STATISTICAL MACHINE LEARNING APPROACH TO LIVER DISEASE PREDICTION

Date	18 November 2022
Team ID	PNT2022TMID01272
Project Name	Project - Statistical Machine Learning Approaches to Liver Disease Prediction
Team Members	Team Lead – Deepthi Shree S Member 1 – Keerthika.L Member 2 – Harita.K.S Member 3 – Deepthi Shree S Member 4 – Deepa.S

PROJECT REPORT

INTRODUCTION

OVERVIEW OF THE PROJECT:

To detect disease, healthcare professionals need to collect samples from patients which can cost both time and money. Often, more than one kind of test or many samples are needed from the patient to accumulate all the necessary information for a better diagnosis. The most routine tests are urinalysis, complete blood count (CBC), and comprehensive metabolic panel (CMP). These tests are generally less expensive and can still be very informative.

The liver has many functions such as glucose synthesis and storage, detoxification,production of digestive enzymes, erythrocyte regulation, protein synthesis, and various other features of metabolism. Chronic liver diseases include chronic hepatitis, fibrosis, and cirrhosis. Hepatitis

can occur from viral infection (e.g., hepatitis c virus) or auto-immune origin. Inflammation from hepatitis infection can cause tissue damage and scarring to occur in the liver. Moderate scarring is classified as fibrosis, while severe liver damage/scarringis classified as cirrhosis. Fibrosis and cirrhosis can also occur from alcoholism and non-alcoholic fatty liver disease.

Common Liver Disorder :

·Fatty liveris a revocable condition where large vacuoles of triglyceride fat acquire in liver cells via the process of limit. It can occur in people with a high level of alcohol consumption as well as in people who never had alcohol.
·Hepatitis(usually caused by a virus spread by excess contamination or direct contact with infected body fluids).
·Cirrhosis of the liver is one of the most serious liver diseases. It is an action used to indicate all forms of diseases of the liver characterized by the significant loss of cells. The liver gradually contracts in size and becomes leathery and hard. The regenerative action continues under liver cirrhosis but the progressive loss of liver cells exceeds cell replacement.
·Liver cancer. The risk of liver cancer is higher in those who have cirrhosis or who had valid types of viral hepatitis; but more often, the liver is the site of secondary (metastatic) cancers spread from other organs.

PURPOSE:

The main purpose of our project is that to predict the liver disease before it gets to serious state of health problem and gets very hard to cure. Knowledge of the level of liver damage in a patient with liver disease (particularly Hepatitis B and Hepatitis C) is a critical factor in determining the optimal course of treatment and to measure the effectiveness of alternative treatments in patients. The effort here expands on earlier work [4] by adding three additional artificial intelligence techniques to predict the degree of liver damage from blood serum results rather than determination from an invasive biopsy. Clinical Decision Support System (CDSS) provide

cost/effective solutions by correlating historical data to assist clinicians in treatment of disease. In this case we are developing a CDSS focused on predicting Fibrosis Stage from blood serum information. It is an information system which uses expert systems and artificial intelligence (AI) technology to support clinical decision. It makes integrated diagnostic and medical advice bases on the collected patients' information, providing reference for the clinical medical physician [6]. Clinical Decision Support Systems are "active knowledge systems which use two or more items of patient data to generate case%specific advice".

LITERATURE SURVEY

1) Prognosis of Liver Disease: Using Machine Learning Algorithms

Author:C.Priya (Assistant Professor) Department of CSE, PERI Engineering College

Data mining classification techniques like Decision Tree, Linear Discriminant, SVM Fine Gaussian and Logistic Regression algorithms are applied. Laboratory parameters of the patients are used as the dataset. Data contains features that can establish a rigorous model using Classification technique. MATLAB2016 is used in this paper for implementing classification algorithm on the dataset. Linear Discriminant algorithm showed the highest prediction accuracy 95.8% and ROC is 0.93. This paper helps in foresee the presence of liver disease using different classification algorithms. The dataset considered consists of lab reports of 574 patients who are advised with LFT. The initial phase includes assorting the information from the available source. This information is of the patients who are predicted of having liver disorder. The dataset is analyzed in the next phase to check which algorithm can be applied to obtain optimal results. The last phase includes classification of data by application of proper algorithm on the dataset and formatting a training model. In the last phase data is tested against the trained model to get proper predictive values. The proposed paper makes use of laboratory parameters of individuals after taking LFT. The model that yields the maximum efficiency is considered. Once the patient is predicted with the liver

disease, he is further advised for imaging tests to determine the existence of tumor lesions and their stage. This paper makes use of the lab test reports of the patients who has undergone Liver Function Test. MATLAB2016 is used but Logistic Regression gave high accuracy of 95.8%. Various predictors are tested by plotting graph that determined the existence of disorder in liver. Further research is proposed for considering the tumor characteristics of the patient once he is diagnosed with liver disorder. Also large dataset can be considered for training the model and algorithms can be determined.

2.Evaluation based Approaches for Liver Disease Prediction using Machine Learning Algorithms

Author: 1.C.Geetha (Assistant Professor) Department of CSE, Bharath Institute of Higher Education and Research Chennai, 2. India Geetha, Department of CSE Dr. MGR University.

This research also aims to compare the classification algorithms and to provide prediction accuracy results. In order to store large-scale information on patient outcomes, procedures, etc. Electronic health records (EHR) are used. The data on the EHR can be organized or unstructured. Electronic health records are stored in a standardised data format using managed language to log patient knowledge as a written text that is hyperlinked in existence. The EHR aims to streamline knowledge about the clinical workflow. Ensemble learning is a well-known method used for prediction by integrating multiple ensemble models of machine learning. Aggregations of various classifiers are J48, C4.5 and Naive Bayes. Ensembles search for better outcomes than all of the simple classifiers. The proposed work aims to enhance the predictive and classification quality of healthcare data by developing a hybrid predictive classifier model using the classifier ensemble. The SVM, Logistic Regression, comprises two main machine learning techniques. Using all the models, the prediction analysis has been implemented and their performance has been assessed. The probability of liver disease prediction attained with an accuracy of 96%. In future, the present scenario can be compared with other techniques such as naive baye's classification, Random forest etc. Also this work can be further focused on implementation of parametric classifications by bio-inspired optimization algorithms.

3.Liver Disease Prediction System using Machine Learning Techniques

AUTHOR: 1.Rakshith.D B, 2.Mrigank Srivastava 3.Ashwani Kumar 4.Gururaj.S P,From the Department of Computer Science and Engineering, Siddaganga Institute of Technology,Tumkur,India.

In this paper we are going discuss how to predict risk of liver disease for a person, based on the blood test report results of the user. In this paper, the risk of liver disease was predicted using various machine learning algorithms. The final output was predicted based on the most accurate machine learning algorithm. Based on the accurate model we designed a system which asks a person to enter the details of his/her blood test report. Then the system uses the most accurate model which is trained to predict, whether a person has risk of liver disease or not. In this project we have taken UCI ILPD Dataset which contains 10 variables that are age, gender, total Bilirubin, direct Bilirubin, total proteins, albumin, A/G ratio, SGPT, SGOT and Alkphos and contains 415 as liver disease patients and 167 as non-liver disease patients. As we got through the next parts of this paper we will explain what proccess as taken place for the selection of best model and building necccessary sytem for the prediction of liver disease. The major outcomes that can be expected through this project is that it increased convenience for predicting a liver disease. So, in such a scenario, our project will be extremely helpful to the society. With the dataset that we used for this project, we got 100 % accuracy for SVM model, and though it might be difficult to get such accuracies with very large datasets, from this projects results, one can clearly conclude that we can predict the risk of liver diseases with accuracy of 90 % or more.

4.LIVER DISEASE DIAGNOSIS USING MACHINE LEARNING

AUTHOR: 1.Prof. Sayalee Deshmukh , 2.Anushka Sawant, 3.Manasi Khopade,4.Pratiksha Kawale, 5.Yashika Palan, Department of Computer Engineering, Bharati Vidyapeeth’s College of Engineering for Women, Pune-411043.

The major goal of this study is to employ classification algorithms to distinguish between liver patients and healthy people. Chemical components (bilirubin, albumin, proteins, alkaline phosphatase) present in the human body, as well as tests such as SGOT and SGPT, determine whether a person is a patient, or whether they need to be diagnosed. Excessive alcohol consumption, inhalation of toxic gases, eating of contaminated food, pickles, and medicines have all contributed to an increase in patients with liver disease. The goal of this research is to analyse prediction algorithms in order to relieve doctors of their workload. Keywords— liver disease, SVM, Random Forest, KNN, ML, python, etc. The major goal of this study is to employ classification algorithms to distinguish between liver patients and healthy people. In this study, the performance of FIVE classification techniques was compared using data from liver patients: Logistic Regression, Support Vector Machines (SVM), K Nearest Neighbour (KNN), Decision Tree and Random Forest (RF). Furthermore, the most accurate model is implemented as a user-friendly Graphical User Interface (GUI) in Python using tkinter package. Doctors and medical practitioners can easily use the GUI as a screening tool for liver disease. The dataset used in this work is The Indian Liver Patient Dataset (ILPD), which was chosen from the UCI Machine Learning repository. It is a representative sampling of the entire Indian population.

5) A Survey on machine learning techniques for the diagnosis of liver disease

AUTHOR:1.Golmei Shaheamlung , 2.Harshpreet , 3.Mandeep.

The motive of this paper is to give a survey and comparative analysis of the entire machine learning techniques for diagnosis and prediction of liver disease in the medical area, which has already been used for the prediction of liver disease by various authors and the analysis are based on Accuracy, Sensitivity, Precision, and Specificity. With the help of this survey and study, it has clearly found and observed that some machine learning algorithm such as Decision tree, J48 and ANN provide better accuracy on detection and prediction of liver diseases. With this in mind and performance we know that different algorithm has different performance based on different scenario but most importantly, the dataset and feature selection is also very important to

get better prediction results. With this survey we found out that the accuracy and performance can be improved by using different combination or hybrid machine learning algorithm and in future we can also work on more parameter's which help to get better performance than the existing technique.

6) Prediction of Liver Disease using Classification Algorithms
AUTHOR:1.Thirunavukkarasu.K, 2.Ajay S. Singh, 3.Md Irfan, 4.Abhishek Chowdhury.

The main aim is to predict liver disease using different classification algorithms. The algorithms used for this purpose of work is Logistic Regression, K-Nearest Neighbour and Support Vector Machines. Accuracy score and confusion matrix is used to compare this classification algorithm. In this paper, Logistic Regression, K-Nearest neighbour and Support Vector Machines are been used for prediction of liver disease. The proposed methods are used to compare classification accuracy of Logistic Regression, K-nearest neighbour and Support Vector Machine. The comparison of all these algorithms been done are based on classification accuracy which is found through confusion matrix. From the experiment, Logistic Regression and K-Nearest Neighbour have the highest accuracy but logistic regression have the highest sensitivity. Therefore, it can be concluded that Logistic Regression is appropriate for predicting liver disease.

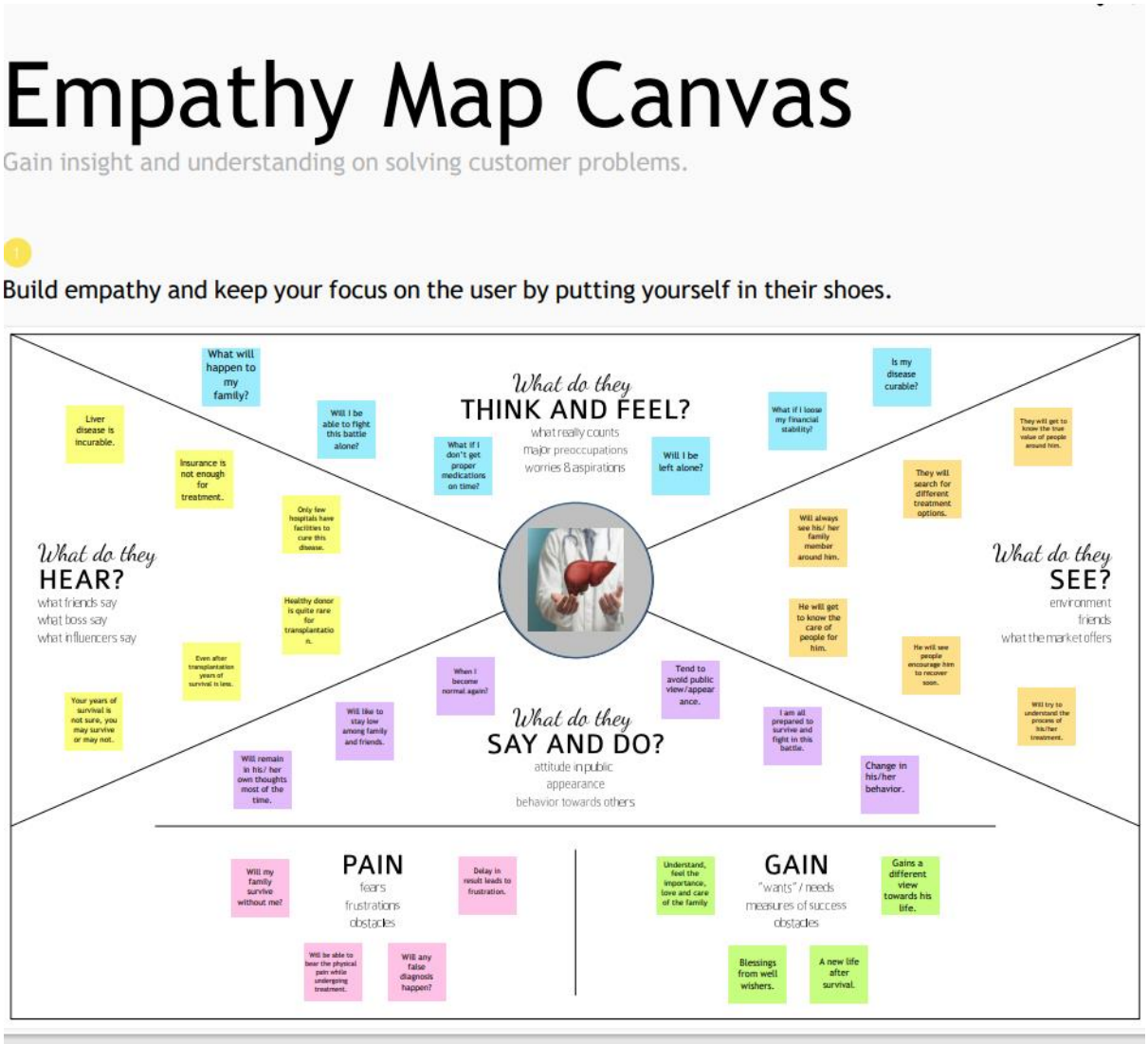
7) Optimizing Liver disease prediction with Random Forest by various Data balancing Techniques
AUTHOR:1.Sateesh Ambesange, 2. Vijayalaxmi A, 3.Rashmi Uppin, 4. Shruthi Patil, 5. Vilaskumar PatiL.

