Heart attack prediction using multi-modal neural networks architecture

Harsh Prajapati

Department of Computer Science and Engineering
National Institute of Technology Rourkela

Heart attack prediction using multi-modal neural networks architecture

Dissertation submitted in partial fulfillment

of the requirements for the degree of

Bachelor in technology

in

Computer Science and Engineering

by

Harsh Prajapati

(Roll Number: 120CS0206)

based on research carried out

under the supervision of

Prof. Dr. Ratnakar Dash

May, 2024

Department of Computer Science and Engineering

National Institute of Technology Rourkela

Department of Computer Science and Engineering
National Institute of Technology Rourkela

Prof. Dr. Ratnakar Dash
Professor

May 5, 2024

Supervisor's Certificate

This is to certify that the work presented in the dissertation entitled *Heart attack prediction using multi-modal neural networks architecture* submitted by *Harsh Prajapati*, Roll Number 120CS0206, is a record of research carried out by him under my supervision and guidance in partial fulfillment of the requirements of the degree of *Bachelor in technology* in *Computer Science and Engineering*. Neither this dissertation nor any part of it has been submitted earlier for any degree or diploma to any institute or university in India or abroad.

Dr. Ratnakar Dash

Declaration of Originality

I, *Harsh Prajapati*, Roll Number *120CS0206* hereby declare that this dissertation entitled *Heart attack prediction using multi-modal neural networks architecture* presents my work carried out as a undergraduate student of NIT Rourkela and, to the best of my knowledge, contains no material previously published or written by another person, nor any material presented by me for the award of any degree or diploma of NIT Rourkela or any other institution. Any contribution made to this research by others, with whom I have worked at NIT Rourkela or elsewhere, is explicitly acknowledged in the dissertation. Works of other authors cited in this dissertation have been duly acknowledged under the sections "Reference" or "Bibliography". I have also submitted my original research records to the scrutiny committee for evaluation of my dissertation.

I am fully aware that in case of any non-compliance detected in future, the Senate of NIT Rourkela may withdraw the degree awarded to me on the basis of the present dissertation.

May 5, 2024
NIT Rourkela

Harsh Prajapati

Acknowledgment

First and foremost, I would like to express my sincere gratitude to my esteemed guide, Prof. Dr. Ratnakar Dash, for his invaluable guidance, unwavering support, and continuous encouragement throughout the course of this project. His profound knowledge and expertise in the field of deep learning and multimodal techniques have been instrumental in shaping this work. I am deeply indebted to the faculty members of the Department of Computer Science and Engineering for their insightful suggestions and constructive feedback, which have significantly enriched the quality of this project. My heartfelt appreciation goes out to my family and friends, whose constant love, understanding, and moral support have been a source of strength and motivation throughout this challenging endeavor. I would also like to extend my gratitude to the authors of the various research papers, books, and online resources that have provided me with the necessary foundation and knowledge to undertake this project successfully. Finally, I am grateful to the almighty for bestowing upon me the perseverance, determination, and ability to complete this project successfully.—

May 5, 2024
NIT Rourkela

Harsh Prajapati
Roll Number: 120CS0206

Abstract

Electrocardiogram (ECG) and blood report play important roles in the diagnosis and prevention of any kind of cardiovascular diseases (CVDs). After the growth of computation power and ample amount of data, detection of CVDs has gathered a huge attention. But the currently available model require only a single source of data. In the medical science, there is no particular test that can confirm the presence of a disease, it is the sequence of tests that lead to the diagnosis of the disease. Hence it is desirabe to design a multi-modal deep learning methods for the prediction of the heart diseases or CVDs. By using a dense and convolution neural network I am extracting the deep-coding features from the datasets respectively. They are merged by using different fusion techniques so that the important feature set contribute in the prediction of the disease. Experimental results denote that the performance of our method is better to those of a single modal methods and alternatives. My method reaches an accuracy of around 87% in predicting the disease in the evaulation set.

