STM JOURNALS

International Journal of
Computer Science Language

Volume 1, Issue 1, 2023
January—June
DOI (Journal): 10.37591/IJCSL

https://journals.stmjournals.com/ijcsl

Review IJCSL

Predicting and Prohibiting the Risk of Heart Failure Using Machine Learning

Sumit Mali[1], Aashish Gaike[2,*], Sanskriti Sharma[3], Shruti Tomake[4], Aishwarya Sapre[5]

Abstract

It is challenging to estimate the likelihood of complex chronic disease while treating conditions like heart failure. The application of machine learning (ML), an area of artificial intelligence (AI), in cardiovascular care is growing quickly. In essence, it defines how computers classify and understand data, or choose a task with or without human intervention. The theoretical underpinnings of machine learning (ML) are models that accept input data (such as images or text) and forecast results using a combination of mathematical optimisation and statistical analysis (e.g., favourable, unfavourable, or neutral). Aiming to deliver results with an accuracy of 97.5%, the Heart Failure Prediction Model is based on machine learning algorithms like Random Forest and Logistic Regression. It includes sections that list the symptoms of cardiovascular diseases, offer preventative measures, and allow users to schedule doctor appointments. In order to create a prediction model based on machine learning, this document tries to capture the system needs and features. By using that, we can estimate the risk of heart disease. Heart disease sym ptoms can be reduced and understood with the help of machine learning (ML). choosing particular qualities. Additionally, the website is a subscription-based model with a unique attribute that includes several courses at various price points that are offered on the website related to their everyday activities. It consists of a yoga instructor and a nutritionist who will advise the customer on correct diet plans and exercise regimens.

Keywords: ML (Machine Learning), AI (Artificial Intelligence), Random forest, Logistic Regression, heart failure prediction models,

INTRODUCTION

According to the World Health Organisation, heart disease and heart failure are responsible for almost 12 million deaths worldwide each year. Heart failure stands out as the world's most common cause of sickness and death. In the discipline of data analysis, predicting heart failure is seen as an important topic of study. Numerous studies have been conducted to investigate the risk factors for cardiovascular disease and calculate the overall risk as a result of the rising incidence of heart failure in many different geographic areas in recent years [1]. Heart disease is also said to be the silent killer that causes death without showing any symptoms. The timely detection of heart disease holds significant significance in identifying necessary lifestyle modifications for high-risk individuals and, subsequently, mitigating potential complications. Machine learning has proven effective in making decisions and making predictions from the vast amounts of data generated by the healthcare industry. The objective of this project is to employ a machine-learning algorithm to analyze patient

***Author for Correspondence**
Aashish Gaike
E-mail: ashishgaike.nbn8802.com@gmail.com

[1]Assistant Professor, Department of Computer sciences and Engineering, NBN Sinhgad School of Engineering, Pune, Maharashtra, India
[2-5]Student, Department of Computer Sciences & Engineering, NBN Sinhgad School of Engineering, Pune, Maharashtra, India, India

Received Date: June 07, 2023
Accepted Date: June 20, 2023
Published Date:

Citation: Sumit Mali, Aashish Gaike, Sanskriti Sharma, Shruti Tomake, Aishwarya Sapre. Predicting and Prohibiting the Risk of Heart Failure Using Machine Learning. International Journal of Computer Science Language. 2023; 1(1): 15–20p.

data and predict the likelihood of future occurrences of heart disease, including the classification of individuals into groups with heart disease, heart failure risks, or without such conditions [2].

MODEL

Machine Learning Model

A mathematical representation of the results of the training process is known as a machine learning model. The study of various algorithms that may develop a model automatically through practice and historical data is known as machine learning. A machine learning model is comparable to software created for computers that can identify patterns or behaviours based on past experience or data. A machine learning (ML) model that captures the patterns found in the training data is produced by the learning algorithm after it analyses the training data for patterns [3].

