International Journal on Recent Researches in Science, Engineering & Technology (IJRRSET)
A Journal Established in early 2000 as National journal and upgraded to International journal in 2013 and is in existence for the last 10 years. It is run by Retired Professors from NIT, Trichy. Journal Indexed in JIR, DIIF and SJIF.
Available online at: www.jrrset.com

ISSN (Print) : 2347-6729
ISSN (Online) : 2348-3105
JIR IF : 2.54
SJIF IF : 4.334
Cosmos: 5.395

Volume 5, Issue 7 - July 2017 - Pages 100-110

# CARDIO-D21

**[1]Prof Sangamaesh S K [2]Prof Prof Pavitra M Gadhar [3]Prof S T Halakatti**
**Department of Computer Science & Engineering**
**RTE Society's Rural Engineering College Hulkoti, Karnataka 582205**

**ABSTRACT:-** The CVD is detected and displayed in the web application. Cleveland Heart data are taken from the UCI. In this application five algorithms are used with an additional Ensemble model .Random Forest, Logistic Regression, Decision Tree Algorithm, Multinomial_nb, Support vector machine, and Ensemble model are performed for detection purposes. After implementing above mentioned algorithms these are further deployed in python flask framework for graphical user interface and the target result of the prediction to be displayed in the application. The research gaps identified in the literature gives rise to three objectives as follows: Develop accurate Hybrid Data Classification Algorithms. Develop an efficient ensemble model for accurate prediction of CVD. Deploy the constructed models for the real time use. The purpose of the proposed research investigation is to deduce and optimize an algorithm which accurately classifies CVD. In this process, various machine learning models will be developed to give accurate measures for classification of CVD .Further, Ensemble model is to be implemented among machine learning models for accurate prediction of CVD and deployed for production. Deploy the constructed models for the real time use

**Key words:-** CVD, Random Forest, Logistic Regression, Decision Tree Algorithm, Multinomial_nb, Support vector machine, Ensemble model.

## 1.INTRODUCTION

Heart disease is one of the major problems world-wide. Early detection of heart diseases and proper medical treatment can save lives in cases of sudden deaths. According to (WHO), about 17 million people in the world die every year due to cardiovascular diseases. This is about 31% of the total deaths globally. Cardio Vascular disease (CVD) or also known as heart disease include blood and heart of the human body. Myocardial infarction (as a heart attack) is also a part of the CVD. Another type of Heart Disease is called Coronary Heart Disease (CHD), in this type of disease, a substance called Plaque develop in the coronary arteries. Due to digital technologies are rapidly growing, healthcare center's store huge amount of data in their database that is very complex and challenging to analysis.

Data mining techniques and machine learning algorithms play vital roles in analysis of different data in medical centers. The techniques and algorithms can be directly used on a dataset for creating some models or to draw vital conclusions, and inferences from the dataset. With the development of computer technology, machine learning knowledge and clinical practice in the field of predicting disease also in unceasingly thorough. Risk factors of coronary heart disease include family history, smoking, high LDL cholesterol levels, high blood pressure, age, and uncontrolled diabetes. To diagnose heart disease severity in patients, current methods that are used include exercise stress tests, chest X-rays, heart scans (CT), cardiac magnetic resonance imaging (MRI), coronary angiograms, and electrocardiograms (EKG).

International Journal on Recent Researches in Science, Engineering & Technology (IJRRSET)
A Journal Established in early 2000 as National journal and upgraded to International journal in 2013 and is in existence for the last 10 years. It is run by Retired Professors from NIT, Trichy. Journal Indexed in JIR, DIIF and SJIF.
Available online at: www.jrrset.com
ISSN (Print) : 2347-6729
ISSN (Online) : 2348-3105
JIR IF : 2.54
SJIF IF : 4.334
Cosmos: 5.395
Volume 5, Issue 7 - July 2017 - Pages 100-110

Early diagnosis of coronary heart disease can be challenging, and computer-aided methods have been created to detect and diagnose heart disease in patients. Increasingly utilized among computer-aided detection methods in medical institutions is machine learning, a technology that analyzes clinical data, processes it, and provides diagnoses for medical conditions. There is a need for machine learning in cardiovascular medicine. A primary difference between statistical methods and machine learning methods is that the former primarily help to understand relationships between a limited and a number of variables, but the latter contribute to identify and engineer features from the data and perform prediction.

