STATISTICAL MACHINE LEARNING APPROACHES TO LIVER DISEASE PREDICTION

(TEAM ID: PNT2022TMID34932)

PROJECT REPORT

Submitted by

THERSAL J. (962819106043)

ASHMITHA M. (962819106007)

SRUTHI R. (962819106042)

DEVI PRIYA L R. (962819106010)

in partial fulfillment for the award of degree

of

BACHELOR OF ENGINEERING

in

ELECTRONICS AND COMMUNICATION ENGINEERING

UNIVERSITY COLLEGE OF ENGINEERING, NAGERCOIL.

ANNA UNIVERSITY : CHENNAI 600 025

NOVERMBER 2022

INDEX

1. INTRODUCTION

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 patient's 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. 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 are:

- **Increased convenience for predicting a liver disease**
- **Reduction in number of deaths due to liver diseases**
- **More accurate diagnosis of liver disease by the doctors**

1.1 Project Overview

Liver Disease is a prominent Disease other than heart attack, which is taking a lot of lives. Since most of the time liver disease is detected at the later stage leading to death. Number of liver patients is increasing because of several reasons like over consumption of alcohol, breathing in injurious gas, consuming polluted water and so on which can affect health parameter. Using a machine learning prediction models, liver diseases can be predicted using those parameters in early stages. In this work to build the machine-learning models, Indian Liver Patient Dataset (ILPD) hosted at UCI.edu is used, which is based on Indian patient and Naive Bayes algorithm is used to predict the disease with different prepossing techniques. Dataset is checked for skewness, outliers and imbalance using univariate and bivariate analysis and then suitable algorithms used to remove outliers and various oversampling and undersampling techniques are used to balance the datal. Further refinement of model is done through hyperparameter tunning using grid search and feature selection. The Final model provides 100% accuracy and also good score across different metrics.

1.2 Purpose

The purpose of this project is used to predict the liver disease using Naive Bayes Algorithm and this is used to predict the liver disease earlier and reduce the death rate.

2. LITERATURE SURVEY

1. In this paper, we have compared several ML methods, such a Logistic Regressio (LR), Decision Tree (DT), Random Fores (RF) and Extra Trees (ET) for the prediction of liver disorders. Issues ignored by the previous researchers have been taken into consideration to improve the prediction accuracy. At thepreprocessing steps ,categorical values are encoded through label encoding. Then we have utilized the Pearson Correlation Coefficient to improve irrelevant features. After the removal of redundant features of, over sampling is used to mitigate the imbalanced class distribution problem, and feature scaling is used to handle the outliers. After the completion of datapreprocessing, LR, DT, RF, and ET classifiers are used for classifying liver disorders patients. For further improvement, the AdaBoost classifiers are used to increase the performance of each classification algorithm.

2. This paper gives us the basic idea of past published paper of detection and diagnosis of liver disease based on different machine learning algorithm. With this survey and study it has clearly find and observed that some machine learning algorithm such as Decision tree, J48 and ANN provide better accuracy on detection and prediction of liver disease. And 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. And also the paper presents a survey on different types of machine learning techniques used by different authors and every machine learning techniques has some good and bad outcomes depend on the datasets and features selection etc. With this survey we found out that the accuracy and performance can be improve by using different combination or hybrid machine learning algorithm and in future we can also work on more parameter which help to getbetter performance than

the existing technique.

3. The scale of patient medical records increases day by day in the health care sector. Data mining is the method of using a computer-based information system (CBIS), using modern tactics, to uncover insights from data. The machine learning method is close to that of data mining. Algorithms in machine learning can be differentiated from either supervised or unsupervised methods of learning. For statistical modelling, supervised learning approaches are commonly used. Predictive modelling is a subset of the area of clinical and business intelligence that is used to identify health risks and also to forecast individuals' potential health status. 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 organised or unstructured. Electronic health records are stored in a standardised data format using managed language to log patient knowledge as a written texts 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[1]. Aggregations of various classifiers are J48, C4.5 and Naive Bayes, etc. [2]. 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 [3][4]. Major issues deliberated on patients with liver disease are not readily detected at starting phase since that can usually operate even though it is partly impaired. An early detection of liver disorders will improve the survival rate of the patient. There is a high probability of liver failure among Indians. India is expected to become the World Capital of Liver Diseases by 2025. Because of the deskbound lifestyle, increased alcohol intake and smoking, the pervasive prevalence of infection inside liver in India is contributing around 100 forms of liver infections are present. It would also be of great value in the medical field to build a computer that will increase the diagnosis of the disease. This systems

can help doctors make correct treatment choices, and the patient queue will also be minimised by liver specialists such as endocrinologists assisted by Automated categorization Methods for Disorders in Liver part. In medical diagnosis and disease prediction, classification techniques are widely common. Michael J Sorich [5] described on chemical datasets, the classification (SVM) provides better prediction results. Paul R Harper [6] stated that an absolute greatest categorization method does not provide forecasted results. However, the unsurpassed achieving algorithm depends on distinct features of the dataset being evaluated.

4. The machine-learning model is capable of predicting diseases, based on a data set, which is built in combination of key health parameters of a person with diseases and without diseases. For building models, an effective data set is needed, with proper representation of disease classifications. In this work, Indian Liver Patient Dataset hosted at ics.uci.edu is used. Several classification machine-learning algorithms are able to classify the liver diseases. 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 Random Forest algorithm is to build a model, since it is trained on several samples of data obtained from splitting data so that the model is not tuned for very specific data. 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. Various balancing techniques discussed and their impacts on performance are tabulated laterin sections.

5. The disease can be predicted based on health parameters, oral conditions - like alcohol, city pollution level, movement, body chemical compositions using advanced AI/ML techniques. ML is the branch of AI in which a machine learns from a dataset and its performance measures improve with real data over time. The different techniques of

ML have been adopted for diagnosis and prophecy of various diseases in the field of medicine. Due to easy access to clinical data, ML algorithms play an important role in medical decision making. Therefore to identify the disease and make a real-time effective decision the design and develop of a ML model will play a major role. Several ML classification algorithms exist to predict the Liver disease. Each algorithm has different ways of learning from the data set and can be refined / performance tuned. The paper focuses on KNN algorithms, steps to be performance to optimise the model, step by step developing several models. The reason for picking a KNN algorithm is, which looks for several nearby values to classify diseases. Which helps to analyse various more effectively, as increased K value looks at several nearest values, before classifying disease. Instead of building models using US/Europe based data sets, the paper works on building ML models effectively using Indian dataset and paper discussed how to analyze and predict with more accuracy step by step - preprocessing of data, Univariate and by variate analysis, feature selection, Feature engineering then ML model is trained using this data. As part of preprocessing data is analyzed and checked each feature distribution, most importantly is the data set is balanced/imbalanced and then appropriate methods used to transform the data to normal distribution and imbalance data set is balanced using various methods.. Feature engineering of the dataset is performed to get the important features for prediction and remove the less contributing features so that computation time of the model can be reduced. Hyper parameter tuning methods, tune the parameter's values of KNN ML model to get high accuracy. The paper builds several models, using all these techniques, to indicate performance importance and finally achieving a final KNN ML model to predict liver disease effectively for Indian Dataset. The several models built to predict.

