

VISVESVARAYA TECHNOLOGICAL UNIVERSITY



BELAGAVI – 590018, Karnataka

INTERNSHIP REPORT

On

“HEART DISEASE PREDICTION”

Submitted in partial fulfilment for the award of degree(21INT63)

BACHELOR OF ENGINEERING IN INFORMATION SCIENCE AND ENGINEERING

Submitted by:

ADARSHA K N

1DT22IS402



Conducted at

VARCONS TECHNOLOGIES PVT LTD



DAYANAND SAGAR ACADEMY OF TECHNOLOGY AND MANAGEMENT

Department of Information Science And Engineering

Accredited by NBA, New Delhi

Udayapura ,Kanakapura road ,Bangalore-560082

**DAYANAND SAGAR ACADEMY OF
TECHNOLOGY AND MANAGEMENT**
Department of Information Science And Engineering
Accredited by NBA, New Delhi
Udayapura ,Kanakapura road ,Bangalore-560082



CERTIFICATE

This is to certify that the Internship titled “**HEART DISEASE PREDICTION** ” carried out by **Adarsha K N (1DT22IS402)** a bonafide student of Dayananda Sagar Academy of Technology And Management, in partial fulfillment for the award of **Bachelor of Engineering**, in **INFORMATION SCIENCE AND ENGINEERING** under Visvesvaraya Technological University, Belagavi, during the year 2022-2023. It is certified that all corrections/suggestions indicated have been incorporated in the report.

The project report has been approved as it satisfies the academic requirements in respect of Internship prescribed for the course Internship / Professional Practice (21CSI85)

Signature of Guide

Signature of HOD

Signature of Principal

External Viva:

Name of the Examiner

Signature with Date

1)_____

2)_____

D E C L A R A T I O N

I am, **Adarsha KN**, Second year student of Information Science And Engineering, Dayananda Sagar Academy Of Technology And Management Bangalore - 560082, declare that the Internship has been successfully completed, in **VARCONS TECHNOLOGIES PVT LTD**. This report is submitted in partial fulfillment of the requirements for award of Bachelor Degree in Information Science And Engineering, during the academic year 2022-2023.

Date : _____ :

Place :

USN: 1DT22IS402

NAME : ADARSHA K N

OFFER LETTER



Varcons Technologies Pvt. Ltd.

Date: 25th October, 2023

Name: Adarsha K N
USN: 1DT22IS402

Dear Student,

We would like to congratulate you on being selected for the **Machine Learning With Python (Research Based)** Internship position with **Varcons Technologies**, effective Start Date **25th October, 2023**. All of us are excited about this opportunity provided to you!

This internship is viewed as being an educational opportunity for you, rather than a part-time job. As such, your internship will include training/orientation and focus primarily on learning and developing new skills and gaining a deeper understanding of concepts of **Machine Learning With Python (Research Based)** through hands-on application of the knowledge you learn while you train with the senior developers. You will be bound to follow the rules and regulations of the company during your internship duration.

Again, congratulations and we look forward to working with you!.

Sincerely,

Spoorthi H C
Director
VARCONS TECHNOLOGIES
213, 2nd Floor,
18 M G Road, Ulsoor,
Bangalore-560001

A C K N O W L E D G E M E N T

This Internship is a result of accumulated guidance, direction and support of several important persons. We take this opportunity to express our gratitude to all who have helped us to complete the Internship.

We express our sincere thanks to our Principal, for providing usadequate facilities to undertake this Internship.

We would like to thank our Head of Dept – Information Science Engineering, for providing us an opportunity to carry out Internship and for his valuable guidance and support.

We would like to thank our Software Services for guiding us during the period of internship.

We express our deep and profound gratitude to our guide, Guide name, Assistant/Associate Prof, for her keen interest and encouragement at every step in completing the Internship.

We would like to thank all the faculty members of our department for the support extended during the course of Internship.

We would like to thank the non-teaching members of our dept, forhelping us during the Internship.

Last but not the least, we would like to thank our parents and friends without whose constant help, the completion of Internship would have not been possible.

ADARSHA KN

1DT22IS402

ABSTRACT

Heart plays significant role in living organisms. Heart disease is one of the most significant causes of mortality in the world today. Prediction of cardio vascular disease is a critical challenge in the area of clinical data analysis. Diagnosis and prediction of heart related diseases requires more precision, perfection and correctness because a little mistake can cause fatigue problem or death of the person, there are numerous death cases related to heart and their counting is increasing exponentially day by day. Machine learning has been shown to be effective in assisting in making decisions and predictions from the large quantity of data produced by the health care industry. Various studies give only a glimpse into predicting heart disease with ML techniques. Here, we design a model that aims at finding significant features by applying machine learning techniques resulting in improving the accuracy in the prediction of heart disease. There are many to do accomplish this task successfully, but how effective are they? Our main aim is to produce an enhanced performance level with the good accuracy level through the prediction model for heart disease with the SVM, KNN, Naïve Bayes, logistic regression & Random Forest algorithms. It is estimated that on an average about 17 million people die of cardiovascular diseases each year, which is about one third of total deaths across the globe. In this proposed project we designed a model to detect and predict the accuracy of heart disease. So with that we have to upgrade with the technology constantly, it is becoming easy to track the behaviour and pattern diseases and get cure at the early stages. To come up with the solution one can make use of technologies with the increase of machine learning so it becomes feasible to automate this process and to save someone's life by detecting the disease at an early stages. Initially, we will collect the data set by users and classify it as trained and testing dataset using different type's algorithm and decision trees. By using the feasible algorithm, we can analyze the larger data-set and user provided current data set. Then augment the accuracy of the result data. Proceeded with the application of processing of some of the attributes provided which can find whether the user is having heart disease in viewing the graphical model of data visualization. The performance of the techniques is gauged based on accuracy, sensitivity, and specificity, precision. The results are indicated concerning the best accuracy for Random Forest is unit 80% respectively.

