Nanotechnology Perceptions
ISSN 1660-6795
www.nano-ntp.com

Improving Heart Disease Prediction with an Advanced Healthcare Monitoring System Using Ensemble Chicken Swarm Optimized-Faster RCNN with Feature Transforming Methodology

Priya Dahiya[1], Richa Goswami[2], Sahil Suri[3], Dr. Sachin Gupta[4], Dr. John Yesudas Valluri[5], Shikhar Gupta[6]

[1]*Assistant Professor, Department of CSE (Artificial intelligence), Noida Institute of Engineering and Technology, Greater Noida, Uttar Pradesh, India, Email Id- priyadahiya@niet.co.in*
[2]*Assistant Professor, Department of Management Studies, Vivekananda Global University, Jaipur, India, Email Id- richa.goswami@vgu.ac.in*
[3]*Centre of Research Impact and Outcome, Chitkara University, Rajpura- 140417, Punjab, India, Email Id- sahil.suri.orp@chitkara.edu.in*
[4]*Chancellor, Department of Management, Sanskriti University, Mathura, Uttar Pradesh, India, Email Id- chancellor@sanskriti.edu.in*
[5]*Assistant Professor, Department of Finance, JAIN (Deemed-to-be Univesity), Bangalore, Karnataka, India, Email Id- dr.vy_john@cms.ac.in*
[6]*Chitkara Centre for Research and Development, Chitkara University, Himachal Pradesh-174103, India, Email Id- shikhar.gupta.orp@chitkara.edu.in*

Introduction: Heart diseases are a major global health issue including several conditions that impact the heart and blood vessels. Prediction approaches, which utilize modern technology and advanced data analysis, have a crucial significance.
Method: In this study, we suggested Ensemble Chicken Swarm Optimized-Faster Recurrent Convolutional Neural Network (ECO-FRCNN) to enhance the prediction rate in the context of heart disease prediction with an advanced healthcare monitoring system.We utilized dataset from kaggle consists of 76 attributes. Data cleaning process is performed to clean the gathered data. Chicken Swarm Optimization (CSO) algorithm is utilized to improve the effectiveness of the suggested Faster RCNN to enhance the predictions in heart disease.
Results: Accuracy and loss values for our proposed method is determined, the comparison analysis is done on various parameters including precision (89, 56 %), accuracy (91,56 %) and F1 score (94,86 %) with other traditional approaches.

Conclusion: Our research demonstrates the efficacy of ECO-FRCNN in enhancing the accuracy of heart disease forecasts. The hybrid of CSO with FRCNN leads to exceptional outcomes. This technique shows potential for enhancing prediction models in the field of heart disease, providing a significant addition to healthcare technology.

Keywords: Ensemble Chicken Swarm Optimized-Faster RCNN (ECO-FRCNN), Improving Heart Disease Prediction, Advanced Healthcare, Monitoring System.

1. Introduction

Heart disease refers to a variance of disorders that impact the cardiovascular system. (1) It continues to be a major source of sickness and mortality worldwide, presenting a substantial public health problem. (2) Heart disease, which includes illnesses like coronary artery disease, heart failure and ventricular fibrillation progresses without noticeable symptoms for a long time. Therefore, early identification is important for effective treatments. The initial occurrence of this condition is influenced by lifestyle factors such as an unhealthy diet, sedentary behaviour, smoking and a high level of alcohol. Due to the varied sources and symptoms, it is crucial to have a comprehensive strategy for preventing and managing the condition. (3) Development in the field of medical research and technology has facilitated the establishment of novel methods to forecast cardiac disease.

Predictive modelling, using Machine Learning (ML) algorithms, examines extensive datasets encompassing medical records, genetic data and lifestyle variables. (4) These algorithms possess the ability to detect patterns and risk variables, enabling healthcare practitioners to evaluate an individual's probability of developing heart disease. (5) The prediction of heart disease entails an integration of several data points, encompassing age, gender and background information. ML models are highly effective at managing this intricacy, providing a more subtle comprehension of an individual's risk profile. These models face constant development as additional data becomes accessible, ensuring their precision and dependability in predicting heart disease among various populations. (6) An important advantage of predictive modelling is its capacity to customize interventions based on individual risk indicators.