2.1 Existing Problem

The main Problem is doctors cannot diagnose on the basis of variation in the test result .In this Application ,by using patient records that includes blood test reports

results,we will determine which patient has liver disease and which ones do not in an accurate and faster way

2.2 References

1, Paper Name: Liver Disease Prediction Using Machine Learning Algorithms: A Comparative Study
Author Name: Md. Fazle Rabbi, S.M. Mahedy Hasan, Arifa Islam Champa, Md. Asif Zaman, Md. Kamrul Hasan
2,Paper Name: A Survey on machine learning techniques for the diagnosis of liver disease
Author Name: Golmei Shaheamlung, Harshpreet Kaur, Mandeep Kaur
3, Paper Name: Evaluation based Approaches for Liver Disease Prediction using Machine Learning Algorithms
Author Name: C. Geetha, AR. Arunachalam
4, Paper Name: Optimizing Liver disease prediction with Random Forest by various Data balancing Techniques
Author Name: Satessh Ambesange, Vijayalaxmi A, Rashmi Uppin, Shruthi Patil,Vilaskumar Patil
5, Paper Name: Liver Diseases Prediction using KNN with Hyper Parameter Tuning Techniques
Author Name: Sateesh Ambesange, Ranjana Nadagoudar, Rashmi Uppin, Vilaskumar Patil, Shruti

2.3 Problem Statement Definition

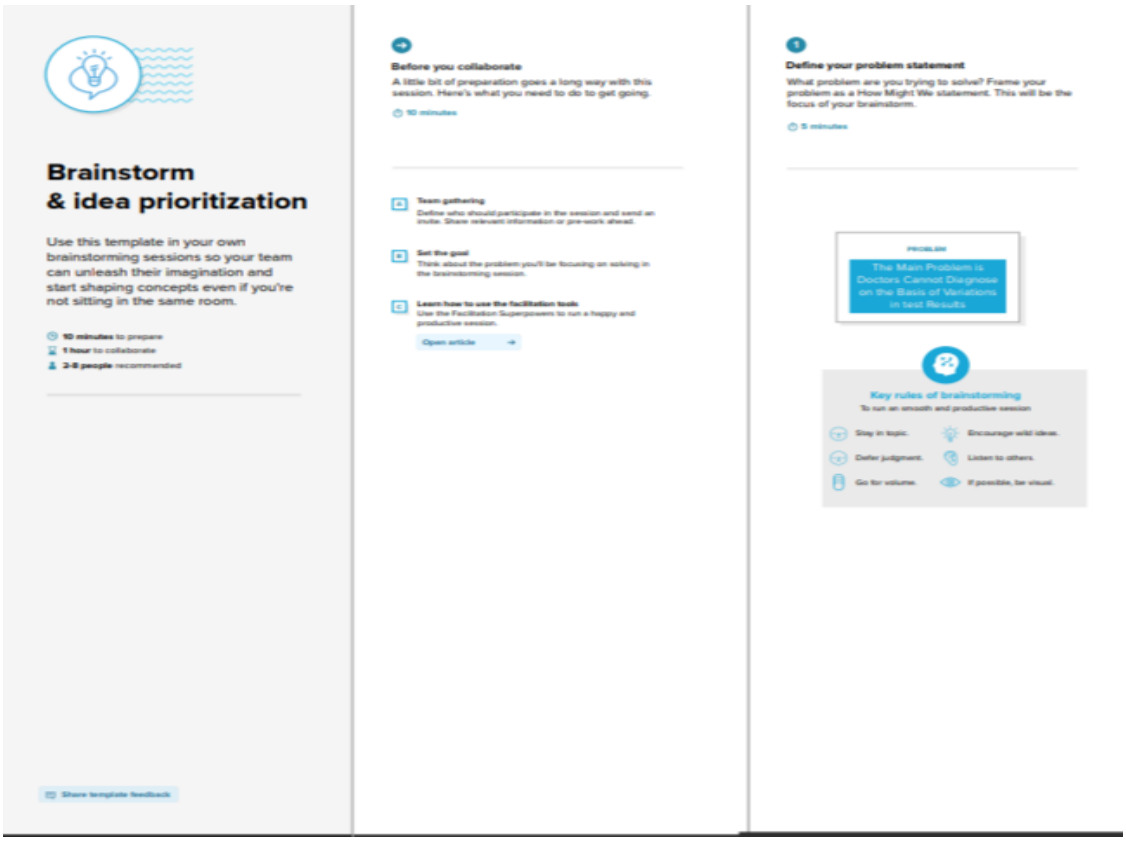
The main problem is doctor cannot diagnosis on the basis of variations in test results.

