A comparative study of classification and prediction of Cardio-Vascular Diseases (CVD) using Machine Learning and Deep Learning techniques

M. Swathy*, K. Saruladha

Department of CSE, Puducherry Technological University, Puducherry, India

Received 8 December 2020; received in revised form 3 May 2021; accepted 18 August 2021

Available online 3 September 2021

Abstract

Cardio-Vascular Diseases (CVD) are found to be rampant in the populace leading to fatal death. The statistics of a recent survey reports that the mortality rate is expanding due to obesity, cholesterol, high blood pressure and usage of tobacco among the people. The severity of the disease is piling up due to the above factors. Studying about the variations of these factors and their impact on CVD is the demand of the hour. This necessitates the usage of modern techniques to identify the disease at its outset and to aid a markdown in the mortality rate. Artificial Intelligence and Data Mining domains have a research scope with their enormous techniques that would assist in the prediction of the CVD priory and identify their behavioural patterns in the large volume of data. The results of these predictions will help the clinicians in decision making and early diagnosis, which would reduce the risk of patients becoming fatal. This paper compares and reports the various Classification, Data Mining, Machine Learning, Deep Learning models that are used for prediction of the Cardio-Vascular diseases. The survey is organized as threefold: Classification and Data Mining Techniques for CVD, Machine Learning Models for CVD and Deep Learning Models for CVD prediction. The performance metrics used for reporting the accuracy, the dataset used for prediction and classification, and the tools used for each category of these techniques are also compiled and reported in this survey.

© 2021 The Korean Institute of Communications and Information Sciences (KICS). Publishing services by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

Keywords: Cardio-Vascular Diseases; Classification; Regression; SVM; Deep Learning; Data Mining; Machine Learning; ANN; Artificial Intelligence

1. Introduction

Cardio-Vascular Disease (CVD) is an overall term referring to the conditions that affect the heart and blood vessels of a human body. This can also include the damage of the arteries in the organs such as the kidneys, heart, eyes and brain [1]. CVD is one of the main causes of death in many developed and developing countries all over the world even with young people. But, the fact is that it can be extensively prevented by leading a healthy lifestyle.

There are four main categories of the Cardio-Vascular Diseases. First and foremost is the Coronary Heart Disease which occurs due to the blockage of blood to heart muscle. This causes an increased strain on the heart and leads to angina, heart attacks and heart failure. The second type is Strokes and Transient Ischaemic Attack (TIA) which occurs due to blockage of blood to the brain and temporary disruption in the blood flow. The third type is Peripheral Arterial Disease which occurs due to blockage of blood to the limbs. This causes worst leg pains, hair loss on legs and feet, weakness in legs and persistent ulcers. The last type is the Aortic Disease which affects the largest blood vessel — Aorta. This has no symptoms, but causes a life-threatening bleeding when there is a chance of burst.

Cardiovascular disease comprises of the coronary artery diseases (CAD) like angina and myocardial infarction (commonly known as a heart attack) and coronary heart diseases (CHD), in which a waxy substance called plaque is developed inside the coronary arteries. Without a fast initiate to recovery, a heart attack can lead to serious health problems and even death, as this stays as a common cause of death worldwide.

Though, the exact cause for the CVD is not clearly found yet, there are lot of possibilities of one getting it. There are several risk factors involved with the chances of getting a CVD. Some of the most prominent factors are High Blood Pressure, Smoking, High Cholesterol, Diabetes, Obesity, Family History, Age, etc.,
Identification of people at risk due to CVD is a cornerstone. Due to many constraints in the manual identification of the heart diseases, scientists have moved towards modern approaches like Data Mining, Machine and Deep Learning methodologies for predicting the disease. These proved [2] to be effective to assist in decisions-making and predictions from enormous amount of data produced by the health care industry.

CVD can be diagnosed using an array of lab tests and imaging studies. However, the primary part of diagnosis is medical and family history of the patient, risk factors and physical examination. Through the statistical data, we can coordinate the findings and predict the presence of disease from results and procedures. Automation with Machine and Deep Learning can enable doctors to make informed decisions.