Random Forest (RF) algorithm is used to predict the disease with different preprocessing techniques. Data set is checked for skewness, outliers and imbalance using univariate and bivariate analysis and then suitable algorithms used to remove outliers and various oversampling and under sampling techniques are used to balance such data's. In this work, Indian Liver Patient Dataset hosted at "ics.uci.edu" is used. Instead of selecting the algorithm, which gives better performance, the paper approaches how to tune the ML module for Random Forest

algorithm in step-by-step ways. The main focus of the paper is to deeply analyze how models can be further tuned beyond one point of saturation due to an imbalanced data set. In this work, ML models are built using various preprocessing techniques to balance the unbalanced data and predicted using RF algorithm.

IDEATION PHASE:

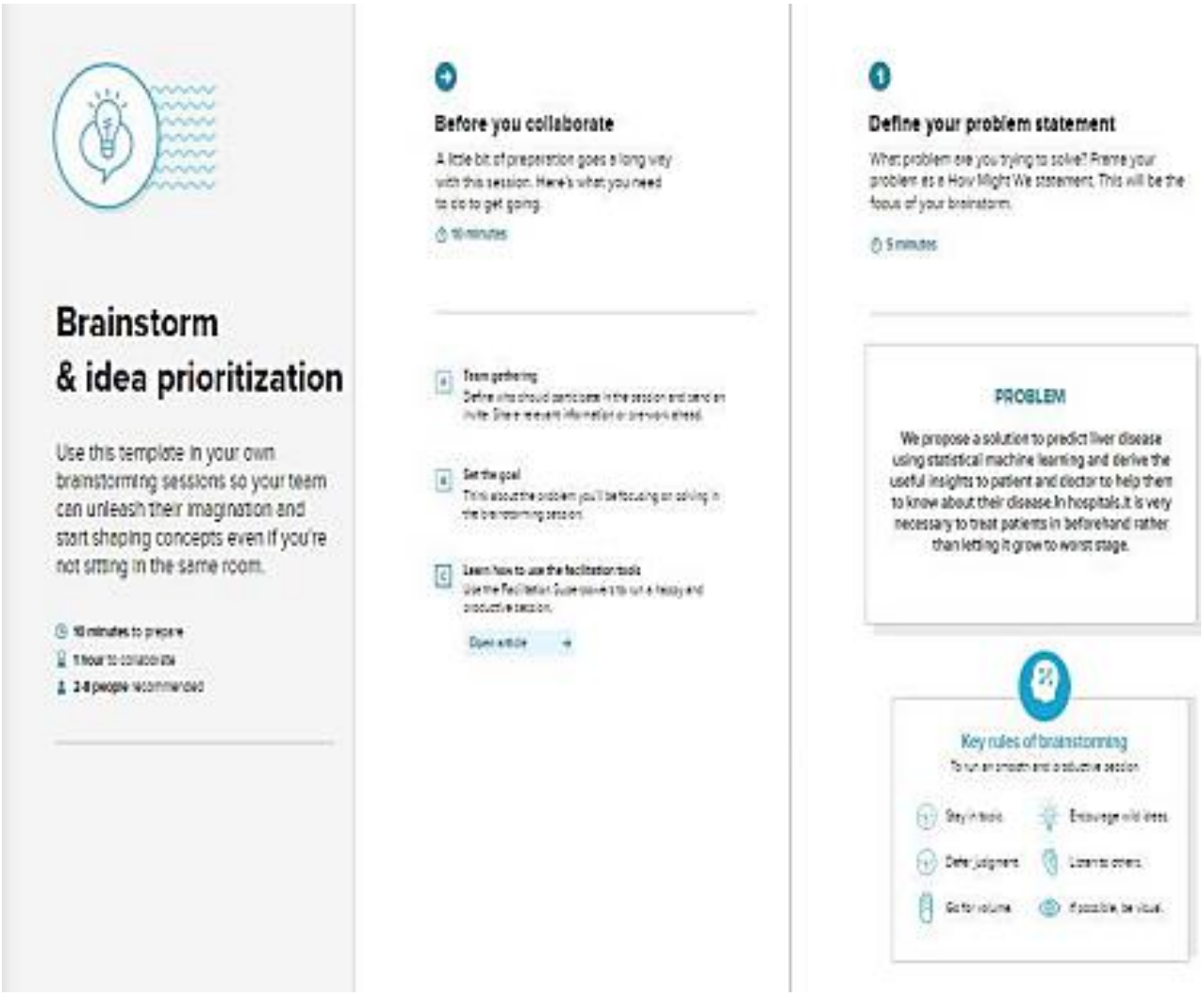
EMPATHY MAP CANVAS:



IDEATION & BRAINSTORMING:

Date	19 September 2022
Team ID	PNT2022TMID01272
Project Name	Statistical Machine Learning Approach to Liver Disease Prediction
Maximum Marks	4 Marks

Step-1: Team Gathering, Collaboration and Select the Problem State



Step-2: Brainstorm, Idea Listing and Grouping

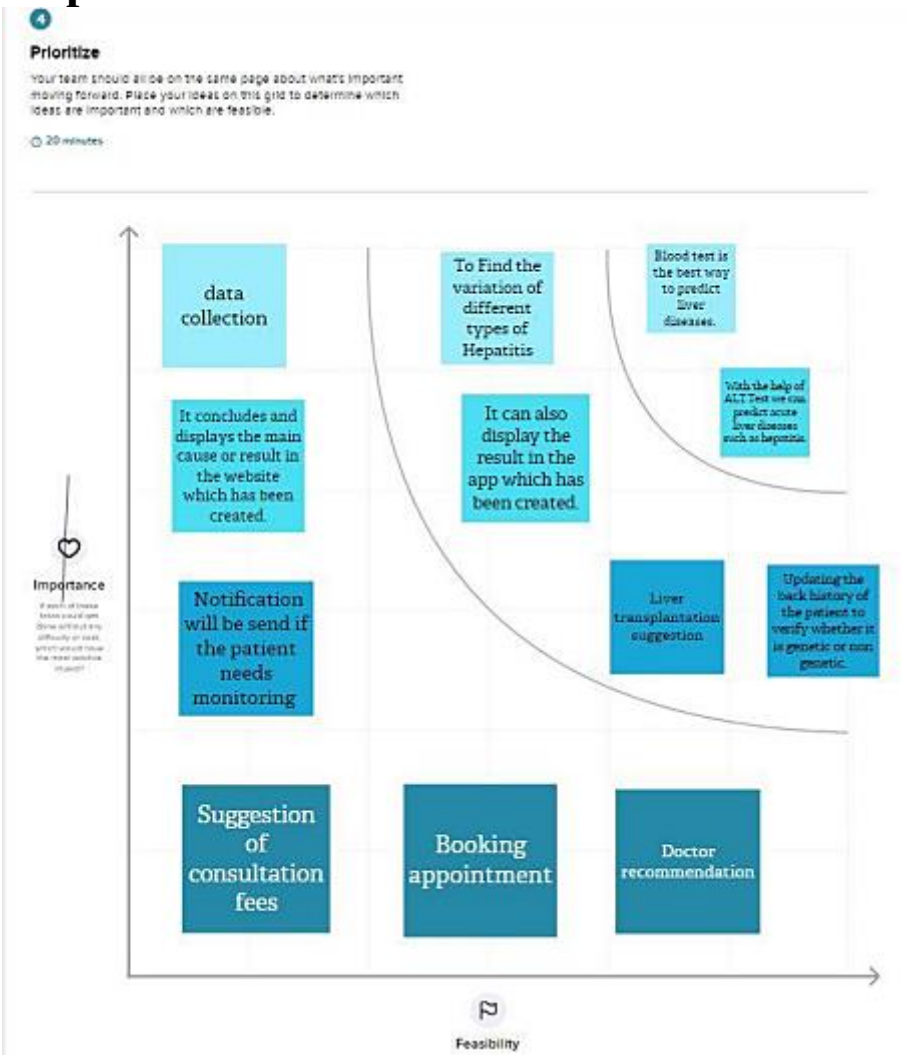
3

Group ideas

Take turns sharing your ideas while clustering similar or related notes as you go. Once all sticky notes have been grouped, give each cluster a sentence-like label. If a cluster is bigger than six sticky notes, try and see if you and break it up into smaller sub-groups.

20 minutes

Step-3: Idea Prioritization



PROPOSED SOLUTION:

Date	13 October 2022
TeamID	PNT2022TMID01272
Project Name	Statistical Machine Learning Approaches to Liver Disease Prediction
MaximumMarks	2Marks

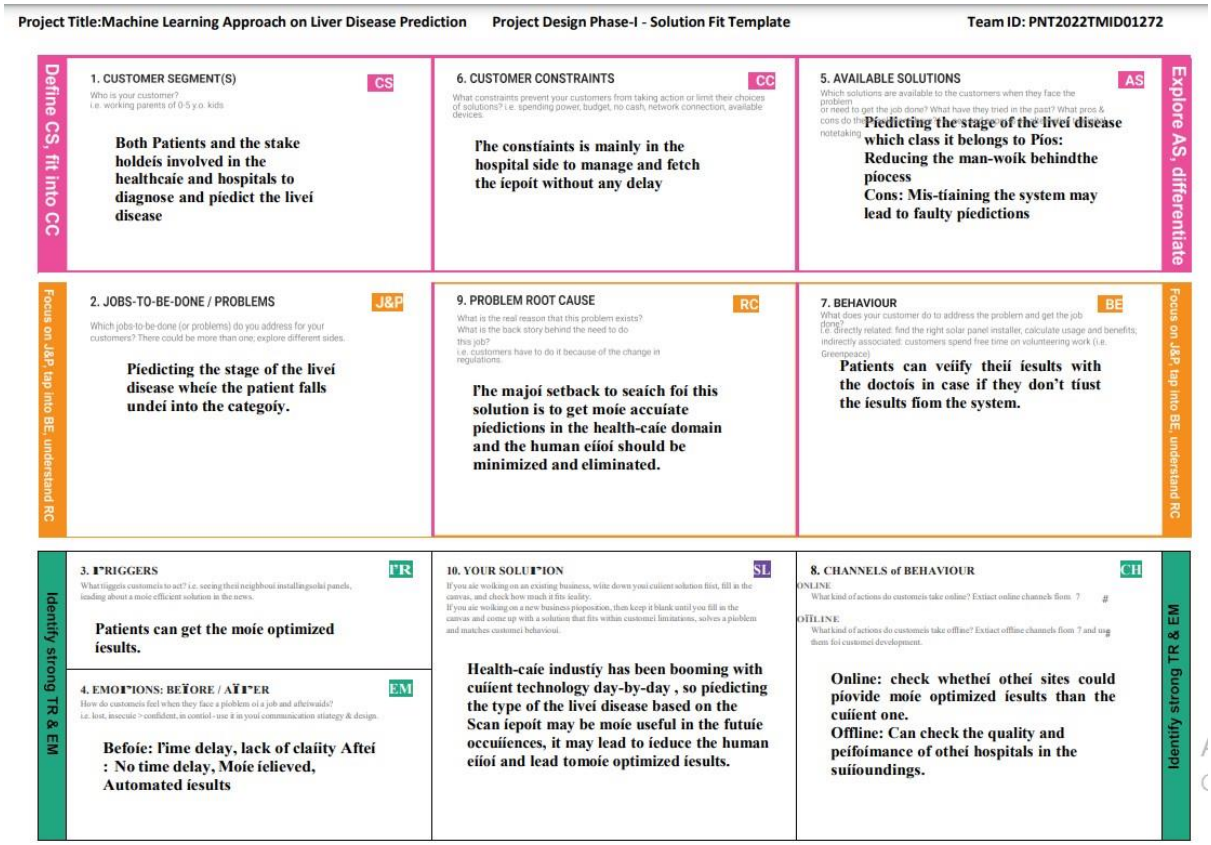
ProposedSolution:

S.No.	Parameter	Description

1.	ProblemStatement(Problemtobes o lved)	We propose a solution to predict liver disease using statistical machine learning and derive the useful insights topatient and doctor to help them in a know about their disease. In hospitals, it is very necessary totreat patients in before hand rather than letting it grow to worse stage.
2.	Idea/Solutiondescription	We propose a solution to build a simple webapplication which takes input as patient dataand returns us output with the prediction of liver disease affected by the patient.The results will be displayed to the end user in a web page.
3.	Novelty/Uniqueness	The innovative and additional perk to make thissolution stronger and the results more reliable,we use machine learning algorithms to developa predictive analysis model which will be usedto make predictions either on the patient's liver disease.Predictionfortheseresultswillbeshowni n theuserfriendly-manner.
4.	SocialImpact/CustomerSatisfaction	The solution can never go unnoticed, though itisnewtothesociety,becauseitis inaproactive way of prediction. It willaddresstheconcernofthekeystakeholders,so it will create the impact in the patient as wellasthesocial side.
5.	BusinessModel(RevenueModel)	The take-away of this project in a businessscope of manner is mean to be plenty, it can bebeneficial for the users (Patients and Doctors)more intriguing way. It is in need for thecommunity of people, where it comes to handyin day-to-day life. It is a part of the life savinganalysisandinsights.