3. IDEATION AND PROPOSED SOLUTION

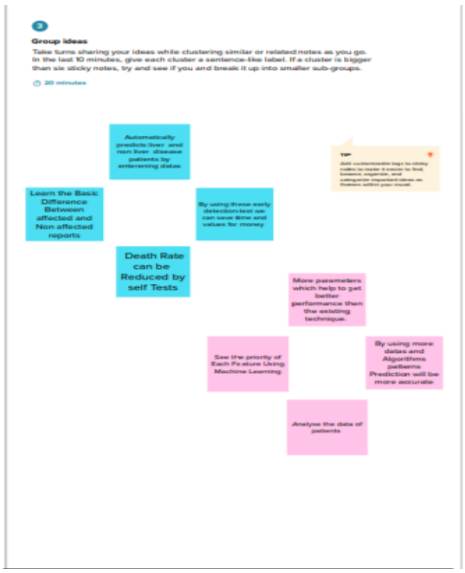
3.1. Empathy Map Canvas

3.2. Ideation and Brainstorming

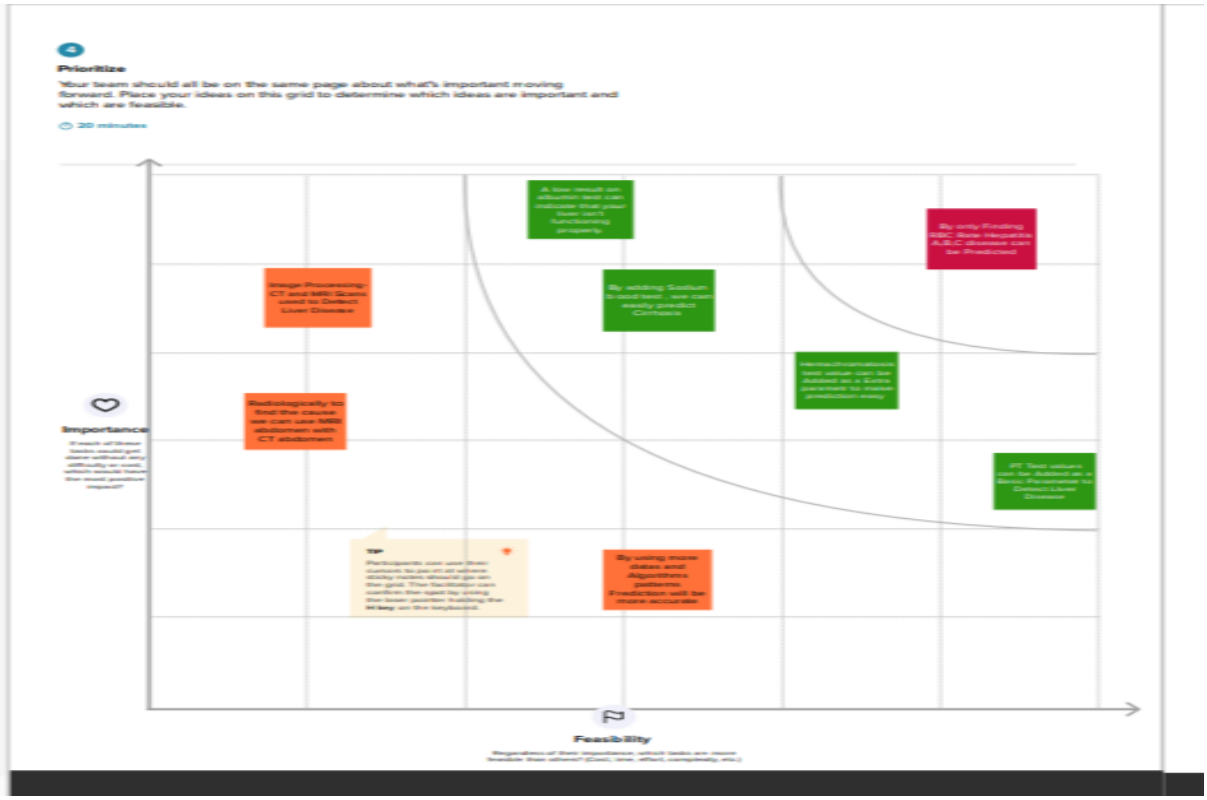
Step 1: Team Gathering, Collaboration and Select the Problem Statement



Step 2: Brainstorm, Idea Listing and Grouping



Step 3: Idea Prioritization



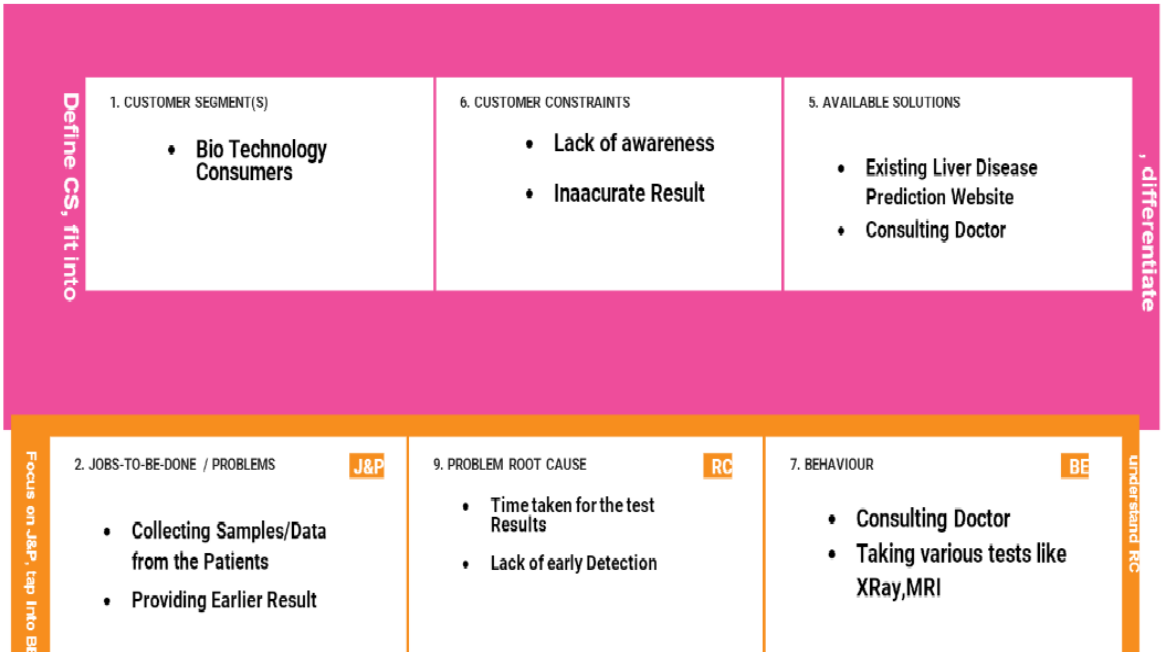
3.3. Proposed Solution

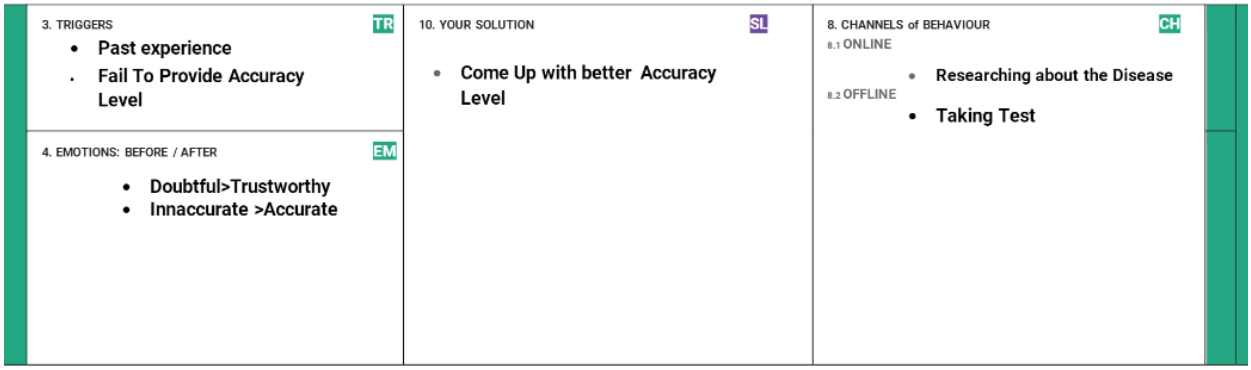
s.no	Parameter	Description
1.	Problem Statement (Problem to be solved)	Nowadays many peoples are affected by Liver disease because of that so many people are losing their life. Early detection of liver disease can be very helpful in the treatment of the disease to fast recover but it is very difficult to identify the liver disease in early

		stages. In some situations, the medical expert is unable to detect the symptom even at an early stage. It is one of the great losses for the patient
2.	Idea / Solution description	The software will detect the patients symptoms and it will find out the disease according to that symptom and it will show the result by ML techniques. We will do this by taking the data set of both normal and abnormal liver and we will train the software in that way by detecting the disease according to the symptom.
3.	Novelty / Uniqueness	In this software model we're not only predicting the disease but also giving some basic precautions like what to do & don't.which makes this model unique
4.	Social Impact / Customer Satisfaction	Many liver disease are left unpredicted.By implementing this software liver disease can be diagnosed in early stages which result in the decrease of death rate
5.	Business Model (Revenue Model)	Currently the global is running with

		newest technology likewise our project will more helpful to medical fields. And the medical institution , clinics and hospitals need to paid for yearly license and get renew yearly to continue the check up using this software
6.	Scalability of the Solution	The software will never show the new types of liver diseases in future because the algorithm and datasets we provided only for the current liver diseases in case of any new liver diseases founded in future it will show as a error in output we need to change the algorithm process to show without error in future.

