Project On

STATISTICAL MACHINE LEARNING APPROACHES TO LIVER DISEASE PREDICTION

powered By IBM India

Submitted By:

VINISHA S
SUSMA T
SHREEDH AR T
VIJAY M
SYED ASIF Z

Project ID: PNT2022TMID18834

of

Department of INFORMATION TECHNOLOGY
SONA COLLEGE OF TECHNOLOGY
SALEM - 636500

College Mentor: **RAJKUMAR S**

Industrial Mentor: **NIDHI**

INDEX

1. INTRODUCTION

1.1 Project Overview

1.2 Purpose

2 LITERATURE SURVEY

2.1 Existing problem

2.2 References

2.3 Problem Statement Definition

3 IDEATION & PROPOSED SOLUTION

3.1 Empathy Map Canvas

3.2 Ideation & Brainstorming

3.3 Proposed Solution

3.4 Problem Solution fit

4 REQUIREMENT ANALYSIS

4.1 Functional requirement

4.2 Non-Functional requirements

5 PROJECT DESIGN

5.1 Data Flow Diagrams

5.2 Solution & Technical Architecture

5.3 User Stories

6 PROJECT PLANNING & SCHEDULING

6.1 Sprint Planning & Estimation

6.2 Sprint Delivery Schedule

6.3 Reports from JIRA

7 CODING & SOLUTIONING

7.1 Feature 1

7.2 Feature 2

7.3 Database Schema (if Applicable)

8 TESTING

8.1 Test Cases

8.2 User Acceptance Testing

9 RESULTS

9.1 Performance Metrics

10 ADVANTAGES & DISADVANTAGES

11 CONCLUSION

12 FUTURE SCOPE

13 APPENDIX

1. INTRODUCTION

1.1. Project Overview

It is essential to diagnose liver disease early in order to save lives and take the necessary steps to control the condition. The ensemble approach has been successfully used in a number of sectors, particularly in the field of medical science. This study examines the early diagnosis of liver illness using a variety of ensemble methodologies.

1.2. Purpose

Our project's objective is to analyse data from liver patients with a focus on the correlationsbetween a vital list of liver enzymes, proteins, age and gender using them to try and predict the likeliness of liver disease.We are creating a model in this instance by utilising different machine learning algorithms to identify the most accurate model. and incorporate into web applications built with flask. By entering certain parameters into the web application, users can predict the disease.

2. LITERATURE SURVEY

2.1 Existing problem

TITLE: " A Comparative Analysis of Unsupervised Machine Techniques for Liver Disease Prediction"

AUTHOR: Varun Vats

YEAR: 2018

He considered three different ML (Machine Learning) algorithms. A comparison of these algorithms had been carried out for evaluating their forecasting accuracy and computing intricacy . These algorithms included AP (Affinity Propagation), K means and DBSCAN. This work was dedicated to the medical dataset based on lever disorders. This work made use of the Silhouette coefficient to measure the comparative efficiency of the considered algorithmic approaches.

TITLE : “Prognosis of Liver Disease using Machine Learning Algorithms”
AUTHOR: Vyshali J Gogi
YEAR : 2018
He stated that the healthcare sector had a lot of data but this data was of no use [17]. This ample data required a leading analytic tool so that the hidden relationship and the valuable knowledge could be determined. The liver disease referred to the medical condition of the human liver-related to the human liver. The liver diseases led to sudden changes in health conditions that governed the functioning of the liver affecting other internal body organs. This work made use of several classification algorithms based on data mining. These algorithms included DT (Decision Tree), LD (Linear Discriminant), SVM Fine Gaussian, and LR (Logistic Regression). This work made use of Lab-based metrics of patients in the form of a liver dataset.

TITLE: “Accuracy Prediction Using Machine Learning Techniques for Indian Patient Liver Disease”
AUTHOR: Auxilia
YEAR: 2018
She stated that the use of medical datasets had attracted the medical experts globally [18]. The use of ML (Machine Learning) algorithms was quite common as a branch of making selection expressively helpful networks for the prediction of diseases by arranging therapy-based datasets. Grouping schemes had been generally employed as a segment of the curative domain for extracting order more efficiently as compared to a signal classification model. The disorders of the Liver malady could be described as liver damage or sickness. Liver disorder can be categorized into several categories. This work made use of standard Indian liver illness patient records as a database for providing support to the researcher

TITLE: “New Fuzzy-ANWKNN algorithm for the successful prediction of liver disorder”
AUTHOR: Pushpendra Kumar
YEAR: 2019
He stated that it was a very difficult task for the doctors to detect the consequences of liver disorders on a person . In general, researchers used datasets based on LFT (Liver Function Test) for implementing classification algorithms so that the predictions about liver disorders could be

generated. The dataset based on ground truth had several problems such as a class imbalance in the liver disorder data. With regard to the majority classes, the classic algorithms of classification generated influenced outcomes. This work presented a new Fuzzy-ANWKNN algorithm for the successful prediction of liver disorder.

TITLE: “To accurately predict liver disorder by means of several data mining algorithms”
AUTHOR: Sanjay Kumar
YEAR: 2018
He described different classification approaches by implementing them on the dataset of patients suffering from liver diseases . The main objective here was to accurately predict liver disorder by means of several data mining algorithms. This work performed the analysis using the dataset of real-time patients to build classification paradigms for the prediction of liver diseases. This work implemented five classification algorithms on the used dataset. This work analyzed different metrics such as precision, recall, and accuracy for determining the efficiency of the implemented classification models.

TITLE: ” Liver disease prediction by using different decision tree techniques”
AUTHOR: Nazmun Nahar and Ferdous Ara
YEAR: 2018
They stated that their research work explores the early prediction of liver disease using various decision tree techniques. The liver disease dataset which is select for this study is consisting of attributes like total bilirubin, direct bilirubin, age, gender, total proteins, albumin, and globulin ratio. The main purpose of this work is to calculate the performance of various decision tree techniques and compare their performance. The decision tree techniques used in this study are J48, LMT, Random Forest, Random tree, REPTree, Decision Stump, and Hoeffding Tree. The analysis proves that Decision Stump provides the highest accuracy than other techniques.

TITLE: “Prediction of liver fibrosis stages by machine learning model: A decision tree approach”
AUTHOR: Heba Ayeldeen
YEAR: 2015
This work made by using an ML algorithm based on DT for predicting the level of liver fibrosis in every patient . The outcomes revealed that DT (Decision Tree) classifier achieved a classification

accuracy of 93.7%. This accuracy rate was higher than the accuracy rate reported by other investigations in the same conditions.

TITLE: Supervised classification and prediction of fibrosis seriousness using ultrasonic images
AUTHOR: C. A. Prajith
YEAR: 2016
He described the growth of scar tissue due to inflammation, infection, or injury so called liver fibrosis.This disease could be the reason for liver cirrhosis. The use of various non-invasive imaging techniques was quite common for the treatment of liver fibrosis. These techniques included MRI, CT, Electrography, and ultrasound. This study was focused on the extraction of texture features from liver images of ultrasound. This work implemented various classification models such as ANN, GMM, and SVM for classifying the risk level of the liver fibrosis. SVM has a specificity of 95 % , the sensitivity of 93.33%, and an accuracy of 94 %.

2.2 PROBLEM STATEMENT DEFINITION

Liver diseases avert the normal function of the liver. It is mainly caused due to the large amount of alcohol consumption. Early prediction of liver disease using classification algorithm can help the doctors to diagnose the existence of liver disease at an early stage which 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. This project examines the data from liver patients concentrating on relationships between them to predict the likeliness of liver disease and building a model to find the best accurate model and integrate to flask based web application using which the user can predict the liver disease by entering their parameters.

2.3.REFERENCES

Biomarkers for prediction of liver fibrosis in patients with chronic alcoholic liver disease written by Sylvie Naveau and Bruno Runyard.
Strategic analysis in prediction of liver disease using classification algorithms written by Piyush Kr Shukla andBinish Khan.

Software based prediction of liver disease with feature selection and classification techniques witten Jagdeep Singh, Sandeep Bagga and Ranjodh Kaur.

Liver disease prediction using SVM and NaÃ¯ve Bayes algorithm written by S Dhayanand.