6.	ScalabilityoftheSolution	Scalability is the measure of the systemperformance against the increase or decreasein user demand. The system can handle theuserrequestandreturntheresultsontime.It does not require much of the Graphicalprocessor unit; it can be even run on thesystem of both doctor and patient.

PROBLEM SOLUTION FIT:



REQUIREMENT ANALYSIS:

FUNCTIONAL REQUIREMENT:

Project Design Phase-II
Solution Requirements (Functional & Non-functional)

Date	20 October 2022
Team ID	PNT2022TMID01272
Project Name	Project - Statistical Machine Learning Approaches to Liver Disease Prediction
Maximum Marks	4 Marks

Following are the functional requirements of the proposed solution.

FR No.	Functional Requirement (Epic)	Sub Requirement (Story / Sub-Task)
FR-1	User Registration	Registration through Website Registration through Application

FR-2	User Input	The dataset should be uploaded(blood test report).
FR-3	Building the system	Dataset will be splitted into training and test dataset and then the model will be trained using training dataset.
FR-4	Prediction model	After getting the dataset from the user the pattern of the blood content will be learned using all the algorithms and then it will predict whether the person is affected by liver disease or not.
FR-5	Algorithm of prediction of disease	Machine learning
FR-6	Output	Result will be shown at the last whether the person is affected or not.

NON-FUNCTIONAL REQUIREMENT:

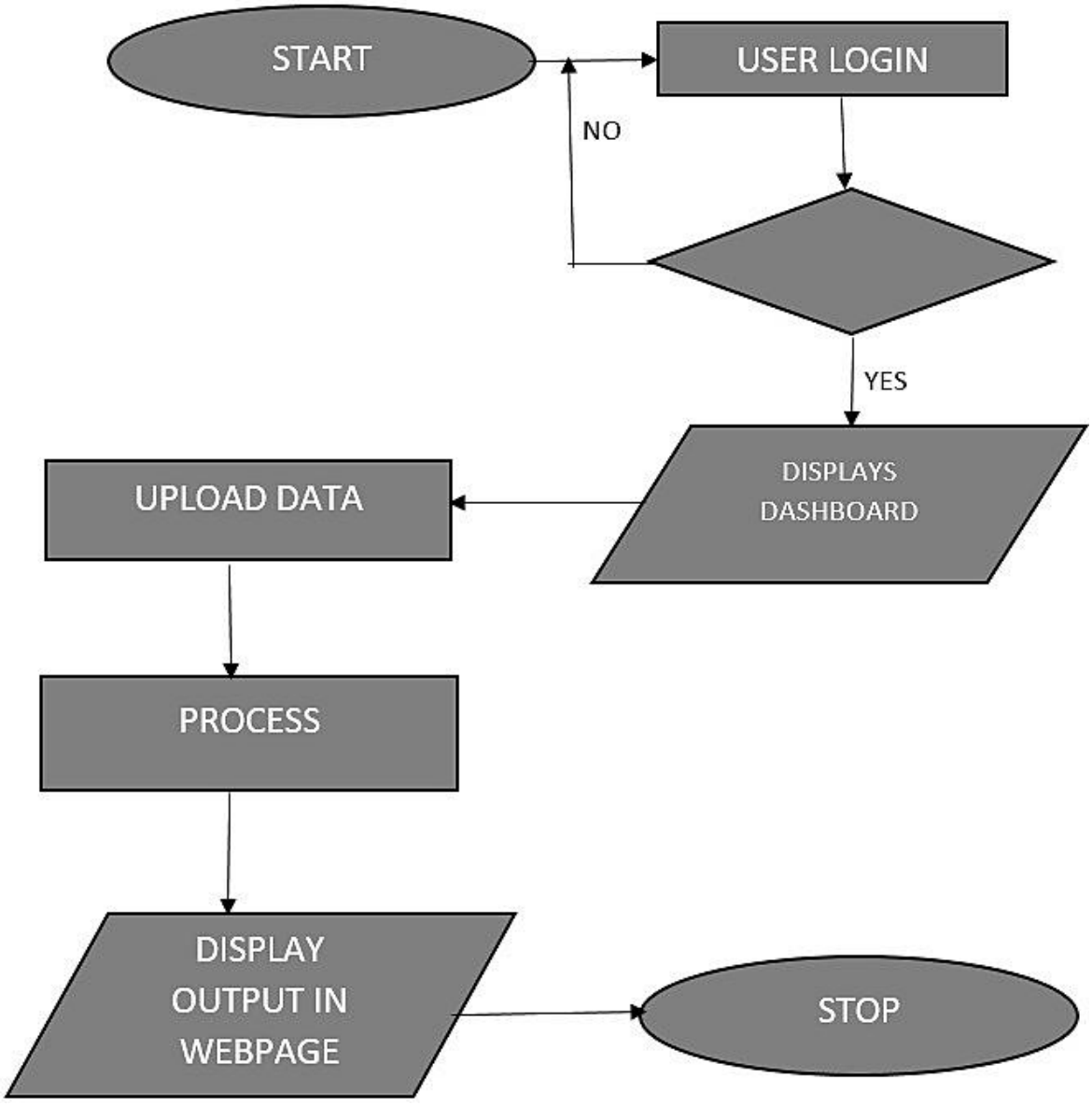
Following are the non-functional requirements of the proposed solution.

FR No.	Non-Functional Requirement	Description
NFR-1	**Usability**	It is very simple and easy way of predicting liver disease in an early stage.
NFR-2	**Security**	Early prediction of disease allows patients to take treatment in early stage and it will save many lives.
NFR-3	**Reliability**	This method will offer better performance and make it more dependable.
NFR-4	**Performance**	This provides more than 90% accuracy. Thus, it has high performance rate.
NFR-5	**Availability**	By having the dataset of the patient such as blood reports disease can be predicted.
NFR-6	**Scalability**	It has more efficiency in detecting liver disease.

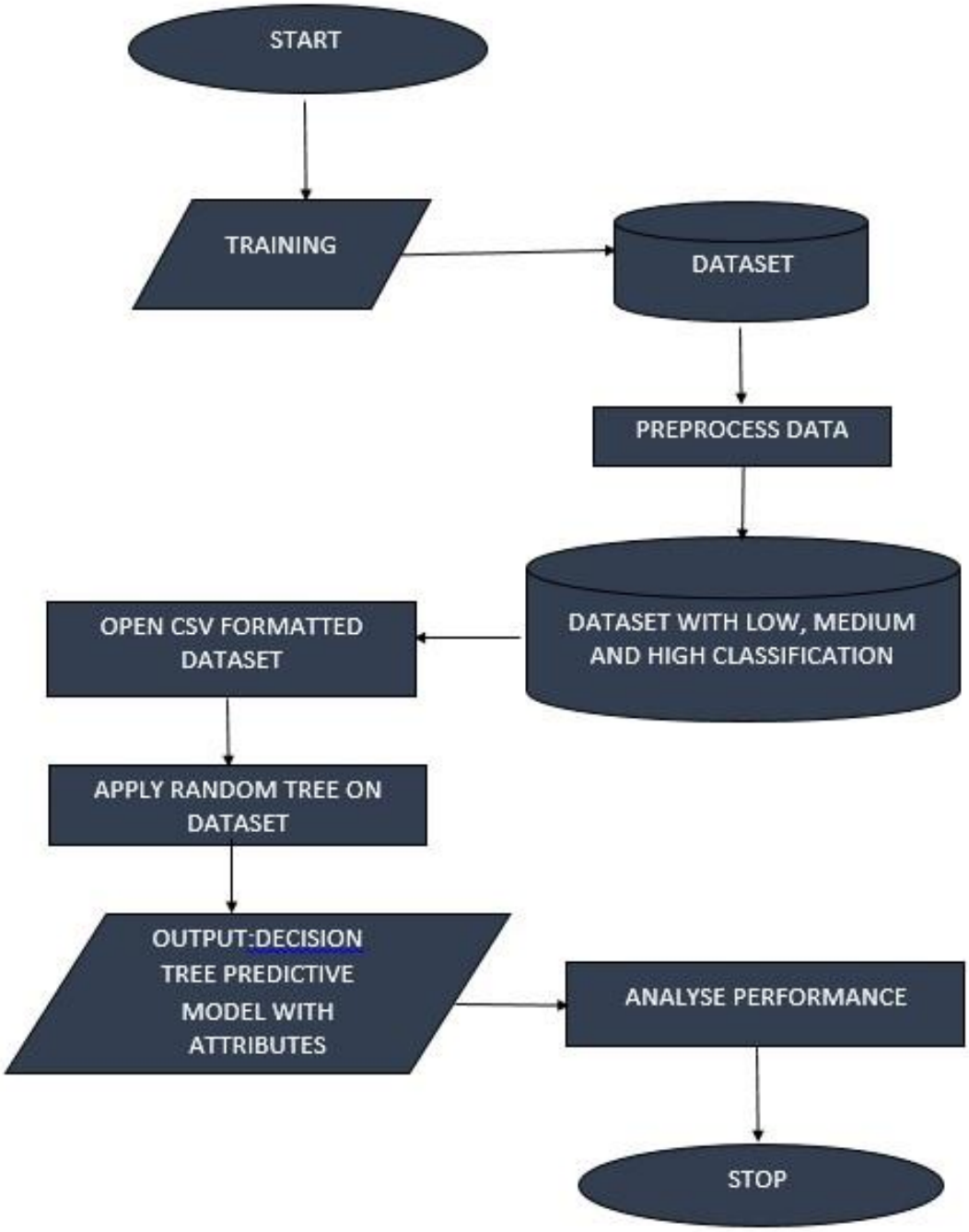
PROJECT DESIGN:
DATA FLOW DIAGRAM:

Date	19 October 2022
Team ID	PNT2022TMID01272
Project Name	Project – Statistical Machine Learning Approaches to Liver Disease Prediction.
Maximum Marks	4 Marks

DATA FLOW DIAGRAMS: DFD 0



PROCESS: DFD1



USER STORIES:

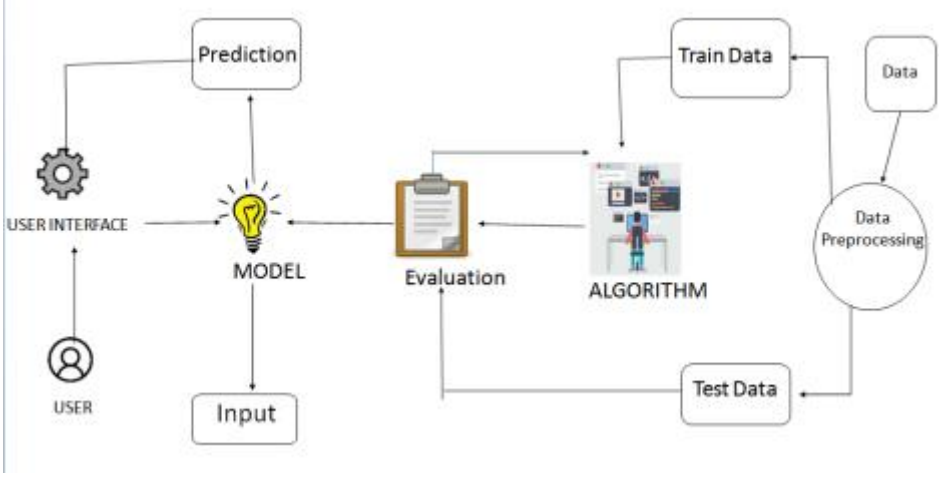
User Type	Functional Requirement (Epic)	User Story Number	User Story / Task	Acceptance criteria	Priority	Release
Customer (Mobile user)	Registration	USN -1	As a user, I can register for the application by entering my email, password, and confirming my password.	I can access my account / dashboard	High	Sprint -1
		USN -2	As a user, I will receive confirmation email once I have registered for the application	I can receive confirmation email & click confirm	High	Sprint -1
		USN -3	As a user, I can register for the application through Face book	I can register & access the dashboard with Facebook Login	Low	Sprint -2
		USN -4	As a user, I can register for the application through Gmail		Medium	Sprint -1
	Login	USN -5	As a user, I can log into the application by entering email & password	Can access the dashboard.	High	Sprint -1
	Dashboard	USN -6	As a user, I Can navigate through different pages using the dashboard	I can various pages	High	Sprint -2

Customer (Web user)	Upload data	USN -6	As a user, I can upload the data that required for finding whether liver disease is there are not.	Can get result based on the information provided.	High	Sprint -3
	Search	USN -7	As a user, I can search for the specialist and best hospital in that respective field.	I can receive information on various doctors and hospitals.	Low	Sprint -4
Administrator	Analyse	USN -8	As an admin, I will analyse the given data.	I can analyse the given data.	High	Sprint -2

SOLUTION & TECHNICAL ARCHITECTURE:

Date	13 October 2022
TeamID	PNT2022TMID01272
Project Name	Statistical Machine Learning Approaches to Liver Disease Prediction
MaximumMarks	2Marks

Solution Architecture:



PROJECT PLANNING & SCHEDULE:

SPRINT PLANNING & ESTIMATION:

Project Planning Template (Product Backlog, Sprint Planning, Stories, Story points)

Date	22 October 2022
Team ID	PNT2022TMID01272
Project Name	STATISTICAL MACHINE LEARNING APPROACH TO LIVER DISEASE PREDICTION
Maximum Marks	8 Marks

