A REVIEW OF LIVER PATIENT ANALYSIS METHODS USING MACHINE LEARNING

1 INTRODUCTION

1.1 Overview

Liver patient analysis using machine learning is a growing field that aims to improve the accuracy and speed of liver disease diagnosis and treatment. In recent years, various machine learning algorithms have been developed for liver disease diagnosis, including decision trees, support vector machines, random forests, artificial neural networks, and deep learning models.

One common approach to liver disease analysis using machine learning is to use laboratory test results and medical history data as input features to train a model that can accurately classify patients as healthy or diseased. For example, researchers have used data from liver function tests, such as alanine aminotransferase (ALT), aspartate aminotransferase (AST), alkaline phosphatase (ALP), and total bilirubin (TBIL), as well as demographic data, such as age and gender, to build predictive models for liver disease.

1.2 Purpose

The use of machine learning in analyzing liver patients can provide significant benefits in terms of accuracy and efficiency. With the vast amounts of data generated from liver patient analysis, traditional methods may not be able to fully leverage the potential of this data. Machine learning algorithms, on the other hand, are designed to handle large amounts of data and can detect patterns and trends that may not be immediately apparent to humans.

Benefits:

1. **Improved accuracy**
2. **Early detection**
3. **Personalized treatment plans**
4. **Improved efficiency**

2 Problem Definition & Design Thinking

2.1 Empathy Map

Template

Empathy map

Use this framework to develop a deep, shared understanding and empathy for other people. An empathy map helps describe the aspects of a user's experience, needs and pain points, to quickly understand your users' experience and mindset.

Share template feedback

Empathy Mapping For Liver patient Analysis :

An empathy map is a tool used to understand and empathize with the experiences,thoughts,feelings and behaviors of patients.It involves gathering insights by observing and listening to patients and then organizing the information into different categories.

Says

What have we heard them say?
What can we magine them saying?

- Ask questions about his condition and treatment
- Tries to stay calm and positive for his family
- Expresses gratitude to his family and medical team
- Admits to feeling Scared and overwhelmed at times

Thinks

What are their wants, needs, hopes, and dreams? What other thoughts might influence their behavior?

- I am worried about the progression of my liver disease and its potential impact on my life.
- I am concerned about the effectiveness of my current treatment and whether I will need more invasive procedures.
- I am constantly thinking about my diet and lifestyle choices and how they may be affecting my liver health.
- I am researching and seeking out information about my condition and treatment options to make informed decisions about my care.

Does

What behavior have we observed?
What can we imagine them doing?

- Medical equipment and machines.
- Needles and Injections.
- Medical Jargon and teminology.
- Beeping of machines and equipment.

Feels

What are their fears, frustrations, and anxieties? What other feelings might influence their behavior?

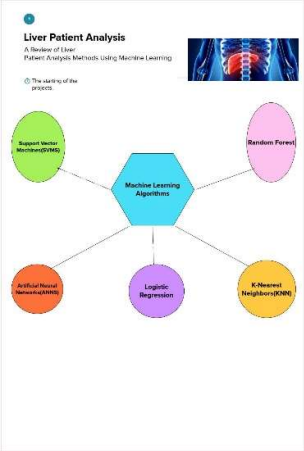
- Anxiety and fear about the unknowns of their condition and its impact on their life.
- Frustration with the limitations and restrictions that come with managing liver disease.
- A sense of isolation or loneliness due to the stigma surrounding liver disease and the need for lifestyle changes.
- Hopeful but cautions about potential treatments and outcomes.

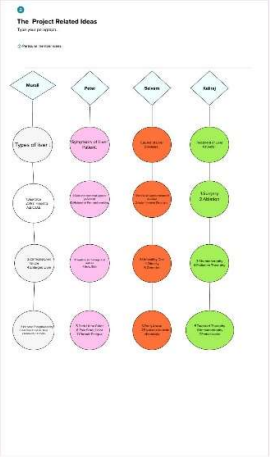
Need some inspiration?

