REPUBLIC OF TÜRKİYE

ALTINBAŞ UNIVERSITY

Institute of Graduate Studies
Electrical and Computer Engineering

A NOVEL INTELLIGENT MACHINE LEARNING SYSTEM FOR CORONARY HEART DISEASE DIAGNOSIS

Haedar ALSAFI

Doctor of Philosophy

Supervisor

Prof. Dr. Osman Nuri UÇAN

Istanbul, 2022

A NOVEL INTELLIGENT MACHINE LEARNING SYSTEM FOR CORONARY HEART DISEASE DIAGNOSIS

Haedar ALSAFI

Electrical and Computer Engineering

Doctor of Philosophy

ALTINBAŞ UNIVERSITY

2022

The thesis titled A NOVEL INTELLIGENT MACHINE LEARNING SYSTEM FOR CORONARY HEART DISEASE DIAGNOSIS prepared by HAEDAR EMAD SHAREF AL-SAFI and submitted on 19/12/2022 has been **accepted unanimously** for the degree of PhD in Electrical and Computer Engineering.

Prof. Dr. Osman Nuri UÇAN

the Supervisor

Thesis Defense Committee Members:

Prof. Dr. Osman Nuri UÇAN	Department of Electrical and Electronic, Altınbaş University	_____
Asst. Prof. Dr. Sefer KURNAZ	Department of Computer Engineering, Altınbaş University	_____
Asst. Prof. Dr. Dogu Çagdas ATILLA	Department of Electrical and Electronic, Altınbaş University	_____
Prof. Dr. Hasan Hüseyin BALIK	Department of Computer Engineering, Yıldız Technical University	_____
Assoc. Prof. Dr. Adil Deniz DURU	Department of Sports and Health Sciences, Marmara University	_____

I hereby declare that this thesis meets all format and submission requirements of a PhD thesis.

Submission date of the thesis to Institute of Graduate Studies: ___/___/___

I hereby declare that all data presented in this graduation project has been obtained in full accordance with academic rules and ethical conduct. I further affirm that, in accordance with the aforementioned standards of academic integrity, all borrowed ideas, arguments, and conclusions have been properly referenced in the text, as have all sources listed in the Reference List.

Haedar ALSAFI

DEDICATION

For the love of Allah, my Maker and my Master; in honor of my wonderful advisor, Prof. Dr. Osman Nuri Uçan (May Allah bless and give him); and in gratitude to Iraq, my motherland.

PREFACE

First and foremost, I want to give my deep appreciation to my advisor, Prof. Dr. Osman Nuri Uçan, for his unwavering encouragement, inspiration, and guidance during my PhD studies and research. All through the process of researching and writing this thesis, his advice was invaluable. The guidance and support I received from my adviser and mentor during my PhD studies was invaluable.

ABSTRACT

A NOVEL INTELLIGENT MACHINE LEARNING SYSTEM FOR CORONARY HEART DISEASE DIAGNOSIS

Alsafi, Haedar

Ph.D., Electrical and Computer Engineering, Altınbaş University,

Supervisor: Prof. Dr. Osman Nuri UÇAN

Date: December / 2022

Pages: 67

As one of the most prevalent forms of heart disease, coronary heart disease (CHD) is a major health concern. Since heart illness can strike at any time, a thorough screening process is essential. In this study, we explore a recent approach to CHD diagnosis that makes use of classifier ensembles, a technique for optimizing machine learning. To improve our framework's efficiency, we employed the Feature-Selector optimization model to pick the optimal subset of CHD traits. To address the issue of unbalanced CHD data, we next employed optimal SMOTE over-sampling, a very effective approach using an optimization model. The classmark estimate from three different optimization learners (random forest, XGBoost API optimization, and the SVM optimization model) is integrated in a stacked architecture. Those with coronary heart disease are utilized to put the recognition model through its paces. Finally, when compared to existing detection models based on optimization model ensembles and individual classifiers, ours exhibited superior accuracy, F1, and ROC-Curve. We achieved 91% accuracy using random forest optimization, and 90 percent

accuracy using the XGBoost API optimization model. In comparison to previous studies published in the literature, this study shows that our proposed model has a significant effect.

Keywords: SMOTE Optimization, CHD Disease, Python, API function, Optimization Machine learning models, Feature Selector optimization.

TABLE OF CONTENTS

Pages

LIST OF TABLES

Pages

LIST OF FIGURES

Pages

ABBREVIATIONS

ACM	:	All-Cause Mortality
AI	:	Artificial intelligence
ANN	:	Artificial Neural Network
BMI	:	Body Mass Index
CAG	:	Comptroller and Auditor General
CHAID	:	Chi-squared Automated Interaction Ietection
CHD	:	Coronary Heart Disease
CV	:	Cross-Validation
CVD	:	Cardiovascular Dysfunction
DM	:	Different Methods
DT	:	Decision Tree
GS	:	Grid Search
HBP	:	High Blood Pressure
IHD	:	Ischemic Heart Disease
KNN	:	Kernel Nearest Neighbor
ML	:	Machine Learning
NB	:	Naive Bayes
PCA	:	Principal component analysis
RBF	:	Radial Base Functions
RF	:	Random Forest
ROC	:	Receiver Operating Characteristic
RTs	:	Random Trees

SMOTE : Synthetic Minority Over-sampling Strategy

SOM : System On Modul

SVM : Support Victor Machine

TARs : Tethered Aerostat Radar System

XGBoost : Xtreme Gradient Boosting

VEB : Ventricular Ectopic Beat

SVEB : Standard Ventricular Ectopic Beat

AAMI : Association for the Advancement of Medical Instrumentation

SaDE : Selfadaptive Differential Evolution

QRS : Qwave Rwave Swave

ECG : Electrocardiogram

LIST OF SYMBOLS

μ : Electron mobility

λ : Wavelength

ω : Angular frequency

φ : Wave Function

∈ : Element of

∑ : Summation

1. INTRODUCTION

The main cause of death in the USA is coronary heart diseases (CHD), which affects roughly 13% of the population. It is impossible for stress the importance of early heart attack identification in preventing cardiac arrests and other medical complications. Machine learning techniques can sometimes be referred to as "data mining." It's been almost fifty years since any of these techniques were initially discovered. [1], improvements in algorithmic programming and the processing capacity of current computers have helped fuel a recent explosion in both interest in and understanding of these systems. Clinical settings have maintained the use of these tools and therapies throughout the past three years in diagnostic radiography, cardiac electrophysiology, coronary heart disease, dermatology, and psychology. [2]. Heart disease is a leading killer worldwide, as reported by the World Health Organization. Arrhythmia is a form of sickness that causes the heart to beat irregularly. A typical machine-learning algorithm will first transform a given job into an optimization issue, and then apply one of several optimization methods to the problem at hand. Several hyper parameters, which are determined in front of the learning process, make up the optimization function and affect how machine learning functions. The goal, before starting the training phase, is to determine a range of hyper parameter values that yields the optimal impacts on the data in an acceptable length of time. Hyper parameter tweaking or optimization describes this process. It's a must-have for making accurate predictions using machine learning. However, hyper parameters cannot be directly correlated with the results methods for computer vision. Through practice, it is feasible to adjust hyper parameters regularly and prepare a range of models with various value mixes, then compare models outcomes to determine the best one from model.[3]. That's why it's so important to figure out how to tweak the hyper parameters of a machine learning algorithm. The two primary categories enhancement of hyper parameters techniques are the manual investigation and the automated search. A manual search is a testing technique that is done by hand. hyper parameter collections. It is based on the principles, the wisdom, and experience of seasoned customers who can identify the crucial variables that have the most influence on the outcomes and then use modeling approaches to analyze the relationship between these variables and the outcome. Manual scanning requires a higher level of technical and practical knowledge on the part of the customer [4]. Non-expert consumers frequently find it difficult to apply. Tuning system parameters is a difficult task to recur. Furthermore, number of

Because people struggle to deal with complicated data and are prone to misinterpreting or misinterpreting forms and interactions in Hyper parameters, the range of potential values widens, making it harder to control. As a result, machine learning has long struggled to find a solution to the problem of obtaining the robotic tuning algorithm to attain high accuracy and dependability. [5]. A hyper parameter tuning trouble is one in which the optimization goal function is unknown or undefined. The Newton process and gradient descent are ineffective traditional optimization methods. In this approach of optimization, Bayesian optimization is a pretty successful optimization procedure. The Bayesian formula is used to combine previous knowledge about the unknown parameter with sample data in order to get posterior knowledge about the function distribution. Then, using this posterior information, we can know where the function gets its best value. Recent technological advancements also accelerated the growth and availability of raw data, allowing for rapid advancements in research and technology [6]. This has provided a huge gap for knowledge discovery and computer engineering science to play a critical role in a broad variety of applications, ranging from everyday civilian life to national security, from business information management to federal decision-making support structures, and from micro scale data exploration to macro-scale knowledge discovery. The topic of imbalanced learning has piqued the attention of academics, business, and government funding agencies in recent years. The central problem with the imbalanced learning dilemma is that imbalanced data has the capacity to dramatically degrade the efficiency of most normal learning algorithms [7]. Most regular algorithms presume or predict equivalent misclassification costs or balanced class distributions. A person with an arrhythmia has a heart that cannot efficiently pump blood throughout the body. The brain, the heart, and other organs are all susceptible to injury from insufficient blood supply. Patients with these conditions have a much better chance of avoiding unexpected death if they receive a prompt diagnosis and aggressive medical treatment. If cardiac arrhythmia is caught early, it is easily treated. The absence of symptoms in early stages of disease allows for a straightforward diagnosis. Cardiac arrhythmia is a common and frustrating condition, and its symptoms are comparable to those of other heart conditions, thus it's clear that an intelligent system is needed to accurately identify this condition. Utilize a neural network and an evolutionary algorithm together to make a more accurate diagnosis of the causes of arrhythmic heart disease. Numerous factors support the neural network in this investigation. For its physical (natural) applications, artificial neural networks' straightforward architecture and high class-classification accuracy make them ideal candidates. Artificial neural

networks are characterized primarily by attributing results to entry vectors that were not present during network training. Due to the vast number of variables and contextual elements impacting the current state of arrhythmia detection, evolutionary algorithms are employed. The diagnosis of this disease is made using a suitable classifier trained on relevant data. The utilization of neural network architecture and evolutionary algorithms is employed for this aim. A neural network is one of the first aspects to be established. In the next step, evolutionary The neural network's weights are calibrated using several techniques. Neural networks using evolutionary algorithms are useful for identifying arrhythmic heart illness and achieving optimal treatments. The goal of study is to identify the causes of arrhythmic heart illness and use that information to provide an accurate diagnosis as soon as possible. As a consequence, when confronted with complex imbalanced data sets, these algorithms struggle to adequately reflect the data's distributive features, resulting in unfavorable accuracies throughout data groups. As extended to real-world domains, the imbalanced learning the dilemma is a recurrent topic of high interest and far-reaching consequences that needs further study [8]. In contrast, hyper parameter optimization methods seek to define the optimal hyper-parameter configurations for a machine learning algorithm's architecture. This research attempts to evaluate the efficiency and time benefits that innovative optimized ML approaches will have over a professionally established procedure by employing one to solve the aforementioned challenging challenge. This is how the rest of the essay is organized: Initiatives like this are discussed in Part 2. In Section 3, we delved into the conceptualization of unique structured machine learning approaches, the development of an optimized (SMOTE) algorithm, and the development of an optimal collection of functions. In Section 4, we discuss the experiments and their results. Section 5 provides a summary and a synthesis of the work done thus far.

1.2 MACHINE LEARNING

Machine learning (ML) is a set of methods, techniques, and programs used to address diagnostic and prognostic issues in a variety of healthcare settings. The relevance of clinical parameters and their differences in disease progression is being explored using machine learning techniques. These include condition growth assessment, medical information extraction for outcome testing, therapy preparation and assistance, and overall patient care.[9]. Machine learning is often used to analyze data in a number of areas, including finding trends in data by adequately managing lost data, evaluating continuous data utilized in the Intensive Care Unit, and advising troubleshooting, which

helps in accurate and reliable tracking. It is proposed that effective application of ML approaches will aid in the introduction of computer-based applications in the healthcare context, allowing medical experts' practice to be facilitated and enhanced, thereby improving the reliability and consistency of medical treatment [10]. We've compiled a list of some of the most important machine learning technologies in medicine. Since datasets are incomplete (missing parameter values), inaccurate (systematic or spontaneous noise in the details), sparse (few and/or non-representable patient records available), and inexact, learning from patient data poses many challenges (inappropriate selection of parameters for the given mission). These characteristics are common in health datasets, and machine learning provides tools for dealing with them. [11]. As the healthcare system becomes more reliant on computational technologies, it stands to reason that machine learning tools will provide useful assistance to doctors in a variety of ways, including the elimination of issues related to human fatigue and habituation, the provision of rapid detection of anomalies, and the facilitation of real-time diagnosis.

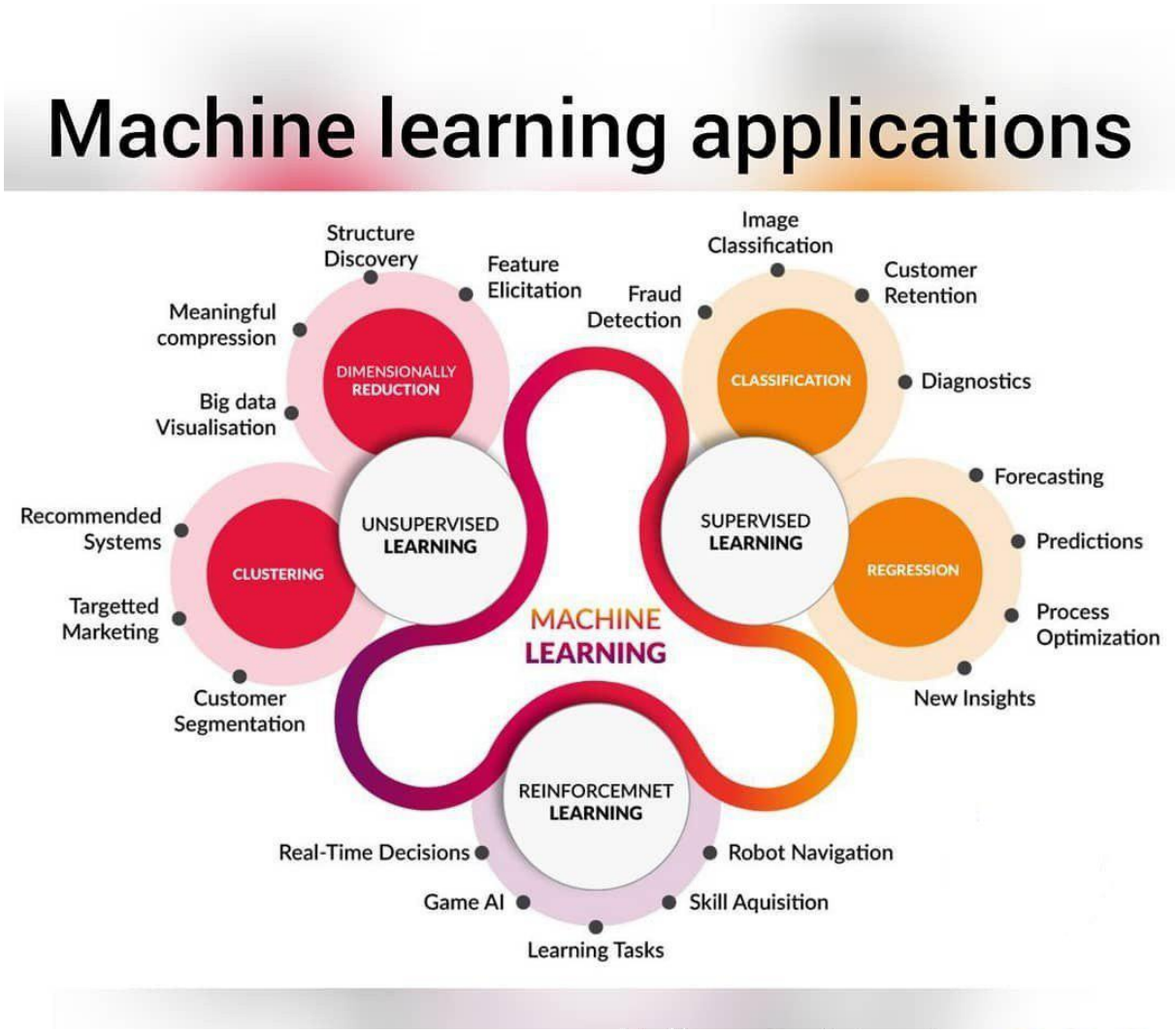

Figure 1. 1: Machine learning application [12].