Table of Contents

Sl no	Description	Page no
1	Company Profile	
2	About the Company	
3	Introduction	
4	System Analysis	
5	Requirement Analysis	
6	Design Analysis	
7	Implementation	
8	Snapshots	
9	Conclusion	
10	References	

1. COMPANY PROFILE

A Brief History of Compsoft Technologies

Compsoft Technologies, was incorporated with a goal "To provide high quality and optimal Technological Solutions to business requirements of our clients". Every business is a different and has a unique business model and so are the technological requirements. They understand this and hence the solutions provided to these requirements are different as well. They focus on clients requirements and provide them with tailor made technological solutions. They also understand that Reach of their Product to its targeted market or the automation of the existing process into e-client and simple process are the key features that our clients desire from Technological Solution they are looking for and these are the features that we focus on while designing the solutions for their clients.

Sarvamoola Software Services. is a Technology Organization providing solutions for all web design and development, MYSQL, PYTHON Programming, HTML, CSS, ASP.NET and LINQ. Meeting the ever increasing automation requirements, Sarvamoola Software Services. specialize in ERP, Connectivity, SEO Services, Conference Management, effective web promotion and tailor-made software products, designing solutions best suiting clients requirements.

Compsoft Technologies, strive to be the front runner in creativity and innovation in software development through their well-researched expertise and establish it as an out of the box software development company in Bangalore, India. As a software development company, they translate this software development expertise into value for their customers through their professional solutions.

They understand that the best desired output can be achieved only by understanding the clients demand better. Compsoft Technologies work with their clients and help them to define their exact solution requirement. Sometimes even they wonder that they have completely redefined their solution or new application requirement during the brainstorming session, and here they position themselves as an IT solutions consulting group comprising of high caliber consultants.

They believe that Technology when used properly can help any business to scale and achieve new heights of success. It helps Improve its efficiency, profitability, reliability; to put it in one sentence " Technology helps you to Delight your Customers" and that is what we want to achieve.

CHAPTER 2

ABOUT THE COMPANY



Compsoft Technologies is a Technology Organization providing solutions for all web design and development, MYSQL, PYTHON Programming, HTML, CSS, ASP.NET and LINQ. Meeting the ever increasing automation requirements, Compsoft Technologies specialize in ERP, Connectivity, SEO Services, Conference Management, effective webpromotion and tailor-made software products, designing solutions best suiting clients requirements. The organization where they have a right mix of professionals as a stakeholders to help us serve our clients with best of our capability and with at par industry standards. They have young, enthusiastic, passionate and creative Professionals to develop technological innovations in the field of Mobile technologies, Web applications as well as Business and Enterprise solution. Motto of our organization is to “Collaborate with our clients to provide them with best Technological solution hence creating Good Present and Better Future for our client which will bring a cascading a positive effect in their business shape as well”. Providing a Complete suite of technical solutions is not just our tag line, it is Our Vision for Our Clients and for Us, We strive hard to achieve it.

Products of Compsoft Technologies.

Android Apps

It is the process by which new applications are created for devices running the Android operating system. Applications are usually developed in Java (and/or Kotlin; or other such option) programming language using the Android software development kit (SDK), but other development environments are also available, some such as Kotlin support the exact same Android APIs (and bytecode), while others such as Go have restricted API access.

The Android software development kit includes a comprehensive set of development tools. These include a debugger, libraries, a handset emulator based on QEMU, documentation, sample code, and tutorials. Currently supported development platforms include computers running Linux (any modern desktop Linux distribution), Mac OS X 10.5.8 or later, and Windows 7 or later. As of March 2015, the SDK is not available on Android itself, but software development is possible by using specialized Android applications.

Web Application

It is a client-server computer program in which the client (including the user interface and client-side logic) runs in a web browser. Common web applications include web mail, online

retail sales, online auctions, wikis, instant messaging services and many other functions. web applications use web documents written in a standard format such as HTML and JavaScript, which are supported by a variety of web browsers. Web applications can be considered as a specific variant of client-server software where the client software is downloaded to the client machine when visiting the relevant web page, using standard procedures such as HTTP. The Client web software updates may happen each time the web page is visited. During the session, the web browser interprets and displays the pages, and acts as the universal client for any web application. The use of web application frameworks can often reduce the number of errors in a program, both by making the code simpler, and by allowing one team to concentrate on the framework while another focuses on a specified use case. In applications which are exposed to constant hacking attempts on the Internet, security-related problems can be caused by errors in the program.

Frameworks can also promote the use of best practices such as GET after POST. There are some who view a web application as a two-tier architecture. This can be a “smart” client that performs all the work and queries a “dumb” server, or a “dumb” client that relies on a “smart” server. The client would handle the presentation tier, the server would have the database (storage tier), and the business logic (application tier) would be on one of them or on both. While this increases the scalability of the applications and separates the display and the database, it still doesn’t allow for true specialization of layers, so most applications will outgrow this model. An emerging strategy for application software companies is to provide web access to software previously distributed as local applications. Depending on the type of application, it may require the development of an entirely different browser-based interface, or merely adapting an existing application to use different presentation technology. These programs allow the user to pay a monthly or yearly fee for use of a software application without having to install it on a local hard drive. A company which follows this strategy is known as an application service provider (ASP), and ASPs are currently receiving much attention in the software industry.

