A Relative Research of Machine Learning Build Heart Disease Forecast Model

Preksha Agrawal, Kailash Soni, Mukesh Kumar Gupta
Department of Computer Science & Engineering, Swami Keshvanand Institute of Technology, Management & Gramothan Jaipur-302017 (INDIA)
Email: **prekshaag84@gmail.com, kailash.100ni.er@gmail.com, mukeshgupta@skit.ac.in**
Received 17.02.2020 received in revised form 16.10.2020, accepted 17.10.2020

Abstract- Nowadays, heart disease is the leading cause of death globally, with an estimated 610000 lives each year due to heart condition. One of the most common causes of heart disease is high blood pressure (HBP), fasting blood sugar (FBS), diabetes, cholesterol, Body mass Index (BMI), heart rate (HR). Diagnosis of heart disease is more prevalent nowadays; this involves a lot of accuracy and uncertainty due to the large-scale data decision-based on doctors may fail in some cases. Data mining is an intelligent diagnostic tool in healthcare. Thus, it is imperative to predict that each menace stage depends on age, sex, blood pressure, diabetes symptoms, what we can do for precaution by diagnosing the disease and proper treatment at the right moment. The purpose of the research work is to develop different predictive models using different forecasting measures and perform comparative analysis. In this work, we have used Cleveland and Statlog datasets with Naive Bayes (NB), K-Nearest Neighbors (KNN), and Logistic Regression (LR), Support vector machine (SVM), Decision Tree (DT), and Random Forest (RF) Classifier to develop various predictors. The experiment result shows that the random forest classifier gives better accuracy and results on both datasets when we perform a comparative analysis of them among all other classification models

Keywords– Heart ailment, forecast classification, facts removal Classifier, Ensemble Learning, Random Forest (RF).

1. INTRODUCTION

The heart consists of strength to pumps blood; arteries to provide blood to the heart muscle, if the blood delivers to a part of the heart strength is completely block heart failure can take place. General symptoms' include High blood pressure (HBP), Fasting Blood Sugar (FBS), Cholesterol, BMI, Heart Rate (HR); Diabetes is the main cause of heart failure and stroke[1], [2].

1.1 Fact Removal Method

In Healthcare industry Fact Removal Method is used for diagnosing and detecting patients' disease before time using healthcare data present on UCI Machine Learning Repository over worldwide[3]. Medical support from data helps experts to diagnose patients' symptoms and provide cost effective treatment for individuals [4].

1.2 Machine Learning Based Approach

Machine learning is a simulated intellect model so as to provide various algorithms to make computers more intelligent to make a computer an intelligent machine[5], [6]. Classification methods are the mainly extensively used algorithm in the healthcare field because it helps predict the status of patients by classifying patients' records and locating the class that matches the new patient records[7], [8]. This paper presents an introduction to the classification algorithm used in our comparative study of heart disease prediction models using Naïve Bayes (NB), KNN, SVM, LR, Decision Tree (DT), Random Forest (RF) classifier on healthcare dataset and diagnostic the patients' ailment from patients is inflated Becomes challenging when training and testing data is from a different domain[9], [10].

Each technique has a different way of creating a classifier, which ensures that these techniques behave differently and produce different results[11]. where the original dataset differs Different training and testing are divided into sets[7][12]. In each validation cycle, an observation is conducted where the remaining observations serve as training sets and they are used to create the classifier model[13].

2. PROBLEM DEFINITION

In this research work, A comparative study of two dataset are used for prediction and analysis of heart disease with examining patient symptoms by means of facts removal categorization methods, to achieve this goal, a literature review to review data mining operations connected to the analysis of heart ailment was done[14].

Six classifiers (e.g., Naive Bayes (NB), Decision Tree (DT), Support Vector Machine (SVM), K-Nearest Neighbor (KNN), Logistic Regression (LR), and Random Forest (RF)) were selected to create the model with the maximum accuracy possible[5], [15]–[17].

We have also explored precision score, recall score, F-score, false negative using confusion matrix for every algorithm used[8], [10], [16], [18].