Machine learning methods thus complement and extend existing statistical methods providing tools and algorithms to understand patterns from large, complex and heterogeneous data. Although classical statistical methods are capable of both discovery and prediction, machine learning methods are suitable and generalize able across a variety of data types and offer analyses and interpretation across complex variables. Additionally, machine learning techniques typically rely on fewer assumptions and provide superior and more robust predictions.

From the various existing systems it is understood that there is a need for the following:

1. There is a need to develop hybrid classification algorithms.
2. There is lack of technique to develop efficient ensemble model for accurate prediction of CVD.
3. There is a need to deploy the constructed models for the real time use

**Table 1**: Different types of heart disease

| | |
|---|---|
| Arrhythmia | The heart beat is improper whether it may irregular, too slow or too fast. |
| Cardiac arrest | An unexpected loss of heart function, consciousness and breathing occur suddenly. |
| Congestive heart failure | The heart does not pump blood as well as it should, it is the condition of chronic. |
| Congenital heart disease | The heart's abnormality which develops before birth. |
| Coronary artery disease | The heart's major blood vessels can damage or any disease occurs in the blood vessels. |
| High Blood Pressure | It has a condition that the force of the blood against the artery walls is too high. |
| Peripheral artery disease | The narrowed blood vessels which reduce flow of blood in the limbs, is the circulatory condition. |
| Stroke | Interruption of blood supply occur damage to the brain. |

The system CVD is detected and displayed in the web application. Cleveland Heart data are taken from the UCI. This dataset consists of 303 cases and 76 attributes/features. 13 features are used out of 76 features. In this application five algorithms are used with an additional Ensemble model .Random Forest, Logistic Regression, Decision Tree Algorithm, Multinomial_nb, Support vector machine, and Ensemble model are performed for detection purposes. After implementing above mentioned algorithms these are further deployed in python flask framework for graphical

International Journal on Recent Researches in Science, Engineering & Technology (IJRRSET)
A Journal Established in early 2000 as National journal and upgraded to International journal in 2013 and is in existence for the last 10 years. It is run by Retired Professors from NIT, Trichy. Journal Indexed in JIR, DIIF and SJIF.
Available online at: www.jrrset.com
ISSN (Print) : 2347-6729
ISSN (Online) : 2348-3105
JIR IF : 2.54
SJIF IF : 4.334
Cosmos: 5.395
Volume 5, Issue 7 - July 2017 - Pages 100-110

user interface and the target result of the prediction to be displayed in the application.

The research gaps identified in the literature gives rise to three objectives as follows:

1. Develop accurate Hybrid Data Classification Algorithms.
2. Develop an efficient ensemble model for accurate prediction of CVD.
3. Deploy the constructed models for the real time use.

## II.METHODOLOGY

A system requirements specification is a description of a software system to be developed. It lays out functional and non-functional requirements, and may include a set of use cases that describe user interactions that the software must provide. System requirements specification establishes the basis on what the product is to do as well as what it is not expected to do. It permits a rigorous assessment of requirements before design can begin and reduces later redesign. It should also provide a realistic basis for estimating product costs, risks, and schedules. Used appropriately, it can help prevent software project failure.

Functional requirements are specific functionality that define what a system is supposed to accomplish. Functional requirements are supported by non-functional requirements (also known as quality requirements), which impose constraints on the design or implementation (such as performance requirements, security, or reliability). Generally, functional requirements are expressed in the form "system must do <requirement>".

**Table 2**: Functional Requirements.

| Functionality | Description |
|---|---|
| **User Roles** | ➢ User should be able to login the system.<br>➢ User should be able to enter the patients' health information to predict CVD.<br>➢ User should be able to view the probability of CVD according to the health information entered in the system.<br>➢ User should be able to get the result accurately. |
| **System Roles** | ➢ The system should be able to authenticate users<br>➢ The system should be able to fit patient's health data into various machine learning models and ensemble model.<br>➢ All the algorithms implemented should have accuracy of more than 75 percent.<br>➢ The system should fit the patient's health information in the machine learning models separately.<br>➢ The system should be able to predict the result and display it to the use. |