Security breaches on these kinds of applications are a major concern because it can involve both enterprise information and private customer data. Protecting these assets is an important part of any web application and there are some key operational areas that must be included in the development process. This includes processes for authentication, authorization, asset handling, input, and logging and auditing. Building security into the applications from the beginning can be more effective and less disruptive in the long run.

Web design

It encompasses many different skills and disciplines in the production and maintenance of websites. The different areas of web design include web graphic design; interface design; authoring, including standardized code and proprietary software; user experience design; and

search engine optimization. The term web design is normally used to describe the design process relating to the front-end (client side) design of a website including writing mark up. Web design partially overlaps web engineering in the broader scope of web development. Web designers are expected to have an awareness of usability and if their role involves creating mark up then they are also expected to be up to date with web accessibility guidelines. Web design partially overlaps web engineering in the broader scope of web development.

Departments and services offered

Compsoft Technologies plays an essential role as an institute, the level of education, development of student's skills are based on their trainers. If you do not have a good mentor then you may lag in many things from others and that is why we at Compsoft Technologies gives you the facility of skilled employees so that you do not feel unsecured about the academics. Personality development and academic status are some of those things which lie on mentor's hands. If you are trained well then you can do well in your future and knowing its importance of Compsoft Technologies always tries to give you the best.

They have a great team of skilled mentors who are always ready to direct their trainees in the best possible way they can and to ensure the skills of mentors we held many skill development programs as well so that each and every mentor can develop their own skills with the demands of the companies so that they can prepare a complete packaged trainee.

Services provided by Compsoft Technologies.

- Core Java and Advanced Java
- Web services and development
- Dot Net Framework
- Python
- Selenium Testing
- Conference / Event Management Service
- Academic Project Guidance
- On The Job Training
- Software Training

CHAPTER 3

INTRODUCTION

In this project author is evaluating performance of various classification/prediction algorithms such as SVM, Naïve Bayes, and Logistic Regression etc to predict heart disease. All these algorithms are good in prediction but accuracy is not good enough. To get better prediction accuracy author is combining two classification algorithms such as Linear Model and Random Forest to build new algorithm called Hybrid Machine Learning to get better prediction accuracy of heart dataset. Hybrid algorithm will form up by using Voting classifier, Internally Voting classifier will build up using Linear Model and Random Forest and while classification voting algorithm will evaluate prediction accuracy of both algorithms and vote for that algorithm which gives better accuracy. So by using hybrid model always we will have better prediction accuracy algorithm which helps in better prediction of heart disease.

It is difficult to identify heart disease because of several contributory risk factors such as diabetes, high blood pressure, high cholesterol, abnormal pulse rate and many other factors. Various techniques in data mining and neural networks have been employed to find out the severity of heart disease among humans. The severity of the disease is classified based on various methods like K-Nearest Neighbor Algorithm (KNN), Decision Trees (DT), Genetic algorithm (GA), and Naive Bayes (NB). The nature of heart disease is complex and hence, the disease must be handled carefully. Not doing so may affect the heart or cause premature death. The perspective of medical science and data mining are used for discovering various sorts of metabolic syndromes. Data mining with classification plays a significant role in the prediction of heart disease and data investigation.

Introduction to ML

Machine learning is programming computers to optimize a performance criterion using example data or past experience. We have a model defined up to some parameters, and learning is the execution of a computer program to optimize the parameters of the model using the training data or past experience. The model may be predictive to make predictions in the future, or descriptive to gain knowledge from data.

The field of study known as machine learning is concerned with the question of how to construct computer programs that automatically improve with experience.

Machine learning is a subfield of artificial intelligence that involves the development of algorithms and statistical models that enable computers to improve their performance in tasks through experience. These algorithms and models are designed to learn from data and make predictions or decisions without explicit instructions. There are several types of machine learning, including supervised learning, unsupervised learning, and reinforcement learning. Supervised learning involves training a model on labeled data, while unsupervised learning involves training a model on unlabeled data. Reinforcement learning involves training a model through trial and error. Machine learning is used in a wide variety of applications, including image and speech recognition, natural language processing, and recommender systems.

Problem Statement

There is ample related work in the fields directly related to this paper. ANN has been introduced to produce the highest accuracy prediction in the medical field. The back propagation multilayer perception (MLP) of ANN is used to predict heart disease. The obtained results are compared with the results of existing models within the same domain and found to be improved

CHAPTER 4

SYSTEM ANALYSIS

Existing System

The traditional detection method mainly depends on the doctor's vision of treating the patient and his level of experience, which usually are delayed, inaccurate and not in-time. After following these methods, it may take time for diagnosing the records and giving the summary and then treating the patient.

Proposed System

In this paper author is evaluating performance of various classification/prediction algorithms such as SVM, Naïve Bayes, and Logistic Regression etc to predictheart disease. All this algorithms are good in prediction but accuracy is not good enough. To get better prediction accuracy author is combining two classification algorithms such as Linear Model and Random Forest to build new algorithm called Hybrid Machine Learning to get better prediction accuracy of heart dataset. Hybrid algorithm will form up by using Voting classifier, Internally Voting classifier will build up using Linear Model and Random Forest and while classification voting algorithm will evaluate prediction accuracy of both algorithms and vote for that algorithm which gives better accuracy. So by using hybrid model always we will have better prediction accuracy algorithm which helps in better prediction of heart disease.

