A Deep Learning Technique based on Generative Adversarial Network for Heart Disease Prediction

Chour Singh Rajpoot[1,a], **and Dr. Praveen Gupta** [2,b,]

1 Asst Professor, Suresh Gyan Vihar University, Jaipur, India
2 GITAM Institute of Technology, GITAM (Deemed to be University), Visakhapatnam, India
E-mail: [a]csrajput22@gmail.com, [b,*]pgupta@gitam.edu (Corresponding author)

Abstract. Globally, heart illnesses are the leading cause of death. A lot of knowledge and expertise is needed to accurately predict the condition. Although the sickness can be foreseen, it takes a long time to cure. This can help patients prevent heart attacks, as well as medical physicians discover the key causes of heart attacks and evade them before they occur. As a result, the rate of death can be reduced by initiating treatment at an early stage. A fast-growing technique, Data Mining (DM) and artificial intelligence (AI) techniques are used to acquire substantial data and forecast the outcome. A generative adversarial network (GAN) is used to forecast the chance of cardiovascular illness in a patient. Using GAN for detection of heart disease early, this technique is worried with temporal data modelling. We associated the results with existing approaches and got good outcomes. In terms of performance evaluation measures, the suggested method outperforms the existing methodologies. The proposed approach achieves an accuracy of about 98.5%.
.

Keywords: Artificial Intelligence; Blood Sugar; Chest Pain; Data Mining; Early Stage of Identification; Generative Adversarial Network; Heart Disease.

ENGINEERING JOURNAL Volume # Issue #
Received Date Month Year
Accepted Date Month Year
Published Date Month Year
Online at https://engj.org/
DOI:10.4186/ej.20xx.xx.x.xx

DOI:10.4186/ej.year.vol.issue.pp

1. Introduction

DM is a procedure of searching through a huge amount of information to identify patterns, important data and predict the output [1]. The Knowledge Discovery in Database (KDD) is implemented using various kinds of data such as relational database, data warehouses, Data Repository and so on. Various techniques are being used in Data Mining and here we use the classification technique [2-3]. Data Mining is a broadly used concept used in various fields such as healthcare, fraud detection, Market Basket Analysis, Customer Relationship Management, Education, Financial Banking etc.., related with other applications data mining is cost-efficient. Data Mining aids the decision-making practice, which is more efficient due to the usage of a large amount of existing data in a short period of time [4]. Data regarding patients, disease diagnoses etc. are generated by the healthcare business today. These strategies are used to uncover patterns in data. Quality of service is a major concern for the healthcare business [5]. Diagnosing disease correctly & providing appropriate therapies to patients are part of the quality of service. Unacceptable outcomes can arise from a poor diagnosis.

According to a WHO report, heart attacks and strokes are responsible for 17 million deaths worldwide. Everywhere in the world, coronary heart disease (CHD) is a foremost reason for death. Among all the countries in the world, India has highest rates of cardiovascular sickness [6]. From 2.26 million in 1990 to 4.77 million in 2020, the number of heart disease-related deaths in India is expected to increase annually. Even very young children can be impacted by HD. Smoking, a poor diet and high BP are entirely chief risk factors for heart disease [7-8]. Coronary is a heart disease that is caused due to inappropriate beating of the heart, whether asymmetrical, too fast or too slow. It happens when voltaic impulses in the heart don't work properly. This disease does not have any noticeable symptoms and hence it difficult to predict [9]. Here, we have used data mining algorithms to determine if the patient has HD or not.

Based on the doctor's knowledge and understanding, a diagnosis is often reached. Unwanted consequences & exorbitant medical costs are the result. As a result, an automatic medical identification scheme is built that takes benefit of acquired data base and decision support scheme. With fewer medical tests and more effective treatments, this method can assist to diagnose the disease.

Patient's current test findings and doctor's knowledge are often used to make a diagnosis in numerous situations. Diagnosis is a hard undertaking that demands a great deal of experience & high skill. Various methods in DM have utilized to determine the severity of cardiac disease in humans. [12] K-Nearest Neighbor Algorithm (KNN), [13] Decision Trees (DT), [14] Genetic algorithm (GA), and [15] Naive Bayes are used to classify the severity of the disease. As a result, cardiac diseases must be managed with care. Therefore, in this research study, a deep learning system is implemented for predicting the HD.