Keywords*: *ECG*; *Multimodal*; *neural networks*; *convolution*; *CVD

Contents

Chapter 1

Introduction

Cardiovascular disease (i.e. CVD) remains a significant global health concern, claiming an estimated 17.9 million lives in 2017 alone. This number makes around 30% of all the global deaths (Li et al., 2021). Physicians have an array of diagnostic tools available to them, ranging from cardiac auscultation (potentially recorded as a phonocardiogram) to electrocardiography, ultrasound imaging, CT scans, myocardial enzyme testing, and angiography. However, the intricate nature of cardiovascular activity and the existence of numerous subtypes of cardiovascular disease, such as arrhythmia, mitral valve prolapse, and coronary artery disease, pose challenges for diagnosis. Currently, there is no single tool that can provide a comprehensive view of cardiovascular disease.

Recognizing this complexity, there is a growing emphasis on the use of multimodal data for diagnosis. Electrocardiography (ECG) and blood tests are particularly prominent due to their cost-effectiveness and convenience. ECG records the heart's electrical activity throughout each cardiac cycle, while blood tests provide insights into the composition of the blood. In clinical practice, it is crucial for physicians to integrate findings from both blood tests and electrocardiograms when formulating a diagnosis. This holistic approach enhances diagnostic accuracy and ensures comprehensive assessment of cardiovascular health.

The dataset is trained using deep learning classification techniques such as convolutional neural networks (CNNs). These techniques improve classification accuracy and help run classification and predictive models. Cardiac prediction systems can extract features from data to reveal hidden patterns in the data, potentially allowing healthcare professionals to improve the quality of their services. In this article, we have listed the necessary parameters for early detection of heart disease. Comparing different results shows the accuracy of the model for different parameters. The results of this work are presented based on the algorithm's evaluation metrics. Finally, we discuss the conclusions and future directions of this study.

Chapter 2

Research Problem

Cardiovascular diseases, especially heart attacks, stand as a predominant global health challenge. The difficulty in predicting the occurrence of a heart attack with precision remains a critical problem. This research project addresses this challenge by proposing a solution based on artificial neural networks. Specifically, the research aims to develop a multimodal neural network capable of handling two distinct types of inputs – tabular blood reports in .csv format and ECG datasets – with a focus on achieving heightened accuracy when both datasets are available.

The choice of this research problem is grounded in the urgent need for more accurate and nuanced methods of predicting heart attacks, a leading cause of mortality worldwide. Conventional risk assessment models often lack the depth required for early and precise detection. The integration of blood reports and ECG data into a single, flexible model is motivated by the complementary nature of these datasets. Blood reports offer comprehensive biochemical insights, while ECG data captures vital cardiac activity patterns.

The significance of this problem is underscored by the potential to revolutionize preventive healthcare. Timely prediction of heart attacks allows for proactive interventions, reducing the burden on healthcare systems and improving patient outcomes. By addressing the limitations of existing models and leveraging the capabilities of artificial neural networks, this research seeks to contribute to the development of a sophisticated yet accessible tool for accurate heart attack prediction.

In essence, the research problem is chosen not only for its scientific and technological challenges but, more importantly, for its potential real-world impact on public health. The proposed multimodal neural network represents a step towards personalized and efficient healthcare, aligning with the broader goals of enhancing preventive medicine in the era of advanced computational methodologies.

Chapter 3

Research Objectives

This project aims to develop a predictive model for diagnosing heart disease utilizing a multi-modal neural network architecture capable of processing diverse inputs. The model processes features from two distinct sources: blood report data stored in an .xlsx format and feature-extracted electrocardiogram (ECG) data. This architecture allows for flexibility in input selection, accommodating either data source individually or in combination to predict the healthiness of the heart. By using the deep learning techniques, the model learns intricate patterns within the data to provide accurate diagnostic assessments. Through rigorous experimentation and validation, this project seeks to contribute to the advancement of early detection and intervention strategies in cardiovascular health.

Objectives:

1. Develop a multi-modal neural network architecture capable of processing blood report .xlsx files and feature-extracted ECG data.

2. Implement data preprocessing techniques to clean and standardize both types of input data.

3. Explore feature engineering methods to extract relevant information from the blood report and ECG data.

4. Design and train the neural network model to predict the presence of heart disease based on the provided inputs.

5. Investigate the impact of using individual inputs versus combined inputs on the model's predictive capabilities.

6. Conduct extensive validation experiments to assess the robustness and generalization ability of the model.

7. Compare the performance of the multi-modal neural network model with existing approaches for heart disease prediction.