Objective of the System

- Understand a wide variety of learning algorithms
- Understand how to evaluate models generated from data.
- Apply, the algorithms to real problems
- Optimize the models learned and report on the expectancy accuracy that can be achievedby applying the models.

CHAPTER 5

REQUIREMENT ANALYSIS

Hardware Requirement Specification

PROCESSOR : Intel i5

RAM : 4GB

HARD DISK : 16GB

Software Requirement Specification

OPERATING SYSTEM : Linux/Windows

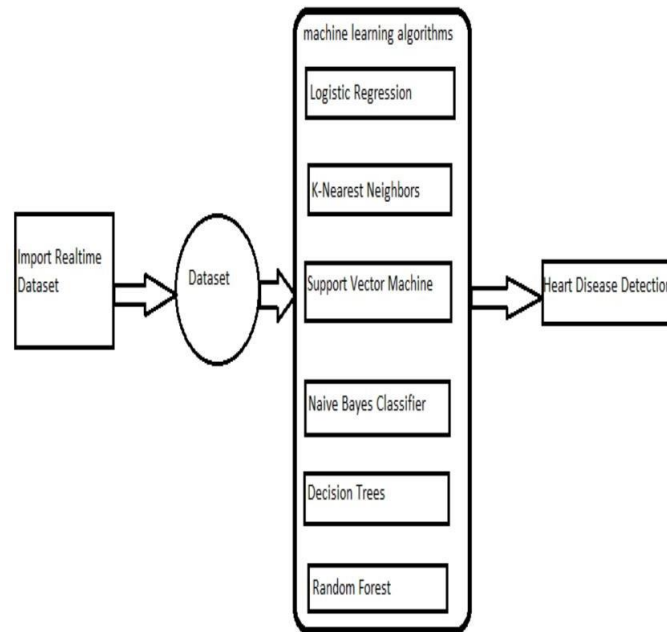
BACK-END : Python 3

OTHER BACKEND LIBRARIES : matplotlib, numpy, pandas, sklearn, Seaborn

CHAPTER 6

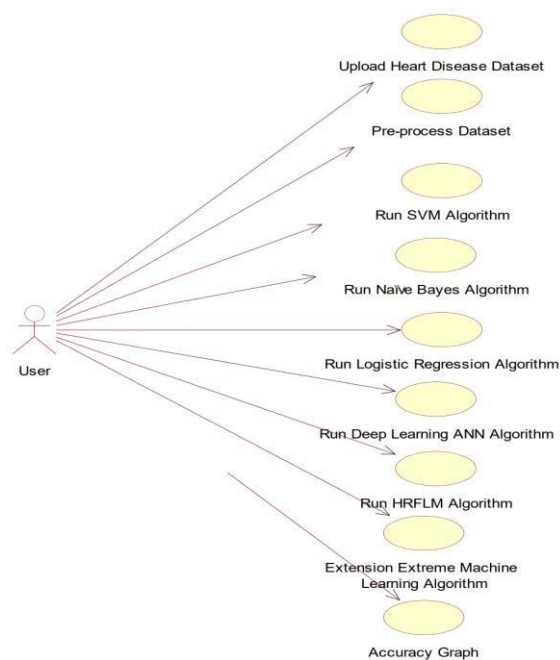
DESIGN ANALYSIS

Date Flow Diagram

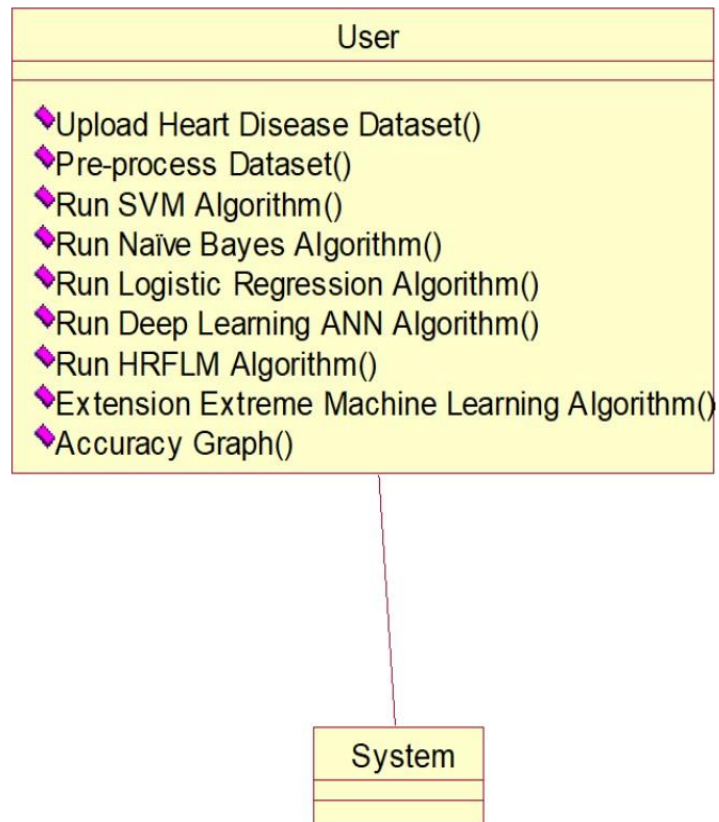


Date Flow Diagram

Use Case Diagram



Class Diagram



CHAPTER 7

IMPLEMENTATION

Implementation is the stage where the theoretical design is turned into a working system. The most crucial stage in achieving a new successful system and in giving confidence on the new system for the users that it will work efficiently and effectively.