Logistic Regression

A common statistical method for forecasting binary outcomes, such as the existence or absence of

A specific condition is logistic regression. Logistic regression can be used to create a model that predicts the likelihood that a patient will experience heart failure based on a variety of risk variables in the instance of heart failure prediction [4].

The steps listed below can be used to create a logistic regression model for heart failure prediction:

- *Information gathering:* Gather information on a variety of risk factors, including age, gender, smoking status, blood pressure, cholesterol levels, diabetes, and family history of heart disease. Additionally, gather information on whether or not each patient has experienced heart failure.
- *Data cleaning and preprocessing:* Remove missing values, handle outliers, scale numerical variables, and encode the data.
- *Model development:* Using the risk variables as predictors and the status of having heart failure as the outcome variable, fit a logistic regression model to the data. The projected likelihood of heart failure can be determined using the model's estimated coefficients for each predictor variable and the intercept term.
- To evaluate the effectiveness of the model, compute measures like accuracy, sensitivity, specificity, and area under the receiver operating characteristic (ROC) curve.
- *Model improvement:* If the model's performance is unsatisfactory, improve it by incorporating new predictors, eliminating old ones, modifying model parameters, or applying more sophisticated machine learning methods.

Building heart failure prediction models using logistic regression is a valuable technique, but it's crucial to remember that the accuracy of the model depends on the caliber of the data and the suitability of the modeling assumptions. When analyzing the findings of a heart failure prediction model, it is always advisable to speak with a medical expert [5].

Random Forest

Random forest is a well-known machine learning approach for classification and regression problems that can also be used to predict heart failure. On randomly picked subsets of the data, random forest constructs an ensemble of decision trees, and the final forecast is formed by aggregating the predictions of individual trees [6].

To build a random forest model for heart failure prediction, the following steps can be followed:

- *Data collection:* Collect data on various risk factors such as age, gender, smoking status, blood pressure, cholesterol levels, diabetes, family history of heart disease, etc. Also, collect data on whether or not each patient has developed heart failure.
- *Data preparation:* Clean and preprocess the data by removing missing values, handling outliers, scaling numeric variables, and encoding categorical variables.

- *Model development:* Fit a random forest model to the data using the risk factors as predictors and heart failure status as the outcome variable. The model estimates the importance of each predictor variable and the best splitting criteria for each decision tree.
- *Model evaluation:* Evaluate the performance of the model by calculating metrics such as accuracy, sensitivity, specificity, and area under the receiver operating characteristic (ROC) curve. Use techniques such as cross-validation to avoid overfitting.
- *Model refinement:* If the model's performance is not satisfactory, refine the model by adjusting hyperparameters such as the number of trees, the depth of each tree, and the size of each subset. Use feature selection techniques to remove irrelevant or redundant predictors.

Random forest is a powerful algorithm for heart failure prediction, as it can handle nonlinear interactions between risk factors and can capture complex decision boundaries. However, it is important to note that random forest is a black-box model, meaning it is not easily interpretable, and the importance of each predictor should be interpreted with caution [7]. It is always a good idea to consult with a medical professional when interpreting the results of a heart failure prediction model.

Heart failure prediction is only one of the regression and classification problems that the random forest algorithm is frequently employed for in machine learning. It works by creating a group of decision trees based on randomly selected subsets of data, and then combining the predictions of each tree to make a final prediction.

To create a random forest model for heart failure prediction, the process involves collecting data on various risk factors, such as age, gender, smoking status, blood pressure, cholesterol levels, diabetes, and family history of heart disease. The data is then preprocessed by removing missing values, dealing with outliers, scaling numeric variables, and encoding categorical variables. Next, a random forest model is fitted to the data using the risk factors as predictors and heart failure status as the outcome variable. The model estimates the importance of each predictor variable and the best splitting criteria for each decision tree. The model's performance is then evaluated by calculating metrics like accuracy, sensitivity, specificity, and the area under the receiver operating characteristic curve, with cross-validation techniques used to avoid overfitting. In the event of unsatisfactory performance, the model can be improved by fine-tuning hyperparameters, including the number of trees, tree depth, and subset size [8]. Additionally, feature selection techniques can be employed to eliminate irrelevant or redundant predictors, further refining the model. While random forest is a powerful algorithm for heart failure prediction, it is important to note that it is a black-box model, meaning it is not easily interpretable. Therefore, the importance of each predictor variable should be interpreted with caution, and it is always recommended to consult with a medical professional when interpreting the results of a heart failure prediction model [9].