Product Backlog, Sprint Schedule, and Estimation (4 Marks)

Sprint	Functional Requirement (Epic)	User Story Number	User Story / Task	Story Points	Priority	Team Members
Sprint -1	Data Collection and Data Preprocessing	USN-1	As a user, I will get the data and take it as a input	10	Medium	KEERTHIKA L
Sprint -2	Visualize the data	USN-2	AS a user, I will get the visualization of the data for further understanding.	20	High	DEEPTHI SHREE S

Sprint -3	Training and Testing the model	USN-3	I will train and test the ML Model with higher accuracy and model performance.	10	Medium	HARITA K S
Sprint -4	Deploying the model into IBM Cloud	USN-4	I will deploy the model into IBM cloud as a web app for making predictions.	20	Medium	DEEPA S

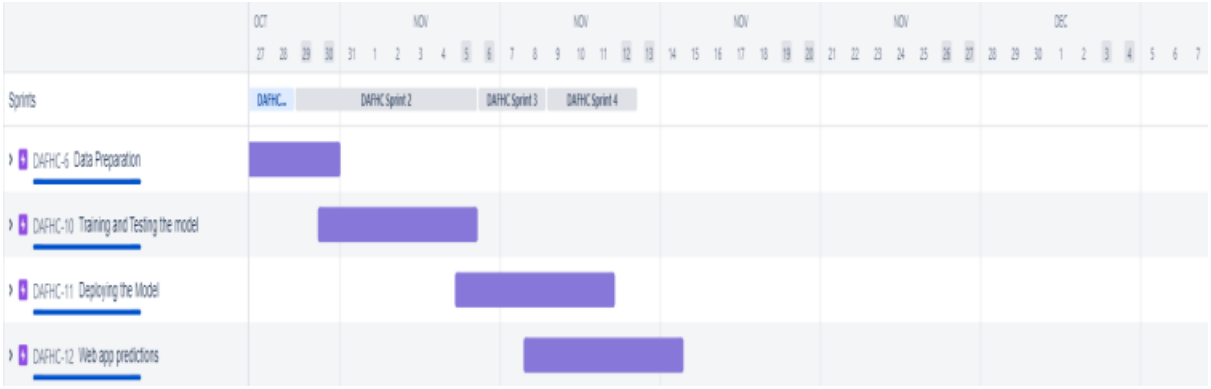
Project Tracker, Velocity & Burndown Chart:

Sprint	Total Story Points	Duration	Sprint Start Date	Sprint End Date (Planned)	Story Points Completed (as on Planned End Date)	Spr (Actual)
Sprint -1	8	7 Days	22 Oct 2022	28 Oct 2022	20	
Sprint -2	8	8 Days	29 Oct 2022	05 Nov 2022	20	
Sprint -3	5	3 Days	06 Nov 2022	08 Nov 2022	20	
Sprint -4	5	4 Days	09 Nov 2022	12 Nov 2022	20	

Velocity:

Imagine we have a 6-day sprint duration, and the velocity of the team is 20 (points per sprint). Let's calculate the team's average velocity (AV) per iteration unit (story points per day)

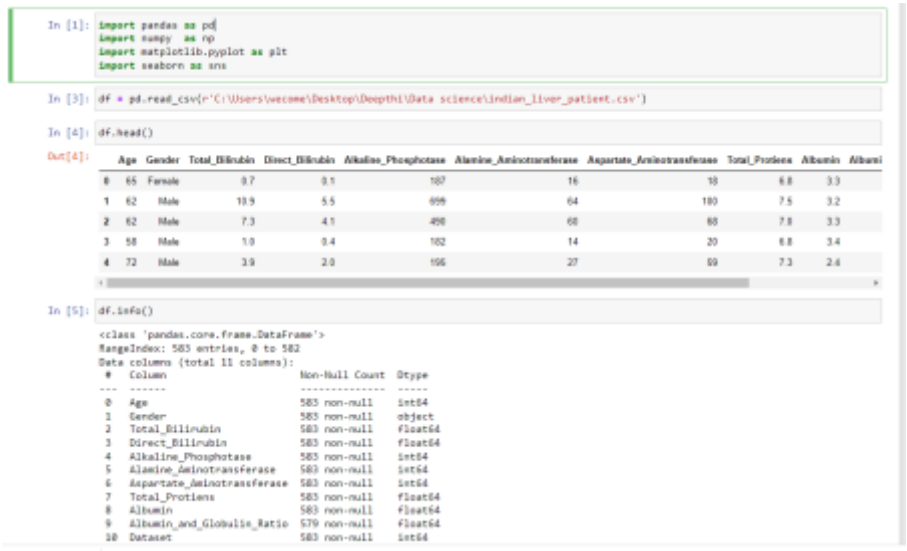
Burndown Chart:



SPRINT DELIVERY SCHEDULE:

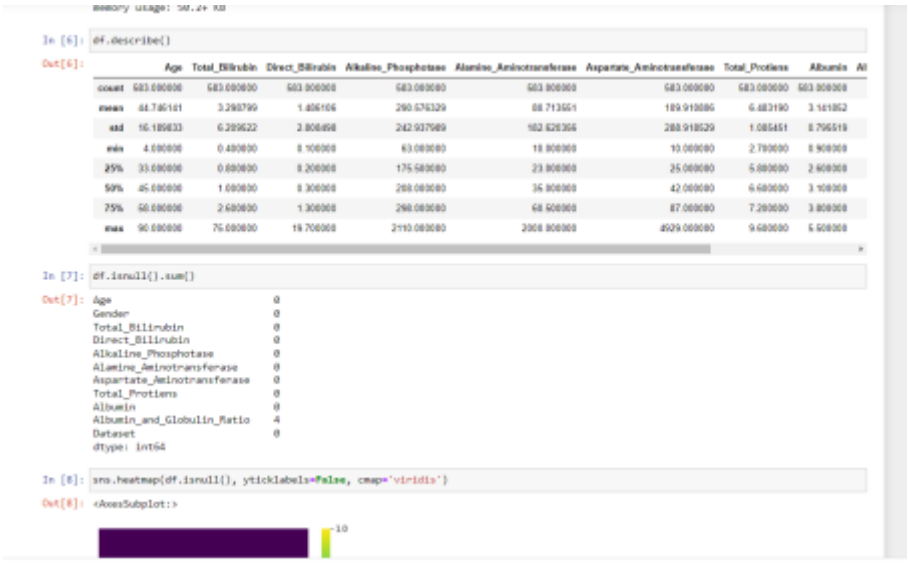
SPRINT-1: PROJECT DELIVERABLES

Data Preparation and Preprocess Reading the dataset:



```
In [1]: import pandas as pd
        import numpy  as np
        import matplotlib.pyplot as plt
        import seaborn as sns

In [3]: df = pd.read_csv(r'C:\Users\welcome\Desktop\Deepthi\Data science\indian_liver_patient.csv')

In [4]: df.head()

Out[4]:
   Age  Gender  Total_Bilirubin  Direct_Bilirubin  Alkaline_Phosphotase  Alamine_Aminotransferase  Aspartate_Aminotransferase  Total_Proteins  Albumin  Albumi
0   65  Female              0.7               0.1                   187                        16                          18             6.8      3.3
1   62    Male             10.9               5.5                   699                        64                         100             7.5      3.2
2   62    Male              7.3               4.1                   490                        60                          68             7.0      3.3
3   58    Male              1.0               0.4                   182                        14                          20             6.8      3.4
4   72    Male              3.9               2.0                   195                        27                          59             7.3      2.4

In [5]: df.info()
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 583 entries, 0 to 582
Data columns (total 11 columns):
 #   Column                      Non-Null Count  Dtype
---  ------                      --------------  -----
 0   Age                         583 non-null    int64
 1   Gender                      583 non-null    object
 2   Total_Bilirubin             583 non-null    float64
 3   Direct_Bilirubin            583 non-null    float64
 4   Alkaline_Phosphotase        583 non-null    int64
 5   Alamine_Aminotransferase    583 non-null    int64
 6   Aspartate_Aminotransferase  583 non-null    int64
 7   Total_Proteins              583 non-null    float64
 8   Albumin                     583 non-null    float64
 9   Albumin_and_Globulin_Ratio  579 non-null    float64
 10  Dataset                     583 non-null    int64
```

Describe dataset:



```
In [6]: df.describe()
```

	Age	Total_Bilirubin	Direct_Bilirubin	Alkaline_Phosphotase	Alamine_Aminotransferase	Aspartate_Aminotransferase	Total_Protiens	Albumin	Al
count	583.000000	583.000000	583.000000	583.000000	583.000000	583.000000	583.000000	583.000000	
mean	44.746141	3.298799	1.486106	290.576329	80.713551	109.910806	6.483190	3.141852	
std	16.189833	6.209522	2.808498	242.937989	182.620356	288.918529	1.085451	0.795519	
min	4.000000	0.400000	0.100000	63.000000	10.000000	10.000000	2.700000	0.900000	
25%	33.000000	0.800000	0.200000	175.500000	23.000000	25.000000	5.800000	2.600000	
50%	45.000000	1.000000	0.300000	208.000000	35.000000	42.000000	6.600000	3.100000	
75%	58.000000	2.600000	1.300000	298.000000	60.500000	87.000000	7.200000	3.800000	
max	90.000000	75.000000	19.700000	2110.000000	2000.000000	4929.000000	9.600000	5.500000	

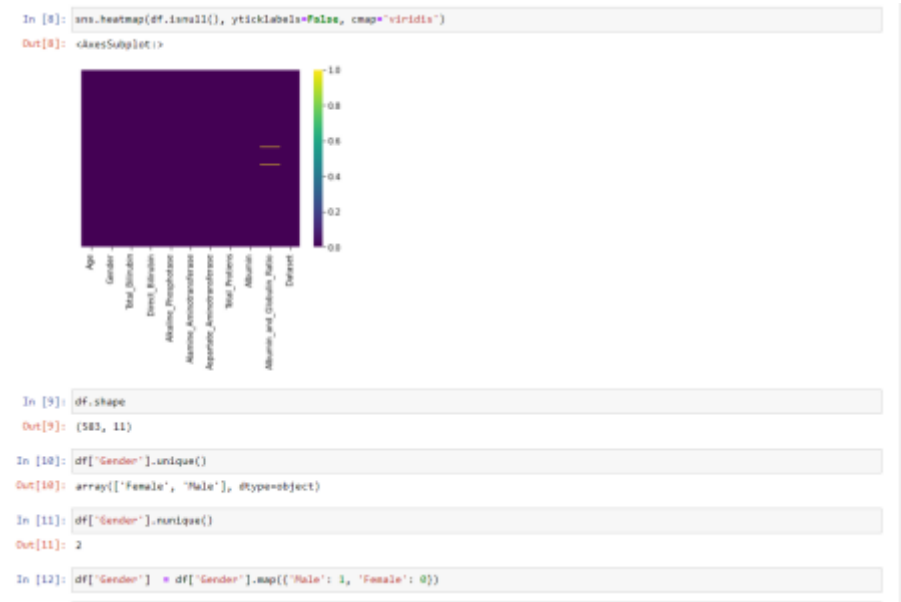
```
In [7]: df.isnull().sum()
```

```
Out[7]: Age                         0
Gender                      0
Total_Bilirubin             0
Direct_Bilirubin            0
Alkaline_Phosphotase        0
Alamine_Aminotransferase    0
Aspartate_Aminotransferase  0
Total_Protiens              0
Albumin                     0
Albumin_and_Globulin_Ratio  4
Dataset                     0
dtype: int64
```

```
In [8]: sns.heatmap(df.isnull(), yticklabels=False, cmap='viridis')
```

```
Out[8]: <AxesSubplot:>
```

Heatmap for Visualizing Null Values:



```
In [8]: sns.heatmap(df.isnull(), yticklabels=False, cmap='viridis')
```

```
Out[8]: <AxesSubplot:>
```

```
In [9]: df.shape
```

```
Out[9]: (583, 11)
```

```
In [10]: df['Gender'].unique()
```

```
Out[10]: array(['Female', 'Male'], dtype=object)
```

```
In [11]: df['Gender'].nunique()
```

```
Out[11]: 2
```

```
In [12]: df['Gender']  = df['Gender'].map({'Male': 1, 'Female': 0})
```

Dropping Null Values from Dataset:

```
df['Gender']  = df['Gender'].map({'Male': 1, 'Female': 0})
```

```
df.head()
```

	Age	Gender	Total_Bilirubin	Direct_Bilirubin	Alkaline_Phosphotase	Alamine_Aminotransferase	Aspartate_Aminotransferase	Total_Protiens	Albumin	Albumi
0	65	0	0.7	0.1	187	16	18	6.8	3.3	
1	62	1	10.9	5.5	699	64	100	7.5	3.2	
2	62	1	7.3	4.1	490	60	68	7.0	3.3	
3	58	1	1.8	0.4	182	14	20	6.8	3.4	
4	72	1	3.9	2.0	195	27	59	7.3	2.4	

```
df.dropna(inplace=True)
```

```
df.shape
```

```
(579, 11)
```

```
df.head()
```

	Age	Gender	Total_Bilirubin	Direct_Bilirubin	Alkaline_Phosphotase	Alamine_Aminotransferase	Aspartate_Aminotransferase	Total_Protiens	Albumin	Albumi
0	65	0	0.7	0.1	187	16	18	6.8	3.3	
1	62	1	10.9	5.5	699	64	100	7.5	3.2	
2	62	1	7.3	4.1	490	60	68	7.0	3.3	
3	58	1	1.8	0.4	182	14	20	6.8	3.4	
4	72	1	3.9	2.0	195	27	59	7.3	2.4	

```
df['Dataset'].unique()
```

```
array([1, 2], dtype=int64)
```