The residual paper is constructed by the study of existing techniques in Section 2, the explanation of proposed methodology in Section 3. The validated results of GAN with existing techniques in terms of all parameters are provided in Section 4. Lastly, the assumption of the work is labelled in Section 5.

.

2. Literature Review

Various researchers have previously conducted a large amount of research in this topic. With the use of powerful ML procedures and big data skills, this part provides an overview of the research that has previously been done in the prediction of cardiac disease. Different research groups have utilized a variety of data analysis approaches for model prediction to identify cardiovascular illnesses. KNN method, Nave Bayes classifier, generalized estimating equations, decision trees and deep neural networks were utilized to carry out the experimental work. When the above-mentioned strategies are applied in combination, the accuracy computed is high in the majority of the study studies [16]. In the forecast of cardiovascular disease, a large amount of research has previously been done.

Please take a moment to explore some of the work that has been done in this area by various research groups. H. Van Pham has created an expert system employing neural networks in [17]. According to a doctor's request, the information base is signified using fuzzy rules and updated in order to make better decisions based on heart disease risk levels. Intelligent machine-learning algorithms have been effectively implemented in decision support systems by a number of researchers. An ontology-based data mining method has been successfully deployed by Qrenawi, Mohammed, and colleagues to a particular dataset of diabetics with cardiac disease [18]. To determine whether there is a relationship among the history of kind two diabetic persons and the laboratory tests recommended by a medical physician, this study was carried out in order to find out. Ontology-based, rule-induction methods are used in a later segment of the research. Ontology-based techniques are used in this research to reduce the amount of properties in the pre-processing phase and aid in the majority of the data mining stage. More than 90 percent accuracy was also attained.

For HD categorization in an IOT-based deep learning scheme medical environment was developed by Nguyen et al. [19]. We breakdown the cardiac signals into wavelet coefficients using the WPD algorithm and then extract features using the WPCA algorithm (wavelet-based kernel PCA). Different nodes were utilized to classify heart disease on back propagation. Isra'a Ahmed et al. [20] built an intelligent medical decision support system based on DM techniques, as an example. To access and study risk factors associated to heart disease in the framework of statistical analysis for identifying HD, this incorporation's major aim was conceived. DT, RF, Nave Bayes, Discriminant Analysis, and SVM were all easily compared. Models were applied on two separate datasets to show that their technique is practicable, and the results were encouraging. Researchers find that all categorization algorithms are fairly analytical and can give a near-correct result, according to their research article. The decision tree, however, surpasses other classifiers, according to the authors of the study. Algorithm of random forest is ranked second in this category. Heart attacks and heart disease can now be predicted before they occur, thanks to advances in modern research.

As Kanchan et al., [21] demonstrate, smartphone technology can help estimate the hazard of heart attack. In a cardiac hospital, they established an android request that is coupled with a clinical database consisting of data from more than 500 patients. Variables like diabetes, hypertension, smoking, dyslipidaemia, stress, obesity as well as present clinic symptoms were examined in the connection of the presence of ischemic heart disease (IHD) with data provided. As a result of this analysis, a notional score was obtained. In order to assign a risk score to IHD, three risk classes were created: low, medium, and high risk. The android app was tested using the data of 89 patients with acute coronary syndrome (ACS). A significant link was established between having a cardiac event and being in the high or low category, with a p value of 0.0001. IHD was discovered in 89 percent of patients in the high-risk category, but not in the low-risk category. Only 12.5 percent of patients in the low-risk category had IHD. High and medium scores were also significantly different (p 0.0001). Eighty-six percent of patients with ACS got high scores, according to a study.

Ammar Aldallal et al. [22] built the application software. With the use of this software, medical practitioners might forecast the onset of non-communicable diseases (NCDs). Bahrain Defense Force Hospital patient records were used to test the software program. In the hospital, physicians were asked to test and execute this application. A medical practitioner is said to be able to make better judgements on patient health risks with the help of this application. A number of researchers have developed sensor-based and hardware-oriented ways to study the factors elaborate in the expansion of HD in addition to software-based approaches.

