Grenze International Journal of Engineering and Technology, June Issue

Performance Evaluation of AI Models for Investigating Post-COVID-19 Cardiovascular Disease

Swati S. Khandalkar[1] and Shweta M. Barhate[2]
[1-2]DECS, RTM Nagpur University, Nagpur, India
Email: swatiskhandalkar@gmail.com, shwetab73@yahoo.com

***Abstract*— Patients often experience post COVID-19 syndrome with varied symptoms which can continue long-term cardiovascular complications. This research work proposes to focus on the consequence of Post Covid-19 infection on cardiovascular health using Artificial Intelligence (AI) models. Dataset has been collected by the researcher from hospital for early-stage prediction of Post Covid-19 cardiovascular diseases. The said dataset is assessed to detect risk factors and predict cardiovascular outcomes. Understanding the Post COVID-19 effect on different cardiovascular disease types is important for early detection, long term monitoring and predictive modelling. This study proposes to use Artificial Intelligence models to predict Cardiovascular risk in Post Covid-19 patients. Most contributing features were recognized such as inflammatory biomarkers, hemoglobin, heart rate, BMI and family history contributing to Post Covid-19 cardiovascular abnormalities. The analysis shows that Recurrent Neural Network attains best predictive results as compared to other machine learning models. The findings suggest that deep learning predictive models give an edge in enhancing early diagnosis of Covid-19 cardiovascular complications.**

***Index Terms*— Cardiovascular disease, Deep Learning, Artificial Intelligence, Machine Learning, Recurrent Neural Network, Risk Prediction.**

I. Introduction

SARS-CoV-2 virus that is COVID-19 has affected the respiratory system and in addition affected multiple long-term complications across organ systems. Amongst several long-lasting complications post COVID-19, cardiovascular disease is observed as the most common between survivors of COVID-19. It is much difficult to know common reason for developing cardiovascular disease and its risks after contracting COVID-19[1]. Problems with difficulties, significantly affect the heart which an essential organ that circulates blood throughout the body. Cardiovascular disease was already a major source of death in underdeveloped nations, especially for younger people, but COVID-19 has made it worse. In response, health care providers are investigating novel diagnostic methods of CVD prediction after Covid-19. Increasing incidences throws light on the link in-between cardiovascular diseases (CVD) and COVID-19. This raises the concerns for long-term health risks to those who are recovering from the viral effects [2].
Current medical studies have reported that a substantial number of recovered COVID-19 patients experience cardiovascular problems, including Myocardial Infraction (MI), Hypertension, Heart Failure, Arrhythmia, Stroke, Coronary heart disease, Thrombosis [3]. These heart problems after COVID-19 happen because the body is still healing even after the virus is gone.

Grenze ID: 01.GIJET.12.2.210_22

Many people continue to have inflammation, mild damage to blood vessels, changes in the immune system and a higher chance of blood clotting. Blood tests like D-dimer, troponin, ferritin and CRP often stay higher than normal. When the levels stay high, the body is stressed and this can cause heart problems. Regular clinical test is often insufficient to identify early risk indicators.

Artificial Intelligence has come up as influential tool for medical diagnosis, prediction and pattern recognition. Machine learning and deep learning models can handle large medical datasets and identify hidden association among clinical variables. These models enable accurate prediction of post COVID cardiovascular disease [4]. AI-driven approaches support early disease detection, risk stratification and outcome prediction. They use clinical variables, demographic information and inflammatory biomarkers. Traditional ML algorithms like Naive Bayes, KNN, and SVM are extensively utilized in classification of diseases. Deep learning models like RNN are useful in gathering temporal patterns of healthcare data.

For learning multiple AI models to predict CVD risks in post COVID patients, statistical and AI techniques are used. This identifies important clinical factors linked to progression of disease. This stresses the importance of improved diagnostic and predictive approach. Results shows that in post COVID-19 patients, deep learning model early identifies CVD complications and its recovery.