Table 1: Literature Survey

Author	Purpose	Techniques used and accuracy
C.Ordonez (2001)[20]	Mining association rules and identifying useful constraints	LM,LAD,LCX,RCA 70%
F.Le Duff (2007)[21]	533 patients who had suffered from cardiac arrest	Naïve Bayes (NB) Mean age is 63 73% men & 27% women.
S. Palaniappan (2008)[4]	Discovery of hidden pattern and relationship using heart disease prediction system	Decision Tree (DT) (80.4%) Neural Network (NN) (85.68%) Naïve Bayes (NB) (86.12%)
L.Parthiban (2008)[22]	Heart Disease Prediction using CANFIS and Genetic Algorithm	Fuzzy Logic (FL), Neural Network
S.Dangare (2012)[23]	Improved Study of Heart Disease Prediction Using Data Mining Classification Technique	Neural Network=100% Decision Tree=99.62% Naïve Bayes=90.74%
S.U. Amin (2013)[24]	Genetic Neural Network Based Data Mining in Prediction of Heart Disease Using Risk Factor	Neural Network (NN) =89% Genetic Algorithms
E.Yilmaz (2013)[25]	Determination of fetal state from cardiotocogram using LS-SVM	Quadratic Support Vector Machine (LS-SVM) = 91.62%
B.Tarley (2017)[26]	Improved Artificial Neural Network for dimension reduction in medical data classification	OLPP Cleveland dataset (92.0%) Hungarian dataset (81.1%) Switzerland dataset (99.0%)
S.D.Desai (2019)[9]	Evaluate the precision of characterization models for the illness	BPNN =85.074% LR=92.58% Neural Network (NN)

Our goal is a comparative study of Machine Learning based Heart disease prediction models to ensure occurrence or nonappearance of disease or which affecting the patient on the basis of less number of attribute[13].

3. RELATED RESEARCH

Disease Detection is the important part of the healthcare industry. Many Researchers are working on heart disease prediction for better result[19]. Table 1 summaries the literature survey of existing research in prediction and detection ailment.

3. PROPOSED METHODOLOGY

3.1 Naïve Bayes Classifier (NB)

NB that uses the probabilistic classifier[11]

$$P(A/B) = \frac{P(B/A) * P(A)}{P(B)}$$

Where A and B are two events, Such as NB) classifiers use probability theory to find the most likely classification of an unseen instance. This algorithm performs positively but poorly with hierarchical data[27]

3.2 K-nearest neighbor Classifier (KNN)

KNN [2] is a recognized technique for classifying a hidden instance, which uses a categorization of its adjacent instances. The essential KNN categorization algorithm mechanism by verdict K training instances that is close to the hidden example using distance procedures[2], [28]. The algorithm after that sets the group for the hidden example taking the most happening class in the adjacent K instances.

3.3 Decision Tree Classifiers (DT)

The DT [16], [29] is a non-invariable supervised learning technique worn for classification and regression. The aim is to generate a replica that predicts the rate of the goal unpredictable by knowledge that is incidental from facts features. Several compensation comprise easy to recognize plus make clear. Trees can be anticipated, requiring very modest data. Additional techniques frequently need facts normalization, creating data variables and extracting unfilled values[8], [21].

3.4 Support Vector Machine Classifier (SVM)

In machine learning, the SVM [10], [25] is a monitored learning model with an Associated Learning algorithm that analyzes the data used for classification and regression analysis. Given a set of training examples, each marked for one of the two categories, SVM training creates a model that provides new examples in one category or another, making it a non-probabilistic classifier goes.

3.5 Logistic Regression Classifier (LR)

LR [9], [17], [30] provides high accurateness plus chart illustration. In this algorithm, facts must be imported primary and then trained. The LR algorithm by the equation is revealed in the graph viewing the differences among the attributes. From the training facts, we encompass to estimate and correspond to finest and predictable coefficients.