3.4. Problem Solution Fit





4. REQUIREMENT ANALYSIS

4.1. Functional Requirements

FR No	Functional Requirements(Epic)	Sub Requirement (Story/Sub Task)
FR 1	User Registration	Registration through Gmail
FR 2	User Confirmation	Confirmation via Email
FR 3	Prediction	Liver Disease can be Predicted more Accurately by using Support Vector Machine Algorithm
FR 4	Hardware Requirements	2GB RAM(minimum) 100GB HDD(minimum) Intel i3 quad core 1.66GHz processor(minimum) Internet Connectivity
FR 5	Software Requirements	Windows 7 or higher Python 3.6.0 or higher Visual Studio Code Flask (python platform) HTML Dataset consisting of Liver Disease Required libraries Jupiter notebook
FR 6	Other requirements	IBM cloud login Chrome extension features
FR 7	Events	Model needs a capability of retrieving and displaying

		accurate result

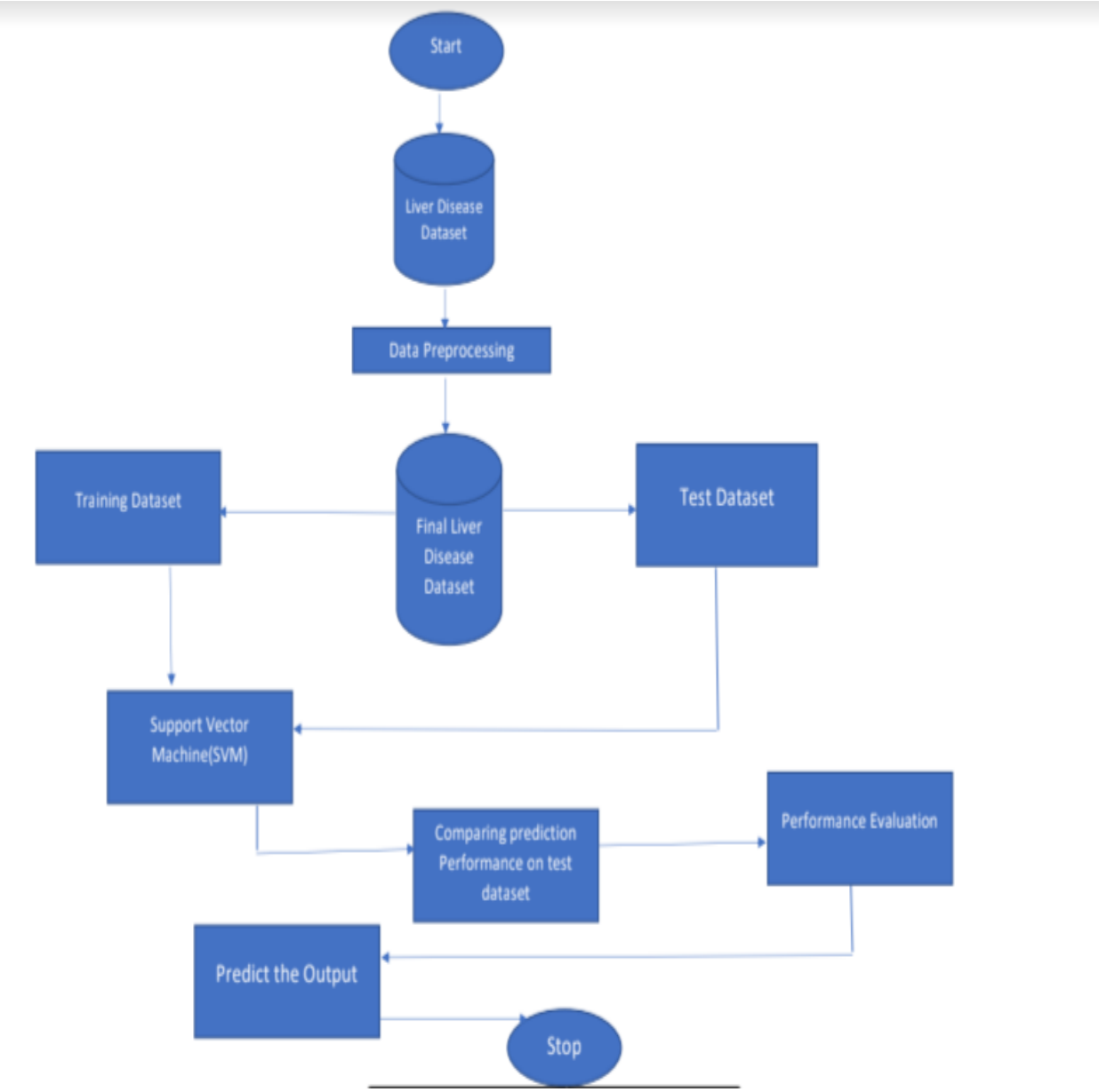
4.2. Non-Functional Requirements

NFR No	Non Functional Requirements	Description
NFR 1	Usability	This system is really used as it can able to detect Liver Disease .By detecting the liver disease early ,death rate is decreased
NFR 2	Security	Assuring all data inside the system or its part will be Protected
NFR 3	Reliability	This Approach gives more accuracy than the existing solution
NFR 4	Performance	The effectiveness of these methods relies on feature collection, training data, and classification algorithms. It must be processed and executed within a fraction of a second using the Machine learning algorithm
NFR 5	Availability	It doesn't have any