II. Primary Objectives

- To study post-COVID-19 issues linked to health problems.
- To analyze how much health factors have been affected in post-COVID-19 individuals and their association with each other.
- To determine key clinical features and inflammatory biomarkers that contribute significantly to cardiovascular risk after COVID -19 infection.
- To develop and evaluate AI-based ML and DL models for classifying and predicting post-COVID cardiovascular complication.
- To design a predictive framework that can assist clinicians in early detection and decision making for post-COVID cardiovascular health.

III. Literature Review

There are many research studies showing that people often develop heart related problems after recovering form COVID-19. The most commonly reported issues include myocarditis, arrythmias, blood clots and high blood pressure. Several studies also highlight that inflammation continues even after recovery and important biomarkers such ad D-dimer, troponin, ferritin and CRP often remain high, increasing the risk of cardiovascular complications. Previous studies highlight the following evidence:

Mumtaz et al. [5] reported that patients of COVID-19 often have elevated inflammatory markers such as IL-6, ferritin, D-dimer and CRP. These increased markers are closely linked to heart injury both during and after the infection. The study found that patients with stronger inflammatory responses were more likely to develop heart problems like myocarditis and heart failure. The author concluded that tracking inflammatory biomarkers can help identify patients who are at greater risk of cardiovascular complication.

Alonso Dominguez et al. [6] used XGBoost to forecast long term COVID-19 from primary medical and laboratory data. Key factors included hemoglobin, oxygen saturation, CRP, APTT, sodium, ling infiltrates and sex with hemoglobin being statistically significant. Moderate accuracy of 63.9% and AUC-ROC of 0.732 was achieved by the model but 78.6% of high recall showed its potential for early detection of high-risk patients.

Naushad et al. [7] assesses Artificial Intelligence use in cardiovascular health after COVID-19. COVID-19 increased heart problems like myocarditis, arrythmias and blood clots, AI helps predict risks, analyze biomarkers, imaging and support remote monitoring. It can improve public health and guide treatment for post COVID patients.

Hovornyan et al. [8] studied 328 COVID-19 patients to assess long term heart complications and developed an AI model for risk prediction. Key predictors included D-dimer, CRP and neutrophil/lymphocyte ratio. The model achieved ROC-AUC 0.891, sensitivity 92.4%, specificity 90.1 and 87.2% accuracy for predicting post-COVID cardiovascular events.

Gupta et al. [9] proposed a stacking ensembled based ML model to predict difficulties in post-COVID-19 clinical data. The model combined multiple classifiers and outperformed individual methods such as SVM, decision tree and random forest. It achieved accuracy of 93.23%, specificity of 95.74%, recall of 92.05%, and

precision of 95.24%. The results show the effectiveness of ensemble learning for predicting post-COVID complication.
Islam et al. [10] studied post-COVID-19 health effects using machine learning techniques on survey data from recovered patients in Bangalesh. The analysis focused on demographic, physiological and neurological factors like age, sleep patterns, gender, anxiety and memory. Statistical test confirmed significant relationships among these variables. Among the evaluated models, the Decision tree achieved the best performance and effectively identified key features using Gini index, information gain and SHAP methods.
Juliet et al. [11] developed a decision-making system using reinforcement learning to predict cardiovascular disease in long COVID patients. The model assesses data of each patient and provides personalized risk assessment. With high sensitivity and specificity, it achieved accuracy of 97.88%.
Yousif et al. [12] predicted cardiovascular complications in 352 post-COVID patients from Iraq by using machine learning algorithm. They observed demographics, clinical data, inflammatory markers and imaging. The model early identified high risk patients for timely intervention with high accuracy, sensitivity and specificity.
Khan et al. [13] studied ECG images for heart effects using a hybrid efficient NET-DBN in post-COVID patients. Level one classified ECG as normal or abnormal and level two classified as myocardial infarction or COVID related issues. The model showed good performance in detecting heart problems after COVID-19 with 94.70% sensitivity, 91.80% specificity and 89.80% accuracy.
Bhushan Shah et al. [14] studied heart problems in COVID-19 affected patients. In 120 healthy people and 92 recovered patients, they studied heart rate variability (HRV) using short-term ECG. This model found HRV as the main feature to distinguish healthy and heart patients.