The research (7) improved healthcare by utilizing the combination of the “Internet of Things (IoT)” and Cloud to deliver effortless and sustained communication. Applying statistical analysis in medicine shifted the responsive strategy to a proactive one, utilized sophisticated artificial intelligence and ML. Researchers (8) introduced a framework named Health-Fog to include “ensemble Deep Learning (DL) into Edge computing devices for a practical application of Heart Disease analysis”. The suggested model's effectiveness was evaluated using the Fog-Bus cloud architecture.

The study (9) constructed a ML model to identify and diagnose cardiovascular disorders. The “K-Nearest Neighbours (KNN)” method performed as the superior choice when compared to other algorithms like “Random Forest (RF), Decision Tree (DT), Support Vector Machine (SVM) and Naive Bayes”. Moreover, a working model was created to verify the accuracy of the findings. The authors (10) suggested improving the accuracy of prediction in diagnosing cardiac illness. They suggested utilizing a framework called

“Internet of Medical Things (IoMT) and an adaptive neuro-fuzzy inference system (ANFIS)”. The outcome of the simulations revealed that the MSSO-ANFIS forecasting algorithm outperformed the other techniques in terms of accuracy.

The research (11) presented an IoT architecture that utilized a “Modified Deep Convolutional Neural Network (MDCNN)” for increasing the accurate range in disease evaluation. Analysed the algorithm's effectiveness by contrasting the suggested MDCNN to conventional DL “neural networks and logistic regression”. The outcomes indicated that the heart disease forecasting technology based on MDCNN outperformed other approaches. Authors (12) suggested an advanced “healthcare framework” for forecasting cardiovascular heart disease using a “Swarm-Artificial Neural Network (Swarm-ANN)” method. The suggested “Swarm-ANN” method demonstrated superior accuracy for identifying cardiovascular illness among individuals using a benchmark dataset.

The study (13) introduced a “heart disease prediction model (HDPM) for a Clinical Decision Support System (CDSS)”. The HDPM utilized “Density-Based Spatial Clustering of Applications with Noise (DBSCAN)” to detect and eliminate outliers for forecasting heart disease. The findings demonstrated that the suggested model performed better than conventional approaches and attained outcomes with precision rate of 98 %.

In this study, we suggested ECO-FRCNN to enhance the prediction rate in the context of heart disease prediction with an advanced healthcare monitoring system. The remainder of this paper is categorized into the subsequent parts: part 3, proposed methodology; part 4, findings and part 5, conclusion.

2. Methodology

Data acquisition

In this research, we obtained a dataset from the Kaggle open data source, primarily from the “Heart-Disease-Dataset database”. It contains information such as “age, chest pain kind, resting blood pressure, glucose levels, and heart rate attained”. It consists of a total of 76 variables. The characteristic defined "goal" represents the existence of a heart condition in an individual and is expressed by a number that varies from 0 to 4. (14)

Data cleaning

In this phase, the dataset performed thorough data cleaning to assure its integrity. The primary objective was to improve the quality of data to ensure precise analysis. By doing meticulous data cleaning, the dataset was improved, establishing a dependable basis for subsequent research and enhancing the reliability of the research's findings.

Chicken Swarm Optimization (CSO) Algorithm

CSO Algorithm strategy is derived on the collective behaviours of chickens and it is designed to optimize model parameters to achieve the highest possible efficiency.

In the entire chicken swarm optimization, the overall amount of individuals in all flocks is denoted as N. The location of everyone in the chicken swarm is expressed as $v_{j,i}(h)$, that indicates the position acquired by the j^{th} individuals in the h^{th} iteration in the i^{th} dimensions.Consequently, the CSO assigns distinct locations to the various kinds of chickens. Specifically, the location update of each flock is modified based on the kind of chicken. The rooster has superior fitness in each subgroup and possesses the ability to detect and locate sustenance over a broad expanse of environments. The location relating to the cock has been modified in the following manner in Equation (1):

$$v_{j,i}(s+1) = v_{j,i}(h) * \left(1 + \mathrm{Rand}\ m(0, \sigma^2)\right),$$

$$\sigma^2 = \begin{cases} 1, & e_j \le e_l \\ \exp\left(\frac{e_l - e_j}{|e_j| + \varepsilon}\right), & e_j > e_l \end{cases} \quad (1)$$