The system can be implemented only after thorough testing is done and if it is found to work according to the specification. It involves careful planning, investigation of the current system and its constraints on implementation, design of methods to achieve the change over and an evaluation of change over methods as a part from planning.

Two major tasks of preparing the implementation are education and training of the users and testing of the system. The more complex the system being implemented, the more involved will be the system analysis and design effort required just for implementation.

The implementation phase comprises of several activities. The required hardware and software acquisition is carried out. The system may require some software to be developed. For this, programs are written and tested. The user then changes over to his new fully tested system and the old system is discontinued.

TESTING

The testing phase is an important part of software development. It is the Information zed system will help in automate process of finding errors and missing operations and also a complete verification to determine whether the objectives are met and the user requirements are satisfied. Software testing is carried out in three steps:

1. The first includes unit testing, where in each module is tested to provide its correctness, validity and also determine any missing operations and to verify whether the objectives have been met. Errors are noted down and corrected immediately.
2. Unit testing is the important and major part of the project. So errors are rectified easily in particular module and program clarity is increased. In this project entire system is divided into several modules and is developed individually. So unit testing is conducted to individual modules.
3. The second step includes Integration testing. It need not be the case, the software whose modules when run individually and showing perfect results, will also show perfect results when run as a whole.

CHAPTER 8

SNAPSHOTS

First ,import all the libraries/packages which are necessary to analyse the dataset

```
import numpy as np
import pandas as pd
import matplotlib as plt

import seaborn as sns
import matplotlib.pyplot as plt
```

Python libraries

Next ,we have to insert the dataset which is present in the local system using pandaslibrary

```
filePath = '/content/heart.csv'
data = pd.read_csv(filePath)
data.head(10)
```

	age	sex	cp	trestbps	chol	fbs	restecg	thalach	exang	oldpeak	slope	ca	thal	target
0	63	1	3	145	233	1	0	150	0	2.3	0	0	1	1
1	37	1	2	130	250	0	1	187	0	3.5	0	0	2	1
2	41	0	1	130	204	0	0	172	0	1.4	2	0	2	1
3	56	1	1	120	236	0	1	178	0	0.8	2	0	2	1
4	57	0	0	120	354	0	1	163	1	0.6	2	0	2	1
5	57	1	0	140	192	0	1	148	0	0.4	1	0	1	1
6	56	0	1	140	294	0	0	153	0	1.3	1	0	2	1
7	44	1	1	120	263	0	1	173	0	0.0	2	0	3	1
8	52	1	2	172	199	1	1	162	0	0.5	2	0	3	1
9	57	1	2	150	168	0	1	174	0	1.6	2	0	2	1

Heart Dataset

Next about rows and columns of the dataset

```
print("(Rows, columns): " + str(data.shape))
data.columns

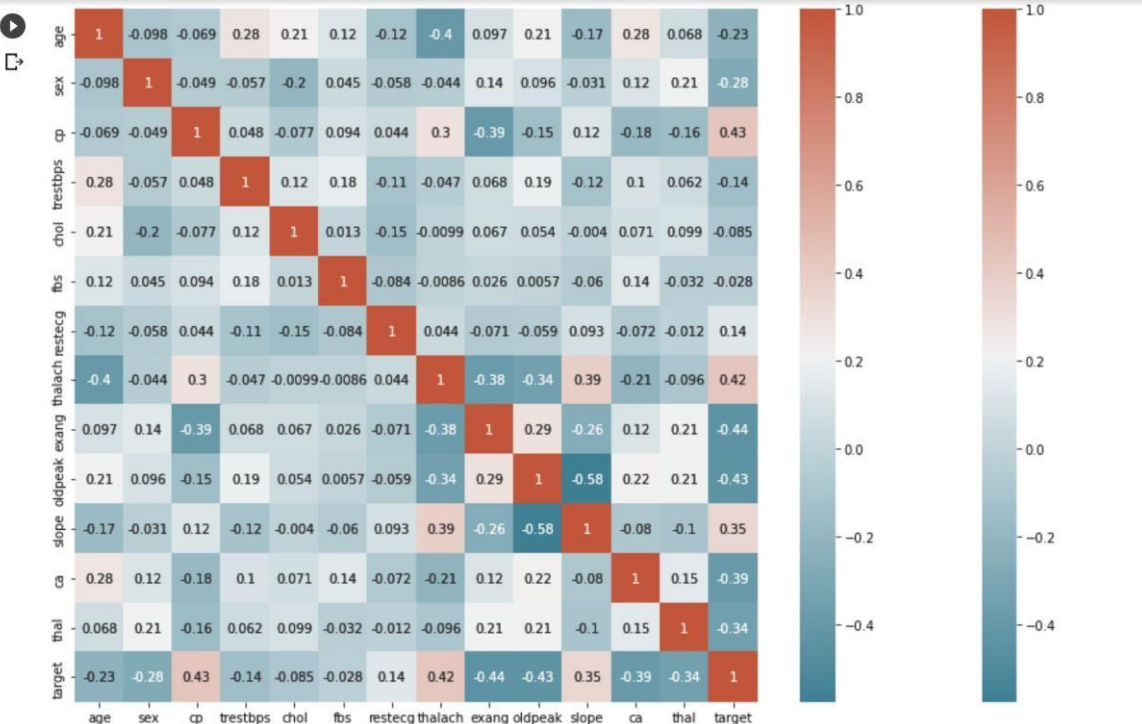
(In [1]: (Rows, columns): (303, 14)
Index(['age', 'sex', 'cp', 'trestbps', 'chol', 'fbs', 'restecg', 'thalach',
       'exang', 'oldpeak', 'slope', 'ca', 'thal', 'target'],
      dtype='object')
```

