



Singaporean Journal of Scientific Research(SJSR)

An International Journal (AIJ)

Vol.17.No.1 2025,Pp.18-28

ISSN: 1205-2421

available at :[www.sjsronline.com](http://www.sjsronline.com)

Paper Received : 04-12-2024 Paper Accepted: 06-01-2025

Paper Reviewed by: 1.Prof. Cheng Yu 2. Dr.Yarab Baig

Editor : Dr. Chen Du Fensidal

## A Hybrid Machine Learning Approach for Heart Attack Prediction

**K. Hemavathi**

Asst. Professor

School of Computer Science

Takshashila University, Ongur, Tindivanam, Villupuram Dist, Tamilnadu, India.

**R. Saranya**

Asst. Professor

School of Computer Science

Takshashila University, Ongur, Tindivanam, Villupuram Dist, Tamilnadu, India.

*Abstract: A heart attack also known as cardiac arrest, diversify various conditions impacting the heart and became one of the chief-reason for death worldwide over the last few decades. Approximately, 31% of total deaths globally are due to CVDs. It constitutes the pinnacle of chronic processes which involve complex interactions between risk factors which can and cannot be improved. Most of the instances or cases of cardiovascular diseases can be allocated to revisable risk factors where most of the instances are considered preventable. ML became the enhancing approach for the evolution of predictive models in health care industries and was decided to test various algorithms to check what extent their prediction scores estimate or ameliorate upon the results acquired. Researchers deploy various machine learning and data mining techniques over a set of enormous data of cardiovascular patients to attain the prediction for heart attacks before their occurrence for helping healthcare industries and professionals. This research comprises various Supervised ML classifiers like, Gradient Boosting, Decision Tree, Random Forest and Logistic Regression that have been used to deploy a model for Myocardial Infarction prediction. It uses the existing datasets from the Framingham database and others from the database of the UCI Heart repository.*

*This research intends to ideate the prediction for probabilities of occurrence of a heart attack in the patients. These classifiers have been deployed in pipeline approach of machine learning to attain the prediction using both ways i.e., without optimizations and feature transformations as well as vice-versa. The results impersonate that the Gradient Boosting classifier is achieving the highest accuracy score in such a way that prediction used by our model is of binary form in where 1 means a chance of heart attack and 0 means no chance. Some of the most influential attributes are chest pain type among which the typical angina is the most influential and asymptotic chest pain is least, cholesterol level in which the level greater than 200mg/dl are more prone, increased heart rate, thal, and age. It is concluded that premature heart attack is preventable in 80% of the total cases just by using a healthy diet along with regular exercises and not using tobacco products also the person who drinks more than 5 glasses of water daily are less likely to develop attacks. The medical checkup of Blood-pressure level, cholesterol level and heart rate on daily basis along with meditation can help you prevent the major heart attacks.*

**Keywords:** *Cardiovascular-Disease, Framingham Model, Gradient Boosting, Machine Learning, Mayo-cardinal Infarction, UCI Model.*

## 1. Introduction

Heart, the main organ of the human body used for pumping blood into the whole body through the vessels of the circulatory system. In the circulatory system, the most crucial role is played by the heart [1]. The circulatory system is the most important part of our body as it is responsible for the transport of blood carrying food, oxygen, water, minerals, and other important substance important for our body throughout our entire body. If the working of the heart is disrupted due to any circumstance and it does not function properly then it may cause serious health issues including death.

Cardiovascular is a term that is used to refer to the pathologies altering or affecting the function or structure of the heart or blood vessel having the most common type of cardiovascular disease as coronary artery disease. The prevalence of the most common cardiovascular diseases (CVDs) represents the pinnacle of incurable processes which involve complex interactions between risk factors which can and cannot be improved. Most of the instances or cases of cardiovascular diseases can be allocated to improvable risk factors where most of the instances are considered preventable.