Non-functional requirement in software system engineering, a software requirement that describes not what the software will do, but how the software will do it, for example, software performance requirements, software external interface requirements, design constraints, and software quality attributes. Non-functional requirements are difficult to test therefore, they are

International Journal on Recent Researches in Science, Engineering & Technology (IJRRSET)
A Journal Established in early 2000 as National journal and upgraded to International journal in 2013 and is in existence for the last 10 years. It is run by Retired Professors from NIT, Trichy. Journal Indexed in JIR, DIIF and SJIF.
Available online at: www.jrrset.com
ISSN (Print) : 2347-6729
ISSN (Online) : 2348-3105
JIR IF : 2.54
SJIF IF : 4.334
Cosmos: 5.395
Volume 5, Issue 7 - July 2017 - Pages 100-110

usually evaluated subjectively. Generally, non-functional requirements are "system shall be <requirement>".

## III.IMPLEMENTATION

### CVD PREDICTION:

The term “heart disease” is often used interchangeably with the term “cardiovascular disease”. Cardiovascular disease generally refers to conditions that involve narrowed or blocked blood vessels that can lead to a heart attack, chest pain (angina) or stroke. Other heart conditions, such as those that affect your heart’s muscle, valves or rhythm, also are considered forms of heart disease.Machine learning (ML) proves to be effective in assisting in making decisions and predictions from the large quantity of data produced by the healthcare industry.

### Algorithms Implemented:

**SVM classifier:**

The objective of the support vector machine algorithm is to find a hyperplane in an N-dimensional space (N — the number of features) that classifies the data point.

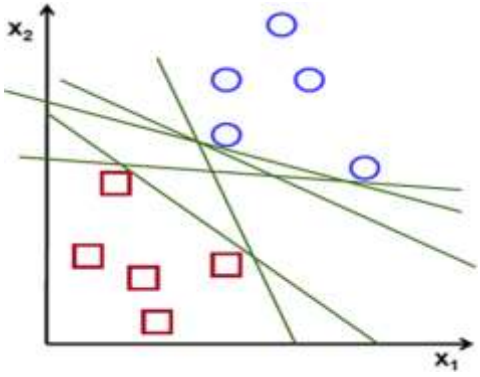



Figure : Possible Hyperplanes

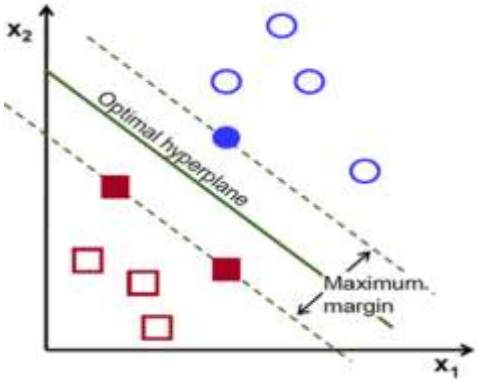



Figure: Optimal Hyperplane

To separate the two classes of data points, there are many possible hyper planes that could be chosen. Our objective is to find a plane that has the maximum margin, i.e the maximum distance between data points of both classes. Maximizing the margin distance provides some reinforcement so that future data points can be classified with more confidence. Support vectors are data points that are closer to the hyper plane and influence the position and orientation of the hyper plane. Using these support vectors, we maximize the margin of the classifier. Deleting the support vectors will change the position of the hyper plane. These are the points that help us build our SVM.

**Kernel SVM:** Simple SVM algorithm can be used to find decision boundary for linearly separable data. However, in the case of non-linearly separable data, such as the one shown in Fig. 2, a straight line cannot be used as a decision boundary. It is method of using linear classifier

International Journal on Recent Researches in Science, Engineering & Technology (IJRRSET)
A Journal Established in early 2000 as National journal and upgraded to International journal in 2013 and is in existence for the last 10 years. It is run by Retired Professors from NIT, Trichy. Journal Indexed in JIR, DIIF and SJIF.
Available online at: www.jrrset.com
ISSN (Print) : 2347-6729
ISSN (Online) : 2348-3105
JIR IF : 2.54
SJIF IF : 4.334
Cosmos: 5.395
Volume 5, Issue 7 - July 2017 - Pages 100-110

to classify non-linear data points. Mathematically, it is Mercer's theorem, which maps non-linear input data points into higher dimension where they can be linearly separable. And kernel is a function which actually perform the above task for us. There are different types of kernel like 'linear', 'polynomial', 'redial basis function' etc. Selecting a right kernel which can best suit your data is obtained by cross-validation.