Rows and columns

Using Calculating Correlation Matrix

```
#used for calculating correlation matrix
corr = data.corr()
plt.subplots(figsize=(15,10))
sns.heatmap(corr, xticklabels=corr.columns, yticklabels=corr.columns, annot=True,
cmap=sns.diverging_palette(220, 20, as_cmap=True))
sns.heatmap(corr, xticklabels=corr.columns,
yticklabels=corr.columns,
annot=True,
cmap=sns.diverging_palette(220, 20, as_cmap=True))
```

<matplotlib.axes._subplots.AxesSubplot at 0x7f1e29fffa10>



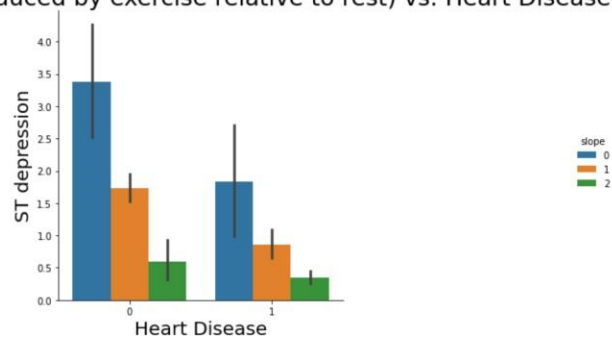
Correlation Matrix

Plotting different graphs

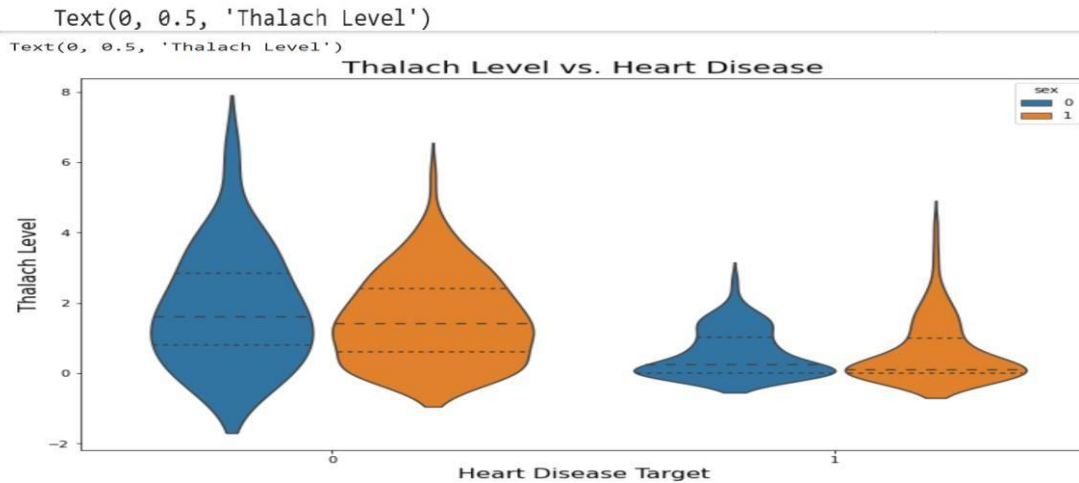
```
sns.catplot(x="target", y="oldpeak", hue="slope", kind="bar", data=data);
plt.title('ST depression (induced by exercise relative to rest) vs. Heart Disease',size=25)
plt.xlabel('Heart Disease',size=20)
plt.ylabel('ST depression',size=20)
```

Text(26.426458333333343, 0.5, 'ST depression')

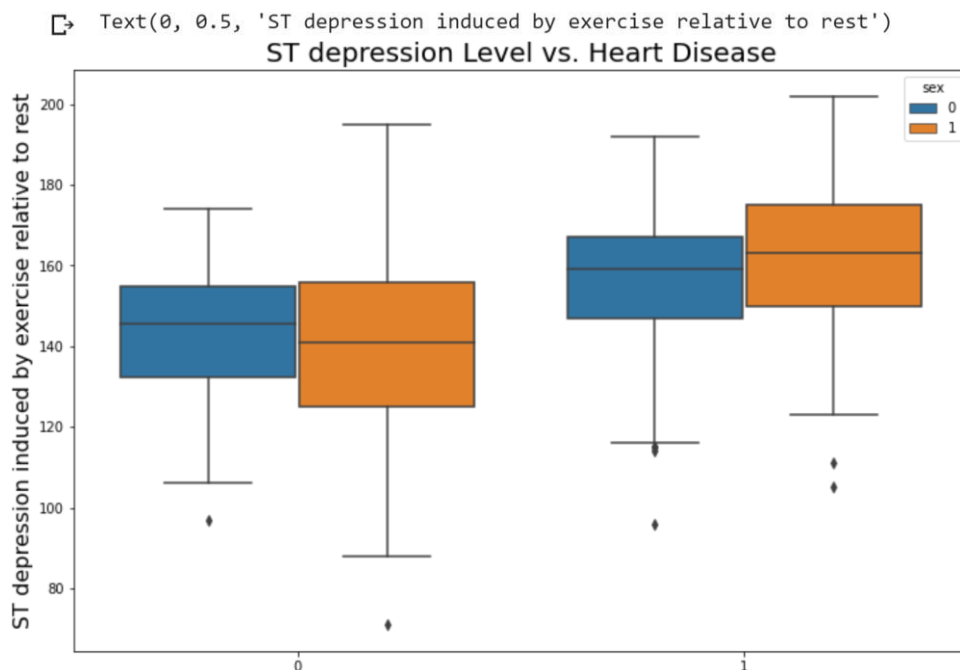
ST depression (induced by exercise relative to rest) vs. Heart Disease



```
[ ] plt.figure(figsize=(12,8))
sns.violinplot(x= 'target', y= 'oldpeak',hue="sex", inner='quartile',data= data )
plt.title("Thalach Level vs. Heart Disease",fontsize=20)
plt.xlabel("Heart Disease Target", fontsize=16)
plt.ylabel("Thalach Level", fontsize=16)
```