Automation in disease predictions can create a single platform from which structured data can be retrieved and efficient care can be given to the patients. Thus, it redefines the level of personalized health-care. Through artificial intelligence and machine learning, computers are taught to recognize patterns in which the disease occurs and convert them as structural data for predicting the same.

Innovations are made in Electronic Health Records (EHR), revenue cycle and operations through AI. In future, it will be integrated with the clinical work-flow with the existing tools empowering the practitioners with real-time data at the point of care.

2. Survey Organization

This survey gives a comparison of various classification and predictions for CVD. It follows a threefold organization with Data Mining Techniques for CVD, Machine Learning Models for CVD and Deep Learning Models for CVD prediction.

This study gives a brief description about the methods and algorithms which are used for predictions, the classification techniques, performance metrics and tools used for evaluation of their model.

2.1. Classification and Data-Mining techniques for CVD

Heart is the vital organ of a human body. It pumps blood to all other organs through the body. If there is any distortion in the circulation of blood (or) insufficiency in blood, it can lead to serious effects such as brain fever or even death that occurs within minutes [3]. A Heart disease is a collective term that refers to the diseases in par with the heart and its associated blood vessel system.

Data Mining (DM) is a field of Computer Science that is used for extraction of useful data from huge data sets which can be used in predictions or the data can be described using techniques such as classification, clustering, association, etc., [4,5]. Data Mining combines machine learning, mathematical analysis and information technology to evaluate large pre-existing databases to extract the hidden patterns among the data. In the healthcare domain, DMT can be used to predict the presence of a disease by evaluating the giant databases and exploring the relationships between the data using the useful trained patterns. This automation of this can be extremely advantageous in the prediction systems. Research based on DMT has already been applied to diabetes, asthma, CVD and AIDS. DMT such as Naive Bayesian (NB) classification, Artificial Neural Networks (ANN), Support Vector Machines (SVM), Decision Trees (DT), Logistic Regression (LR), etc. are used in the medicinal research. Decision tree gives a procedural approach for classification of categorical data based on the features or attributes. Hence, it stands the most widely used method for classification in DMT for processing large amount of data.

DMT basically applies to Predictive Analytics. This refers to extracting uncovered patterns from vast datasets and uses them to build an appropriate network for prediction. The general workflow of prediction systems is given below (Fig. 1).

This research work used the Artificial Neural Network (ANN) [6] which is an effective way to design the heart disease prediction system. Similarly [7] utilizes DMT for the cerebral-vascular disease prediction system. It uses Decision Tree and Bayesian classifiers for classification and a Multi-layer perceptron with Back Propagation (BP) for training the model. The result of this work indicates that, DMT can make the prediction of CVD more efficiently. Advanced DMT can even help in identifying the hidden patterns and relationships which often goes unexploited.

The healthcare and medicinal industry deal with large volume of data. Many of them are not structured and viable to identify the existing patterns. Making the process of discovering the relationships can make an effective decision-making system [8]. Choosing the appropriate data needed for implementing the above procedure from the huge loads of data is the most needed work. DMT would be the best approach to derive the useful details in depth on all the different perspectives [9,10].

The prediction of CVD is the hot topic of research in the milestone of medical industry. DMT based prediction systems can help in determining the disease at the starting stage itself which can minimize the risk associated. The research work [11], proposed a slight modification in the Weighted Associated Classifier (WAC) which showed an accuracy of 81.51%. Similarly [12], proposed a rule based discovery model with Associative Classification Mining to construct the classification systems. This uses an algorithm called MAFIA (Maximal Frequent Item-set Algorithm) to determine the frequent patterns. Medical Data Mining is a part of healthcare with lot of imprecision and uncertainty. This work [13] proves the working efficiencies of ID3 and CART algorithms on decision trees and concludes stating CART is more efficient than ID3 both theoretically and practically. The decision making capability [14] can be enhanced using the K-Means Clustering.

According to WHO estimation till 2030, very nearly 23.6 million individuals will pass because of Heart Malady [15]. CVD though is the major cause, can be controlled and prevented. To minimize the seriousness caused, analysis is very important. The most complex part is to choose the right ailment [16]. Gives the summary of recent works done in the field of data mining related to CVD. It gives a beautiful conclusion.
stating instead of relying upon a very specific DMT, hybrid or combination techniques can produce more efficient and accurate results.