Table 2: Attribute Description

S.No.	Feature Name	Feature Information
1	Age	Age of patients in existence
2	Sex	1= male 0=female
3	Cp	Value 1: Typical Angina Value 2: Atypical Angina Value 3: Non-Angina Pain Value 4:Asymptotamatic
4	Trestbps	Resting Blood Pressure in mm Hg in admittance to the sanatorium
5	Chol	Serum Cholesterol of patients calculated in mg/dl
6	Fbs	Fasting blood Sugar of the patient. If >120 mg/dl Value 1 =accurate Value 0 =fake
7	Restecg	Value 0 = Usual Value 1 = having ST-T Wave irregularity Value 2 =Showing likely or exact Left Ventricular Hypertrophy by Estes' criterion
8	Thalach	Maximum Heart pace Achieved of Patient
9	Exang	Value 1= positively Value 0 = negatively
10	Oldpeak	ST Depression induced by exercise relative to rest
11	Slope	Value 1 = Up Sloping Value 2 =Flat Value 3 = Down Sloping
12	Ca	quantity of foremost vessel{0-3}
13	Thal	Value 1 = Usual Value 2 = Permanent Defected Value 3 = Reversible Defect
14	Target	Numerical value 0 to 4

3.6 Random Forest Classifier (RF)

RF [5], [31] is an ensemble algorithm to constructs a set of decision trees as of a random example of training sets. It repeats the procedure with numerous random samples as well as makes a last choice based on majority selection. RF algorithm is effectual in management of absent values [6], [32].

In this work, we used the Cleveland and Statlog datasets with Naive Bayes, KNN, Logistic regression, SVM, Decision Tree, Random Forest classifier to create various speculation models. The test result shows that the Random Forest (RF) Classifier provides better accuracy and has an effect on both sets of data when we perform a comparative analysis of it among all other classification models.

3.7 Dataset Structure and description

In this paper, we used two dataset for study. Cleveland dataset contain 303 records, 297 complete and 6 with missing/unknown values and Statlog dataset having 270 complete records[5]. Both datasets have 76 attributes, but we have to use 14 features in an attempt to measure which disease affecting the patients'[5], [33]. The attribute num is the heart disease diagnosis attribute. It was classified as presence and absence[8], [34]. If it is presence, then value of num would be low or medium or high or very high. If it is absence, then value of num would be zero[35].

3.8 Data Collection

The heart disease dataset can be easily downloaded from the machine available at the UCI repository[11], [19]. We use 14 of them to obtain accurate results with a small number of attributes[12], [36][37], [38]. The "Target" field refers to the occurrence of heart disease in the patient (target). It is numeral valued from 0 (no occurrence) to 4 (occurrence)[23]. The main advantage of using this dataset is that it can be communal with other manufactures or health experts. Advice from other manufactures may be helpful in improving the results of the proposed framework[18]. Table 2 describe the characteristics and their possible data types or value-selected in the heart disease dataset.

Algorithm for Performance Evaluation

(1) Load data and Import Libraries for calculate number of records and missing value from large infinite data.
(2) Convert large data into categorical valued data for feature selection & fix missing value by applying Data Pre-Processing.
(3) Use the correlation matrix for depicting the relation between feature attribute targets and find the strong and negative correlation between them.
(4) Apply Visualization technique for Data Analysis of feature selection of relevant data using Standard Scalar to fit & and transform the data into 0 and 1 form which is easy for predicting disease.
(5) Split the dataset into 80% training& 20% testing data set to fit the parameter and assess the performance of model respectively.
(6) Apply Machine Learning algorithm on both dataset and compare the performance using performance matrix for depicting which is better algorithm.
(7) After modelling and predicting with Machine Learning Classifier Check the Accuracy Score, Precision, Recall, and F-Measure of both dataset.
(8) Evaluate Result of both dataset and calculate performance measure by plotting graph of different classifier.