1.3 DEFINITIONS

1.3.1 Coronary Heart Disease:

Is the constriction of the heart's auxiliary arteries, which reduces blood flow and oxygen levels. Coronary heart disease is sometimes referred to as coronary artery problem.

1.3.2 Blood Pressure: Hypertension occurs when blood pressure against blood vessel walls is consistently high.

1.3.3 Cholesterol: is a waxy substance that may be detected in blood; it is produced mostly in the liver but can also be derived from meals like red meat, butter, and eggs.

1.3.4 Data Mining: is the process of isolating useful data from massive datasets in order to address issues by identifying patterns and connections.

1.4 PROBLEM STATEMENT

Medical diagnosis is an inherently difficult process that necessitates extreme accuracy when taking into account a variety of variables. Furthermore, given the amount of experience and information needed for reliable results, CHD prediction is a far more difficult task. In technical terms, every data collection with an uneven distribution of its groups is said to be imbalanced. However, the general consensus in the field is that imbalanced data refers to data sets with severe, if not serious, imbalances. We provide an illustration from biomedical applications to illustrate the real-world consequences of the imbalanced learning issue. Consider the CHD dataset, which is a list of attributes gleaned from medical exams conducted on a group of patients and has been commonly utilized in the evaluation of algorithms that solve the imbalanced learning issue. Learn the machine is techniques allow to the application of intelligent optimization methods to a variety of datasets in order to uncover valuable insights. The reprogrammable capacity of machine learning to explore, process, and analyze datasets allows it to appeal to decision-makers in fields like medical diagnosis. Since detecting CHD necessitates training a model based on historical data, machine learning appears to be a suitable technology to address this problem.

1.5 AIM OF THESIS

The key goal for this research is to learn how to cope with CHD by applying data mining classification methods to minimize the number of time doctors spend diagnosing patients and to

reduce costs. In order to determine the right approach to cope with the lack of data in medical databases, the method employs several data mining methods in its classification processes.

The study's theory is that inaccuracy in data derived from medical sources may be a significant cause of errors during coronary heart disease diagnosis. The patient's misinformation to the physician, as well as the lack of and inaccuracy of facts, are some of the factors that contribute to the disease's progression.

Coronary heart failure cannot be entirely minimized or removed without addressing the underlying causes of the disease. Furthermore, several commentators believe that 40 percent of the recent downturn in coronary heart disease is attributed to better care rather than a reduction in disease incidence.

1.6 THESIS CONTRIBUTION

The Proposes of this study a novel intelligent ML diagnosis model to identify the different types of (CHD) diseases using a medical dataset. We will enhance different types of classification models to achieve better results and the best model for disease diagnosis problems like CHD.

• To begin, we may use powerful pre-processing techniques, To cope with missing values in coronary CHD datasets, one oversampling technique is the Synthetic Minority Over-sampling Strategy (SMOTE).

• To accelerate training and ensure high classification accuracy, we chose a newly improved extraction of features model (Feature-Selector) as the most effective method for selecting relevant input characteristics.

• After carefully adjusting the sampling size and utilizing cutting-edge methods for choosing pertinent features, we are currently testing the use of robust random forest, robust SVM tuning, robust XGBoost API models, and robust hyper-parameter tuning to develop a novel intelligent CHD disease diagnosis framework.

• To ensure that the outcomes of our innovative approach are comparable to those of other existing approaches, we employ a wide range of machine learning and computer programming approaches. For CHD data sets, we have substantial experimental results.

2. LITERATURE REVIEW

2.1 INTRODUCTION

In most countries, coronary heart disorders (CHDs) constitute the main cause of avoidable death, according to a global survey of scientific, technical, and medical journals. This highlights the need for a CHD diagnostic paradigm or framework to be developed in order to cut down on management costs. Therefore, prophylactic action is crucial. Data mining, AI, and machine learning are only some of the cutting-edge approaches presented by several researchers for predicting and diagnosing CHD. Information is extracted from CHD using a variety of Techniques for data mining, also including artificial neural network, Bayesian theory, decision trees, and genetic algorithms.

2.2 CORONARY HEART DISEASE

Coronary heart disease is caused by problems with the heart's major arteries that carry blood and oxygen (CHD). The most prevalent causes of cardiac failure include ischemia, coronary disease, and cardiovascular dysfunction of various sorts (IHD). Both men and women of all ages die most frequently from cardiovascular disease and coronary heart disease in many nations, but mainly the United States and China.[13]. CHDs are normally caused by plaque deposits in the heart arteries. The type CHD and its impact on the arteries of the heart are illustrated in the diagram below.

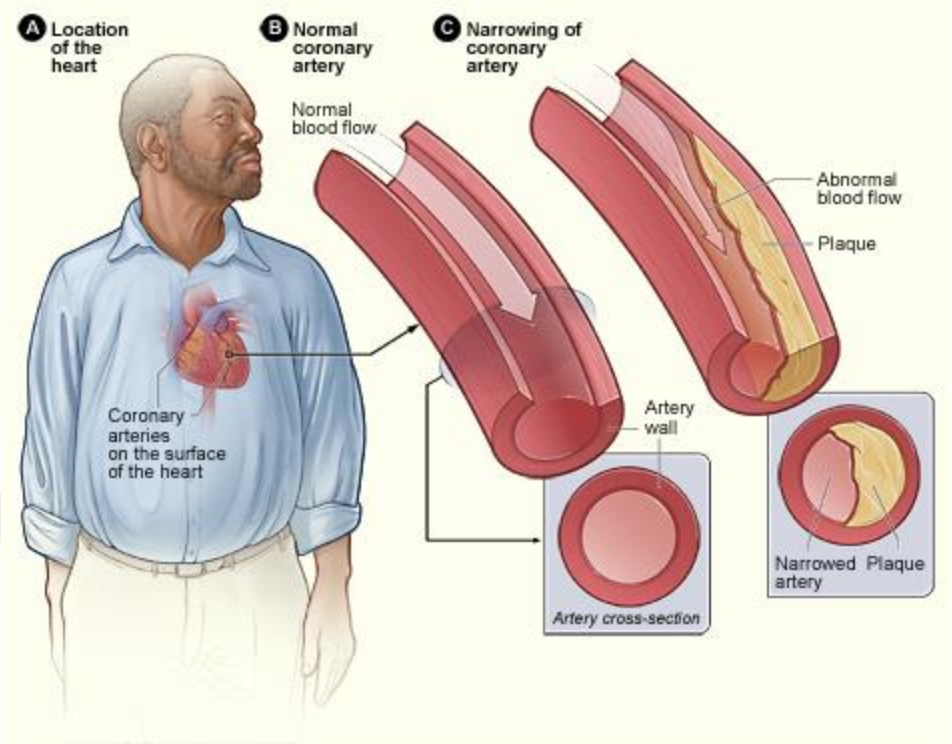

Figure 2. 1: Coronary heart disease (CHD) [14].

The coronary arteries get progressively more constricted and the blood supply to the heart is cut off when plaque accumulates there. This decline could result in chest pressure, shortness of breath, and other CHD symptoms [15]. Full blockages may result in a heart attack, with obesity, caused mostly by not moving about enough, eating poorly, and maintaining a healthy weight are all forms of inactivity, hypertension, and excessive cholesterol.

2.3 BACKGROUND OF CHD

The community is particularly interested in and worried about CHD diagnosis and research due of the tremendous human loss, particularly in European nations. The availability of accurate medical identification and appropriate medication administration is a significant concern for healthcare organizations. Poor professional choices have substantial and unacceptably negative consequences. More sophisticated technology, such as a computer-based information system, is required to assist clinical decision-making in the healthcare environment [16]. Research in machine learning (ML) may be used to medical datasets to uncover previously concealed information and insights using a variety of physical and categorization methods. With this information at their disposal, medical professionals may make more informed clinical decisions,

As a result, patient care improves, patients' life satisfaction improves, and more accurate diagnoses are made. Various researchers have developed a number of systems to assist experts and physicians in detecting heart failure. Basic strategies are more general, such as NB, DT, and KNN. Data mining-based classification algorithms used in CHD diagnosis and data mining techniques include kernel density, sequences lowest enhancement, bagging algorithms, direct kernel self-organizing graphs, neural networks, and SVM. Data mining classification algorithms are used to aid in the diagnostic process.

2.4 LITERATURE STUDIES

Kalia Orphanoua, Arianna Dagliati et al [17]. In this paper, the authors describe a Naive Bayes classification model that takes into account laws of temporal connectivity (TARs). Using TARs, which show how often a certain temporal pattern emerges in a patient's log, three classifiers are developed and compared. A longitudinal dataset is used to include all available classifiers into a final diagnosis (CHD). A common kind of heart dysfunction, arrhythmia causes abnormal heart rhythms. Due to the lack of early symptoms, cardiac arrhythmias are frequently detected when they have already progressed to a severe state. In order to effectively treat this disease, an early diagnosis is essential. This calls for a smart system capable of detecting cardiac arrhythmias in their earliest stages. The main cause of death in the United States is heart disease, making it the most pressing health concern. In contrast, heart disease mortality can be reduced with early detection and treatment.

Enrique Puertas, Juan-Jose Beunza et al [18]. This study aims to assess the internal consistency and predictive efficacy of many supervised ML algorithms for clinical event prediction. Two computational processing systems were used to gather data, and then the results were compared. SVM (AUC = 0.75) is the most reliable algorithm.

Delila Mehanovi, Zerina Maeti, and Dino Keo [19] created a model for predicting heart disease utilizing ANN, KNN, and SVM algorithms. Each of these algorithms has been employed by a number of different researchers. Both multi-class and binary classification benefit from majority voting, with the former yielding an accuracy of 61.16 percent and the latter of 87.37 percent.

Kranthi K. Kolli et al [20]. Repeat after me: "Donghee Hana, Donghee Hana, Donghee Hana, Donghee Hana, Donghee Hana, Donghee Hana" They evaluated the accuracy of several risk

prediction strategies with that of a machine learning-driven prediction algorithm based on information gathered from yearly health checkups to see whether or not it was possible to accurately forecast all-cause mortality (ACM). For reclassifying patients from low to intermediate risk, ML technique fared better than the other methods. The decision tree that has been suggested takes into account the length of time that passes between heartbeats, as well as the presence or absence of an irregular cardiac rhythm and the pattern of the CHD. In this research, we employ multiple heart rates and investigate how decision tree classification can be used to classify arrhythmic heart rates. It takes a lot of effort and is prone to mistakes when trying to manually extract the proposed features for CHD. Since this is the case, this research presents a smart decision tree classification approach for CHD data, and the outcomes demonstrate the model's excellent accuracy. Cardiac arrhythmia is an irregular activity of the heart that has been studied. In this study, stochastic forest classification was used to classify the heart problem.

Bayu Adhi Tama et al [21]. This work examines classifier ensembles, a modern CHD recognition system based on machine learning. Using a two-tier ensemble is possible. For another ensemble classifier, certain ensemble classifiers act as the basic classifier. Three ensemble learners' predictions are combined in a layered design: Gradient boosting machine, random tree, and heavy gradient boosting. Feature extraction using discrete wavelet transform is a linear operator that breaks down signals into existing components (wavelets) in the band of different frequencies. In this method, attribute values can be divided into two different subsets. In this case, the decision branch will be different in two branches. According to the characteristics of each branch, the examined property will be placed in that branch. Random forest is a classification technique based on classification trees.

Bijoy K. Menon et al., Shakiru A. Alaka1 et al [22]. They suggested that functional outcomes following a stroke might be predicted using patient estimates of their stroke-related functional risk. It is being researched if machine learning algorithms can estimate long-term functional outcomes following a stroke.

Hossein Ahmadi, Mehrbakhsh Nilashi, et al [23]. The aim of this study is to create a predictive algorithm for heart disease prediction using ML algorithms. Unsupervised and guided learning approaches are used to create the proposed paradigm. For missing meaning imputation, two

imputation methods are used in this study: PCA, SOM (Fuzzy SVM), and PCA (Principal Factor Analysis) (PCA).

Javad Hassannataj Joloudari and colleagues et al [24]. They primary objective of this study is to improve the precision of coronary heart disease diagnoses by establishing a hierarchy of the most important predicting indicators. In this research, we offer a new, more effective machine learning method. Decision trees based on C5.0, RTs, SVMs, and CHAID (Chi-squared Automated Interaction Detection) were all employed in this investigation. In order to detect an irregular heartbeat, an electrocardiographic signal is analyzed. Professionals in the medical field are crucial to the process of identifying arrhythmic forms. It is not only time-consuming to analyze the arrhythmias, but there is also a chance of making a mistake in the diagnosis. Intelligent systems are required to address such issues. Parameters for heart arrhythmias include Premature ventricular contraction, atrial fibrillation, failure to convert to normal sinus rhythm, and a rapid heart rate are all symptoms of heart blockages (left, right, or both). The objective of this study is to determine which characteristics are crucial for making an arrhythmia diagnosis. Fuzzy logic relies on a set of input variables, most of which are carefully chosen properties. The database uses index numbers to categorize the rhythm of each heartbeat recorded in its many data files. Approximately 95.42 percent of cardiac arrhythmias can be detected with the suggested approach.