But the tough procedure here is to find the appropriate combination of the algorithms to form the accurate model. The frequently used list of datasets with DMT and ML models are listed in the table (see Table 1).

From the results of [17] it can be concluded that the accuracy of prediction depends on the DMT used, Data Set handled by the model and the number of attributes. According to [18], if 102 cases are analysed, SVM has a highest accuracy of 90.5% and Logistic Regression has the lowest of 73.9%. Survey of 1000 patients showed SVM with 92.1%, ANN with 91% and Decision Tree with 89.6% accuracy. It can also be inferred that results with large sensitivity rate and specificity rate but are with lower accuracy will be abstained from the results, which makes the model a highly efficient.

2.1.1. Observation and Inferences: Data-Mining & Classification Techniques

Data Mining is a generalized term that includes many techniques to extract meaningful information without having pre-conceived notations about what will be discovered.

1. In general the most widely used method for classification is Decision Tree and Naive Bayes Classifiers. Observations show that, Neural Network based classification has more performance than the above two methods.

2. On the other hand, other data mining methods such as Clustering, Association Rule based, Time Series based can also be analysed for the usage of predictions.

3. Almost all the models discussed above uses categorical data for their classification and prediction. In real-time, the usage of continuous data can be more advantageous for analysis.

4. Further, we should also think upon extending the models to ensemble based algorithms.

2.2. Machine Learning models for CVD

Machine Learning is the basic practice of using algorithms to make predictions by parsing data and learning from it. These models have the capability of learning by itself from prior experience or from historical data. These algorithms can figure out extract the important tasks to be performed by generalizing from examples provided to them as training sets.

Different types of ML algorithms have evolved. These are grouped by either learning style (i.e. supervised learning, unsupervised learning, and semi-supervised learning) or by similarity or by their functioning (i.e. classification, regression, decision tree, clustering, deep learning, etc.). All machine learning algorithms comprise of three different components namely:

1. Representation: Set of classifiers in Computer understandable form
2. Evaluation: The objective defined for the classifier model (algorithm) - Scores
3. Optimization: Search method for the most scored classifier.

The primary goal of ML algorithms are to generalize beyond the training samples provided to them, which involves the successful interpretation of data that it has never noticed before.

The main difference that holds between ML and AI is that ML works for increasing accuracy and AI works for the increasing chance of success.

The basic three types of ML Techniques are: Supervised Learning, Unsupervised Learning and Reinforcement Learning. The selection of the algorithm and the learning type can be made by different approaches like depending on the task accomplished, (or) the amount of data involved (or) the different types of data that are available. This exhibits a dynamic role that plays out in applications of medical diagnostics as it involves creating self-learning algorithms. CVD prediction involves supervised learning technique as labelled data is required for training the model.

In case of prediction of CVD, regular diagnosis is very important in the initial stages of treatment which in turn reduces the risks associated with it. The most common and vital diagnostic tests include echocardiography (echo), cardiac magnetic resonance imaging (MRI), and computed tomography (CT) [19]. High quality cardiac images are produced by MRI and CT scans, which are not preferred for predictions as they have prolonged acquisition time, limited availability and involve the use of radiations.

Electrocardiogram (ECG) is a graphical representation that is produced by repolarization and depolarization of the ventricles and atria. Though there are several advancements made in the field of prediction and diagnostics, the accurate way of avoiding heart attack is not known as there are no proper symptoms associated with it. To discover a disease that forms the main causes of death such as HIV, Cancer, CVD, machine learning [20] can be used. It is a great consequence to research.

CVD are caused due to the deposits of fat (cholesterol) in the inner walls of arteries which narrows down or blocks coronary arteries. An efficient heart disease prediction system can be a beneficial way to exactly predict the diseases and save the patient’s life. The system model presented in [21] can interpret human patterns and accurately determine trends in the patients’ records.

The efficient functioning of heart is very vital for human life. In the prediction of heart-based diseases, automation
Table 1
Comparison of datasets used with the prediction models.