Johanna O'Donnell et al., [23] suggest a similar sensor-based technique. Heart failure can be diagnosed using the data acquired from multi-sensor patch data, according to the authors of the paper. These data were acquired from people suffering from heart failure using multi-sensor patch technology. A first analysis of sleep patterns and actions of heart failure patients by strictness of HD can be made with these data, according to the study authors. A 13 heart failure patients were requested to wear chest-worn multi-sensors for roughly seven consecutive days. 11 of the 13 patients had high-quality multisensor data, which was used in the analysis. Heart failure patients of varying severity showed a possible variation in heart rate, sleep angle, and wake time activity. Larger studies are needed to evaluate the function of action and sleep as indicators of heart failure.

In a recent study, Mehmood et al. [24] predicted the likelihood of a heart attack based on data extracted from the UCI repository. When it comes to mining information for prediction, authors stressed the need of using attribute extraction approaches. By applying attribute extraction techniques, the researchers say they can derive patterns that can be used to detect HD earlier. In this study effort, several strategies in Artificial Neural Networks (ANNs) are explained. However, principal component analysis raised the Accuracy to 97.7%.

Alizadeh-dizaj et al. [25] also use the data mining technique. A DT based on the risk factors that distress it was used to predict the probability of stroke in suspected stroke patients. Eleven hundred and eighty-four records comprise the primary database. We have employed Classification Trees, Nave Bayes and Neural Network methods in the modeling phase. According to this study, the most efficient indicators in predicting potential heart disease were physical inactivity, and high blood pressure. These criteria can be used as a ideal for forecasting stroke risk in patients. 95.52% accuracy was claimed for the model. A novel sample with specified characteristics can be used to determine the stroke risk by applying the guidelines. T. John Peter et al. [26] suggested a method to assess the presentation of diverse classification methods such as K-NN, DT, NB on a HD dataset. Authors observed that the Nave Bayes classifier performed better than other techniques. Data dimension was reduced using attribute selection methods in order to improve performance.

3. Proposed Methodology

3.1. Dataset Default

A well-established dataset and a existing machine learning method called Generative adversarial networks are used in this study to predict cardiac disease. This section contains a full discussion of the technique that has been offered.

Table 1. Dataset Description.

Attribute	Possible values	Description
Age	Valid numbers	when disclosed to the hospital
Sex	0: female/1: male	
Chol	125–565	cholesterol Level in mg/dl
Fbs	0: false/1: true;	Blood sugar level is greater 120 mg/dl?
restecg	0: normal 2: represent the left ventricular hypertrophy 1: with ST-T	Results of electrocardiography while resting
Thal	6 as fixed defect; 7 as reversible Defect; 3 as normal;	
Cp	1:typical 2:non-anginal 3:atypical 4: asymptomatic	Kind of chest pain
trestbps	90–200	BP while resting in mm Hg
Num	1 agrees to greater than 50% tapering diameter	Finding for the actuality of heart disease, status
old peak	0–7	ST induction of depression by exercise relative to rest
Slope	1 depicts upsloping 2 depicts flat 3 shows a down-sloping	The slope of the peak exercise AST segment
Ca	0–3	Number of major vessels (0–3) colored by fluoroscopy
thalach	71–202	Max heart rate
exang	1 true; 0 false	Angina induced by exercise

3.2. Data Preparation

Before conducting an experiment, it is required to prepare the data in accordance with its nature, as was indicated in the preceding section. As part of data preparation, a third-party software or SKLearn, a Python library, may be used to fetch and save the dataset into the file system.

In the dataset, there was a notable quantity of noise and several irregularities. A variety of Python-based programs were written to clean the data and make it consistent. In order to convert the raw data into a purified feature matrix, many programs are built to do different tasks, such as data purification and replacement of abnormalities.

3.3. Mainstream Voting with LASSO Shrinkage

A technique incorporating the regression approach, the least absolute shrinkage or LASSO, is used to develop the prediction accuracy and interpretability of a model by regularizing and selecting variables.