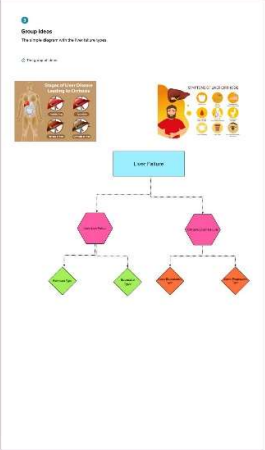
See a finished version of this template to kickstart your work.

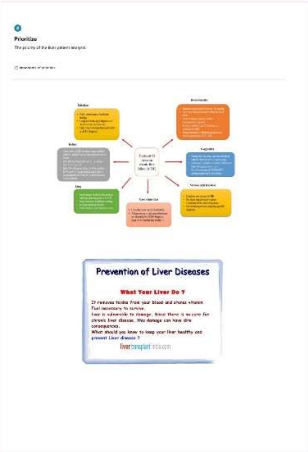
Open example →

2.2 Ideation & Brainstorming Map









3.RESULT:

3.1Data Model

Object name	Fields in the Object

Patient	

Field lable	Data type
Patient_ID	Text
Age	Number
Gender	Text
Total_Bilirubin	Number
Direct_Bilirubin	Number
Alkaline_Phosphatase	Number
Alamine_Aminotransferase	Number
Aspartate_Aminotransferase	Number

Patient	

Field lable	Data type
Total_Proteins	Number
Albumin	Number
Albumin and Globulin_Ratio	Number
Liver_Disease	boolean

3.2 Activity & Screenshot

Activity 1: Importing Libraries

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

```
In [2]: # for displaying all feature from dataset:
        pd.pandas.set_option('display.max_columns', None)
```

```
In [3]: # Reading Dataset:
        dataset = pd.read_csv("Dataset/Liver_data.csv")
        # Top 5 records:
        dataset.head()
```

Out[3]:

	Age	Gender	Total_Bilirubin	Direct_Bilirubin	Alkaline_Phosphotase	Alamine_Aminotransferase	Aspartate_Aminotransferase
0	65	Female	0.7	0.1	187	16	18
1	62	Male	10.9	5.5	699	64	100
2	62	Male	7.3	4.1	490	60	68
3	58	Male	1.0	0.4	182	14	20
4	72	Male	3.9	2.0	195	27	59

In [5]:
```
# Shape of dataset:
dataset.shape
```

Out[5]:
```
(583, 11)
```

In [6]:
```
# Cheaking Missing (NaN) Values:
dataset.isnull().sum()
```

Out[6]:
```
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.

In [7]:
```
# Mean & Median of "Albumin_and_Globulin_Ratio" feature:
print(dataset['Albumin_and_Globulin_Ratio'].median())
print(dataset['Albumin_and_Globulin_Ratio'].mean())
```