**Logistic regression classifier:**

Logistic regression is a classification algorithm used to assign observations to a discrete set of classes. Unlike linear regression which outputs continuous number values, logistic regression transforms its output using the logistic sigmoid function to return a probability value which can then be mapped to two or more discrete classes.

**Types of logistic regression**

- Binary (Pass/Fail)
- Multi (Cats, Dogs, Sheep)
- Ordinal (Low, Medium, High)

In this project, binary logistic regression is used as there are only two class labels (1 or 0).

**Binary logistic regression:**

**Sigmoid activation:** In order to map predicted values to probabilities, we use the sigmoid function. The function maps any real value into another value between 0 and 1.

**Math:** $$S(z) = \frac{1}{1+e^{-z}}$$

- s(z) = output between 0 and 1 (probability estimate)
- z = input to the function (your algorithm's prediction e.g. mx + b)
- e = base of natural log

**Graph:**

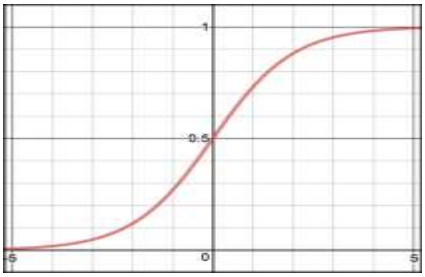



Figure : Logistic Regression Curve

**Decision boundary:**

Our current prediction function returns a probability score between 0 and 1. In order to map this to a discrete class (true/false, cat/dog), we select a threshold value or tipping point above which we will classify values into class 1 and below which we classify values into class 2. i.e, if $p \geq 0.5$, class=1 and if $p < 0.5$, class=0.

**Making predictions:**

International Journal on Recent Researches in Science, Engineering & Technology (IJRRSET)
A Journal Established in early 2000 as National journal and upgraded to International journal in 2013 and is in existence for the last 10 years. It is run by Retired Professors from NIT, Trichy. Journal Indexed in JIR, DIIF and SJIF.
Available online at: www.jrrset.com
ISSN (Print) : 2347-6729
ISSN (Online) : 2348-3105
JIR IF : 2.54
SJIF IF : 4.334
Cosmos: 5.395
Volume 5, Issue 7 - July 2017 - Pages 100-110

Using our knowledge of sigmoid functions and decision boundaries, we can now write a prediction function. A prediction function in logistic regression returns the probability of our observation being positive, True or "Yes". We call this class 1 and its notation is P (class=1). As the probability gets closer to 1, our model is more confident that the observation is in class 1.

**Random Forest classifier:**

Random forest is a supervised learning algorithm which is used for both classification as well as regression. But however, it is mainly used for classification problems. As we know that a forest is made up of trees and more trees means more robust forest. Similarly, random forest algorithm creates decision trees on data samples and then gets the prediction from each of them and finally selects the best solution by means of voting. It is an ensemble method which is better than a single decision tree because it reduces the over-fitting by averaging the result

**Working of Random Forest Algorithm:**

We can understand the working of Random Forest algorithm with the help of following steps:

- **Step 1** − First, start with the selection of random samples from a given dataset.
- **Step 2** − Next, this algorithm will construct a decision tree for every sample. Then it will get the prediction result from every decision tree.
- **Step 3** − In this step, voting will be performed for every predicted result.
- **Step 4** − At last, select the most voted prediction result as the final prediction result.

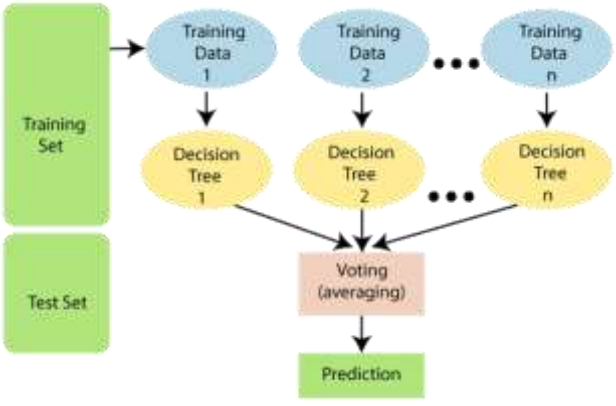



Figure :Random Forest Algorithm

**Decision Tree Classifier:**

Decision Trees (DTs**)** are a non-parametric supervised learning method used for classification and regression. The goal is to create a model that predicts the value of a target variable by learning simple decision rules inferred from the data features.