The expression Rand $m(0, \sigma^2)$ generates a random number with a "Gaussian distribution" and a median of 0, with a "standard deviation" of σ. The variable ε is a small integer used to avoid division by zero. The variable $e_j$represents the fitness value of individualj, whereas $e_l$represents the fitness value of personl. An individual, denoted as l, is picked at random from a group of roosters, with the condition that l is not equal to j. The position of the hen is modified according to the following procedure in Equation (2):

$$v_{j,i}(h+1) = v_{j,i}(h) + d_1 * \mathrm{rand} * \left(v_{q1,i}(h) - v_{j,i}(h)\right) + d_2 * \mathrm{rand} * \left(v_{q2,i}(h) - v_{j,i}(h)\right)$$

$$d_1 = \exp\left(\frac{(e_j - e_{q1})}{\mathrm{abs}(e_1) + \varepsilon}\right)$$

$$d_1 = \exp\left(e_{q2} - e_j\right) \quad (2)$$

Let rand be a random integer ranging from 0 to 1. Let q1 represent the cock of the j^{th} hen and q2 represent any individual among all the cocks and hens in the group. It is important to note that q1 cannot be equal to q2. The location associated with the chick has been modified in the manner that follows in Equation (3):

$$v_{j,i}(h+1) = v_{j,i}(h) + E * \left(v_{n,i}(h) - v_{j,i}(h)\right) \quad (3)$$

Let n denote the hen associated with the j^{th} chick and E indicate the follow-up factor, which indicates that the chick follows the hen to locate food.

FR-CNN architecture

FR-CNN detects and analyses significant characteristics in medical data. The "region proposal network (RPN)" is employed to construct the "Region of Interest (ROI)" instead of using the slower sliding window and limited search technique. The framework of FR-CNN is illustrated in Figure 1.

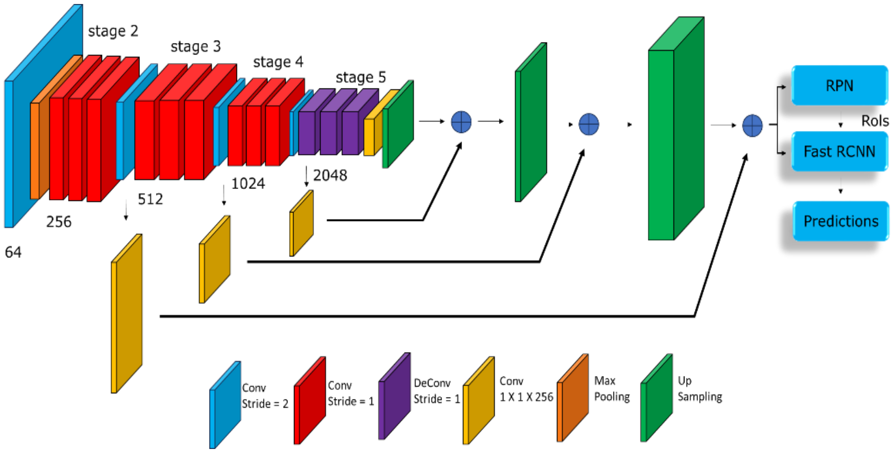

Figure 1.FR-CNN architecture

(Source: Author)

The RPN employs a 3×3 convolution kernel in a sliding window to execute convolutional calculations on the characteristic map. It identifies the central points of the sliding window and maps them to the corresponding positions in the original input data. Each point relates to the source data to build K anchor points with various sizes and diverse aspect ratios. The assignment of positive and negative samples is determined based on the "Intersection over Union (IOU)" value comparing the produced anchor and the ground truth (GT).

The soft-max algorithm is employed in the network of classifications to categorize the "foreground and background of the anchor", while the $\mathrm{smooth}_{k,1}$function is utilized to determine the deviation between the positive anchor andthe GT. Subsequently, the ROIs could be verified. The computation of IOU is described in Equation (4):

$$\mathrm{IoU} = \frac{\mathrm{GT} \cap \mathrm{Anchor}}{\mathrm{GT} \cup \mathrm{Anchor}} \tag{4}$$

GT refers to the actual frame, whereas Anchor refers to the frame used as a reference. The soft-max value and $\mathrm{smooth}_{k,1}$ function can be denoted as equations (5) and (6) respectively:

$$T_j = \frac{\exp(b_j)}{\sum_{j=1}^{m} \exp(b_j)} \tag{5}$$

$$\mathrm{smooth}_{k,1}(w) = \begin{cases} 0.5w^2, |w| < 1 \\ |w| - 0.5, |w| \geq 1 \end{cases} \tag{6}$$