Lewlyn L. R. Rodrigues et al [25]. The use of machine learning techniques in this work is in the realm of medicine. The purpose of this study was to investigate, using machine learning, the impact of age, BMI, frequent cigarette use, weekly alcohol intake, blood pressure, and systolic blood pressure on hypertension, CHD, and simulated outcomes. The results show that the proposed system has a reliable capacity to classify electrocardiogram signals and help physicians to accurately diagnose heart disorders. The use of random forest classification is more efficient than decision-making methods for classifying the electrocardiogram signal and has better results. If your heart rate is erratic, you may have cardiac arrhythmia. It's possible that this is merely a brief lull in the action, and that it has no effect on the heart's regular rhythm, or it could cause the heart to beat abnormally rapidly or slowly. Here, we present an artificial neural network-based technique for automatically distinguishing between these two types of arrhythmias. An intelligent system with two steps for diagnosis of cardiac arrhythmia has been proposed. In the first step, the Fisher score is used to select the attribute and reduce the dimensions of the data set attribute space.

Second, a subset of the features chosen in the first step are employed in conjunction with the least squares of the help vector machine to make a diagnosis of cardiac arrhythmias.

Jikuo Wang et al [26]. In this research, they presented a two-stage stacking-based paradigm, with level 1 serving as the base level and level 2 as the meta-level. The projections of the base-level classifiers are used to choose the meta-level feedback. To find the classifier with the lowest correlation, the person correlation coefficient and maximum knowledge coefficient are determined first. And, using the enumeration algorithm, the best mixing classifiers are found, yielding the best answer. We use the Z-Alizadeh Sani CHD list, which has 303 cases that have been confirmed by CAG. Cardiac arrhythmias are difficult to diagnose and require the attention of a trained professional. As a result, the concept of using automation to identify cardiac arrhythmias emerged. A fresh approach to identifying and categorizing cardiac arrhythmias using the integration of violet and neural networks.

J. Jeyaranjani et al [27]. They developed an effective ANN model for CHD detection that seeks to enhance existing classification through the use of an efficient feature selection procedure. This improves the prediction model's accuracy while lowering the error rate. Among all supervised learning algorithms, the ANN is regarded as the classification problem model that performs the best. One of the most observed techniques for choosing a characteristic is the Fisher score. This method's initial step involves calculating the Fisher score for each property. A subset of attributes with a higher Fisher score is selected. Support vector machines are supervised learning that has a structure based on the principle of minimization. As a consequence, this study utilizes a NN model to intelligently forecast the class of cardiac disease based on the most effective characteristics. Output is calculated to show the proposed model's correctness. Precision, sensitivity, reliability, consistency, and error rate are the metrics for performance analysis[28].

We focused our research on developing cutting-edge optimization machine learning strategies to aid in the implementation of a revolutionary idea that could be distinct and more efficient in terms of outcome and accuracy than prior approaches for dealing with unbalanced CHD large data sets. The sickness is primarily categorized as either normal or abnormal in this investigation. The signals from electrocardiograms were utilized to train and evaluate three distinct neural network models. Some information may be absent while examining arrhythmic data. Whenever a data attribute is missing, the next best match for that class is substituted for it. When training models

for artificial neural networks to detect cardiac arrhythmias, Learning the law of motion is paired with the inverse algorithm. An artificial neural network consists of connections of artificial neurons that study the target function. After pre-processing, feature extraction is performed and then arrhythmia is diagnosed using artificial neural networks.

2.5 MODEL PARAMETER TUNING

The testing approach optimizes (or "tunes") model parameters by running data through the model's operations, comparing the resulting approximation to the real value for each data event, evaluating the precision, and updating until the correct values are found. By running the whole training job, calculating the overall precision, and making adjustments, hyper parameters are fine-tuned.[29]. In both instances, you're changing the model's structure in an attempt to find the right solution to your dilemma. You'll have to make physical changes to the hyper parameters over the course of several training runs if you don't use an advanced technology like AI Platform Training hyper parameter tuning. The method of deciding the right hyper parameter settings is made smoother and less tedious with hyper parameter tuning[30].

2.5.1 How Hyper-Parameter Tuning Works

Several trials are conducted in a single training effort to do hyper parameter tuning. With the hyper parameter values set to the upper and lower bounds you choose, each trial is a complete run of our training application. [31]. The AI Platform Consulting training service maintains track of each trial's outcomes and adjusts for potential trials. When the job is done, we can get a rundown of all the trials as well as the most successful value setup depending on the parameters you determine. For hyper parameters to be tuned, there must be clear communication between your training program and the AI platform training provider. The training application contains a list of all the data that your model needs [32]. You must first specify the hyper parameters (variables) you want to modify, as well as a goal value for each of them. When Hyper parameter tuning against similar models, AI Platform Training will progress over time and render hyper parameter tuning more effective by modifying solitary the objective feature or introducing a different input support.

2.5.2 Hyper-Parameter Tuning Optimizes

We assign a particular target variable, also known as the hyper-parameter metric, and hyper parameter tuning optimizes it. A typical metric is the model's accuracy, as measured from an inspection transfer [33]. You should decide whether you want to tune the formula to optimize or

reduce your metric, and the metric must be a numeric number. We call the hyper-parameter metric while we first start a job with Hyper-parameter tuning. This is the name you give the scalar description you employ in your training software.

2.6 SUPPORT VICTOR MACHINE ALGORITHM

Since 1963, V. N. Vapnik and Alexy Y. Chervonenkis have been industrializing the SVM supervised methodology in its original form. It is used in the nonlinear classification method by adding Bernhard E. Boser and Isabelle M. Guyon's Kernel Trick to Maximum-Margin Hyper Planes. It's often used in a variety of fields because it's so good at binary classification. It is, however, mostly used for binary data classification problems [34]. We may claim that SVM is light and ideal for use with healthcare datasets. The definition of decision thresholds, which describe decision limits, underpins the SVM ranking. A category of objects with distinct class memberships is separated at the judgment stage. For the best separation of classes, SVM searches for vectors (or "vector support") that define the commas. SVM offers options for both binary and multiclass objectives. In linear support vector machines, it is necessary to determine both the highest and lowest hyperplane margins that differentiate the training dataset [35]. If it is possible to separate the training data, two hyperplanes are selected that will allow for the separation of the data. The difference between them is regarded as a margin and there are no points between them. The general SVM classification strategy is illustrated in the illustration below.

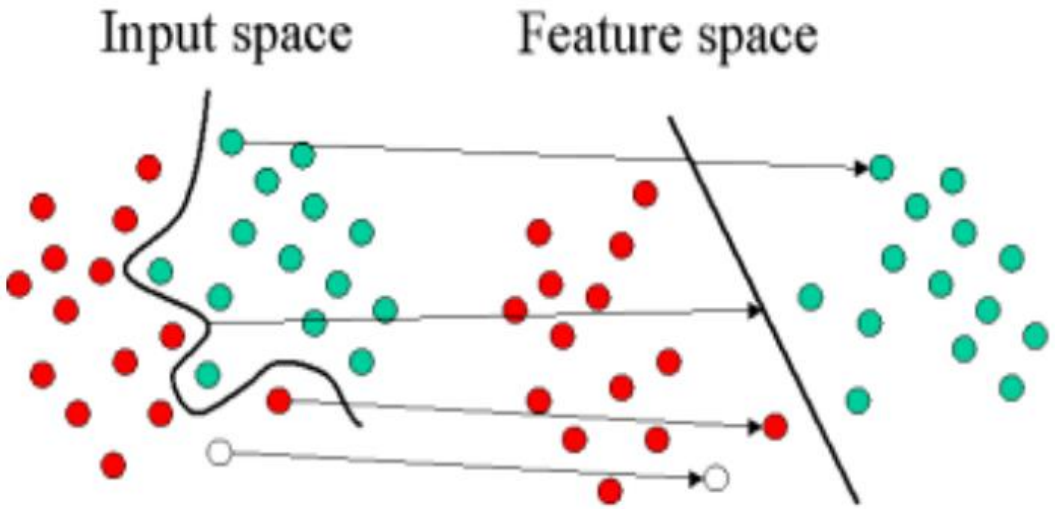

Figure 2. 2: Form of the SVM classifier [36].

The benefits of SVM include pliability in choosing a kernel type and robustness with a limited number of datasets, as well as the freedom to do what we like with the right kernel. When the mechanism is uncertain, it works well. Furthermore, When there are more measurements than there are useable samples, the SVM algorithm determines which alternative is preferable. The drawbacks of SVM include its poor performance and slowness during the testing process, as well as its high difficulty and need for a significant amount of memory for quadratic programs. In addition, SVM can be computationally costly.

2.7 RANDOM FOREST CLASSIFIER

A supervised classification technique, the Random Forest data mining algorithm is described. This classifier, considering its apparently self-explanatory name, functions to construct a forest and randomize it. In this approach, each decision tree serves as a specialist in a different subset of the training results, and the nodes of each tree are divided according to a different randomly selected feature from the dataset. The method will establish models that are unrelated to one another if a training dataset is entered with an endpoint and features in DT, indicating that you prepared a collection of laws. [37]. Predictions maybe created using these laws. Phase one of the Random Forest algorithm involves randomly selecting f features from m features where km, phase two involves computing node d by finding a better split point among the f features, and phase three

involves splitting the node into child nodes. At last, it attempts to grow a forest by cycling through the steps n times to produce n leaves. To further understand how the random forest classifier works, consider the diagram below.

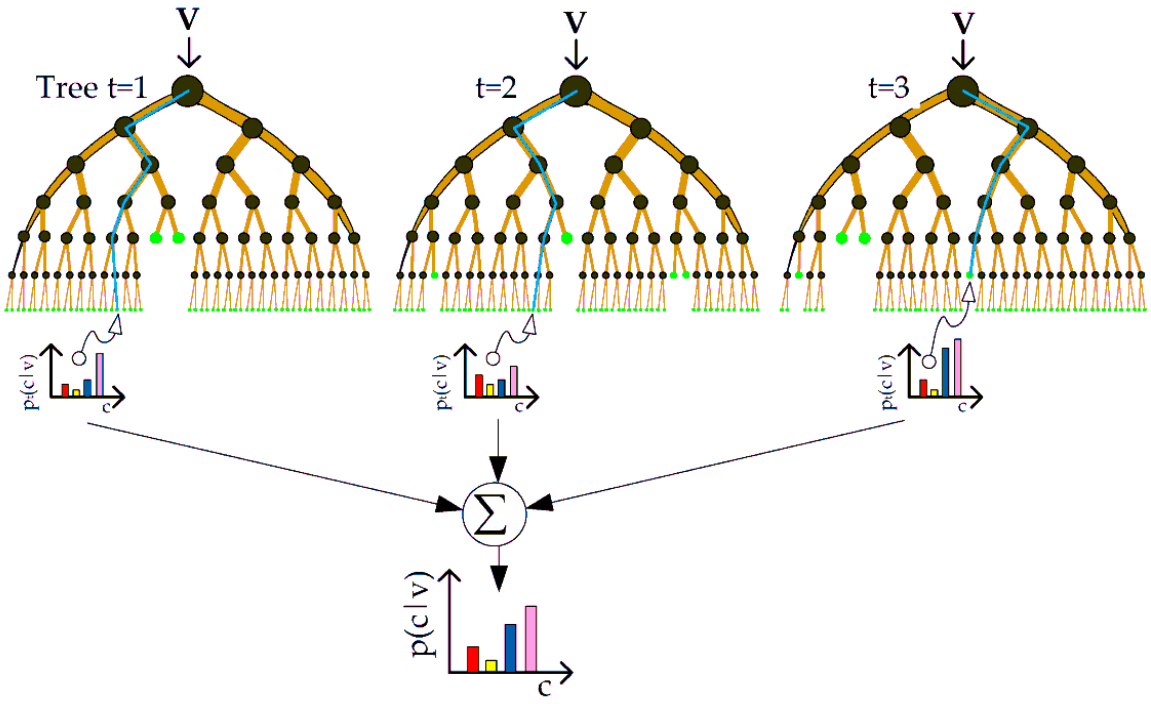

Figure 2. 3: Working form of the random forest [38].

The only way to know for sure if our kid would enjoy animated movies is to watch them with him and build a list of all the ones we've enjoyed over the years so we can use it as a benchmark against which to judge his potential interest. Then rules can be generated using a decision tree algorithm. The features of this film will then be entered to see whether it would cater to the kid. The formation of rules and the estimation method of these nodes are used to gather knowledge [39]. The procedure of the Random Forest is more complicated than that of the Decision Tree, which is the fundamental distinction between the two. to identify the root node. The Decision Tree strategy prevents the slicing attribute node from running haphazardly. We went with the Random Forest method since it can be applied to both classification and regression. Additionally, this classifier would be able to tolerate any missing values and would not over-fit the device.

2.8 XGBOOST CLASSIFIER

When used to a predictive function, "boosting" or "optimization" techniques improve its accuracy by chaining together many instances of the same algorithm and then averaging their results and assigning weights to each one to lower the prediction's error. An example sequence like this can improve prediction accuracy beyond that of the simplest function in some situations. [40]. The XGBoost technique is a supervised ML package for building fast and effective Tree models are being improved. Its speed is due to the concurrent processing occurring in the background. Scalability and efficiency have contributed to its rapid growth and widespread use in recent years. What follows is a schematic representation of the XGBoost algorithm.

Figure 2. 4: XGBoost algorithm [41].

These benefits stem from the fact that the method is not only simple to pick up and use, but also has a library that works with R and python, and that it is computationally efficient and accurate to boot. [42]. The free version of this method has limited storage space and comes with hefty charges, as this is a cloud service.

3. MATERIAL AND METHOD

3.1 PYTHON IN MACHINE LEARNING

Python is a programming language that is both powerful and easy to learn. It has good high-level data structures and an object-oriented programming interface that is simple but efficient.[43]. Python is a perfect language for scripting and rapid software creation on a number of platforms because of its elegant syntax, dynamic typing, and interpreted nature. New C or C++ functions and data types can be easily applied to the Syntax parsing in Python (or other languages callable from C). One further way Python may be put to use is as an extension language for tweaking existing software. Compared to a shell, Python's structure and support for complex structures are far better developed. On the other hand, it has much more error checking than C, and since it is a higher-level language, it has built-in high-level data categories such as standardized arrays and dictionaries, which will take days to incorporate in C. Python is an interpreted language that eliminates the need for compilation and connecting, saving you time during software creation [44]. The interpreter can be used interactively, making it possible to experiment with language functionality, write throw-away programs, and validate functions throughout the development of bottom-up programs. Since it is written in C, it can also be used as a desk calculator. Python is accessible to a far wider issue domain than Awk or even Perl, due to its more generic data forms, and many items are at least as straightforward in Python as they are in those languages [45]. Python lets you break the software down into components that can be reused in other Python programs. Python encourages you to write programs that are both lightweight and understandable.