**Decision tree working:**

International Journal on Recent Researches in Science, Engineering & Technology (IJRRSET)
A Journal Established in early 2000 as National journal and upgraded to International journal in 2013 and is in existence for the last 10 years. It is run by Retired Professors from NIT, Trichy. Journal Indexed in JIR, DIIF and SJIF.
Available online at: www.jrrset.com
ISSN (Print) : 2347-6729
ISSN (Online) : 2348-3105
JIR IF : 2.54
SJIF IF : 4.334
Cosmos: 5.395
Volume 5, Issue 7 - July 2017 - Pages 100-110

The decision of making strategic splits heavily affects a tree's accuracy. The decision criteria are different for classification and regression trees. Decision trees use multiple algorithms to decide to split a node into two or more sub-nodes. The creation of sub-nodes increases the homogeneity of resultant sub-nodes. In other words, we can say that the purity of the node increases with respect to the target variable. The decision tree splits the nodes on all available variables and then selects the split which results in most homogeneous sub-nodes.

**Steps in ID3 algorithm:**

1. It begins with the original set S as the root node.
2. On each iteration of the algorithm, it iterates through the very unused attribute of the set S and calculates **Entropy (H)** and **Information gain (IG)** of this attribute.
3. It then selects the attribute which has the smallest Entropy or Largest Information gain.
4. The set S is then split by the selected attribute to produce a subset of the data.
5. The algorithm continues to recur on each subset, considering only attributes never selected before.

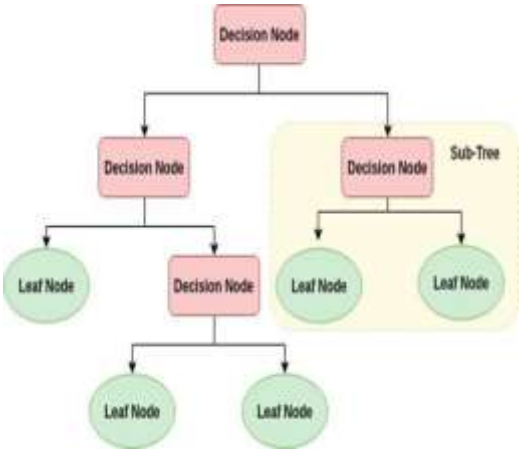



**Figure: Decision tree classification**

**Multinomial Naïve Bayes classifier:**

MultinomialNB implements the naive Bayes algorithm for multinomially distributed data, and is one of the two classic naive Bayes variants used in text classification (where the data are typically represented as word vector counts, although tf-idf vectors are also known to work well in practice). The distribution is parameterized by vectors $\theta_y = (\theta_{y1}, \ldots\ldots, \theta_{yn})$

