International Journal of Computer Applications (0975 - 8887)
Volume 183 - No.3, May 2021

Diagnosis of Cardiovascular Diseases using Artificial Intelligence Techniques: A Review

Tazeen Tasneem
Department of Computer Science & Engineering
Rajshahi University of Engineering & Technology
Rajshahi, Bangladesh

Mir Md. Jahangir Kabir
Department of Computer Science & Engineering
Rajshahi University of Engineering & Technology
Rajshahi, Bangladesh

Shuxiang Xu
Discipline of ICT, School of Technology
Environments and Design,
University of Tasmania, Australia

Tabeen Tasneem
Department of Computer Science & Engineering
Rajshahi University of Engineering & Technology
Rajshahi, Bangladesh

ABSTRACT

In the last couple of decades, many techniques have been introduced for medical support system. One alarming field in medical health care is cardiovascular disease as millions of deaths occur every year because of this. Thus, diagnosis of heart disease has always been one of the most important issues. For predicting and diagnosis of cardiovascular disease, skilled and experienced physicians are needed. As this is an era of technology, researchers have been proposed many algorithms and learning techniques for assisting the physicians. The aim of this research work is to thoroughly analyze these algorithms and methods. This article has explored the used datasets, feature selection techniques and missing value imputation methods, and finally compared their performances.

General Terms

Artificial Intelligence, Classification

Keywords

Cardiovascular disease, Feature selection, Missing value imputation, Artificial Neural Network, Classification

1. INTRODUCTION

Cardiovascular disease (CVD) is a broad term that may allude to any condition that has effects on heart [1]. According to the World Health Organization (WHO), one of the most dangerous diseases all around the world is CVD. As per the organization, around 17.9 million individuals died from these diseases in 2016 [2] which is 31% of all global deaths. CVD caused 256,800 deaths in Bangladesh in the year 2016. Like the low- and middle-income countries, CVD is also a threat to the developed countries as well. In the same year, 840,600 people died because of CVD in the United States of America [3].

Hyperlipidemia, stable angina, unstable angina, myocardial infarction are some diseases that are included in CVD [4–6]. Although many people have symptoms, most of the patients do not have any symptoms of these diseases before having a heart attack. So, these diseases are hard to diagnose in the primary stage. For the betterment of the patients, it demands early diagnosis of the diseases.

A lot of people either find it very difficult to control the risk factors that cause CVD or are completely oblivious to the fact that they are at high risk. Risk factors can be of two types - behavioral and physiological. Physiological factors are related to an individual's physical fitness that may include blood pressure, diabetes etc. Behavioral factors are something that are related to the habit of an individual such as smoking and unhealthy diet - and these kinds of behavioral factors can be changed by changing habits. However, not all factors can be changed, for example age and family history of an individual [1].

In early days, the patients had to depend only on the experience and expertise of the physicians and that used to take a lot of time. But with the advancement of artificial intelligence and machine learning, enormous models have been proposed for assisting the physicians to diagnose the diseases. Various techniques have been used to achieve this goal. Some of the strategies include data mining techniques for the early stage diagnosis of the diseases, some strategies use machine learning techniques as well [7–11].

Data mining as well as machine learning have diverse applications in medical field and healthcare, some of which are - detecting the causes of diseases, coming up with better medical solutions for the patients, identifying proficient treatment methods and many more [12]. Different techniques are applied to find out the hidden information, pattern, relation among the patients, their medical condition and treatments in a cogent manner. The techniques mainly include Support Vector Machine(SVM), naive Bayesian theorem, Artificial Neural Network (ANN), clustering, association rule mining, decision tree, fuzzy logic and sometimes hybrid methods. As CVD is not a single disease rather a group of diseases - it covers a

vast area in medical science. For this reason, a lot of research works have been conducted in this field. The objectives of this paper are -

(i) Focus attention on the works that have been done so far in this field.
(ii) Present a comparative analysis of the papers.
(iii) Finally, a future research direction of this work is provided at the end of the paper.

As mentioned earlier, there can be many types of diseases that are part of CVD and each of the diseases has different risk factors. Thus, in section 2, various CVDs have been introduced and in the next section, section 3, the risk factors of the diseases have been described. Section 4 discusses the strategy based on which the articles have been selected for this review process. Section 5 gives detailed information about the dataset used by those selected articles. Section 6 and Section 7 discuss the feature selection strategies and missing value imputation techniques, respectively. Different types of base classifiers have been canvassed in Section 8. Section 10 compares the performance of all the proposed models, Section 11 focuses on some challenges and pitfalls of AI technology and gives directions about some future works, and finally Section 12 concludes this research work.

2. TYPOLOGY OF HEART DISEASE

Often, there is a misconception that CVD is a single disease. But it is not a single disease rather a bunch of critical diseases that affect an individual's heart, blood vessels and cardiac muscle. These include - coronry heart disease, angina, stroke, rheumatic heart disease and so on [13]. Every year hundreds of thousands of people die because of these diseases. Figure 1 shows that, in the year of 2016 CAD was the primary cause of death. Stroke is the second major reason for human death followed by high blood pressure, heart failure and many more [14].

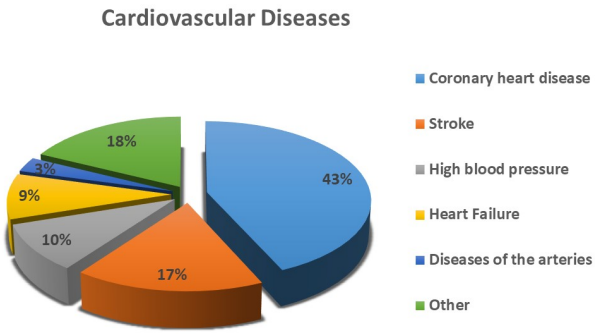

Fig. 1: Deaths caused by different cardiovascular diseases

Coronary Heart Disease. Human heart has several arteries, commonly known as the coronary artery, that provide oxygen, blood and nutrients to the heart. When these arteries become too narrow, coronary heart disease starts developing. When the cholesterol starts to accumulate on the wall of arteries, the arteries start becoming narrow. So gradually the blood flow through the arteries reduces. As a result, the heart may not get sufficient amount of oxygen which is necessary for its working. Eventually, it may lead to heart attack [15].

Angina. Coronary artery disease has several symptoms, angina is one of them. Angina is not a disease but a symptom. It can be described as pain or cramp in the heart caused for not having enough oxygen [15]. Rather than *pain* the word *discomfort* can explain angina more accurately because not all people sense angina as painful. Most of them sense angina as painful when it is severe. Angina can be categorized as stable, unstable and variant angina. Angina itself is not a life-threatening condition, but it strongly indicates that a person may have heart disease.

Stroke. Stroke is a state of the brain which occurs if the blood flow to and within the brain stops because of blockage or rupture in the blood vessel. With the reducing flow of blood, the flow of oxygen and nutrients toward the brain also decreases. As a result, the brain cells start dying within minutes. The part of brain where all these happen - stops working and the body parts that are controlled by that part of the brain, also cease to work [15].

Rheumatic Heart Disease. This disease occurs when acute rheumatic fever damages one or more heart valves. *Damaged heart valves* refers to a scenario when the blood flows backward through the valves as the valves are not closing properly or the blood may not flow at all as the valves are not opening [15]. This generally occurs 10-20 years after the primary illness and not everybody who had rheumatic fever, ends up with rheumatic heart disease [15]. Although this disease has been eradicated more or less in Europe and North America, it is prominent in Central and South Asia, the Middle East, Africa, the South Pacific along with some other developed countries [16].

3. RISK FACTORS

From section 2 it is clear that, cardiovascular disease is a cluster of diseases. Heart disease is a subset of cardiovascular disease. Stroke is another example that supports this statement. It is a cardiovascular disease that does not affect heart but brain. As there are many variants of CVD, there are numerous risk factors for it [17]. Often the risk factors are very complex as some of the risk factors are correlated. So for diagnosing CVD, the risk factors have to be studied thoroughly.
The risk factors responsible for CVD can be grouped as non-modifiable risk factors and modifiable risk factors [17]. As the name suggests, if a person has non-modifiable risk factors, there is nothing in his grasp to avoid CVD. Examples include age, family history, ethnicity and so on [18]. Unlike non-modifiable risk factors, a person can change modifiable risk factors to elude CVD. These include smoking, obesity, high cholesterol, high lipid on blood and many more.

3.1 Non-modifiable Risk Factors

Family History. Human nature is controlled by the genetic elements which people get from their parents. Similarly, there are some genetic elements which define if a person would have CVD from his or her family tree. Normally, this occurs if a person's first degree relative has CVD in early age. This means if that person's father or brother had CVD before turning 55, or mother or sister

Table 1. : List of papers along with their IDs

Paper ID	Author	Paper ID	Author	Paper ID	Author	Paper ID	Author
PID01	Son et al. [19]	PID19	Das et al. [20]	PID37	Xu et al. [21]	PID54	Usman et al. [22]
PID02	Orphanou et al. [23]	PID20	Nguyen et al. [24]	PID38	Latha et al. [25]	PID55	Chen et al. [26]
PID03	Comak et al. [27]	PID21	Muthukaruppan [28]	PID39	Das et al. [29]	PID56	Amin et al. [30]
PID04	Polat et al. [31]	PID22	Nahar et al. [32]	PID40	Escamila et al. [33]	PID57	Nahato et al. [34]
PID05	Chen et al. [35]	PID23	Sengur [36]	PID41	Christo et al. [37]	PID58	Abdar et al. [38]
PID06	Vivekanandan et al. [39]	PID24	Sengur et al. [40]	PID42	Khourdifi et al. [41]	PID59	Shah et al. [42]
PID07	Vivekanandan [43]	PID25	Turkoglu [44]	PID43	zen et al. [45]	PID60	Khan et al. [46]
PID08	Anooj [47]	PID26	Yan et al. [48]	PID44	Shah et al. [49]	PID61	Tama et al. [50]
PID09	Omurlu et al. [51]	PID27	Polat et al. [52]	PID45	Hedeshi et al. [53]	PID62	Sudha [54]
PID10	Samuel et al. [55]	PID28	Senthil [56]	PID46	Gestel et al [57]	PID63	Peker et al. [58]
PID11	Babaolu [59]	PID29	Alizadehsani [60]	PID47	Babu et al. [61]	PID64	Li et al. [62]
PID12	Sekar et al. [63]	PID30	Arabasadi [64]	PID48	Huang et al. [65]	PID65	Shilaskar et al. [66]
PID13	Alneamy [67]	PID31	Olaniyi et al. [68]	PID49	Inbarani et al [69]	PID66	Shao et al. [70]
PID14	Haq et al. [71]	PID32	Das et al. [1]	PID50	Lee et al. [72]	PID67	Nilashi et al. [73]
PID15	Liu et al. [74]	PID33	zen et al. [75]	PID51	Kim et al. [76]	PID68	Alneamy et al. [77]
PID16	Kahramanli et al. [78]	PID34	Nguyen et al. [79]	PID52	Ali et al. [80]	PID69	Le et al. [81]
PID17	Abdar et al. [82]	PID35	Nguyen et al. [83]	PID53	Dutta et al. [84]	PID70	Mustaqeem [85]
PID18	Avci [86]	PID36	Beheshti et al. [87]				

had CVD before turning 65, it is more likely for the person to suffer from CVD [15]. Apart from this, high cholesterol, high blood pressure and type II diabetics in family history also increases one's possibility of having CVD [15]. If one has family history regarding CVD, it does not indicate that CVD is unavoidable for him but indicates that it is more likely for him to have a CVD.

Age. Aging is a bitter truth of human life. The risk of developing CVD increases with the increase of age. Older people (commonly over 50) are more likely to have CVD [15, 88].

Ethnicity. Descendants of some particular region like South Asia, Africa and Caribbean area are more likely to be at a greater risk of developing CVD. The reason behind this is yet to know [18, 88, 89].

3.2 Modifiable Risk Factors

Smoking. Smoking tobacco is one of the most prominent reasons for CVD. This independent risk factor causes sudden deaths of the patients who are suffering from coronary heart disease. Along with other risk factors, smoking increases the risk of coronary heart disease significantly. Nicotine in cigarette is responsible for raising blood pressure and the carbon monoxide in it reduces the amount of oxygen in blood stream. Heart and blood vessels can be damaged badly by smoking, which in turn boost up the risk of heart attack [15].

Cholesterol. Cholesterol is a fat-like, waxy stuff generally made by the liver and found in certain foods. Normally, the amount of cholesterol liver makes is sufficient for human body, but when an individual intakes more cholesterol from food than the body needs, it builds up plaque on the arteries. In consequence, the blood flow to the heart, kidney, brain and other parts of the body decreases. Cholesterol can be broken down into two categories - low-density lipoprotein (LDL) cholesterol and high-density lipoprotein (HDL) cholesterol. LDL cholesterol is considered to be harmful for human health as it causes different CVDs by creating plaque on the artery walls. On the other hand, HDL cholesterol is good for health as higher levels of this cholesterol yields safeguarding against CVD [15].

