The early inquisition of cardiovascular problems using advanced machine learning Techniques

Adeed Sayeed Abbas [1] Mohammed Waheeduddin Hussain[2] Ravichandran.M [3]

[1] Research Scholar, Department of Computer Science and Engineering, Shadan College of Engineering and Technology, Hyderabad, Telangana, India – 500086.

[2] Professor, Department of Information Technology, Shadan College of Engineering and Technology, Hyderabad, Telangana, India – 500086.

[3] Professor, Department of Computer Science and Engineering, Shadan College of Engineering and Technology, Hyderabad, Telangana, India – 500086.

-- ***--

Abstract: - Health care professionals face a challenging task of predicting and detecting heart diseases. Many hospitals have expensive mechanisms and machines in order to detect heart disease in a patient. Heart diseases have now become a common illness and needs to be diagnosed in a cost-effective way so that what health care facilities are reachable to the common man. Angiography is one of those expensive techniques that is used to detect heart disease. However, angiography is generally done at a very later stage almost just before an operation to detect the blocks in the arteries of the heart. If heart disease can be predicted at a very early stage, the patient's health can be safeguarded, and medical expenses could be reduced significantly when a patient takes appropriate steps to reduce the risk of heart disease. Some of the main and common reasons that are supposed to cause heart diseases are consuming alcohol, smoking, high calorie diets data rich in empty carbohydrates and lifestyle factors such as lack of exercise. The data generated by the healthcare industry has been studied over the years and machine learning has proved to be an effective in making decisions and predicting the results. During the study some of the algorithms such as logistic regression, decision trees, support vector machines, artificial neural networks etc have been used to create machine learning models for the heart disease data set. we have also recorded the performance of various algorithms and summarized them. We also aim to deploy a practical machine learning model using the best algorithm that suits the scenario to a web application where a user can register himself and give his inputs about his health condition so that he can get predictions a weather he's susceptible to heart disease or not.

Keywords:- *Machine learning, supervised learning, , heart disease, health care services.*

---***---

I INTRODUCTION

One of the most important body parts that is necessary for proper functioning of the whole body after the brain is heart. The functionality of the heart is to supply oxygenated blood to all the body parts and supply deoxygenated blood to the lungs for the purpose of oxygenation[1]. Any issue with the heart can be life threatening and might lead to major surgeries. there are several types of issues or diseases that might occur with the heart such as failure of the heart, blockage of coronary arteries , structural defects in the heart, Heart rhythm disorders etc. It is very difficult to diagnose a heart disease because the symptoms are not readily visible and a patient needs highly qualified medical practitioners, lab technicians and costly equipment

diagnose any issue with the heart[2]. Angiography is one such techniques which is expensive and invasive in nature. These issues are generally diagnosed at a very later stage almost just before a heart operation like bypass surgery or implantation of stents[3]. These costly procedures are not affordable by the common man and we observe a significant rise in heart diseases India existing population[4]. Hence it would be very advantageous if the heart disease is can be diagnosed and predicted at a very early stage so that the patients at risk can take appropriate actions to be hale and healthy. The medical expenditures would also be reduced significantly[5].

Machine learning is one of the techniques using which heart diseases can be predicted at a very early stage with high accuracy[6]. the medical datasets created in due course of time after studying a huge number of patients can be fed to the machine learning models in order to predict the pre-existence of heart disease at a very early stage[7]. Some of the common attributes or features of data that are supposed to cause or have a relationship with the existence of heart disease are:

1. age of the patient
2. gender of the patient
3. blood pressure
4. type of chest pain
5. cholesterol levels
6. blood sugar levels
7. heart rate
8. ECG results etc

In our project, we have used the heart disease data set from the UCI repository to create an effective and accurate heart disease prediction model. We have analyzed the data using the techniques of exploratory data analysis and understood the relationship between various attributes and the existence of heart disease. We have also fed the data set to multiple efficient and effective machine learning algorithms like logistic regression, K nearest neighbor, support vector machine, Naïve Bayes, decision trees and random forest algorithms. The results and graphical analysis has been summarized in the pages below. We have also tried to host this heart disease prediction service on a web server so that it is accessible to the public as

an attempt to reduce the medical costs. A user needs to give his health parameters in order to get a prediction of whether he has a heart disease or not.