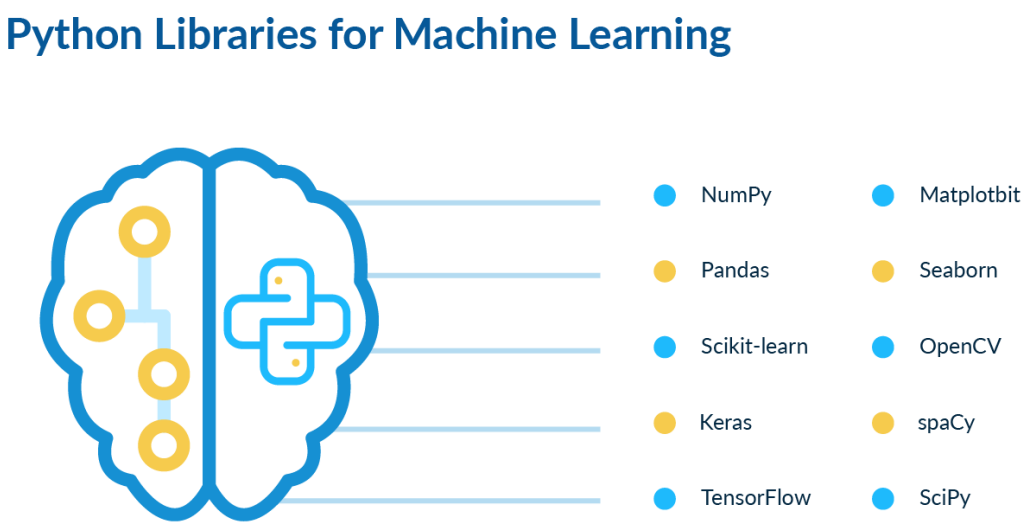

Figure 3. 1: Python libraries for machine learning [46].

3.2 ALGORITHM SELECTION

A key stage is determining which particular learning algorithm to use. Until early testing is considered good, the classifier (which maps unlabeled cases to classes) is not ready for widespread use. [47]. Classifiers are often judged based on their prediction accuracy (the ratio of correct to total predictions). There are at least three ways to determine a classifier's accuracy. One strategy involves using two-thirds of the data set for training purposes and the remaining one-third for estimating output. The training array is partitioned into independent and uniformly sized subsets, and the classifier is then trained on the union of all the subsets..[48]. Thus, the error rate of the classifier may be approximated as the mean error rate of the subsets. Leave-one-out validation refers to a type of cross-validation where one hypothesis is disregarded. In the test set, there is just one occurrence. [49]. This kind of verification inevitably comes with a higher price tag. This approach requires a lot of processing power, but it provides the most accurate estimate of the classifier's error rate.

3.3 FORMAL TECHNIQUES AND METHODS

Medical diagnosis is an inherently challenging job that necessitates severe accuracy when taking into account a number of variables [50]. Furthermore, given the degree of skill and experience needed for an accurate outcome, CHD prediction is a far more challenging job. According to a WHO poll, only 67 percent of medical practitioners can accurately forecast heart failure. Heart disorder affects one out of every twenty adults in New Zealand (roughly 180,000 people) and claims one life every 90 minutes [51]. With the number of people dying from heart disease predicted to increase in the coming years, there is a major study potential for forecasting CHD. ML techniques allow the application of intelligent methods to a variety of datasets in order to uncover valuable insights. The reprogrammable capacity of machine learning to explore, process, and analyze datasets allows it attractive to decision-makers in fields like medical diagnosis. Since detecting CHD necessitates training a model based on historical data, machine learning appears to be an effective technology to solve this problem.

3.4 CHD DATA COLLECTION

Information from a recent cardiovascular review focusing on Framingham, Massachusetts residents is freely accessible through the Kaggle platform. This score is then used to determine whether or not the patient has a 10% chance of (CHD) [52]. The information in the dataset is about the patients. There are over 4,000 records in all, including 15 distinct characteristics. Variables are those that may be modified. Each personality trait carries the possibility of becoming a risk factor. Socioeconomic, behavioral, and diagnostic conditions are all risk factors.

Table 3. 1. The CHD data attributes description.

Attributes	**Description**
age	Age of the patient
BMI	Body Mass Index
BPMeds	if the patient was taking blood pressure medicine or not
cigsPerDay	the average daily number of cigarettes smoked by the individual
currentSmoker	yes or no
diabetes	whether or not the patient has diabetes
diaBP	diastolic blood pressure
Gender	M or F
glucose	glucose level
heartRate	Heart rate, for example, is a discrete feature.
prevalentHyp	whether or not the patient was hypertensive
prevalentStroke	if the patient has ever experienced a stroke
sysBP	systolic blood pressure
TenYearCHD	10-year risk of (CHD) yes or no
totChol	rate of free cholesterol

3.5 PRE-PROCESSING AND FEATURE-SELECTOR

The aim of feature extraction and preprocessing is to reduce the number of elements in a dataset by creating new facets from existing ones (and then discarding the authentic features). After that, the current reduced collection of facets must be able to summarize the majority of the details present in the unique set of functions. An accumulation of the authentic collection may be used to construct a summarized model of the specific elements in this manner. To delete missed values, use the scikit-learn library's Basic Imputer pre-processing class. The replacement method and target value (which need not be NaN) are both under our control thanks to this flexible class (such as mean, median, or mode). [53] Feature-Selector would be included in this area. This tool's function is to successively apply several feature selection techniques. The number of chosen features might be reduced, for instance, by first removing related characteristics and then using recursive feature removal. In order to extract the most useful information from CHD data, we built a feature selection pipeline that can identify and produce just the most relevant and informative characteristics while filtering out the rest. [54] Our group developed the (BorutaPy) algorithm. To put it simply, the role meaning attribute aids tree-based models in differentiating attributes that may be utilized to create predictions. [55]. After a feature selection process, we selected 11 features out of 15 as the best features for our latest integrated machine learning models.

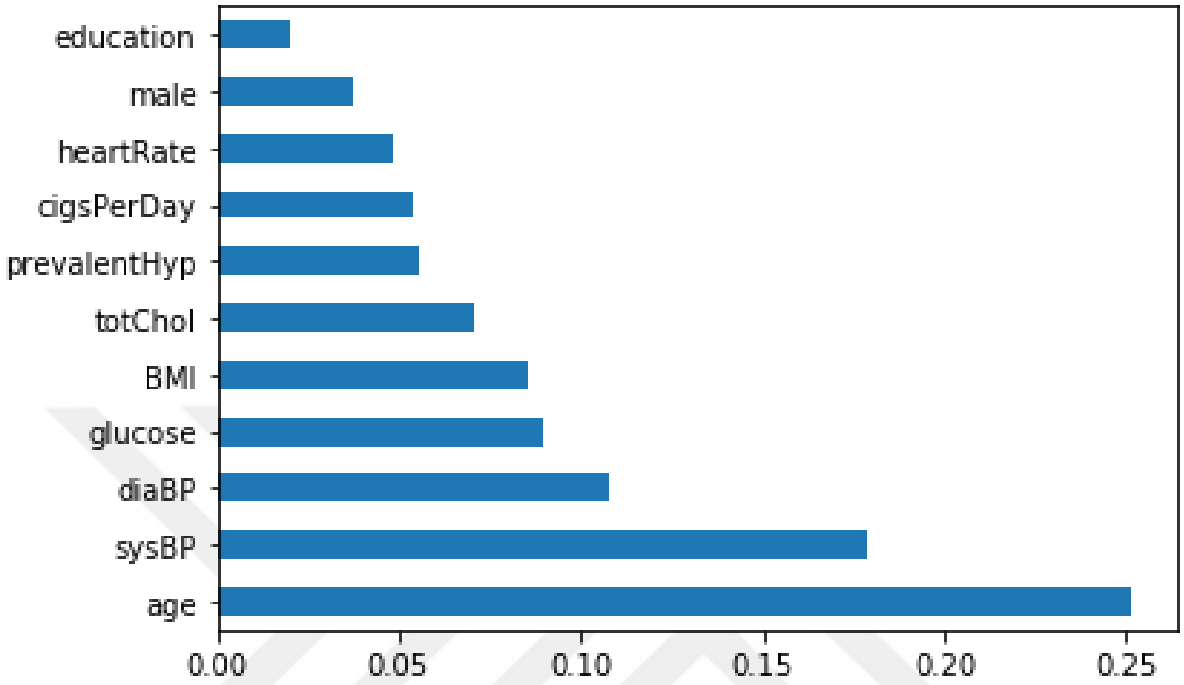

Figure 3. 2: Showed the importance of our selected features.

3.6 OVER SAMPLING MODEL

Since several natural processes seldom allow such findings, CHD data has an unbalanced distribution and is indistinguishable by class. CHD disorders, for example, may contribute to medical data with limited diagnostic classes in a community. Models of machine learning may have issues [56]. The problem with the uneven grouping is that there aren't enough instances of the minority class for the paradigm to learn the decision boundary with confidence. Oversampling instances from the underrepresented group is one strategy for addressing this problem. Minimal Synthetic Over-sampling is one popular approach to create additional instances. [57]. SMOTE is an approach to evaluating function spaces that clusters data points together based on their Euclidean distance from one another. Through the use of the SMOTE method and an original optimization strategy, we were able to significantly enhance the performance of our over-sampling model.

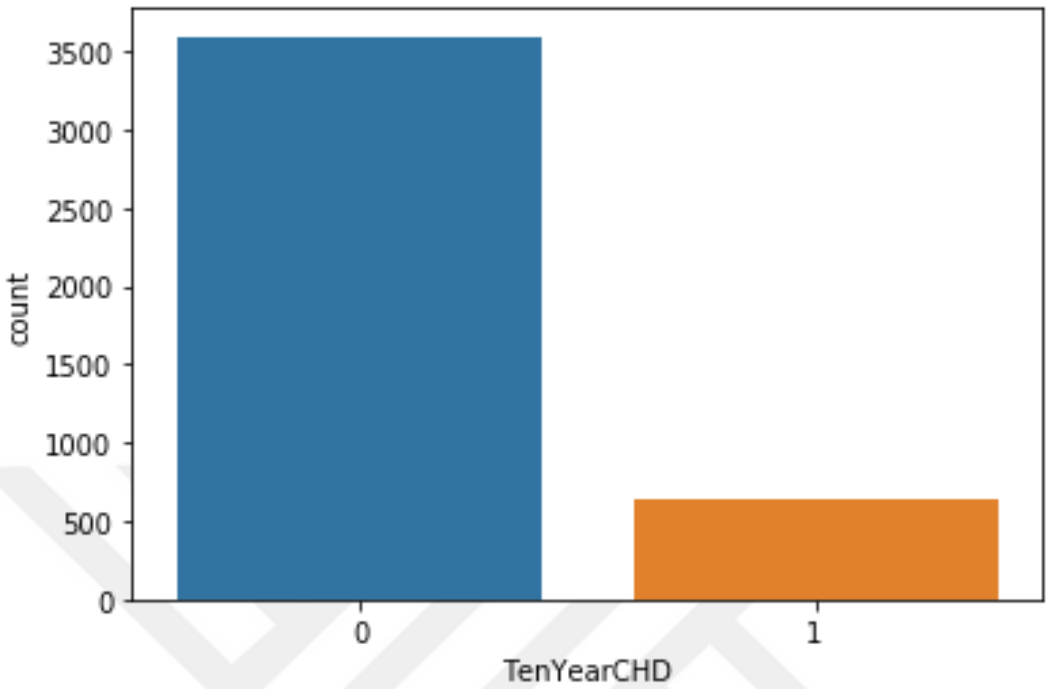

Figure 3. 3: The distribution of predicted classes before optimized SMOTE model.

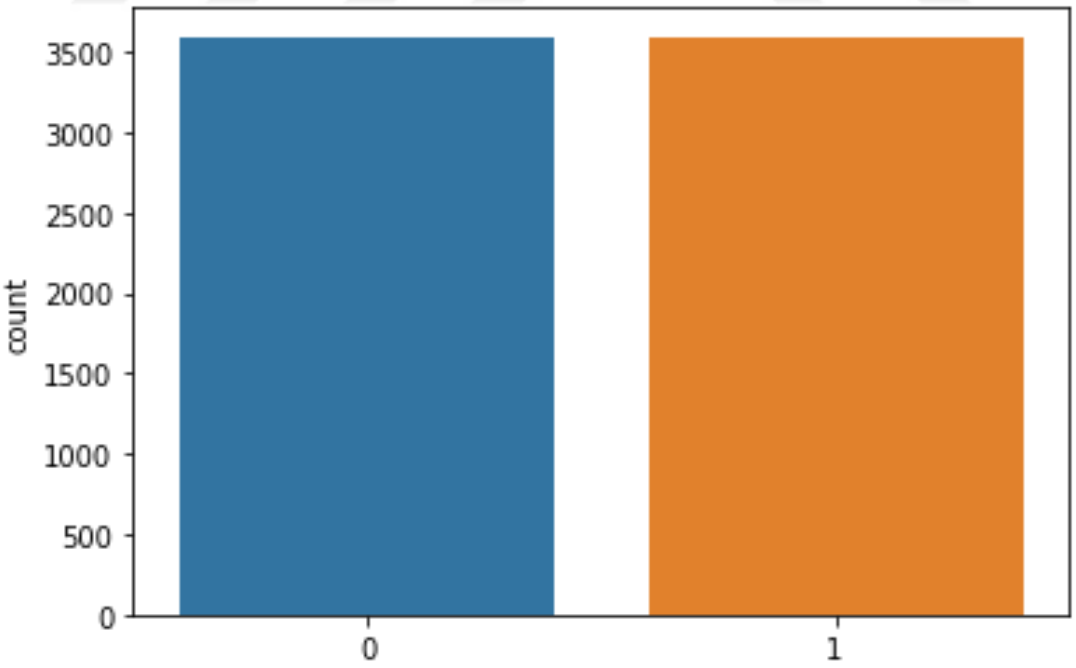

Figure 3. 4: The distribution of predicted classes after optimized SMOTE model.

3.7 HYPER-PARAMETER TUNING

One of the distinguishing characteristics of the most effective machine learning algorithms is their ability to automatically adjust hundreds, if not millions of "learnable" parameters in order to spot patterns in data. A tree-based model's computational thresholds for rendering predictions on the left or right branch, as well as the decision factors at each node, are all parameters that may be trained (decision trees, random forests, XGBoost, for exampleIn neural networks, the weights assigned to each link in order to boost or dampen activity as it passes from one layer to the next are known as "learnable parameters." We will adjust five hyper-parameters, for instance, for our XGBoost tests. [58]. An extensive and vital area of study, automated hyper-parameter optimization is devoted to this very problem. The heart of the technique is the algorithm that chooses the next set of hyper-parameters to test in light of the validation output backdrop, given a variety of hyper-parameters, learning rates, loss functions, and epochs to work with. [59].