<table>
<thead>
<tr>
<th>S. No</th>
<th>Data-set</th>
<th>No. of I/P attributes</th>
<th>Data mining/Machine learning techniquesa</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>DT</td>
</tr>
<tr>
<td>1</td>
<td>Cleveland database</td>
<td>13</td>
<td>✓</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td>14</td>
<td>✓</td>
</tr>
<tr>
<td>3</td>
<td></td>
<td>15</td>
<td>✓</td>
</tr>
<tr>
<td>4</td>
<td>UCI repository</td>
<td>13</td>
<td>✓</td>
</tr>
<tr>
<td>5</td>
<td>Publicly available heart dataset</td>
<td>13</td>
<td>✓</td>
</tr>
<tr>
<td>6</td>
<td>Disease dataset</td>
<td>6</td>
<td>✓</td>
</tr>
<tr>
<td>7</td>
<td>Kaggle</td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>8</td>
<td>Deep Learning techniques</td>
<td></td>
<td>✓</td>
</tr>
</tbody>
</table>


Table 2
Comparison of evaluation metrics with deployment models.

<table>
<thead>
<tr>
<th>S. No</th>
<th>Evaluation metric</th>
<th>Data mining/Machine learning modelsa</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>DT</td>
</tr>
<tr>
<td>1</td>
<td>Sensitivity</td>
<td>✓</td>
</tr>
<tr>
<td>2</td>
<td>Specificity</td>
<td>✓</td>
</tr>
<tr>
<td>3</td>
<td>Accuracy</td>
<td>✓</td>
</tr>
<tr>
<td>4</td>
<td>Precision</td>
<td>✓</td>
</tr>
<tr>
<td>5</td>
<td>Confusion matrix</td>
<td>✓</td>
</tr>
<tr>
<td>6</td>
<td>Efficiency</td>
<td>✓</td>
</tr>
<tr>
<td>7</td>
<td>Lift chart</td>
<td>✓</td>
</tr>
<tr>
<td>8</td>
<td>Deep Learning models</td>
<td></td>
</tr>
</tbody>
</table>


plays an important role with a swift examination of accurate result. This research work [22] incorporates the classes of Heart Disease through Support Vector Machine (SVM). They initially analyse the historical data of the patient and fetch the real-time ECG values. Prediction is then made with these as input through SVM.

In the prevailing lifestyle, everyone is more tensed and is tested to have high blood pressure and sugar levels at a very young age. The less attention shown towards the quality of food taken and the own medications they tend to possess can lead to major threat of heart diseases [23]. This system extracts hidden knowledge associated with CVD diseases. The two important notices are, continuous data is preferred instead of categorical data and the integration of data mining and text mining through the machine learning model. The evaluation metrics commonly used to test the prediction model with DMT and ML Techniques is discussed (see Table 2).

The most popular machine learning algorithm called Naïve-Bayes (NB) forms the basis of several other algorithms and data processing ways.

This algorithm uses the Bayesian Rule that calculated the predictive capabilities through probabilistic approach [24]. This helps in exploring new ways of knowledge-oriented training, classification and prediction. Data Mining combined with Naïve Bayes can give efficient results of prediction. In this [25] research work, DSHDPS is deployed as web-based questionnaire application. Based on the user responses, the model can discover and extract the hidden relationships associated with the heart disease. This can be the most significant way of prediction of heart diseases.

2.2.1. Observation and inferences: Machine Learning techniques

Different types of Machine Learning oriented research papers have been analysed. The general conclusion drawn from the implemented models are listed below:

1. More featured medical attributes can be used for providing better model with more accuracy and performance.
2. Integrate data-mining and text-mining with the existing models for constructing efficient prediction systems.
3. Continuous information can be used in place of categorical information to build a heart disease system with early detection.
(4) Execution can be improved by using Genetic Algorithms and Swarm Intelligence Techniques to provide more concentration on the feature selection and input parameters.

2.3. Deep Learning models for CVD prediction

Deep Learning is referred to as the subset of ML and AI that has a network which capable of learning unsupervised forms of data. In Machine Learning, when the learning algorithm is not working properly, to make it accurate we feed more amounts of data for training the model. This may lead to issues with scalability and the learning time of the model increases exponentially. To overcome the data handling issues, we can switch over to Deep Learning techniques (DLT) which is capable of learning better representation of unstructured (or) unlabelled data with multiple levels of abstraction.

