chapter 2: The Rise of Artificial Intelligence in Healthcare Applications", Artificial Intelligence in Healthcare, DOI: https: //doi.org/10.1016/B978-0-12-818438- 7.00002-2, 2020, pp. 25-60.
2. Thomas Davenport and Ravi Kalakota, "The Potential for Artificial Intelligence in Healthcare", *Royal College of Physicians Future Healthcare Journal*, June 2019, Vol. 6, Issue 2, pp. 94 – 98, DOI: 10.7861/futurehosp.6-2-94.
3. Konieczny L and Roterman I., "Personalized precision medicine", Bio-Algorithms Med-Syst 2019, 15.
4. Sarkar Siddique and James C.L. Chow, "Machine Learning in Healthcare Communications", MDPI Encyclopedia 2021, 1, pp.220-239, https://doi.org/10.3390/encyclopedia1010021.
5. Guler S, Hurton S, C winn M and Molinari M, "Levels in Decision Making and Techniques for Clinician", International Journal of Digestive Diseases. 2015; 1(1:2).

6. Weng, W. H., Wagholikar, K. B, McCray, A. T, Szolovits, P. and Chueh, H.C, "Medical Subdomain Classification of Clinical Notes Using A Machine Learning Based Natural Language Processing Approach", BMC Med. Inform. Decis. Mak, 2017, 17, pp.1-13.
7. Blease C, Kaptchuk TJ, Bernstein MH, Mandl KD, Halamka JD and Desroches CM, "Artificial intelligence and the future of primary care: exploratory qualitative study of UK general practitioners' views", J. Med Internet Res 2019, 21:e12802.
8. Sayantan Khanra, Amandeep Dhir, A. K. M. Najmul Islam and Matti Mäntymäki, "Big Data Analytics in Healthcare: A Systematic Literature Review", Enterprise Information Systems, 14:7, 878-912, DOI: 10.1080/17517575.2020, 1812005, Oct 2020.
9. William R. Hersh, "Healthcare Data Analytics", Hoyt, RE, Yoshihashi, A, Eds.(2014), Health Informatics: Practical Guide for Healthcare and Information Technology Professionals, Sixth Edition, Pensacola, FL, Lulu.com., pp. 1-13, Chapter 3.
10. Israel Júnior Borges do Nascimento , Milena Soriano Marcolino, Hebatullah Mohamed Abdulazeem, Ishanka Weerasekara, Natasha Azzopardi-Muscat, Marcos André Gonçalves and David Novillo-Ortiz, "Impact Of Big Data Analytics On People's Health: Overview of Systematic Reviews and Recommendations for Future Studies", J Med Internet Res. 2021 Apr 13;23(4):e27275, DOI: 10.2196/27275.
11. Lidong Wang and Cheryl Ann Alexander, "Big Data Analytics in Healthcare Systems", International Journal of Mathematical, Engineering and Management Sciences Vol. 4, No. 1, pp. 17–26, 2019.
12. Chandan K. Reddy and Charu C. Aggarwal, "Healthcare Data Analytics", Chapman and Hall/CRC Press, June 2020, pp. 1-35.
13. Rakhi Ray, "Advances in Data Mining: Healthcare Applications", International Research Journal of Engineering and Technology (IRJET), Volume: 05, Issue: 03, Mar-2018, pp. 3738-3742.
14. Revanth Sonnatti, "Improving Healthcare Using Big Data Analytics", International Journal of Scientific & Technology Research, Volume 6, Issue 03, March 2017, pp. 142-146.
15. M. Saqlain, W. Hussain, N. A. Saqib, Nazar, and M.A. Kha,"Identification of Heart Failure by Using Unstructured Data of Cardiac Patients," in Proc. 45th International Conference on Parallel Processing Workshops, pp.426-431, 2016.
16. Jin Yang, Yuanjie Li, Qingqing Liu, Li Li, Aozi Feng, TianyiWang, Shuai Zheng, Anding Xu and Jun Lyu, "Brief introduction of medical database and datamining technology in big data era", *J Evid BasedMed.* 2020, Vol. 13, pp. 57–69, wileyonlinelibrary.com/journal/jebm, DOI: 10.1111/jebm.12373.
17. R. Pallavi Reddy, Ch. Mandakini and Ch. Radhika, "A Review on Data Mining Techniques and Challenges in Medical Field", International Journal of Engineering Research & Technology (IJERT), Vol. 9 Issue 08, August-2020, pp. 329 – 333.
18. Senerath Mudalige Don Alexis Chinthaka Jayatilake and Gamage Upeksha Ganegoda1, "Involvement of Machine Learning Tools in Healthcare Decision Making", Computational Decision-Making Tools for Healthcare, Volume 2021 | Article ID 6679512 | https://doi.org/10.1155/2021/6679512, pp. 1-34.
19. Jignesh Patel and Ramesh K Goyal, "Applications of Artificial Neural Networks in Medical Science", Artificial Intelligence in Drug Discovery and Development Project, Current Clinical Pharmacology, Vol. 2(3), pp. 217-26, October 2007, DOI: 10.2174/157488407781668811, PubMed Source.
20. Senthilkumar Mohan, Chandrasegar Thirumalai and Gautam Srivastava, "Effective Heart Disease Prediction Using Hybrid Machine Learning Techniques", DOI: 10.1109/ACCESS.2019.2923707, Vol. 7, July 2019, pp. 81542 – 81554.
21. Abderrahmane Ed-daoudy and Khalil Maalmi, "Real-Time Machine Learning for Early Detection of Heart Disease using Big Data Approach", IEEE Xplore, International Conference on Wireless Technologies, Embedded and Intelligent Systems (WITS), 2019.
22. Kun-Hsing Yu, Andrew L. Beam and Isaac S. Kohane, "Artificial Intelligence in Healthcare", Nature Biomedical Engineering, Vol. 2, October 2018, www.nature.com/natbiomedeng, pp. 719-731.
23. Bernard Marr, "First FDA approval for clinical Cloud-Based Deep Learning in Healthcare", 2017, https://www. forbes.com/ sites/ bernardmarr/ 2017/01/ 20/ first- fda- approval- for- clinical- cloud- based- deep- learning- in healthcare/#7a0ed8dc161c (accessed 1 Jun 2017).
24. Julie Grollier, Damien Querlioz and Mark D. Stiles, "Spintronic Nano devices for Bioinspired Computing", Journals & Magazines, Proceedings of the IEEE, DOI: 10.1109/JPROC.2016.2597152, Oct 2016, Vol. 104, Issue 10, pp. 2024-2039.