High Blood Pressure. It escalates the workload of heart, thus thickens the muscle of the heart and also stiffens it. As the heart muscle stiffens, heart has to work hard to pump blood and this overload causes heart attack, stroke, heart failure and kidney failure most often. The risk of these diseases increases even more when high blood pressure is present followed by obesity, high

blood cholesterol levels, smoking or diabetes [15].

Obesity. Obese individuals or in other word people with excess body fat - specially fat around the waist - have more possibility of developing CVD. They are also at risk even if other risk factors are absent. Obesity increases the workload of heart, raises blood pressure and LDL cholesterol level, and also lowers the HDL cholesterol level. All together, these can boost up the risk of CVD [15].

Diabetes. Human body needs energy for doing each and every work in day to day life. They get the energy from glucose. It is necessary to move glucose from the food to body cells for energy and a hormone called insulin helps the job to get done. If someone has diabetes, his/her body neither makes sufficient insulin nor can it use its own insulin, or both. Diabetes is a condition that makes the sugar build up in the blood. High level of sugar causes damage to the artery walls and also paves the way of building up fatty deposits. When these fatty substances heap up in the arteries, they may lead to coronary heart disease [15].

Other Risk Factors. There are many more risk factors for CVDs. Having an unhealthy, poor nutrient diet seriously affects the heart by increasing blood pressure, LDL cholesterol and decreasing HDL cholesterol. It also affects the sugar and salt levels in the blood. Drinking too much alcohol often leads to CVD by raising blood pressure. So if anyone wants to drink alcohol, it should be moderate. Physical inactivity leads to heart diseases as well and it is the root of many other risk factors of CVD. Example includes - overweight, high blood pressure, high cholesterol, diabetes and many more. A person who is physically active has many benefits such as - reduced cholesterol levels, decreased obesity, well-functioned heart to pump blood around the body [15].

4. ARTICLE SELECTION STRATEGY

The papers were found out after a thorough search in IEEE Xplore, ScienceDirect, ACM Digital Library, Hindawi, Springer and PubMed. These databases provide hundreds of thousands of journals and conference papers. Only the articles which contain some specific word in title, abstract or keyword segment were selected for further evaluation. The databases were searched using words *diagnosis/prognosis/classification* of *cardiovascular disease/coronary artery disease/heart valve disease/heart failure* using *machine learning/artificial intelligence/data mining/neural network*. Numerous articles popped up having these words in title or abstract or keyword segment. From these numerous articles only those articles were selected which -

(1) were published in Q1, Q2 or Q3 ranked journals, or A*, A or B ranked conferences. A few papers that did not belong to these categories were selected because of their highly influential content.
(2) propose new methods to diagnosis/prognosis/classify CVD/Coronary artery disease(CAD)/heart valve disease/heart failure.
(3) implement existing methods or technologies for diagnosing CVD/CAD/heart valve disease/heart failure.

Some articles were promising but could not be selected for some issues.

(1) Papers containing no classification/diagnosis methods but only medical issues (i.e why these diseases are increasing day by day, what is the reason behind these diseases and so on)
(2) Review papers were not considered.
(3) Duplicate papers were discarded

Table 1 demonstrates the selected papers. A large number of research articles have been selected for this reviewing work. Each selected research article has been assigned an unique ID. Figure 2

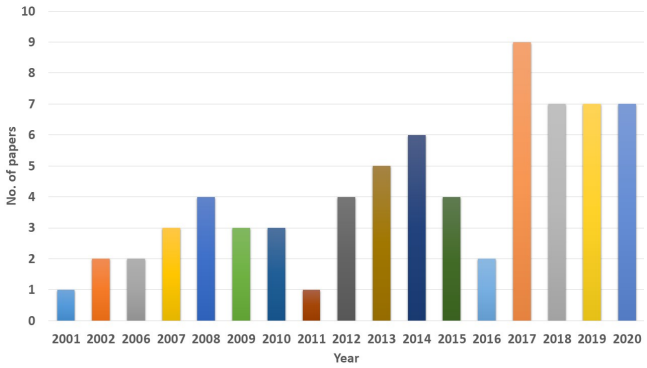

Fig. 2: Number of papers in each year from the timeline 2001 to 2020

provides comprehensive information about these selected articles. AI based CHD detection has been commenced long ago. But effectively this started from 2001. There is barely any research work until 2008. There is a substantial growth in paper from 2008.

5. DATASETS

In total, 20 datasets have been considered which were collected from different sources. Some datasets were collected from online dataset repositories and some other were collected from different hospitals and medicals of different countries, for example Turkey, South Korea, China, Croatia, Pakistan and so on. Only the datasets which have been used by authors for the diagnosis of CVD, CAD and valvular heart disease are considered here. To make the review process more clear, each dataset is given an unique ID, which has been shown in table 2. So in the remaining of this paper, datasets' ID will be used to describe any of these datasets. The number of instances and attributes of these datasets differ considerably. Table 3 has enlisted each and every dataset along with their corresponding attributes and instances. *Dataset 06*, the smallest dataset, contains only 84 instances and 9 attributes, used for diagnosis of CVD. On the other hand, the largest dataset, *Dataset 17*, has 36 attributes and 37079 instances.

Dataset 01 is the most used dataset for heart disease diagnosis, named *Cleveland Dataset*. This dataset has been used in thirteen papers for diagnosis of heart disease. As a matter of fact, this dataset is a subset of 14 attributes of *Heart Disease Dataset* available at [90]. The original dataset contains 75 attributes. Cleveland dataset attributes are of three types - categorical, integer and real, and contains some missing values. Among its 303 instances, 297 instances are free of missing values. As the authors were not willing to handle missing values, they used 297 instances instead of 303 instances [1, 55, 67, 71].

Dataset 02 was collected from Southwest Hospital and Dajiang Hospital located in China [48]. This dataset has total 352 cases and several missing values, so the authors adopted the substituting

Table 2. : Assigned Dataset ID

Dataset ID	Paper ID
Dataset 01	PID06, PID07, PID08, PID10, PID13, PID14, PID16, PID20, PID21, PID22, PID27, PID28, PID32, PID38, PID40, PID41, PID42, PID44, PID45, PID47, PID48, PID50, PID52, PID56, PID58, PID59, PID60, PID61, PID62, PID65, PID66, PID67, PID68, PID69
Dataset 02	PID26
Dataset 03	PID17, PID29, PID30, PID54, PID58, PID60, PID61
Dataset 04	PID03, PID18, PID19, PID23, PID24, PID25
Dataset 05	PID20, PID44, PID49, PID65
Dataset 06	PID05
Dataset 07	PID09
Dataset 08	PID11
Dataset 09	PID15, PID31, PID39, PID41, PID43, PID46, PID54, PID56, PID57, PID61, PID63, PID67
Dataset 10	PID12
Dataset 11	PID08, PID21, PID30, PID40, PID44, PID45, PID54, PID59, PID61, PID65, PID66
Dataset 12	PID08, PID30, PID44, PID45, PID59, PID65, PID66
Dataset 13	PID30, PID45, PID65, PID66
Dataset 14	PID02
Dataset 15	PID37
Dataset 16	PID51
Dataset 17	PID53
Dataset 18	PID54
Dataset 19	PID54, PID64
Dataset 21	PID70

mean method to handle this. This dataset has been used to classify five heart diseases - CAD, Rheumatic valvular heart disease, Hypertension, Chronic cor pulmonale, Congenital heart disease which has 86, 82, 71, 60 and 53 cases, respectively.

Another dataset, *Dataset 03*, available at UCI machine learning repository under the name *Z-Alizadeh Sani dataset* has 303 instances among which 216 CAD affected patients and 87 normal healthy people [90]. There are 54 attributes(integer and real type) and no missing values. This dataset has been used by the authors for diagnosing coronary artery/heart disease [60, 64, 82].

Dataset 04 was collected from the Cardiology Department of Firat Medical Center, Turkey. In this dataset, every single audio Doppler Heart Sound signals were obtained from the Acuson Sequoia 512 Model Doppler Ultrasound. The Doppler ultrasonic flow transducer was used in a continuous operating mode of 2 MHz. For collecting these signals, each person's aortic and mitral valves were studied. Among 215 samples (132 men and 83 women), 56 people had normal aortic valves and 54 people had abnormal aortic valves, and 39 people had normal mitral valves and 66 people had abnormal mitral valves.

Table 3. : Dataset information about attributes and instances

Dataset ID	Attributes	Instances
Dataset 01	13	303
Dataset 02	40	352
Dataset 03	54	303
Dataset 04	91	215
Dataset 05	44	267
Dataset 06	9	84
Dataset 07	8	1254
Dataset 08	23	480
Dataset 09	13	270
Dataset 10	38	2267
Dataset 11	13	294
Dataset 12	13	123
Dataset 13	13	200
Dataset 14	-	1427
Dataset 15	8	7360
Dataset 16	16	4146
Dataset 17	36	37079
Dataset 18	7	209
Dataset 19	12	131
Dataset 20	18	1070

Another dataset that can be found in UCI machine learning repository is *Dataset 05*, named SPECTF Dataset [90]. Single Proton Emission Computed Tomography (SPECTF) images were processed to extract 44 continuous attributes or features that best describes the original images. All the continuous attributes have integer values ranging 0 to 100. Among 267 instances, 55 instances belong to class 0 and 221 instances belong to class 1, and none contain any missing values.

Dataset 06, the smallest dataset, contains only 84 samples among which 60 samples were from healthy individuals and 24 samples were from cardiovascular patients. The dataset does not have missing values.

The samples of *Dataset 07* were collected from the Cardiology Clinic of Trakya University Medical Faculty, Turkey between the time period of January, 2002 and February, 2003. Among 1254 samples, coronary artery disease was present in 865 samples and absent in 380 samples. There is no missing value in the dataset. The authors did not mention the type of the attributes.

In *Dataset 08*, 480 samples were gathered from the patients who underwent EST and coronary angiography. Two experienced cardiologists evaluated the EST results and angiographic images. The authors did not mention from where they collected these samples or if there is any missing values [59].

Dataset 09 is another dataset that can be found in UCI machine learning repository under the name Statlog(Heart) Dataset [90]. The dataset has 270 instances and 13 attributes; the types of the

attributes are - real, ordered, binary and nominal. No missing value has been found in this dataset.
Dataset 10 contains 38 attributes and 2267 samples. The samples of cardiovascular patients were collected from two universities of China and healthy samples were gathered randomly. There is no missing value in the dataset. The dataset has been used to diagnose CVD.
Dataset 11, Dataset 12 and Dataset 13 are the three datasets that have been used in Heart Disease Dataset along with *Dataset 01*. The name of these three datasets are - Hungarian Dataset, Switzerland Dataset and VA long beach dataset, respectively and all these three datasets can be found in UCI machine learning repository. These datasets are not as popular as Dataset 01 (Cleveland Dataset) because they contain missing values in huge amount. Datasets with this much missing values cannot be used to train any network. So, imputation is needed for these datasets.
Dataset 14, named Stulong, was collected from a diachronic study of atherosclerosis primary prevention. Table 3 does not contain the number of instances of this dataset because the authors did not mention it. After feature selection process, they found 16 features responsible for CVD. To impute missing values, missforest method has been used.
Dataset 15 has been collected from the hospitals located in north and south China. This is one of the largest datasets used by the authors containing 7360 instances. The dataset had some missing values which is very much normal in case of this type of large datasets; the authors have just discarded those instances which had missing values. This dataset has been used for the classification of presence or absence of coronary artery disease.
Another dataset which has been created by taking data from survey is *Dataset 16*. The data for this dataset have been collected from the 6th Korea National Health and Nutrition Examination Survey (KNHANES-VI). The authors have used this dataset for finding out from which coronary artery disease an individual is suffering. Among 4146 samples, 3031 had low and 1115 had high CVD risk.
Dataset 17 is the largest dataset containing 37079 instances and 36 attributes. It is a highly imbalanced dataset used for predicting the frequency of coronary artery disease. The data were collected from National Health and Nutritional Examination Survey (NHANES) in the time period from 1999-2000 to 2015-2016. It contains 35,779 instances of non-CAD people and 1300 instances of CAD patients which makes the dataset imbalanced. The aim of the authors was to propose a model which can generate highly accurate result trained by imbalanced dataset; so they needed a highly imbalanced dataset and so used this one.
Another dataset, available in online dataset repository, is *Dataset 18*, named Eric. This dataset contains 7 attributes and 209 instances and has been used for classifying normal people and those people who have risk of suffering from heart diseases.
Dataset 19 is a dataset that is available in UCI machine learning repository like some other datasets. It contains 132 instances and 12 attributes. Some of the instances contain missing values. This dataset has been used in two articles, PID54 and PID64, and in both articles, the authors have used this dataset for heart disease classification.
Dataset 20 was collected from a hospital located in Pakistan. The patients who were admitted to that hospital for heart disease and also the people who went the hospital for regular check-up were considered to create this dataset. This dataset has 19 attributes and 1070 instances and contains some missing values.
Some datasets have common attributes like age, sex, chest pain and so on. For example, Cleveland dataset, Hungarian dataset, Switzerland dataset, Statlog dataset and some other datasets contain same features. On the contrary, some datasets have unique features such as Z-Alizadeh Sani Data Set contains some features like current smoker, ex-smoker and many more. However, some datasets contain many irrelevant and redundant features which make the training process long. So, feature selection is needed for these datasets. Another problem is that, some of the datasets have missing values. Missing values sometimes can create huge burden if they are not handled properly. In some papers, authors have just discarded the samples containing missing values which may affect the classification greatly.