Prediction and analysis of liver diseases using data mining written Shambel Kefelgen, Pooj Kamat

1.H. Jin , S. Kim and J.Kim , Decision Factors on Effective Liver Patient Data Prediction, International

Journal of Bio-Science and Bio-Technology

2. Reetu and N.Kumar , Medical Diagnosis for Liver Cancer Using Classification Techniques, International Journal of Recent Scientific. Volume 6. Issue, 6, pp 4809-4813.

3. Sivakumar D , Manjunath Varchagall , and Ambika L Gusha S “Chronic Liver Disease Prediction Analysis Based on the Impact of Life Quality Attributes.” (2019). International Journal of Recent Technology and Engineering (IJRTE) ISSN: 2277-3878, Volume- 7, Issue-6S5, April 2019

4. Sindhuja, D., and R. Jemina Priyadarsini. "A survey on classification techniques in data mining analyzing liver disease disorder." International Journal of Computer Science and Mobile Computing 5.5 :

483-488.

5.Sontakke, Sumedh et al. “Diagnosis of liver diseases using machine learning.” 2017 International Conference on Emerging Trends & Innovation in ICT (ICEI) (2017): 129-133.

Data Mining Techniques for optimization of liver disease classification.” Sadiyah Noor Novit Alfisahrin, Teddy Mantoro Electron-ic ISBN: 978-1-4799-2758-6 DOI: 10.1109/ACSAT.2013.81-IEEE

6. S. A. Gonzalez dan E. B. Keeffe, “Acute liver failure,” dalam Handbook of Liver Disease Third Edition, Philadelphia, Elsevier, 2012, pp. 20-33

7. M. Hassoon, M. S. Kouhi, M. Zomorodi-Moghadam and M. Abdar, "Rule Optimization of Boosted C5.0 Classification Using Genetic Algorithm for Liver disease Prediction," 2017 International Conference on Computer and Applications (ICCA), Doha, 2017, pp. 299-305. doi: 10.1109/COMAPP.2017.8079783

8. D.A. Saleh F. Shebl M. Abdel-Hamid et al. "Incidence and risk

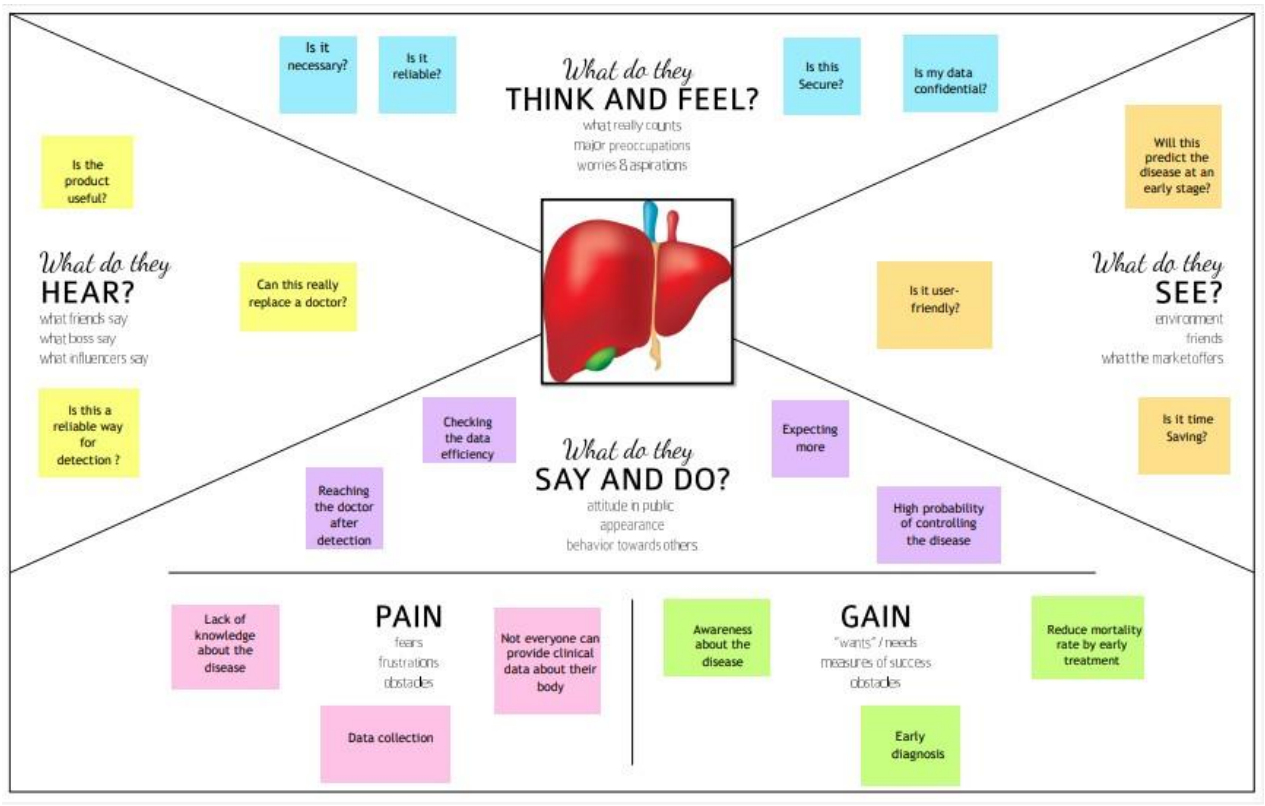
factors for hepatitis C infection in a cohort of women in rural Egypt"Trans. R. Soc. Trop. Med. Hyg. vol. 102 pp. 921928 2008. https://doi.org/10.1016/j.trstmh.2008.04.011

9. Performance analysis of classification algorithms on early detetion of Liver disease “Moloud Abdar, Mariam Zomorodi- Moghadam, Resul Das, I-Hsien Ting, doi: 10.1016/ j.eswa. 2016.08.065

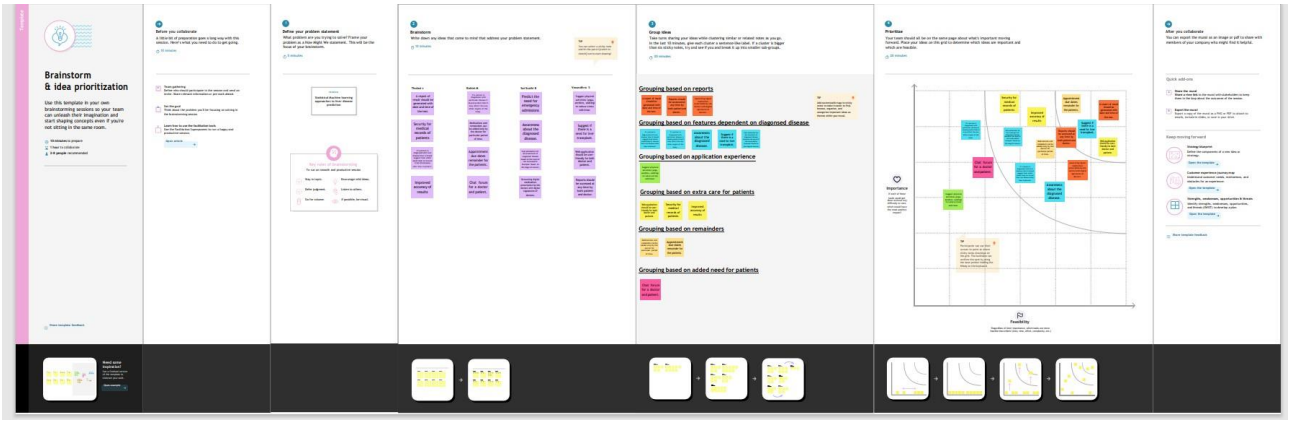
10. A Critical Study of Selected Classification Algorithms for Liver Disease Diagnosis “ Bendi Venkata Ramana1, Prof. M.Surendra Prasad Babu2, Prof. N. B. Venkateswarlu3 -(IJDMS), Vol.3, No.2, May 2012