DLT utilizes hierarchical level of artificial neural networks (built like a human brain) to carry out the exact process of ML. This is one of the biggest advantages of DLT as it processes the data in a non-linear fashion while others process in a linear way. This makes the system quickly adapts the healthcare domain as it offers the ability to analyse data with a greater speed and precision. It also has an added benefit of being able to take decisions with a significantly less involvement of human trainers. Deep Learning requires less pre-processing of data when compared to ML and DMT. The DLT network itself is capable of filtering and normalization tasks, which is done by human programmers in other MLT. According to [26], clinical decisions are made based on the heuristic’s experiences and on the doctors’ intuition.

The knowledge based on the hidden data can be used to efficiently diagnose a heart disease thus reducing medical errors and it also decreases the diagnostic time. Through regular practices, it can also enhance the patient’s satisfaction and safety. Diagnosing a disease is the most crucial part because it is proven to be based on the doctor’s knowledge and experiences. But, training a machine to act like a human and making the machine learn the algorithms to carry over can in turn make it more [27] accurate and time efficient. The accuracy of the prediction is the prime concern in the predictive methodology and models. DLT offers a wide range of applications to represent trust-worthy text analytics in spite of the biased and skewed data. So many types of DLT are offering roles in decision making and predictive analytics because it combines the more advanced methods in order to increase the power and creates new method to benefit for prediction and prevention. Some of their applications are Personalized Treatment for Diabetes(Type II) and Cancer Patients, Using radiology for Image Classification for Cancer, Tumour and Lung treatment, Drug Discovery and Data Augmentation using GAN, Treatment Identification for Cancer and HIV, etc., In today’s world, many people are found to be living with heart diseases [28] without any awareness of themselves. Once they are predicted in advance, an accurate treatment can be provided to reduce the consequence. DLT can be used for Chat Bots and Medical Imaging Solutions which can identify the patterns and symptoms associated with specific kinds of diseases.

Neural Network which is a DLT follows a graph topology. It is a parallel, distributed information processing structure consisting of multiple nodes. Each node corresponds to the neurons and the weight associated with them corresponds to the edge. It has many hierarchical layers which are finite in order to decrease the time of problem solving. Every node has a single output connection that branches into many connections. The layers of NN typically include one input and output with multiple hidden layers. Neuro-Fuzzy is a combination of Neural Networks and Fuzzy Logics which can be used to solve wide range of real-time problems with an ease. This can well suit the limitations pertaining to other models of the automated techniques.

Artificial Neural Networks (ANN) is used in many areas of medicine [29], such as cardiology, Electroencephalography, Pulmonology, Genetics, Clinical chemistry, Pathology, Ophthalmology, Obstetrics and Gynecology. Considerable research are also being taken over the heart diagnosis. Prediction of CVD from various factors and symptoms is a multi-layered issue [30] which may lead to false presumptions and unpredictable effects.

DLT is so adept for Image Processing that many of the AI researchers are using Neural Networks to create medical images — to read them, analyse them and use them for the prediction of diseases. Convolution Neural Networks (CNN) is a type which is particularly suited to analyse the MRI Scans and X-rays. These can operate more effectively on larger images also thus surpassing the diagnosticians’ accuracy on imaging studies. In prediction of CVD, there is a demand to accurately detect the ED and ES frames using an automated image-driven method. CNN and RNN (Recurrent Neural Networks) has gained an enormous growth in the medicinal applications such as CVD Prediction Systems, Tumour Detection, Cancer Detection, Gene Classification, Neural Cells Classification, etc., The main advantages of this is compared to other techniques, this can automatically detect the important features for prediction networks without human intervention [19]. CNN can be used for Image Feature Extractions while RNN can be used for learning about the temporal dependencies between them. These are also computationally efficient when compared with other techniques for prediction. These models can also suit the time-series kind of data [31,32] like the patient’s diagnostic history, EHR, E-Prescriptions etc., for predictions as the information is remembered throughout the network. The final output from these networks can be given to Regression Modules for prediction.

ResNet and DenseNet are a form of NN. The former uses skip-connections to forward the features from one layer to another whereas the latter uses the entire information by concatenating them from the preceding layers and pass them as input features to the next layer. Effective and accurate diagnosis can only lead to appropriate treatment for the patient. This can be completely done on deep study of CV analysis of the patient. The [33] table represents the performance metrics of various algorithms.
Fig. 2. Prediction/classification workflow.