6. FEATURE SELECTION

The datasets that have been used by the authors for diagnosing different cardiovascular diseases, often contain some features or attributes that are of no use for the diagnosis job. Moreover, some datasets contain redundant attributes. These counterproductive as well as redundant features have negative influence as they increase the diagnosing time i.e, training time and also make the model overfit. So, having many features that are useless, irrelevant and redundant can be very burdensome [91, 92].
Feature selection is a process that removes the unnecessary, irrelevant (or partially relevant) and redundant features/attributes from a dataset and makes a subset of the original dataset that is useful for pattern mining and knowledge discovery. As the subset contains only relevant and useful attributes, it helps to reduce training time as well as prevents the model to be overfit for unseen samples [93]. Feature selection algorithms do not amend the representation of data, they just create a subset of attributes from the dataset [91]. If the feature selection algorithm is capable of finding a suitable subset of features, then it can make the model perform better and acquire greater accuracy than it had before [94]. For this benefit, authors opt for feature selection algorithms before training their model.
Feature selection algorithms select features automatically that contribute the most to predict the output level. These algorithms give scores to each and every feature according to the importance to predict the output level. Then the best features are kept while discarding others. No irrelevant or partially relevant features remain present after feature selection. The model is trained based on the most relevant features and so gets higher accuracy, less training time and no over fitting.
Not all papers that have been considered in this literature, used feature selection methods. There is an obvious reason behind this. A majority of authors who had been working on cardiovascular diseases (absence or presence of CVD or finding out the risk factors of CVD), considered Cleveland Heart Dataset (Dataset 01) which is already reprocessed; that means, there is an available version of this dataset where all unnecessary, irrelevant or partially relevant features are discarded. Not only Dataset 01 but also some other datasets, used by a number of authors, is reprocessed. As these authors had already been using reprocessed datasets, they did not need any feature selection method.
The authors who had worked on raw or not processed datasets, and also the authors who were willing to find risk factors or more interrelated features, used feature selection methods. Table 4 shows some feature selection methods used in different papers. From this table, it can be easily observed that only few papers have used just one feature selection technique before applying the updated dataset to any classifier. Most of authors have used hybrid feature selection techniques where sometimes they combined two or more feature selection algorithms, and sometimes they have used feature

optimization algorithms and/or feature weighting algorithms along with feature selection techniques.
A bunch of authors had worked on classification of heart valve disease. For this, they considered Doppler Heart Sound (DHS). So, instead of feature selection method, they had to use feature extraction methods on DHS signals. For feature extraction - discrete wavelet transform and wavelet entropy were used in PID18; wavelet decomposition and wavelet entropy were used in PID19; wavelet decomposition, STFT(short time furier transformation) and after these two, wavelet entropy was used in PID23, PID24 and PID25.

7. HANDLING MISSING VALUES

Missing value means an unrecorded data of a feature for a sample in a dataset. It is an obvious problem that occurs in almost all datasets. This is mentioned as *problem* because a dataset containing missing values can dwindle the analytical power of a system, and can generate distorted output/classification which leads to inoperative conclusions. So, researchers need to handle missing values to avoid these problems [95, 96].
Some of the datasets, that have been used in the articles mentioned in Table 1, contain missing values. To get rid of this, some authors just simply discarded the samples which contains missing values. But, not all time, this is a solution. If any dataset contains a huge amount of missing values, they cannot be just discarded because this will lead to discarding a lot of samples which is not feasible as other attribute values of those samples might be helpful for training the system. Again, that dataset cannot be used for training any system because a system cannot be trained using a dataset containing that much missing values. In this case, imputation of missing value is used. Some of the authors have used missing value imputation techniques.
Table 5 shows the techniques used by the authors to handle missing values. The articles that are not included in this table either just discarded the instances containing missing values or have not mentioned how they dealt with this problem. Majority of the authors have chosen the simplest way - discarded the missing value contained instances. The reason behind this is Dataset 01 is the most used dataset and this dataset contains only six samples with missing values. So this is feasible to discard these samples. But for other datasets, corresponding authors have used several techniques for the imputation of missing values.

8. METHODOLOGY

The articles considered for this work, used several supervised and unsupervised learning algorithms for classifying CVD, CAD, heart valve disease and so on. Each and every author has worked with some existing classifiers and fine-tuned them to work better on these diseases. Many of them have also used some hybrid methods combining two or more existing methods. The following part of this section has described some of the existing classification algorithms.

8.1 ANN and MLP

The most used predictor that has been used for classification is Artificial Neural Network(ANN). It was originated to replicate the way human brain works to evaluate information [97]. ANN is a supervised machine learning technique that amalgamates artificial neurons for processing data to find out necessary information from them [98]. It has adjustable fine-tuning operator (weight) that changes with external and/or internal information that circulates through the network. ANN has three main components - input layer, hidden layer and output layer. Input layer takes external data values and transmits them forward by assigning random weights which get updated in next layer(s). There can be several hidden layers between input and output layer depending on the need. This layer is optional; not all ANN has this layer. The sum of the products inputs and weights of each neuron are calculated and if the calculated value satisfies a threshold value, a neuron fires output which may work as input of next neuron. After such calculations in each hidden unit of each hidden layer, output layer finally outputs the predicted class. As it has self-learning capabilities, the more data are fed to ANN the more accurate becomes its classification. Figure 3 shows an Artificial Neural Network.

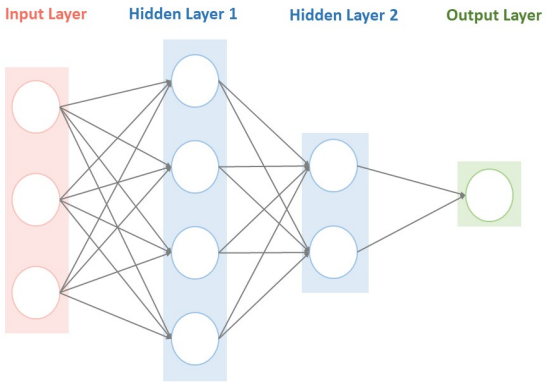

Fig. 3: Example of an ANN which have 1 input layer, 2 hidden layers and 1 output layer

Depending on the architecture, mathematical operations and required parameters ANN can be of several types - feedforward neural network, radial basis function neural network, Kohonen self-organizing neural network, convolutional neural network and many others. Multilayer Perceptron (MLP) is an extension of feedforward neural network. Feedforward neural network is the simplest form of ANN where it can have several hidden layers or no hidden layer at all. But in case of MLP, it contains at least one hidden layer. It can learn from non-linear functions and often trained by backpropagation algorithm.

8.2 SVM

Support Vector Machine (SVM) is a supervised binary classifier. It can be used for both classification and regression. SVM does the classification job by determining a decision boundary or hyperplane. The hyperplane is not influenced by all data points but by support vectors. Support vectors are nothing but the co-ordinates of single observation. The hyperplane is measured by keeping largest margin possible from the support vectors of both classes [99]. SVM classifies N-dimensional data objects by determining a single hyperplane. Here, N is the number of input features. Initially not all N-dimensional data objects can be separated using a single hyperplane but if the N-dimensional data objects are mapped to a higher-dimension they become separable by a single hyperplane. SVM follows this technique and keeps increasing the dimension until the objects are separable.

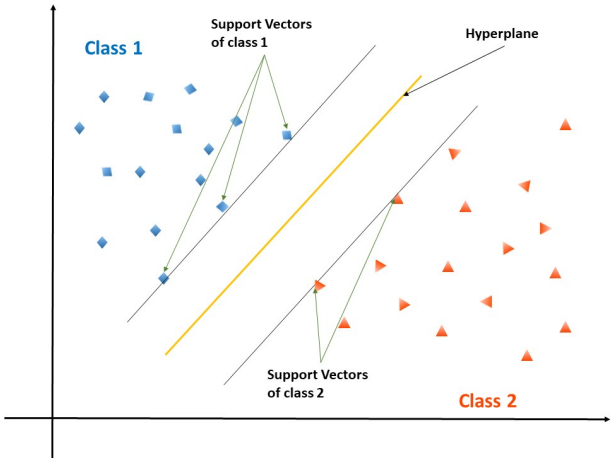

Fig. 4: A binary SVM classifying two classes using a hyperplane

8.3 K-NN

K-Nearest Neighbour (K-NN) is the simplest supervised learning method used for both classification and regression but mostly used for classification problems. It is a non-parametric learning algorithm and depends only on memory. When a new data point is considered for predicting its class label, this algorithm uses similarity distance measure [100]. At first k nearest neighbours are selected

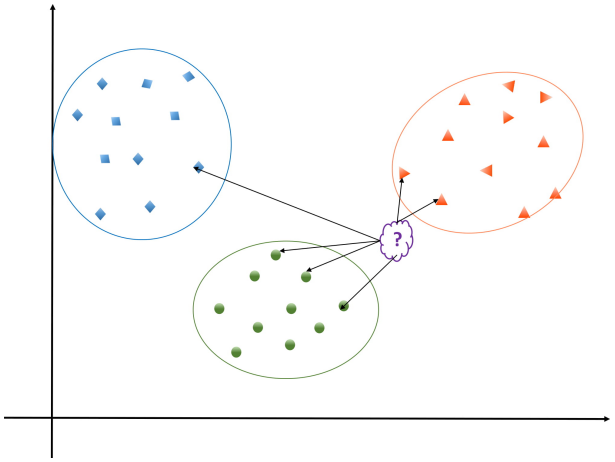

Fig. 5: K-Nearest Neighbour classifier

from the data point and votes are taken from that neighbouring data points measured by a distance function. Euclidean, Manhattan, Minkowski distance measures are used for continuous features and Hamming distance measure is used for categorical feature. k can be any positive integer value. The class label which gets the most vote, becomes the label of that data point [40].

8.4 Naive Bayes Classifier

Naive Bayes classifier is a supervised probabilistic classifier based on Bayes' theorem. The central objective of this classifier is the conditional independence of each feature present in a dataset i.e. this classifier assumes that the presence of an feature is not dependent on the presence of any other feature in a dataset [71]. For a feature vector, it calculates the posterior probability from likelihood, class prior probability and evidence prior probability. The class which has the largest posterior probability, will be assigned as class label to the feature vector [101]. The formula for calculating posterior probability -

$$\text{P(outcome|evidence)} = \frac{P(evidence|outcome)P(outcome)}{P(evidence)}$$

Where,
P(outcome|evidence) = posterior probability
P(evidence|outcome) = probability of likelihood of evidence
P(outcome) = class prior probability
P(evidence) = evidence prior probability

8.5 Decision Tree

Decision tree is a supervised learning algorithm which is very convenient to show the paths that lead to various possible outcomes. It is comprise of several nodes and branches. Decision tree is like any biological tree with an exception - the root is in the top and leaves are in the bottom. Nodes of a decision tree represent conditions. Branches coming out from a node represent different decisions. Finally, leaf nodes represent the outcome of the entire path [102]. Because of this simplicity, decision trees are widely used for classification and regression problems.
There are several variants of decision trees. ID3 (Iterative Dichotomiser 3) creates a multiway tree but it is designed only for categorical features. The tree is allowed to grow to maximum size and then pruning is generally applied to reduce the size of the tree and forestall overfitting for unseen data. C4.5 is a descendant of ID3 and removes the limitation of features being only categorical. It generates a ruleset from the output of ID3 and then accuracy of each rule is evaluated. Another version, C5.0, generates smaller ruleset than C4.5 and hence use less memory. CART, similar to C4.5, supports numerical features and does not generate any ruleset.

8.6 Fuzzy Logic System

Computer can understand only binary numbers and so binary system takes some precise input and produces only two outputs; either TRUE(1) or FALSE(0). This resembles human's decision YES and NO. But in real world, not all decisions can be only YES and NO. Here comes the Fuzzy Logic System, introduced by Zadeh, output of which includes a possible range between 1 and 0 [103]. Fuzzy logic system produces sustainable yet specific output with respect to vague, distorted, or ambiguous input. It has four main parts in its architecture - fuzzification module, rule base, inference engine and defuzzification module.