3. IDEATION & PROPOSED SOLUTION

Empathy Map



3.2 Ideation and Brainstorming

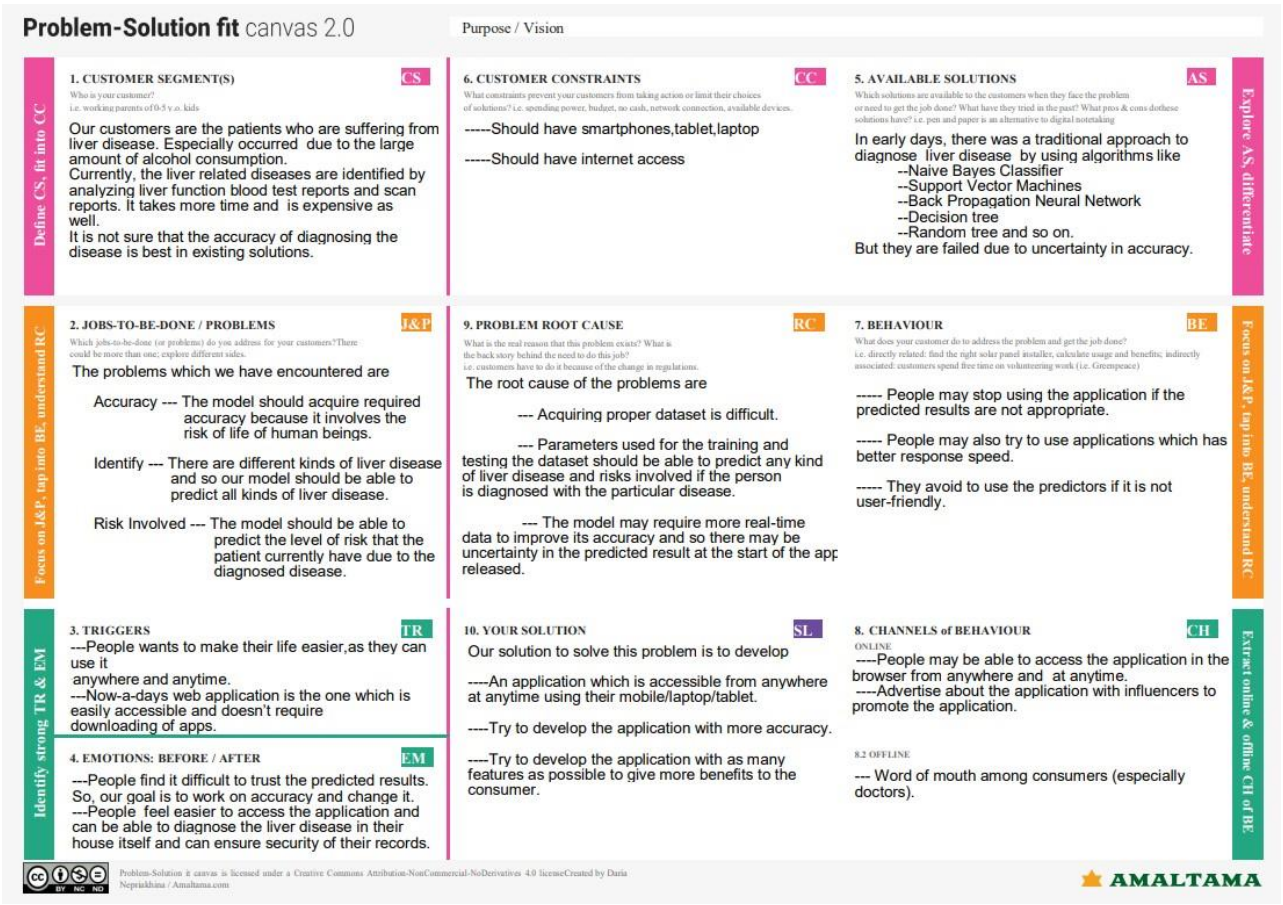


3.3 Proposed Solution

S.No	Parameter	Description
1.	Problem Statement (Problem to be solved)	Discovering the existence of liver diseases at early stage is a complex task for doctors. The challenge is to predict the liver disease patient fast and accurate and to diagnose the patients in early stage .
2.	Idea / Solution description	Machine learning model which uses statistical data to predict the liver disease of the patients.
3.	Novelty / Uniqueness	Accurately classifies the intensity of the liver disease from the patients concentrating on relationship between a key list of enzymes, proteins, age and gender using them to predict the likeliness of the liver disease
4.	Social Impact / Customer Satisfaction	• Capable of predicting the liver disease in early stage • Works accurately and precisely to predict the liver disease • Doctors can be able to diagnose the live patients in early stage to save many lives
5.	Business Model (Revenue Model)	• This system can be integrated with any Health sector domain, It solves the complex process of predicting the liver disease of patients and makes ease to

		the doctors to diagnose the liver disease. • The user can be able to get consulting with doctors
6.	Scalability of the Solution	• Can be extended to predict many classification of diseases in early stage • This can be integrated to with any hospitals and health sectors to get patient records securely through APIs

3.4 Proposed Solution Fit



4. REQUIREMENT ANALYSIS

4.1 FUNCTIONAL REQUIREMENTS

Functional Requirements:

Following are the functional requirements of the proposed solution.

FR No.	Functional Requirement (Epic)	Sub Requirement (Story / Sub-Task)
FR-1	User Registration	The user can register the application by entering their E-mail id, Password, and Confirming the Password
FR-2	User Confirmation	The system gives an approval message to their respective mail id after the user account is activated.
FR-3	Website Entry &Data Management	Collecting user data and storing it in the Database will be used for future reference
FR-4	Predict the Disease	The user can easily predict the disease easily just by entering the Parameter data
FR-5	Result	The user can view their result immediately after predicting the disease
FR-6	Good Network Connection	The User should have a stable internet connection to accessthe functionality of our project and view their results via a web application.