Table 3
Comparison table – Prediction models with deployment tools.

<table>
<thead>
<tr>
<th>Models</th>
<th>Tools used</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>WEKA</td>
</tr>
<tr>
<td>DT</td>
<td>✓</td>
</tr>
<tr>
<td>NB</td>
<td>✓</td>
</tr>
<tr>
<td>NN/A-NN</td>
<td>✓</td>
</tr>
<tr>
<td>FL</td>
<td>✓</td>
</tr>
<tr>
<td>GA</td>
<td>✓</td>
</tr>
<tr>
<td>SVM</td>
<td>✓</td>
</tr>
<tr>
<td>PCA</td>
<td>✓</td>
</tr>
<tr>
<td>CL</td>
<td>✓</td>
</tr>
<tr>
<td>J48</td>
<td>✓</td>
</tr>
<tr>
<td>RF</td>
<td>✓</td>
</tr>
</tbody>
</table>


2.3.1. Observation and inferences: Deep Learning techniques

Present NN based models suit only for specific or minimal kind of heart diseases. Hence, NN based systems should be expanded to suit wide range of heart based diseases.

1) With respect to ANN, we can make changes in the architecture and train algorithms for achieving more accurate results.

2) Generally, for 15 attributes, the Multi-Layer Perceptron Neural Networks with Back Propagation provides better results than other models.

3) Mostly, in all DL models, accuracy is considered to be the performance metric. But, we can also try to consider other metrics (such as sensitivity, specificity, efficiency, etc..) based on the demands of the diagnosis.

4) In future, Fuzzy Logics can be incorporated with NN to include more discrete valued attributes for prediction of CVD.

Table 4
Pros and cons of WEKA.

<table>
<thead>
<tr>
<th>Advantages</th>
<th>Disadvantages</th>
</tr>
</thead>
<tbody>
<tr>
<td>Freely available and portable</td>
<td>Algorithm does not cover sequential modelling</td>
</tr>
<tr>
<td>Easy to use GUI and command line</td>
<td>It cannot suit multi-relational data mining</td>
</tr>
<tr>
<td>Provides access to SQL – Databases through JDBC</td>
<td>It is memory bound</td>
</tr>
</tbody>
</table>

3. Tools for prediction/classification models

There are several tools available for evaluation of the proposed prediction models. Some of the widely used tools are listed below. The tools suitable for different kind of prediction models is also discussed (see Table 3).

3.1. Waikato Environment for Knowledge Analysis (WEKA)

It is a Data-Mining/ Machine-Learning tool that can be used to apply the algorithm directly on the dataset or through JAVA Code. It contains tools for data pre-processing, classification, regression, clustering, association rules and visualization. It is open-source software in JAVA issued under GNU. (General Public Licensed) (see Table 4).

3.2. TANGARA

It is free Data Mining software for academic and research purposes. It contains tools for exploratory data analysis, statistical learning, machine learning, and databases. It allows analysing both real and synthetic data and also allows adding our own data mining methods, to compare their performances. It is open source software in C++ issued under GNU.
their basic level and further efforts and resources are needed to use a combination of methods to achieve an accurate model. According to our research perspective and methods to effectively predict CVD. Each method has its specific kinds of automation used along with their algorithms from CVD.

Machine Learning helps to assess the risk of patients suffering cardiac imaging and analysis and intelligent Robots. Development of Sensor Technology has furthered the application of AI in CVD — includes precisional medicine, clinical prediction, cardiac imaging and analysis and intelligent Robots. Development of Sensor Technology has furthered the application of AI. Machine Learning helps to assess the risk of patients suffering from CVD.

In order to diagnose a heart disease, several kinds of lab tests and imaging studies are required. The latest research includes examination of risks of heart attacks and its possible recovery of open heart surgeries with angioplasty and stenting in patients with diabetes and blockages in more than one coronary artery. Researchers are exploring the use of diagnostic technologies in detection of heart diseases.