IV. MATERIALS AND METHODS

This research investigates the outcome of post COVID-19 conditions on CVD. The data was pre-processed and models were prepared to recognize risk factors and assess its prediction performance.
This section defines the methods used to comparative analysis of RNN with traditional machine learning algorithms for prediction of cardiovascular disease. Standard classification models were used such as Naive Bayes, KNN and SVM as a base model. Figure 1. shows the proposed methodology for predicting cardiovascular diseases. The steps are as follows –

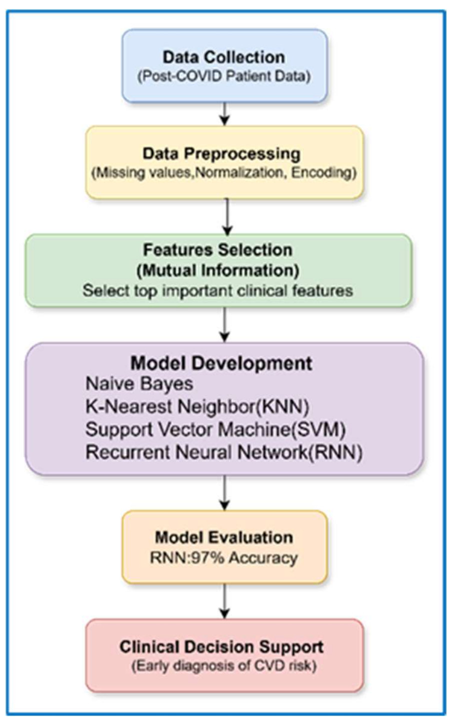


Figure 1. Structure of Proposed Model

A. Data Collection

The aim of this study is to study CVD in post COVID-19 patients keeping their data private. The study uses an observational design by combining self-generated synthetic data and hospital data. Hospital records of post COVID-19 patients were collected for data collection. All personal information was removed, and the data were anonymized to protect patient privacy. Synthetic data were generated to increase size and balance the distribution of cardiovascular outcomes. This research studied total 1,384 post-COVID male and female patients from young to old. The final dataset includes only safe, non-identifiable information.
The dataset includes:

- Demographic (Age, Gender)
- Clinical (Height, Weight, BMI, FBS, Sys_BP, Dia_BP, Diabetes, Total_Cholesterol, Thyroid, ECG, Chest_Pain, Max_Heart_Rate)
- Lifestyle (Alcohol, Smoking)
- Laboratory (Hemoglobin, CRP, IL6, Ddimer, Troponin_T)
- Outcome (CVD_Diagnosis)

B. Data Processing

Missing values were checked and filled or removed. Categorical variables were converted into numbers. Continuous variables were scaled. Outliers and unrealistic values were removed to improve data quality.

C. Feature Selection

In this study, Mutual Information was used to analyse feature importance. Mutual Information [15] measure the amount of information a feature provides for CVD outcome. It aids to recognize the association in-between the target class and input variables. Features with higher Mutual Information scores have a stronger impact on predicting cardiovascular disease. This method was used to rank the features and identify the most important clinical and inflammatory factors related to post -COVID heart risk. The contribution of each feature to the prediction was evaluated. Features such as CRP, IL6, Ddimer, Troponin_T, Max_Heart_Rate, BMI and Hemoglobin were analyzed to see which had the most impact on CVD outcomes.

D. Model Development

Four models were used in this study to predict cardiovascular disease in post-COVID-19 patients.

Naive Bayes
Naive Bayes is a probabilistic machine learning model based on Baye's theorem. This considers all input features as independent of each other. The model calculates the probability of a patients having cardiovascular disease based on given input values [16]. Naive Bayes is fast, easy to implement and works well with medical datasets.