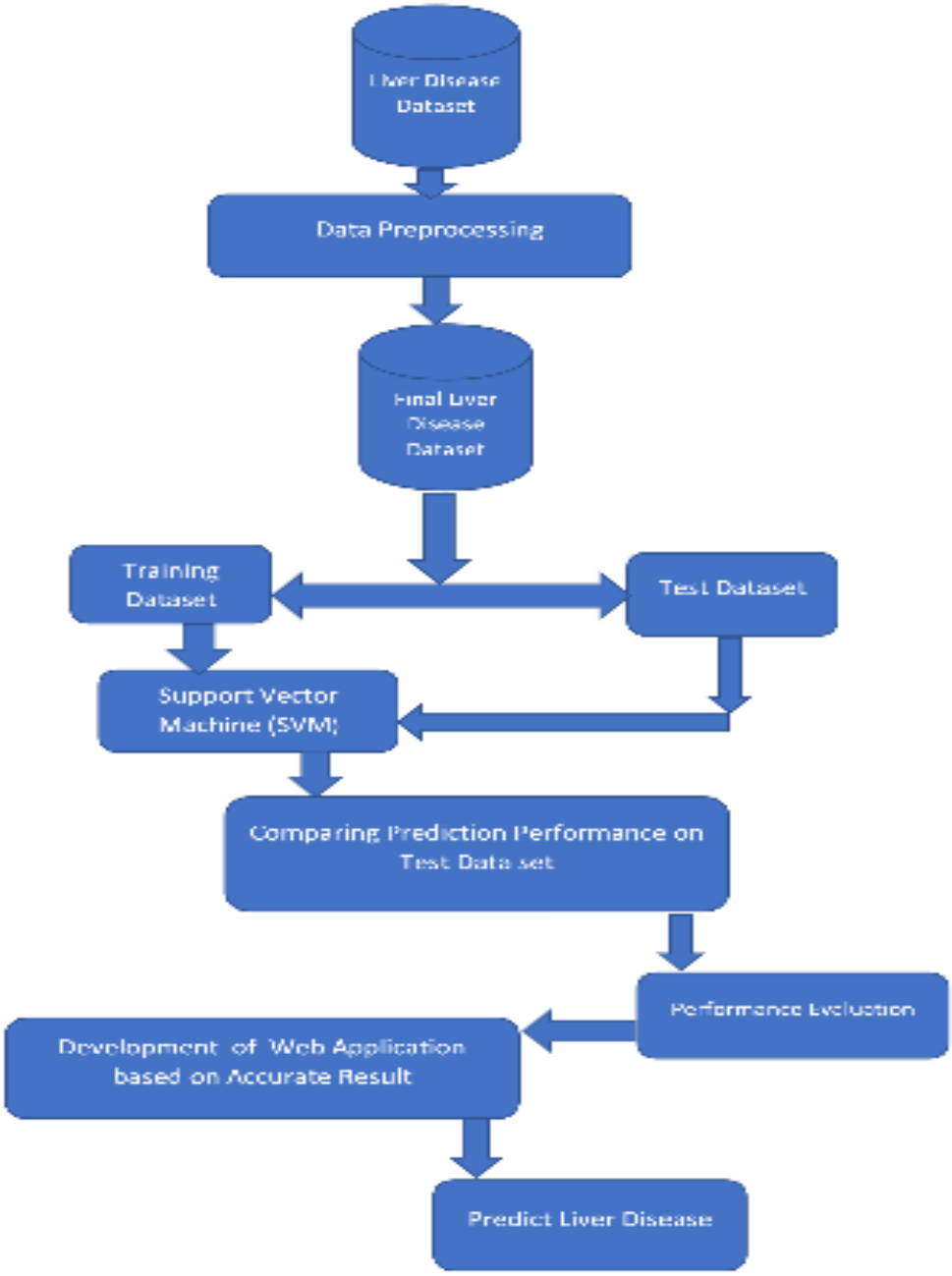
		restrictions , it is available for all individual user
NFR 6	Scalability	It is acceptable to fit them over any place and any resources.

5. PROJECT DESIGN

5.1 Dataflow Diagrams



5.2. Solution Architecture



5.3. Technical Architecture

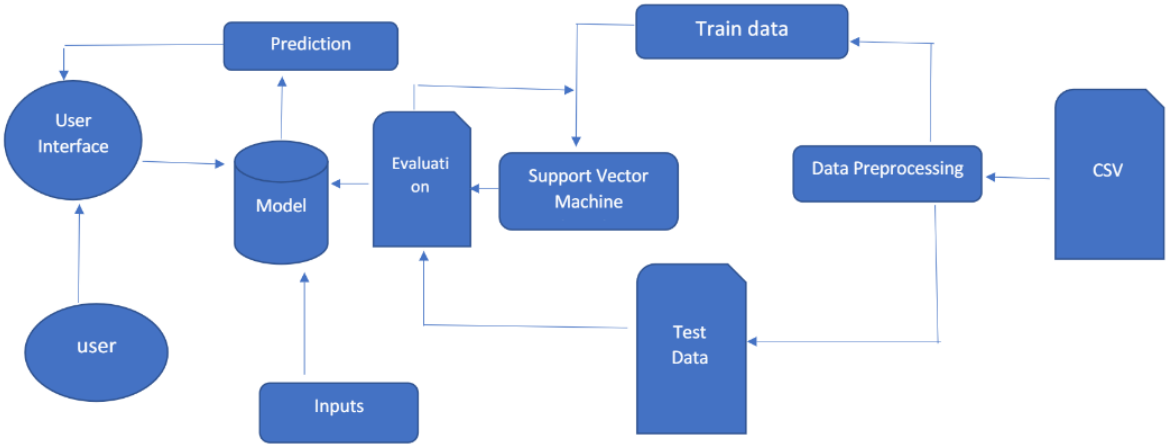


Table-1: Components & Technologies

S.No	Component	Description	Technology
1	User Interface	How user interacts with application e.g. Web UI, Mobile App, Chatbot etc.	HTML, CSS, JavaScript
2	Application Logic-1	Logic for a process in the application	Python
3	Application Logic-2	Logic for a process in the application	IBM Watson STT service
4	File Storage	Files are stored in cloud	IBM Block Storage or Other Storage Service or Local

			Filesystem
5	Machine Learning Model	Prediction of Liver Disease	Support Vector Machine Algorithm
6	Infrastructure (Server / Cloud)	IBM Cloud App Configuration is a centralized feature-management and configuration service on IBM Cloud	IBM Cloud Foundry, Kubernetes, etc.

Table-2: Applications Characteristics

S.No	Characteristics	Description	Technology
1	Open-Source Frameworks	There are no open-source frameworks in this application.	Technology of Opensource framework
2	Security Implementations	Block chain technology is used for Security implementation its private framework protects all data..	Block chain
3	Scalable Architecture	Users are Provided with medical services online	IBM cloud
4	Availability	Available for	Technology used

		everyone , no Restrictions	
5	Performance	Predicted Result is more accurate	Support Vector Machine Algorithm

5.3 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
	verify / Login	**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
	Monitoring	**USN-3**	As a user, I can monitor the	I can do it from any	High	Sprint 2

			account process to access .	place		
	Dashboard	**USN-4**	All the login process and activities done will be displayed on the dashboard	Helpful for reminding the actions	Medium	Sprint 2
Customer (Web user)	**Getting information**	**USN-5**	As a user, I need to gather the information from the real scenario	I can collect all the Information	High	Sprint 1
Customer Care Executive	**Analysing**	**USN-6**	As a User, I need to analyse the information and get into the decision to predict the disease	Helpful to Predict the disease early	High	Sprint 2
Administrator	**Ordering**	**USN-7**	As a user, I would order my officers to help the user to predict the disease	Iam the higher Authority and I can order them	High	Sprint 2
			As a user, I will	I will try to	High	Sprint 1

		USN-8	observe the account. If any prediction of disease goes wrong, I will surely take actions to avoid them	manage the situation		
	Obeying Orders	**USN- 9**	As a user I need to obey my higher officers command and take measures mentioned by them	I am in the way to obey the orders	Medium	Sprint 2

6. PROJECT PLANNING &SCHEDULING

6.1. Sprint Planning & Estimation

Sprint	**Functional Requirement (Epic)**	**User Story Number**	**User Story / Task**	**Story Points**	**Priority**	**Team Members**
Sprint-1	User Input	USN-1	Test datas are given as input for prediction	10	High	M.Ashmitha, L.R.Devi Priya

Sprint-1		USN-2	Model compares the given data with the Liver disease affected data	10	High	J.Thersal, R.Sruthi
Sprint-2	Prediction	USN-3	Model predicts the liver disease using Machine Learningalgorithm Support Vector Machine(SVM)	10	High	J.Thersal, M.Ashmitha
Sprint-3	Classifier	USN-4	Model sends all the output to the classifier and produces the final result.	10	High	R.Sruthi, L.R.Devi Priya
Sprint-4	Announcement	USN-5	Model then displays whether the patient is affected by liver disease or not	10	High	J.Thersal, R.Sruthi
	Events		This model needs the capability of displaying accurate result	10	High	M.Ashmitha, L.R.Devi Priya,