Traditional efforts for preventing cardiovascular diseases have focused on modifiable behaviors governed by individuals [2]. The use of physical inactivity, tobacco, a poor diet, and obesity majorly contribute to known risk factors for cardiovascular diseases, such as mellitus, diabetes, hypertension, and the development of atherosclerosis. As per the current scenario approximately 17.9 million people every year lose their lives due to cardiovascular diseases (CVDs), apparent mainly as attacks and strokes. The above scenario of death is 31% of total deaths globally [3]. Heart-diseases or cardiovascular diseases (CVD) are the type of diseases that involve the heart and blood vessels.

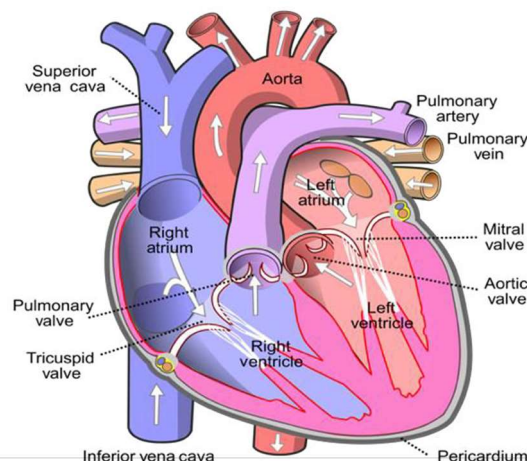
### A. Coronary Artery Disease

In coronary arteries, blockage formation due to fat, cholesterol also called plaques. These plaques on being damaged release platelets which cause blood to clot. They accumulate around the plaque and hence restrict the blood flow, which eventually damages the heart muscle [4] [5]. This damage may become more substantial if the blood flow to any segment of the heart is completely blocked [6]. Coronary artery disease is the tapering or blockage of the coronary arteries which is usually caused by atherosclerosis which builds up plaques in the artery-blocking blood flow. The most serious complication of coronary artery disease is myocardial infarction which is best known as a heart attack [7]. This disease at its primary stage is a preventable condition but preventing its quick and early treatment, it can lead to serious health problems and in the worst case even death [8].

Environmental agents also have an important role in the growth and ferocity of cardiovascular diseases although the individual often has very little control over them. The cardiovascular system are very much prone to a variety of different environmental agents which includes solvents, tobacco, pesticides, smoke, and other inhaled or ingested pollutants, as well as extremes in noise and temperature [9]. Exposure to environmental pollutants occurs via three general routes: inhalation, ingestion, and absorption. Condemnation to environmental agents mainly occurs via three general routes which are inhalation, ingestion, and absorption. In 2014, the World Health Organization proclaimed air pollution as the deadliest single environmental health risk. In 2012, more than 3.5 million deaths in people who were aged >60 years were recognized to outdoor air pollution out of which 80% of these deaths were the consequences of cardiovascular diseases [10].

## B. Heart Attack

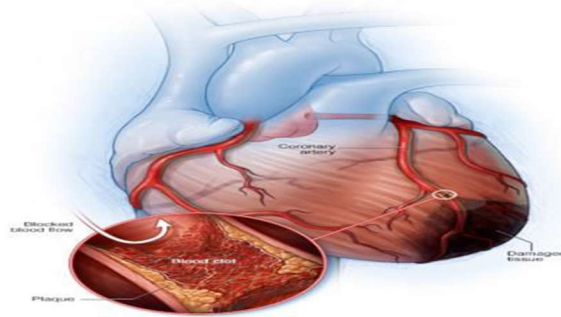
Heart Attack consequences from one of these CV diseases. STEMI, an attack, which occurs due to atherosclerosis, restricting blood flow to a wide area of the heart. This led to continuous damage to the heart muscle due to which the functioning of the heart is completely stopped and may cause death. This attack is severe and needs rapid attention. NSTEMI, an attack that occurs due to partial blockage of coronary arteries, restricting blood flow severely. Although being less dangerous, it led to permanent damage of a restricted part of the heart which is not receiving the blood flow. Coronary artery spasm, the silent heart attacks. Sometimes due to the contraction of arteries connected to the heart, blood flow to major parts of the heart is restricted



respect to other attacks, it never causes any permanent damage to the heart [11][12].