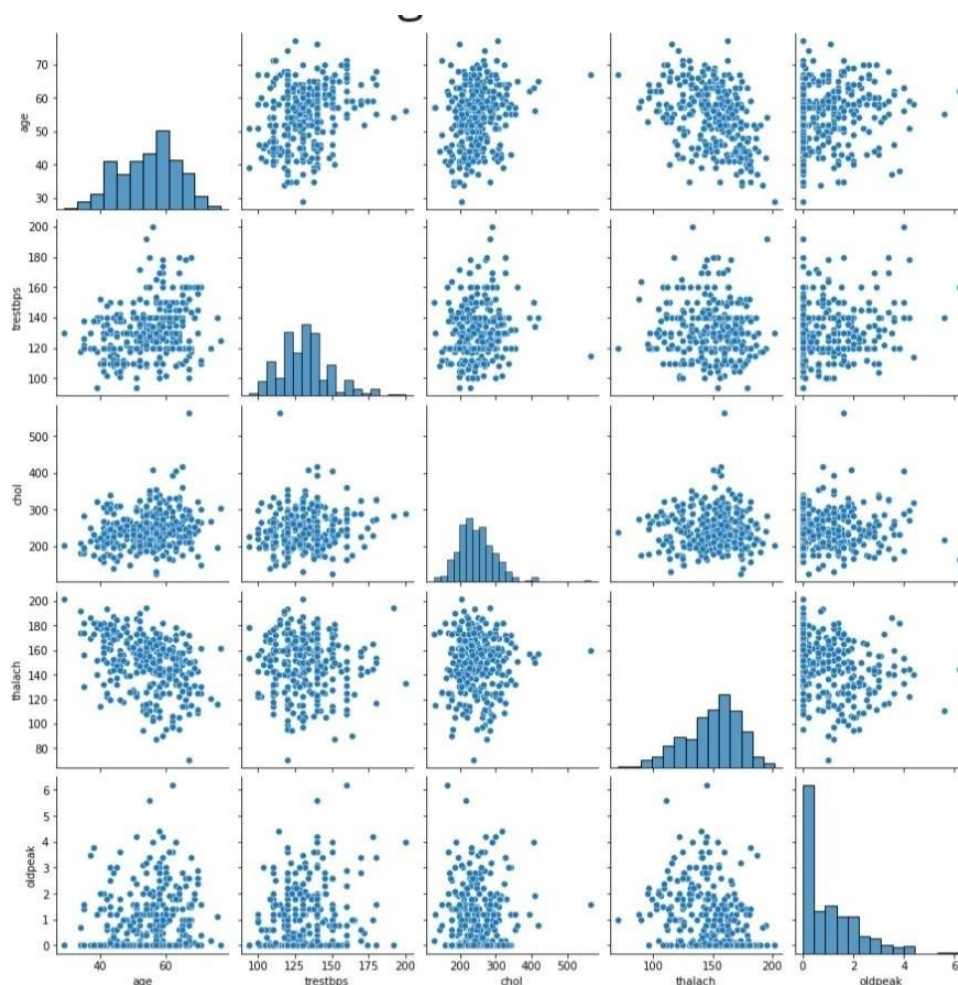


```
▶ plt.figure(figsize=(12,8))
sns.boxplot(x= 'target', y= 'thalach',hue="sex", data=data )
plt.title("ST depression Level vs. Heart Disease", fontsize=20)
plt.xlabel("Heart Disease Target",fontsize=16)
plt.ylabel("ST depression induced by exercise relative to rest", fontsize=16)
```



```
▶ subData = data[['age', 'trestbps', 'chol', 'thalach', 'oldpeak']]
sns.pairplot(subData)
```

<seaborn.axisgrid.PairGrid at 0x7f1e26649610>



Graphs of dataset

```
[ ] from sklearn.model_selection import train_test_split
    x_train, x_test, y_train, y_test = train_test_split(X,y,test_size = 0.2, random_state = 1)

[ ] from sklearn.preprocessing import StandardScaler
    sc = StandardScaler()
    x_train = sc.fit_transform(x_train)
    x_test = sc.transform(x_test)
```