8.7 Model Infographics

Above 95% of heart disease detection papers considered for this research have used supervised classification algorithms. Only a few papers reported application of unsupervised algorithms such as clustering. Altogether lots of algorithms have been used for detection of different heart diseases. Every single author has used one or more base classifiers in accordance with some other algorithms to do the classification job more accurately. Many authors have suggested some new variants of some algorithms for this.
Figure 6 shows different supervised and unsupervised algorithms that have been applied on different datasets. It can be observed that most of the techniques have been applied on Dataset 01, and the

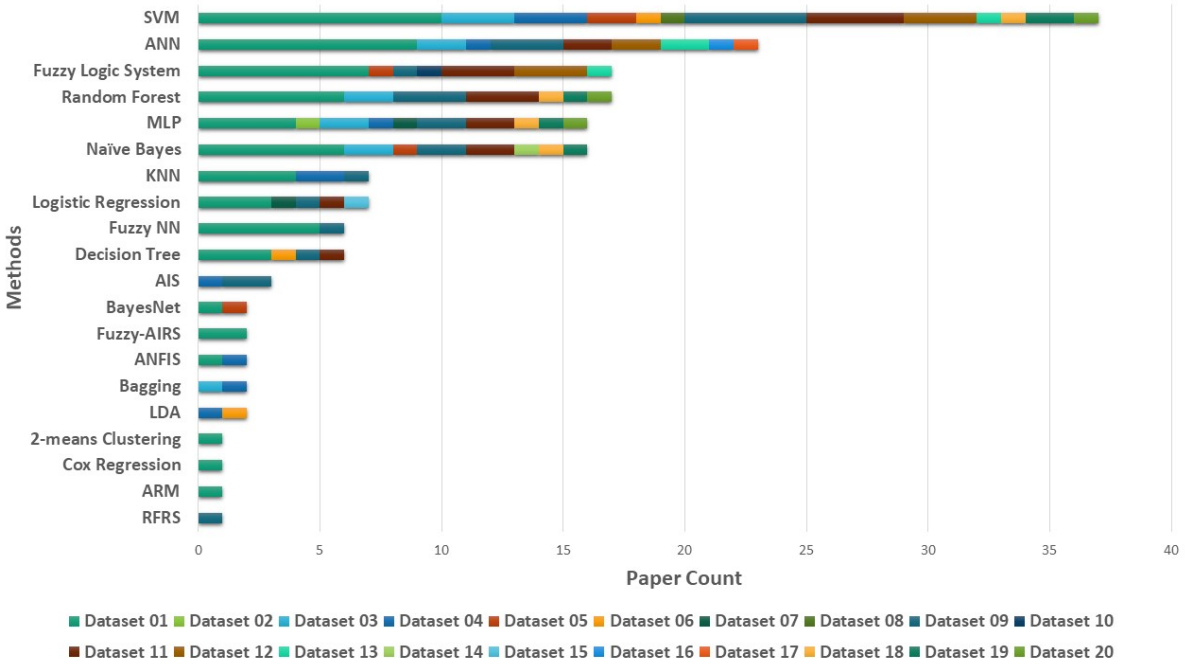

Fig. 6: Datasets that were selected for applying different techniques for heart disease detection

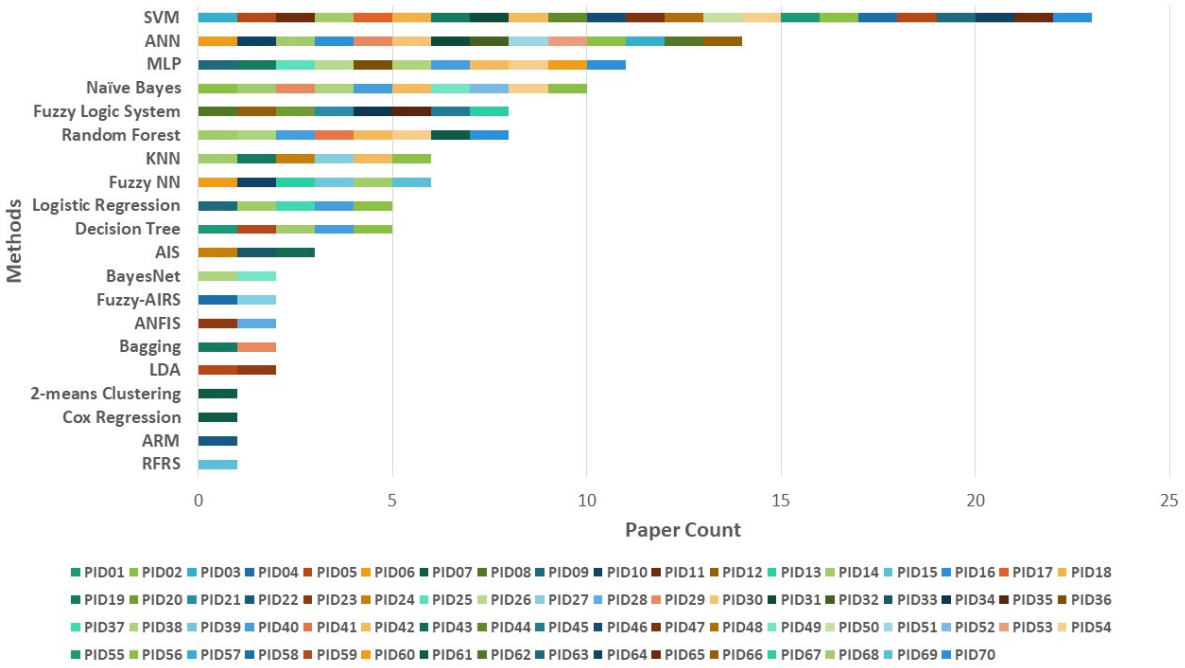

Fig. 7: Most commonly used AI techniques for heart disease ditection

first two most used techniques are SVM and ANN. Fuzzy Logic System and Random Forest are the next two most used algorithms that have been applied on different datasets. Figure 7 shows the statistics from a different angle. It shows how many papers have used a single algorithm for classifying different types of heart diseases. In essence, SVM and ANN are the mostly used algorithms that have been vastly applied on different datasets as well as used by several papers.

Over the year, the popularity of different methods changes. From Figure 8, it can be observed that SVM was applied for the first time in 2001 and has been used as a top classifier till today. One year after SVM, in 2002 MLP was considered as base classifier for the first time and it is one of those classifiers which are still being used as base classifier. Figure 8 also shows that, in 2007 several algorithms were applied initially in this classification matter and this cycle of using different techniques for this particular problem keeps going henceforth till today. Some classifiers were popular for a certain period but after that, they lost their popularity. For example ANFIS and AIS were used in the early years as classifier but after 2009 none of these algorithms have been used.

9. PERFORMANCE EVALUATION METRICS

Building a predictive model is not the main goal of any author. A predictive model has to be robust and give high performance for unseen samples as well. Hence and accordingly, a model has to be evaluated before predicting real world unseen data. There are dozens of evaluation metrics for this job. The evaluation process is often accomplished by splitting a dataset into training set and testing set, and applying these metrics on both of them.
The most popular evaluation metric that has been used in almost all articles is accuracy. It is a very simple metric and easy to understand. It is simply the ratio of correctly predicted samples to the total number of sample.

$$\text{Accuracy} = \frac{\textit{Number of samples correctly predicted}}{\textit{Total number of samples used for prediction}}$$

For its simplicity it has been widely used but the main concern about this metric is, it cannot be used for evaluating imbalanced datasets. In case of imbalance datasets it gives false perception of achieving high accuracy. So, for evaluating the efficiency of a predictive model, other evaluation metrics have to used besides accuracy.
Sensitivity (or, Recall) can be used along with accuracy. It is the rate of actual positive samples predicted as positive (true positive rate) and indicates how efficiently positive class was predicted by the model. It is also known as True Positive Rate (TRP).

$$\text{Sensitivity} = \frac{\textit{True Positive (TP)}}{\textit{True Positive (TP)} + \textit{False Negative (FN)}}$$

There is another metric, specificity, almost similar with sensitivity. But it is the rate of actual negative samples predicted as negative (true negative rate). Hence it is complement to sensitivity. It shows how efficiently negative class was predicted by the model. It is also known as True Negative Rate (TNR).

$$\text{Specificity} = \frac{\textit{True Negative(TN)}}{\textit{True Negative(TN)} + \textit{False Positive(FP)}}$$

Sensitivity and specificity, both can be calculated using confusion matrix. A confusion matrix summarizes the prediction result made by a model. It calculates true positive, false positive, true negative and false negative values to show a model's prediction performance.
Before applying any of these evaluation methods for measuring the efficiency of a model there is another important task to do and that is data splitting. Datasets are often splitted into training, test and validation sets; but sometimes one step is omitted and datasets are divided into training and test sets only. Training set is used to train the model by setting the values of different parameters along with the optimization of error function. Validation set is used for avoiding overfitting. Test set helps to find out how the model will perform

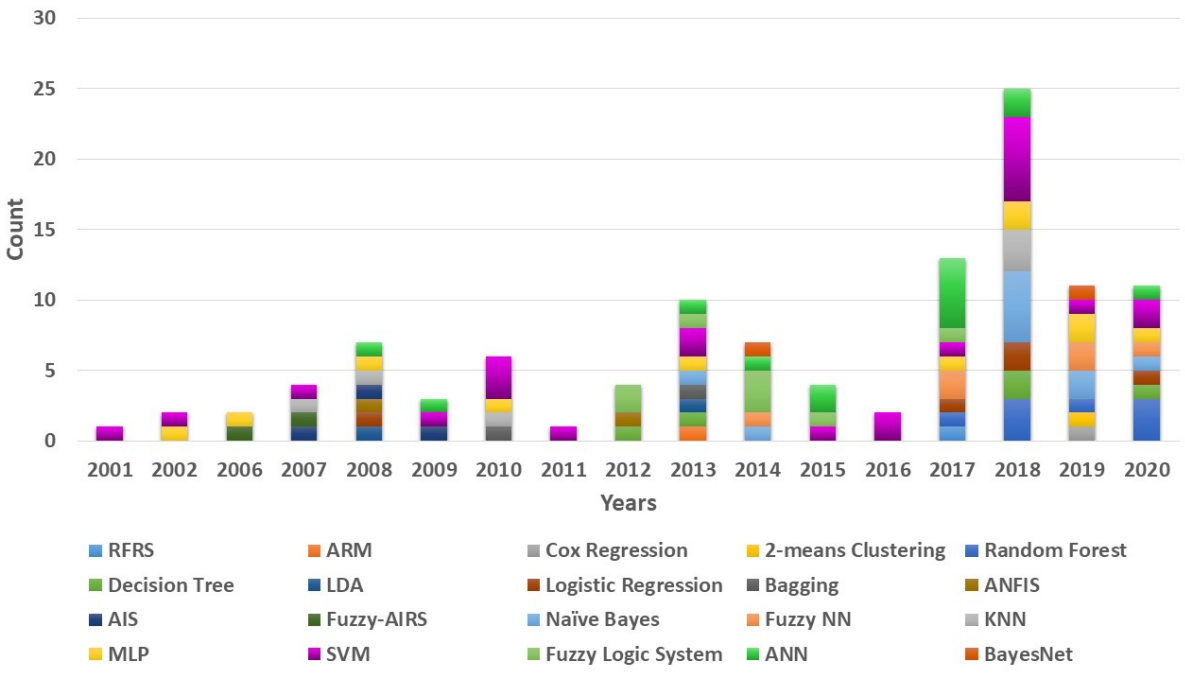

Fig. 8: Number of methods used in the timeline 2001-2020

for unseen samples. Evaluation metrics are applied for the test set. If no splitting is done in a dataset, proposed model has to be trained with the entire dataset and then there will be no data or sample left for testing the model before applying it in real world problem. So, data splitting is a foremost step for training any model. There are several methods for data splitting, such as cross-validation, hold-out, bootstrapping, jackknife and so on.

10. MODEL PERFORMANCE

Authors of different articles proposed divergent models for the prediction of different heart diseases. Two types of work have been noted here. A group of authors have applied existing algorithms as mentioned in Section 8 for this work. But majority of the authors have applied these algorithms with some innovations. They have blended these algorithms with some other proposed or existing algorithms to get higher performance. Both types of work have been appreciated for this research.

There is another point to discuss here and that is which disease has been considered by the authors for their articles. A bunch of authors have worked on cardiovascular diseases, but again several authors have worked on some particular cardiovascular diseases(CVD), for instance coronary artery disease(CAD) and heart valve disease. Among the considered papers, 52 of them are about classifying CVD or finding out risk factors of CVD. Classification of CAD and heart valve disease have 12 papers and 6 papers respectively. So, this section will be divided into three subsection for these three categories.