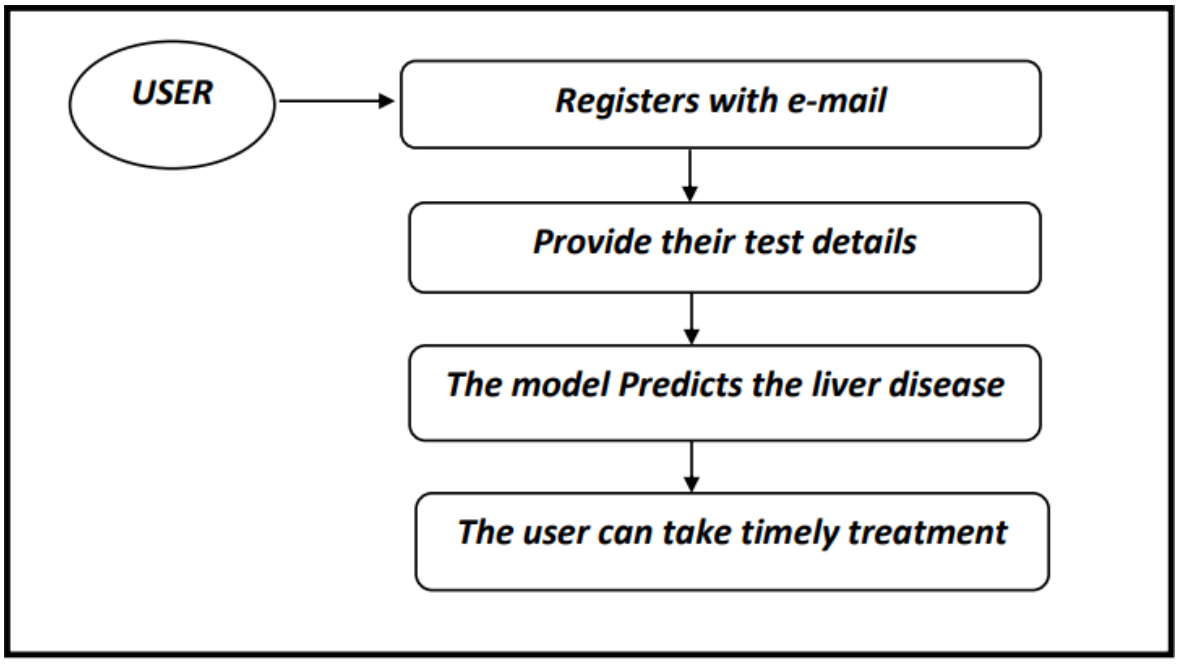
4.2 NON FUNCTIONAL REQUIREMENTS

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

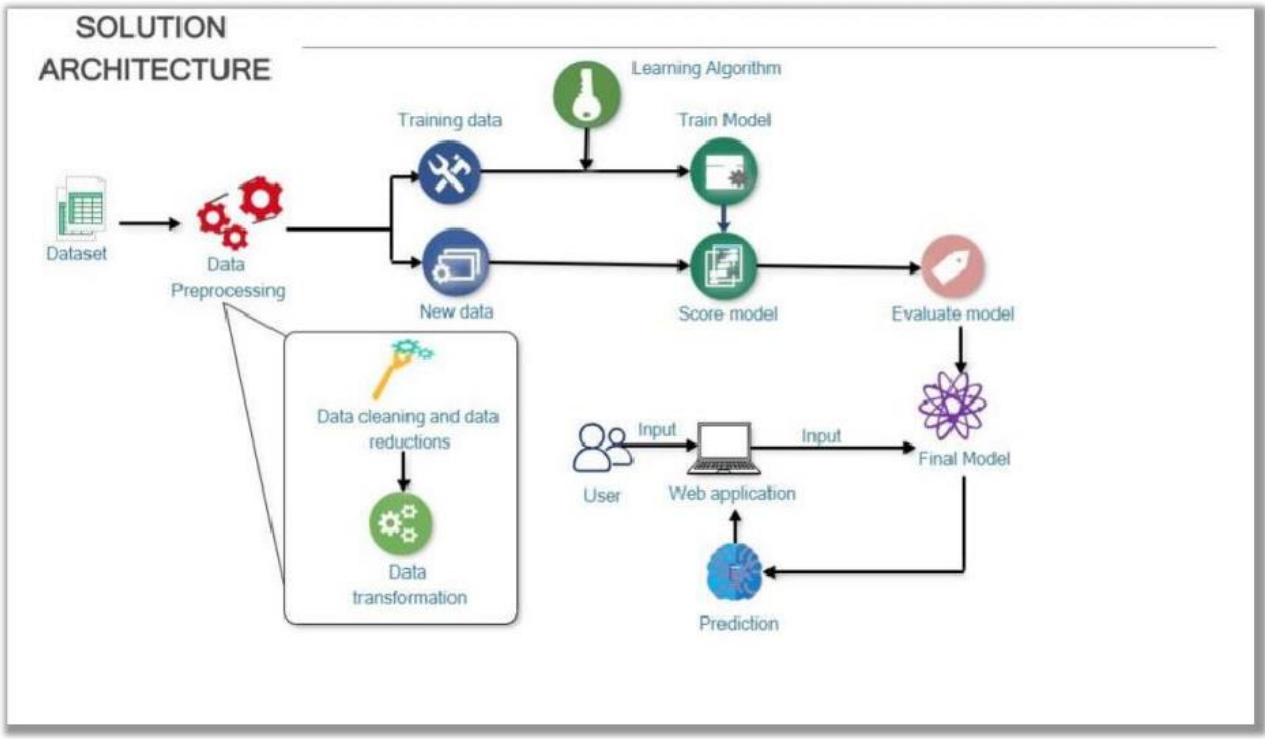
FR No.	Non-Functional Requirement	Description
NFR-1	**User Friendly**	The entire system can be accessed through a web Application at any time and use the platform whenever they want. It is very easy to use.
NFR-2	**Safe and Secure**	The user data will be stored in a database so only the authorized person can able to access it so the user's data is safe and secured.
NFR-3	**Reliability**	It is a highly reliable platform other than authorized person no can access it.
NFR-4	**Performance**	It is a quality attribute that describes the responsiveness of the system to the various user interactions with it.
NFR-5	**Availability**	The application is available all the time so the user can predict liver disease at anytime
NFR-6	**Scalability**	It is more scalable because the application can be used in any kind of operating system either in the small or large

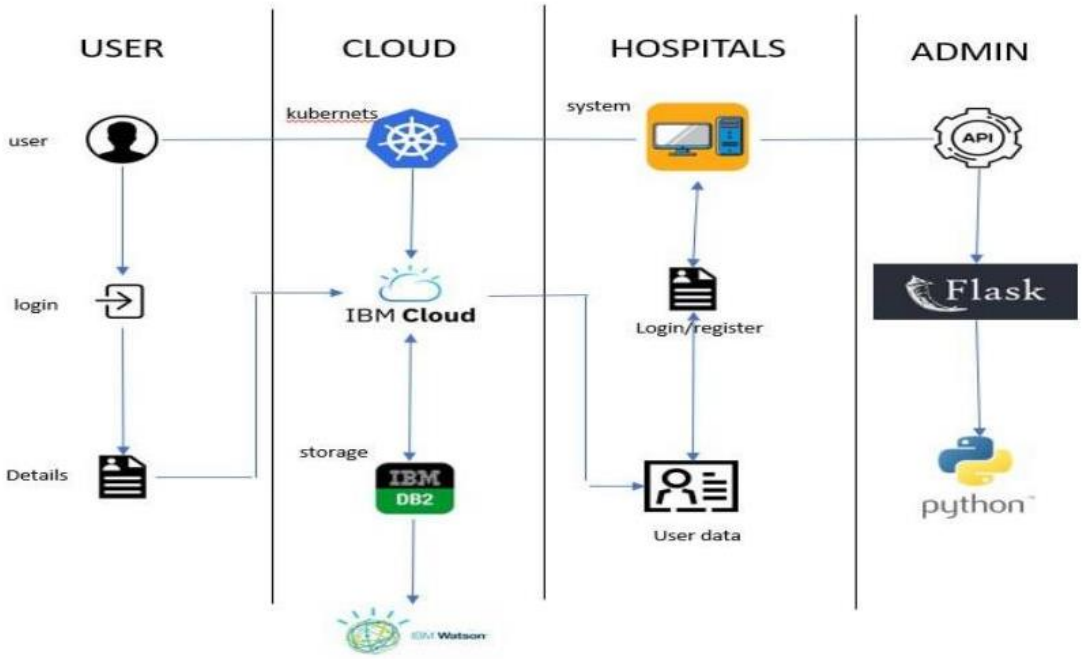
5. PROJECT DESIGN

5.1 DATA FLOW DIAGRAM



5.2 SOLUTION AND TECHNICAL ARCHITECTURE


SOLUTION ARCHITECTURE
Learning Algorithm
Training data
Train Model
Dataset
Data Preprocessing
New data
Score model
Evaluate model
Data cleaning and data reductions
Data transformation
Input
Input
User
Web application
Final Model
Prediction


USER
CLOUD
HOSPITALS
ADMIN
user
kubernets
system
API
login
IBM Cloud
Login/register
Flask
Details
storage
IBM DB2
User data
python
IBM Watson

5.3 USER STORIES

Use the below template to list all the ser stories for product

User Type	Functional Requirement (Epic)	User Story Number	User Story / Task	Acceptance criteria	Priority	Release
Customer (Web User)	Registration	USN-1	As a user, I can register for the application by entering my email,a password & confirmation password	I can access my account dashboard	High	Sprint-1
		USN-2	As a user, I will get a confirmation email once I have registered	I can receive a confirmation mail	High	Sprint-1
	Login	USN-3	As a user, I can login for the application through email & password	I can login and access the application	Medium	Sprint-2
	Dashboard	USN-4	As a user, I can access the dashboard	I can give personal information	High	Sprint-1
	Prediction of Result	USN-6	As a user, I can predict disease using web application	I can get a result & try to get early treatment	High	Sprint-1

6. PROJECT PLANNING & SCHEDULING

6.1. SPRINT PLANNING AND ESTIMATION

Sprint	Functional Requirement (Epic)	User Story Number	User Story / Task	Story Points	Priority	Team Members
Sprint-1	Registration	USN-1	As a user, I can register for the application by entering my email, password, and confirming my password.	5	High	Vinisha S
Sprint-1		USN-2	As a user, I will receive confirmation email Once I have registered for the application	5	High	Susma T
Sprint-1	Login	USN-3	As a user, I can log into the application by entering email & password	10	High	Vijay M
Sprint-2	Input Necessary Details	USN-4	As a user, I can give Input Details to Predict Likeliness of Liver Disease.	15	High	Shreedhar T
Sprint-2	Data pre-processing	USN-5	Transform raw data into suitable format for prediction.	5	High	Vinisha S
Sprint-3	Prediction of Liver Disease	USN-6	As a user, I can predict Liver Disease Using machine learning model.	15	High	Syed Asif Z
Sprint-3	.	USN-8	As a user, I can get accurate prediction of Liver disease.	5	Medium	Susma T
Sprint-4	Review	UNS-8	As a user, I can give feedback of The application	20	High	Vinisha S