The equation (7) demonstrates the loss function used for RPN.

$$K(\{o_j\}, \{s_j\}) = \frac{1}{M_{cls}} \sum_j k_{cls}\left(o_j, o_j^*\right) + \lambda \frac{1}{M_{reg}} \sum_j O_j^* \, K_{reg}\left(s_j, s_j^*\right) \tag{7}$$

Let's define the following variables, j as the location of the present anchor, pi as the predicted probability that the j^{th} anchor box is a foreground, s_j^* as the actual probability that the j^{th} anchor box is a foreground, $s_j$as the estimated offset, s_j^* as the true offset, k_{cls}as the classification loss and M_{reg}as the regression loss. The value of λ should be chosen such that $\frac{1}{M_{cls}}$ and $\frac{1}{M_{reg}}$ are approximately equal. This ensures that the classification loss and regression loss are balanced. A value of λ equal to 1 is recommended.

The loss function utilized in FR-CNN is displayed in Equation (8):

$$K\,(o, v, s^{v}, u) = K_{cls}(o, v) + \lambda\,[v \geq 1]K_{reg}(s^{v}, u) \tag{8}$$

In this context, $O = (O_0, \ldots O_l)$ stands for the prediction category index, v represents the true category index, where v = 0 means that the grouping is background. The variables s and u demonstrates the coordinates of the prediction box and the GT box, respectively. The border loss is calculated when the target box is foreground, which is indicated by $[v \geq 1]$. The hyper-parameter λ is used to equalize the classification and regression loss, depending on previous experience, an amount of 1 is usually selected for λ.

The CSO method improves the process of selecting features and optimizing model parameters, while the FR-CNN utilizes its ability to identify detailed patterns from intricate medical imaging data. Collectively, they provide a harmonious combination that enhances the precision of cardiac disease forecasting and the computational efficiency of the overall procedure. This hybrid approach demonstrates the possibility of combining nature-inspired algorithms with modern deep learning applications, providing an enhanced and resilient solution for early diagnosis and specific treatment in heart disease. The methodology of the suggested ECO-FRCNN is mentioned in Algorithm 1.

Algorithm 1: Suggested ECO-FRCNN

Initialize CSO variables (population size, iterations, etc.)
Generate initial chicken population
Foreach chicken in the population:
a. Convert features to input format suitable for FR-CNN
b. Use FR-CNN to predict heart disease likelihood
c. Evaluate fitness based on prediction accuracy
For iteration in range(total iterations):
Update chicken positions using CSO algorithm
For each chicken in the population:
a. Convert features to input format suitable for FR-CNN
b. Use FR-CNN to predict heart disease likelihood
c. Evaluate fitness based on prediction accuracy
Select chickens with the best fitness values
Choose the chicken with the best position
Convert features to input format suitable for FRCNN
Use FRCNN for the final heart disease prediction
Output the final prediction and relevant information

3. Results and Discussion

The system configurations consisted of a Windows 11 operating system. We utilized

Python 3.10 to implement our recommended technique. The system has been equipped with an Intel HD Graphics. Different types of metrics including "precision, accuracy and F1 score", were analysed throughout the assessment phase of the results. This phase performed a comparative analysis between those findings and existing methodologies, including "Naive Bayes weighted (NB) [15], KNN [16], DT [15].

Precision refers to the amount of precision in correctly identifying positive predictions. It computes the proportion of accurately anticipated positive instances out of the total expected positive cases. A higher level of precision suggests a reduced occurrence of false positives. Table 1 and Figure 2 present the precision percentages for both the suggested and existing methods. The ECO-FRCNN technique produces a precision rate of 89,56 %, performed better in heart disease predictions. The precision is determined by using the following equation (9).

$$\text{Precision} = TP / (TP + FP) \tag{9}$$

Table 1. Performance evaluation (Source: Author)

Methods	Precision (%)	Accuracy (%)	F1-Score (%)
NB[15]	82,34	86	89,21
KNN[16]	74,9	70,12	73,45
DT[15]	81,24	77,61	89,13
ECO-FRCNN [Proposed]	89,56	91,56	94,86