Automation applied to the field of prediction can lead to a high benefit of asset in the medicinal industry. AI applied in CVD — includes precisional medicine, clinical prediction, cardiac imaging and analysis and intelligent Robots. Development of Sensor Technology has furthered the application of AI. Machine Learning helps to assess the risk of patients suffering from CVD.

The above discussed three-fold approach confines to the specific kinds of automation used along with their algorithms and methods to effectively predict CVD. Each method has its own kind of pros and cons. According to our research perspective, we can either relate them and choose a specific method or use a combination of methods to achieve an accurate model.

In conclusion, the prediction methods listed are clearly in the field of Prediction of Cardio-Vascular Diseases using Machine and Deep Learning Techniques. The lifestyle has changed over the past few years leading to a lot of health complications which goes unnoticed in major of the population. Taking right measures at the right time can lead to save an individual’s life.

Heart Diseases form a vital part of mortality rate. These have no specific symptoms for their occurrences. This can be prevented by making custom lifestyle changes such as stop smoking, having a controlled BP, checking cholesterol rate, keeping diabetes under control, diet patterns and exercises and maintaining pressures and stress. The existing techniques and the workflow prediction is depicted as a overview in the above figure (Fig. 2).

<table>
<thead>
<tr>
<th>Advantages</th>
<th>Disadvantages</th>
</tr>
</thead>
<tbody>
<tr>
<td>Platform independent &amp; Easy to use</td>
<td>It can be slow as MATLAB is an interpreted language</td>
</tr>
<tr>
<td>Predefined functions and Toolboxes</td>
<td>It is not a free - software</td>
</tr>
<tr>
<td>Device independent plotting</td>
<td>Consumes large amount of memory</td>
</tr>
<tr>
<td>Good visualization of results</td>
<td>Consumes large amount of time – making real time applications very complicated.</td>
</tr>
</tbody>
</table>

### 3.3. Matrix Laboratory (MATLAB)

It provides user accurate solutions for problems with flexible graphics. It is highly interactive and programmable environment. It supports all mathematical computations, visualizations and programming. It is composed of High-Level programming language which is similar to C (see Table 5).

### 4. Future research perspectives

This survey paper is a consolidation of works done in the field of Prediction of Cardio-Vascular Diseases using Machine and Deep Learning Techniques. The lifestyle has changed over the past few years leading to a lot of health complications which goes unnoticed in major of the population. Taking right measures at the right time can lead to save an individual’s life.

Heart Diseases form a vital part of mortality rate. These have no specific symptoms for their occurrences. This can be prevented by making custom lifestyle changes such as stop smoking, having a controlled BP, checking cholesterol rate, keeping diabetes under control, diet patterns and exercises and maintaining pressures and stress. The existing techniques and the workflow prediction is depicted as a overview in the above figure (Fig. 2).

In order to diagnose a heart disease, several kinds of lab tests and imaging studies are required. The latest research includes examination of risks of heart attacks and its possible recovery of open heart surgeries with angioplasty and stenting in patients with diabetes and blockages in more than one coronary artery. Researchers are exploring the use of diagnostic technologies in detection of heart diseases.

Automation applied to the field of prediction can lead to a high benefit of asset in the medicinal industry. AI applied in CVD — includes precisional medicine, clinical prediction, cardiac imaging and analysis and intelligent Robots. Development of Sensor Technology has furthered the application of AI. Machine Learning helps to assess the risk of patients suffering from CVD.

The above discussed three-fold approach confines to the specific kinds of automation used along with their algorithms and methods to effectively predict CVD. Each method has its own kind of pros and cons. According to our research perspective, we can either relate them and choose a specific method or use a combination of methods to achieve an accurate model.

In conclusion, the prediction methods listed are clearly in their basic level and further efforts and resources are needed to gather more data with length follow-ups to derive population-specific methods that addresses all the concerns of existing prediction models. This can also end up in producing personalized risk assessments in the future. The future still is in the hands of medical professionals who are now being supported by the technology to understand their needs and reduce the stress they experience upon.

### Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

### Acknowledgments

This work was supported by AICTE Research funding under the Research Promotion scheme-2019- VIDE NO:F.No.8-35/RIFD/RPS. The authors acknowledge the support of Dr.V.Akila of Department of Computer Science and Engineering, Paducherry Technological University, for her assistance in improving the quality of the manuscript.

### References