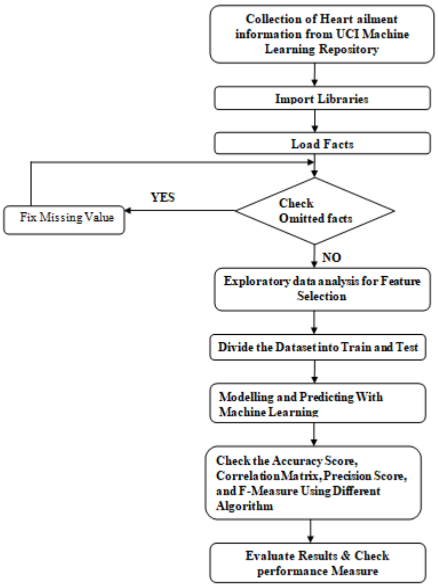

Figure 1: Flow Chart of Heart Disease Prediction

3.9 Data Pre-Processing

In this experiment, six missing values of instances in the Cleveland dataset were cleaned and transformed by data pre-processing, only 297 instances were taken for this study[31], [37]. For the pre-processing of data, there are several steps such as cleaning, normalization, transformation, etc[18].

3.10 Cleveland Heart Disease Dataset

	age	sex	chest_pain_type	...	num_major_vessels	thalassemia	target
56	48	1	0	...	0	1	Yes
283	40	1	0	...	0	3	No
273	58	1	0	...	1	3	No
242	64	1	0	...	2	2	No
217	63	1	0	...	3	3	No

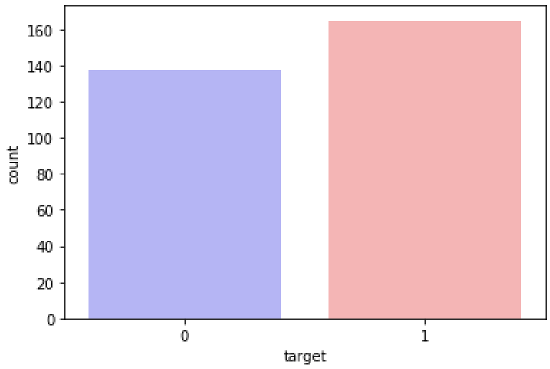

Figure 2: Diseases vs. Non-Disease Analysis

1: 165 (Disease Present)
0: 138 (Disease Not Present)

3.10.1 Correlation Matrix

A correlation matrix is a chart to shows association coefficients among variables. Every cell in the table shows the correlation between two variables. [18], [35].

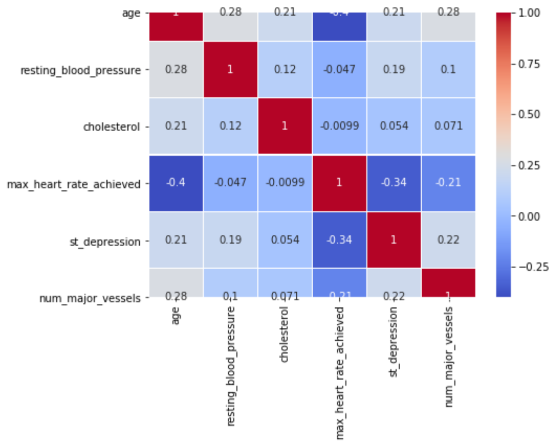

Figure 3: Correlation Matrix

3.11 Statlog Heart Disease Dataset

	age	sex	chest	...	number_of_major_vessels	thal	target
5	65	1	4	...	0	7	absent
137	56	1	4	...	1	3	present
85	42	1	1	...	2	3	absent
44	59	1	3	...	1	6	present
45	58	1	3	...	0	3	absent

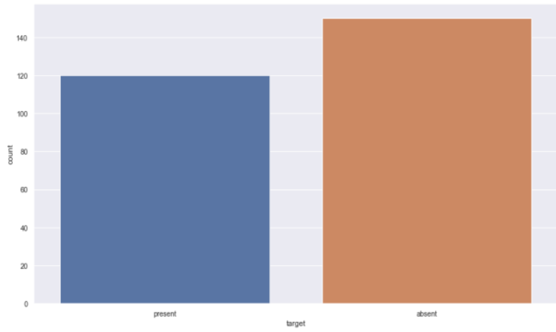

Figure 4: Diseases vs. Non-Disease Analysis

Present: 120
Absent 150

4 EVALUATION AND RESULT

$$\text{Precision} = \frac{TP}{TP + FP} \quad \ldots\ldots\ldots\ldots 1$$

$$\text{Recall} = \frac{TP}{TP + FN} \quad \ldots\ldots\ldots\ldots 2$$

$$F1 = \frac{2 * \text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}} \quad \ldots\ldots\ldots\ldots 3$$

$$\text{Accuracy} = \frac{TP}{TP + FP + FN + TN} \quad \ldots\ldots\ldots\ldots 4$$

$$\text{Specificity} = \frac{TN}{TN + FP} \quad \ldots\ldots\ldots\ldots 5$$

(1) Positive (P): Observation is optimistic
(2) Negative (N): Observation is not optimistic