3.8 XGBOOST API MODEL

The XGBoost (eXtreme Gradient Boosting) implementation of the gradient boosted trees algorithm is a common and effective open-source implementation. Gradient boosting is a supervised learning algorithm that incorporates predictions from a set of simpler and weaker models to try to forecast a target variable with acceptable precision. Due to its efficient handling of a range of data styles, correlations, and distributions, as well as the amount of Hyper parameters that can be fine-tuned, the XGBoost algorithm performs well in machine learning competitions.[60]. XGBoost may be used to solve problems including regression, sorting (binary and multiclass), and rating. We'll go into how to fine-tune the core parameters of our XGBoost API optimization model in this thesis. We should run a huge grid search with all the criteria together to find the best response in an ideal universe with infinite resources and no time limits [61]. With smaller datasets, we may be able to do so, but as data size increases, training time increases and the cost of each step in the tuning process increases. As a result, knowing the role of the parameters, as well as focusing on the interventions we expect to have the biggest influence on our results, is crucial. We'll fine-tune hyper-parameters that have a huge effect on output in this segment.

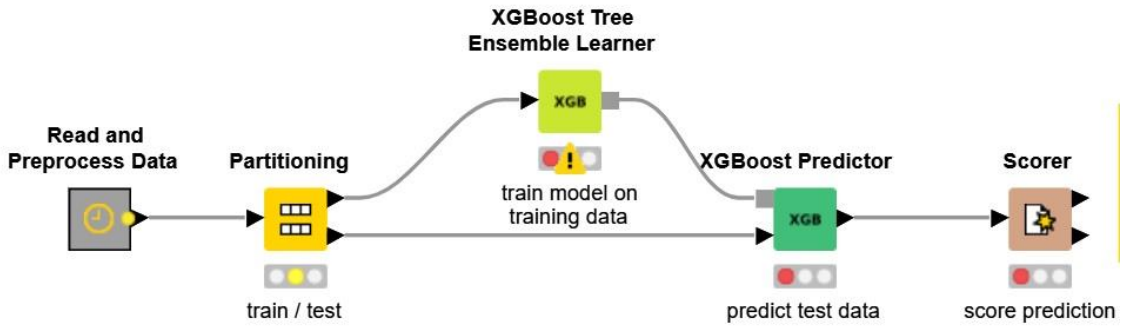

Figure 3. 5: XGBoost model structure [62].

3.9 RANDOM FOREST PARAMETER TUNING

A machine learning method called Random Forest uses decision trees. Random Forest's machine learning strategy is both practical and versatile. It's widely used because of its adaptability and usability. While the hyper-parameters aren't fine-tuned, it performs well on a variety of classification functions. All aspects of the forest, from its layout to its density and even its degree of randomization, are under RF's watchful eye (e.g., the number of variables listed as candidate splitting variables for each break or the sampling scheme used to produce the data).[63]. It's worth mentioning that the majority of hyper parameters are "tuning parameters," meaning their values must be adjusted to obtain the best performance for the specified dataset. Optimality is described by a particular value parameter that must be chosen in advance. Over fitting is a common occurrence in parameter tuning, in which dynamic rule parameter values tend to over-fit the training data, resulting in prediction laws that are excessively precise to the training data and implement well for this data but not for independent data. When tuning, utilizing a comparison dataset or cross-validation techniques may help avoid selecting sub-optimal parameter values. The problem of selecting random forest algorithm parameters is examined from two perspectives in this review. Its first section reviews the literature on the collection of various radio frequency parameters, while the second section compares various tuning techniques and software packages for obtaining optimum hyper parameter values. Here, we'll examine the efficacy of Random Forest and the key factors that can be adjusted to optimize results. [64]. Hyper-parameter tuning is expected since each data set has its own characteristics. Hyper-parameter tuning can be accomplished in a variety of ways. Hyper-parameters can be tweaked after the base model has been created and validated to improve certain metrics like the model's accuracy or f1 efficiency.

Repeatedly choosing parameters from the grid, the random forest model is fitted once the map is constructed with scikit-learn and Randomised-Search-CV. [65]. We won't have the correct criterion, but we'll pick the better one from among the many that are being fitted and evaluated.

3.10 SVM PARAMETER TUNING MODEL

It stands to reason that the SVM, and almost certainly the kernel, can be tweaked to improve quality, given they both include parameters that can affect training and classification performance. As a first step in optimizing a model, lowering the penalty for slack variables (C) and determining the optimal trade-off between representing misclassifications and broadening the definition are two common approaches. Misclassified samples will be heavily punished with high C values, leading to a (hyper-plane) that drastically lowers classification errors without compromising generalization[66]. C=1 always contributes to a low margin of safety in SVM behavior. However, misclassifications that might lead to an improper separation are only mildly punished at low levels. Linear functions, polynomial functions, and radial base functions are all examples of different kinds of functions (RBF). Polynomial and RBF are helpful for non-linear hyper-plane. Python's (Scikit-learn) C module's regularization method is used to maintain regularization. The "misclassification" or "error" penalty parameter is denoted by the letter "C." [67]. The term "misclassification" or "error" indicates how much error the SVM optimization can take.

4. EXPERIMENTAL RESULT

4.1 MODELS RESULTS

For CHD unbalanced findings, we employed optimal SMOTE over-sampling, which considers the extremely efficient strategy. We have implemented a train-test split on our dataset, allocating 30% of the CHD dataset for analysis and utilizing the remaining 70% for training and cross-validation (5 CV). We have pre-processed our dataset using techniques like attribute significance and missing values. The oversampling procedure yielded a sample size of 7192 for the distribution of the predictable class (10-Year CHD), with 3596 records coming from each of the two categories (0 and 1).

4.1.1 XGBoost API Optimization Model

With Hyper parameter customization and the API capability, we were able to increase the accuracy of our bespoke XGBoost model to 89% when applied to CHD data that had been over-sampled. Upon optimization, the optimal settings are as follows: (Learning Rate: 0.01, Gamma: 1, Max Depth: 10, Subsample: 1.0, Lamda: 4.5, Number of Trees: 500). Results from a classification study using our set up XGBoost models are shownin table 4.1.

Table 4. 1. The results of the optimized XGBoost API model**.**

Class No.	Precision	Recall	F-Score	Support
Cardio 0	0.85	0.97	0.92	1065
Cardio 1	0.96	0.84	0.90	1095
Avg.	0.91	0.90	0.90	2159

In addition, we display the ROC-Curve of optimization process in the figure 4.1.

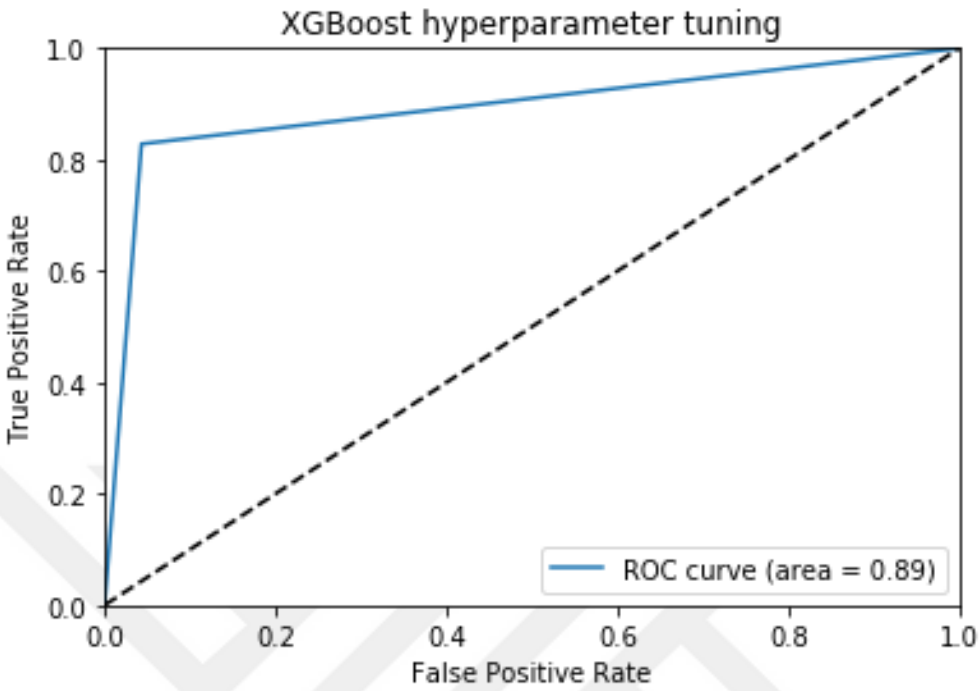

Figure 4. 1: The ROC-Curve of optimized XGBoost API model.

4.1.2 Optimized Random Forest.

The configured parameters had a top depth of 3 (also known as Bootstrap, contained in the 'Bootstrap Accuracy', which is false), along with which the model attained an accuracy of 90% (also seen in Table 4.2, with CHD features set to 7).

Table 4. 2. The results of optimized random forest model with CHD data.

Class No.	Precision	Recall	F-Score	Support
Cardio 0	0.87	0.95	0.91	1065
Cardio 1	0.95	0.86	0.90	1065
Avg.	0.91	0.91	0.91	2159

Using CHD data records, figure 4.2 displays the area under the ROC curve for this model at a confidence level of 90%.

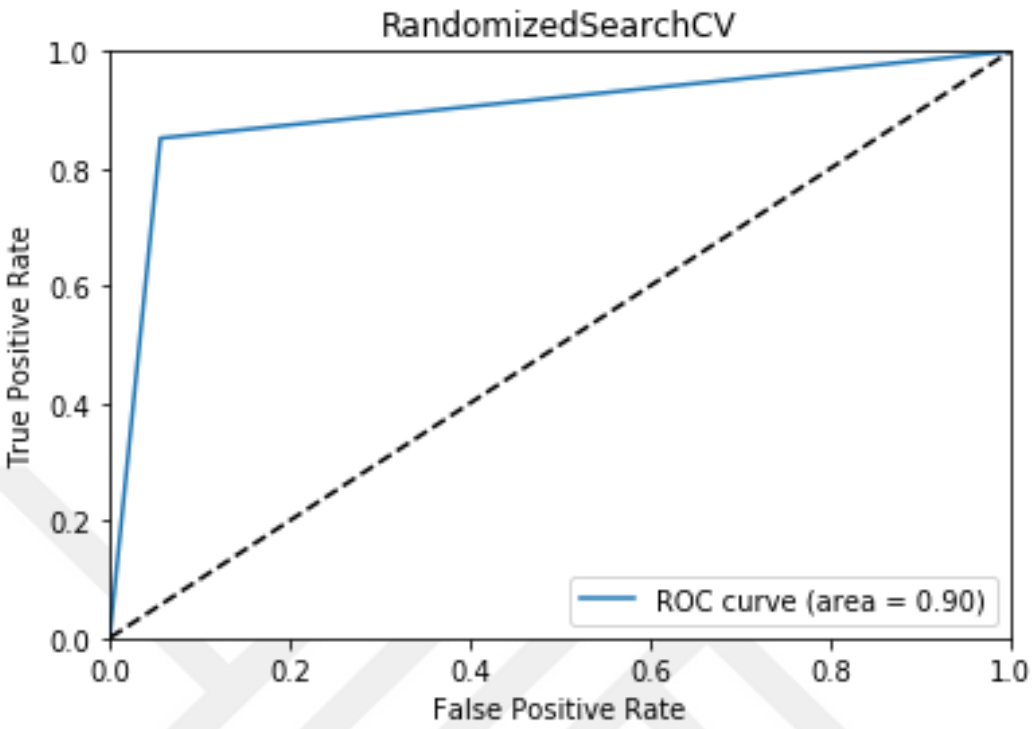

Figure 4. 2: The ROC-Curve of optimized random forest model.

4.1.3 SVM Hyper-Parameter Optimization

Here, we optimized the hyper-parameters of the SVM to build a powerful model. It was found that the most effective strategy for searching in a grid with consistent spacing was Grid Search (GS) with kernel adjustment. Using the optimal development-stage parameter combination, such as ('C': 1000, 'gamma': 0.001, and 'kernel': 'rbf'), we were able to attain an accuracy of 87%. Classification results for the set up SVM model are displayed alongside the results in Table 4.3.

Table 4. 3. The results of SVM optimization model with CHD data.

Class No.	Precision	Recall	F-Score	Support
Cardio 0	0.93	0.82	0.87	1065
Cardio 1	0.85	0.94	0.89	1065
Avg.	0.89	0.88	0.88	2159

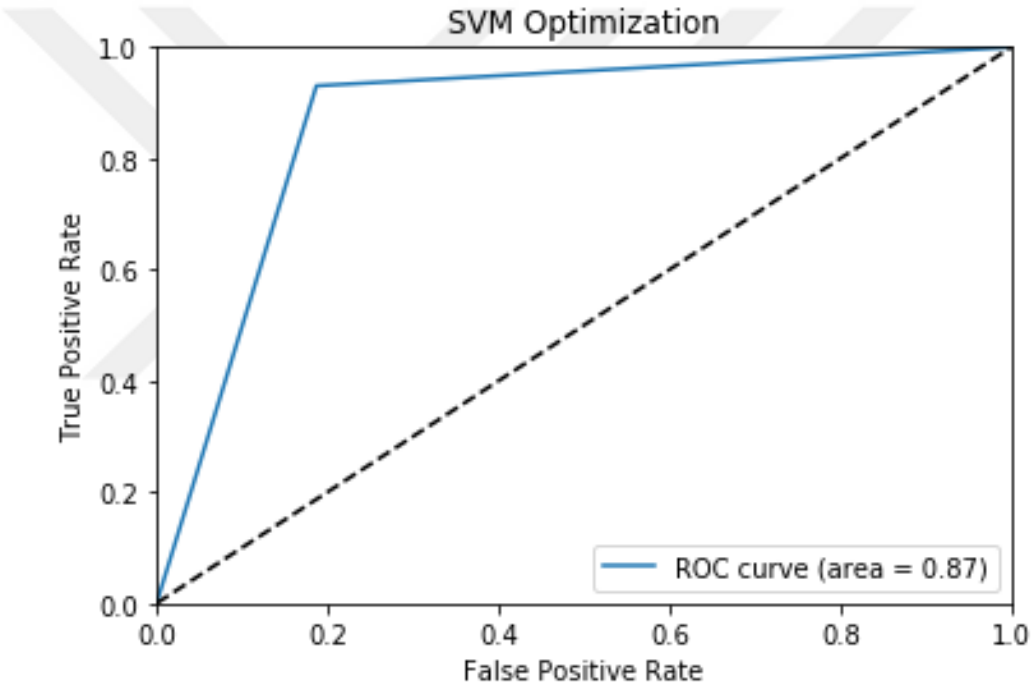

Figure 4. 3: The ROC-Curve of SVM optimization model.

4.2 COMPARISON

Our study centered on developing state-of-the-art optimization machine learning methodologies to aid a ground-breaking concept that, compared to existing methods for handling asymmetrical CHD datasets, would yield significantly better outcomes and higher accuracy.

Table 4. 4. This study's findings were linked to those in previous research.