#### - Symptoms

Major symptoms of the occurrence of heart attack are tightfistedness or affliction majorly in the chest, neck, back, and arms, tiredness, dizziness, abnormal heartbeat, and consternation. Risk factors including unchangeable factors like age, sex, family background, and changeable factors like smoking, high cholesterol, high blood pressure, fatness, deficiency of proper diet as well as exercise, and a huge amount of stress. Arterial reclamation to medication, ECG, and bypass surgery is the most used treatment methods in case of heart attacks [13][14].



**Figure 2: Blockage in Coronary Artery [5]**

#### - Factors causing Heart Attacks

The comestible factors that assist or protect against the growth of coronary heart diseases ensuing to the arrival of atherosclerosis. The primary cause or initiator of atherosclerosis is unrevealed, despite one conjecture is now acquiring attention that the primary event is free radical damage to cholesterol in circulating low density lipoproteins and it is possible to acknowledge a seven dietary factor having two promoter and five protective feature for developing coronary heart disease [15]. Myocardial infarction, most serious complication of coronary heart disease, represents an amalgamation of having two distinct effects of dietary factors. There are evidences that lowering serum homocysteine containing folic acid reduces the risk of cardiovascular diseases is largely observational. Drug Treatment for preventing the cardiovascular disease events and heart strokes has been limited to single risk factors which targets the small amount of patients' population with values in the appendage of the risk factor distribution, and to reducing the risk factors to average population values [16].

### C. Artificial Intelligence

Artificial intelligence can be termed as the proficiency of a computer program to mimic the human brain by developing intelligence. It gives computers the ability to think which makes them more intelligent. Alan Mathison Turing first proposed the concept of artificial

intelligence more than sixty years ago. According to him if a human cannot distinguish between the human and computer based on the answers or solutions or responses given by a machine then the machine is called as “intelligent”.

Its importance’s can be felt by witnessing the last two decades which showed the advances of the artificial intelligence and robotics and the future can be expected to be more speculated. The term artificial intelligence was first coined in 1956. AI is based on the principle that the machine tries to mimic and execute tasks which are simplest to those that are even more complex by using human intelligence.

#### - **Machine Learning**

Machine Learning can be called as a subdivision of Artificial Intelligence in which the program learns through data or input. It primarily focused on two principles i.e., in what way a computer system be developed so that it automatically improves the results through its previous experiences and the fundamental statistical computational-information-theoretic laws those regulating all learning systems, including humans, computers and organizations. It was designed to solve a specific problem by developing an intelligent system that is able to solve that problem without being categorically programmed for it. The learning of the program is done by deriving knowledge from a large amount of data which is useful in making predictions [20].

#### **D. Classification**

In Machine Learning, Classification is a function which uses ML algorithms to learn for assigning class label to cases from problem set. Classification is described as predictive modeling of any problem with a labelled class to be predicted for a given dataset. For example, to determine whether a human is male or female, to determine whether a mail is spam or not, to classify any given handwritten character i.e., whether it is known or unknown character. Supervised learning algorithms are the most basic machine learning algorithm which is trained on labeled data. This algorithm is extremely robust when used in the right circumstances. It usually performs the prediction tasks since the goal to achieve is to either foresee or classify the target from a certain outcome of interest i.e., availability or unavailability of the heart problems.

#### - **Decision Tree**

A **Decision tree** is a non-parametric supervised machine learning algorithm used for solving classification and regression problems which uses decision support tool that uses a flowchart-like graph structures in making decision. In decision tree, each internal node refers a test on a feature taken via the pathways along which the branches describing the classification rules to attain the final result (i.e., class label) defined by each leaf node. The decision tree looks like, learning by splitting the dataset into subsets based on value test which is carried on an attribute. This process is continued on every extracted subset recursively which is handled until the subset at a node have the same value of the target variable, or adds no value to the prediction when splitting.