Heatmap to check if there is any Null Value:

```
df['Dataset'].unique()
```

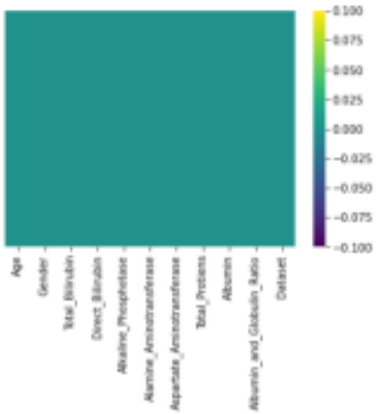
```
array([1, 2], dtype=int64)
```

```
df['Dataset'].value_counts()
```

```
1    414
2    165
Name: Dataset, dtype: int64
```

```
sns.heatmap(df.isnull(),yticklabels=False,cmap='viridis')
```

```
<AxesSubplot:>
```




```
df.corr()['Dataset']
```

EDA : Exploratory Data Analysis

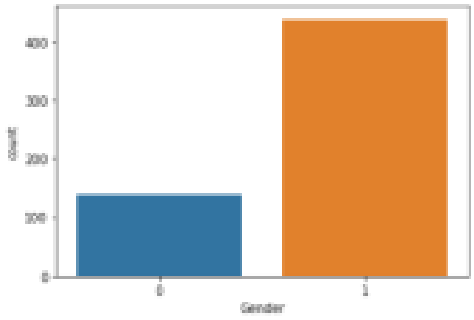
Uni – variate Analysis:

```
df.head()
```

	Age	Gender	Total_Bilirubin	Direct_Bilirubin	Alkaline_Phosphotase	Alamine_Aminotransferase	Aspartate_Aminotransferase	Total_Protiens	Albumin	Albumi
0	65	0	0.7	0.1	187	16	18	6.8	3.3	
1	62	1	10.9	5.5	699	64	100	7.5	3.2	
2	62	1	7.3	4.1	490	60	68	7.0	3.3	
3	58	1	1.0	0.4	182	14	20	6.8	3.4	
4	72	1	3.9	2.0	195	27	59	7.3	2.4	

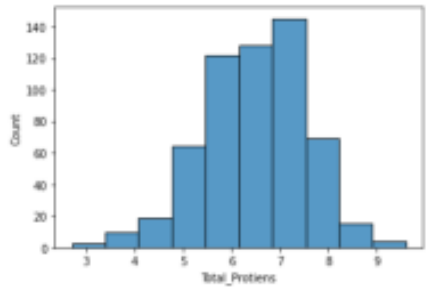
```
sns.countplot(x='Gender',data=df, dodge=True)
```

```
<AxesSubplot:xlabel='Gender', ylabel='count'>
```



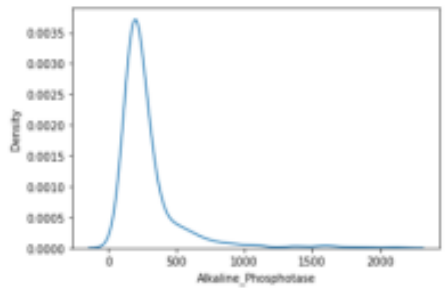

```
sns.histplot(x='Total_Protiens',data=df,bins=10)
```

```
<AxesSubplot:xlabel='Total_Protiens', ylabel='Count'>
```




```
sns.kdeplot(x='Alkaline_Phosphotase', data=df)
```

```
<AxesSubplot:xlabel='Alkaline_Phosphotase', ylabel='Density'>
```




```
sns.boxplot(x='Albumin and Globulin Ratio',data=df)
```

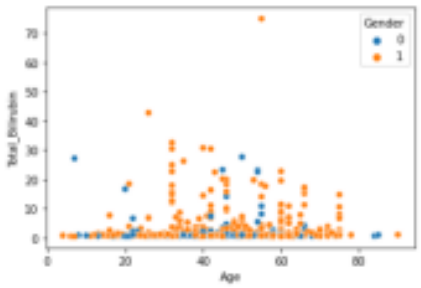
Bi – variate Analysis:

```
df.head()
```

	Age	Gender	Total_Bilirubin	Direct_Bilirubin	Alkaline_Phosphotase	Alamine_Aminotransferase	Aspartate_Aminotransferase	Total_Proteins	Albumin	Albumi
0	65	0	0.7	0.1	187	16	18	6.8	3.3	
1	62	1	10.9	5.5	699	64	100	7.5	3.2	
2	62	1	7.3	4.1	490	60	68	7.0	3.3	
3	58	1	1.0	0.4	182	14	20	6.8	3.4	
4	72	1	3.9	2.0	195	27	59	7.3	2.4	

```
sns.scatterplot(x='Age',y='Total_Bilirubin',data=df,hue='Gender')
```

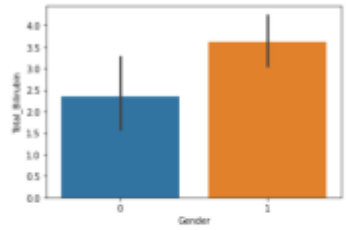
```
<AxesSubplot:xlabel='Age', ylabel='Total_Bilirubin'>
```




```
sns.barplot(x='Gender',y='Total_Bilirubin',data=df)
```

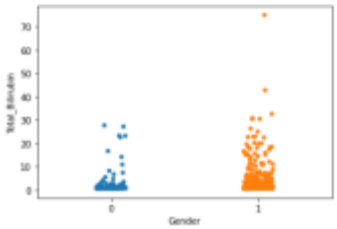
```
sns.barplot(x='Gender',y='Total_Bilirubin',data=df)
```

```
<AxesSubplot:xlabel='Gender', ylabel='Total_Bilirubin'>
```



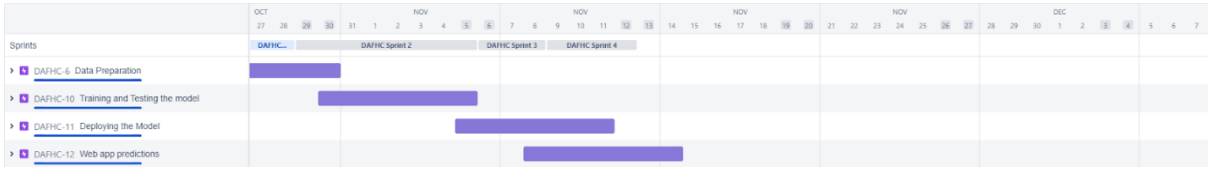

```
sns.stripplot(x='Gender',y='Total_Bilirubin',data=df)
```

```
<AxesSubplot:xlabel='Gender', ylabel='Total_Bilirubin'>
```



Multi – variate Analysis:

REPORTS FROM JIRA



CODING & SOLUTIONS

FEATURE 1: Jupyter Notebook

In Jupyter Notebook, dataset is uploaded and read. After importing the dataset into notebook to extract the useful information and crop the unwanted information we do data preprocessing. After this, Heatmap is use to check the null values in the dataset. If any null values present, dropna command is used to rectify the null values. Next step is to detect the outliers. To detect the outliers we use flooring and capping technique. Here values above upper limit is considered as outliers and values below the lower limits is considered as outliers. We replace

those outliers with upper limit and lower limit values. Next we perform Data Visualization. In data visualization we do EDA - Exploratory Data Analysis, where we do three types of analysis. They are: 1. Uni-Variate Analysis, 2. Bi-Variate Analysis, 3. Multi-Variate Analysis. Next the dataset is split into train and test. Later, dataset is trained in three algorithms to check which algorithm gives best accuracy. Then, the algorithm which gives best accuracy is converted in pickle format to be deployed.

FEATURE 2: Visual Studio Code

After getting the pickle format, inorder to deploy it as a web application run using local host site we use a python library called Flask. It is the main application file used to route the HTML files and get into form and processs it after it unpickles our model. It is used to classify the target with input data by getting data from the prediction form. We have four HTML forms. They are: 1. Index Page, 2. Home Page, 3. Prediction Page, 4. Result Page(No Chance, Chance). In prediction.html, it contains the form which is used to get the input data used.

TESTING:

TEST CASES:
HTML CODING:

1.INDEX PAGE:

```
<!DOCTYPEhtml>
<html>

<head>
    <metacharset="utf-8">
<style>
    body {
    background-image:
url('https://png.pngtree.com/thumb_back/fh260/background/20220217/pngtree-
fresh-wind-hand-painted-illustration-background-for-preventing-liver-disease-
image_946993.jpg');
    background-repeat: no-repeat;
    background-attachment: fixed;
    background-size: cover;
}
</style>
    <title>HOME PAGE</title>

    <link  rel="stylesheet"href="/css/style.css"type="text/css"/>

</head>
```

```
<body>

    <sectionid="main">
        <centre>
            <nav>

            <spanclass="menu-space"></span>
            <h2>Liver Patient Analysis</h2>
            <ul class="menu">

                <li><ahref="C:\Users\wecome\Documents\Liver_disease\templates\
Homepage2.html">Home</a></li>
                <li><ahref="C:\Users\wecome\Documents\Liver_disease\templates\
prediction.html">Prediction</a></li>
             </ul>
             </nav>
        </centre>
    </section>

    <divclass="content">

        <divclass="main-text">

            <h3>Introduction</h3>
            <p style="font-size: large";>
                 <pstyle="color:rgb(1, 1, 1)">
                 <strong>
                 Liver diseases avert the normal function of the liver.
                 Mainly due to the large amount of alcohol consumption liver
disease arises. Early prediction of liver disease using classification
algorithms is an efficacious task that can help the doctors to diagnose the
disease within a short duration of time.
                 Discovering the existence of liver disease at an early stage
is a complex task for the doctors.
                 The main objective of this project is to analyze the
parameters of various classification algorithms and compare their predictive
accuracies so as to find out the best classifier for determining the liver
disease.
                 With a growing trend of sedentary and lack of physical
activities, diseases related to liver have become a common encounter nowadays.
                 In rural areas the intensity is still manageable, but in
urban areas, and especially metropolitan areas the liver disease is a very
common sighting nowadays.
                 Liver diseases cause millions of deaths every year. Viral
hepatitis alone causes 1.34 million deaths every year.
                 Problems with liver patients are not easily discovered in an
early stage as it will be functioning normally even when it is partially
damaged.
                 An early diagnosis of liver problems will increase patients
survival rate.
                 Liver failures are at high rate of risk among Indians.
                 It is expected that by 2025 India may become the World
Capital for Liver Diseases.
                 The widespread occurrence of liver infection in India is
contributed due to deskbound lifestyle, increased alcohol consumption and
smoking.
```

```
                There are about 100 types of liver infections.
                With such alarming figures, it is necessary to have a concern
towards tackling these diseases.
                Afterall, we cannot expect a developed and prosperous nation,
with unhealthy youths.
                </strong>
                </p>

            |</p>
        </div>
    </div>

</body>

</html>
```

2.HOME PAGE:

```
<Html>
<head>
<style>

body {
background-image: url('https://media.istockphoto.com/id/1414176792/photo/top-
view-photo-of-pink-silk-ribbon-symbol-of-breast-cancer-awareness-and-
stethoscope-on.jpg?s=612x612&w=0&k=20&c=7sjsU3Xu_kmxV349N-
_LxWEAjYC8cBWkAFvW5lF3WfA=');
background-repeat: no-repeat;
background-attachment: fixed;
background-size: cover;
}

</style>
</head>
<body>

<title>home </title>
<h1><b>LIVER DISEASE </b></h1></head>
<body>
<h3><b>SYMTOMS:</b></h3>
<p> 1.Skin and eyes that appear yellowish (jaundice)<br>
    2.Abdominal pain and swelling<br>
    3.Swelling in the legs and ankles<br>
    4.Itchy skin<br>
    5.Dark urine color<br>
    6.Pale stool color<br>
    7.Chronic fatigue<br>
    8.Nausea or vomiting<br>
    9.Loss of appetite<br>
    10.Tendency to bruise easily<br>
</p><br>

<h3><b>TOTAL BILIRUBIN: </b></h3>
<p> This is a blood test that measures the amount of a substance called
bilirubin.
    This test is used to find out how well your liver is working.
```

```
    It is often part of a panel of tests that measure liver function.
    A small amount of bilirubin in your blood is normal, but a high level may
be a sign of liver disease.
</p>

<h3><b>DIRECT_BILIRUBIN:</b></h3>
<p> In the liver, bilirubin is changed into a form that your body can get rid
of.
    This is called conjugated bilirubin or direct bilirubin.
    This bilirubin travels from the liver into the small intestine.
    A very small amount passes into your kidneys and is excreted in your
urine.
    This bilirubin also gives urine its distinctive yellow color.
</p><br>

<h3><b>ALKALINE_PHOSPHOTASE:</b></h3>
<p>An alkaline phosphatase level test (ALP test) measures the amount of
alkaline phosphatase enzyme in your bloodstream.
</p><br>

<h3><b>ALBUMIN:</b></h3>
<p> An albumin blood test is used to check your general health and to see how
well your liver and kidneys are working.
    If your liver is damaged or you're not well nourished, your liver may not
make enough albumin.
</p><br>
</body>
</html>
```