Author	Objective	Methods	Accuracy
Juan-Jose Beunza et al.[14]	machine-learning algorithms for clinical event prediction (risk CHD disease)	Decision tree	84%
		Boosted decision tree	85%
		Random forests	79%
		SVM	69%
This study with imbalanced CHD data	New Novel Intelligent Optimization System for Imbalanced CHD Data	XGBoost API Optimization	89%
		Optimized Random Forest	90%
		Feature Selector	85%
		SVM Optimization	87%

5. CONCLUSION AND FUTURE WORK

5.1 THESIS SUMMARY

In the United States, coronary heart disease (CHD), which accounts for approximately 13 percent of all fatalities, is now the main cause of death due to cardiovascular conditions. It is impossible to exaggerate the significance of detecting heart abnormalities in their early stages in order to lessen the impact on cardiovascular health and stay away from cardiac arrests. Data mining is often used to describe the processes and procedures used in machine learning. Any one of these methods was first detailed in print well over fifty years ago. In spite of this, there has been a dramatic surge in both interest and knowledge of them in recent years. This is mostly attributed to huge advancements in algorithmic programming and the expanding processing power of modern computers. In this research project, we investigate a unique approach to the detection of coronary heart disease (CHD) that takes use of optimal machine learning methodologies, such as classifier ensembles.

The dataset is available for no cost on the Kaggle website, and it was collected as part of a recent study of cardiovascular disease among residents of Framingham, Massachusetts. The purpose of these subgroups is to identify whether patients have a 10% chance or higher of (CHD). There is patient-related information in the database. The total number of papers is around 4000, and there are 15 distinguishing features. It's possible to change the values of variables. Each characteristic may serve as a warning sign in some situations. Some of these risk factors are societal and others are physiological or medical in nature.

We applied an extremely effective technique, improved SMOTE over-sampling, to correct for CHD data imbalances. We also employed a train-test split to partition our dataset, allocating 30% of the CVD dataset to testing and 70% to training and cross-validation (5 CV). We also pre-processed our data using methods like feature significance and missing values removal. After using the over-sampling technique, the distribution of the predicted class (10-Year CHD) yielded 7192 records, split evenly between classes 0 and 1, with 3596 records in each.

In this work, we provide a novel CHD disease detection model that is optimized primarily by a machine learning optimization approach. We employed an integrated SMOTE over-sampling method in our optimization system and a Feature-Selector optimization model to choose the

optimal subset of features for our optimization framework, ensuring optimal performance. The suggested methodology utilizes three distinct optimization learners—random forest, XGBoost API machine, and SVM optimization—to get optimal results.

In order to provide a reasonable comparison to existing research, the suggested detection approach was tested using skewed CHD data. Rarely seen in the literature, we developed a two-stage statistical significance technique to evaluate the value of benchmarked classifiers' performance. We found that our suggested model surpassed the best existing CHD detection approaches in terms of accuracy, F1, and area under the ROC-Curve. When applied to these data sets, the results mirrored the most significant discoveries to yet.

5.2 CONCLUSION

In this study, we present a novel CHD disease optimization detection model that places particular emphasis on the optimization machine learning procedure. To pick the optimal collection of high-performing features for our optimization framework, we employed a Feature-Selector optimization model and the SMOTE over-sampling approach. There are three types of optimization learners incorporated in the proposed solution: random tree, XGBoost API machine, and SVM optimization. The suggested detection approach was validated using imbalanced CHD data, providing a useful benchmark against which to compare future investigations. Not commonly seen in the literature, we calculated the worth of benchmarked classifiers' outputs using a two-stage statistical significance technique. Based on experimental data, our suggested model has better accuracy, F1, and ROC-Curve value than state-of-the-art CHD detection methods. When applied to the relevant data sets, the findings mirrored the most up-to-date observations available.

5.3 FUTURE WORK

In the future, we'll concentrate on creating a setting conducive to medical facilities operating in smart cities. The process of transferring data between centers in smart cities, especially medical data, needs speed in data transfer and ensuring data security for patients. Therefore, we will use private generation networks, which are the fastest at the present time, to ensure the speed of data transfer. As for data security, we will use the blockchain to secure data movement within the network for health centers. This requires a script that ensures smooth work and very accurate results. In order to improve task completion and do away with the repetitive, time-consuming

routine, this work requires significant study since smart city implementation requires high precision. The blockchain is utilized in the context of encrypted work, such as Bitcoin, since it is thought to be difficult to hack. Therefore, its use in protecting patient data gives reliability to patients who live in smart cities. This allows to work within a safe environment. In addition, identification of other heart diseases and analysis of ECG data will be included. Where there is a major defect when transmitting ECG signals. Heart disease is a leading killer worldwide, as reported by the World Health Organization. Arrhythmia is a form of sickness that causes the heart to beat irregularly. A person with an arrhythmia has a heart that cannot efficiently pump blood throughout the body. The brain, the heart, and other organs are all susceptible to injury from insufficient blood supply. Patients with these conditions have a much better chance of avoiding unexpected death if they receive a prompt diagnosis and aggressive medical treatment. If cardiac arrhythmia is caught early, it is easily treated. Numerous factors support the use of a neural network in this investigation. For its physical (natural) applications, artificial neural networks' straightforward architecture and high class-classification accuracy make them ideal candidates. We offer an artificial neural network-based approach for identifying arrhythmia from an ECG recording. The sickness is primarily categorized as either normal or abnormal in this investigation. The signals from electrocardiograms were utilized to train and evaluate three distinct neural network models. Some information may be absent while examining arrhythmic data. Whenever a data attribute is missing, the next best match for that class is substituted for it. When training models for artificial neural networks to detect cardiac arrhythmias, Learning the law of motion is paired with the inverse algorithm. Cardiac arrhythmia means that the heart rhythm is abnormal. The normal rhythm of the heart begins at the sinus node and spreads to the ventricles after transmission to the atrioventricular node. The features studied in the study incorporate both rapid and slow heart rates, as well as supraventricular and ventricular tachycardia, as well as complete and incomplete block of the ventricles. The data set used consists of 20 histories of patients with arrhythmias that were manually examined. In the proposed method, electrocardiogram signals are first pre-processed to eliminate noise. The aberrations in the signals are then detected and the QRS set very important feature in identifying arrhythmias is identified. An artificial neural network consists of connections of artificial neurons that study the target function. After pre-processing, feature extraction is performed and then arrhythmia is diagnosed using artificial neural networks. It is important to note that the suggested neural network is a a three-layer "forward" neural network

(input, hidden, and output layers). Synaptic connections in the brain is trained using 20 datasets that contain QRS characteristics. The proposed system has achieved 98.48% accuracy.

The results produced by an electrocardiogram reflect the heart's electrical activity. P waves, QRS complexes, T waves, and U waves all play critical roles in this signal. Arrhythmic alterations are a symptom of cardiac illness that can be detected by monitoring the ECG for any irregularities. Cardiac arrhythmias are difficult to diagnose and require the attention of a trained professional. As a result, the concept of using automation to identify cardiac arrhythmias emerged. A fresh approach to identifying and categorizing cardiac arrhythmias using the integration of violet and neural networks. Time and frequency information from ECG signal recordings is extracted using discrete violet conversion. The output is fed into a neural network's training and evaluation processes as an input vector. While many heart arrhythmia algorithms have been developed over the past few years, most researchers have only examined a small sample size of data (20 records from the standard database, or 420 cases). First, the ECG signals are acquired and preprocessed to get rid of the recorded sounds and noise, which are the broad processes included in the current investigation. Then, we take just 5 pulses from the signal and pull out their time-frequency and discrete violet conversion features. The heart arrhythmia classification is done after the input vector has been normalized. The simulation findings demonstrate that the designed system, which employs a multilayer perceptron network as the classifier, is very accurate, achieving an accuracy of more than 97% when testing on a database of simulated heart rhythms.

Cardiac arrhythmia means that the heart rhythm is abnormal and is diagnosed using an electrocardiogram signal. Eight different forms of cardiac arrhythmias are classified using an electrocardiogram signal classification technique for the diagnosis of cardiac arrhythmias, which combines one reduction approach, principal component analysis, and potential neural network classification. The electrocardiogram signal shows the function and rhythm of the heart and is one of the most important tools for diagnosing various heart diseases. In the last few decades, the use of automatic cardiac arrhythmia classification using computational intelligence techniques has received much attention. The use of three methods of feature extraction, feature selection and classification construction has been investigated in this study. In the feature extraction method, a number of signals are extracted as effective features. The dimensions of the retrieved characteristics are decreased to enhance the performance of the suggested system. Principal component analysis and linear analysis methods are used to reduce feature dimensions. The heart

arrhythmia classification uses 200 data samples from the 1980s. The highest classification accuracy for the proposed system is 99.71%.

The most reliable method for identifying cardiac problems is an ECG reading. Exact readings of the heart's electrical activity can be obtained. It takes a lot of time to manually identify arrhythmic ECG patterns. One of the major areas of study in clinical cardiology is the automatic identification and classification of arrhythmias. For the purpose of identifying cardiac arrhythmias, we describe a multi-stage clustering approach that combines clustering by greatest margin using evolutionary algorithms. The database that was employed contains information on five distinct types of arrhythmia, ventricular fusion, premature contraction block, including normal sinus rhythm, heart rate that is normal and premature atrial contraction. The method for identifying cardiac arrhythmias is made up of three parts: preprocessing, feature extraction, and arrhythmia classification. To begin, the unfiltered ECG signal is acquired and processed to determine the individual waveforms. Each record is processed in advance by having the audio stripped out. The second step is feature extraction, which aims to find the best coefficients determined to describe the electrocardiogram signal. The final phase employs a hierarchical clustering technique with the purpose of identifying cardiac arrhythmia. Seventy percent of the data, or 130 records, were used for training, while the remaining 30 percent, or 55 records, were used for testing the system. Three methods of sensitivity, properties and accuracy have been used to evaluate the performance of the system. Sensitivity, properties and accuracy were 82.4%, 98.8% and 97.4%, respectively.

a computer diagnostic system for classifying five heart rates is presented. Investigations include normal heart rate, premature block contraction, premature lumbar contraction, left block, and right block. A spectral correlation matrix of each heart rate is shown. Since the majority of the components in this matrix are almost zero, the feature is reduced by employing the whole matrix throughout the feature extraction procedure. The main purpose of reducing the specified space is to find the relevant criteria that cause maximum classification accuracy. In this paper, two methods are used to reduce the feature space, which include principal component analysis and feature selection. The filter-based principal component analysis method uses a linear combination of input properties and variance. In the feature selection method, each feature is evaluated independently. The database used includes 30 minutes of heart rate recording. 363 data were used in the present study. The current study looked into the parameters of preterm heart rate, preterm ventricular contractions, preterm atrial contraction failure, and preterm heart rate. The suggested system has

been tested using the 10-segment validation approach. Sensitivity, specificity, and accuracy for the suggested technique are 99.20%, 98.50%, and 97.50%, respectively.

Provides a fresh method for categorizing heart rhythm disorders. In this investigation, we used an inverted neural network in conjunction with a correlation-based feature selection method. Cardiac arrhythmias were the focus of this research, which sought to categorize them into two groups: those with and without the condition. The UCI arrhythmia database has been used to test intelligent decision systems that incorporate reliability characteristics. Artificial neural networks for categorization are one tool in this subject. For improved and more precise classification, many different feature selection methods have been applied. The goal of this work is to establish a connection between feature selection and prospective linear selection. Accuracy, specificity, and sensitivity are used to grade the categorization outcomes. 420 samples are available in the utilized database. The datasets used can be broadly classified into three types: training (68%), validation (16%), and test (12%). (16 percent). Because even with the best classification, the system may have poor performance if the features are not selected, feature extraction and reduction is one of the crucial processes of classification. The initial stage of every model's development process is the preparation of the data. Common neural networks don't progress in stages. The suggested neural network examines this issue by incorporating a scaling factor that minimizes the weight settings of the network. This study project is divided into two primary stages: the feature extraction phase is the first, and the feature reduction phase is the second. Both phases make use of feature selection correlation and classification through the utilization of backward extended neural networks. In this study, we offer experimental results showing that, on average, 100 simulations lead to a classification accuracy of 87.71 percent.

Diagnosing heart abnormalities such as arrhythmias using electrocardiogram signals takes hours and is tedious and time consuming. The use of automated computer-based techniques to diagnose arrhythmias requires a large amount of medical information. Diagnoses cardiac arrhythmias by classifying them as normal and abnormal. In the recorded data, standard electrocardiogram signals had 12 features to detect cardiac arrhythmias. The benefit of the suggested strategy is that it reduces the feature space by employing various feature reduction techniques. Feature space is reduced based on the classification of the decision tree. The examined study parameters include mean, absolute error, mean root error, relative absolute error, and F measurement. Consolidated data were used in the current investigation. The practice of combining data from several sources to provide

better outcomes is known as data integration. This step involves using the removal of various features. The resulting data set contains redundancy in the attribute, so there is space to filter this redundancy in the attribute. The pseudo-general code for this step of data synchronization is done using the feature deletion technique. The present study used a basic classification based on the decision tree. During this process, a rule is created, deleting the items it covers and re-creating rules for the remaining items until no properties are left and all the properties are categorized. The data used contains 279 samples. 90% of the data were used for system training and 10% of the data were used for system testing. The accuracy of classification based on the area under the curve is 91.11% [68].

A broad approach for computerized diagnostics employing decision tree, artificial neural network, and nearest neighbor K methods is suggested to properly diagnose cardiac abnormalities. This study utilized information from two distinct arrhythmia and fibrillation databases. 1200 heart rates were analyzed in 360 samples that were stored in the database. At first, the incoming ECG signals are pre-processed to exclude any extraneous data. The Fourier transform technique is useful for resolving frequencies but lacks localization time and incorporates information that occurs at various periods. To accomplish this separation, the signals are subjected to a discrete wavelet transform. Using principal component analysis, we are able to narrow the scope of the issue under investigation. The system has been verified using the 10-segment validation strategy. MATLAB is used to run virtual experiments on the proposed approach. An efficiency of 99.45 percent has been achieved by using the nearest neighbor K. With this in mind, the suggested diagnostic system offers a great deal of dependability for its intended use by medical professionals. The use of this approach to the detection of further cardiovascular diseases is possible. The benefits of the suggested approach mainly include of: 1) When compared to earlier approaches, the proposed method is more precise. 2) With such a high rate of success in classifications, the system's accuracy speaks for itself. 3) The data is retrievable as the mean efficiency of 10 separate tests. 4) The proposed approach is hands-off, non-invasive, and minimally doctor-interactive. 5) The proposed approach is generalizable to a variety of cardiac conditions. Presents a multi-arrhythmic diagnosis of ECG using convolutional neural network based on sampling attention time, which uses slpbd data set and its accuracy is 92.97%.