#### - **Logistic Regression**

**Logistic Regression** is a parametric supervised machine learning algorithm which is a linear model used for solving classification problem which predicts outcome of a categorical

dependent variable from the set of predictor or independent variables. It is a very simple method and much more efficient for binary and linear classification problem where response variable tends to be binary having continuous explanatory variables despite having regression on its name. This model can be generalized further for multi-class classification problems. It models the liability for classification problems with two possible outcomes for target. It uses logistic equations for finding the results between 0 and 1. Logistic function (or, sigmoid function) is defined as:

- **Gradient Boosting Classifier**

**Gradient Boosting Classifier** is one of the Ensemble learning techniques. The different types of gradient boosting classifier algorithms have been increased with several interesting proposals naming its different types. Here, boosting means integrating multiple weak learning algorithms with lower accuracy to a stronger model predicting with higher accuracy. It follows mistakes done by predecessor's algorithms for maintain the upcoming algorithms to attain higher accuracy. In Gradient Boosting, each predictor tries to overcome the mistakes done by its predecessor's algorithms. It always fits a new predictor for overcoming the errors done by its predecessor's algorithms.

- **Random Forest**

**Random Forrest** is a supervised learning algorithm, as its name look alike, it is an integration of large number of individual decision trees created during training time which functions as Ensemble learning (can be termed as nearest neighbor predictor) used to solve classification as well as regression problem. In the crowd of such decision trees, the individual tree whose class prediction is upvoted most or the class having the largest mode value becomes the prediction of model. In other words, initially it develops many individual decision trees and then integrates them to obtain a higher accuracy and better prediction. It finds a natural balance between the two extreme problems of high variance along with high bias by averaging them.

## E. Problem

According to the WHO (World Health Organization) the 32% of all global deaths were due to cardiovascular diseases out of which, approximately 85% deaths were due to heart attacks. These surveys also represent that approximately 48% of US population have any kind of heart disease and around every 19 deaths one person is due to heart problems. According to the National Crime Records Bureau (NCRB) in India, the percentage of deaths due to heart attacks is increased by approximately 53% in last five years and approximately 33% of deaths in India were due to heart attacks and other problems. Although in many cases the patient may be saved by doctor even after minor or major attacks, but also in most of heart attack cases patient gets died due to unavailability of medical treatment under minimal time due to the incapability of patient to inform anyone during the attack period and also for their family members as being unknown from this kind of emergency condition to any of their relatives. And due to which, this kind of sudden attacks may lead to the death of patient even before consulting to any doctor. So, we cannot prefer consulting doctors after having attacks as the perfect solution for precaution from occurrence of heart attacks.

## F. Approaches

Hence, to overcome and reduce the chances of death, AI and ML comes in role to predict the chances of upcoming heart attacks according to the health conditions as well as previous medical conditions of the patient. Although there are numerous ml-models have been already developed but most of the models are prepared for heart disease prediction system instead of being specialized for heart attacks only. Many other models have been deployed using various algorithms and over various datasets along which many are simpler whereas many are complex. In this project we applied various ML-algorithms over two datasets to take a survey and to achieve a better accuracy to deploy a model for predicting the upcoming heart attack scenario earlier than the actual attack time to prevent the patient from going to a critical condition which can lead him/her to death.

## G. Objective

For the both datasets, the risk factors provided in it associated with heart attacks for the patients to find out whether they have any risk of cardiovascular problems in upcoming years or not.

Based on the datasets provided:

- To predict the probability of suffering of any patient from heart attacks in the upcoming future.
- Identifying the major factors influencing heart attacks.
- Identifying cholesterol level with higher possibilities of heart attacks.

## 2. Related Works

As per initial prospects, ML can enhance the evolution of predictive models in health care industries, it was decided to test various algorithms to check what extent their prediction scores estimate or ameliorate upon the results acquired in the authentic Framingham model. This model is one of the most important cardiac arrest risk prediction tables from the point of view of clinical practice. Several CV-risk prediction algorithms have been proposed till now. The performance of various algorithms that have been used for calculating CV-risk can still be considered as a problem. Particularly, different scores like, Framingham Score, Systematic Coronary Risk Evaluation tend to underrate cardiac risk in patients. Comparative study has been done for CV risk prediction using some well-known ML algorithms like k-nearest neighbor, support vector machine, classification, gradient boosting, logistic regression and regression tree and random forest[3]. Among the comparison between various features selection ML algorithms, for heart attack prediction Naive Bayes, SVM, and KNN the most optimistic classifiers.