3.PREDICTION PAGE:

```
<html>
<head>
<title>  LIVER PATIENT PREDICTION </title>

<h1> Prediction page: </h1>

</head>

<style>
.one{
  text-align: center;
}
body{
  background-color: gold;
}
</style>

<body>
  <divclass="container">
    <h2class='container-heading'><spanclass="heading_font">Liver Disease
Prediction</span></h2>
</div>
<centre>
<imgsrc="C:\Users\wecome\Documents\Liver_disease\static\images\liver.jpg"  wid
th="200"height="100"/>
```

```
<br>
<divclass = "ml-container">
<form  action="{{url_for('prediction')}}"  method="POST">
<br>
<br>
<label> Age: </label>
<inputid="first"name="age"required="require"size="20"/><br>
<br>
<labelfor="Gender"required="require">Choose Gender:</label>
<selectid="Gender"name="Gender">
  <optionvalue="0">Female</option>
  <optionvalue="1">Male</option>
</select><br>
<br>
<label> Total_Bilirubin: </label>
<inputid="second"name="totalbilirubin"required="require"size="20"/><br><br>

<label> Direct_Bilirubin: </label>
<inputid="third"name="directbilirubin"required="require"size="20"/><br><br>

<label> Alkaline_Phosphotase: </label>
<inputid="fourth"name="alkalinephosphotase"required="require"size="20"/><br><b
r>

<label> Alamine_Aminotransferase: </label>
<inputid="fifth"name="alamine"required="require"size="20"/><br><br>

<label> Aspartate_Aminotransferase: </label>
<inputid="sixth"name="aspartate"required="require"size="20"/><br><br>

<label> Total_Proteins: </label>
<inputid="seventh"name="totalproteins"required="require"size="20"/><br><br>

<label> Albumin: </label>
<inputid="eigth"name="albumin"required="require"size="20"/><br><br>

<label> Albumin_and_Globulin_Ratio: </label>
<inputid="ninth"name="albuminandglobulin"required="require"size="20"/><br><br>

<buttonid="sub"type="submit">Predict</button>


</form>
</div>
</centre>
<style>
body{
  font-family: Arial, Helvetica,sans-serif;
    text-align: center;
    margin: 0;
    padding: 0;
    width: 100%;
  height: 100%;
  display: flex;
```

```
  flex-direction: column;
}
.container-heading{
  margin:0;
}
.heading_font{
  color: black;
  font-family: 'Pacifico', cursive;
  font-size: 50px;
  font-weight: normal;
}
  #first{
    border-radius: 14px;
    height: 25px;
    width: 150px;
    font-size: 20px;
    text-align: center;
  }
  #second{
    border-radius: 14px;
    height: 25px;
    width: 150px;
    font-size: 20px;
    text-align: center;
  }
  #third{
    border-radius: 14px;
    height: 25px;
    width: 150px;
    font-size: 20px;
    text-align: center;
  }
  #fourth{
    border-radius: 14px;
    height: 25px;
    width: 150px;
    font-size: 20px;
    text-align: center;
  }
  #fifth{
    border-radius: 14px;
    height: 25px;
    width: 150px;
    font-size: 20px;
    text-align: center;
  }
  #sixth{
    border-radius: 14px;
    height: 25px;
    width: 150px;
    font-size: 20px;
    text-align: center;
  }
  #seventh{
    border-radius: 14px;
    height: 25px;
```

```
    width: 150px;
    font-size: 20px;
    text-align: center;
  }
  #eigth{
    border-radius: 14px;
    height: 25px;
    width: 150px;
    font-size: 20px;
    text-align: center;
  }
  #ninth{
    border-radius: 14px;
    height: 25px;
    width: 150px;
    font-size: 20px;
    text-align: center;
  }

  #sub{
    width: 120px;
    height: 43px;
    text-align: center;
    border-radius: 14px;
    font-size: 18px;
  }
</style>
</body>
</html>
```

4.NO CHANCE PAGE:

```
<!DOCTYPEhtml>
<htmllang="en">

<head>
    <metacharset="UTF-8">
    <metaname="viewport"content="width=device-width, initial-scale=1.0">
    <title>Liver Disease Result</title>
</head>

<body>

    <divclass="container">
        <formaction="{{ url_for('prediction')}}" method="post">
            <h2 class='container-heading'><spanclass="heading_font">Liver
Disease Prediction</span></h2>

        <br><br><br><br><br><br><br>

        <!-- Result -->
        <divclass="results">
                <h1><spanclass='safe'>  Congratulation!  <br><br>You DON'T
have LIVER DISEASE.</span></h1>
        </div>
        </form>
```

```
    </div>


<style>

/* Color */
body{
    font-family: Arial, Helvetica,sans-serif;
    text-align: center;
    margin: 0;
    padding: 0;
    width: 100%;
    height: 100%;
    display: flex;
    flex-direction: column;
}


/* Heading Font */
.container-heading{
    margin: 0;
}

.heading_font{
    color: black;
    font-family: 'Pacifico', cursive;
    font-size: 50px;
    font-weight: normal;
}
</style>
</body>
</html>
```

5.CHANCE PAGE:

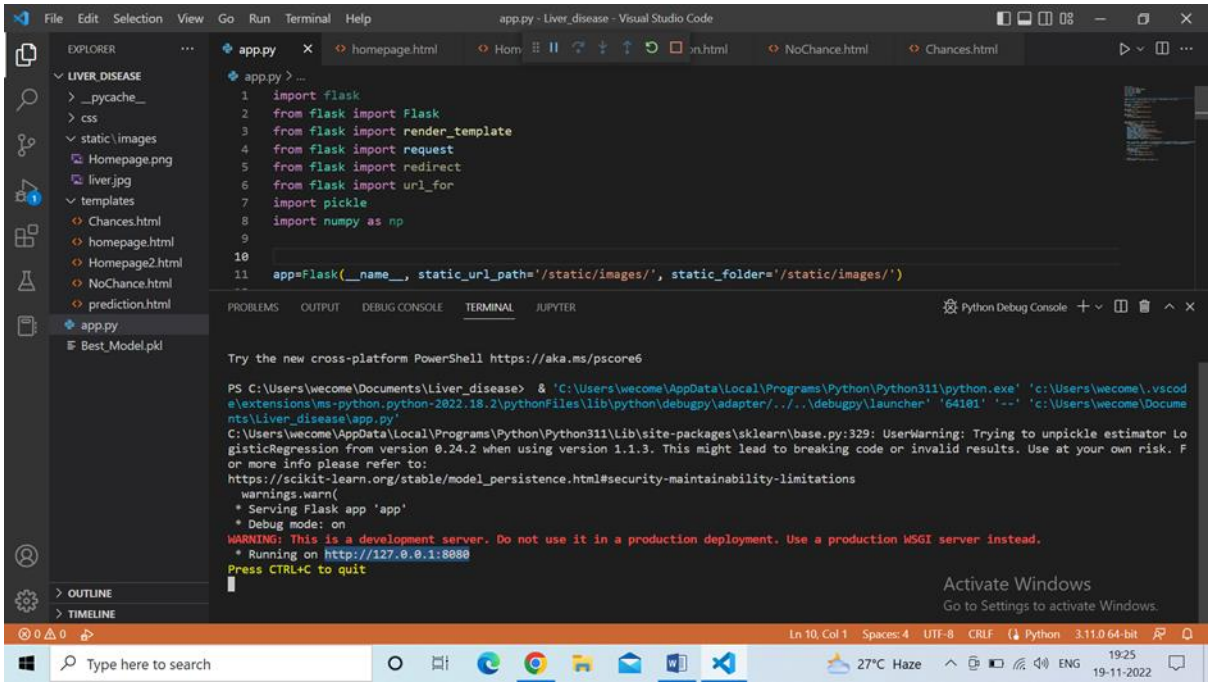
```
<!DOCTYPEhtml>
<htmllang="en">

<head>
    <metacharset="UTF-8">
    <metaname="viewport"content="width=device-width, initial-scale=1.0">
    <title>Liver Disease Result</title>
</head>

<body>

    <divclass="container">
        <formaction="{{ url_for('prediction')}}" method="post">
            <h2 class='container-heading'><spanclass="heading_font">Liver
Disease Prediction</span></h2>

        <br><br><br><br><br><br><br>

        <!-- Result -->
        <divclass="results">
```

```
                <h1><spanclass='danger'>Oops! ☹<br><br>You have LIVER DISEASE
<br><br>Please Consult a Doctor.</span></h1>

        </div>
        </form>

    </div>
    <div>
            <br><br><br><br><br><br><br><br><br><br>

    </div>


<style>

/* Color */
body{
    font-family: Arial, Helvetica,sans-serif;
    text-align: center;
    margin: 0;
    padding: 0;
    width: 100%;
    height: 100%;
    display: flex;
    flex-direction: column;
}


/* Heading Font */
.container-heading{
    margin: 0;
}

.heading_font{
    color: black;
    font-family: 'Pacifico', cursive;
    font-size: 50px;
    font-weight: normal;
}
</style>
</body>
</html>
```

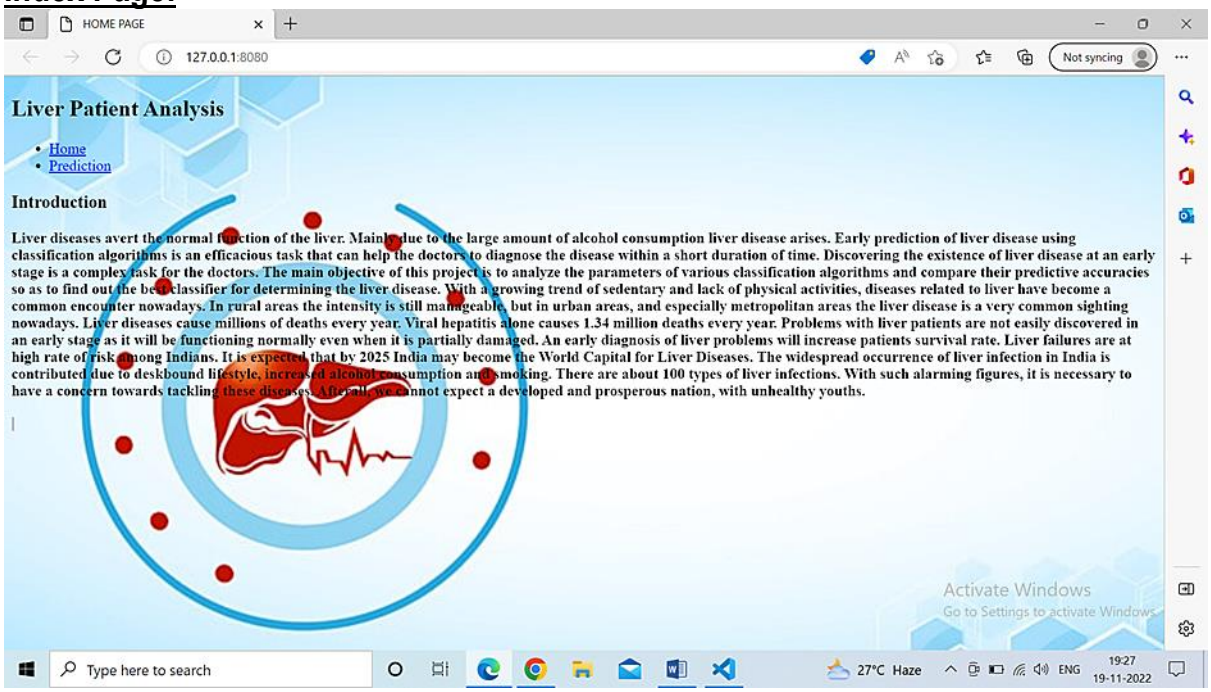
USER ACCEPTANCE TESTING:

RESULTS:

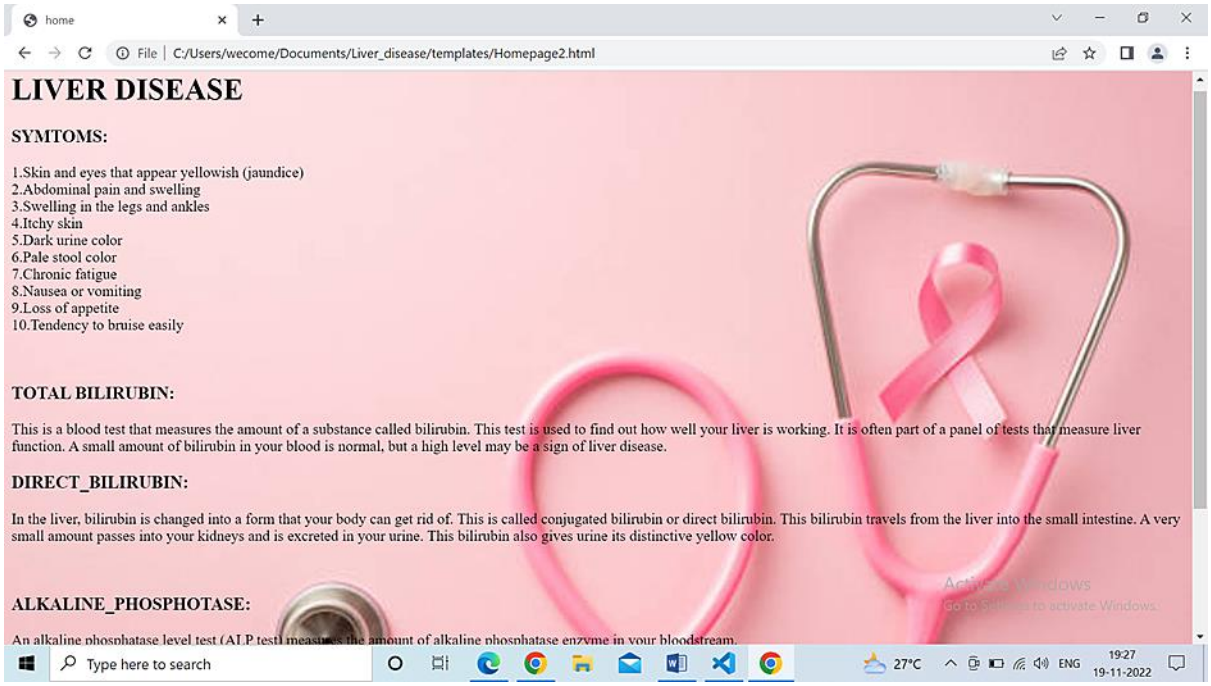
Local Host Site:



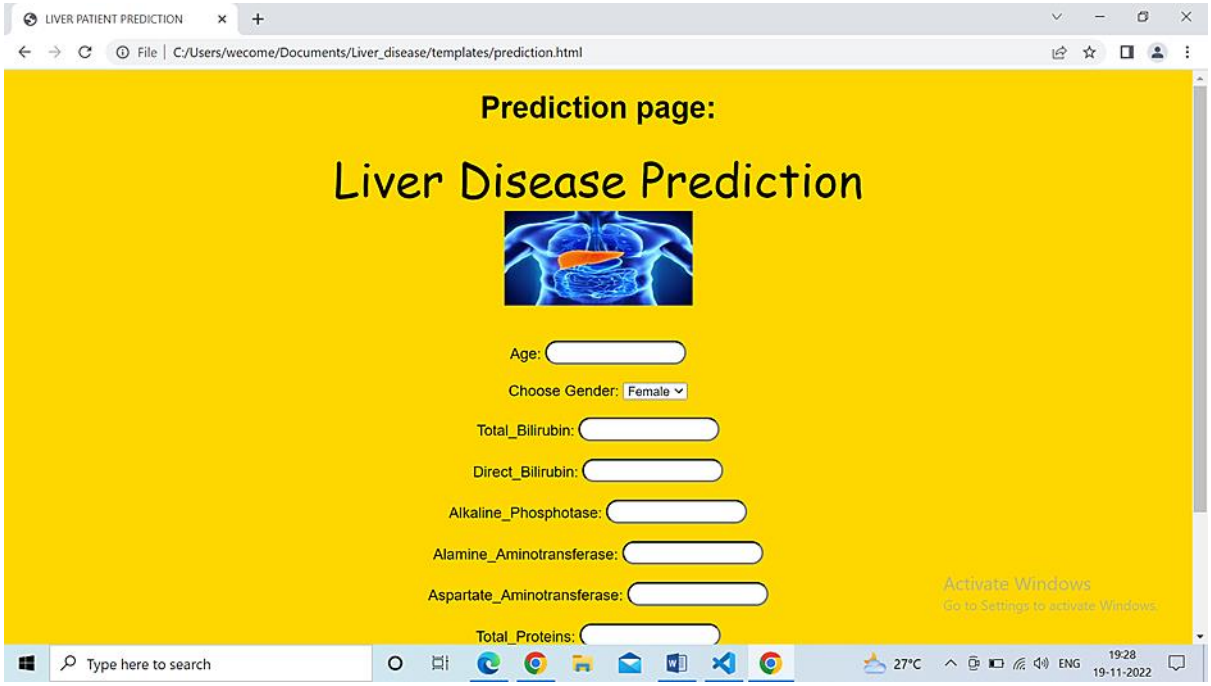
Index Page:

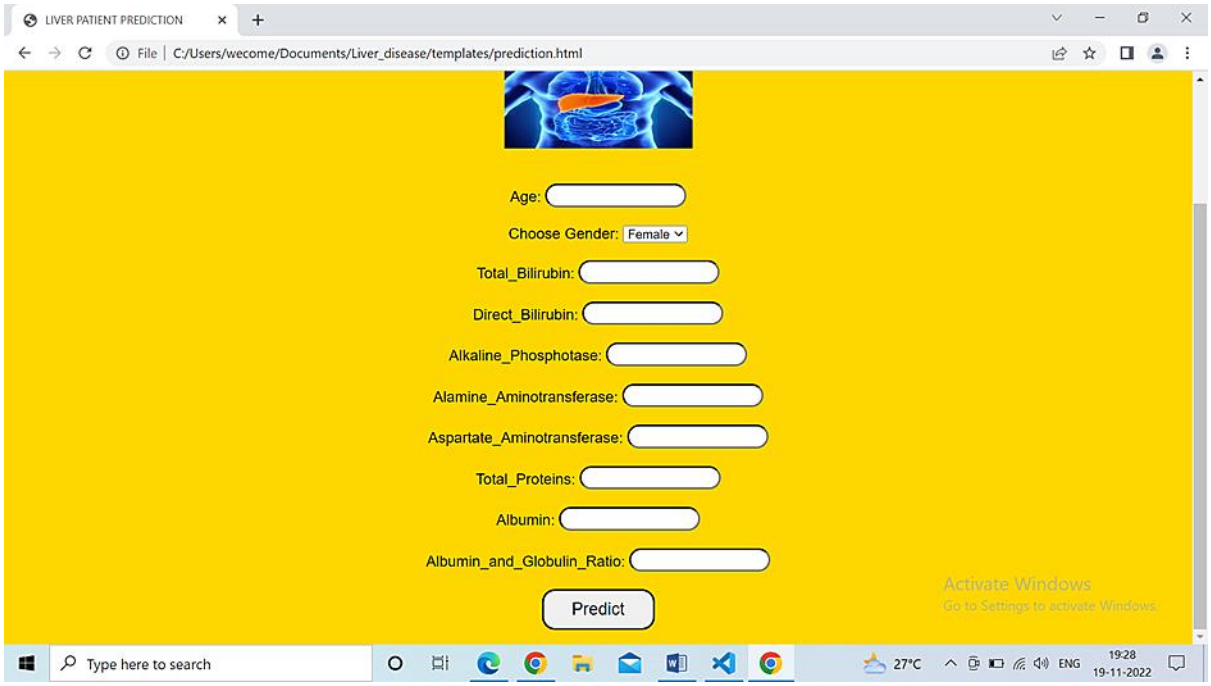


Home Page:



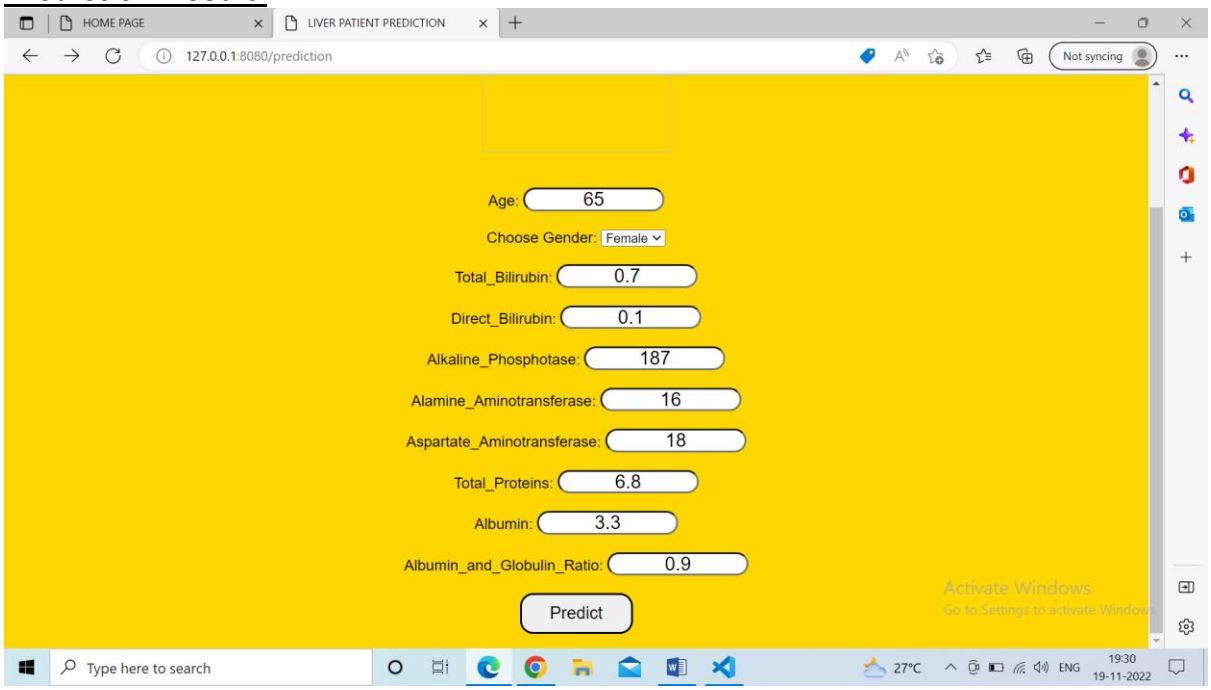
Prediction Page:



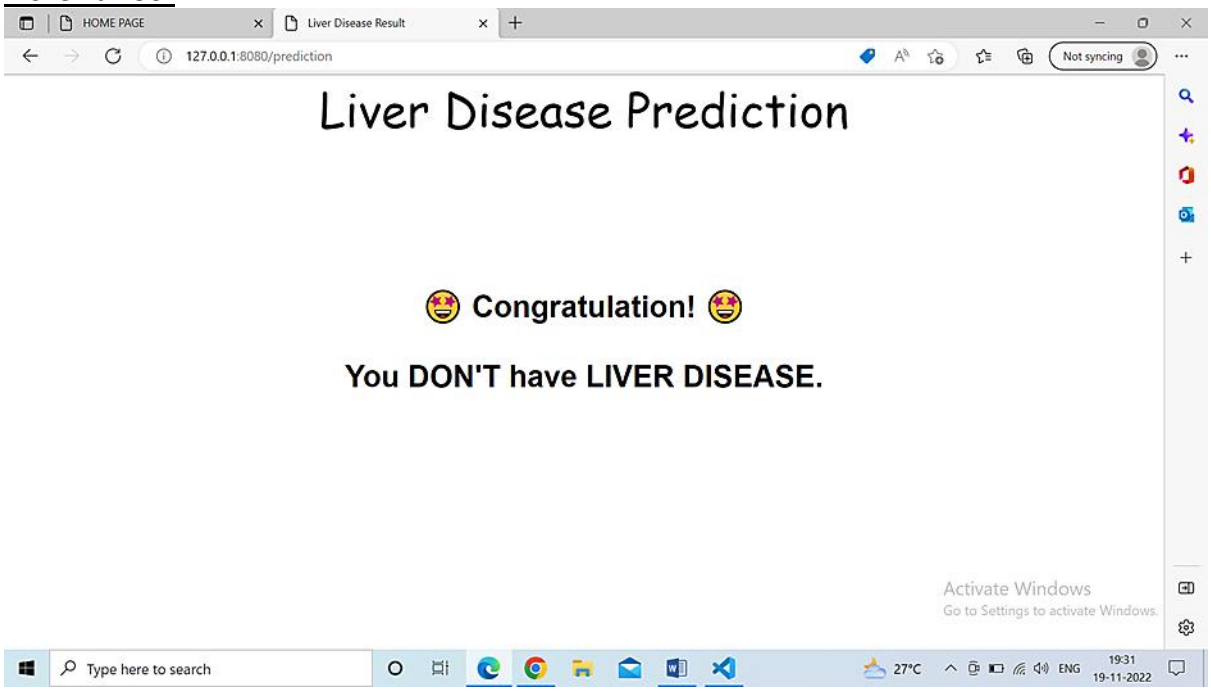


PERFORMANCE METRICS:

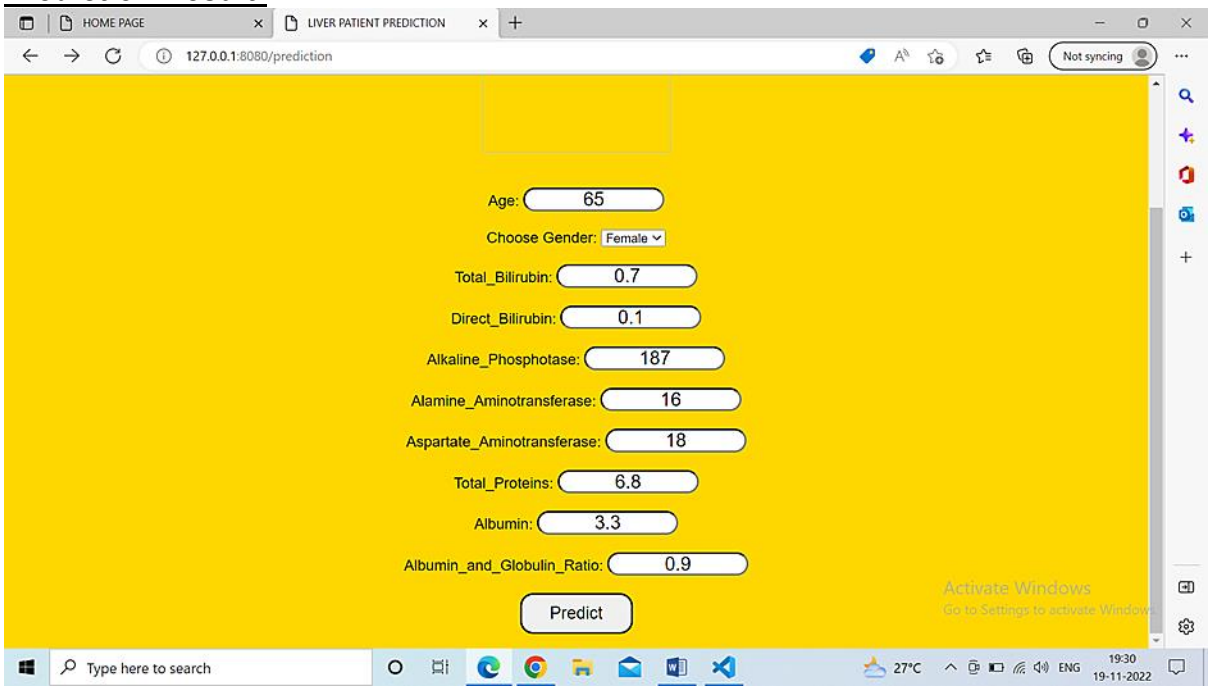
Prediction Result:



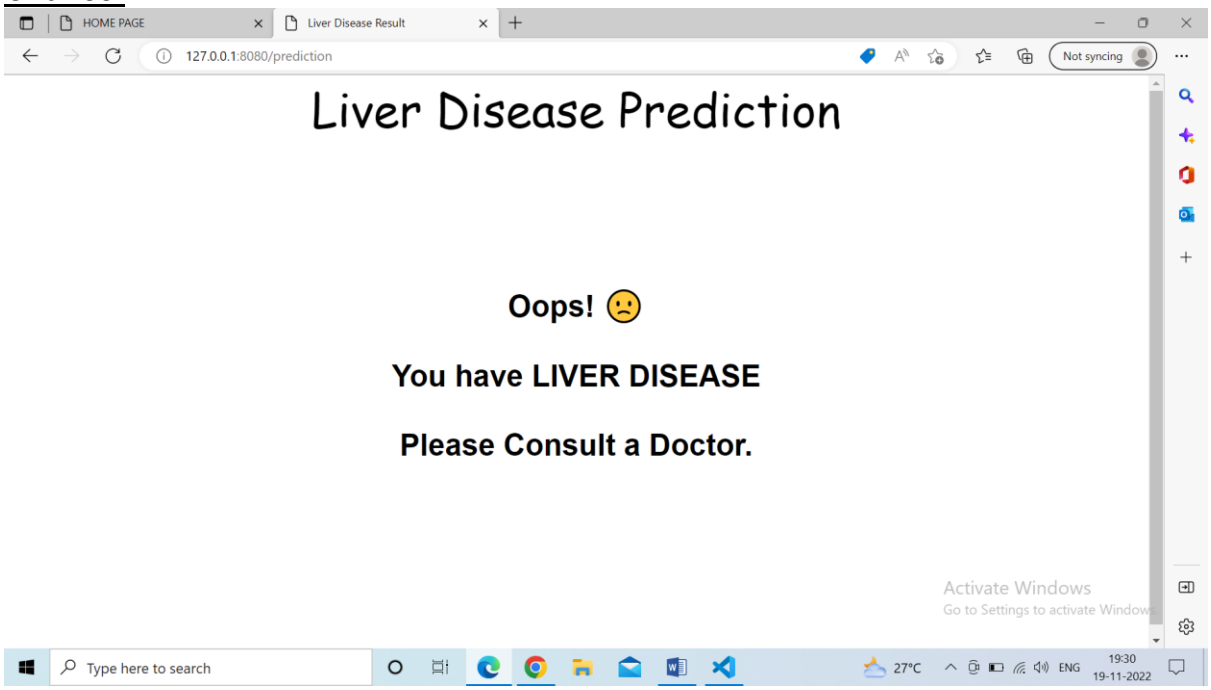
No Chance:



Prediction Result:



Chance:



Conclusion:

In this paper, we proposed and built a machine learning based on a hybrid classifier to be used as a classification model for liver diseases diagnosis to improve performance and experts to identify the chances of disease and conscious prescription of further treatment healthcare and examinations. In future work. the use of fast datasets technique like Apache Hadoop or Spark can be incorporated with this technique. In addition to this, we can use distributed refined algorithms like Forest Tree implemented in Apache Hadoop to increase scalability and efficiency.

FUTURE SCOPE:

In future, along with this analysis of the input data, using Machine Learning algorithms for predictive analysis may be more effective. The predictions and classification of user data can be done by using the ML algorithms easily and a full fledged application can be created for analysis and predictions of the Patient

data can be done near future. And the results can be directly send to the patient email address without delay.

SOURCE CODE:

App.py:(Main)

```
importflask
fromflask importFlask
fromflask importrender_template
fromflask importrequest
fromflask importredirect
fromflask importurl_for
importpickle
importnumpy asnp


app=Flask(__name__, static_url_path='/static/images/',
static_folder='/static/images/')

model = pickle.load(open('Best_Model.pkl', 'rb'))

@app.route('/',methods=['GET'])
defHome():
    returnrender_template('homepage.html')

@app.route('/nochance')
defresult():
    returnrender_template("NoChance.html")

@app.route('/chance')
defresult1():
    returnrender_template("Chances.html")



@app.route('/prediction', methods=['POST','GET'])
defprediction():
    ifrequest.method == "GET":
        returnrender_template("prediction.html")
    age = request.form["age"]
    gender = request.form["Gender"]
    totalbilirubin = request.form['totalbilirubin']
    directbilirubin = request.form['directbilirubin']
    alkalinephosphotase = request.form['alkalinephosphotase']
    alamine = request.form['alamine']
    aspartate = request.form['aspartate']
    totalproteins = request.form['totalproteins']
    albumin = request.form['albumin']
    albuminandglobulin = request.form['albuminandglobulin']


    data = [(float(age),  float(gender),float(totalbilirubin),
float(directbilirubin), float(alkalinephosphotase),
    float(alamine), float(aspartate), float(totalproteins), float(albumin),
float(albuminandglobulin))]
```

```
    model = pickle.load(open('Best_Model.pkl', 'rb'))

    prediction = model.predict(data)[0]
    print(prediction)
    if (prediction == 1):
        returnrender_template('Chances.html')
    else:
        returnrender_template('NoChance.html')


if __name__ == '__main__':
    app.run(debug=True, port=8080, use_reloader=False)
```

Index Page:

```
<!DOCTYPE html>
<html>

<head>
   <meta charset="utf-8">
<style>
   body {
   background-image: url('https://png.pngtree.com/thumb_back/fh260/background/20220217/pngtree-fresh-wind-hand-painted-illustration-background-for-preventing-liver-disease-image_946993.jpg');
   background-repeat: no-repeat;
   background-attachment: fixed;
   background-size: cover;
}
</style>
   <title>HOME PAGE</title>

   <link  rel="stylesheet" href="/css/style.css" type="text/css" />

</head>

<body>

   <section id="main">
      <centre>
         <nav>

         <span class="menu-space"></span>
         <h2>Liver Patient Analysis</h2>
         <ul class="menu">

            <li><a
href="C:\Users\wecome\Documents\Liver_disease\templates\Homepage2.html">Home</a></li>
            <li><a
href="C:\Users\wecome\Documents\Liver_disease\templates\prediction.html">Prediction</a></li>
          </ul>
          </nav>
      </centre>
   </section>
```

```
  <div class="content">

    <div class="main-text">

      <h3>Introduction</h3>
      <p style="font-size: large";>
        <p style="color:rgb(1, 1, 1)">
        <strong>
        Liver diseases avert the normal function of the liver.
        Mainly due to the large amount of alcohol consumption liver disease arises. Early prediction of liver disease using classification algorithms is an efficacious task that can help the doctors to diagnose the disease within a short duration of time.
        Discovering the existence of liver disease at an early stage is a complex task for the doctors.
        The main objective of this project is to analyze the parameters of various classification algorithms and compare their predictive accuracies so as to find out the best classifier for determining the liver disease.
        With a growing trend of sedentary and lack of physical activities, diseases related to liver have become a common encounter nowadays.
        In rural areas the intensity is still manageable, but in urban areas, and especially metropolitan areas the liver disease is a very common sighting nowadays.
        Liver diseases cause millions of deaths every year. Viral hepatitis alone causes 1.34 million deaths every year.
        Problems with liver patients are not easily discovered in an early stage as it will be functioning normally even when it is partially damaged.
        An early diagnosis of liver problems will increase patients survival rate.
        Liver failures are at high rate of risk among Indians.
        It is expected that by 2025 India may become the World Capital for Liver Diseases.
        The widespread occurrence of liver infection in India is contributed due to deskbound lifestyle, increased alcohol consumption and smoking.
        There are about 100 types of liver infections.
        With such alarming figures, it is necessary to have a concern towards tackling these diseases.
        Afterall, we cannot expect a developed and prosperous nation, with unhealthy youths.
        </strong>
        </p>

      |</p>
    </div>
  </div>
</body>
</html>
```

Home Page:

```
<Html>
<head>
<style>

body {
background-image: url('https://media.istockphoto.com/id/1414176792/photo/top-view-photo-of-pink-silk-ribbon-symbol-of-breast-cancer-awareness-and-stethoscope-on.jpg?s=612x612&w=0&k=20&c=7sjsU3Xu_kmxV349N-_LxWEAjYC8cBWkAFvW5lF3WfA=');
```

```
background-repeat: no-repeat;
background-attachment: fixed;
background-size: cover;
}

</style>
</head>
<body>

<title>home </title>
<h1><b>LIVER DISEASE </b></h1></head>
<body>
<h3><b>SYMTOMS:</b></h3>
<p> 1.Skin and eyes that appear yellowish (jaundice)<br>
   2.Abdominal pain and swelling<br>
   3.Swelling in the legs and ankles<br>
   4.Itchy skin<br>
   5.Dark urine color<br>
   6.Pale stool color<br>
   7.Chronic fatigue<br>
   8.Nausea or vomiting<br>
   9.Loss of appetite<br>
   10.Tendency to bruise easily<br>
</p><br>

<h3><b>TOTAL BILIRUBIN: </b></h3>
<p> This is a blood test that measures the amount of a substance called bilirubin.
   This test is used to find out how well your liver is working.
   It is often part of a panel of tests that measure liver function.
   A small amount of bilirubin in your blood is normal, but a high level may be a sign of liver disease.
</p>

<h3><b>DIRECT_BILIRUBIN:</b></h3>
<p> In the liver, bilirubin is changed into a form that your body can get rid of.
   This is called conjugated bilirubin or direct bilirubin.
   This bilirubin travels from the liver into the small intestine.
   A very small amount passes into your kidneys and is excreted in your urine.
   This bilirubin also gives urine its distinctive yellow color.
</p><br>

<h3><b>ALKALINE_PHOSPHOTASE:</b></h3>
<p>An alkaline phosphatase level test (ALP test) measures the amount of alkaline phosphatase enzyme in your bloodstream.
</p><br>

<h3><b>ALBUMIN:</b></h3>
<p> An albumin blood test is used to check your general health and to see how well your liver and kidneys are working.
```

If your liver is damaged or you're not well nourished, your liver may not make enough albumin.
</p>
</body></html>

Prediction Page:

```
<html>
<head>
<title>  LIVER PATIENT PREDICTION </title>

<h1> Prediction page: </h1>

</head>

<style>
.one{
  text-align: center;
}
body{
  background-color: gold;
}
</style>

<body>
  <div class="container">
    <h2 class='container-heading'><span class="heading_font">Liver Disease Prediction</span></h2>
</div>
<centre>
<img src="C:\Users\wecome\Documents\Liver_disease\static\images\liver.jpg"  width ="200" height="100"/>


<br>
<div class = "ml-container">
<form  action="{{url_for('prediction')}}"  method="POST" >
<br>
<br>
<label> Age: </label>
<input id="first" name="age" required="require" size="20"/> <br>
<br>
<label for="Gender" required="require">Choose Gender:</label>
<select id="Gender" name="Gender">
  <option value="0">Female</option>
```

```
  <option value="1">Male</option>
</select><br>
<br>
<label> Total_Bilirubin: </label>
<input id="second" name="totalbilirubin" required="require" size="20"/> <br>
<br>

<label> Direct_Bilirubin: </label>
<input id="third" name="directbilirubin" required="require" size="20"/> <br>
<br>

<label> Alkaline_Phosphotase: </label>
<input id="fourth" name="alkalinephosphotase" required="require" size="20"/>
<br> <br>

<label> Alamine_Aminotransferase: </label>
<input id="fifth" name="alamine" required="require" size="20"/> <br> <br>

<label> Aspartate_Aminotransferase: </label>
<input id="sixth" name="aspartate" required="require" size="20"/> <br> <br>

<label> Total_Proteins: </label>
<input id="seventh" name="totalproteins" required="require" size="20"/> <br>
<br>

<label> Albumin: </label>
<input id="eigth" name="albumin" required="require" size="20"/> <br> <br>

<label> Albumin_and_Globulin_Ratio: </label>
<input id="ninth" name="albuminandglobulin" required="require" size="20"/>
<br> <br>

<button id="sub" type="submit">Predict</button>


</form>
</div>
</centre>
<style>
body{
        font-family: Arial, Helvetica,sans-serif;
    text-align: center;
    margin: 0;
    padding: 0;
```

```
    width: 100%;
        height: 100%;
        display: flex;
        flex-direction: column;
}
.container-heading{
  margin:0;
}
.heading_font{
  color: black;
  font-family: 'Pacifico', cursive;
  font-size: 50px;
  font-weight: normal;
}
  #first{
    border-radius: 14px;
    height: 25px;
    width: 150px;
    font-size: 20px;
    text-align: center;
  }
  #second{
    border-radius: 14px;
    height: 25px;
    width: 150px;
    font-size: 20px;
    text-align: center;
  }
  #third{
    border-radius: 14px;
    height: 25px;
    width: 150px;
    font-size: 20px;
    text-align: center;
  }
  #fourth{
    border-radius: 14px;
    height: 25px;
    width: 150px;
    font-size: 20px;
    text-align: center;
  }
  #fifth{
```

```
    border-radius: 14px;
    height: 25px;
    width: 150px;
    font-size: 20px;
    text-align: center;
   }
   #sixth{
    border-radius: 14px;
    height: 25px;
    width: 150px;
    font-size: 20px;
    text-align: center;
   }
   #seventh{
    border-radius: 14px;
    height: 25px;
    width: 150px;
    font-size: 20px;
    text-align: center;
   }
   #eigth{
    border-radius: 14px;
    height: 25px;
    width: 150px;
    font-size: 20px;
    text-align: center;
   }
   #ninth{
    border-radius: 14px;
    height: 25px;
    width: 150px;
    font-size: 20px;
    text-align: center;
   }

   #sub{
    width: 120px;
    height: 43px;
    text-align: center;
    border-radius: 14px;
    font-size: 18px;
   }
  </style>
```

```
</body>
</html>
```

GITHUB & PROJECT DEMO LINK:

GITHUB LINK:
https://github.com/IBM-EPBL/IBM-Project-4364-1658729878

DEMO LINK:
https://drive.google.com/drive/folders/1mTfBtW9atpdqcuwTFDafiNzuE2i4vgGl?usp=share_link