6.2. SPRINT DELIVERY SCHEDULE

Sprint	Total Story Points	Duration	Sprint Start Date	Sprint End Date (Planned)	Story Points Completed (as on Planned End Date)	Sprint Release Date (Actual)
Sprint-1	10	5 Days	21 Oct 2022	31 Oct 2022	10	31 Oct 2022
Sprint-2	15	10 Days	1 NOV 2022	05 Nov 2022	15	05 Nov 2022
Sprint-3	15	10 Days	07 Nov 2022	12 Nov 2022	15	12 Nov 2022
Sprint-4	10	7 Days	14 Nov 2022	19 Nov 2022	10	19 Nov 2022

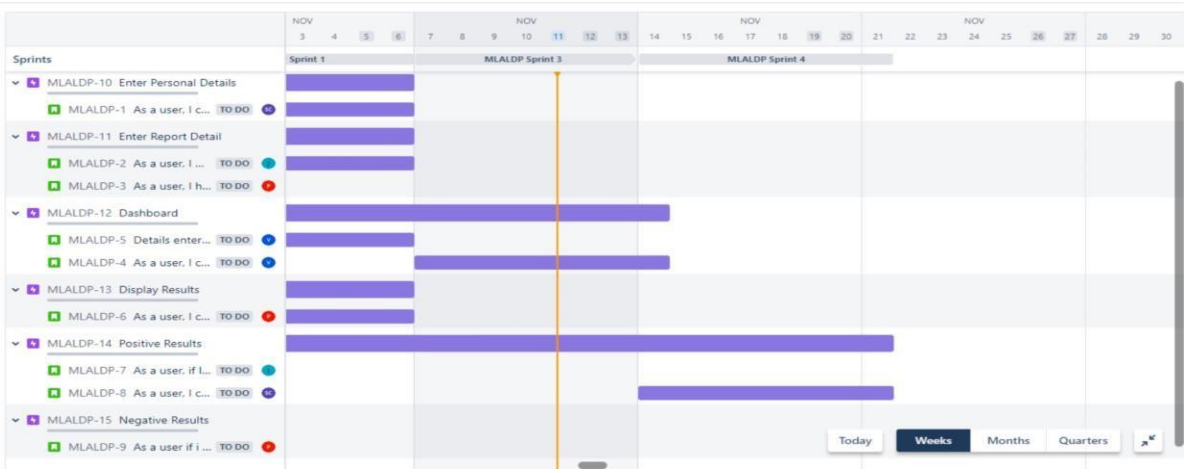
Velocity:

Imagine we have a 10-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)

$$AV = \frac{sprint\ duration}{velocity}$$

AV= 6 / 20 = 0.3

6.3. Reports from JIRA


NOV
3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30
Sprints
Sprint 1
MLALDP Sprint 3
MLALDP Sprint 4
MLALDP-10 Enter Personal Details
MLALDP-1 As a user, I c... TO DO
MLALDP-11 Enter Report Detail
MLALDP-2 As a user, I ... TO DO
MLALDP-3 As a user, I h... TO DO
MLALDP-12 Dashboard
MLALDP-5 Details enter... TO DO
MLALDP-4 As a user, I c... TO DO
MLALDP-13 Display Results
MLALDP-6 As a user, I c... TO DO
MLALDP-14 Positive Results
MLALDP-7 As a user, if I... TO DO
MLALDP-8 As a user, I c... TO DO
MLALDP-15 Negative Results
MLALDP-9 As a user if i ... TO DO
Today
Weeks
Months
Quarters

7. CODING & SOLUTIONING

7.1 FEATURE 1

```
# Importing Libraries:
import pandas as pd
import numpy as np
import seaborn as sns
import matplotlib.pyplot as plt
```

```
# for displaying all feature from dataset:
pd.pandas.set_option('display.max_columns', None)
```

```
# Reading Dataset:
dataset = pd.read_csv("Dataset/Liver_data.csv")
# Top 5 records:
dataset.head()
```

```
# Last 5 records:
dataset.tail()
```

```
# Shape of dataset:
dataset.shape
```

```
(583, 11)
```

```
# Cheaking Missing (NaN) Values:
dataset.isnull().sum()
```

```
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
```

- 'Albumin_and_Globulin_Ratio' feature contain 4 NaN values.

```
# Mean & Median of "Albumin_and_Globulin_Ratio" feature:
print(dataset['Albumin_and_Globulin_Ratio'].median())
print(dataset['Albumin_and_Globulin_Ratio'].mean())
```

```
0.93
0.9470639032815201
```

```
# Filling NaN Values of "Albumin_and_Globulin_Ratio" feature with Median :
dataset['Albumin_and_Globulin_Ratio'] = dataset['Albumin_and_Globulin_Ratio'].fillna(dataset['Albumin_and_Globulin_Ratio'].median())
```

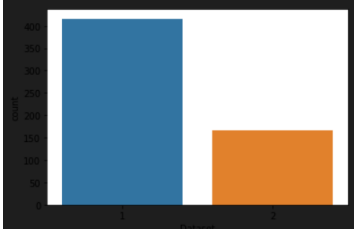
```
# Datatypes:
dataset.dtypes
```

```
Age                           int64
Gender                       object
Total_Bilirubin             float64
Direct_Bilirubin            float64
Alkaline_Phosphotase          int64
Alamine_Aminotransferase      int64
Aspartate_Aminotransferase    int64
Total_Protiens              float64
Albumin                     float64
Albumin_and_Globulin_Ratio  float64
Dataset                       int64
dtype: object
```

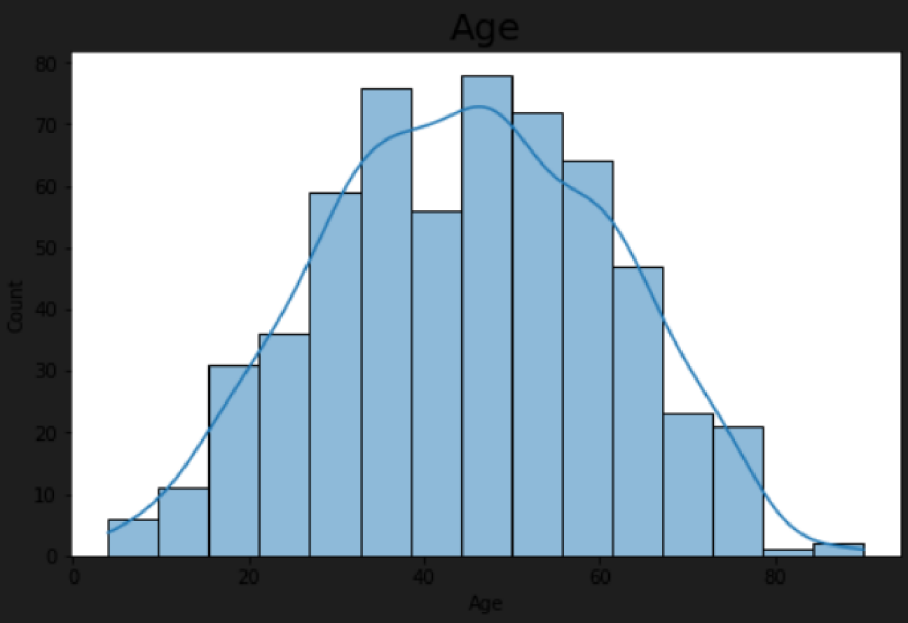
```
# Description:
dataset.describe()
```

```
# Target feature:
print("Liver Disease Patients      :", dataset['Dataset'].value_counts()[1])
print("Non Liver Disease Patients  :", dataset['Dataset'].value_counts()[2])

# Visualization:
sns.countplot(dataset['Dataset'])
plt.show()
```

```
Liver Disease Patients      : 416
Non Liver Disease Patients  : 167

C:\Users\LENOVO\Anaconda3\lib\site-packages\seaborn\_decorators.py:36: FutureWarning: Pass the following variable as a keyword arg: x. From version 0.12, the only valid positional
argument will be `data`, and passing other arguments without an explicit keyword will result in an error or misinterpretation.
  warnings.warn(
```



```python
# Histrogram of Age:
plt.figure(figsize=(8,5))
sns.histplot(dataset['Age'], kde=True)
plt.title('Age', fontsize=20)
plt.show()
```