It aids in the removal of limits and the selection of variables by shrinking the data values to a central point Pc. Multi-collinearity models are well-suited to this type of regression approach. As a result of variable elimination, particular coefficients are excluded from the model after flattering zero. A unique purpose is achieved by the LASSO solutions because of their quadratic nature:

$$\sum_{j=1}^{n}(y_j - \sum_j x_{jk\gamma k})^2 + \lambda \sum_{k-j}^{q} |\gamma k| \quad (1)$$

After the shrinking process, a subset of values (signified by) becomes zero, which is clearly understood by the regression model. A parameter is included in the equation above, which indicates the amount of shrinkage. When =0, no parameters are deleted from the model. Regarding the growth in, additional coefficients are fixed to zero and removed from the model as a result. As decreases, the variance increases and as rises, the bias increases. A variable or factor's relevance in terms of its support to the difference under consideration is expressed as its value. Considered irrelevant, this variable will be disregarded. An unbalanced dataset can lead to deceptive findings from LASSO, which can lead to an inaccurate selection of relevant variables when LASSO is conducted on the entire dataset, according to the authors.

It is possible to reduce the effect of imbalance by randomly subsampling the dataset and re-iterating the LASSO algorithm numerous times. In most of the iterations, the nonzero variables are selected by majority vote based on values. Using a scenario, let's try to comprehend what you're talking about. Every instance of LASSO contains equal numbers of CHDs and non-CHDs. As an example, if there are 45 variables, then we can get the 45th instance. A variable v is included in further analysis by counting the number of nonzero illustrations of that variable and a threshold value physically determined. Alternatively,

$$x(\gamma)\begin{cases}0 & if\ \gamma = 0\\ 1 & otherwise\end{cases}$$

$$[x(\gamma_1, d), x(\gamma_2, d), x(\gamma_3, d) \ldots\ldots x(\gamma_N, d)]1 \geq \frac{m}{a} \Rightarrow d\ is\ selected \quad (2)$$

3.4. Generative adversarial networks

An example of a generative model is the GAN [28], which was meant to generate samples directly from a chosen data distribution without explicitly modeling the probability density function. Generated by the generator G and discriminated by the discriminator D, this neural network comprises of two neural networks. A Gaussian or uniform distribution p(z) is frequently selected for simplicity as the prior distribution for G, z. If x r is derived from the genuine data distribution p r, then the output of G should be visually comparable to x r (x). As x g=G(z; _g), we indicate the nonlinear mapping function learned by G. An actual or produced sample is fed into D. It returns a single result that indicates if the input sample is real or bogus. y 1 = D(x; _d) denotes the mapping learned by D and parametrized by d. After successful training, the resulting samples form a distribution p g (x), which is intended to approximate p r (x). GAN setup is depicted in Figure 1. A 2D CT slice of lung nodule is generated by G in this case.

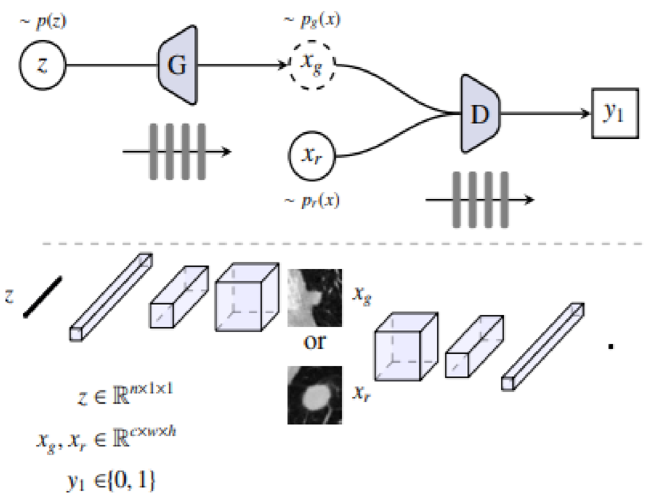

Figure 2: Schematic view of the GAN for synthesis.