Another approach used for the cardiac infarction risk prediction involves partitioning the dataset in a random way and deploying various old data mining approaches like J48, REPTREE, Naïve Bayes, Bayes Net, and CART. The deployed model was able to answer more complex queries in cases of cardiac infarction prediction. The research was carried out in February 2021 to make a model which uses an algorithm by leveraging random under sampling, clustering, and oversampling techniques which were better known as under

sampling-clustering-oversampling algorithm (shortly, UCO algorithm). The difference in this algorithm was that the data that was used for training on machine learning algorithms were nearly very balanced data. This algorithm was very good in extracting features which were then tested on different classifiers giving the best predicting performance with an accuracy of 70.29%, precision of 70.05%, 1-Recall of 75.59%, 0-Recall of 63.95% of the random forest. There was research carried out for predicting the heart attack early with taking an account of chest pain with 24 other attributes. The data mining algorithms used were decision tree and random forest classifier to analyze the heart attack datasets. The deep classification was done by a decision tree algorithm where the random forest was used for the classification of targets.

There was another technique used for the prediction of heart disease risk which involves partitioning the dataset randomly using mean based splitting approach into smaller subsets and then using homogeneous ensemble created using various classification and regression trees models using an accuracy based weighted aging classifier ensemble.

### 3. Datasets and Pre-Processing

Dataset, generally said to be as a collection of data organised for specific purposes. Here in this research, we have used two different datasets. The first one dataset we used is Framingham dataset, which is published by collaboration of Boston University, National Institute of Heart (NIH), National Heart Lung and Blood Institute (NHLBI). This dataset is completely dedicated for identifying factors for identifying cardiovascular diseases, especially attacks and strokes. And the second one dataset we used is Heart dataset from UCI Machine Learning Repository (University of California, Irvine, School of Information and Computer Sciences). This dataset is a combination of four different databases from “Hungarian Institute of Cardiology, Budapest”, “University Hospital, Zurich, Switzerland”, “University Hospital, Basel, Switzerland”, “V.A. Medical Center, Long Beach and Cleveland Clinic Foundation”. Both the above datasets are also available on *Kaggle Website*.

#### A. Framingham Dataset

Attributes of the Framingham dataset are classified in four types like Demographics, Behavioral, Previous medical history-based information, Current medical condition-based information.

Demographic Attributes:

- Sex: Categorised as either 0 or 1 such as 0 = female and 1 = male
- Age: Age of the patient at the current time of examination
- Education: It is an inessential data, because any medical issue doesn't occur as per someone's education level Behavioral:

Behavioral:

- Current Smoker: Categorised as either 0 or 1 depending upon whether a patient smoke currently or not i.e., 1 = yes and 0 = no

#### Preprocessing

Data Pre-Processing is defined as transforming or encoding the data in such a state so that



is can be easily parsed by the machines for generating accurate information. In other words, it should be transformed in such a form so that is can be easily interpreted by different algorithms with producing higher accurate results. It is not necessary to be complete pure data in each and every dataset. There is always some missing data in each and every dataset in “NULL” form due to which the dataset becomes redundant and hence leads the models to predict results with poor accuracies. Hence, to overcome these poor accuracies and to attain higher and better accuracies, data pre-processing came in genre. We usually clean the tuples having missing values by either dropping those tuples from dataset or by imputing mean or median values of respective column or some other hyperparametric optimization to attain the imputable values for replacing those missing values. Since the both datasets used in our model consist numeric data only.

Hence, in our proposed model we are using mean and median imputation approaches for imputing missing values in data set for attaining its consistency to achieve higher accuracy. Mean imputation is the way of replacing missing values (i.e., ‘NA’ or ‘NULL’) data in dataset by mean of that parameter. And median imputation is the way of replacing missing values (i.e., ‘NA’ or ‘NULL’) data in dataset by median of that parameter. Even in this mean and median imputation there is always a confusion that when we should use mean imputation and when we should use median imputation. It can be described as whenever the parameter represents a normal distribution then we can use any one of both mean and median imputation. But if the parameter represents skewed distribution instead of normal distribution, then the median imputation is preferred over mean imputation.