AIM OF THE PROJECT

The main of the project is to predict the risk of heart disease based on patient's health parameters. We have used the Cleveland Heart disease dataset from the UCI repository and fed it to the proposed model that is built using advanced machine learning techniques

SCOPE OF THE PROJECT

The scope of the project is limited to the compute the accuracy of the proposed model and predict the risk of heart disease for a patient. The admin of the system analyzes the dataset, compares the accuracy of the dataset on multiple algorithms and creates a suitable machine learning model by training and testing the proposed model with training data. The users of the system can register themselves on the system, get predictions based on health parameters and save the results to the database.

II LITERATURE SURVEY

2.1 "An Optimized Stacked Support Vector Machines Based Expert System for the Effective Prediction of Heart Failure"

Providing health care services especially related to heart is difficult in developing countries due to the dearth of diagnostic tools as they are costly and lack of qualified health care professionals[8]. The situation is becoming worse day by day as the people who areat the risk of heart disease are increasing day by day. Some of the conventional methods used for diagnosis of heart disease such as angiography are costly, invasive and have many side effects associated with it. It also needs costly equipment and qualified medical experts to diagnose the blockage of arteries through angiography[9].

Research Paper   ,

When the existing systems are failing, machine learning has come up with non-invasive heart disease prediction mechanism in which the health data gathered over the years can be used to feed the machine learning models to generate accurate predictions of whether a patient is that a risk of heart disease or not.
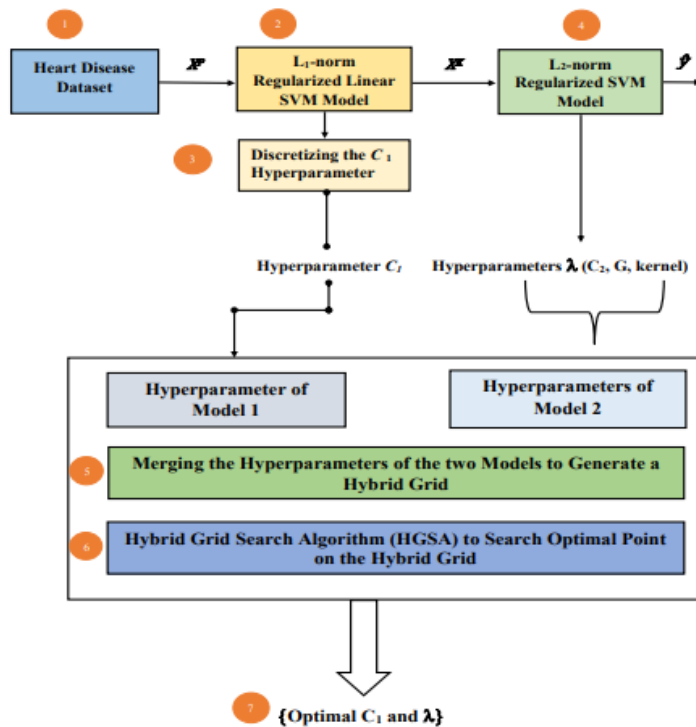
In this paper, the concept of Stacked SVMs has been implemented in order to diagnose heart disease[10]. The purpose of the first SVM is to remove the unnecessary features that do not affect the preexistence of heart disease and the purpose of the second SVM is to get trained and predict the risk of heart disease. Hybrid grid search algorithm has been used in this process for maximal accuracy.