6.2. Sprint Delivery Schedule

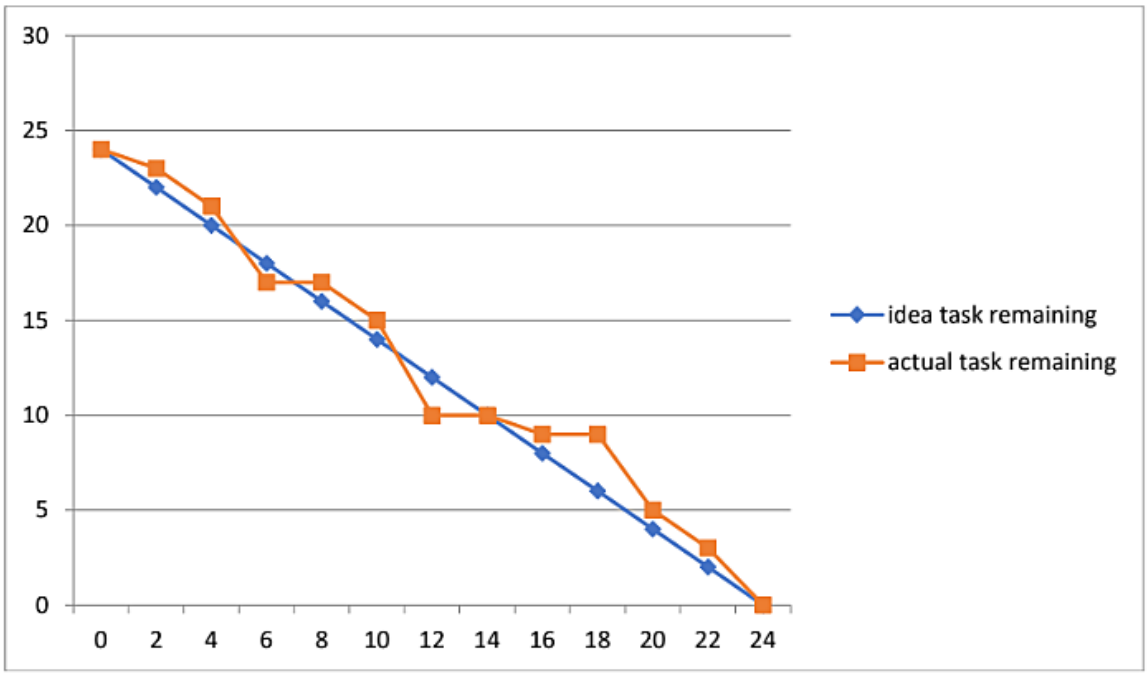
Sprint	Total Story Points	Duration	Sprint Start Date	Sprint end Date (Planned)	Story points Completed(as on Planned End Points)	Sprint Release Date (Actual)
Sprint-1	20	6 Days	24 Oct 2022	29 Oct 2022	20	20 Oct 2022
Sprint-2	20	6 Days	31 Oct 2022	05 Nov 2022	20	05 Nov2022
Sprint-3	20	6 Days	07 Nov 2022	12 Nov 2022	20	12 Nov2022
Sprint-4	20	6 Days	14 Nov 2022	19 Nov2022	20	19 Nov2022

Velocity

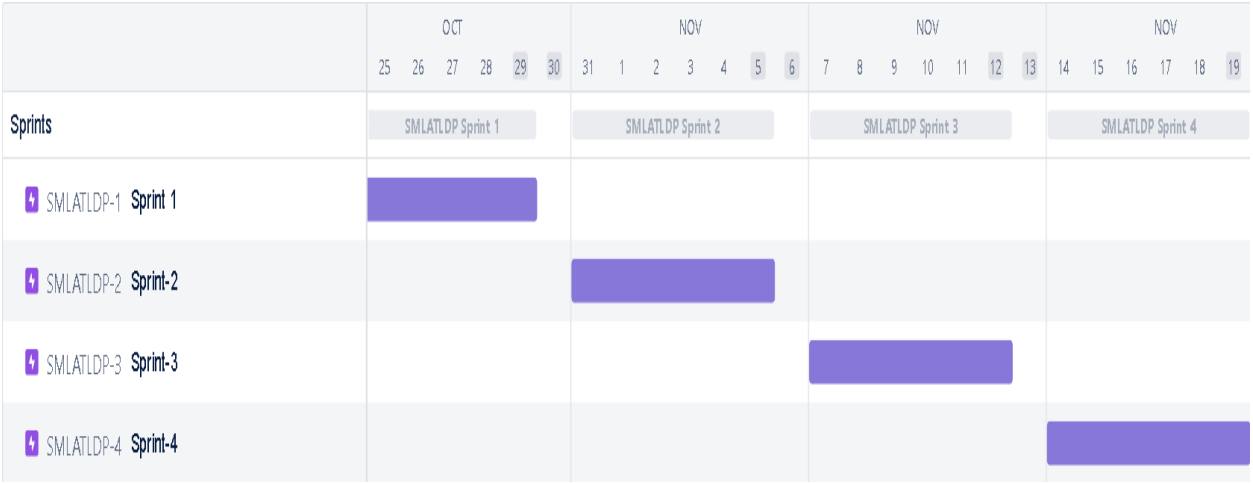
Imagine we have a 10-day spirit duration ,and the velocity of the team is 20(points per spirit). Let's calculate the team average velocity(AV) per iteration unit (story points per day)

AV= Spirit Duration/Velocity =20/10 =2

Burndown Chart



6.3 Reports from JIRA



7.CODING AND SOLUTIONING

7.1 Feature

Login Page: The login page ask the user to enter the data of test result.

Result Page: The result page tells whether the person has liver disease or not.

7.2 Code s

home: index

```
<!DOCTYPE html>
<html lang="en">
  <head>
    <meta charset="UTF-8">
    <meta http-equiv="X-UA-Compatible" content="IE=edge">
    <meta name="viewport" content="width=device-width, initial-scale=1.0">
    <title>Document</title>
    <style>
      table
      {
        width:100%;
      }
      td
      {
        height:50px;
        width:50px;
      }
    </style>
  </head>
  <body>
    <form method="POST" action="/predict">
      <table>
        <tr>
```

```
            <td>
              <label for="age">Age :</label>
            </td>
            <td>
              <input type="number" name="age" step=0.1/>
            </td>
          </tr>
          <tr>
            <td>
              <label for="gender">Gender :</label>
            </td>
            <td>
              <input type="number" name="gender" step=0.1/>
            </td>
          </tr>
          <tr>
            <td>
              <label for="total_bilirubin">Total Bilirubin :</label>
            </td>
            <td>
              <input type="number" name="total_bilirubin" step=0.1/>
            </td>
          </tr>
          <tr>
            <td>
              <label for="alkaline_phosphotase">Alkaline Phosphotase :</label>
            </td>
            <td>
              <input type="number" name="alkaline_phosphotase" step=0.1/>
            </td>
          </tr>
          <tr>
            <td>
              <label for="alamine_aminotransferase">Alamine Aminotransferase
:</label>
            </td>
            <td>
```