SUBSCRIPTION MODEL

The major objective of the subscription-based approach is to bring all the factors that lower the risk of heart failure under one roof. User Interface and Admin Interface are the two interfaces that make up the Subscription Model. Brief explanations of both interfaces are provided in the sections below.

User Interface

The project provides both free and premium features. The prediction model for heart failure is available without charge. Dietary and nutritional courses are available through purchasing a subscription in an effort to assist people in adopting a healthy lifestyle, which eventually lowers the risk of heart failure. Dietary courses have the distinct selling feature of offering tailored training for pregnant women as well as senior citizens [10]. Access to fitness courses is also provided by subscription; these courses are tailored based on three categories. Men's courses, women's courses, and general courses are among these categories. Access to unique blogs that focus on health and discuss the advantages of living

a healthy lifestyle is also provided by subscription. In addition to all the features listed above, subscription services allow users to book online appointments with nutritionists and cardiologists.

Admin Interface

Real-time visitors, total subscribers, doctors in the database, total admins with managerial access, total courses and course category, as well as blog information, are all displayed on the admin dashboard. The admin interface provides the ability to handle all managerial tasks in a single location. The administrative interface's features include:

- All doctor visits scheduled online with BeSure can be viewed by registered admins.
- Admins have the ability to change existing records of cardiologists or dietitians, add new information about these professionals, and delete an entire record from the database.
- In addition to the aforementioned operations, admin has the ability to change the current courses, which includes creating new categories for courses and changing the existing courses inside a certain category. An current course or entire course category may also be totally deleted by the admin.
- The administrator has the option to update existing blogs or even delete them in order to inform users of the current industrial situation about heart-related difficulties.

The entire website is created using HTML, CSS, JavaScript, and PHP, and we used SQL to store all the data to keep it user-friendly and simple to grasp as shown in Figures 1–3.

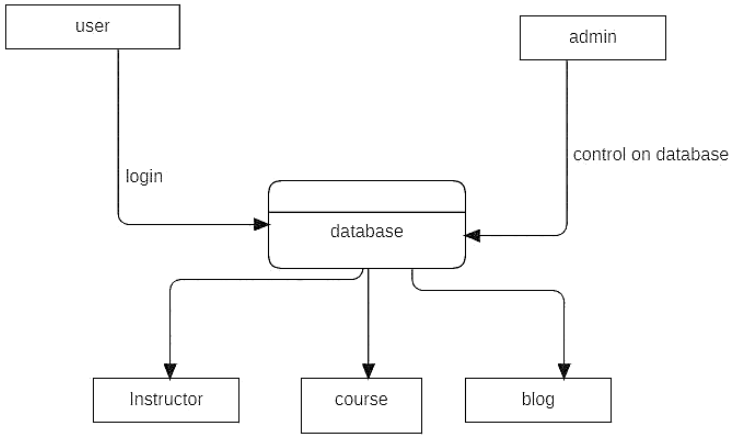

Figure 1. Database Flow diagram.