The method it will be. After the user inputs their ECG signal data, the program uses an optimization strategy based on an extraction operation feature using a differential evolution algorithm. Two key

topics in discussions of differential evolution are the regulation of parameters and the choice of evolutionary strategies. First, it determines the values for F, the chance of a CR crossing, and the number of NPs in the population. However, many approaches perform better than others when applied to various issues. Therefore, the best course of action should be chosen. Population variety, early development potential, and subsequent convergence are all impacted by the parameters chosen. The trade-off between divergence and convergence in differential evolution is largely determined by the evolutionary approach chosen. There is a clear disparity in the polling capacities and convergence tendencies of the various evolutionary techniques. There are a number of ways in which combination operations might influence a global optimization procedure. Traditional binomial combinations serve a useful purpose, but they are more reliant on the compound's underlying coordinate system, which limits their use in some situations. The structure of the population is also a good barometer of the algorithm's efficacy. Due to the high probability of extinction of efficient alleles in a small population, the generation of fit people will be stunted. But if there are too many people in the population, the algorithm is less likely to find what it's looking for. Improved differential evolution performance is increasingly being focused on as a result of early convergence, with a focus on controlling parameters and improving strategy, and on the performance of composition and population structure. Parameter-compatible jDE, JADE with the "flow to pbest / 1" strategy and adaptive parameters, SaDE with the adaptive difference strategy, CoDE with relationships for the experimental individual algebra strategy and the control parameters, EPSDE with the mutation strategy and the control parameter, TDE with the triangular mutation strategy and the ultra-consistent multi-S, and so on are just a few examples. Additionally, evolutionary differentiation based on hybrid operations, like ODE that makes use of orthogonal hybrid operators, and CoBiDE that makes use of covariance learning and a two-dimensional parameter distribution to better its predictions. as well as evolutionary differentiation based on population structure, such as SPSRDEMMS for multivariate strategies, MPEDE Population multiplication and strategy sets, master-slave model, island model, and cellular model.

Differential evolution is comparable to greedy evolution algorithms that employ global optimization and real number encoding in this way. The evolutionary stage involves recurrent rounds of mutation, combination, and selection until the conditions for their cessation are satisfied. Quality is measured by the fit function's output, and the top performer is noted. Population growth over G generations is shown by (D, Gχ.) where NP is the population size and D is the dimension

of the solution space. The following components make up a human being: D, which may be written as the following equation: (5.1)

$$x_i^G = \{x_{i,1}^G, \dots, x_{i,,D}^G\}, i \in \{1,2, \dots, NP\} \quad (5.1)$$

In this regard, $x_{i,j}^G \in (x^L, x^H)$ and x^L and x^H the extremes of the repeated samples' values should be used as examples. The independent sample x_i^G generates an individual type of v_i^G employing mutation as a technique in the original population. "DE / rand / 1" means that DE chooses the mutation's perturbation at random. Relationships are characterized by Expression (5.2).

$$v_i^G = x_{r1}^G + F.(x_{r2}^G - x_{r3}^G), r_1, r_2, r_3 \in \{1,2, \dots, NP\} \quad (5.2)$$

In this regard, $r_1 \neq r_2 \neq r_3$ and r_1, r_2 and r_3 are random mutants. The range [0,1] is used to determine the dimensions F that meet the condition. By expanding the genetic diversity of the newly formed individuals, blending procedures aim to produce novel mixtures of people from the main population. The binomial combination strategy has received the differential evolution algorithm's blessing. Equation describes the combination operation (5.3).

$$u_{i,j}^G = \begin{cases} v_{i,j}^G, & if\ rand_j \leq CR\ or\ j = j_{rand} \\ x_{i,j}^G, & othewise, j = 1, \dots, D \end{cases} \quad (5.3)$$

Where $rand_j \in [0,1]$ is chosen from a range of [1,2,...,D] with a probability of [0,1]. Children are always better than or on par with their parents, X_i., thanks to the selection process, which is primarily motivated by a survival of the fittest mentality. The new user interface is accepted when the new user's fit function performs better than the impartial individual. In any case, X_i. still exists in the subsequent population and continues to carry out mutation and combination operations as a goal in the subsequent iterative computation to ensure that the population constantly evolves in the direction of the ideal solution. The selection procedure minimizes the fit function's value in accordance with Equation (5.4).

$$x_r^{G+1} = \begin{cases} u_i^G, & if\ f(u_i^G \leq f(x_i^G) \\ x_i^G, & othersiwe \end{cases} \quad (5.4)$$

In this study, the fitting function f(x) is utilized to make a diagnosis of coronary heart disease; the investigation is over as soon as the ailment is identified.

The identification and classification of cardiac arrhythmias is possible with the use of deep learning techniques like convulsive neural networks. The attributes will be presented as a matrix s that contains relative frequencies for two signals, one with a gray surface value of I and another with a value of j, separated by a distance d and an angle. By utilizing the input signal window W(x,y,c) for each d and , as well as the input matrix s(i,j,d,) for the convolution neural network and its parameters, we can get the following: we can train a deep neural network to recognize images.

The following is a broad description of what it means:

- The matrix is fed the count of instances when gray area I is tilted toward gray. s as an input. j, so that $\boldsymbol{W(x_1, y_1) = i}$ and $\boldsymbol{W(x_2, y_2) = j}$ and the relation$\boldsymbol{(x_2, y_2) = (x_1, y_1) + (dcos\theta, dsin\theta}$.
- Inputs to the differential evolution optimization method
- Confusion reigns in the neural network's inputs and among its neurons.
- Convolutional neural networks use a torsion layer, a fully connected layer, and a pulling layer in their hidden or intermediate layer.
- Neurons and the input layer of the neural network receive the extracted characteristics as input.
- An appropriate filter weight distribution for use in the torsion layer is $w[0] * x[0] + w[1] * x[1] + w[2] * x[2]$. Note that this is a Dilation-style filter. The filter's initial weight, which is the filter's content, will be a 3 x 3 x 3 matrix whose dimensions are adjustable within the range of the extracted attributes.
- When supplied as a stimulus to the torsion layer, the sigmoid function takes on a nonlinear form.
- Maximum pooling speed in the pooling layer is utilized for basic pooling.
- After a predetermined amount of iterations of training in the convulsive neural network's hidden layers, the condition will be resolved if the feature classes can be correctly

classified. After all other diagnostic methods have been exhausted, the ECG signal will be used to make the diagnosis of cardiac arrhythmia.

To use the lattice, it is required to define the torsions involved. Thresholding wavelet coefficients, using adaptive filters, and limiting the range of action potentials are three common approaches. This study takes a different strategy by employing thresholds inside the action potential range. This cutoff point is calculated using the following equation: (5.5).

$$\begin{cases} \sigma_n = median\{\frac{|x|}{0.6745} \\ Threshold = 3.5\ \sigma_n \end{cases} \quad (5.5)$$

Microelectrode signal (raw signal) is denoted by x in (5.5), whereas the estimated noise standard deviation is denoted by σ_n. If the signal's standard deviation is utilized instead, a bigger threshold value is calculated, leading to the erroneous elimination of some torsions. Once the cutoff point is decided upon, the angles are arranged according to their maximum values. Finding dynamic barriers with torsions relies heavily on proper alignment of these obstacles. Like any other network, this neural network has to be trained. Our goal in this lesson is to locate a mapping of the form $f: R^n \rightarrow R$ in relation (5.6).

$$f(v) = \sum_i^n w_i \varphi(\left||v - C_i|\right|) \quad (5.6)$$

According to Equation (5.6), $v \in R^n$ is a 32-point vector for input, and the Gaussian basis function $\varphi(0)$ is defined as Equation (5.7).

$$\varphi(v) = exp(\frac{-v^2}{2\sigma^2}) \quad (5.7)$$

For each training sample, the error is then calculated using the slope of the associated equation and random beginning values for the weights, as shown in Equation (5.8).

$$e_i = t_i - y_i = t_i \sum_{j-1}^N w_j \varphi(\left|\left|v_i - C_j\right|\right|) \quad (5.8)$$

Thus, the sum of the network's errors across all P training input vectors is equal to $E = \frac{1}{2}\sum_{i=1}^{P}|e_i|^2$. Training is considered complete when E is less than the threshold error. This parameter is adjusted manually at the start of each task. If it doesn't work, a gradient slope is used to recalculate the weights. Each torsion's class membership is learned by a process of training using a multi-channel convolutional neural network. The layered architecture of the network must be defined. There are many layers involved in the training process, including a visible "input" layer that contains neurons and a "hidden" "training" layer. This inside tier features three interconnected twisting, tugging, and gluing levels. The output layer undergoes a last quality assurance procedure. Coronary artery disease detection is a difficult problem and a potential field for exploration. An, X_{nvar}. array representing the current location of the torsion layer in the convolution neural network is used to determine if a coronary artery is present or absent in an optimization problem with N_{var}. In order to define this array, we need the following equation: (5.9).

$$Convolve = [x_1, x_2, \dots, x_{Nvar}] \tag{5.9}$$

The cardiac function (,f_p.) in the Convolve is evaluated to determine how suitable (or how much gain) the present torsion layer is. This means that we may establish the linkages (5.10) and (5.11).

$$profit = f_p(Convolve) = f_p(x_1, x_2, \dots, x_{Nvar}) \tag{5.10}$$

$$j\left(x^{(i)}, \dots, x^{(n)}, \theta^{(1)}, \dots, \theta^{(n)}\right) = \frac{1}{2}\sum_{(i,j):r(i,j)=1}\left(\left(\theta^{(j)}\right)^T x^{(i)} - y^{(i,j)}\right)^2 + \frac{\lambda}{2}\sum_{i=1}^{n_m}\sum_{k=1}^{n}(x_k^{(i)})^2 + \frac{\lambda}{2}\sum_{i=1}^{n_u}\sum_{k=1}^{n}(\theta_k^{(j)})^2 \tag{5.11}$$

Specifically, the function of the overall aim in the context of detecting disturbances or failing to do so is described as a relation (5.11). In order to better diagnose coronary heart disease, it is recommended that the aforementioned function be reduced to an absolute minimum. Down fact, one way to reduce the number of unnecessary subsections is to zero in on the precise region at issue, so weakening the connection between them (5.12).

$$\min_{\substack{x^{(1)},\dots,x^{(n_m)} \\ \theta^{(1)},\dots,\theta^{(n_u)}}} j\left(x^{(1)}, \dots, x^{(n)}, \theta^{(1)}, \dots, \theta^{(n)}\right) \tag{5.12}$$

As can be seen, the multi-channel convolution approach relies on a deep neural network with a topology optimized for maximizing the presence or absence of coronary heart activity. In this research, it was shown that adding a negative sign to the cost function multiplied by a deep multi-channel convolutional neural network was adequate to address minimization difficulties. The convolution technique first employs a Convolve matrix of size $N_{pop} * N_{var}$ is produced. Each of these convolves is then given a different set of pulling layers. The typical number of elements in a pooling layer is between two and five. These figures are utilized to determine the maximum and minimum amounts of polishing time allotted to each torsion component throughout all training repetitions. The use of linked layers within a given region is another hallmark of convolutional structures based on deep structure. As a result, the highest amplitude of a convolution neural network's linked layers is known as $Max_{Connected\ Layer}$). The maximum allowable range of data in an optimization problem var_{high} and the lower limit var_{low}, There will be a number of $Max_{Connected}$ Levels equal to the number of depth layers. The problem depends on two parameters: the maximum and minimum number of existing levels of educational data. As a result, the definition of the $Max_{Connected}$ Layer is the following: (5.13).

$$Max_{Connected\ Layer} = \alpha \times \frac{Number\ of\ current\ layers}{Total\ number\ of\ layers} \times \left(Var_{high} - Var_{low}\right) \quad (5.13)$$

Equation determines the Ma,x-Connected Layer's maximum value (5.13). Layers are θ in relation (5.13), while the estimator's value is λ in relation (5.12).

The deep convolutional neural network's torsional segments each have a radian deflection and only traverse (λ%) of the total identified areas to the best possible destination at the moment. These two parameters assist each torsion portion of the deep convolutional neural network in its quest for additional environment data. The value of () is a random integer between 0 and 1, whereas the value of λ is a number between π/6 and π/6. After the migration of all the torsional sections of the deep convolutional neural network to the target point and the determination of the new habitats for each, each torsional section of the deep convolutional neural network will have a number of additional layers. Following the specification of a $Max_{Connected\ Layer}$ for each torsion section of the deep convolutional neural network, communication with the layers may then commence. The prior stage decides the amount of layers that will be used for each segment of torsion that will be created.

The deep convolution neural network's construction creates a hidden layer after reading the data from the input layer. This layer is made up of 20-channel torsion layers and a 3x3 filter with an atherosclerotic rate of 3 (r = 3), which is 3x3x3 in the maps.

Fusion may be seen in the torsion layer. In addition to this, its atherosclerosis rate is 7, which corresponds to the value r = 7, and the filter that has been applied to it is 333. In this segment as well, there are ten different strata. Next, the 1000 data equivalent to the training layers (i.e. 100%) are used in a random pooling and max pooling layer to apply training and classification operations in the polishing layer. Because the data has been "trained," it may be used to evaluate several potential mappings. Similar to its original 333 structure, the filter applied to it has an atherosclerosis rate of 11. The third layer of the data test is the three-dimensional layer, which is likewise influenced by the torsion layer because it is a completely linked layer. A cardiac coronary is displayed in the output layer thanks to the data test done in this layer, and classes are then generated to show off this particular heart region.

In recent decades, there has been a lot of research on automatically detecting arrhythmias using an electrocardiogram. ECG heart rate analysis became feasible with the ongoing development and improvement of open source ECG databases like MIT-BIH and slpdb. The AAMI standard categorizes heart rates into five groups based on the waveform of the heartbeat: normal ectopic, extracerebral (SVEB), ventricular ectopic (VEB), fusion, and unknown beat. An effort has been made in this study to offer a smart medical diagnosis system based on reliable slpdb data. In this procedure, the ECG signal data is input into the software first. Then, all of the steps necessary to achieve the goal of detection—feature extraction using a differential evolution optimization method, classification, and the use of a convolution neural network for learning—are performed concurrently. The results demonstrate that the suggested strategy increases accuracy in the identification of cardiac arrhythmias by a percentage point when compared to conventional approaches. In this research, we consider three different types of algorithms for detecting coronary heart disease: a deep convolutional neural network trained with the bat algorithm, a deep convolutional neural network trained with the differential evolution algorithm., and a deep convolutional neural network trained with the genetic algorithm. The reliability of the various methods offered has been assessed using a variety of criteria, and their relative performance in controlled experiments has been compared (with both sets of data and both sets of operators using the same parametric rate). The suggested approaches achieve high rates of accuracy (94.02

percent), sensitivity (93.54 percent), and features per second (92.30 percent). Equal to 89.35%, 89.31%, and 88.50% in the sensitive area. Three different feature percentages total 81.21 percent, 80.56 percent, and 80.49 percent. As an added bonus, two primary publications propose a technique using an ANN and using deep neural network cardiac arrhythmia diagnostic seizures have all the same study data. The accuracy of their solutions is 93.18 and 92.97 percent. It is obvious that the suggested technique, which uses a deep convolution neural network based on the bat algorithm, gives greater accuracy when comparing the outcomes of the two referred methods with those of the way offered in this study (94.02 percent). related to an improvement of between 0.84 and 1.05 percent. Considering that a lack of further data or clinical data in the country is one of the most significant challenges facing this research, the work is presented in a research context. In addition, the processing of "big data," or a massive amount of data, requires robust infrastructure.