(3) True Positive (TP): Observation is optimistic, and is predicted to be optimistic.
(4) False Negative (FN): Observation is optimistic, but is predicted pessimistic.
(5) True Negative (TN): Observation is pessimistic, and is predicted to be pessimistic.
(6) False Positive (FP): Observation is pessimistic, but is predicted optimistically.

3.11.1 Correlation Matrix

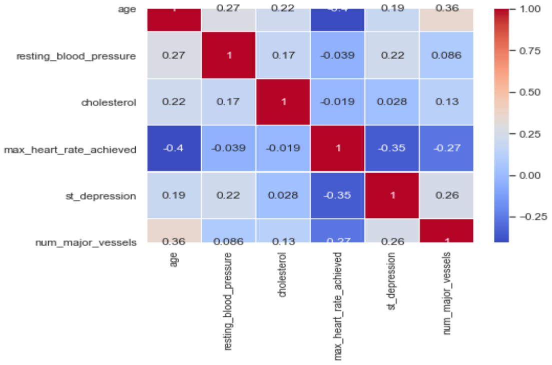

Figure 5: Correlation Matrix

4.1 Performance Measure

4.1.1 Cleveland Heart Disease Dataset

Table 3: Attribute Description

Performance Measure	Decision Tree	Naive Bayes	KNN	SVM	Logistic Regression	Random Forest
Accuracy	0.786885	0.901639	0.868852	0.859259	0.852459	0.918032
Support	61	61	61	61	61	61
Precision	0.79	0.90	0.87	0.85	0.85	0.88
Recall	0.79	0.90	0.87	0.85	0.85	0.87
F- Measure	0.79	0.79	0.87	0.85	0.85	0.87

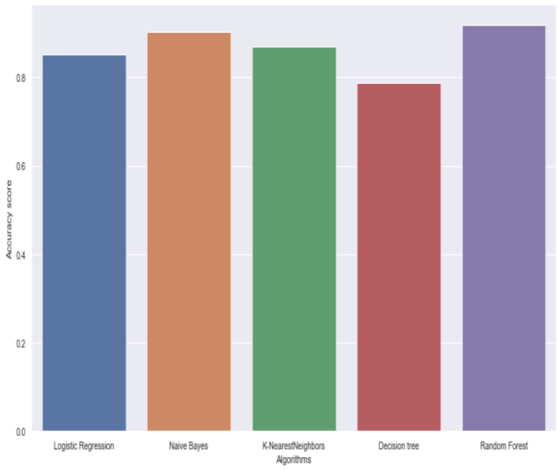

Figure 6: Accuracy Score Using Cleveland dataset

4.1.2 Statlog Heart Disease Dataset

Table 4: Attribute Description

Performance Measure	Decision Tree	Naive Bayes	KNN	SVM	Logistic Regression	Random Forest
Accuracy	0.851851	0.740740	0.6851851	0.759259	0.777777	0.8703703
Support	54	54	54	54	54	54
Precision	0.85	0.74	0.69	0.76	0.78	0.87
Recall	0.84	0.74	0.68	0.75	0.78	0.86
F- Measure	0.85	0.74	0.69	0.76	0.78	0.87