10.1 Model Performance on CVD

Enormous methods have been proposed for the classification of CVD or identifying the risk factors of CVD. The articles that have been used for this are enlisted below -

PID01	PID22	PID41	PID57
PID04	PID26	PID42	PID59
PID05	PID27	PID43	PID60
PID06	PID28	PID44	PID61
PID07	PID31	PID46	PID62
PID08	PID32	PID47	PID63
PID10	PID33	PID48	PID64
PID12	PID34	PID49	PID65
PID13	PID35	PID50	PID66
PID14	PID36	PID52	PID67
PID15	PID38	PID54	PID68
PID16	PID39	PID55	PID69
PID20	PID40	PID56	PID70

PID22 is reported for identifying risk factors of CVD. All the other papers have done classification task. Most of the authors have used the evaluation metrics discussed in Section 9.

The result of each paper will be analyzed in a step-by-step procedure. For classifying CVD, some authors have applied only renowned classifiers, some applied/proposed ensemble techniques, and some other authors have applied feature selection/feature extraction/feature weighting techniques followed by some base/ensemble classifiers. After performing the classification task, many authors get very high accuracy; on the contrary, some authors do not get that much high accuracy.

For analyzing the efficiency of the articles, they have been partitioned into five categories. Table 6 enlisted those articles which get classification accuracy less than 80%. Only 5 articles among 52 articles achieve classification accuracy less than 80%. Table 7 shows the articles having classification accuracy between 80% and 85%. 16 articles have achieved accuracy between this range. Table 8 listed 18 papers with classification accuracy in the range 85% to 90%. Table 9 shows the 22 papers which have achieved accuracy in between 90% and 95%. Finally, Table 10 shows the papers with best achieved classification accuracy. Each of the papers get accuracy more than 95%. If the tables are observed carefully, it can be seen that some papers are enlisted in more than one table. The reason is in many papers the authors have applied more than one technique for classifying CVD. Hence they get different accuracies for different methods, and so enlisted in more than one table.

After inspecting Table 6, Table 7, Table 8, Table 9 and Table 10 it can be observed that 93 techniques have been applied for classifying CVD by the authors of 52 articles. The lowest performance is achieved by PID08. It's accuracy (57.851%) along with sensitivity (52.473%) and specificity (68.75%) is lower than any other articles considered for this research work. They used their proposed feature selection technique, also handled missing values carefully and used weighted fuzzy rule base for classification but still got such a low performance. They applied these techniques on three datasets, and recorded three different performances; but for this research work only the best recorded performance has been taken.

5 articles with 7 techniques achieved accuracy less than 80% and 10 articles with 12 techniques achieved accuracy higher than 95%. The techniques which achieved accuracy in between 80% and 95% can be considered as average accuracy, i.e., the achieved accuracy is not bad but not that much good either. 74 techniques have achieved this average accuracy. SVM, ANN, MLP, fuzzy, ensemble and other techniques have been applied for this.

Among the 12 techniques, which have shown excellent performance (i.e., achieved accuracy is higher than 95%), 4 of them have used random forest algorithm as base classifier along with some other feature selection and feature weighting techniques. In total, 8 techniques have used random forest algorithm as base classifier and 4 of them have shown this level of high performance. So, it is evident that random forest algorithm helps to achieve higher accuracy most of the time.

The highest performance is achieved by PID54 and the accuracy is 100%. The applied algorithm is CSA-NB and COA-NB. Both the algorithms have been applied on the dataset *Echocardium* and achieved the mentioned accuracy. They have applied the same techniques on other four datasets but could not achieve this much high accuracy. On this note, the authors have not mentioned the sensitivity and specificity. So the only metric that has been used for measuring efficiency is accuracy.

If sensitivity is considered, three articles have shown 100% sensitivity - PID10 (method applied Fuzzy_AHP + FFN), PID40 (method applied ChiSqSelector + PCA and RF) and PID41 (method applied co-operative co-operation + RF on statlog dataset). Considering specificity, two articles have shown 100% specificity - PID04 (method applied Fuzzy-AIRS-Knn based system) and PID05 (applied method Decision tree)

10.2 Model Performance on CAD

A bunch of articles have proposed several methods for classifying CAD. 12 articles among the selected articles have worked for the classification of CAD and the articles are -

PID02	PID17	PID30	PID51
PID09	PID21	PID37	PID53
PID11	PID29	PID45	PID58

These 12 articles have done the classification task in 25 different approaches. A number of authors have proposed a single technique and applied the technique on several datasets. Some other authors have considered a single dataset and carried out several methods on that dataset. The performance of each of these approaches has been recorded carefully in Table 11.
From Table 11 it can be observed that, according to accuracy, the highest performance is achieved by PID58 and the accuracy is 98.6%. The authors have used NE-nu-SVC + feature selection + multi-step balancing technique for the classification task and applied this method on two datasets - Cleaveland dataset and Z-Alizadeh Sani dataset. But the mentioned accuracy has been achieved from Cleaveland dataset. On that matter, the authors have not mentioned neither the sensitivity nor the specificity of this method. PID09 has recorded highest sensitivity (98.9%) using the method Self-organizing feature maps (SOFM). If specificity is considered, PID29 has achieved the highest specificity of 95.4% by applying the method - Weights by SVM + Naive Bayes.

10.3 Model Performance on Heart Valve Disease

Like CVD and CAD, heart valve disease also sought attention of a bunch of authors. The articles that have worked on heart valve disease can be enlisted as -

PID03	PID19	PID24
PID18	PID23	PID25

Several methods have been applied on different datasets for classifying heart valve disease. Many base classifiers as well as hybrid techniques have been applied for this purpose. Table 12 showed the efficiency achieved by each paper regarding this matter.
Among the six articles, PID19 has shown the best performance. It has achieved 98.4% accuracy along with 97.3% sensitivity and 100% specificity. No paper regarding heart valve disease has shown this much high performance but this one.

11. DISCUSSION AND FUTURE WORK

Section 10 shades light on how the models work for classifying CVD, CAD and heart valve disease. This is clear that, no particular model is perfect for all the datasets. This review work has spotted enormous cases where a specific model has been applied on different datasets but for one dataset it has achieved very high accuracy, on the contrary, for other datasets it could not achieve that much high accuracy. Another important factor is feature selection. It has been found that the articles which have applied feature selection strategy, have achieved high accuracy most of the time. 8 among 9 articles summarized by Table 10 have used feature selection techniques.

11.1 Challenges of Using AI Techniques

Though AI based techniques are the mostly and variously used techniques now-a-days because of their high performance, there are several limitations to some extents. During this research work some factors have been identified that limit the capability of these techniques.

(1) Every technique has its own domain where it performs extraordinarily. One method can perform remarkably on a particular dataset, conversely it can perform poorly on other datasets. Therefore after collecting/making a dataset it is very dilemmatic to select a method because nobody knows which method will work fine on that dataset.

(2) When data are being collected for a dataset, there is a huge possibility to record irrelevant data. Hence applying any feature selection technique is an obvious step before starting the classification task. Again, choosing a good method for selecting relevant features is very bewildering.

(3) Handling missing values is another issue for these algorithms. Some algorithms cannot handle missing values on its own. If missing values are present in training set, then sometimes these algorithms cannot be trained using that training set and sometimes after training the algorithm generates faulty results. So, pre-processing is a very necessary step.

(4) In general, AI based algorithms need to be trained with large datasets for producing better result. The datasets also have to be unbiased. Lots of time and resources are needed for collecting this type of datasets.

(5) Training is the main task of any AI based algorithm. How accurately an algorithm classifies objects solely depends on the training dataset. If an algorithm is trained with some inaccurate data or biased data, then the algorithm will eventually produce misleading result. Even after noticing the errors, it takes a lot of time and resources to fix the problem.

11.2 Future Work

AI techniques perform great in some aspect. But there are some directions where improvement can be done.

(1) The authors of the articles, considered for this research work, have used 20 datasets. Some of them are collected from online dataset repositories and the remaining datasets are collected manually by the authors from different clinics of different countries. Section 5 provides detailed information about each dataset and by observing the details it can be summarize that the datasets are collected from some certain countries. Many continentals/countries, like Australia, Africa, South Asia and Europe, have no database regarding CVD/CAD/heart valve disease. Ethnicity is a primary risk factor of these diseases. So, databases should be created using the data of those countries also, so that in future authors can apply their AI models on those datasets and determine corresponding result.

(2) Association rule mining is a very useful tool for unleashing relevance among features. Except PID22, no paper has applied this technique for the classification task. Evolutionary algorithm, one of the branches of association rule mining, can be used for this purpose as this algorithm does not stuck in local minima. Hence it can improve the classification accuracy.

(3) Deep learning has brought revolutionary changes in the field of AI. It has improved the performance of different tasks such as robotics, games, speech recognition, computer vision and what not. But no top-quality work has been done in the field of classifying CVD/CAD/heart valve disease. Application of deep learning can improve the classification performance tremendously. If there is any work in this regard in the future, researchers will be able provide a systematic comparative study to new users.

12. CONCLUSION

In this review work, a large number of articles have been investigated that classified CVD/CAD/heart valve disease. We performed a thorough search in many databases and search engines, and after this searching process, these quality articles had been collected. In our research work, we have brought out the risk factors which are responsible for CVD/CAD/heart valve disease. Then the datasets were investigated comprehensively and detailed information was recorded about each dataset.

This research work also shades light on how the features and missing values have been handled by the authors. Feature selection strategy has been used in almost 56% articles. The percentage should be higher as applying appropriate feature selection strategy improves results to a great extent.

In case of missing values, only 13% articles have used any missing value imputation technique; all the other articles have just discarded it which can affect their model performance severely.

Finally, the performance of each model proposed by the authors has been recorded. The performance demonstrated that SVM, ANN and MLP are the most used algorithms considered for this scenario; but best performance is recorded when these methods are combined with other algorithms. Despite achieving great results, there are some shortcomings that must be addressed in coming years.

13. REFERENCES

[1] Resul Das, Ibrahim Turkoglu, and Abdulkadir Sengur. Effective diagnosis of heart disease through neural networks ensembles. *Expert Syst. Appl.*, 36:7675–7680, 05 2009.

[2] WHO. Cardiovascular diseases (cvds). `https://www.who.int/`, last accessed on 03-04-2020.

[3] WHO. Noncommunicable diseases (ncd) country profiles. 2018.

[4] Yaowapa Maneerat, K. Prasongsukarn, Surachet Benjathummarak, and W. Dechkhajorn. Intersected genes in hyperlipidemia and coronary bypass patients: Feasible biomarkers for coronary heart disease. *Atherosclerosis*, 252:e183–e184, 09 2016.

[5] Takahiro Nakashima, Teruo Noguchi, Seiichi Haruta, Yusuke Yamamoto, Shuichi Oshima, Koichi Nakao, Yasuyo Taniguchi, Junichi Yamaguchi, Kazufumi Tsuchihashi, Atsushi Seki, Tomohiro Kawasaki, Tatsuro Uchida, Nobuhiro Omura, Migaku Kikuchi, Kazuo Kimura, Hisao Ogawa, Shunichi Miyazaki, and Satoshi Yasuda. Prognostic impact of spontaneous coronary artery dissection in young female patients with acute myocardial infarction: A report from the angina pectoris-myocardial infarction multicenter investigators in japan. *International Journal of Cardiology*, 207, 01 2016.

[6] James Zebrack, Jeffrey Anderson, Chloe Maycock, Benjamin Horne, Tami Bair, and Joseph Muhlestein. Usefulness of high-sensitivity c-reactive protein in predicting long-term risk of death or acute myocardial infarction in patients with unstable or stable angina pectoris or acute myocardial infarction. *The American journal of cardiology*, 89:145–9, 01 2002.

[7] Barry Robson and Srinidhi Boray. Implementation of a web based universal exchange and inference language for medicine: Sparse data, probabilities and inference in data mining of clinical data repositories. *Computers in Biology and Medicine*, 2193:82–102, 09 2015.

[8] Seyed Shenas, Bijan Raahemi, Mohammad Tekieh, and Craig Kuziemsky. Identifying high-cost patients using data mining techniques and a small set of non-trivial attributes. *Computers in Biology and Medicine*, 53, 10 2014.

[9] Jae-Kwon Kim, Jong-Sik Lee, Dong-Kyun Park, Yong-Soo Lim, Youngho Lee, and Eun-Young Jung. Adaptive mining prediction model for content recommendation to coronary heart disease patients. *Cluster Computing*, 17, 09 2013.

[10] U Rajendra Acharya, Oliver Faust, Nahrizul Adib Kadri, Jasjit Suri, and Wenwei Yu. Automated identification of normal and diabetes heart rate signals using nonlinear measures. *Computers in Biology and Medicine*, 43:1523–9, 10 2013.

[11] Carlo Barbieri, Flavio Mari, Andrea Stopper, Emanuele Gatti, Pablo Escandell-Montero, Jos Martnez-Martnez, and Jos Martn-Guerrero. A new machine learning approach for predicting the response to anemia treatment in a large cohort of end stage renal disease patients undergoing dialysis. *Computers in Biology and Medicine*, 61, 03 2015.