8. Contribute to the advancement of early detection and intervention strategies in cardiovascular health

Chapter 4

Lirerature Review

1. **Heart Disease Detection using Machine Learning Technique [1]**:
   In conducting the aforementioned experiment, it is apparent that the outcomes from Naïve Bayes and decision tree algorithms may vary. Consequently, it's necessary to compare both algorithms for each prediction to ensure accuracy. Additionally, relying solely on a single algorithm without data preprocessing capabilities may lead to suboptimal accuracy. Therefore, a hybrid approach incorporating multiple algorithms such as k-means, ID3, and Naïve Bayes is proposed to enhance accuracy.

2. **Heart Disease Prediction Using Machine Learning: A Systematic Literature Review:[2]**
   This paper conducts a systematic literature review (SLR) spanning the years 2021 to 2023, focusing on the utilization of machine learning in predicting heart disease. The SLR encompasses three phases: planning, execution, and documentation. In the planning phase, the necessity of the review is established, research questions are formulated, and a search strategy is devised using the PICOC technique.

3. **Multimodal Deep Neural Network with Image Sequence Features for Video Captioning:[3]**
   This research paper focuses on the challenge of automatically describing video clips, which has applications in areas like video search, interaction with machines, and aiding physically handicapped individuals. The authors compare the effectiveness of a model called S2VT with a recently proposed approach called NeuralTalk2 for image captioning. S2VT uses a sequence-to-sequence model with Long Short-Term Memory (LSTM) networks to process video frames and generate textual descriptions. While effective, S2VT sometimes produces inaccurate sentences, possibly due to limited data. To address this, the authors integrate S2VT with NeuralTalk2 to create MDNNiSF, aiming to improve accuracy by leveraging additional image caption data. The paper suggests that combining these approaches can enhance video captioning accuracy and presents experimental results using Microsoft Research datasets.

4. **Multimodal learning using convolution neural network and Sparse Autoencoder** [4]

   In this study, they created a system to classify patterns using a combination of convolutional neural networks (CNN) and sparse autoencoders (SAE), bringing together information from MRI and FDG-PET scans. Their main goal was to compare this fusion approach with using each scan alone, emphasizing the benefits of combining different types of information. The experiments showed that the method effectively captures local 3D patterns from both MRI and FDG-PET, and it significantly improves the accuracy of Alzheimer's disease prediction compared to using each scan separately. This improvement was observed in both the fusion of modalities and when using CNN with pretraining from SAE.

   Results:

Method	Only MRI	Only PET	proposed PET MRI
Simple CNN	80.62	81.93	84.7
SAE + CNN	85.24	85.53	91.14

Chapter 5

Methodology

5.1 Data collection

The dataset containing blood reports of heart patients was obtained from Kaggle. It includes variables such as age, gender, pain level, cholesterol levels, blood glucose levels, and resting blood pressure. Prior to analysis, the dataset underwent preprocessing, which involved removing records with missing or irrelevant values and averaging when necessary. Any remaining NaN values were also eliminated during the cleaning process. The dataset was then divided into training and test sets, with 25% reserved for testing.

5.2 Pre-processing for ECG dataset:

The dataset is downloaded from PTB-XL(physionet) dataset repository. The ECG has been processed to extract the signals in mV in a table containing the peaks of the signals. The dataset has 783 columns.The dataset has various columns such as p_peaks, r_peaks,p_area,q_area and many columns.

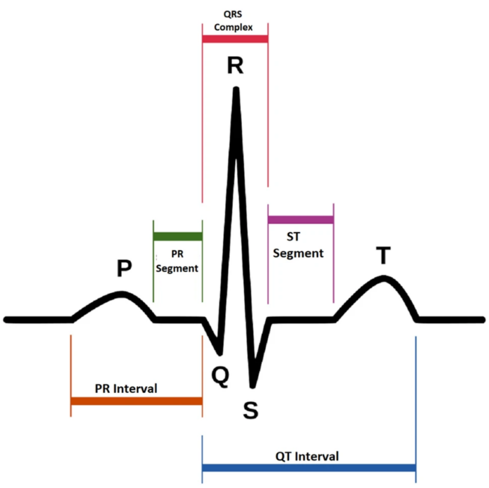

Figure 5.1: features from ECG signals (referred from [5])

The model is also trained for finding out most correlated 50 features among the 783 columns in the dataset.