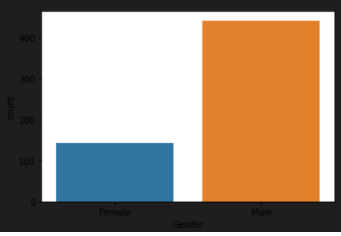



```python
dataset.head()
```

```python
# Gender feature:
print("Total Male   :", dataset['Gender'].value_counts()[0])
print("Total Female :", dataset['Gender'].value_counts()[1])

# Visualization:
sns.countplot(dataset['Gender'])
plt.show()
```

Python

```
Total Male   : 441
Total Female : 142

C:\Users\LENOVO\Anaconda3\lib\site-packages\seaborn\_decorators.py:36: FutureWarning: Pass the following variable as a keyword arg: x. From version 0.12, the only valid positional
argument will be `data`, and passing other arguments without an explicit keyword will result in an error or misinterpretation.
  warnings.warn(
```



```
# Printing How many Unique values present in each feature:
for feature in dataset.columns:
    print(feature,":", len(dataset[feature].unique()))
```

```
Age : 72
Gender : 2
Total_Bilirubin : 113
Direct_Bilirubin : 80
Alkaline_Phosphotase : 263
Alamine_Aminotransferase : 152
Aspartate_Aminotransferase : 177
Total_Protiens : 58
Albumin : 40
Albumin_and_Globulin_Ratio : 69
Dataset : 2
```

```
# Label Encoding
dataset['Gender'] = np.where(dataset['Gender']=='Male', 1,0)
```

```
dataset.head()
```

```python
# Correlation using Heatmap:
plt.figure(figsize=(12,8))
sns.heatmap(dataset.corr(), annot=True, cmap='YlGnBu')
plt.show()
```

There is Multi-Collinearity found on our dataset.

```python
dataset.columns
```

Python

```
Index(['Age', 'Gender', 'Total_Bilirubin', 'Direct_Bilirubin',
       'Alkaline_Phosphotase', 'Alamine_Aminotransferase',
       'Aspartate_Aminotransferase', 'Total_Protiens', 'Albumin',
       'Albumin_and_Globulin_Ratio', 'Dataset'],
      dtype='object')
```

1. Multicollinearity betwwen **'Total_Bilirubin'** and **'Direct_Bilirubin'** is **0.87%**
2. Multicollinearity betwwen **'Alamine_Aminotransferase'** and **'Aspartate_Aminotransferase' **is **0.79%**
3. Multicollinearity betwwen **'Total_Protiens'** and **'Albumin'** is **0.78%**
4. Multicollinearity betwwen **'Albumin'** and **'Albumin_and_Globulin_Ratio'** is **0.69%**

Usually we drop that feature which has above 0.85% multicollinearity between two independent feature. Here we have only 'Total_Bilirubin' and 'Direct_Bilirubin' feature which has 0.87% mutlicollinearity. So we drop one of the feature from them and other independent feature has less multicollinearity, less than 0.80% So we keep that feature.

```python
# Droping 'Direct_Bilirubin' feature:
dataset = dataset.drop('Direct_Bilirubin', axis=1)
```

Python

```python
dataset.columns
```

Python

```python
# Droping 'Direct_Bilirubin' feature:
dataset = dataset.drop('Direct_Bilirubin', axis=1)
```

Python

```python
dataset.columns
```

Python

```
dex(['Age', 'Gender', 'Total_Bilirubin', 'Alkaline_Phosphotase',
     'Alamine_Aminotransferase', 'Aspartate_Aminotransferase',
     'Total_Protiens', 'Albumin', 'Albumin_and_Globulin_Ratio', 'Dataset'],
    dtype='object')
```

```python
sns.distplot(dataset['Albumin'])
```

Python

```
\Users\LENOVO\Anaconda3\lib\site-packages\seaborn\distributions.py:2557: FutureWarning: `distplot` is a deprecated function and will be removed in a future version. Please adapt
ur code to use either `displot` (a figure-level function with similar flexibility) or `histplot` (an axes-level function for histograms).
warnings.warn(msg, FutureWarning)

xesSubplot:xlabel='Albumin', ylabel='Density'>
```

7.2 FEATURE 2:

```html
<!DOCTYPE html>
<html lang="en">
<head>
    <meta charset="UTF-8">
    <title>Liver Prediction Model</title>
</head>
<body>
    <div class="container">
        <h2 class='container-heading'><span class="heading_font">Liver Disease Prediction</span></h2>
    </div>
    <div class="ml-container">
        <form action="{{ url_for('predict') }}" method="POST">
            <br>
        <center>
            <table> <tr>
            <td><h3>Age</h3></td>
            <td><input id="first" name="Age" placeholder=" " required="required"></td>
            </tr><tr>
            <td><h3>Gender</h3></td>
            <td> <select>  <option value="Male" id="second">Male</option>   <option value="female" id="second">Female</opti
            </td>  </tr>  <br>   <tr>
            <td><h3>Total Bilirubin</h3></td>
            <td><input id="third" name="Total_Bilirubin" placeholder=" " required="required"></td>
            </tr>  <tr>
           <td><h3>Direct Bilirubin</h3></td>
            <td><input id="third" name="Total_Bilirubin" placeholder=" " required="required"></td> </tr>
            <br> <tr>
            <td><h3>Alkaline Phosphotase</h3></td>
            <td><input id="fourth" name="Alkaline_Phosphotase" placeholder=" " required="required"></td>
            </tr> <br> <tr>
            <td><h3>Alamine Aminotransferase</h3></td>
            <td><input id="fifth" name="Alamine_Aminotransferase" placeholder=" " required="required"></td>
```

```
            <tr>
            <td><h3>Aspartate Aminotransferase</h3></td>
            <td><input id="sixth" name="Aspartate_Aminotransferase" placeholder=" " required="required"></td>   </tr>    <b
            <td><h3>Total Protiens</h3></td>
            <td><input id="seventh" name="Total_Protiens" placeholder=" " required="required"></td> </tr>
            <br> <tr>
            <td><h3>Albumin</h3></td>
            <td><input id="eight" name="Albumin" placeholder=" " required="required"></td>
            </tr>  <br>  <tr>
            <td><h3>Albumin and Globulin Ratio</h3></td>
            <td><input id="ninth" name="Albumin_and_Globulin_Ratio" placeholder=" " required="required"></td>
            </tr>  <br>  <br>  <br>  <tr>
            <td><button id="sub" type="submit ">Predict</button></td>
            </tr>
        </table>   </center>   </form>  </div>
<style>
/* Background Image */
body
{
background-color: blueviolet;
height: 100%;

/* Center and scale the image nicely */
background-position: center;
background-repeat: no-repeat;
background-size: 100% 100%;

}

/* Color */
body{
    font-family:'Lucida Sans', 'Lucida Sans Regular', 'Lucida Grande', 'Lucida Sans Unicode', Geneva, Verdana, sans-serif;
```

```
<!DOCTYPE html>
<html lang="en">
<head>
    <meta charset="UTF-8">
    <meta name="viewport" content="width=device-width, initial-scale=1.0">
    <title>Liver Disease Result</title> </head>
<body>  <div class="container">
        <form action="{{ url_for('predict')}}" method="post">
            <h2 class='container-heading'><span class="heading_font">Liver Disease Prediction</span></h2>
        <br><br><br><br><br><br><br>
        <!-- Result -->
        <div class="results">
            {% if prediction==2 %}
                <h1><span class='danger'>Oops! 😟<br><br>You have LIVER DISEASE <br><br>Please Consult a Doctor.</span></h
            {% elif prediction==1 %}
                <h1><span class='safe'>🤩 Congratulation! 🤩<br><br>You DON'T have LIVER DISEASE.</span></h1>       </div
        </form></div> <div>   <br><br> <br><br><br><br><br><br><br><br>
    </div><style>
/* Background Image */
body
{
background-color: blueviolet;
height: 100%;
/* Center and scale the image nicely */
background-position: center;
background-repeat: no-repeat;
background-size: 100% 100%;
}

/* Color */
body{
    font-family:'Lucida Sans', 'Lucida Sans Regular', 'Lucida Grande', 'Lucida Sans Unicode', Geneva, Verdana, sans-serif;
```