[12] Dr. Subhash Pandey. Data mining techniques for medical data: A review. 11 2016.

[13] WHO. Cardiovascular disease. `https://www.who.int/cardiovascular_diseases/about_cvd/en/`.

[14] Heart disease and stroke statistics. 2019.

[15] Stanley Davidson, Stuart Ralston, Ian Penman, Mark Strachan, and Richard Hobson. *Davidson's principles and practice of medicine a textbook for students and doctors*. Elsevier, 23 edition, 2018.

[16] World Heart Federation. Rheumatic heart disease: A preventable, treatable form of cardiovascular disease. `https://www.world-heart-federation.org/programmes/rheumatic-heart-disease/`, last accessed on 10-04-2020.

[17] Prevention of cardiovascular disease. 2007.

[18] Lawrence D. Chilnick. Heart disease: An essential guide for the newly diagnosed. 2008.

[19] Chang-Sik Son, Yoon-Nyun Kim, Hyung-Seop Kim, Hyoung-Seob Park, and Min-Soo Kim. Decision-making model for early diagnosis of congestive heart failure using rough set and decision tree approaches. *J. of Biomedical Informatics*, 45(5):9991008, October 2012.

[20] Resul Das and Abdulkadir Sengur. Evaluation of ensemble methods for diagnosis of valvular heart disease. *Expert Systems with Applications*, 37:5110–5115, 07 2010.

[21] Hongzeng Xu, Zhiying Duan, Chi Miao, Song Geng, and Yuanzhe Jin. Development of a diagnosis model for coronary artery disease. *Indian heart journal*, 69(5):634 – 639, 2017.

[22] Ali Muhammad Usman, Umi Kalsom Yusof, and Syibrah Naim. Cuckoo inspired algorithms for feature selection in heart disease prediction. *International Journal of Advances in Intelligent Informatics*, 4:95–106, 2018.

[23] Kalia Orphanou, Arianna Dagliati, Lucia Sacchi, Athena Stassopoulou, Elpida Keravnou, and Riccardo Bellazzi. Incorporating repeating temporal association rules in nave bayes classifiers for coronary heart disease diagnosis. *Journal of Biomedical Informatics*, 81, 03 2018.

[24] Long Nguyen Cong, Phayung Meesad, and Herwig Unger. A highly accurate firefly based algorithm for heart disease prediction. *Expert Systems with Applications*, 06 2015.

[25] C. Beulah Christalin Latha and S. Carolin Jeeva. Improving the accuracy of prediction of heart disease risk based on ensemble classification techniques. *Informatics in Medicine Unlocked*, 16:100203, 2019.

[26] Chih-Wen Chen, Yi-Hong Tsai, Fang-Rong Chang, and Wei-Chao Lin. Ensemble feature selection in medical datasets: Combining filter, wrapper, and embedded feature selection results. *Expert Systems*, n/a(n/a):e12553.

[27] Emre Comak, Ahmet Arslan, and Ibrahim Turkoglu. A decision support system based on support vector machines for diagnosis of the heart valve diseases. *Computers in biology and medicine*, 37:21–7, 02 2007.

[28] Muthukaruppan M S and M.J. Er. A hybrid particle swarm optimization based fuzzy expert system for the diagnosis of coronary artery disease. *Expert Systems with Applications*, 39:1165711665, 10 2012.

[29] Himansu Das, Bighnaraj Naik, and H.S. Behera. Medical disease analysis using neuro-fuzzy with feature extraction model for classification. *Informatics in Medicine Unlocked*, 18:100288, 2020.

[30] Mohammad Shafenoor Amin, Yin Kia Chiam, and Kasturi Dewi Varathan. Identification of significant features and data mining techniques in predicting heart disease. *Telematics and Informatics*, 36:82 – 93, 2019.

[31] Kemal Polat, Salih Gne, and Slayman Tosun. Diagnosis of heart disease using artificial immune recognition system and fuzzy weighted pre-processing. *Pattern Recognition*, 39:2186–2193, 11 2006.

[32] Jesmin Nahar, Tasadduq Imam, Kevin Tickle, and Yi-Ping Phoebe Chen. Association rule mining to detect factors which contribute to heart disease in males and females. *Expert Systems with Applications*, 40:10861093, 03 2013.

[33] Anna Karen Grate-Escamila, Amir Hajjam El Hassani, and Emmanuel Andrs. Classification models for heart disease prediction using feature selection and pca. *Informatics in Medicine Unlocked*, 19:100330, 2020.

[34] Kindie Biredagn Nahato, Khanna Nehemiah Harichandran, and Kannan Arputharaj. Knowledge mining from clinical datasets using rough sets and backpropagation neural network. *Computational and Mathematical Methods in Medicine*, 2015, 2015.

[35] Hui Chen, Chao Tan, Zan Lin, Tong Wu, and Yuanbo Diao. A feasibility study of diagnosing cardiovascular diseases based on blood/urine element analysis and consensus models. *Comp. in Bio. and Med.*, 43(7):865–869, 2013.

[36] Abdulkadir Sengur. An expert system based on linear discriminant analysis and adaptive neuro-fuzzy inference system to diagnosis heart valve diseases. *Expert Syst. Appl.*, 35:214–222, 07 2008.

[37] V. R. Elgin Christo, H. Khanna Nehemiah, J. Brighty, and Arputharaj Kannan. Feature selection and instance selection from clinical datasets using co-operative co-evolution and classification using random forest. *IETE Journal of Research*, 0(0):1–14, 2020.

[38] Moloud Abdar, U. Rajendra Acharya, Nizal Sarrafzadegan, and Vladimir Makarenkov. Ne-nu-svc: A new nested ensemble clinical decision support system for effective diagnosis of coronary artery disease. *IEEE Access*, 7:167605–167620, 2019.

[39] T. Vivekanandan and N Ch Sriman Narayana Iyenger. Optimal feature selection using a modified differential evolution algorithm and its effectiveness for prediction of heart disease. *Computers in Biology and Medicine*, 90, 09 2017.

[40] Abdulkadir Sengur and Ibrahim Turkoglu. A hybrid method based on artificial immune system and fuzzy k-nn algorithm for diagnosis of heart valve diseases. *Expert Systems with Applications*, 35:1011–1020, 10 2008.

[41] Youness Khourdifi and Mohamed Bahaj. Heart disease prediction and classification using machine learning algorithms optimized by particle swarm optimization and ant colony optimization. *International Journal of Intelligent Engineering and Systems*, 12:242–252, 2019.

[42] Syed Muhammad Saqlain Shah, Faiz Ali Shah, Syed Adnan Hussain, and Safeera Batool. Support vector machines-based heart disease diagnosis using feature subset, wrapping selection and extraction methods. *Computers and Electrical Engineering*, 84:106628, 2020.

[43] T. Vivekanandan and Swathi Narayanan. A hybrid risk assessment model for cardiovascular disease using cox regression analysis and a 2-means clustering algorithm. *Computers in Biology and Medicine*, 113:103400, 08 2019.

[44] Ibrahim Turkoglu, Ahmet Arslan, and Erdogan Ilkay. An expert system for diagnosis of the heart valve diseases. *Expert Syst. Appl.*, 23:229–236, 10 2002.

[45] Seral zen and Salih Gne. Effect of feature-type in selecting distance measure for an artificial immune system as a pattern recognizer. *Digital Signal Processing*, 18(4):635 – 645, 2008.

[46] Younas Khan, Usman Qamar, Muhammad Asad, and Babar Zeb. Applying feature selection and weight optimization techniques to enhance artificial neural network for heart disease diagnosis. In *Intelligent Systems and Applications*, pages 340–351. Springer International Publishing, 2020.

[47] P.K. Anooj. Clinical decision support system: Risk level prediction of heart disease using weighted fuzzy rules. *Journal of King Saud University - Computer and Information Sciences*, 24(1):27 – 40, 2012.

[48] Hongmei Yan, Yingtao Jiang, Jun Zheng, Chenglin Peng, and Qinghui Li. A multilayer perceptron-based medical decision support system for heart disease diagnosis. *Expert Systems with Applications*, 30:272–281, 02 2006.

[49] Syed Muhammad Saqlain Shah, Muhammad Sher, Faiz Ali Shah, Imran Khan, Muhammad Usman Ashraf, Muhammad Awais, and Anwer Ghani. Fisher score and matthews correlation coefficient-based feature subset selection for heart disease diagnosis using support vector machines. *Knowledge and Information Systems*, 58:139–167, 2018.

[50] Bayu Adhi Tama, Sun Im, and Seungchul Lee. Improving an intelligent detection system for coronary heart disease using a two-tier classifier ensemble. *BioMed Research International*, 2020, 2020.

[51] Imran Omurlu, Mevlut Ture, and A. Kurum. Comparing performances of logistic regression, classification and regression tree, and neural networks for predicting coronary artery disease. *Expert Systems with Applications*, 34:366–374, 01 2008.

[52] Kemal Polat, Seral zen, and Salih Gne. Automatic detection of heart disease using an artificial immune recognition system (airs) with fuzzy resource allocation mechanism and k-

nn (nearest neighbour) based weighting preprocessing. *Expert Systems with Applications*, 32:625–632, 02 2007.

[53] N. Ghadiri Hedeshi and Mohammad Saniee Abadeh. Coronary artery disease detection using a fuzzy-boosting PSO approach. *Comput. Intell. Neurosci.*, 2014:783734:1–783734:12, 2014.

[54] M. Sudha. Evolutionary and neural computing based decision support system for disease diagnosis from clinical data sets in medical practice. *Journal of Medical Systems*, 41, 11 2017.

[55] Oluwarotimi Sauel, Mojisola Asogbon, Arun Sangaiah, Fang Peng, and Li Guanglin. An integrated decision support system based on ann and fuzzy ahp for heart failure risk prediction. *Expert Systems with Applications*, 68:163172, 10 2016.

[56] A.V. Kumar. Diagnosis of heart disease using fuzzy resolution mechanism. *Journal of Artificial Intelligence*, 5:47–55, 01 2012.

[57] T. V. Gestel, J. A. K. Suykens, G. Lanckriet, A. Lambrechts, B. D. Moor, and J. Vandewalle. Bayesian framework for least-squares support vector machine classifiers, gaussian processes, and kernel fisher discriminant analysis. *Neural Computation*, 14(5):1115–1147, 2002.

[58] Musa Peker. A decision support system to improve medical diagnosis using a combination of k-medoids clustering based attribute weighting and svm. 40(5), 05 2016.

[59] smail Babaolu, Oguz Findik, and Erkan lker. A comparison of feature selection models utilizing binary particle swarm optimization and genetic algorithm in determining coronary artery disease using support vector machine. *Expert Syst. Appl.*, 37:3177–3183, 04 2010.

[60] Roohallah Alizadehsani, Jafar Habibi, Mohammad Javad Hosseini, Hoda Mashayekhi, Reihane Boghrati, Asma Ghandeharioun, Behdad Bahadorian, and Zahra Sani. A data mining approach for diagnosis of coronary artery disease. *Computer methods and programs in biomedicine*, 111, 03 2013.

[61] Gokulnath Chandra Babu and Shantharajah S. Periyasamy. An optimized feature selection based on genetic approach and support vector machine for heart disease. *Cluster Computing*, 22:14777–14787, 2019.

[62] Der-Chiang Li, Chiao-Wen Liu, and Susan C. Hu. A fuzzy-based data transformation for feature extraction to increase classification performance with small medical data sets. *Artificial Intelligence in Medicine*, 52(1):45 – 52, 2011.

[63] Booma devi Sekar and Mingchui Dong. Function formula oriented construction of bayesian inference nets for diagnosis of cardiovascular disease. *BioMed research international*, 2014:376378, 08 2014.

[64] Zeinab Arabasadi, Roohallah Alizadehsani, Mohamad Roshanzamir, Hossein Moosaei, and Ali Yarifard. Computer aided decision making for heart disease detection using hybrid neural network - genetic algorithm. *Computer Methods and Programs in Biomedicine*, 141, 01 2017.

[65] Guang-Bin Huang, Xiaojian Ding, and Hongming Zhou. Optimization method based extreme learning machine for classification. *Neurocomputing*, 74(1):155 – 163, 2010. Artificial Brains.

[66] Swati Shilaskar and Ashok Ghatol. Feature selection for medical diagnosis : Evaluation for cardiovascular diseases. *Expert Systems with Applications*, 40(10):4146 – 4153, 2013.

[67] Jamal Alneamy and Rahma Alnaish. Heart disease diagnosis utilizing hybrid fuzzy wavelet neural network and teaching learning based optimization algorithm. *Advances in Artificial Neural Systems*, 2014, 09 2014.