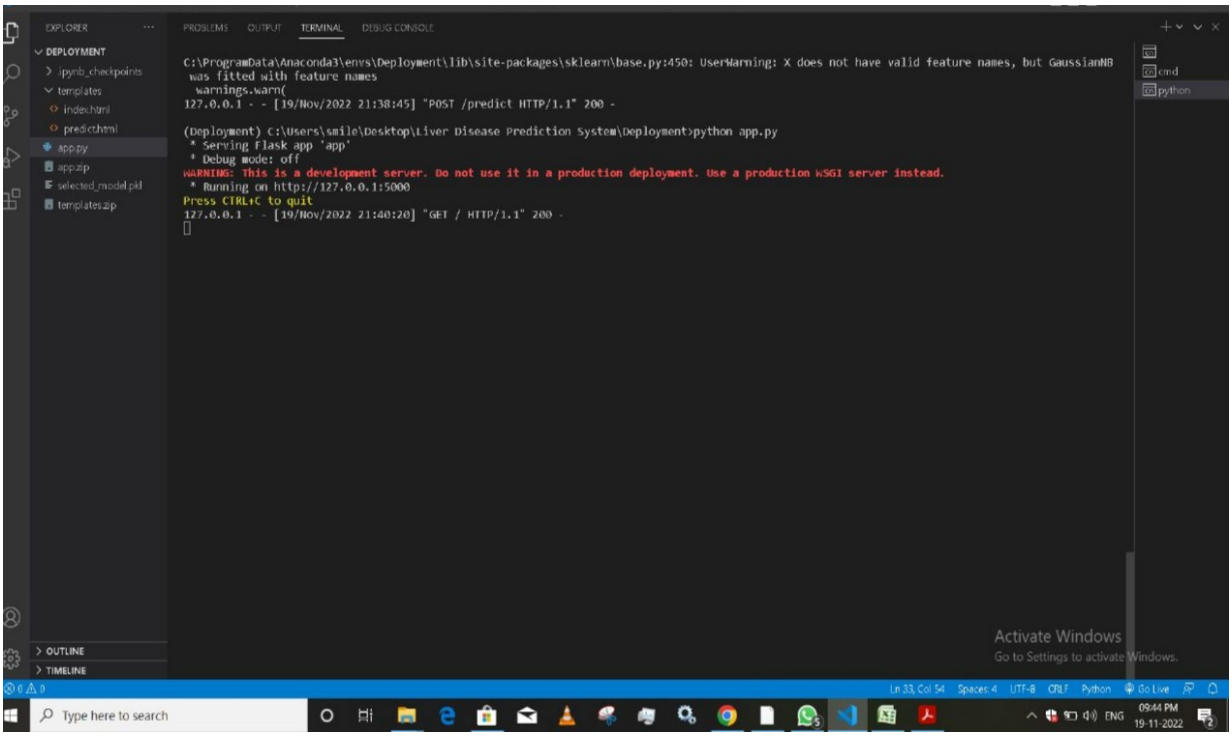
```
                <input type="number" name="alamine_aminotransferase" step=0.1/>
            </td>
        </tr>
        <tr>
            <td>
                <label for="aspartate_aminotransferase">Aspartate Aminotransferase
:</label>
            </td>
            <td>
                <input type="number" name="aspartate_aminotransferase" step=0.1/>
            </td>
        </tr>
        <tr>
            <td>
                <label for="total_protiens">Total Protiens :</label>
            </td>
            <td>
                <input type="number" name="total_protiens" step=0.1/>
            </td>
        </tr>
        <tr>
            <td>
                <label for="albumin">Albumin :</label>
            </td>
            <td>
                <input type="number" name="albumin" step=0.1/>
            </td>
        </tr>
        <tr>
            <td>
                <label for="albumin_and_globulin_ratio">Albumin and Globulin Ratio
:</label>
            </td>
            <td>
                <input type="number" name="albumin_and_globulin_ratio" step=0.1/>
            </td>
        </tr>
```

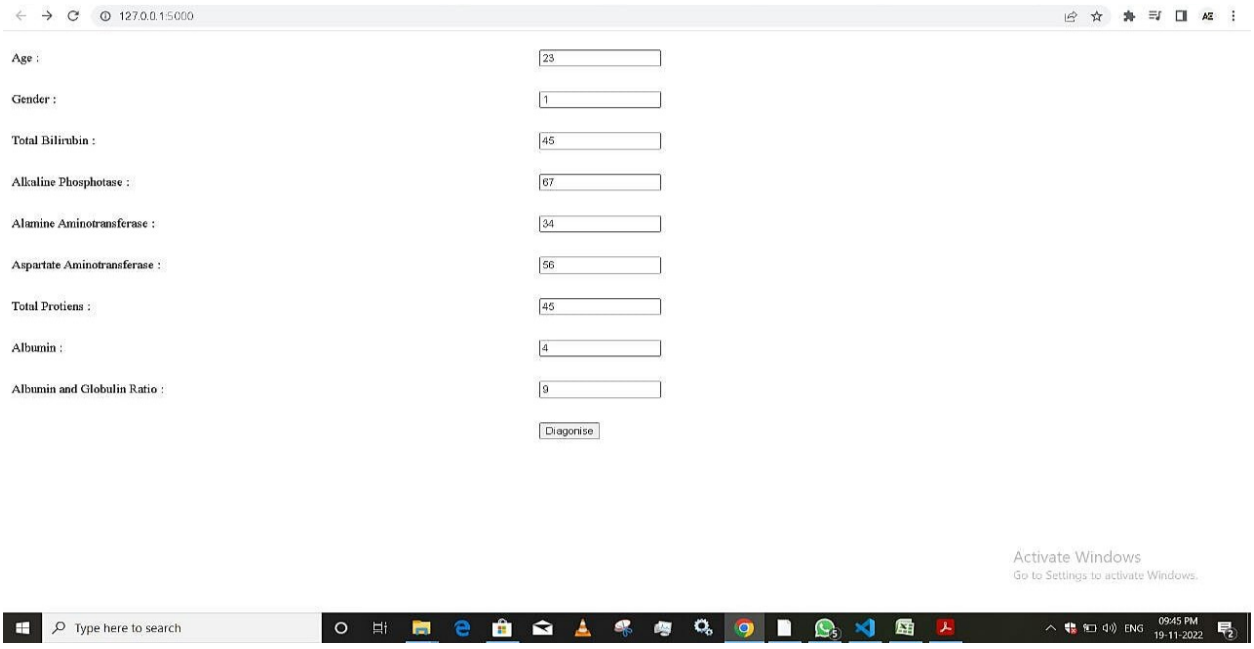
```
            <tr>
                <td>
                </td>
                <td>
                    <button type="submit">Diagonise</button>
                </td>
            </tr>
        </table>
    </form>
</body>
</html>
```

Predict

```
<!DOCTYPE html>
<html lang="en">
<head>
    <meta charset="UTF-8">
    <meta http-equiv="X-UA-Compatible" content="IE=edge">
    <meta name="viewport" content="width=<device-width>, initial-scale=1.0">
    <title>Document</title>
</head>
<body>
    <h1>The predicted result is</h1>
    <h1>{{predict}}</h1>
    <a href="/">Go Back</a>
</body>
</html>
```