#### 4. Methodology

In our proposed model we are using two major datasets of the field of Cardiovascular analysis. The first one dataset is Framingham dataset and the other one is UCI heart dataset. We are using four ML classifiers which are Logistic Regression, Decision Tree Classifier, Random Forest Classifier, and Gradient Boosting Classifier. Here, we are using iterative approach for deploying our model over both the datasets. Both the first datasets are loaded iteratively and checked for the missing values and further pre-processed by imputation of the missing values Then dataset is split into two parts, one is training data and other is holdout data for testing purpose. The ratio of splitting of dataset into training and holdout part is 9:1 i.e., 90% of dataset will be used for training purpose of model whereas remaining 10% will be used for testing purpose of model. not number text heads-the template will do that for you.

##### 4.1 Work Flow of Proposed Model

After splitting and pre-processing, we are iteratively deploying each classifier over the dataset for their respective predictions and accuracies attained. For each classifier, we are generating four pipelines by optimizing and enhancing some features from previous pipeline. Pipeline in machine learning is an approach to code and for automating its workflow to build the model. Out of total sixteen pipelines, each classifier has four pipelines in order, first without any optimization then, second with hyperparameters optimization (i.e., marked as “**hpo-1**”) then, third with previous hyperparameters optimization and feature engineering (i.e., marked as “**hpo-1 + fe**”)

and at last, with previous hyperparameters optimization, feature engineering and final updating again with hyperparameters optimization (i.e., marked as “**hpo-1 + fe + hpo-2**”). After these four iterations on pipelines of each classifier, the best pipeline of each classifier is selected. And then the final four pipelines, best one pipeline of each classifier are analyzed together as per their performance and accuracies. And then the final best of all sixteen pipelines is selected to deploy the final model for further predictions in future. The complete workflow designed for this study as flowchart.

## 5. Conclusion

In this research, various Supervised ML classifiers namely, Random Forest, Decision Tree, Gradient Boosting, and Logistic Regression have been used to deploy a model for Myocardial Infarction prediction. In spite of having many inconsistencies in both the datasets, various feature transformers have been used to improve the consistencies of datasets and to attain an average accuracy equal to 85.5% and a recall rate of 82% in Framingham dataset based on Gradient Boosting classifier, and more enhancing the recall rate reaches to 89.1% when training of Gradient Boosting classifier over UCI Heart dataset. After these all approach our final model is deployed for cardiac arrest prediction using Gradient Boosting Classifier. We can additionally diversify this research integrating semi-supervised and deep learning techniques.

The results impersonate that the Gradient Boosting classifier is achieving the highest accuracy score in such a way that prediction used by our model is of binary form in where 1 means a chance of heart attack and 0 means no chance. Some of the most influential attributes are chest pain type among which the typical angina is the most influential and asymptotic chest pain is least, cholesterol level in which the level greater than 200mg/dl are more prone, increased heart rate, thal, and age. It is concluded that premature heart attack is preventable in 80% of the total cases just by using a healthy diet along with regular exercises and not using tobacco products also the person who drinks more than 5 glasses of water daily are less likely to develop attacks. The medical checkup of Blood-pressure level, cholesterol level.

## Reference

1. H. Animesh, K. M. Subrata, G. Amit, M. Arkomita, and A. Mukherje, “Heart Disease Diagnosis and Prediction Using Machine Learning,” *Heart Disease Diagnosis and Prediction Using Machine Learning and Data Mining Techniques: A Review.* Advances in Computational Sciences and Technology 10(7): 2137–59. ht,” *Adv. Comput. Sci. Technol.*, vol. 10, no. 7, pp. 2137–2159, 2017, [Online]. Available: <http://www.ripublication.com>.
2. H. S. Buttar, T. Li, and N. Ravi, “Prevention of cardiovascular diseases: Role of exercise, dietary interventions, obesity and smoking cessation,” *Exp. Clin. Cardiol.*, vol. 10, no. 4, pp. 229–249, 2005.
3. I. D. Mienye, Y. Sun, and Z. Wang, “An improved ensemble learning approach for the prediction of heart disease risk,” *Informatics Med. Unlocked*, vol. 20, p. 100402, Jan. 2020, doi: 10.1016/J.IMU.2020.100402.