2.2 "An Intelligent Learning System Based on Random Search Algorithm and Optimized Random Forest Model for Improved Heart Disease Detection"

With the changing lifestyles and unhealthy diets, heart failure and existence of heart disease has become quite common nowadays. However diagnosing a heart disease is very costly affair because it needs costly medical equipment and qualified medical professionals and invasive procedures to detect anomalies in the functioning of the heart. This has become a challenging activity in underdeveloped countries due to the lack of sufficient resources and the population there may not be able to afford such high expenditures for health care. Hence research has been progressed to develop intelligent systems that can predict heart disease is at very early state. But the problem with existing automated systems is that the models created may not be good enough for the new data that is being predicted or tested for.

In this paper, a new technique has been proposed to diagnose the heart failure. This system has implemented random search algorithm [RSA] to find out the features that are relevant or have a relationship with the preexistence of heart disease. The main model uses random forest algorithm and the whole process has been optimized using grid search algorithm. This novel system has proved Toby more efficient than the existing automated systems.

.

III SYSTEM ANALYSIS

Problem Statement:

Most of the existing systems with respect to pre investigation of heart diseases cannot identify without invasive methods like angiography. The cost associated with these procedures is also huge. we aim to reduce these costs and try to effectively predict the pre-existence of heart diseases without any invasive procedures using machine learning techniques.

Existing system:

Angiography is an invasive technique used for the prediction of heart diseases. However, there are some drawbacks associated with angiography technique. It is an expensive procedure and physicians must analyze so many factors to diagnose a patient hence this process makes physician job very difficult, so these limitations motivate to develop a

non-invasive method for prediction of heart disease. These conventional methods deal with medical reports of the patients moreover these conventional methods are time-consuming, and it may give erroneous results because these conventional methods are performed by humans.

Disadvantages:
- Invasive technique
- Prone to errors

Proposed System:

We propose to develop and deploy a web application which can predict the susceptibility of heart disease of a patient based on various health parameters. The machine learning model would be deployed to web where the user can upload and view his health parameters and receive predictions. We aim to build an automated prediction system which can accurately classify if the person is susceptible to heart diseases or not.

Advantages:
- Automated system
- Non-invasive
- Can predict heart diseases at early stage

IV IMPLEMENTATION

Below is the proposed modular implementation of the project.
It consists of two modules:
1. Admin
2. User

Admin Module:

The admin of the system is responsible for the activities like:
1. Uploading the dataset
2. Data Analysis of the dataset
3. Splitting the dataset for training and testing
4. Training the model for multiple algorithms
5. Review the performance of the algorithms on the given dataset
6. Create the model using Linear SVM algorithm.
7. View user details and predictions

User Module:

The user of the system can utilize the machine learning services that are offered like:
1. Logging into the system
2. Save one's health related parameters to the database
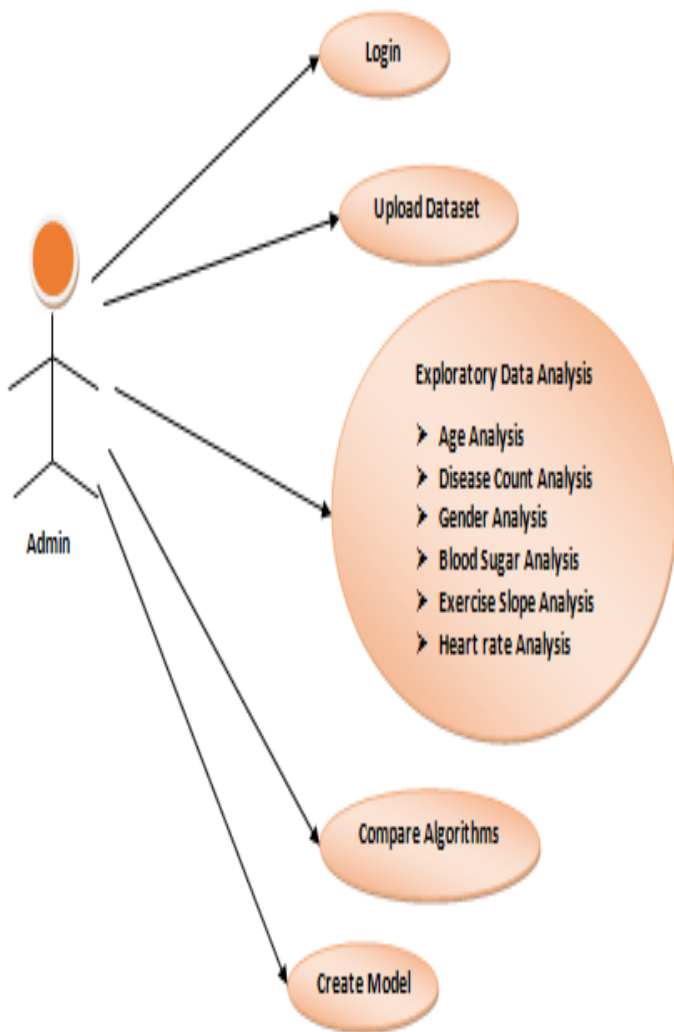3. Receive