In this case, D's purpose is to distinguish between the two groups of photos, while G's goal is to confuse D as much as possible. An intuitive interpretation of the situation is that G is acting in the capacity of a counterfeiter, while D is acting in the capacity of an officer trying to discover the forgeries. If the created data is accurate or not, we can see G as getting a return signal from D. So G modifies its restrictions to yield an output image that can mislead D, based on the gradient information that has been transferred from D to G in reverse. Mathematically, D and G's training objectives can be represented as follows:

$$\mathcal{L}_D^{GAN} = \max_D \mathbb{E}_{x_r \sim p_r(x)}[\log D\,(x_r)] + \mathbb{E}_{x_g \sim p_g(x)}[\log\,(1 - D(x_g))]$$

$$\mathcal{L}_D^{GAN} = \min_D \mathbb{E}_{x_r \sim p_r(x)}[\log\,(1 - D(x_g))] \quad (3)$$

In this case, D's purpose is to distinguish between the two groups of photos, while G's goal is to confuse D as much as possible. An intuitive interpretation of the situation is that G is acting in the capacity of a counterfeiter, while D is acting in the capacity of an officer trying to discover the forgeries [28]. If the created data is accurate or not, we can see $p_r(x)$ and $p_g(x)$ as receiving a return signal from D. So G modifies its factors to yield an output image that can mislead D, based on the gradient information that has been transferred from D to G in reverse. Mathematically, D and G's training objectives can be represented as follows:

4. Results and Discussion

We set up an experimental environment to carry out this research and assess the outcomes of the study. Intel i7 4th generation processor, clocking at 2.3 GHz with four megabytes of memory. The workstation is equipped with 16 gigabytes of DDR3 RAM and a 1 terabyte (TB) SATA hard drive rotating at 7 K RPM. Microsoft Windows 10 Pro is used as the base OS for this project. Version 2018a of MATLAB.

4.1. Performance Metrics

Using some performance metrics, one may appraise the classifier performance. In machine learning, a classifier's performance can be evaluated based on a variety of parameters. Here are some of these characteristics explained.

Precision

When evaluating a classifier's performance, precision is used to determine how accurate it is. There are fewer false positives if accuracy is high.

$$Precision = \frac{TP}{TP+FP} \quad (4)$$

Recall

As a measure of completeness, it is used to determine if a classifier is accurate. Les false negatives recall is higher; conversely, when the recall is lower, the false negatives are greater. Precision often suffers as a result of an improvement in recall.

$$Recall \frac{TP}{TP+FN} \quad (5)$$

F-Score

F-score is the grouping of accuracy and recall:

$$F1\ Score = \frac{Recall \times Precision}{Recall + Precision} \quad (6)$$

If recall and precision are multiplied by each other, the F1 score may be calculated.

4.2. Experimental analysis

Using the UCI standard repository [27], we built up an experimental setup with the subsequent parameters in order to observe and analyse the performance of the GAN algorithm. Feature matrices are created once the dataset has been presented as a feature matrix. The following phase is to divide the dataset into classes. Firstly, we separated the dataset into two classes, HD and

No-HD, based on the presence or absence of heart disease. Data are classified using GAN, and a classifier model is placed on the file scheme as a result of the classification process. Table 2 provides the presentation of projected GAN model for binary class.

Table 2 Presentation of the GAN for binary class

Metric	Value
Precision	0.9705
F1-score	0.9673
Accuracy	0.9712
Recall	0.9638

The above table 2 shows that the GAN network achieved 97.05% of precision, 96.38% of recall, 96.73% of F-score and 97.12% of accuracy for two classes, i.e. No_HD and HD. Using the UCI standard repository [27], we built up an experimental setup with the subsequent parameters in order to detect and analyze the performance of the GAN algorithm. Feature matrices are created once the dataset has been presented as a feature matrix. The next stage is to divide the dataset into classes. Firstly, we divided the dataset into two classes, HD and No-HD, based on the presence or absence of heart disease. Data are classified using GAN, and a classifier model is placed on the file system as a result of the classification process.

Table 3: Performance overview of the GAN ideal with four classes

Metric	Value
Precision	0.8769
Accuracy	0.8767
Recall	0.8374
F1-score	0.8514

Overall of 114 records were classed as type 1 records, according to the classification. On the other hand, 22 records were categorized into the type 3 class. Type 4: A entire of 47 records are included in this category. 8 records were falsely labeled for type 1 disease, whereas 26 records were falsely identified for type 2 sickness. There are six erroneous categories for type 3 disease, but there are none for type 4. The GAN network achieved 87.69% of precision, 83.74% of recall, 85.14% of F1-score and 87.67% of accuracy.