The correlation coefficient is calculated by using the following formula.

$$\rho_{xy} = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2 \sum_{i=1}^{n}(y_i - \bar{y})^2}}$$

The target file for this dataset contains the list of diseases in abbreviated format.For example ”['NSR', 'LVH3', 'LPAREN', 'RAVL', 'COMMA', 'CORNPROD', 'RPAREN', 'NST', 'AB'],” But for this project we are targeting only binary classification. Hence an analysis was done to identify the important diseases that are more fatal to health of heart. A list of five diseases were made and the target was set to 1, if the person contained any 3 of those diseases.

5.3 Multi-Modal implementation

The first dataset is the blood report of the person in the tabular format. The input is fed to a artifical neural network architecture. After that the prediction are calculated. Similarly for the case of second dataset, which is the ECG dataset with signal extracted - a model containing 1D convolution layers are used to find patterns in the signals. It is tested to work better than the vanilla ANN architecture. The final layers from each of the model are removed. Hence the first model has 8 neurons just behind the output neuron. and the second model has 128 neurons just behind the output neuron. Hence when data is fed to both the models they return a list of features of the respective size. [6]

After the inputs are processed and features are extraced they are stored as a list of features and fused together according to some criteria. The most simple method to fuse the data is to concatenate the features and pass to the next level. The features are fused and then the combined features are passed to the final model to get the prediction.

The parameters such as the activation functions, epochs, number of hidden layers and neurons, etc are modified and tested to produce the best results.

Chapter 6

Proposed approach

6.1 Model Structure

I have chosen ’1’ to represent bad heart health of a person and ’0’ for a healthy heart. The blood report dataset is pre-processed and passed through the ANN architecture network. The features from the model is extracted and stored in a list. Similarly the ECG dataset is pre-processed and passed through an CNN architecture network. Before passing them through their respective models they were properly cleaned and augmented to give better results. The input layer 1 and input layer are trying to extract the correlated features from the datasets. Individually the feature vectors from both of the models may not be correlated.

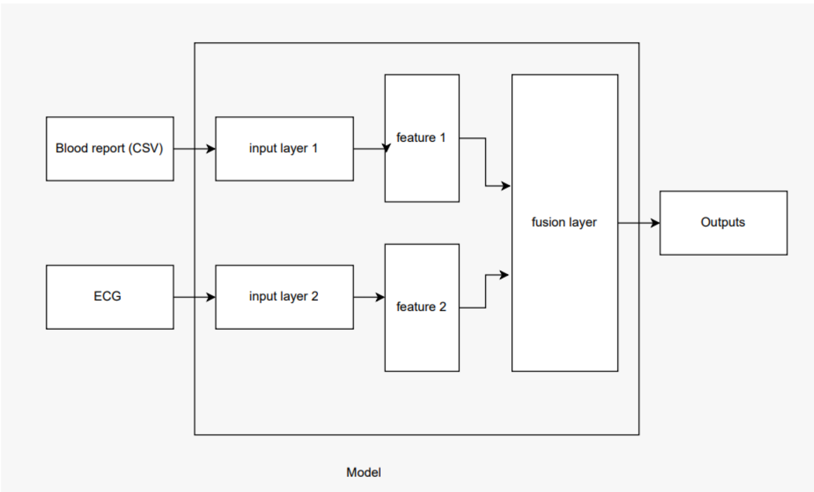

Figure 6.1: Proposed model

The features from the model is extracted and stored in a another list. [7]
The list is then concatenated using different methods. The results are calculated using different type of important features are retained using correlation method.

correlation(x, y) = $\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})\frac{}{\sqrt{\sum_{i=1}^{n}(x_i-\bar{x})^2\sum_{i=1}^{n}(y_i-\bar{y})^2}}$

6.2 Noise reduction

Normalization of features can be performed so that the final model can be trained with less noise. The last layers of both the model for the blood report and the ECG data are removed and in-between features are extracted. The extracted features are then normalized so that they can be passed to the next layer. The final model after the fusion is an ANN binary classifier model which is predicting the final outputs.
Most of the features in the ECG model were not correlated or ha negative correlation. The columns that were not contributing to the prediction had to be removed. The matrix for correlation was made and the columns that were not contributing were analysed. After the analysis it was seen that mostly 50 columns in the input dataset were contributing features and rest of them were outliers or noises.