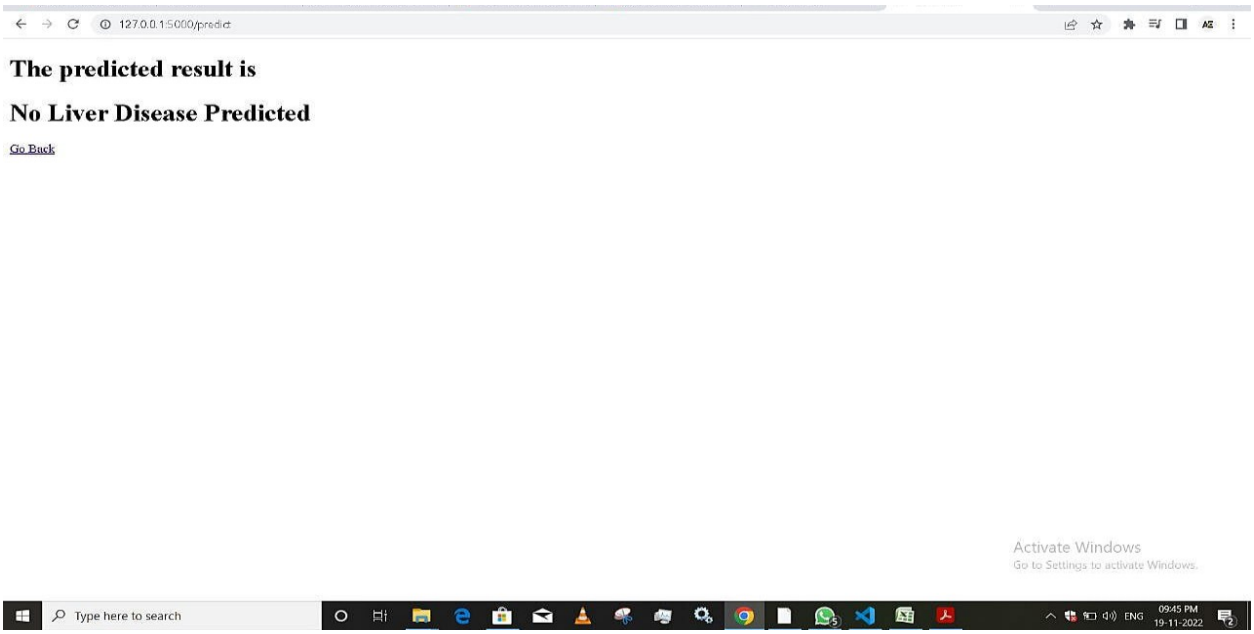
8.TESTING


EXPLORER
DEPLOYMENT
ipynb_checkpoints
templates
index.html
predict.html
app.py
app.zip
selected_model.pkl
templates.zip
OUTLINE
TIMELINE
PROBLEMS
OUTPUT
TERMINAL
DEBUG CONSOLE
C:\ProgramData\Anaconda3\envs\Deployment\lib\site-packages\sklearn\base.py:450: UserWarning: X does not have valid feature names, but GaussianNB was fitted with feature names
warnings.warn(
127.0.0.1 - - [19/Nov/2022 21:38:45] "POST /predict HTTP/1.1" 200 -
(Deployment) C:\Users\smile\Desktop\Liver Disease Prediction System\Deployment>python app.py
* Serving Flask app 'app'
* Debug mode: off
WARNING: This is a development server. Do not use it in a production deployment. Use a production WSGI server instead.
* Running on http://127.0.0.1:5000
Press CTRL+C to quit
127.0.0.1 - - [19/Nov/2022 21:40:20] "GET / HTTP/1.1" 200 -
cmd
python
Activate Windows
Go to Settings to activate Windows.
Type here to search
09:44 PM
19-11-2022


127.0.0.1:5000
Age :
23
Gender :
1
Total Bilirubin :
45
Alkaline Phosphotase :
67
Alamine Aminotransferase :
34
Aspartate Aminotransferase :
56
Total Protiens :
45
Albumin :
4
Albumin and Globulin Ratio :
9
Diagonise
Activate Windows
Go to Settings to activate Windows.
Type here to search
09:45 PM
19-11-2022

9. RESULTS

In this Project,we found that Liver disease can be Predicted using the test result values.



10. ADVANTAGES AND DISADVANTAGES

Advantages:

1. **No medical expertise required:** You dont need to have any knowledge of medical science and Liver Diseases to predict the Liver disease using this

application. All you need to do is enter the details being asked, which are already present in the blood test report and then you will get the result of prediction.

2. **High Accuracy:** This system predicts the results with 100% accuracy for the dataset that we have use while creating this application. While the accuracy might be different in some cases, It will still be high enough to be trustworthy at a large scale.

3. **Immediate results:** The result here are predicted within seconds of entering the details. You dont need to wait for a doctor to come, unlike in traditional method.

Disadvantages:

1. Due to any network issue there will be a delay in getting the predicted result
2. It is difficult to implement these techniques in some rural area
3. There is a possibility of entering wrong data so that the predicted result goes wrong

11. CONCLUSION

Diseases related to liver is becoming more and more common with time. With continous technological advancement, these are only going to increase in the future. Although people are becoming more conscious of health nowadays and are joining yoga classes, dance classes; still the sedentary life style and luxuries that are continously being introduced and enhanced; the problem is going to lastlong.

So in such scenario, our project will be extremely helpful to the society. with the dataset that we used for this project, we got 100% accuracy for Naive Bayes algorithm and through 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 the accuracy of 90% or more.

Today almost everybody above the age of 12 years has smartphones with them,and so

we can incorporate these solutions into an android app or ios app.Also it can be incorporated into a website and these app and website will be highly beneficial for a large section of society

12. FUTURE SCOPE

In the future ,we can apply different deep learning and transfer learning algorithms with various feature selection techniques for classifying liver patients and we can use in another set of data and check for the prediction accuracy. And also, we can work on more parameters which help to get better performance .

13. APPENDIX

Source Code

```
        import flask
from flask import request,render_template
from flask_cors import CORS
import joblib
import sklearn

app=flask.Flask(__name__,static_url_path='')
CORS(app)

@app.route('/',methods=['GET'])
def SendHomePage():
   return render_template('index.html')

@app.route('/predict',methods=['POST'])
def predictResult():
   a=float(request.form['age'])
```

```
    b=float(request.form['gender'])
    c=float(request.form['total_bilirubin'])
    d=float(request.form['alkaline_phosphotase'])
    e=float(request.form['alamine_aminotransferase'])
    f=float(request.form['aspartate_aminotransferase'])
    g=float(request.form['total_protiens'])
    h=float(request.form['albumin'])
    i=float(request.form['albumin_and_globulin_ratio'])

    x=[[a,b,c,d,e,f,g,h,i]]
    model=joblib.load('selected_model.pkl')
    result=model.predict(x)[0]
    if(result==2):
        res="Liver Disease Predicted"
    else:
        res="No Liver Disease Predicted"
    return render_template('predict.html',predict=res)

if __name__=='__main__':
    app.run()
```

Github:
https://github.com/IBM-EPBL/IBM-Project-39863-1660557322