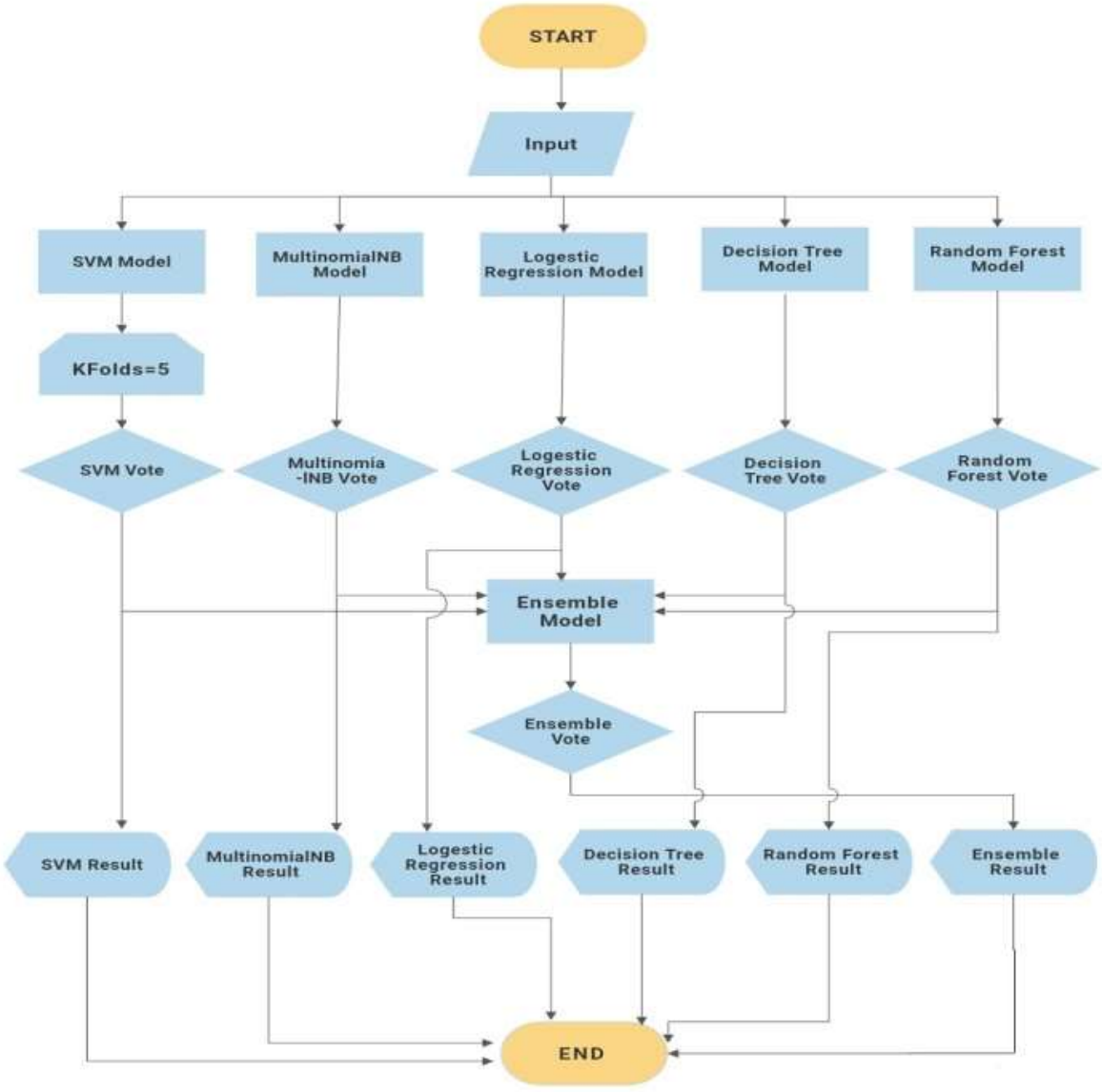
For each class $y$, where $n$ is the number of features (in text classification, the size of the vocabulary) and $\theta_{yi}$is the probability $P(x_i|y)$ of feature$i$appearing in a sample belonging to class $y$.The parameters $\theta_y$is estimated by a smoothed version of maximum likelihood, i.e. relative frequency counting:

$$\widehat{\boldsymbol{\theta}}_{\boldsymbol{yi}} = \frac{\boldsymbol{N_{yi}} + \boldsymbol{\alpha}}{\boldsymbol{N_y} + \boldsymbol{\alpha n}}$$

International Journal on Recent Researches in Science, Engineering & Technology (IJRRSET)
A Journal Established in early 2000 as National journal and upgraded to International journal in 2013 and is in existence for the last 10 years. It is run by Retired Professors from NIT, Trichy. Journal Indexed in JIR, DIIF and SJIF.
Available online at: www.jrrset.com
ISSN (Print) : 2347-6729
ISSN (Online) : 2348-3105
JIR IF : 2.54
SJIF IF : 4.334
Cosmos: 5.395
Volume 5, Issue 7 - July 2017 - Pages 100-110

Where $N_{yi} = \sum_{x \in T} x_i$ is the number of times feature *i* appears in a sample of class *y*in the training set $T$ , and $N_y = \sum_{i=1}^{n} N_{yi}$ is the total count of all features for class *y*.