6.3 Late Fusion

Late fusion is a technique often employed in multi-modal systems, where information from different sources or modalities is combined at a later stage in the processing pipeline. Late fusion allows for greater flexibility in incorporating diverse sources of information, such as blood reports and ECG data, into the prediction model. Each modality can be processed independently, enabling the model to adapt to varying data availability scenarios.It enables the fusion of features extracted from different modalities at a higher-level representation. Instead of directly combining raw data, features extracted from each modality can be fused using fusion techniques such as concatenation, summation, or weighted combination. This allows the model to leverage complementary information from multiple modalities, potentially improving prediction performance.

6.4 CNN and Dropout approach

ECG data is inherently sequential, with time-series information representing the electrical activity of the heart. A 1D CNN is well-suited for capturing temporal patterns within ECG signals, as it can learn hierarchical representations of features across different time scales through its convolutional and pooling layers. This hierarchical representation enables the model to automatically extract relevant features from the raw ECG data, potentially capturing subtle abnormalities indicative of heart disease. Dropout layers are a technique used in neural networks to prevent overfitting, a common problem where the model performs well on the training data but poorly on new, unseen data. During training, dropout randomly ”drops out” (i.e., sets to zero) a fraction of neurons in the network, effectively making it smaller and simpler. This forces the network to learn more robust features, as it cannot rely too heavily on any single neuron. Dropout essentially acts as a form of regularization, helping the model generalize better to new data by reducing the risk of overfitting. Incorporating dropout layers in the CNN architecture helps improve model generalization by reducing the reliance on specific features and encouraging the network to learn more robust representations from the data.

$$y[i] = \sigma\left(\sum_{k=0}^{K-1} w[k] \cdot x[i+k] + b\right) \tag{6.1}$$

1D convolution function

Chapter 7

Results and Discussion

The blood report model was trained with the data and showed an accuracy of 89% on training set and 82% on testing set. The ECG processed data was passed in the convoultion network model that showed an accuracy of 88% on training set and 79% on testing set. The multimodal function which can take both the inputs together was trained on the both the data and showed an training accuracy of 91% and testing accuracy of 87%.

Method used	Only blood report	Only ECG	multimodal blood and ECG
vanilla ANN	79%	82.2%	85.7%
1D CNN + dropout	-	88.24%	84.4%
feature reduced CNN	82%	85%	87%