4. M. Hortmann *et al.*, “The mitochondria-targeting peptide elamipretide diminishes circulating HtrA2 in ST-segment elevation myocardial infarction,” *Eur. Hear. J. Acute Cardiovasc. Care*, vol. 8, no. 8, pp. 695–702, 2019, doi: 10.1177/2048872617710789.
5. Mankad R; Staff MayoClinics, “Heart attack,” *MayoClinic*, 2020. <https://www.mayoclinic.org/diseases-conditions/heart-attack/symptoms-causes/syc-20373106>.
6. P. Severino *et al.*, “Ischemic heart disease pathophysiology paradigms overview: From plaque activation to microvascular dysfunction,” *Int. J. Mol. Sci.*, vol. 21, no. 21, pp. 1–30, 2020, doi: 10.3390/ijms21218118.
7. A. Segura-Galindo, F. Javier Del Cañizo-Gómez, I. Martín-Timón, C. Sevillano-Collantes, and F. Javier Del Cañizo Gómez, “Type 2 diabetes and cardiovascular disease: Have all risk factors the same strength?,” 2014, doi: 10.4239/wjd.v5.i4.444.
8. P. B. Lockhart and Y.-P. Sun, “Diseases of the Cardiovascular System,” in *Burket’s Oral Medicine*, John Wiley & Sons, Ltd, 2021, pp. 505–552.
9. T. Mü nzel, M. R. Miller, M. Sørensen, J. Lelieveld, A. Daiber, and S. Rajagopalan, “Reduction of environmental pollutants for prevention of cardiovascular disease: it’s time to act,” doi: 10.1093/eurheartj/ehaa745.
10. M. Ferrante *et al.*, “Air Pollution in High-Risk Sites—Risk Analysis and Health Impact,” in *Current Air Quality Issues*, InTech, 2015.
11. A. W. R. N. Kandola, “Types of heart attack: What you need to know,” *Medical News Today*, 2018. <https://www.medicalnewstoday.com/articles/321699>.
12. H. Yasue, Y. Mizuno, and E. Harada, “Coronary artery spasm—Clinical features, pathogenesis and treatment—,” *Proc. Japan Acad. Ser. B Phys. Biol. Sci.*, vol. 95, no. 2, pp. 53–66, 2019, doi: 10.2183/pjab.95.005.
13. G. D. Sandler, David A and Aspenson, D Erik and Johnsen, “Oklahoma Heart Institute,” *Citeseer*, vol. 2, no. 1, 2005.
14. R. Fass and S. R. Achem, “Noncardiac chest pain: Epidemiology, natural course and pathogenesis,” *J. Neurogastroenterol. Motil.*, vol. 17, no. 2, pp. 110–123, 2011, doi: 10.5056/jnm.2011.17.2.110.
15. M. S. Ellulu, I. Patimah, H. Khaza’ ai, A. Rahmat, Y. Abed, and F. Ali, “Atherosclerotic cardiovascular disease: a review of initiators and protective factors,” *Inflammopharmacology*, vol. 24, no. 1, pp. 1–10, 2016, doi: 10.1007/s10787-015-0255-y.
16. R. Hajar, “Risk factors for coronary artery disease: Historical perspectives,” *Hear. Views*, vol. 18, no. 3, p. 109, 2017, doi: 10.4103/heartviews.heartviews\_106\_17.
17. J. A. Perez, F. Deligianni, D. Ravi, and G.-Z. Yang, “Artificial Intelligence and Robotics,” pp. 1–56, 2018, [Online]. Available: <https://arxiv.org/ftp/arxiv/papers/1803/1803.10813.pdf>.
18. M. Fatima and M. Pasha, “Survey of Machine Learning Algorithms for Disease Diagnostic,” *J. Intell. Learn. Syst. Appl.*, vol. 09, no. 01, pp. 1–16, 2017, doi: 10.4236/jilsa.2017.91001.
19. M. I. Jordan and T. M. Mitchell, “Machine learning: Trends, perspectives, and prospects,” vol. 349, no. 6245, 2015.