The smoothing priors $\alpha \geq 0$ accounts for features not present in the learning samples and prevents zero probabilities in further computations. Setting $\alpha = 1$ is called Laplace smoothing, while $\alpha < 1$ is called Lidstone smoothing.



International Journal on Recent Researches in Science, Engineering & Technology (IJRRSET)
A Journal Established in early 2000 as National journal and upgraded to International journal in 2013 and is in existence for the last 10 years. It is run by Retired Professors from NIT, Trichy. Journal Indexed in JIR, DIIF and SJIF.
Available online at: www.jrrset.com
ISSN (Print) : 2347-6729
ISSN (Online) : 2348-3105
JIR IF : 2.54
SJIF IF : 4.334
Cosmos: 5.395
Volume 5, Issue 7 - July 2017 - Pages 100-110

## V.RESULTS

**Support vector machine classifier:**

```
[Parallel(n_jobs=-1)]: Using backend LokyBackend with 4 concurrent workers.
[Parallel(n_jobs=-1)]: Done  15 out of  15 | elapsed:    9.3s finished
              precision    recall  f1-score   support

           0       0.83      0.65      0.73        46
           1       0.71      0.87      0.78        45

    accuracy                           0.76        91
   macro avg       0.77      0.76      0.76        91
weighted avg       0.77      0.76      0.76        91
```

**Figure:** SVM Classification Report

**Logistic regression classifier:**

| | precision | recall | f1-score | support |
|---|---|---|---|---|
| 0 | 0.88 | 0.65 | 0.75 | 46 |
| 1 | 0.72 | 0.91 | 0.80 | 45 |
| accuracy | | | 0.78 | 91 |
| macro avg | 0.80 | 0.78 | 0.78 | 91 |
| weighted avg | 0.80 | 0.78 | 0.78 | 91 |

**Figure:** Logistic regression Classification Report

**Random Forest classifier:**

| | precision | recall | f1-score | support |
|---|---|---|---|---|
| 0 | 0.91 | 0.63 | 0.74 | 46 |
| 1 | 0.71 | 0.93 | 0.81 | 45 |
| accuracy | | | 0.78 | 91 |
| macro avg | 0.81 | 0.78 | 0.78 | 91 |
| weighted avg | 0.81 | 0.78 | 0.78 | 91 |

**Figure: Random Forest Classification Report**

**Decision Tree Classifier:**

| | precision | recall | f1-score | support |
|---|---|---|---|---|
| 0 | 0.91 | 0.65 | 0.76 | 46 |
| 1 | 0.72 | 0.93 | 0.82 | 45 |
| accuracy | | | 0.79 | 91 |
| macro avg | 0.82 | 0.79 | 0.79 | 91 |
| weighted avg | 0.82 | 0.79 | 0.79 | 91 |

**Figure:Decision Tree Classification Report**

**Multinomial-NB Classifier:**

| | precision | recall | f1-score | support |
|---|---|---|---|---|
| 0 | 0.81 | 0.57 | 0.67 | 46 |
| 1 | 0.66 | 0.87 | 0.75 | 45 |
| accuracy | | | 0.71 | 91 |
| macro avg | 0.74 | 0.72 | 0.71 | 91 |
| weighted avg | 0.74 | 0.71 | 0.71 | 91 |

International Journal on Recent Researches in Science, Engineering & Technology (IJRRSET)
A Journal Established in early 2000 as National journal and upgraded to International journal in 2013 and is in existence for the last 10 years. It is run by Retired Professors from NIT, Trichy. Journal Indexed in JIR, DIIF and SJIF.
Available online at: www.jrrset.com
ISSN (Print) : 2347-6729
ISSN (Online) : 2348-3105
JIR IF : 2.54
SJIF IF : 4.334
Cosmos: 5.395
Volume 5, Issue 7 - July 2017 - Pages 100-110

**Figure: Multinomial-NB Classification Report**

**Ensemble model:**

| | precision | recall | f1-score | support |
|---|---|---|---|---|
| 0 | 0.94 | 0.63 | 0.75 | 46 |
| 1 | 0.72 | 0.96 | 0.82 | 45 |
| accuracy | | | 0.79 | 91 |
| macro avg | 0.83 | 0.79 | 0.79 | 91 |
| weighted avg | 0.83 | 0.79 | 0.79 | 91 |