[68] Ebenezer Olaniyi, Oyebade Oyedotun, and Adnan Khashman. Heart diseases diagnosis using neural networks arbitration. *International Journal of Intelligent Systems and Applications*, 7:75–82, 11 2015.

[69] H. Hannah Inbarani, Ahmad Taher Azar, and G. Jothi. Supervised hybrid feature selection based on pso and rough sets for medical diagnosis. *Computer Methods and Programs in Biomedicine*, 113(1):175 – 185, 2014.

[70] Yuehjen E. Shao, Chia-Ding Hou, and Chih-Chou Chiu. Hybrid intelligent modeling schemes for heart disease classification. *Applied Soft Computing*, 14:47 – 52, 2014.

[71] Amin Haq, Jian Li, Muhammad Memon, Shah Nazir, and Ruinan Sun. A hybrid intelligent system framework for the prediction of heart disease using machine learning algorithms. *Mobile Information Systems*, 2018:1–21, 12 2018.

[72] Yuh jye Lee and O. L. Mangasarian. Ssvm: A smooth support vector machine for classification. *Computational Optimization and Applications*, pages 5–22, 2001.

[73] Mehrbakhsh Nilashi, Othman bin Ibrahim, Hossein Ahmadi, and Leila Shahmoradi. An analytical method for diseases prediction using machine learning techniques. *Computers and Chemical Engineering*, 106:212 – 223, 2017.

[74] Xiao Liu, Xiaoli Wang, Qiang Su, Mo Zhang, Yanhong Zhu, and Qian Wang, Qiugen Wang. A hybrid classification system for heart disease diagnosis based on the rfrs method. *Computational and Mathematical Methods in Medicine*, 2017, 01 2017.

[75] Seral zen and Salih Gne. Attribute weighting via genetic algorithms for attribute weighted artificial immune system (awais) and its application to heart disease and liver disorders problems. *Expert Syst. Appl.*, 36:386–392, 01 2009.

[76] Jae Kwon Kim and Sanggil Kang. Neural network-based coronary heart disease risk prediction using feature correlation analysis. *Journal of Healthcare Engineering*, 2017, 2017.

[77] Jamal Salahaldeen Majeed Alneamy, Zakaria A. Hameed Alnaish, S.Z. Mohd Hashim, and Rahma A. Hamed Alnaish. Utilizing hybrid functional fuzzy wavelet neural networks with a teaching learning-based optimization algorithm for medical disease diagnosis. *Computers in Biology and Medicine*, 112:103348, 2019.

[78] Humar Kahramanli and Novruz Allahverdi. Design of a hybrid system for the diabetes and heart diseases. *Expert Syst. Appl.*, 35:82–89, 07 2008.

[79] Thanh Nguyen, Abbas Khosravi, Douglas Creighton, and Saeid Nahavandi. Classification of healthcare data using genetic fuzzy logic system and wavelets. *Expert Systems with Applications*, 42, 10 2014.

[80] Liaqat Ali, Shafqat Ullah Khan, Noorbakhsh Amiri Golilarz, Imrana Yakubu, Iqbal Qasim, Adeeb Noor, and Redhwan Nour. A feature-driven decision support system for heart failure prediction based on x^2 statistical model and gaussian naive bayes. *Computational and Mathematical Methods in Medicine*, 2019, 2019.

[81] Tien-Loc Le, Tuan-Tu Huynh, Lo-Yi Lin, Chih-Min Lin, and Fei Chao. A k-means interval type-2 fuzzy neural network for medical diagnosis. *International Journal of Fuzzy Systems*, 21:22582269, 2019.

[82] Moloud Abdar, Wojciech Ksiazek, U Rajendra Acharya, Ru-San Tan, Vladimir Makarenkov, and Pawel Plawiak. A new machine learning technique for an accurate diagnosis of coronary artery disease. *Computer methods and programs in biomedicine*, 179:104992, 2019.

[83] Thanh Nguyen, Abbas Khosravi, Douglas Creighton, and Saeid Nahavandi. Medical data classification using interval type-2 fuzzy logic system and wavelets. *Applied Soft Computing*, 30, 05 2015.

[84] Aniruddha Dutta, Tamal Batabyal, Meheli Basu, and Scott T. Acton. An efficient convolutional neural network for coronary heart disease prediction. *Expert Systems with Applications*, 159:113408, 2020.

[85] Anam Mustaqeem, Syed Muhammad Anwar, Abdul Rashid Khan, and Muhammad Majid. A statistical analysis based recommender model for heart disease patients. *International Journal of Medical Informatics*, 108:134 – 145, 2017.

[86] Engin Avci. A new intelligent diagnosis system for the heart valve diseases by using genetic-svm classifier. *Expert Syst. Appl.*, 36:10618–10626, 09 2009.

[87] Zahra Beheshti, Siti Mariyam Shamsuddin, Ebrahim Beheshti, and Siti Yuhaniz. Enhancement of artificial neural network learning using centripetal accelerated particle swarm optimization for medical diseases diagnosis. *Soft Computing*, 18, 11 2013.

[88] Silverstein, B. Silverstein Alvin, Silverstein Nunn Virginia, and Laura. Heart disease. 2006.

[89] NCBI. Ethnic differences in cardiovascular disease. `https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1767706/`, last accessed on 10-04-2020.

[90] Uci machine learning repository. `http://archive.ics.uci.edu/ml/datasets.php`.

[91] Mohammad Ashraf Ottom, Girija Chetty, Dat Tran, and Dharmendra Sharma. Hybrid approach for diagnosing thyroid, hepatitis, and breast cancer based on correlation based feature selection and nave bayes. volume 7666, pages 272–280, 11 2012.

[92] Dharmendra Modha and W. Spangler. Feature weighting in k-means clustering. *Machine Learning*, 52:217–237, 09 2003.

[93] Tao Wang, Zhenxing Qin, Zhi Jin, and Shichao Zhang. Handling over-fitting in test cost-sensitive decision tree learning by feature selection, smoothing and pruning. *Journal of Systems and Software*, 83:1137–1147, 07 2010.

[94] Roohallah Alizadehsani, Moloud Abdar, Mohamad Roshanzamir, Abbas Khosravi, Parham Kebria, Fahime Khozeimeh, Saeid Nahavandi, Nizal Sarrafzadegan, and U Rajendra Acharya. Machine learning-based coronary artery disease diagnosis: A comprehensive review. *Computers in Biology and Medicine*, 111:103346, 07 2019.

[95] Jiri Kaiser. Dealing with missing values in data. *Journal of Systems Integration*, 5:42–51, 01 2014.

[96] Runmin Wei, Jingye Wang, Mingming Su, Erik Jia, Shaoqiu Chen, Tian-Lu Chen, and Yan Ni. Missing value imputation approach for mass spectrometry-based metabolomics data. *Scientific Reports*, 8, 12 2018.

[97] Yuntian Chen, Haibin Chang, Jin Meng, and Dongxiao Zhang. Ensemble neural networks (enn): A gradient-free stochastic method. *Neural Networks*, 110:170 – 185, 2019.

[98] Xindong Wu, Vipin Kumar, Ross Quinlan, Joydeep Ghosh, Qiang Yang, Hiroshi Motoda, G. Mclachlan, Shu Kay Angus Ng, Bing Liu, Philip Yu, Zhi-Hua Zhou, Michael Steinbach, David Hand, and Dan Steinberg. Top 10 algorithms in data mining. *Knowledge and Information Systems*, 14, 12 2007.

[99] B.E. Boser, I.M. Guyon, and V.N. Vapnik. A training algorithm for optimal margin classifiers. *ACM press*, pages 144–152, 1992.

[100] Maher Alaraj, Munir Majdalawieh, and Maysam F. Abbod. Improving binary classification using filtering based on k-nn proximity graphs. *Journal of Big Data volume*, 7, 03 2020.

[101] Wei Zhang and Feng Gao. An improvement to naive bayes for text classification. *Procedia Engineering*, 15:2160 – 2164, 2011. CEIS 2011.

[102] Monika Kabir, Mir Md. Jahangir Kabir, Shuxiang Xu, and Bodrunnessa Badhon. An empirical research on sentiment analysis using machine learning approaches. *International Journal of Computers and Applications*, 0(0):1–9, 2019.

[103] L.A. Zadeh. Fuzzy sets. *Information and Control*, 8(3):338 – 353, 1965.

Table 4. : Feature selection methods used in different articles

Paper ID	Feature Selection Method
PID01, PID57	Rough Set
PID02	Temporal Association Rule
PID05, PID64, PID67	Principal Component Analysis (PCA)
PID06, PID07	Modified Differential Evolution
PID08	Authors' own method (no assigned method name)
PID11	Binary Particle Swarm Optimization (BPSO)
PID14	Relief, MRMR(minimal redundancy maximal relevance), LASSO(Least Abolute Shrinkage and Selection Operator)
PID15	RFRS(ReliefF + Rough Set)
PID17	GA-PSO (Genetic Algorithm + Particle Swarm Optimization)
PID20	Firefly algorithm + Rough Set
PID29, PID30	Weight by SVM
PID35, PID36	Wavelet Transformation
PID39	Integrated Component Analysis (ICA), Principal Component Analysis (PCA)
PID40	ChiSqSelector + Principal Component Analysis (PCA)
PID41	Co-operative Co-operation
PID42	Fast Correlation-Based Feature Selection (FCBF) + Particle Swarm Optimization (PSO) + Ant Colony Optimization (ACO)
PID44	Mean Fisher score-based feature selection algorithm (MFSFSA) + Forward feature selection algorithm (FFSA) + Reverse feature selection algorithm (RFSA)
PID47	Genetic Algorithm (GA)
PID49	PSO based Relative Reduct(PSO-RR) + PSO based Quick Reduct (PSO-QR)
PID51	Neural Network (NN) based feature correlation
PID52	χ^2 statistical model + forward best-first search for selection
PID53	Least Absolute Shrinkage and Selection Operator (LASSO) + Majority-voting
PID54	Cuckoo Search Algorithm (CSA), Cuckoo Optimization Algorithm (COA)
PID55	Principal Component Analysis (PCA), Genetic Algorithm (GA), C4.5, Ensemble of these methods
PID56	Brute force method
PID58	Genetic Search Algorithm
PID59	Mean Fisher based feature selection algorithm(MFFSA) + Accuracy based feature selection algorithm(AFSA) + PCA
PID60	(weights by Support Vector Machines, weights by Gini Index, Information Gain and Principal Component Analysis) + (Particle Swarm Optimization, Evolution Strategy, Backward and Forward weight optimization)
PID61	Particle Swarm Optimization(PSO)
PID62	Rough set + Genetic Algorithm
PID65	Forward Feature Inclusion, Back-elimination Feature Selection, Forward Feature Selection
PID66	Logistic Regression (LR), Multivariate Adaptive Regression Splines (MARS), Rough Set (RS)
PID70	Information gain

Table 5. : Techniques used for handling missing values in different articles

Paper ID	Dataset ID	Imputation Techniques for Missing Values
PID02	Dataset 14	missForest
PID08	Dataset 01 Dataset 11 Dataset 12	Average value of that attribute class
PID15	Dataset 09	Replaced probabilistically
PID16	Dataset 01	Average of corresponding class
PID21	Dataset 01 Dataset 11 Dataset 12	Nearest neighbour hot deck
PID26	Dataset 02	Substituting mean
PID36	Dataset 09	Mean of corresponding attribute value
PID38	Dataset 01	Method is not mentioned
PID39	Dataset 09	Mean of the attribute class

Table 6. : The articles with accuracies less than 80% for classifying CVD

Paper ID	Year	Used method	Dataset splitting	Accuracy(%)	Sensitivity(%)	Specificity(%)
PID08	2012	Weighted Fuzzy Rule (Cleaveland Dataset)	10-fold cross-validation	57.851	52.473	68.75
PID12	2014	Hierarchical Bayesian fuzzy inference nets (HBFIN)	75% for MF, 25% for validating BIN	72.25	NR	NR
PID34	2014	Fuzzy standard additive model + GA (GSAM) + Wavelet Transformation	5-fold cross-validation	78.78	NR	NR
PID48	2010	SVM (Gaussian Kernel)	Randomly generated training-testing partition	76	NR	NR
		SVM (ELM Kernel)		75.32	NR	NR
		ELM		76.25	NR	NR
PID65	2013	Forward feature selection + SVM (SPECTF dataset)	5-fold cross-validation	78	NR	NR

For some articles, only the method that gives highest accuracy is considered. It is also observed that some papers have used only the metric ”accuracy” for measuring the efficiency of their proposed model. So the corresponding cells of sensitivity and specificity contain NR (Not Reported).