predictions about the health status.
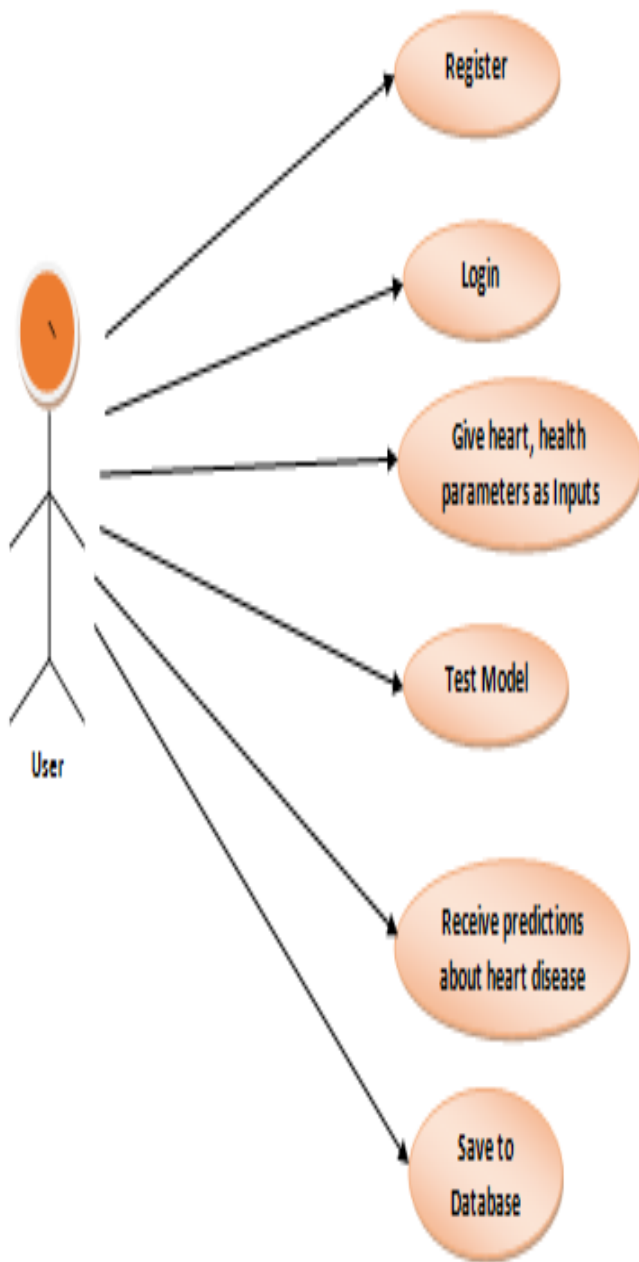
V SYSTEM DESIGN
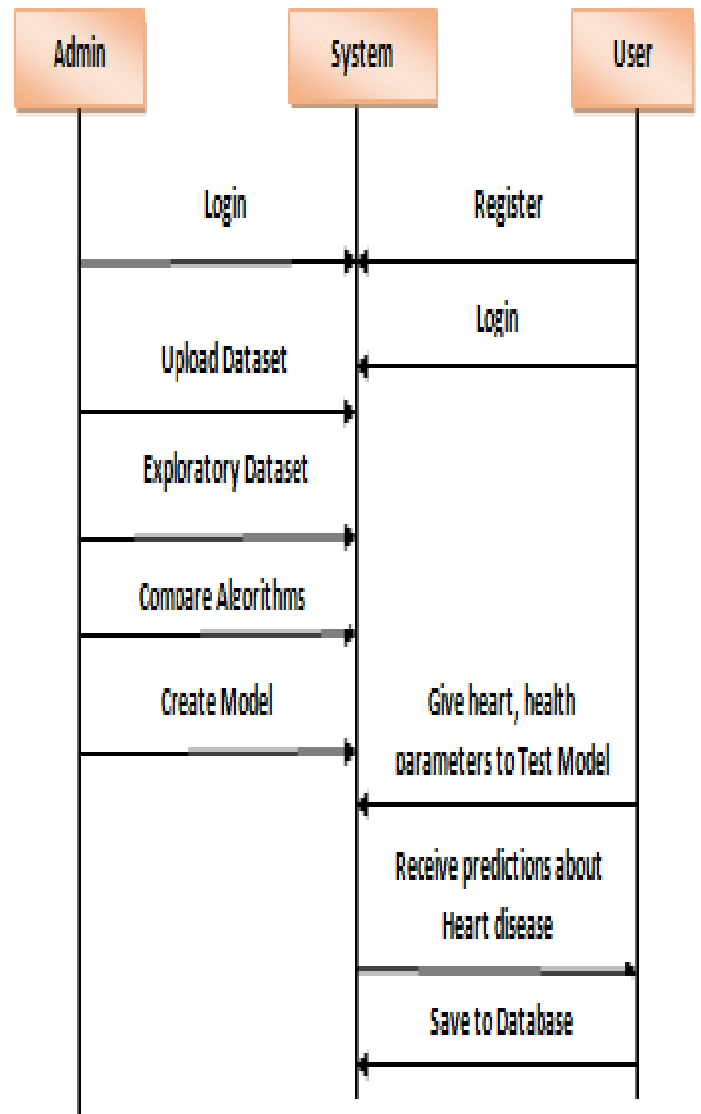
Data Flow Diagram: Admin
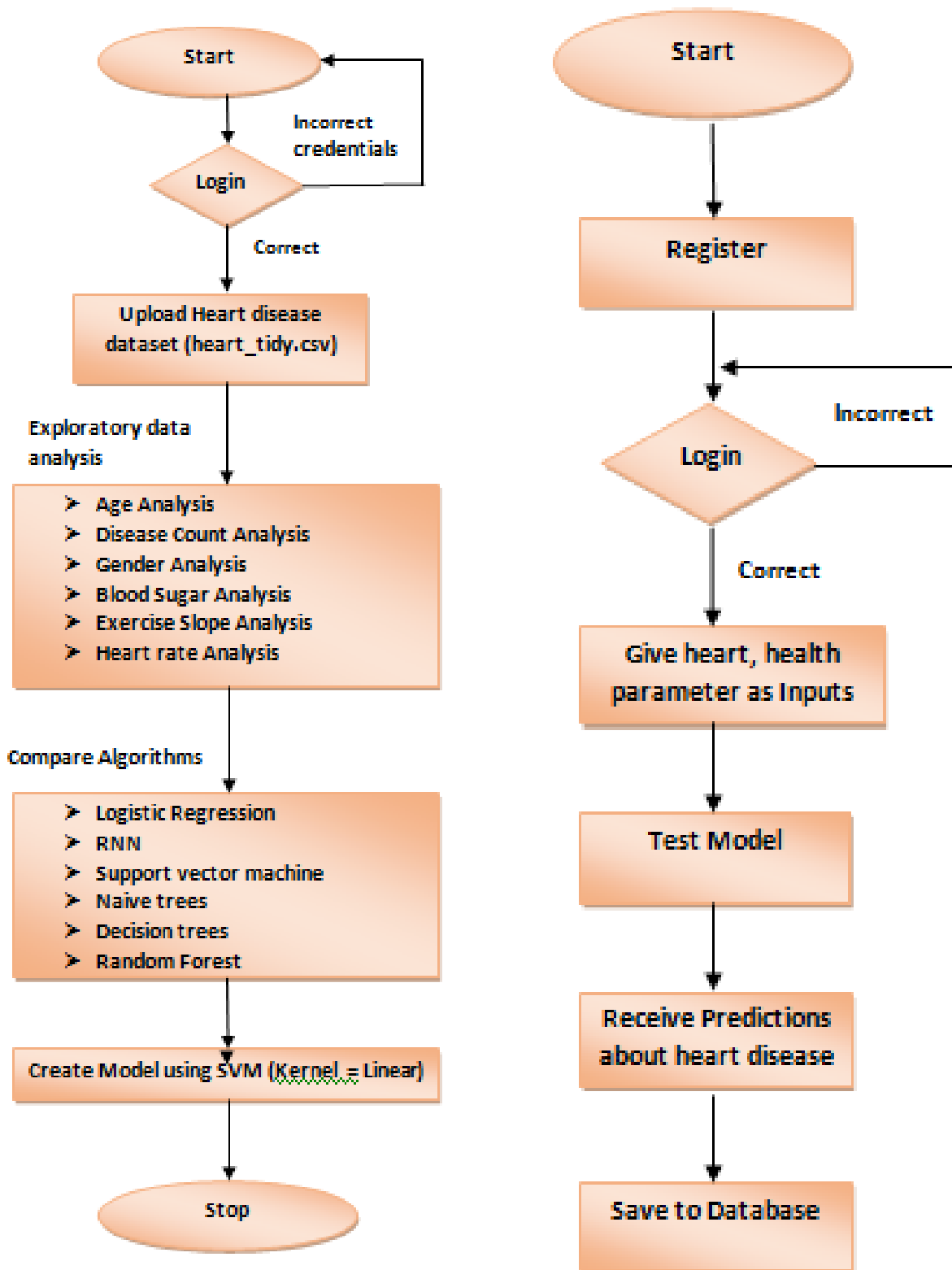
Data flow Diagram: User