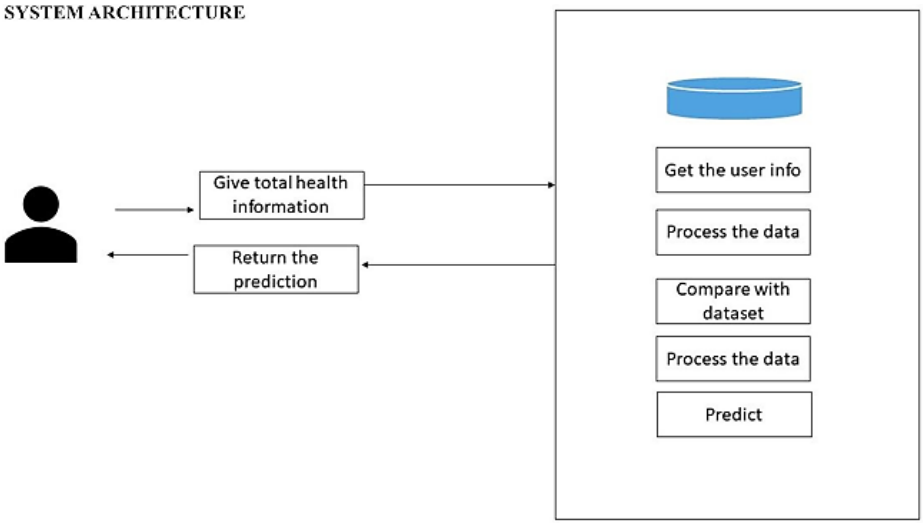

Figure 2. Health Information System Architecture.