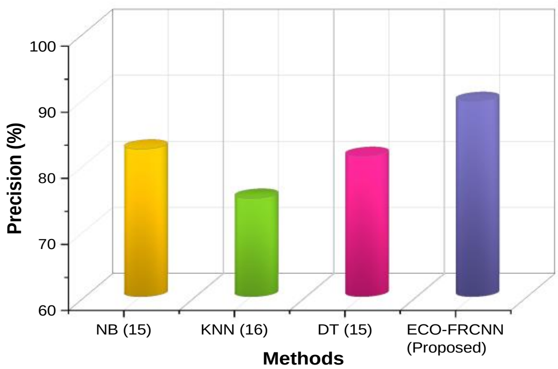

Figure 2. Precision (%)

(Source: Author)

Accuracy pertains to the model's capacity to discern instances with disease and instances without disease. Accuracy is the amount of predicted instances to the overall instances, serving as a complete representation of the model's overall accuracy in its predictions. Table 1 and Figure 3 present the accuracy rates for the suggested and existing approaches. The ECO-FRCNN technique demonstrates an accuracy level of 91,56 %, shown superior performance in heart disease predictions compared to other approaches. The accuracy has been calculated using equation (10).

$$\text{Accuracy} = \text{Total Predictions} / \text{Correct Predictions} \times 100\% \tag{10}$$

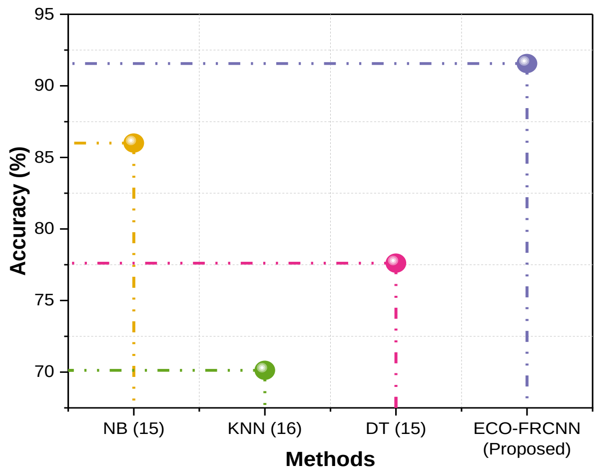

Figure 3. Accuracy (%) (Source: Author)

The F1 score, a scalar range from 0 to 1, provides a balanced evaluation of recall and precision in a prediction model. Within the framework of predicting heart disease, it measures the model's capability to predict positive instances while reducing both “false positives and false negatives”. The F1 scores for the proposed and traditional approaches are mentioned in Table 4 and Figure 4. The ECO-FRCNN strategy achieves an F1 score rate of 94,86 %, demonstrated superior performance in heart disease predictions comparison to other traditional methodologies. The F1 score has been computed using equation (11).

$$\text{F1} - \text{score} = 2 * (\text{Precision} * \text{Recall}) / (\text{Precision} + \text{Recall}) \tag{11}$$

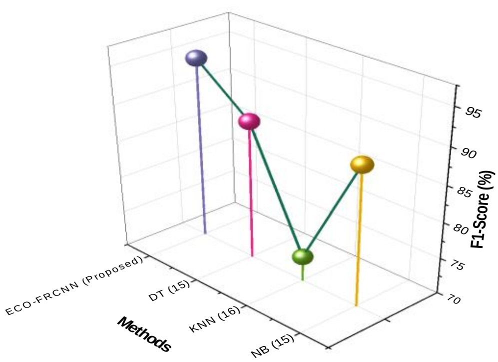

Figure 4. F1 score (%) (Source: Author)

4. Conclusion

Heart disease is a major global health problem including several conditions that impact the hearts. Prediction of heart disease approaches. In this research, we recommended ECO-FRCNN to improve the prediction rate in heart disease prediction. CSO algorithm was used to improve the effectiveness of the suggested FRCNN to enhance the predictions in heart disease. The comparison analysis was done on various parameters including precision (89, 56 %), accuracy (91, 56 %) and F1 score (94, 86 %) with other traditional approaches. The experimental outcomes illustrated the efficiency of the suggested approach. The generalize ability of the proposed method findings is limited due to the special characteristics of the dataset. Future study may explore the real-time implementation, expansion of datasets to enhance and validate the scalability of ECO-FRCNN.