KNN
K-Nearest Neighbors (KNN) is a machine learning model for classification that doesn't consider any specific data distribution(non-parametric). First, this calculates the distance between points using methods like Euclidean, Manhattan or Hamming distance. The k value decides how many neighbors to check for classifying a point. If k=1, the point takes the class of its nearest neighbor. Small k gives high variance but low bias, while larger k gives low variance and low bias [17]. KNN can also handle missing values by estimating them during classification.

SVM
Support Vector Machine (SVM) is a supervised machine learning model. This could be used for classification and regression purpose. SVM separates data into classes by creating a line or a hyperplane. It works for both linear and nonlinear problems. SVM tries to find the best line or hyperplane that keeps the maximum distance from the closest points of each class [18]. This helps the model make more accurate predictions. It is widely used in medical data and other fields for reliable classification.

RNN
Recurrent Neural Network (RNN) is a deep learning model that work with sequential data. It is very unique as its connections come from hidden layers. These hidden layers are repeated multiple times and each time it applies the same weights and biases to the input step. The network runs its processes in a loop pattern. This changes and updates the hidden states and save data to its internal memory. RNN model is updated and rebuilt every time the data is passed through the chain of RNN.

RNNs easily understands the data during training process. It can handle both categorical and numerical data. It needs small amount of data preparation for validation of the model with statistical tests. It also works well with large data sets [19, 20]. RNNs in their statistical algorithms, uses cluster points of data within feature groups. Different types of variables make large data complex thereby making process of classification and clustering difficult [20]. RNNs capture the updates and record them in their neuron units due to the storage facility that they possess. This is a unique quality of RNN in deep learning models. RNNs has the ability to collect data from randomly long sequences while making predictions.

E. Evaluation Metrics

Model performance was measured using accuracy, precision, recall, f1-score and confusion matrix [21]. These metrics shows how well the model predicts cardiovascular disease cases. Some of them are as follows:

$$\text{Accuracy} = \frac{TP+T}{TP+TN+FP+FN} \tag{1}$$

$$Recall = \frac{TP}{TP+F} \tag{2}$$

$$Precision = \frac{TP}{TP+FP} \tag{3}$$

$$F1\,Score = 2 * \frac{precision * recall}{precision + recal} \tag{4}$$

Above equation represents TP, TN, FP and FN as True Positives, True Negatives, False Positives and False Negatives respectively. True Positives and true negatives determine whether a post COVID-19 patient had CVD or not. While false predictions made by the machine learning model is determined by the number of false positives and false negatives.

V. RESULTS AND DISCUSSION

This study compared Naive Bayes, KNN, SVM and RNN models to predict post COVID-19 cardiovascular disease risk. The performance of each model was evaluated to identify the most effective approach. Feature importance was analysed using Mutual Information as shown in Figure 2. This shows the top 10 important features selected using Mutual Information. It measures how much each feature helps in predicting cardiovascular disease. Troponin-T has the highest score, so it is the most important feature. COVID-19, D-dimer and CRP also have high scores, showing strong influence on CVD prediction. IL6 and Hemoglobin have moderate importance. These features are related to inflammatory and oxygen levels in blood. Max_Heart_Rate and BMI contribute less but still help the model. Family history and smoking have lower scores, but they still add useful information.