**Figure : Ensemble model Classification Report**

## VI.CONCLUSION

In this project, we addressed the problem of prediction of cardiovascular heart disease by implementing various machine learning algorithms like Support Vector Machine, Logistic regression, Random Forest classifier, Decision Tree Classifier, Multinomial Naïve Bayes classifier. Different types of in built functions have been defined for accurate prediction of CVD. The main focus of our project was on the accurate prediction on the given data. One of the main contributions of our work is to express this task as a combinatorial optimization problem with constraints, and to propose methods to solve it based on ensemble learning technique which uses max voting technique in order to combine all 5 above mentioned models. In this project even prediction is provided by machine learning models individually using various performance measures. In particular, we proposed this project on web application using various web technologies to provide applicable graphical user interface. The proposed application uses MongoDB for database operations .Finally, the prediction results are displayed on web application.

## REFERENCES

[1].S. Mohan, C. Thirumalai, and G. Srivastava, "Effective heart disease prediction using hybrid machine learning techniques," vol. 7. IEEE, 2019, pp. 81 542–81 554.
[2].J. Bektas, T. Ibrikçi, and I. TurkayOzcan, "The impact of imputation procedures with machine learning methods on the performance of classifiers: An application to coronary artery disease data including missing values," Biomedical Research, vol. 29, no. 13, pp. 2780–2785, 2018.
[3].W. Shahzad, Q. Rehman, and E. Ahmed, "Missing data imputation using genetic algorithm for supervised learning," International Journal of Advanced Computer Science and Applications (IJACSA), no. 3, 2017.
[4].K. Deeba and B. Amutha, "Classification algorithms of data mining," Indian Journal of Science and Technology, vol. 9(39), 2016.
[5].S. Joshi and M. K. Nair, "Prediction of heart disease using classification based data mining techniques," in Computational Intelligence in Data Mining-Volume 2. Springer, 2015, pp. 503–511
[6]. H. D. Masethe and M. A. Masethe, "Prediction of heart disease using classification

International Journal on Recent Researches in Science, Engineering & Technology (IJRRSET)
A Journal Established in early 2000 as National journal and upgraded to International journal in 2013 and is in existence for the last 10 years. It is run by Retired Professors from NIT, Trichy. Journal Indexed in JIR, DIIF and SJIF.
Available online at: www.jrrset.com
ISSN (Print) : 2347-6729
ISSN (Online) : 2348-3105
JIR IF : 2.54
SJIF IF : 4.334
Cosmos: 5.395
Volume 5, Issue 7 - July 2017 - Pages 100-110

algorithms," in Proceedings of the world Congress on Engineering and computer Science, vol. 2, 2014, pp. 22–24.

[7].S. Sa, "Intelligent heart disease prediction system using data mining techniques," International Journal of healthcare & biomedical Research, vol. 1, pp. 94–101, 2013.

[8].T. R. Patil, S. Sherekar et al., "Performance analysis of naive bayes and j48 classi- fication algorithm for data classification," International journal of computer science and applications, vol. 6, no. 2, pp. 256–261, 2013

[9].D. Tomar and S. Agarwal, "A survey on data mining approaches for healthcare," International Journal of Bio-Science and Bio-Technology, vol. 5, no. 5, pp. 241–266, 2013.

[10].Norma latiffitriyani , Muhammad syafrudin ,Ganjaralfian (member, ieee), and jongtae rhee1, "HDPM: An Effective Heart Disease Prediction Model for a Clinical Decision Support System "Department of Industrial and Systems Engineering, Dongguk University, Seoul 04620, South Korea,2020

[11].Jikuowang, changchunliu, liping li, wang li, lianke yao1,han li1, and huanzhang,"A Stacking-Based Model for Non-Invasive Detection of Coronary Heart Disease" School of Control Science and Engineering, Shandong University, Jinan 250061, China,2020

[12**]**.Chunyanguo , jiabingzhang , yang liu , yayingxie ,Zhiqianghan , and jiansheyu "Recursion Enhanced Random Forest With an Improved Linear Model (RERF-ILM) for Heart Disease Detection on the Internet of Medical Things Platform "Anesthesiology Department, The Affiliated Hospital of Inner Mongolia Medical University,