References

1. Zakir M, Ahuja N, Surksha MA, Sachdev R, Kalariya Y, Nasir M, et al. Cardiovascular complications of diabetes: from microvascular to macrovascular pathways. Cureus. 2023 Sep 24;15(9). https://doi.org/10.7759/cureus.45835.
2. Shernazarov F, Zuhridinovna JD. Microcirculation disorders in the vascular system of the bulbar conjunctiva in the initial manifestations of cerebral blood supply deficiency. Science and innovation. 2022;1(Special Issue 2):515-22. https://doi.org/10.5281/zenodo.7367530.
3. MacLeod M, Papi A, Contoli M, Beghé B, Celli BR, Wedzicha JA, et al. Chronic obstructive pulmonary disease exacerbation fundamentals: diagnosis, treatment, prevention and disease impact. Respirology. 2021 Jun;26(6):532-51. https://doi.org/10.1111/resp.14041.
4. Shang N, Khan A, Polubriaginof F, Zanoni F, Mehl K, Fasel D, et al. Medical records-based chronic kidney disease phenotype for clinical care and “big data” observational and genetic studies. Npj Digital Medicine. 2021 Apr 13;4(1):70. https://doi.org/10.1038/s41746-021-00428-1.
5. Mannoh I, Hussien M, Commodore-Mensah Y, Michos ED. Impact of social determinants of health on cardiovascular disease prevention. Current opinion in cardiology. 2021 Sep 1;36(5):572-9. https://doi.org/10.1097/HCO.0000000000000893.
6. Patro SP, Nayak GS, Padhy N. Heart disease prediction by using novel optimization algorithm: A supervised learning prospective. Informatics in Medicine Unlocked. 2021 Jan 1; 26:100696. https://doi.org/10.1016/j.imu.2021.100696.
7. Nancy AA, Ravindran D, Raj Vincent PD, Srinivasan K, Gutierrez Reina D. Iot-cloud-based smart healthcare monitoring system for heart disease prediction via DL. Electronics. 2022 Jul 22;11(15):2292. https://doi.org/10.3390/electronics11152292.
8. Tuli S, Basumatary N, Gill SS, Kahani M, Arya RC, Wander GS, Buyya R. HealthFog: An ensemble DL based Smart Healthcare System for Automatic Diagnosis of Heart Diseases in integrated IoT and fog computing environments. Future Generation Computer Systems. 2020 Mar 1; 104:187-200. https://doi.org/10.1016/j.future.2019.10.043.
9. Gupta A, Yadav S, Shahid S, Venkanna U. HeartCare: IoT based heart disease prediction system. In2019 International Conference on Information Technology (ICIT) 2019 Dec 19 (pp. 88-93). IEEE. https://doi.org/10.1109/ICIT48102.2019.00022.
10. Khan MA, Algarni F. A healthcare monitoring system for the diagnosis of heart disease in the IoMT cloud environment using MSSO-ANFIS. IEEE Access. 2020 Jul 2; 8:122259-69. https://doi.org/10.1109/ACCESS.2020.3006424.
11. Khan MA. An IoT framework for heart disease prediction based on MDCNN classifier. IEEE

Access. 2020 Feb 17; 8:34717-27. https://doi.org/10.1109/ACCESS.2020.2974687.
12. Nandy S, Adhikari M, Balasubramanian V, Menon VG, Li X, Zakarya M. An intelligent heart disease prediction system based on swarm-artificial neural network. Neural Computing and Applications. 2023 Jul;35(20):14723-37. https://doi.org/10.1007/s00521-021-06124-1.
13. Tan L, Yu K, Bashir AK, Cheng X, Ming F, Zhao L, et al. Toward real-time and efficient cardiovascular monitoring for COVID-19 patients by 5G-enabled wearable medical devices: A DL approach. Neural Computing and Applications. 2021 Jul 4:1-4. https://doi.org/10.1007/s00521-021-06219-9.
14. https://www.kaggle.com/datasets/ineubytes/heart-disease-dataset
15. Nagavelli U, Samanta D, Chakraborty P. Machine learning technology-based heart disease detection models. Journal of Healthcare Engineering. 2022 Feb 27;2022. https://doi.org/10.1155/2022/7351061
16. Sujatha P, Mahalakshmi K. Performance evaluation of supervised machine learning algorithms in prediction of heart disease. In2020 IEEE international conference for innovation in technology (INOCON) 2020 Nov 6 (pp. 1-7). IEEE. https://doi.org/10.1109/INOCON50539.2020.9298354