Training and testing the model using Logistic Regression

```
from sklearn.metrics import classification_report
from sklearn.linear_model import LogisticRegression
modell = LogisticRegression(random_state=1)
modell.fit(x_train, y_train)
y_pred1 = modell.predict(x_test)
print(classification_report(y_test, y_pred1))
```

	precision	recall	f1-score	support
0	0.77	0.67	0.71	30
1	0.71	0.81	0.76	31
accuracy			0.74	61
macro avg	0.74	0.74	0.74	61
weighted avg	0.74	0.74	0.74	61

Logistic Regression

Training and testing the model using KNeighbors Classifier

```
] from sklearn.metrics import classification_report
from sklearn.neighbors import KNeighborsClassifier
model2 = KNeighborsClassifier()
model2.fit(x_train, y_train)
y_pred2 = model2.predict(x_test)
print(classification_report(y_test, y_pred2))
```

	precision	recall	f1-score	support
0	0.78	0.70	0.74	30
1	0.74	0.81	0.77	31
accuracy			0.75	61
macro avg	0.76	0.75	0.75	61
weighted avg	0.76	0.75	0.75	61

KNeighbors

ClassifierTraining and testing the model using SVC

```
▶ from sklearn.metrics import classification_report
from sklearn.svm import SVC
model3 = SVC(random_state=1)
model3.fit(x_train, y_train)
y_pred3 = model3.predict(x_test)
print(classification_report(y_test, y_pred3))
```

↗

	precision	recall	f1-score	support
0	0.80	0.67	0.73	30
1	0.72	0.84	0.78	31
accuracy			0.75	61
macro avg	0.76	0.75	0.75	61
weighted avg	0.76	0.75	0.75	61

SVC

Training and testing the model using Gaussian NB

```
▶ from sklearn.metrics import classification_report
from sklearn.naive_bayes import GaussianNB
model4 = GaussianNB()
model4.fit(x_train, y_train)
y_pred4 = model4.predict(x_test)
print(classification_report(y_test, y_pred4))
```

↗

	precision	recall	f1-score	support
0	0.79	0.73	0.76	30
1	0.76	0.81	0.78	31
accuracy			0.77	61
macro avg	0.77	0.77	0.77	61
weighted avg	0.77	0.77	0.77	61

Gaussian NB

Training and testing the model using Decision Tree

```

] from sklearn.metrics import classification_report
  from sklearn.tree import DecisionTreeClassifier
  model5 = DecisionTreeClassifier(random_state=1)
  model5.fit(x_train, y_train)
  y_pred5 = model5.predict(x_test)
  print(classification_report(y_test, y_pred5))

```

	precision	recall	f1-score	support
0	0.68	0.70	0.69	30
1	0.70	0.68	0.69	31
accuracy			0.69	61
macro avg	0.69	0.69	0.69	61
weighted avg	0.69	0.69	0.69	61

Decision Tree

Training and testing the model using Random Forest

```

from sklearn.metrics import classification_report
from sklearn.ensemble import RandomForestClassifier
model6 = RandomForestClassifier(random_state=1)
model6.fit(x_train, y_train)
y_pred6 = model6.predict(x_test)
print(classification_report(y_test, y_pred6))

```

	precision	recall	f1-score	support
0	0.88	0.70	0.78	30
1	0.76	0.90	0.82	31
accuracy			0.80	61
macro avg	0.82	0.80	0.80	61
weighted avg	0.81	0.80	0.80	61

Random Forest

CHAPTER 9

CONCLUTION

At the end of this project, we have acquired the result of an accurate value of using a random forest algorithm with new enhancements. In comparison to existing modules, this proposed module is applicable for the dataset and provides more accurate results. The Random forest algorithm will provide better performance with many training data, but speed during testing and application will still suffer. Usage of more pre-processing techniques would also assist. In this project, we have seen that the accuracy of Random Forest Algorithm is best when compared to other algorithms.

The package was designed in such a way that future modifications can be done easily. The following conclusions can be deduced from the development of the project:

- ❖ Automation of the entire system improves the efficiency
- ❖ It provides a friendly graphical user interface which proves to be better when compared to the existing system.
- ❖ It gives appropriate access to the authorized users depending on their permissions.
- ❖ It effectively overcomes the delay in communications.
- ❖ Updating of information becomes so easier
- ❖ System security, data security and reliability are the striking features.
- ❖ The System has adequate scope for modification in future if it is necessary.

10.REFERENCE

Machine Learning with Python: Design and Develop Machine Learning and DeepLearningTechnique using real world code examples, Abhishek Vijayvargia, 1st Edition, 2019. Python GUI Programming - A Complete Reference Guide: Develop responsive and powerfulGUI applications with PyQt and Tkinter, Alan D. Moore, B. M. Harwani, 2019.

Machine Learning for Beginners: The Definitive Guide to Neural Networks, Random Forests, and Decision Trees, Jennifer Grange, 2017.

A. S. Abdullah and R. R. Rajalaxmi, “A data mining model for predicting the coronary heart disease using random forest classifier,” in Proc. Int. Conf. Recent Trends Comput. Methods, Commun. Controls, Apr. 2012, pp. 22–25.

A. H. Alkeshuosh, M. Z. Moghadam, I. Al Mansoori, and M. Abdar, “Using PSO algorithm for producing best rules in diagnosis of heart disease,” in Proc. Int. Conf. Comput. Appl. (ICCA), Sep. 2017, pp. 306–311.

<https://www.kaggle.com/datasets?fileType=csv>