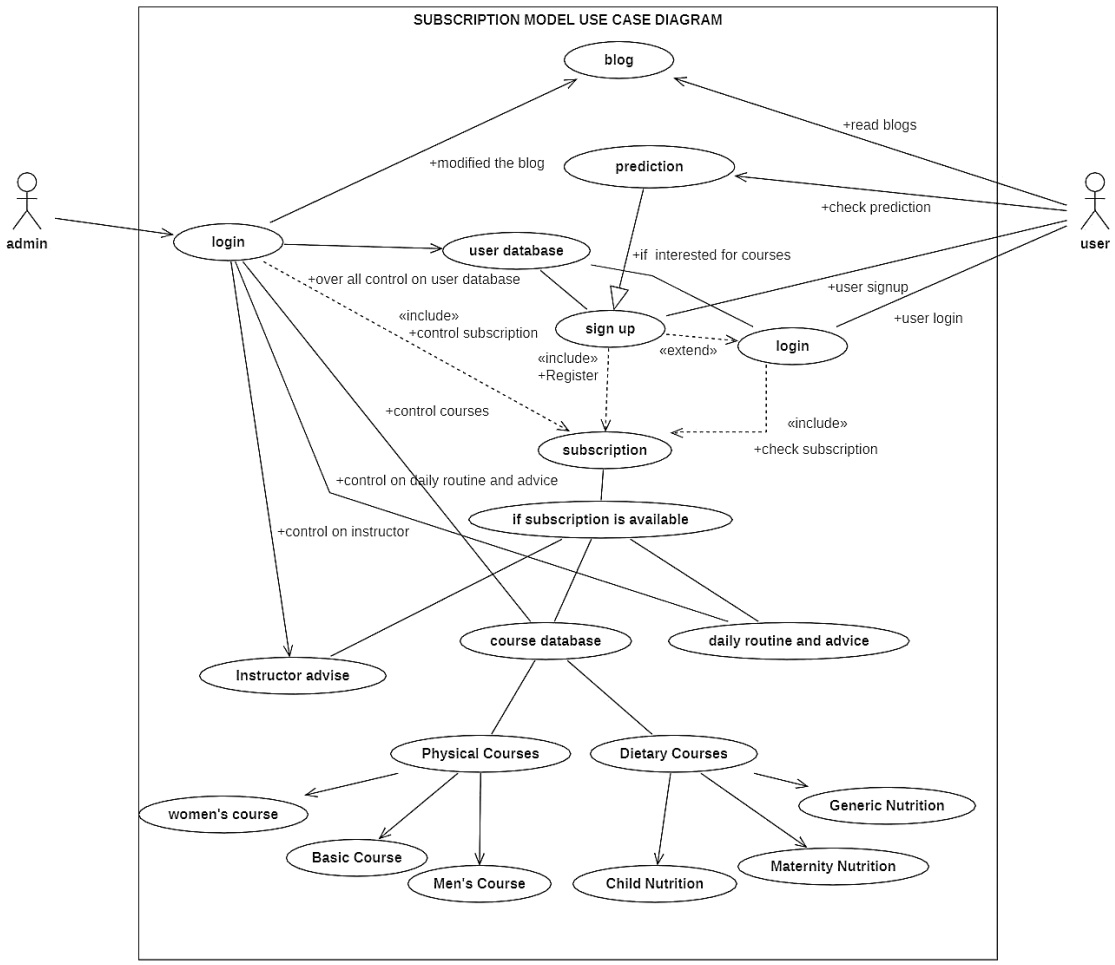

Figure 3. Subscription Model Use Case Diagram.

CONCLUSIONS

That's how we have studied clinical data of patients and got the motivation to develop this website to predict their heart health, along with that we have tried to give precautionary solutions by adding Physical exercises like Yoga, cardio and Diet plans.

Acknowledgments

It is indeed a great pleasure and moment of immense satisfaction for us to present a project on "Predicting the risk of Heart Failure using Machine Learning" amongst a wide panorama that provided us inspiring guidance and encouragement, we take the opportunity to thank those who gave us the indebted assistance. We wish to extend our cordial gratitude with profound thanks to our internal guide Prof Sumit Mali for her everlasting guidance. It was his inspiration and encouragement which helped us in completing our project. Our sincere thanks and deep gratitude to the Head of Department, Prof. Shailesh Bendale and other faculty members; but also, to all those individuals involved both directly and indirectly for their help in all aspects of the project. At last, but not least we express our sincere thanks to our Institute's Principal Dr. ShivPrasad P. Patil, for providing us with infrastructure and technical environment. Head of the department Prof. Shailesh Bendale, and Project Guide Prof. Sumit Mali.

REFERENCES

1. Singh A, Kumar R. Heart disease prediction using machine learning algorithms. In 2020 international conference on electrical and electronics engineering (ICE3) 2020 Feb 14 (pp. 452–457). IEEE.

2. Anuja Bhosale, Gayatri Gadas, et al. Detection of Phishing Websites Using Machine Learning. International Journal of Advanced Research in Computer and Communication Engineering. 2022;11(6):490–494.
3. Beck DE, Roberts PL, Rombeau JL, Stamos MJ, Wexner SD. Colon and Rectal Trauma and Rectal Foreign Bodies. 2009.
4. Singhal N, Konguvel E. An Approach for Restaurant Management System during Covid-19. In 2022 International Conference on Computer Communication and Informatics (ICCCI) 2022 Jan 25 (pp. 1–6). IEEE.
5. Chang V, Bhavani VR, Xu AQ, Hossain MA. An artificial intelligence model for heart disease detection using machine learning algorithms. Healthcare Analytics. 2022 Nov 1;2:100016.
6. Greenberg B, Brann A, Campagnari C, Adler E, Yagil A. Machine Learning Applications in Heart Failure Disease Management: Hype or Hope?. Current Treatment Options in Cardiovascular Medicine. 2021 Jun;23(6):35.
7. Ho IM, Cheong KY, Weldon A. Predicting student satisfaction of emergency remote learning in higher education during COVID-19 using machine learning techniques. Plos one. 2021 Apr 2;16(4):e0249423.
8. Wang T, Hansen KR, Loving J, Paschalidis IC, van Aggelen H, Simhon E. Predicting Antimicrobial Resistance in the Intensive Care Unit. arXiv preprint arXiv:2111.03575. 2021 Nov 5.
9. Machine Learning Optimization: Basics & 7 Essential Techniques. Aporia. 2023. Available from: https://www.aporia.com/learn/machine-learning-model/machine-learning-optimization-basics-7-essential-techniques/
10. learning machine. Random Forest. Learning Machine Learning. 2023. Available from: https://learningmachine.hashnode.dev/random-forest