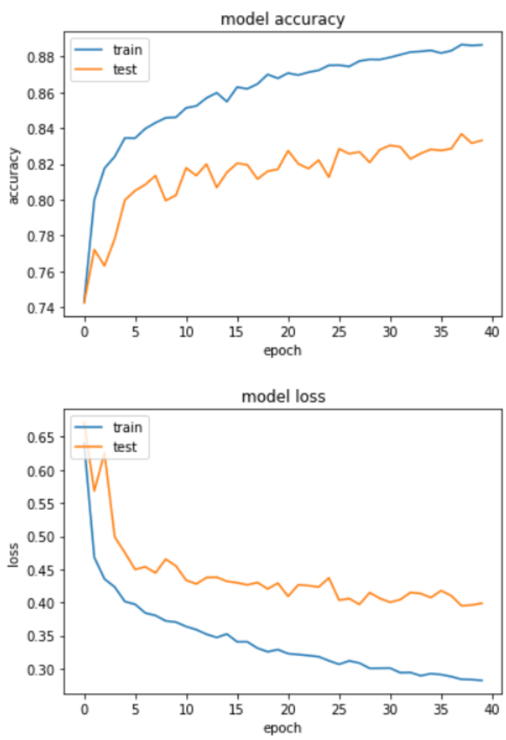

Figure 7.1: training and validation loss for CNN + dropout

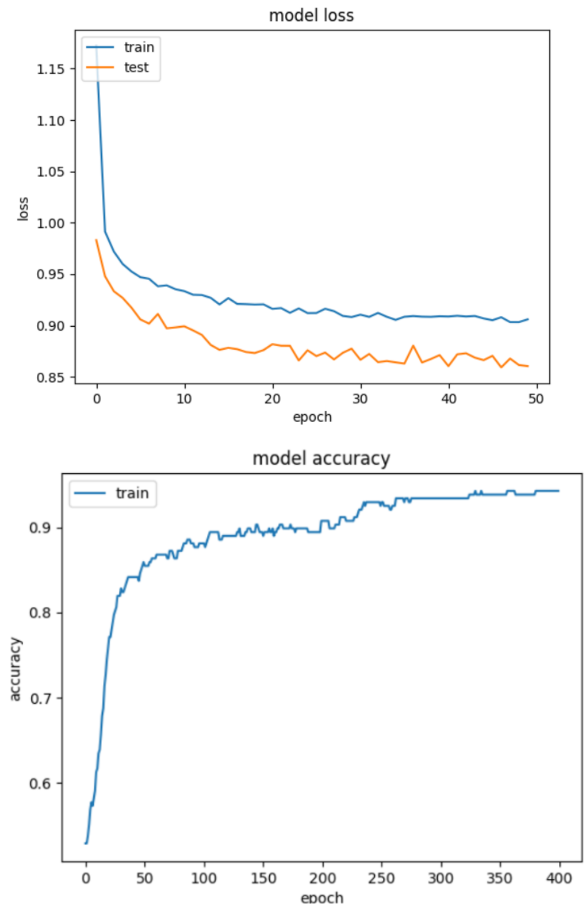

Figure 7.2: training and validation loss for feature-reduced CNN

Chapter 8

Conclusion

In conclusion, this project introduces a pioneering approach to predicting heart disease utilizing a multi-modal neural network architecture capable of processing diverse inputs. By integrating blood report data stored in .xlsx format with feature-extracted electrocardiogram (ECG) data, the model demonstrates versatility in handling various input combinations. The ability to accommodate both individual and combined inputs enhances the model's flexibility and predictive accuracy. Through experimentation and validation, the model showcases promising results, offering a valuable tool for early detection and intervention in cardiovascular health. Moving forward, this research sets the stage for further exploration into multi-modal neural network architectures and their application in advancing heart disease prediction methodologies. By using deep learning techniques and diverse data sources, this project contributes to the ongoing efforts aimed at improving healthcare outcomes and reducing the burden of heart disease worldwide.

Scope for Further Research

The scope for further research in our heart disease prediction project is vast and offers numerous opportunities for exploration and improvement.

Enhancing Model Performance: Continuously refine and optimize the multi-modal neural network architecture to improve prediction accuracy, sensitivity, and specificity. Experiment with different network architectures, activation functions, regularization techniques, and optimization algorithms to achieve better performance.
Feature Engineering: We canlook into advanced methods to create better features from blood reports and ECG data. We can also explore adding other types of data like medical images, genetic information, or patient details to improve predictions.
Data Augmentation: Augment the existing dataset with synthetic samples to address class imbalance and improve model generalization. Explore techniques such as SMOTE (Synthetic Minority Over-sampling Technique) or data augmentation for ECG signals to generate additional training samples.
Deployment: Conduct extensive validation studies to assess the model's performance in real-world clinical settings.

References

[1] Nichenametla, R., Maneesha, T., Hafeez, S., and Krishna, H., 2018. "Prediction of heart disease using machine learning algorithms". *International Journal of Engineering and Technology(UAE),* **7**, 05, pp. 363–366.

[2] Prakoso, S. D., Permanasari, A. E., and Pratama, A. R., 2023. "Heart disease prediction using machine learning: A systematic literature review". In 2023 10th International Conference on Information Technology, Computer, and Electrical Engineering (ICITACEE), pp. 155–159.

[3] Oura, S., Matsukawa, T., and Suzuki, E., 2018. "Multimodal deep neural network with image sequence features for video captioning". In 2018 International Joint Conference on Neural Networks (IJCNN), pp. 1–7.

[4] Vu, T. D., Yang, H.-J., Nguyen, V. Q., Oh, A.-R., and Kim, M.-S., 2017. "Multimodal learning using convolution neural network and sparse autoencoder". In 2017 IEEE International Conference on Big Data and Smart Computing (BigComp), pp. 309–312.

[5] Aziz, S., A. S. . A.", 2021.". In ECG-based machine-learning algorithms for heartbeat classification Sci Rep 11, 18738 (2021).

[6] Patel, T. S., Patel, D. P., Sanyal, M., and Shrivastav, P. S., 2023. "Prediction of heart disease and survivability using support vector machine and naive bayes algorithm". *bioRxiv*.

[7] Latha, C., and Jeeva, S. C., 2019. Improving the accuracy of prediction of heart disease risk based on ensemble classification techniques, informatics in medicine unlocked, vol. 16.