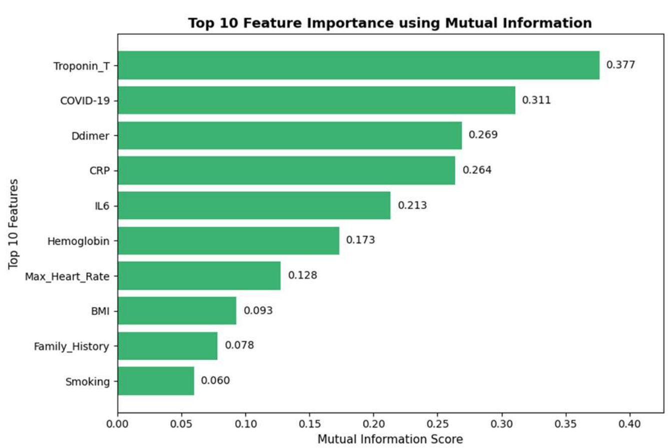
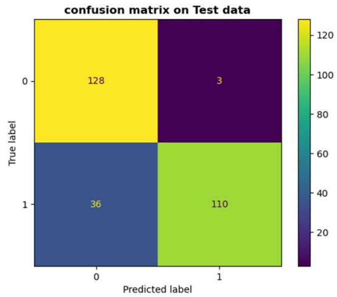
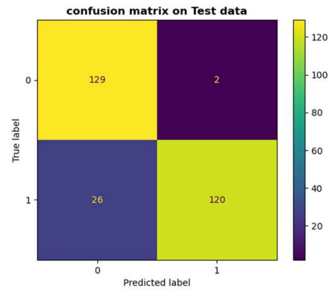
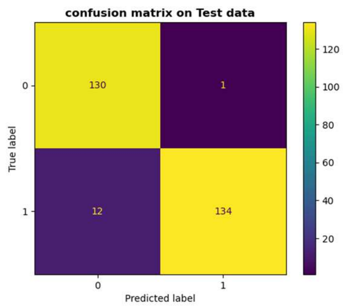


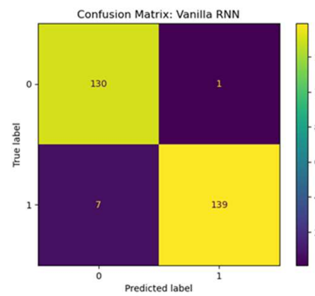
Figure 2. Features Importance





(a) Naive Bayes (b) KNN





(c) SVM (d) RNN

Figure 3. Values obtained for the confusion matrices using different algorithms

TABLE I: ANALYSIS OF MACHINE AND DEEP LEARNING ALGORITHMS

Model Name	Accuracy	Precision	Recall	f1-score
Naive Bayes	0.86	0.97	0.75	0.85
K-Nearest Neighbor (KNN)	0.90	0.98	0.82	0.90
Support Vector Machine (SVM)	0.95	0.99	0.92	0.95
Recurrent Neural Network (RNN)	0.97	0.99	0.95	0.97

Confusion matrices were used to analyse correct and incorrect predictions as shown in Figure 3. For Naive Bayes, the model correctly identified many low-risk cases but missed some high-risk cases which reduced recall. KNN model improved classification by reducing misclassification and showed better balance between TP and TN. The SVM model achieved high TP and TN values, showing strong performance in separating risk classes. The RNN model performed best, with the highest correct predictions and the lowest errors, indicating better learning of complex patterns in the data. Overall, the confusion matrices confirmed that RNN provides the most accurate and reliable predictions for post-COVID CVD risk compared to other models.

Model performance was compared using accuracy, precision, recall and F1-score as shown in table I. Naive Bayes achieved an accuracy of 86% but its recall value was lower compared to other models, indicating that some positive cases were not correctly identified. In cardiovascular risk prediction, missing positive cases is critical, as it may delay necessary medical intervention. KNN improved performance with 90% accuracy and better recall than Naive Bayes. This means it detected more positive cases. But, KNN depends on nearby data points. So, it may not perform as well when the data has complex relationships.

SVM performed well with 95% accuracy and balanced precision and recall. This means it classified both positive and negative cases well. It works effectively even when the data has many features and complex patterns. Amongst all models, the RNN achieved the best results with 97% accuracy, 99% precision, 95% recall and f1-score 97%. It identified most positive cases and made very few mistakes. The better performance is because RNN can learn complex patterns in the data selected using Mutual Information.

Overall results shows that clinical biomarkers and COVID-19 status are the key predictors. These results support clinical findings that COVID -19 increases inflammation and stress which can damage the heart. Performance of RNN model is improved by these features which ultimately supports medical findings that are related to heart disease.