4.3. Binary class Comparison with other classification Techniques

In the performance of GAN model is compared with other deep learning (DL) and machine learning (ML) techniques in terms of accuracy for binary class. The validated results are assumed in Table 4.

Table 4: Performance Analysis of Proposed GAN model with other classifiers for Binary Class.

Classification technique	Accuracy
Random forest (RF)	0.6516
Decision tree (DT)	0.8559
(SVM)	0.8791
(ANN)	0.8461
(RNN)	0.8605
(LSTM)	0.9195
Convolutional Neural Network (CNN)	0.9578
GAN	**0.9712**

From the above table, it is clearly stated that the proposed GAN model achieved better classification accuracy than other techniques. Among them Random Forest achieved less accuracy (i.e. 65.16%), whereas DT, RNN, ANN and SVM achieved nearly 84% to 87% of accuracy. The other techniques includes LSTM and CNN achieved nearly 91% to 95% of classification accuracy, however the proposed GAN model achieved 97.12% of accuracy that shows higher performance than other techniques. We utilized LASSO to appropriately classify the predictor variables for subsequent investigation because of the importance of risk factors and their correlation with correlated variables. Table 5 compares the accuracy of GAN with existing approaches for multi-class classification.

Table 5: Performance Analysis of GAN model on Multi class Comparison with other classification Techniques

Classification technique	Accuracy
Random forest	0.6497
Decision tree	0.7043
SVM	0.7061
ANN	0.7543
RNN	0.7634
LSTM	0.8004
CNN	0.8543
GAN	**0.8767**

While comparing with two-classes, the all techniques achieved less performance on multi-class data. For instance, the RF achieved 64.97% of accuracy, where DT and SVM achieved only 70% of accuracy. The RNN and ANN achieved nearly 76% of accuracy, where LSTM and CNN achieved nearly 83% of accuracy. However, the proposed GAN model achieved 87.67% of accuracy, which is less when compared with two-class data. The next section will describes the overall accuracy of GAN model with existing techniques.

4.4 Comparative Analysis of Proposed GAN with Other Methods

As suggested by the proposed methodology, the trained classifier is then utilized to predict the entire testing dataset, as shown in the following figure. A performance comparison is shown in Fig. 3 between the suggested framework and other research in terms of accuracy results. Performance (i.e. accuracy): As can be seen from the above image, the proposed strategy outperforms the previous techniques.

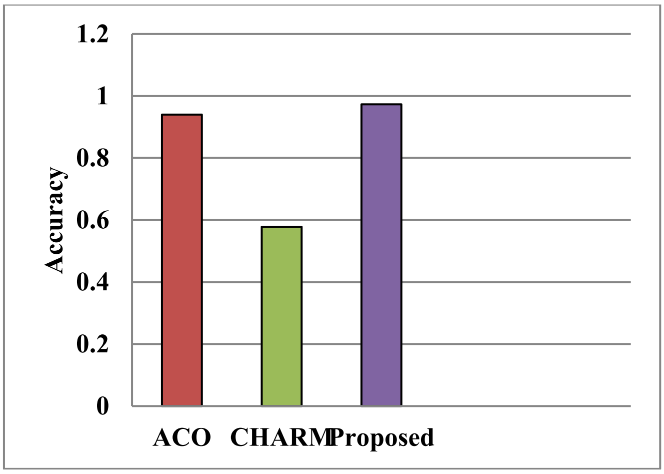

Fig 3 Proposed model with existing Methods Performance comparison.

Among the existing techniques, the risk factor based approach [21] achieved poor performance (i.e. 86.7%), where the Rmonto ontology-driven data mining approach [18] achieved 90% of accuracy. The other techniques includes ANN-based attribute extraction technique [24] and Decision tree-based approach [25] achieved nearly 94% to 95% of overall accuracy. But, the proposed GAN approach achieved 97.12% of accuracy. This shows that our proposed GAN model achieved better performance than all existing ML and DL techniques.