8. TESTING

8.1. TEST CASES

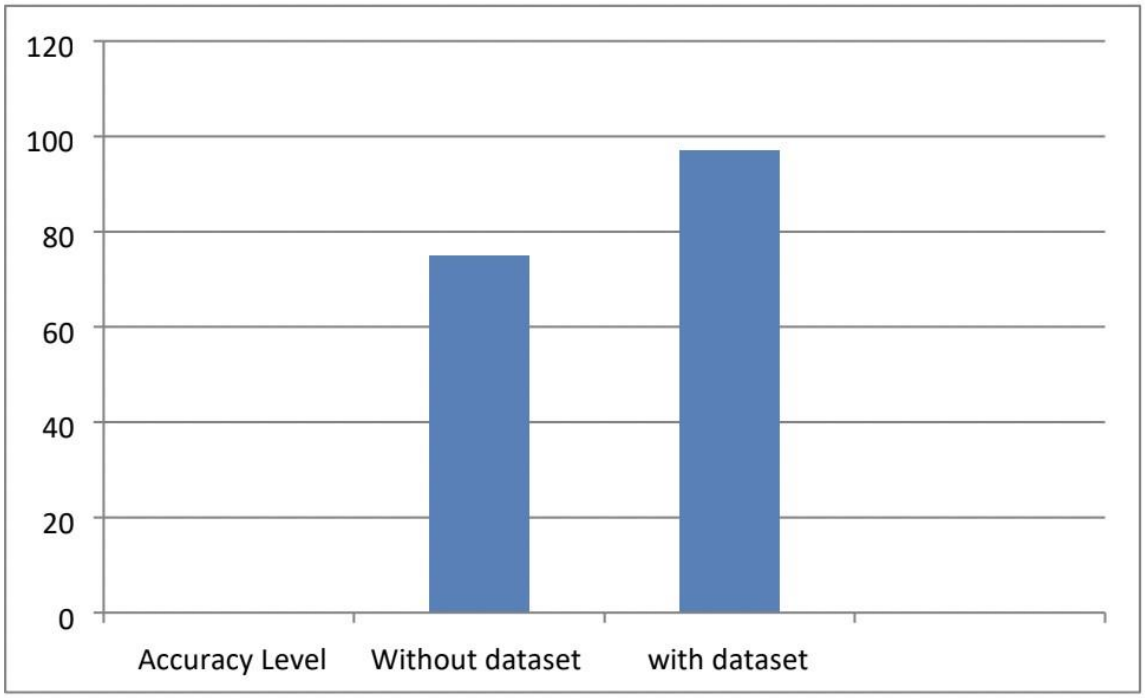
- The home page and the result page is tested .It is working well without issues.
- The app was tested for functionality.
- The scoring end point application is slower than the normal flask app.

8.2. USER ACCEPTANCE TESTING

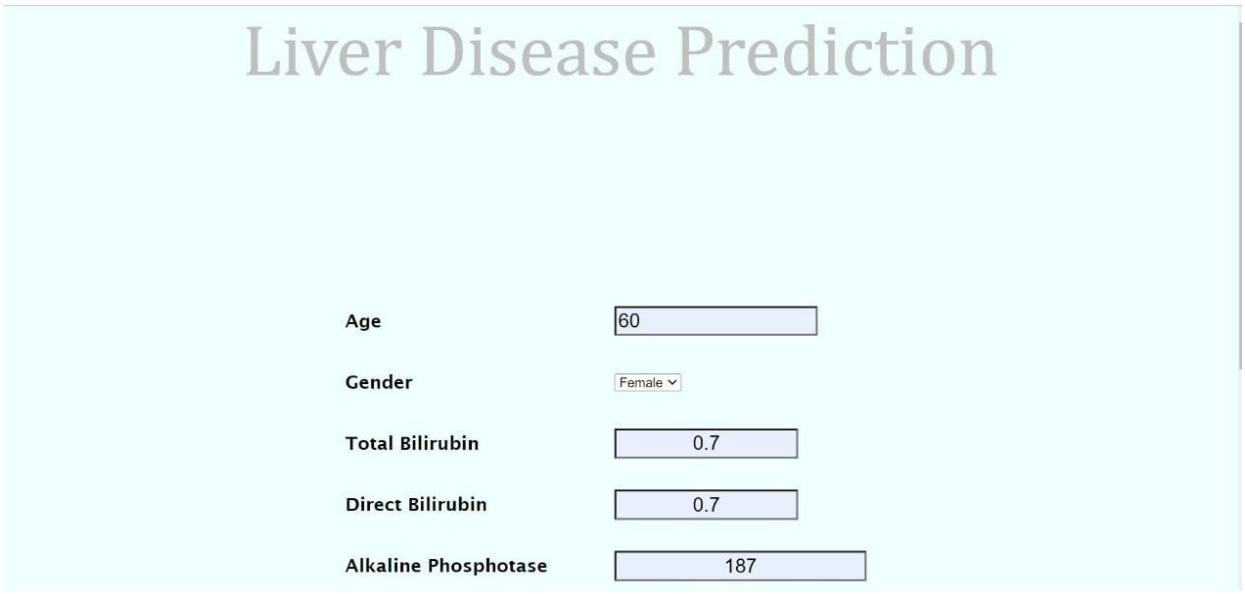
Section	Total Cases	Not Tested	Fail	Pass
Print Engine	8	0	0	8
Client Application	48	0	0	48
Security	2	0	0	2
Outsource Shipping	3	0	0	3
Exception Reporting	7	0	0	7
Final Report Output	4	0	0	4
Version Control	3	0	0	3

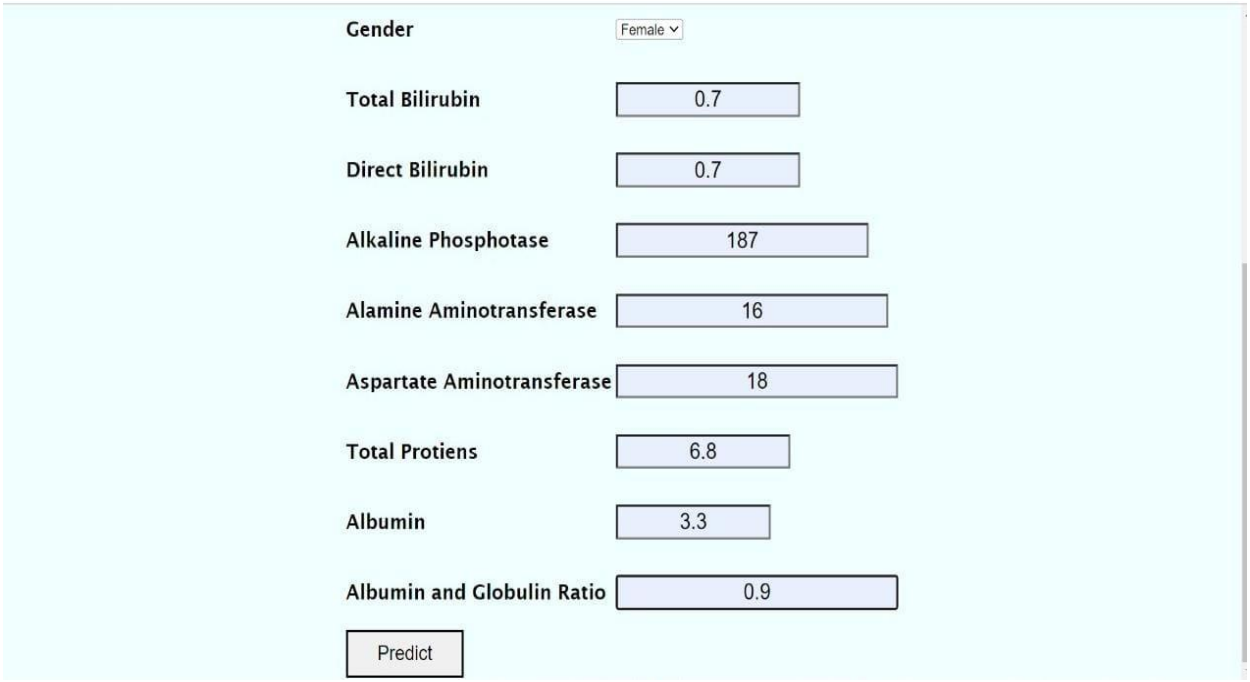
9. RESULTS

9.1 PERFORMANCE METRICS



9.2 OUTPUTS


Liver Disease Prediction
Age
60
Gender
Female
Total Bilirubin
0.7
Direct Bilirubin
0.7
Alkaline Phosphotase
187