Use Case Diagram: Admin
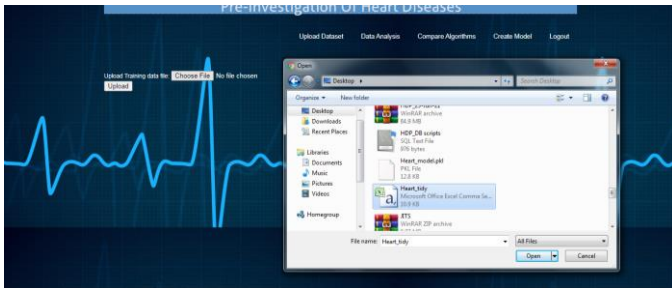
Use Case: User

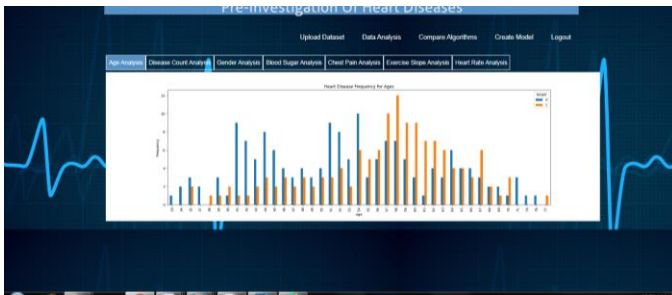**Sequence Diagram:**

VI PROJECT EXECUTION
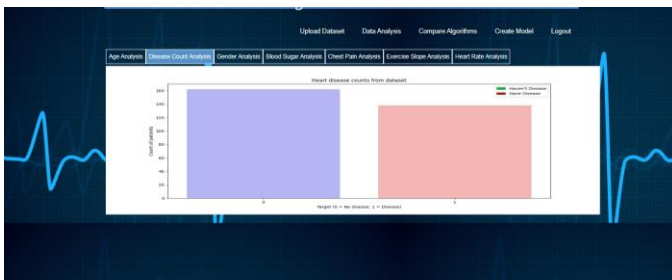
Upload Dataset:

Data Analysis:

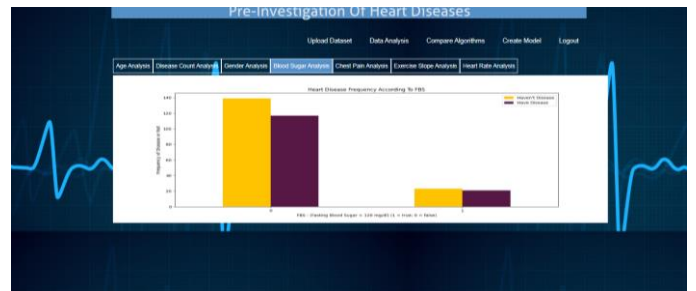- **Age Analysis:**

- **Disease Count Analysis:**

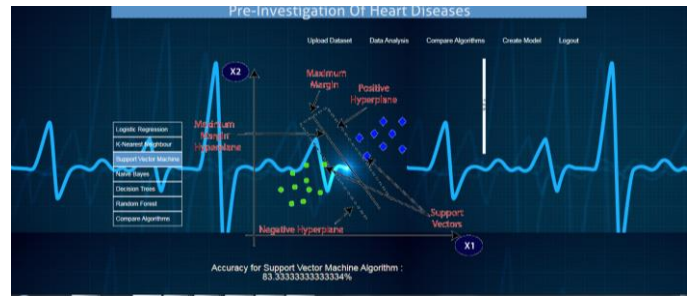- **Gender Analysis:**

- **Blood Sugar Analysis:**

- **Chest pain Analysis:**

- **Exercise Slope Analysis:**

- **Heart Rate Analysis:**

Compare Algorithms:
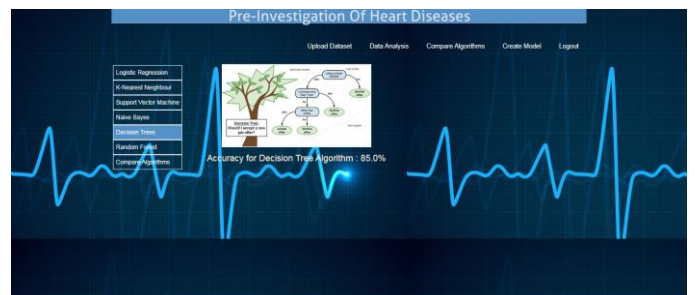
- **Logistic Regression:**

- **K-Nearest Neighbor:**

- **Support Vector Machine:**
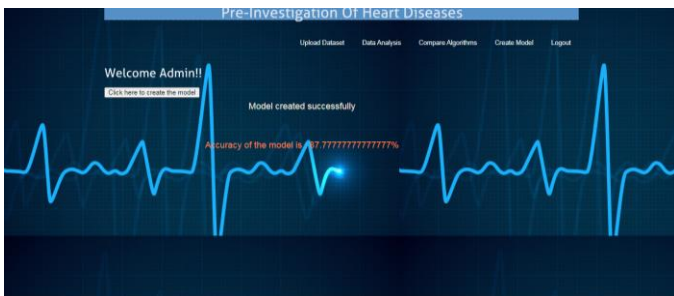

- **Naive Bayes:**

- **Decision Trees:**

- **Random Forest**

- **Compare Algorithms:**

Create Model:

Result page:

User Registration:

View Details:

User Login:

VII CONCLUSION

Nowadays heart disease has become a serious issue. To prevent this at a very early stage we need to predict them through an automated system. These automated systems are very helpful to people to know about their health issues and the predictions that are obtained through these automated systems helps the physicians to diagnose the patients effectively. During the study some of the algorithms such as logistic regression,

Predict:

Research Paper

decision trees, support vector machines, artificial neural networks etc have been used to create machine learning models for the heart disease data set. we have also recorded the performance of various algorithms and summarized them. We observe that Linear SVM after feature selection and parameter tuning has resulted in the best accuracy. We have deployed the model to a flask web application where the user should be able to enter their health parameters and get the predictions. A successful model that is deployed to the web would be helpful for both doctor and general public.