5. Conclusion

If heart disease can be predicted at an earlier stage, it may prevent deaths from heart attacks. Before the onset of cardiovascular disease, a good classification procedure can help the physician predict its presence. It uses the existing dataset available at UCI's repository and GAN to predict a possible heart disease. Certain cardiac test factors as well as common human habits are included in this data set. In total, 97.12 percent of the model's predictions are correct. Death rates can be drastically reduced if early diagnosis and treatment are provided. In the future, we aim to additional improve this research work by predicting the occurrence of heart-related diseases by finding the other risk factors of early heart diseases. Though it is matter of medical doctors to identify the disease and provide solution but the above research and other research can assist them for the better decision.

References

[1] Jain, N. and Srivastava, V., 2013. Data mining techniques: a survey paper. IJRET: International Journal of Research in Engineering and Technology, 2(11), pp.2319-1163.

[2] Larose, D.T. and Larose, C.D., 2014. Discovering knowledge in data: an introduction to data mining (Vol. 4). John Wiley & Sons.

[3] Kesavaraj, G. and Sukumaran, S., 2013, July. A study on classification techniques in data mining. In 2013 fourth international conference on computing, communications and networking technologies (ICCCNT) (pp. 1-7). IEEE.

[4] Tomar, D. and Agarwal, S., 2013. A survey on Data Mining approaches for Healthcare. International Journal of Bio-Science and Bio-Technology, 5(5), pp.241-266.

[5] Jothi, N. and Husain, W., 2015. Data mining in healthcare–a review. Procedia computer science, 72, pp.306-313.

[6] Saxena, K. and Sharma, R., 2016. Efficient heart disease prediction system. Procedia Computer Science, 85, pp.962-969.

[7] Bhatla, N. and Jyoti, K., 2012. An analysis of heart disease prediction using different data mining techniques. International Journal of Engineering, 1(8), pp.1-4.

[8] Dangare, C.S. and Apte, S.S., 2012. Improved study of heart disease prediction system using data mining classification techniques. International Journal of Computer Applications, 47(10), pp.44-48.

[9] Chen, A.H., Huang, S.Y., Hong, P.S., Cheng, C.H. and Lin, E.J., 2011, September. HDPS: Heart disease prediction system. In 2011 computing in cardiology (pp. 557-560). IEEE.

[10] Soni, J., Ansari, U., Sharma, D. and Soni, S., 2011. Predictive data mining for medical diagnosis: An overview of heart disease prediction. International Journal of Computer Applications, 17(8), pp.43-48.

[11] Sa, S., 2013. Intelligent heart disease prediction system using data mining techniques. International Journal of healthcare & biomedical Research, 1, pp.94-101.

[12] Mohan, S., Thirumalai, C. and Srivastava, G., 2019. Effective heart disease prediction using hybrid machine learning techniques. IEEE access, 7, pp.81542-81554.

[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] Parthiban, L. and Subramanian, R., 2008. Intelligent heart disease prediction system using CANFIS and genetic algorithm. International Journal of Biological, Biomedical and Medical Sciences, 3(3).

[15] Repaka, A.N., Ravikanti, S.D. and Franklin, R.G., 2019, April. Design and implementing heart disease prediction using naives Bayesian. In 2019 3rd International conference on trends in electronics and informatics (ICOEI) (pp. 292-297). IEEE.

[16] Rathnayakc, B.S.S.; Ganegoda, G.U.: Heart diseases prediction with data mining and neural network techniques. In: 2018 3rd International Conference for Convergence in Technology (I2CT), pp. 1–6 (2018).

[17]. Van Pham, H.: A proposal of expert system using deep learning neural networks and fuzzy rules for diagnosing heart disease. In: Frontiers in Intelligent Computing: Theory and Applications, ed: Springer, pp. 189–198 (2020).