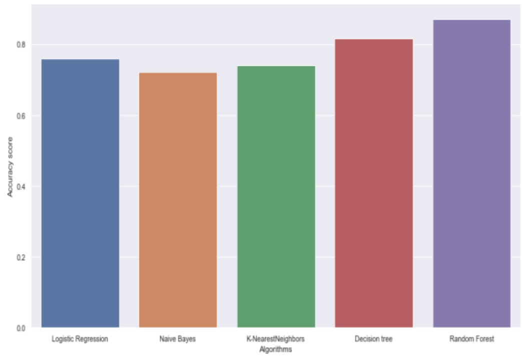

Figure 7: Accuracy Score Using Statlog dataset

5 CONCLUSION AND FUTURE WORK

This study concluded that if we do not have information regarding patients' disease and symptoms, an individual cannot survive for a long duration. Exact detection can help in maintaining the HBP, diabetes, cholesterol, BMI and more. The precise information support experts to diagnose the disease at an early stage and provide the right handling at the time.

Specifically, six classifiers were utilized and implemented using the Sypder tool to emulate health decision making with improved accuracy. The main purpose of our research work is to build a prediction model and perform comparative analysis using performance measures. The Experiment result shows that Random Forest Classifier gives better accuracy and result on both datasets when we do a comparative analysis of them among all other classification models.

In the future, Automated Heart disease prediction systems may be implemented in remote areas such as rural areas to replicate human specialist for diagnosis.

REFERENCES

[1] T. Marikani and K. Shyamala, "Prediction of Heart Disease using Supervised Learning Algorithms," Int. J. Comput. Appl., vol. 165, no. 5, pp. 41–44, 2017, doi: 10.5120/ijca2017913868.
[2] G. Shanmugasundaram, V. Malar Selvam, R. Saravanan, and S. Balaji, "An Investigation of Heart Disease Prediction Techniques," 2018 IEEE Int. Conf. Syst.

Comput. Autom. Networking, ICSCA 2018, pp. 1–6, 2018, doi: 10.1109/ICSCAN.2018.8541165.

[3] S. Ambekar and R. Phalnikar, "Disease Risk Prediction by Using Convolutional Neural Network," Proc. - 2018 4th Int. Conf. Comput. Commun. Control Autom. ICCUBEA 2018, pp. 1–5, 2018, doi: 10.1109/ICCUBEA.2018.8697423.

[4] S. Palaniappan and R. Awang, "Intelligent heart disease prediction system using data mining techniques," AICCSA 08 - 6th IEEE/ACS Int. Conf. Comput. Syst. Appl., pp. 108–115, 2008, doi: 10.1109/AICCSA.2008.4493524.

[5] N. S. Chandra Reddy, S. Shue Nee, L. Zhi Min, and C. Xin Ying, "Classification and Feature Selection Approaches by Machine Learning Techniques: Heart Disease Prediction," Int. J. Innov. Comput., vol. 9, no. 1, May 2019, doi: 10.11113/ijic.v9n1.210.

[6] D S. C. Jeeva, "Improving the accuracy of prediction of heart disease risk based on ensemble classification techniques," Informatics Med. Unlocked, vol. 16, no. November 2018, p. 100203, 2019, doi: 10.1016/j.imu.2019.100203.

[7] A. N. Repaka, S. D. Ravikanti, and R. G. Franklin, "Design And Implementing Heart Disease Prediction Using Naives Bayesian," 2019 3rd Int. Conf. Trends Electron. Informatics, no. Icoei, pp. 292–297, 2019, doi: 10.1109/icoei.2019.8862604.

[8] M. Gandhi and S. N. Singh, "Predictions in heart disease using techniques of data mining," 2015 1st Int. Conf. Futur. Trends Comput. Anal. Knowl. Manag. ABLAZE 2015, pp. 520–525, 2015, doi: 10.1109/ABLAZE.2015.7154917.

[9] S. D. Desai, S. Giraddi, P. Narayankar, N. R. Pudakalakatti, and S. Sulegaon, "Back-propagation neural network versus logistic regression in heart disease classification," in Advances in Intelligent Systems and Computing, 2019, vol. 702, pp. 133–144, doi: 10.1007/978-981-13-0680-8_13.