VI. CONCLUSIONS

This research highlights a strong association between CVD and post COVID-19 circumstances. Large number of Covid-19 patients have greater risk for heart issues because of changes in inflammatory and cardiac markers. The conclusive results confirms that even after recovery, COVID-19 can influence heart health. The research provides AI based comparative and comprehensive analysis in forecasting CVD risk amongst post COVID-19 patients. Amongst all the tested models, RNN performed as the best model with 97% accuracy as compared to Naive Bayes, KNN, and SVM in predicting post COVID-19 CVD risk. This confirms that RNN is the best model in identifying post COVID-19 heart problems.

This research also aids to interrelate different health factors, inflammatory biomarkers and clinical features that increase heart risk post COVID-19. To evaluate and predict post COVID-19 heart complications, ML and DL models were developed. Finally, the research designed a prediction framework to support doctors in early detection and faster as well as better decision making in heart issues of post COVID-19 patients.

In the future, this work can use larger datasets and longer follow-up data. Data from more hospitals can also be added. Additionally, explainable AI methods like SHAP and LIME can be used for deeper clinical interpretability. This may aid doctors predict heart risk early and improve patient outcome.

REFERENCES

[1] D Srinuvasa Rao, Desidi Narsimha Reddy, Lova Naga Babu Ramisetti, M S Radha Manga Mani, Harikrishna Pathipati, Siva Kumar Pathuri and Mohana Prasad Mendu, "Post-COVID Cardiovascular Risk Assessment using an Efficient Machine Learning Model", 3rd International Conference for Advancement in Technology (ICONAT), pp.1-7, 2025.

[2] Nidhi Priyadarshini and Phillip Smith, "The Impact of COVID-19 on Cardiovascular Disease: A Machine Learning Predictive Study", World Journal of Cardiovascular Diseases, 15(2), pp. 19-47, 2025.

[3] Ting Zhang, Zhimao Li, Qimin Mei, Joseph Harold Walline, Zhaocai Zhang, Yecheng Liu, Huadong Zhu and Bin Du, "Cardiovascular Outcomes in Long COVID-19: A Systematic Review and Meta-Analysis", Frontiers in Cardiovascular Medicine, 1(1), pp. 1-9, 2025.

[4] Evandro Carvalho de Andrade, Luana Ibiapina C C Pinheiro, Placido Rogerio Pinheiro, Luciano Comin Nunes, Mirian Caliope Dantas Pinheiro, Maria Lúcia Duarte Pereira, Wilson Correia de Abreu, Raimir Holanda Filho, Marum Simao Filho, Pedro Gabriel C D Pinheiro, Rafael Espindola Comin Nunes, "Hybrid Model for Early Identification Post-Covid-19 Sequelae", J Ambient Intell Humaniz Comput.,1(1), pp. 1-12,2023.

[5] Ayesha Mumtaz, Erum Rehman, Mohammad Anisur Rahaman and Shazia Rehman, "Inflammatory Biomarkers and Cardiac Injury in COVID-19 Patients", Frontiers in Public Health, 10(1), pp. 1-12. 2022.

[6] Jacobo Alonso-Domínguez, Jose A. Gonzalez-Novoa, Aida Lopez-Lopez, Ines Martinez-Barros, Alexandre Perez-Gonzalez, Virginia Leiro-Fernandez, Jose Aguayo-Arjona, Susana Teijeira, Cesar Veiga and Eva Poveda, "Artificial Intelligence-Based Identification of Key Risk Factors for Long COVID from Early Clinical Data", Journal of Infection and Public Health, 19(1), pp.1-7, 2025.

[7] Zayera Naushad, Jaya Malik, Abhishek Kumar Mishra, Shilpy Singh, Dharmsheel Shrivastav, Chetan Kumar Sharma, Ved Vrat Verma, Ravi Kant Pal, Biswajit Roy and Varun Kumar Sharma, "Artificial Intelligence in Cardiovascular Health: Insights into Post-COVID Public Health Challenges",32(1), pp. 475–494, 2025.

[8] Anastasiia Hovornyan and T. Ilashchuk, "Analysis of Long-Term Cardiovascular Outcomes of COVID-19. Artificial Intelligence Model for Risk Evaluation", European Journal of Preventive Cardiology, 30(1), pp. 223-233, 2023.