DOI:10.4186/ej.year.vol.issue.pp

[18]. Qrenawi, M.I.; Al Sarraj, W.: Identification of cardiovascular diseases risk factors among diabetes patients using ontological data mining techniques. In: 2018 International Conference on Promising Electronic Technologies (ICPET), pp. 129–134 (2018).
[19]. Nguyen, T.-H.; Nguyen, T.-N.; Nguyen, T.-T.: A deep learning framework for heart disease classification in an IoTs-based system. In: A Handbook of Internet of Things in Biomedical and Cyber Physical System, ed: Springer, pp. 217–244 (2020).
[20]. Isra'a Ahmed Zriqat, A.M.; Altamimi, M.A.: A comparative study for predicting heart diseases using data mining classification methods.
[21]. More, K.; Raihan, M.; More, A.; Padule, S.; Mondal, S.: A12176 Smart phone based "heart attack" risk prediction; innovation of clinical and social approach for preventive cardiac health. J. Hypertens. 36, e321, 2018.
[22]. Aldallal, A.; Al-Moosa, A.A.A.: Using data mining techniques to predict diabetes and heart diseases
[23]. O'Donnell, J.; Velardo, C.; Shah, S.A.; Khorshidi, G.S.; Salvi, D.; Rahimi, K.; et al.: Physical activity and sleep analysis of heart failure patients using multi-sensor patches. In: 2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), pp. 6092–6095 (2018).
[24]. Awan, S.M.; Riaz, M.U.; Khan, A.G.: Prediction of heart disease using artificial neural network. VFAST Trans. Softw. Eng. 13, 102–112, 2018.
[25]. Alizadeh-dizaj, G.: Risk Prediction and Stratification of Patients with Stroke Using Data Mining Techniques. Tabriz University of Medical Sciences, School of Management and Medical Informatics (2018).
[26]. Somasundaram, T.J.P.K.: An empirical study on prediction of heart disease using classification data mining techniques, presented at the IEEE-International Conference On Advances In Engineering, Science And Management (ICAESM -2012), (2015).
[27] https://archive.ics.uci.edu/ml/machine-learning-databases/heartdisease/
[28] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. WardeFarley, S. Ozair, A. Courville, and Y. Bengio. Generative Adversarial Nets. In Advances in Neural Information Processing Systems 27, pages 2672–2680. Curran Associates, Inc., 2014

Second B. Chour Singh Rajpoot was born in karauli , Rajasthan, India 1988. he received the B.Tech degrees in Computer Engineering from Rajasthan Technical University Kota in year 2010 and M.Tech. degrees in computer engineering from the Rajasthan Technical University Kota in year 2019. He is research scholar at Manipal Unversity, Jaipur since july 2020 and working the field of machine learning applications.

He has work experience of 10 years in teaching at higher educaton at engineering educational institution level in india. He has teaching experience to undergraduate students. He has attended various faculty development programs to improve the skills. And currently working on the writing of the book chapter in machine learning applications with the help of other senior coolegues. He is keenly interested and worked enthustically to organize varous technical events in the institute. He is currently working on various projects in the field of machine learning , data science and applications of the machine learning. His research interest are machine learning , deep learning, Artificial Intelligency and related fields. He has wrtten 7 research papers for the national and international journals.

First A. Dr. Praveen Gupta was born in jaipur in 1969. He has completed bachelor of engineerng from Nagpur Unversity in year 1991. he has completed M. Tech in year 2004 from punjabi university in information technology and completed doctrate degree in Computer Engineering in year 2014. He has completed many certification programs from various IIT and NIIT.

He has worked with mid size software companies in india on java ,j2ee, oracle and web technology for more than 8 years. He has 16+ years of experience in computer software teaching and research in technical educational institues in india. He is currently working as professor in department of computer engineering in jaipur. his research interest are machne learning , data science , cloud and programming. He has written many research papers in journals and conferences, He has attended various faculty development programs and written 3 books in programming subjects of c , C++ and java . He has written a one chapter Title " Location and Availability Protections in Smart Mobility "Chapter No 5 Titled "Cybersecurity: Ambient Technologies, IoT, and Industry 4.0 Implications". CRC Press of Taylor and Francis) (ISBN: 9781003145042). Mr Gupta is also working on a patent with his team title of the invention is : "Modeling, Optimizing and diagnosis of Chiller Systems using Machine Learning". Patent is under process. He has worked on the agile way of software development implimentation in the students among the new learners or techncal students. Which was a successful effort and teaching and learning and published one paper based on that.