REFERENCES

[1] M. D. Smith and A. Coleman-Jensen, "Food insecurity, acculturation and diagnosis of CHD and related health outcomes among immigrant adults in the USA," *Public Health Nutr.*, vol. 23, no. 3, pp. 416–431, 2020.

[2] P. Rajpurkar *et al.*, "CheXNet: Radiologist-Level Pneumonia Detection on Chest X-Rays with Deep Learning," *arXiv*, pp. 3–9, 2017.

[3] M. Grewal, M. M. Srivastava, P. Kumar, and S. Varadarajan, "RADnet: Radiologist level accuracy using deep learning for hemorrhage detection in CT scans," in *Proceedings - International Symposium on Biomedical Imaging*, 2018, vol. 2018-April, pp. 281–284.

[4] J. J. Beunza *et al.*, "Comparison of machine learning algorithms for clinical event prediction (risk of coronary heart disease)," *Journal of Biomedical Informatics*, vol. 97, no. July. Elsevier, p. 103257, 2019.

[5] P. Rajpurkar, A. Y. Hannun, M. Haghpanahi, C. Bourn, and A. Y. Ng, "Cardiologist-Level Arrhythmia Detection with Convolutional Neural Networks," *arXiv*, 2017.

[6] D. S. W. Ting *et al.*, "Development and validation of a deep learning system for diabetic retinopathy and related eye diseases using retinal images from multiethnic populations with diabetes," *JAMA - J. Am. Med. Assoc.*, vol. 318, no. 22, pp. 2211–2223, 2017.

[7] T. Alhanai, M. Ghassemi, and J. Glass, "Detecting depression with audio/text sequence modeling of interviews," in *Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH*, 2018, vol. 2018-Septe, pp. 1716–1720.

[8] J. Wu, X. Y. Chen, H. Zhang, L. D. Xiong, H. Lei, and S. H. Deng, "Hyperparameter optimization for machine learning models based on Bayesian optimization," *J. Electron. Sci. Technol.*, vol. 17, no. 1, pp. 26–40, 2019.

[9] I. Kononenko, "Machine learning for medical diagnosis: History, state of the art and perspective," *Artif. Intell. Med.*, vol. 23, no. 1, pp. 89–109, 2001.

[10] M. L. Giger, "Machine Learning in Medical Imaging," *J. Am. Coll. Radiol.*, vol. 15, no. 3, pp. 512–520, 2018.

[11] M. L. Giger, "Machine Learning in Medical Imaging," *J. Am. Coll. Radiol.*, vol. 15, no. 3, pp. 512–520, 2018.

[12] D. Castellanos, "Machine Learning Explained in a Simple Way | Medium." [Online]. Available: https://medium.com/@castellanos.johann/machine-learning-explained-in-a-simple-way-a2dd09237999.

[13] E. Walsh, "Pathological prediction: a top-down cause of organic disease," *Synthese*, vol. 199, no. 1–2, pp. 4127–4150, 2021.

[14] "Atherosclerosis 2011.jpg Wikimedia Commons." [Online]. Available: https://commons.wikimedia.org/wiki/File:Atherosclerosis_2011.jpg.

[15] J. T. Willerson, A. Maseri, and P. W. Armstrong, "Coronary Heart Disease Syndromes: Pathophysiology and Clinical Recognition," in *Cardiovascular Medicine*, 2007, pp. 667–698.

[16] R. Nakanishi *et al.*, "Machine Learning Adds to Clinical and CAC Assessments in Predicting 10-Year CHD and CVD Deaths," *JACC Cardiovasc. Imaging*, vol. 14, no. 3, pp.

615–625, 2021.

[17] K. Orphanou, A. Dagliati, L. Sacchi, A. Stassopoulou, E. Keravnou, and R. Bellazzi, "Incorporating repeating temporal association rules in Naïve Bayes classifiers for coronary heart disease diagnosis," *J. Biomed. Inform.*, vol. 81, no. March, pp. 74–82, 2018.

[18] J. J. Beunza *et al.*, "Comparison of machine learning algorithms for clinical event prediction (risk of coronary heart disease)," *Journal of Biomedical Informatics*, vol. 97. Elsevier Inc., p. 103257, 2019.

[19] D. Mehanović, Z. Mašetić, and D. Kečo, "Prediction of heart diseases using majority voting ensemble method," in *IFMBE Proceedings*, 2020, vol. 73, no. May 2019, pp. 491–498.

[20] D. Han *et al.*, "Machine learning based risk prediction model for asymptomatic individuals who underwent coronary artery calcium score: Comparison with traditional risk prediction approaches," *J. Cardiovasc. Comput. Tomogr.*, vol. 14, no. 2, pp. 168–176, 2020.

[21] B. A. Tama, S. Im, and S. Lee, "Improving an Intelligent Detection System for Coronary Heart Disease Using a Two-Tier Classifier Ensemble," *Biomed Res. Int.*, vol. 2020, 2020.

[22] S. A. Alaka *et al.*, "Functional Outcome Prediction in Ischemic Stroke: A Comparison of Machine Learning Algorithms and Regression Models," *Front. Neurol.*, vol. 11, no. August, pp. 1–9, 2020.

[23] M. Nilashi *et al.*, "Coronary Heart Disease Diagnosis Through Self-Organizing Map and Fuzzy Support Vector Machine with Incremental Updates," *Int. J. Fuzzy Syst.*, vol. 22, no. 4, pp. 1376–1388, 2020.

[24] J. H. Joloudari *et al.*, "Coronary artery disease diagnosis; ranking the significant features

using a random trees model," *Int. J. Environ. Res. Public Health*, vol. 17, no. 3, 2020.

[25] L. L. R. Rodrigues *et al.*, "Machine learning in coronary heart disease prediction: Structural equation modelling approach," *Cogent Eng.*, vol. 7, no. 1, 2020.

[26] J. Wang *et al.*, "A stacking-based model for non-invasive detection of coronary heart disease," *IEEE Access*, vol. 8, pp. 37124–37133, 2020.

[27] J. Jeyaranjani, T. Dhiliphan Rajkumar, and T. Ananth Kumar, "WITHDRAWN: Coronary heart disease diagnosis using the efficient ANN model," *Mater. Today Proc.*, no. January, 2021.

[28] H. E. S. Alsafi and O. N. Ocan, "A novel intelligent machine learning system for coronary heart disease diagnosis," *Appl. Nanosci.*, no. Ting 2017, 2021.

[29] P. Schratz, J. Muenchow, E. Iturritxa, J. Richter, and A. Brenning, "Hyperparameter tuning and performance assessment of statistical and machine-learning algorithms using spatial data," *Ecol. Modell.*, vol. 406, no. April 2018, pp. 109–120, 2019.

[30] O. N. Ucan, H. Emad, and S. Alsafi, "Cardiac Coronary Detection of Heart Signal with a Combined Approach of Multi-channel Deep Convolution Learning with Differential Evolution Algorithm," *submit*.

[31] L. Wu, G. Perin, and S. Picek, "I Choose You: Automated Hyperparameter Tuning for Deep Learning-based Side-channel Analysis," *IEEE Trans. Emerg. Top. Comput.*, no. Report 2020/1293, pp. 1–12, 2022.

[32] P. Kaul, D. Golovin, and G. Kochanski, "Hyperparameter tuning in Cloud Machine Learning Engine using Bayesian Optimization," pp. 1–8, 2019.

[33] Google Cloud, “Overview of hyperparameter tuning | AI Platform Training,” 2021. [Online]. Available: https://cloud.google.com/ai-platform/training/docs/hyperparameter-tuning-overview.

[34] M. Peker, “A decision support system to improve medical diagnosis using a combination of k-medoids clustering based attribute weighting and SVM,” *J. Med. Syst.*, vol. 40, no. 5, 2016.

[35] Z. Stawska, “SUPPORT VECTOR MACHINE IN GENDER RECOGNITION,” *Inf. Syst. Manag.*, vol. 6, no. 4, pp. 318–329, 2017.

[36] Anonim, “Support Vector Machines Introductory Overview,” 2020. [Online]. Available: https://docs.tibco.com/data-science/GUID-1D973B41-0C45-4873-97A1-D58828E542F4.html.

[37] E. Alickovic and A. Subasi, “Medical Decision Support System for Diagnosis of Heart Arrhythmia using DWT and Random Forests Classifier,” *J. Med. Syst.*, vol. 40, no. 4, pp. 1–12, 2016.

[38] X. Sun, X. Lin, S. Shen, and Z. Hu, “High-resolution remote sensing data classification over urban areas using random forest ensemble and fully connected conditional random field,” *ISPRS Int. J. Geo-Information*, vol. 6, no. 8, 2017.

[39] R. Ani, J. Jose, M. Wilson, and O. S. Deepa, “Modified rotation forest ensemble classifier for medical diagnosis in decision support systems,” in *Advances in Intelligent Systems and Computing*, 2018, vol. 564, pp. 137–146.

[40] M. A. Fauzan and H. Murfi, “The accuracy of XGBoost for insurance claim prediction,”

Int. J. Adv. Soft Comput. its Appl., vol. 10, no. 2, pp. 159–171, 2018.

[41] H.Alsafi, "A Simple XGBoost Tutorial Using the Iris Dataset -Python," vol. 12, 2019.

[42] L. Zhang and X. Zhao, "An adaptive video steganography based on intra-prediction mode and cost assignment," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2017, vol. 10082 LNCS, pp. 518–532.

[43] B. Klaus and P. Horn, *Building Machine Learning Systems with Python*. Mal: UMA, 2018.

[44] E. Upton and G. Halfacree, "An Introduction to Python," in *Raspberry Pi® User Guide*, 2017, pp. 153–179.

[45] L. Stein, *Machine Learning in Python®*. Mal: UMA, 2015.

[46] "Python for Machine Learning Libraries, ML Algorithms and Functionality | LITSLINK Blog." [Online]. Available: https://litslink.com/blog/python-for-machine-learning.

[47] Aized Amin Soofi and Arshad Awan, "Classification Techniques in Machine Learning: Applications and Issues," *J. Basic Appl. Sci.*, vol. 13, no. August, pp. 459–465, 2017.

[48] R. Konieczny and R. Idczak, "Mössbauer study of Fe-Re alloys prepared by mechanical alloying," *Hyperfine Interact.*, vol. 237, no. 1, pp. 1–8, 2016.

[49] U. Shruthi, V. Nagaveni, and B. K. Raghavendra, "A Review on Machine Learning Classification Techniques for Plant Disease Detection," in *2019 5th International Conference on Advanced Computing and Communication Systems, ICACCS 2019*, 2019, pp. 281–284.

[50] A. Sankari Karthiga, M. Safish Mary, and M. Yogasini, “Early Prediction of Heart Disease Using Decision Tree Algorithm.,” *Int. J. Adv. Res. Basic Eng. Sci. Technol.*, vol. 3, no. 3, pp. 1–16, 2017.

[51] F. Thabtah and D. Peebles, “A new machine learning model based on induction of rules for autism detection,” *Health Informatics J.*, vol. 26, no. 1, pp. 264–286, 2020.

[52] C. Ganteng, “Cardiovascular Study Dataset,” 2021. [Online]. Available: https://www.kaggle.com/christofel04/cardiovascular-study-dataset-predict-heart-disea.

[53] J. Gatto, R. Lanka, Y. Iwashita, and A. Stoica, “Single Sample Feature Importance: An Interpretable Algorithm for Low-Level Feature Analysis,” pp. 1–13, 2019.

[54] W. La Cava, C. Bauer, J. H. Moore, and S. A. Pendergrass, “Interpretation of machine learning predictions for patient outcomes in electronic health records,” *AMIA ... Annu. Symp. proceedings. AMIA Symp.*, vol. 2019, pp. 572–581, Mar. 2019.

[55] B. Schlegel and B. Sick, “Design and optimization of an autonomous feature selection pipeline for high dimensional, heterogeneous feature spaces,” in *2016 IEEE Symposium Series on Computational Intelligence, SSCI 2016*, 2017, no. Section V.

[56] H. Fujita and A. Selamat, *Advancing Technology Industrialization Through Intelligent Software Methodologies, Tools and Techniques*. 2019.

[57] L. Gautheron, A. Habrard, E. Morvant, and M. Sebban, “Metric learning from imbalanced data,” in *Proceedings - International Conference on Tools with Artificial Intelligence, ICTAI*, 2019, vol. 2019-Novem, pp. 923–930.

[58] L. Li, K. Jamieson, G. DeSalvo, A. Rostamizadeh, and A. Talwalkar, “Hyperband: A novel

bandit-based approach to hyperparameter optimization," *J. Mach. Learn. Res.*, vol. 18, pp. 1–52, 2018.

[59] F. Hutter, "Meta-learning," in *Studies in Computational Intelligence*, vol. 498, 2014, pp. 233–317.

[60] Y. Wang and X. Sherry Ni, "A XGBOOST RISK MODEL VIA FEATURE SELECTION AND BAYESIAN HYPER-PARAMETER OPTIMIZATION," *Int. J. Database Manag. Syst.*, vol. 11, no. 01, pp. 01–17, 2019.

[61] S. Putatunda and K. Rama, "A comparative analysis of hyperopt as against other approaches for hyper-parameter optimization of XGBoost," in *ACM International Conference Proceeding Series*, 2018, pp. 6–10.

[62] "Forest Cover Type Classification with XGBoost." [Online]. Available: https://hub.knime.com/knime/spaces/Examples/latest/04_Analytics/16_XGBoost/01_Classify_Forest_Covertypes_with_XGBoost.

[63] P. Probst, M. N. Wright, and A. L. Boulesteix, "Hyperparameters and tuning strategies for random forest," *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, vol. 9, no. 3. pp. 1–19, 2019.

[64] D. Sun, H. Wen, D. Wang, and J. Xu, "A random forest model of landslide susceptibility mapping based on hyperparameter optimization using Bayes algorithm," *Geomorphology*, vol. 362, p. 107201, 2020.

[65] D. Sun, J. Xu, H. Wen, and D. Wang, "Assessment of landslide susceptibility mapping based on Bayesian hyperparameter optimization: A comparison between logistic regression

and random forest," *Eng. Geol.*, vol. 281, no. December, p. 105972, 2021.

[66] L. Ali *et al.*, "An Optimized Stacked Support Vector Machines Based Expert System for the Effective Prediction of Heart Failure," *IEEE Access*, vol. 7, pp. 54007–54014, 2019.

[67] A. Rojas-Dominguez, L. C. Padierna, J. M. Carpio Valadez, H. J. Puga-Soberanes, and H. J. Fraire, "Optimal Hyper-Parameter Tuning of SVM Classifiers with Application to Medical Diagnosis," *IEEE Access*, vol. 6, pp. 7164–7176, 2017.

[68] S. Jadhav, S. Nalbalwar, and A. Ghatol, "Feature elimination based random subspace ensembles learning for ECG arrhythmia diagnosis," *Soft Comput.*, vol. 18, no. 3, pp. 579–587, 2014.