[9] Aditya Gupta, Vibha Jain and Amritpal Singh, "Stacking Ensemble-Based Intelligent Machine Learning Model for Predicting Post-COVID-19 Complications", New Generation Computing, 40(1), pp. 987–1007, 2022.

[10] Muhammad Nazrul Islam, Md Shofiqul Islam, Nahid Hasan Shourav, Iftiaqur Rahman, FaizAl Faisal, Md Motaharul Islam and Iqbal H. Sarker, "Exploring Post COVID 19 Health Effects and Features with Advanced Machine Learning Techniques", 14(1), pp.1-25, 2024.

[11] Diana Juliet and Banumathi J, "Prescriptive Analytics Decision-Making System for Cardiovascular Disease Prediction in Long COVID Patients using Advanced Reinforcement Learning Algorithms", The Journal of X-Ray Science and Technology, 33(5), 2025.

[12] Maitham G. Yousif and Hector J. Castro, "Predicting Cardiovascular Complications in Post-COVID-19 Patients using Data Driven Machine Learning Models", Medical Advances and Innovations journal, 1(3), pp. 1-12, 2023.

[13] Mohammed Abdul Basith Ali Khan and Edara Sreenivasa Reddy, "Post-COVID Effect on Heart After Recovery Based on Hybrid Efficientnet-DBN with Multilevel Classification using ECG Images", EngMedicine, 1(2), pp.1-12, 2024.

[14] Bhushan Shah, Shekhar Kunal, Ankit Bansal, Jayant Jain, Shubhankar Poundrik, Manu Kumar Shetty, Vishal Batra, Vivek Chaturvedi, Jamal Yusuf, Saibal Mukhopadhyay, Sanjay Tyagi, Girish Meenahalli Palleda, Anubha Gupta and Mohit Dayal Gupta, "Heart Rate Variability as A Marker of Cardiovascular Dysautonomia in Post-COVID-19 Syndrome using Artificial Intelligence", Indian Pacing and Electrophysiology Journal, 22(2), pp. 70-76, 2022.

[15] Hosam El-Sofan, Belgacem Bouallegue and Yasser M. Abd El-Latif, "A Proposed Technique for Predicting Heart Disease using Machine Learning Algorithms and an Explainable AI Method", Scientific Reports,14(1), pp. 1-18, 2024.

[16] Balaji Shesharao Ingole, Vishnu Ramineni, Nikhil Bangad, Koushik Kumar Ganeeb, Priyankkumar Patel, "Advancements in Heart Disease Prediction: A Machine Learning Approach for Early Detection and Risk Assessment", International Journal of Research and Analytical Reviews (IJRAR), 11(4), pp. 164-172, 2024.
[17] Debangana Mahapatra aand Hemlata Joshi, "Prediction of Heart Disease Using Machine Learning for Framingham Dataset", Mapana-Journal of Science, 23(3), pp. 65-80, 2024.
[18] Ghulab Nabi Ahmad, Shafiullah, Hira Fatima, Mohamed Abbas, Obaidur Rahman, Imdadullah and Mohammed S. Alqahtani, "Mixed Machine Learning Approach for Efficient Prediction of Human Heart Disease by Identifying the Numerical and Categorical Features", Applied Science, 12(7449), pp. 1-33,2022.
[19] S. Krishnan, P. Magalingam, and R. B. Ibrahim, "Advanced Recurrent Neural Network with Tensorflow for Heart Disease Prediction," International Journal of Advanced Science and Technology, 29(5), pp. 966-977, 2020.
[20] N. S. R. Pillai, K. K. Bee, and J. Kiruthika, "Prediction of Heart Disease using RNN Algorithm," International Research Journal of Engineering and Technology (IRJET), 6(3), pp. 4452-4458, 2019.
[21] Israa Nadheer, "Heart Disease Prediction System using Hybrid Model of Multilayer Perception and Xgboost Algorithms", BIO Web of Conferences 97(1), pp. 1-9, 2024.