[10] S. O. Choi et al., "Stability Analysis Of Jointed Rock Slopes With The Barton-Bandis Constitutive Model In Udec Sung O . Choi1 , So-Keul Chung2 1) Korea Institute of Geoscience & Mineral Resources , Daejon , Korea 2) Korea Institute of Geoscience & Mineral Resources , Daej," Int. J. Rock Mech. Min. Sci., vol. 41, no. 3, pp. 1–6, 2004, doi: 10.1109/ICCIT.2007.204.

[11] A. Patel, S. Gandhi, S. Shetty, and B. Tekwani, "Heart Disease Prediction Using Data Mining," Int. Res. J. Eng. Technol., vol. 4, no. 1, pp. 4–6, 2017.

[12] H. S. N. Murthy and M. Meenakhi, "Approach for Early Prediction of Coronary Heart Disease," Proc. Int. Conf. Circuits, Commun. Control Comput., no. November, pp. 21–22, 2014.

[13] C. B. C. Latha and S. C. Jeeva, "Improving the accuracy of prediction of heart disease risk based on ensemble classification techniques," Informatics Med. Unlocked, 2019, doi: 10.1016/j.imu.2019.100203.

[14] A. Golande and T. Pavan Kumar, "Heart disease prediction using effective machine learning techniques," Int. J. Recent Technol. Eng., vol. 8, no. 1 Special Issue 4, pp. 944–950, 2019.

[15] K. Vembandasamy, R. Sasipriya, and E. Deepa, "Heart Diseases Detection Using Naive Bayes Algorithm," Int. J. Innov. Sci. Eng. Technol., vol. 2, no. 9, pp. 441–444, 2015.

[16] T. P. R and J. Thomas, "Proceedings of IEEE International Conference on Circuit, Power and Computing Technologies, ICCPCT 2017," Proc. IEEE Int. Conf. Circuit, Power Comput. Technol. ICCPCT 2017, 2017.

[17] M. T., D. Mukherji, N. Padalia, and A. Naidu, "A Heart Disease Prediction Model using SVM-Decision Trees-Logistic Regression (SDL)," Int. J. Comput. Appl., vol. 68, no. 16, pp. 11–15, Apr. 2013, doi: 10.5120/11662-7250.

[18] A. U. Haq, J. P. Li, M. H. Memon, S. Nazir, R. Sun, and I. Garciá-Magarinõ, "A hybrid intelligent system framework for the prediction of heart disease using machine learning algorithms," Mob. Inf. Syst., vol. 2018, 2018, doi: 10.1155/2018/3860146.

[19] P. S. Kohli and S. Arora, "Application of machine learning in disease prediction," 2018 4th Int. Conf. Comput. Commun. Autom. ICCCA 2018, pp. 1–4, 2018, doi: 10.1109/CCAA.2018.8777449.

[20] C. Ordonez et al., "Mining constrained association rules to predict heart disease," Proc. - IEEE Int. Conf. Data Mining, ICDM, pp. 433–440, 2001, doi: 10.1109/icdm.2001.989549.

[21] F. Le Duff, C. Muntean, M. Cuggia, and P. Mabo, "Predicting survival causes after out of hospital cardiac arrest using data mining method.," Medinfo, vol. 11, no. Pt 2, pp. 1256–1259, 2004.

[22] L. Parthiban and R. Subramanian, "Intelligent Heart Disease Prediction System using CANFIS and Genetic Algorithm," Int. J. Biol. Med. Sci., vol. 3, no. 3, pp. 157–160, 2008.

[23] C. S.Dangare and S. S. Apte, "Improved Study of Heart Disease Prediction System using Data Mining Classification Techniques," Int. J. Comput. Appl., vol. 47, no. 10, pp. 44–48, 2012, doi: 10.5120/7228-0076.

[24] S. U. Amin, K. Agarwal, and R. Beg, "Genetic neural network based data mining in prediction of heart disease using risk factors," in 2013 IEEE Conference on Information and Communication Technologies, ICT 2013, 2013, pp. 1227–1231, doi: 10.1109/CICT.2013.6558288.

[25] E. Yilmaz and Ç. Kilikçier, "Determination of fetal state from cardiotocogram using LS-SVM with particle swarm optimization and binary decision tree," Comput. Math. Methods Med., vol. 2013, 2013, doi: 10.1155/2013/487179.