Table 7. : The articles with accuracies between the range 80% and 85% for classifying CVD

Paper ID	Year	Used method	Dataset splitting	Accuracy(%)	Sensitivity(%)	Specificity(%)
PID06	2017	Modified DE + Fuzzy_AHP + FFNN	NR	83	84	89
PID14	2018	mRMR + Nave Bayes	10-fold cross-validation	84	77	90
PID31	2015	Backpropagation Neural Network	60%-40% training-testing partition	85	NR	NR
PID35	2015	IT2FLS (GCCD) + FCM + GA + WT	5-fold cross-validation	81.01	85	77
		IT2FLS (KMIP) + FCM + GA + WT		80.71	85	77
PID36	2013	CAPSO-MLP	80%-20% training-testing partition	81.85	74.63	90.21
PID39	2020	LNF-PCA	75%-25% training-testing partition	82.89	75.5	NR
PID42	2018	FCBF + PSO + ACO + SVM	70%-30% training-testing partition, 10-fold cross-validation	83.55	NR	NR
PID43	2007	AIS + Hybrid similarity measure	10-fold cross-validation	83.95	NR	NR
		AIS + Euclidean distance		83.21	NR	NR
		AIS + Manhattan distance		80.74	NR	NR
PID44	2018	FFSA + RBF kernel-based SVM (Cleaveland dataset)	65%-35% training-testing partition	81.19	72.92	88.68
		FFSA + RBF kernel-based SVM (Hungarian dataset)		84.52	73.91	92.63
		FFSA + RBF kernel-based SVM (SPECTF dataset)		82.7	86.4	40
PID46	2002	LS-SVM	2/3-1/3 training-testing partition	84.3	NR	NR
		SVM		83.4	NR	NR
		GP (Gaussian Process)		84.1	NR	NR
PID49	2014	SPSO-QR + Nave Bayes	80%-20% training-testing partition	83.46	83.3	NR
PID55	2020	PCA + GA + C4.5 + ensemble + SVM	5-fold cross-validation	83	NR	NR
PID59	2020	MFFSA + AFSA + PCA + SVM (Cleaveland Dataset)	60%-20%-20% training-validation-testing partition, 10-fold cross-validation	83.1	NR	NR
		MFFSA + AFSA + PCA + SVM (Hungarian Dataset)		83.9	NR	NR
		MFFSA + AFSA + PCA + SVM (Cleaveland + Hungarian + Switzerland Dataset)		84.2	NR	NR
PID65	2013	Forward feature selection + SVM (Heart disease dataset)	5-fold cross-validation	85	NR	NR
PID66	2014	MARS-LR	60%-40% training-testing partition	83.93	NR	NR
		RS-LR		83.93	NR	NR
PID70	2017	IG + MLP	66%-34% training-testing partition	83.78	NR	NR

For some articles, only the method that gives highest accuracy is considered. Some information were not present in some articles. Those are represented with NR (Not Reported).

Table 8. : The articles with accuracies between the range 85% and 90% for classifying CVD

Paper ID	Year	Used method	Dataset splitting	Accuracy(%)	Sensitivity(%)	Specificity(%)
PID05	2013	FLDA	mixture of bagging and cross-validation	85.5	65.7	93.5
PID14	2018	Relief + Logistic regression	10-fold cross-validation	89	77	98
		LASSO + SVM		88	75	96
PID16	2008	FNN + ANN	10-fold cross-validation	87.4	93	78.5
PID20	2015	CFARS-AR + Interval type-2 fuzzy logic system (SPECTF dataset)	30%-70% training-testing partition	87.2	94.2	68.9
		CFARS-AR + Interval type-2 fuzzy logic system (Heart disease dataset)	NR	88.3	84.9	93.3
PID27	2007	K-nn + Fuzzy-AIRS	10-fold cross-validation	87	92.3	78.57
PID31	2015	SVM	60%-40% training-testing partition	87.5	NR	NR
PID32	2009	Neural Network ensembles	70%-30% training-validation partition	89.01	80.95	95.91
PID33	2009	GA-AWAIS	10-fold cross-validation	87.43	NR	NR
PID38	2019	Majority vote with NB, BN, RF and MLP	10-fold cross-validation	85.48	NR	NR
PID42	2018	FCBF + PSO + ACO + NB	70%-30% training-testing partition, 10-fold cross-validation	86.15	NR	NR
PID47	2018	SVM	NR	88.34	NR	NR
PID49	2014	SPSO-RR + Nave Bayes	80%-20% training-testing partition	88.88	88.9	NR
PID50	2001	SSVM	10-fold cross-validation	86.13	NR	NR
PID54	2018	CSA-SVM (Eric Dataset)	NR	89.9	NR	NR
PID56	2018	Vote with Nave Bayes and Logistic Regression	10-fold cross-validation	87.41	NR	NR
PID61	2020	Random forest + Gradient boosting machine + XGBoost (Cleaveland dataset)	10-fold cross-validation	85.71	NR	NR
PID63	2016	kmAW + SVM	50%-50% training-testing partition	89.29	NR	NR
PID70	2017	IG + SVM	66%-34% training-testing partition	86.48	NR	NR

For some articles, only the method that gives highest accuracy is considered. Some information were not present in some articles. Those are represented with NR (Not Reported).

Table 9. : The articles with accuracies between the range 90% and 95% for classifying CVD

Paper ID	Year	Used method	Dataset splitting	Accuracy(%)	Sensitivity(%)	Specificity(%)
PID05	2013	SVM	mixture of bagging and cross-validation	92	82.1	96
PID07	2019	Cox regression + 2-means clustering	NR	91	NR	NR
PID10	2017	Fuzzy_AHP + FFNN	65%, 20%, 15% training, validating and testing set	91.1	100	84
PID13	2014	TLBO + FWNN	5-fold cross-validation	90.2909	NR	NR
PID15	2017	ReliefF + RFRS + C4.5 ensemble classifier	Jacknife; 70%-30% training-testing partition	92.59	93.33	87.5
PID26	2006	MLP	4 fold cross-validation	91.5	NR	NR
		MLP	holdout (3/4 for training, 1/4 for holdout; 4 holdout runs)	92	NR	NR
		MLP	bootstraping (0.632 bootstraping, 5 bootstraping sample)	91.1	NR	NR
PID28	2012	ANN + Fuzzy logic system	NR	91.83	NR	NR
PID41	2020	Co-operative co-operation + RF (Cleaveland Dataset)	10-fold cross-validation	93.4	96	90.48
PID42	2018	FCBF + PSO + ACO + MLP	70%-30% training-testing partition, 10-fold cross-validation	91.65	NR	NR
PID44	2018	FFSA + RBF kernel-based SVM (Switzerland dataset)	65%-35% training-testing partition	92.68	97.44	50
PID52	2019	χ^2-GNB	70%-30% training-testing partition	93.33	87.8	97.95
PID54	2018	CSA-SVM (Hungarian Dataset)	NR	94.22	NR	NR
		CSA-SVM (Statlog Dataset)		94	NR	NR
		CSA-SVM (Z-Alizadeh Sani dataset)		94	NR	NR
PID57	2015	RS-BPNN	10-fold cross-validation	90.4	94.67	90.37
PID59	2020	MFFSA + AFSA + PCA + SVM (Switerland Dataset)	60%-20%-20% training-validation-testing partition, 8 fold cross-validation	92.1	NR	NR
PID61	2020	Random forest + Gradient boosting machine + XGBoost (Statlog dataset)	10-fold cross-validation	93.55	NR	NR
		Random forest + Gradient boosting machine + XGBoost (Hungarian dataset)		91.18	NR	NR
PID62	2017	ASIC - BPNN	10-fold cross-validation	93.04	NR	NR
PID63	2016	kmAW + SVM	10-fold cross-validation	90.82	NR	NR
PID64	2011	PCA + SVM		91.37	NR	NR
PID67	2017	EM + PCA + Fuzzy rule-base (Statlog dataset)	10-fold cross-validation	91.4	NR	NR
		EM + PCA + Fuzzy rule-base (Cleaveland)		92.8	NR	NR
PID68	2019	TLBO-FFWNN	10-fold cross-validation	91.1	74.09	70.43
PID69	2019	KIT2FNN	70%-30% training-testing partition	93.81	94.08	93.58
PID70	2017	IG + RF	66%-34% training-testing partition	91.89	NR	NR

For some articles, only the method that gives highest accuracy is considered. Some information were not present in some articles. Those are represented with NR (Not Reported).

Table 10. : The articles having accuracies above 95% for classifying CVD

Paper ID	Year	Used method	Dataset splitting	Accuracy(%)	Sensitivity(%)	Specificity(%)
PID01	2012	RST + Decision Tree	10-fold cross-validation	97.5	97.2	97.7
PID04	2006	fuzzy-AIRS-knn based system	10-fold cross-validation	96.03	92.3	100
PID05	2013	Decision Tree	mixture of bagging and cross-validation	97.6	93	100
PID40	2020	ChiSqSelector + PCA and RF	70%-30% training-testing partition	98.7	100	NR
PID41	2020	Co-operative co-operation + RF (Statlog Dataset)	10-fold cross-validation	96.8	100	93.33
PID42	2018	FCBF + PSO + ACO + K-nn	70%-30% training-testing partition, 10-fold cross-validation	99.65	NR	NR
		FCBF + PSO + ACO + RF		99.6	NR	NR
PID54	2018	CSA-NB (Echocardium Dataset)	NR	100	NR	NR
		COA-NB (Echocardium Dataset)		100	NR	NR
PID60	2019	Feedforward MLP	90%-10% training-testing partition, 10-fold cross-validation	95.01	NR	NR
PID61	2020	Random forest + Gradient boosting machine + XGBoost (Z-Alizadeh Sani dataset)	10-fold cross-validation	98.13	NR	NR

For some articles, only the method that gives highest accuracy is considered. Some information were not present in some articles. Those are represented with NR (Not Reported).

Table 11. : The articles that have classified CAD along with their performances

Paper ID	Year	Used method	Dataset splitting	Accuracy(%)	Sensitivity(%)	Specificity(%)
PID02	2018	temporal association rule + nave bayes classifier	10-fold cross-validation	NR	83	NR
PID09	2008	Logistic Regression (LR)	60% for training, 20% for testing, 20% for cross-validation set	79.5	92.3	45.6
		Classification and Regression Tree (CART)		79.9	92.3	47.1
		Multilayer Perceptron (MLP)		79.1	91.7	45.6
		Radial Basis Function		76.7	89.5	42.6
		Self-organizing feature maps (SOFM)		73.9	98.9	7.4
PID11	2010	GA + SVM	5-fold cross-validation	79.17	NR	NR
		BPSO + SVM		81.46	NR	NR
PID17	2019	N2Genetic + nuSVM	10-fold cross-validation	93.08	NR	NR
PID21	2012	PSO + Fuzzy expert system	80%-20% training-testing partition	93.3	93.2	93.3
PID29	2013	Weights by SVM + Nave Bayes	10-fold cross-validation	75.51	67.59	95.4
		Weights by SVM + Bagging SMO		93.4	95.83	87.36
		Weights by SVM + SMO		94.08	96.3	88.51
		Weights by SVM + NN		88.11	91.2	80.46
PID30	2017	GA + Feed Forward Neural Network (Z-Alizadeh Sani dataset)	10-fold cross-validation	93.85	97	92
		GA + Feed Forward Neural Network (Long-beach-va dataset)		78	93	33
		GA + Feed Forward Neural Network (Hungarian dataset)		87.1	85	88
		GA + Feed Forward Neural Network (Cleaveland dataset)		89.4	88	91
		GA + Feed Forward Neural Network (Switzerland dataset)		76.4	78	50
PID37	2017	Multivariate logistic regression	NR	79	65.8	70.9
PID45	2014	Fuzzy boosting + PSO	10-fold cross-validation	85.76	90.02	82.31
PID51	2017	NN-FCA	70%-30% training-testing partition	82.51	NR	NR
PID53	2020	LASSO + Convolutional NN	NR	81.78	77.3	81.8
PID58	2016	NE-nu-SVC + feature selection + multi-step balancing (Cleaveland dataset)	10 fold cross-validation	98.6	NR	NR
		NE-nu-SVC + feature selection + multi-step balancing (Z-Alizadeh Sani dataset)		94.66	NR	NR

For some articles, only the method that gives highest accuracy is considered. Some information were not present in some articles. Those are represented with NR (Not Reported)

Table 12. : The articles that have classified Heart Valve Disease along with their performances

Paper ID	Year	Used method	Dataset splitting	Accuracy(%)	Sensitivity(%)	Specificity(%)
PID03	2007	LS-SVM	NR	NR	94.5	90
PID18	2009	GSVM	40%-60% training-testing partition	95	NR	NR
PID19	2010	SVM + Adaboost	40%-60% training-testing partition	98.4	97.3	100
PID23	2008	LDA + ANFIS	40%-60% training-testing partition	90	95.9	94
PID24	2008	AIS + Fuzzy k-nn	40%-60% training-testing partition	93.6	95.9	96
PID25	2002	BPNN	NR	94	NR	NR

For some articles, only the method that gives highest accuracy is considered. Some information were not present in some articles. Those are represented with NR (Not Reported)