Future work: Feature selection and prediction, these two are essential for every automated system. By choosing features efficiently, we can achieve better results in predicting heart disease. We have summarized some algorithms which are useful while selecting the features, like hybrid grid search algorithm and random search algorithm, etc. So, in the future, it is better to use search algorithms for selecting the features and then applying machine learning techniques for prediction will give us better results in the prediction of heart disease.

REFERENCES

[1] L. Ali et al., "An Optimized Stacked Support Vector Machines Based Expert System for the Effective Prediction of Heart Failure," IEEE Access, vol. 7, pp. 54007–54014, 2019, doi: 10.1109/ACCESS.2019.2909969.

[2] A. Javeed, S. Zhou, L. Yongjian, I. Qasim, A. Noor, and R. Nour, "An Intelligent Learning System Based on Random Search Algorithm and Optimized Random Forest Model for Improved Heart Disease Detection," IEEE Access, vol. 7, pp. 180235– 180243, 2019, doi: 10.1109/ACCESS.2019.2952107.

[3] M. Gjoreski, A. Gradisek, B. Budna, M. Gams, and G. Poglajen, "Machine Learning and End-to-End Deep Learning for the Detection of Chronic Heart Failure from Heart Sounds," IEEE Access, vol. 8, pp. 20313–20324, 2020, doi: 10.1109/ACCESS.2020.2968900.

[4] L. Ali, A. Rahman, A. Khan, M. Zhou, A. Javeed, and J. A. Khan, "An Automated Diagnostic System for Heart Disease Prediction Based on χ2 Statistical Model and Optimally Configured Deep Neural Network," IEEE Access, vol. 7, pp. 34938–34945, 2019, doi: 10.1109/ACCESS.2019.2904800.

[5] M. R. Ahmed, S. M. Hasan Mahmud, M. A. Hossin, H. Jahan, and S. R. Haider Noori, "A cloud based four-tier architecture for early detection of heart disease with machine learning algorithms," 2018 IEEE 4th Int. Conf. Comput. Commun. ICCC 2018, pp. 1951–1955, 2018, doi: 10.1109/CompComm.2018.8781022.

[6] "types of heart disease." [Online]. Available: https://www.heartandstroke.ca/heart/what-is-heart-disease/typesof-heart-disease.

[7] J. Schmidhuber, "Deep Learning in neural networks: An overview," Neural Networks, vol. 61, pp. 85–117, 2015, doi: 10.1016/j.neunet.2014.09.003.

[8] N. H. Farhat, "Photonit neural networks and learning mathines the role of electron-trapping materials," IEEE Expert. Syst. their Appl., vol. 7, no. 5, pp. 63–72, 1992, doi: 10.1109/64.163674.

[9] A. K. M Sazzadur Rahman, M. Mehedi Hasan, S. Asaduzzaman, M. Asaduzzaman, and S. Akhter Hossain, "An analysis of computational intelligence techniques for diabetes prediction Machine Learning View project An analysis of computational intelligence techniques for diabetes prediction," Int. J. Eng. &Technology, vol. 7, no. 4, pp. 6229–6232, 2018, doi: 10.14419/ijet.v7i4.28245.

[10] G. H. Tang, A. B. M. Rabie, and U. Hägg, "Indian hedgehog: A mechanotransduction mediator in condylar cartilage," J. Dent. Res., vol. 83, no. 5, pp. 434–438, 2004, doi: 10.1177/154405910408300516.

.