[26] B. Tarle and S. Jena, "Improved artificial neural network for dimension reduction in medical data classification," Proc. - 2nd Int. Conf. Comput. Commun. Control Autom. ICCUBEA 2016, pp. 0–5, 2017, doi: 10.1109/ICCUBEA.2016.7860033.

[27] K. J. D'Souza and Z. Ansari, "Big Data Science in Building Medical Data Classifier Using Naïve Bayes Model," Proc. - 7th IEEE Int. Conf. Cloud Comput. Emerg. Mark. CCEM 2018, pp. 76–80, 2019, doi: 10.1109/CCEM.2018.00020.

[28] A. Rairikar, V. Kulkarni, V. Sabale, H. Kale, and A. Lamgunde, "Heart disease prediction using data mining techniques," Proc. 2017 Int. Conf. Intell. Comput. Control. I2C2 2017, vol. 2018-Janua, no. October, pp. 1–8, 2018, doi: 10.1109/I2C2.2017.8321771.

[29] S. Habib, "Heart Failure Risk Prediction and Medicine Recommendation using Exploratory Data Analysis," 2019 1st Int. Conf. Adv. Sci. Eng. Robot. Technol., vol. 2019, no. Icasert, pp. 1–6, 2019.

[30] M. T., D. Mukherji, N. Padalia, and A. Naidu, "A Heart Disease Prediction Model using SVM-Decision Trees-Logistic Regression (SDL)," Int. J. Comput. Appl., vol. 68, no. 16, pp. 11–15, Apr. 2013, doi: 10.5120/11662-7250.

[31] N. Kumar and S. Khatri, "Implementing WEKA for medical data classification and early disease prediction," 3rd IEEE Int. Conf., pp. 1–6, 2017, doi: 10.1109/CIACT.2017.7977277.

[32] S. Tuli et al., "HealthFog: An ensemble deep learning based Smart Healthcare System for Automatic Diagnosis of Heart Diseases in integrated IoT and fog computing environments," Futur. Gener. Comput. Syst., vol. 104, pp. 187–200, 2020, doi: 10.1016/j.future.2019.10.043.

[33] A. F. Otoom, E. E. Abdallah, Y. Kilani, A. Kefaye, and M. Ashour, "Effective diagnosis and monitoring of heart disease," Int. J. Softw. Eng. its Appl., vol. 9, no. 1, pp. 143–156, 2015, doi: 10.14257/ijseia.2015.9.1.12.

[34] T. Karayilan and Ö. Kiliç, "Prediction of Heart disease using neural network," 2nd Int. Conf. Comput. Sci. Eng. UBMK 2017, no. 978, pp. 719–723, 2017, doi: 10.1109/UBMK.2017.8093512.

[35] M. Raihan et al., “Smartphone based ischemic heart disease (heart attack) risk prediction using clinical data and data mining approaches, a prototype design,” 19th Int. Conf. Comput. Inf. Technol. ICCIT 2016, pp. 299–303, 2017, doi: 10.1109/ICCITECHN.2016.7860213.

[36] J. Hippisley-Cox, C. Coupland, Y. Vinogradova, J. Robson, M. May, and P. Brindle, “Derivation and validation of QRISK, a new cardiovascular disease risk score for the United Kingdom: Prospective open cohort study,” Br. Med. J., vol. 335, no. 7611, pp. 136–141, 2007, doi: 10.1136/bmj.39261.471806.55.

[37] A. Chauhan, A. Jain, P. Sharma, and V. Deep, “Heart Disease Prediction using Evolutionary Rule Learning,” Int. Conf. "Computational Intell. Commun. Technol. CICT 2018, no. Cict, pp. 1–4, 2018, doi: 10.1109/CIACT.2018.8480271.

[38] M. A. Jabbar, P. Chandra, and B. L. Deekshatulu, “Prediction of risk score for heart disease using associative classification and hybrid feature subset selection,” Int. Conf. Intell. Syst. Des. Appl. ISDA, pp. 628–634, 2012, doi: 10.1109/ISDA.2012.6416610.