Gender
Female
Total Bilirubin
0.7
Direct Bilirubin
0.7
Alkaline Phosphotase
187
Alamine Aminotransferase
16
Aspartate Aminotransferase
18
Total Protiens
6.8
Albumin
3.3
Albumin and Globulin Ratio
0.9
Predict

Liver Disease Prediction

{% if prediction==2 %}

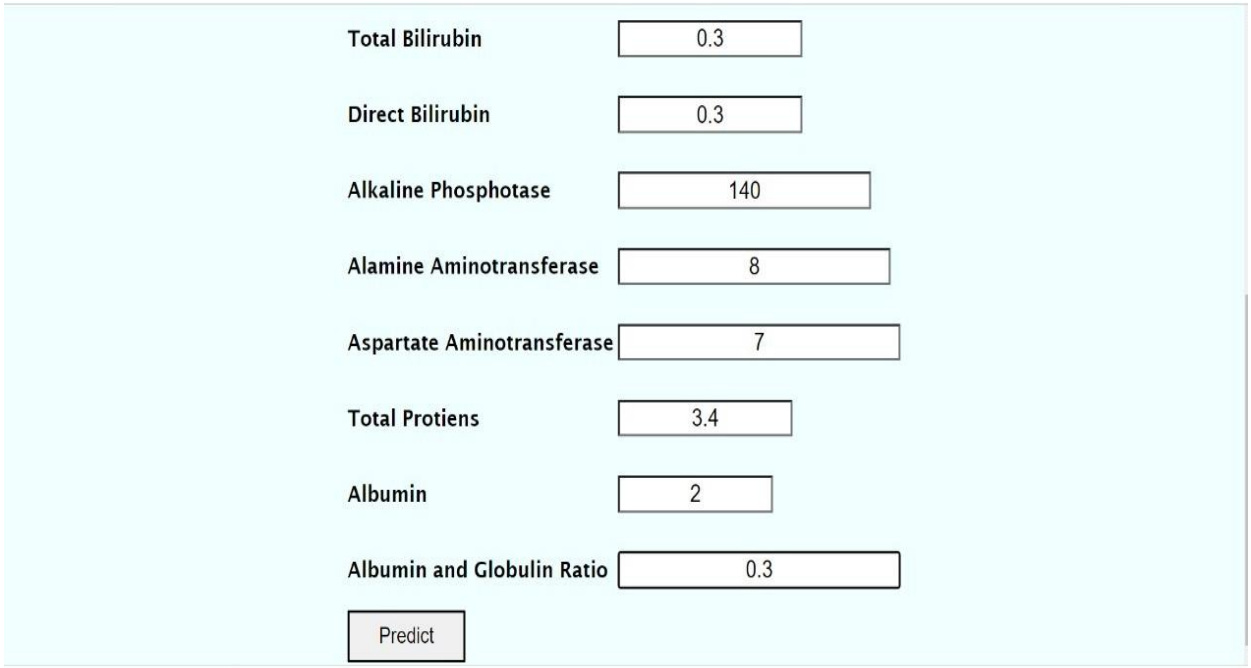
Oops! 😟

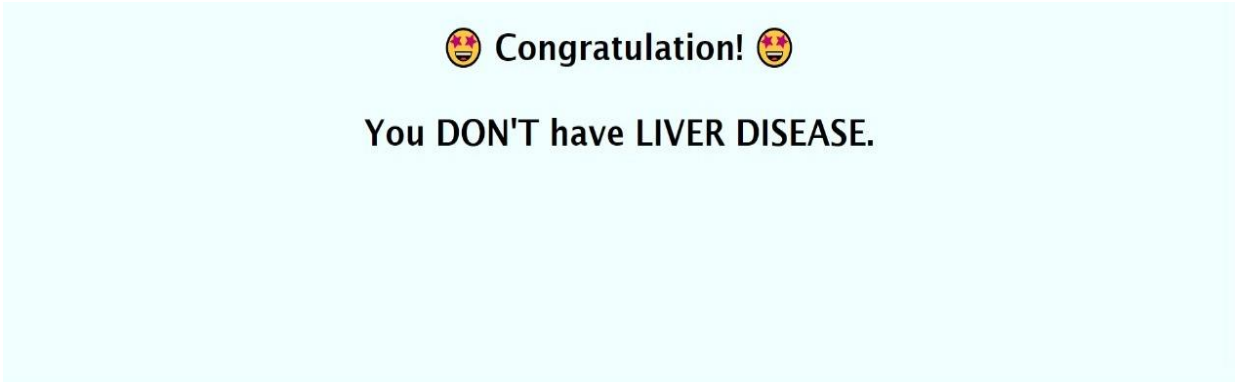
You have LIVER DISEASE

Please Consult a Doctor.

Liver Disease Prediction

Age	60
Gender	Female
Total Bilirubin	0.3
Direct Bilirubin	0.3


Total Bilirubin
0.3
Direct Bilirubin
0.3
Alkaline Phosphotase
140
Alamine Aminotransferase
8
Aspartate Aminotransferase
7
Total Protiens
3.4
Albumin
2
Albumin and Globulin Ratio
0.3
Predict


Congratulation!
You DON'T have LIVER DISEASE.

10. ADVANTAGES

- This helps in early diagnosis of liver disease
- It makes the process simple and easier.
- We are able to monitor the patient in an effective manner
- We can also help patients who are not in a condition to directly consult a doctor.

DISADVANTAGES

- There are inaccuracies which cannot be avoided.
- For people who have difficulty in accessing internet or those who do not have electronic gadgets , this is not a feasible solution.

11. CONCLUSION

We conclude by suggesting this web application for liver disease prediction . The web application is supported by the Machine Learning and IBM Watson cloud which stands for the complex image prediction and user information storage.

Help healthcare management and professionals to explore better results in numerous clinical applications, such as medical image processing, language processing, and tumor or cancer cell detection, by finding appropriate features.

Several statistical and machine learning approaches (e.g., simulation modeling, classification, and inference) have been used by researchers and lab technicians for better prediction. The clinical results are more data-driven than model-dependent.

12. FUTURE SCOPE

- In biomedical science, accuracy and speed are two important factors that should be considered chiefly in dealing with any disease.

- In this regard, Machine Learning techniques can be of great help to physicians.With advances, several machines have entered in our lives.

- One of the most famous areas where computers as the mostly used machines can be helpful is knowledge extraction with the help of a machine (machine learning)

13. APPENDIX

application.py

```
from flask
import
Flask,
render_temp
late,
request
              import pickle
              app=Flask(__name___,template_folder='template')


              @app.route('/')
              def home():
                  return render_template('home.html')


              @app.route('/index')
              def index():
                  return render_template('index.html')


              @app.route('/data_predict',methods=['POST'])
              def data_predict():
                  age=request.form['Age']
                  gender=request.form['Gender']
```

```
    tb=request.form['Total_Bilirubin']
    db=request.form['Direct_Bilirubin']
    ap=request.form['Alkaline_Phosphotase']
    aa1=request.form['Alamine_Aminotransferase']
    aa2=request.form['Aspartate_Aminotransferase']
    tp=request.form['Total_Protiens']
    a=request.form['Albumin']
    agr=request.form['Albumin_and_Globulin_Ratio']



data=[[float(age),float(gender),float(tb),float(db),float(ap),float(aa1),float(
aa2),float(tp),float(a),float(agr)]]


    model=pickle.load(open('liver_analysis.pkl','rb'))
    prediction=model.predict(data)[0]


    if prediction==1:
        return render_template("Chance.html")
    else:
        return render_template("noChance.html")

if __name__=='__main__':
    app.debug=True
    app.run(host='0.0.0.0',port=5000)
```

Github Repo:

https://github.com/IBM-EPBL/IBM-Project-26943-1668681072