```
0.93
0.9470639032815197
```

In [8]:
```
# Filling NaN Values of "Albumin_and_Globulin_Ratio" feature with Median :
dataset['Albumin_and_Globulin_Ratio'] = dataset['Albumin_and_Globulin_Ratio'].fillna(dataset['Albumin_a
```

In [9]:
```
# Datatypes:
dataset.dtypes
```

Out[9]:
```
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
```

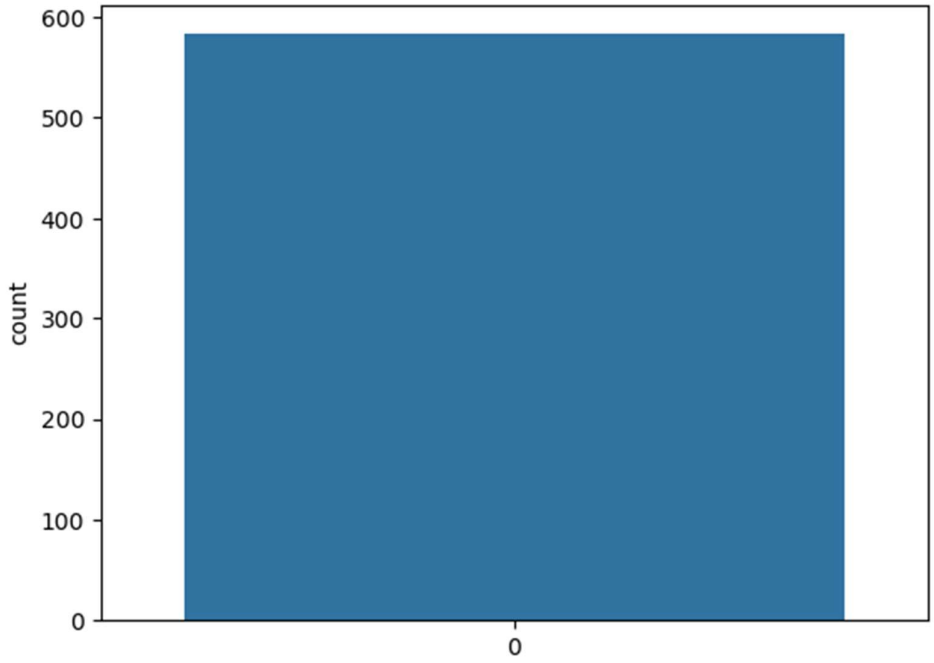
In [10]:
```
# Description:
dataset.describe()
```

Out[10]:

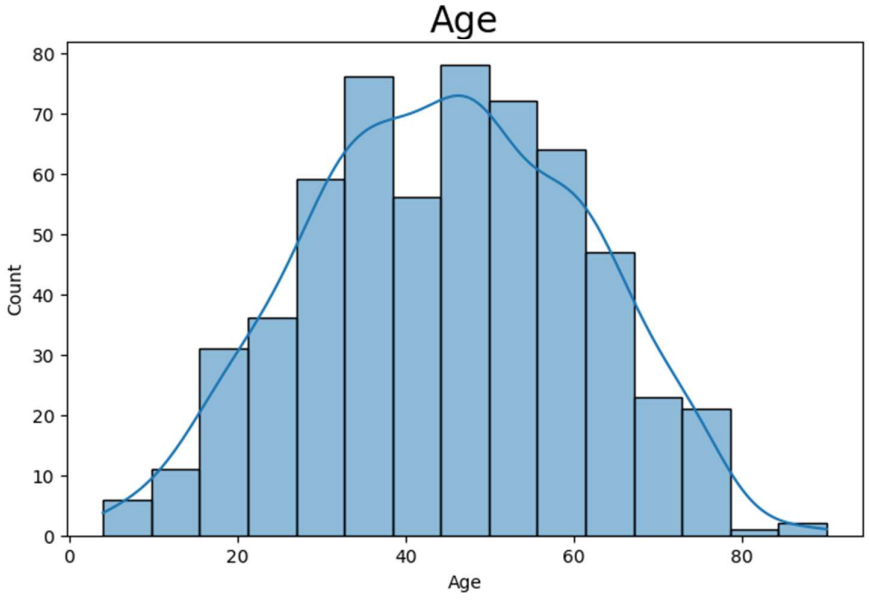
	Age	Total_Bilirubin	Direct_Bilirubin	Alkaline_Phosphotase	Alamine_Aminotransferase	Aspartate_Aminotransferas
count	583.000000	583.000000	583.000000	583.000000	583.000000	583.00000
mean	44.746141	3.298799	1.486106	290.576329	80.713551	109.91080
std	16.189833	6.209522	2.808498	242.937989	182.620356	288.91852
min	4.000000	0.400000	0.100000	63.000000	10.000000	10.00000
25%	33.000000	0.800000	0.200000	175.500000	23.000000	25.00000
50%	45.000000	1.000000	0.300000	208.000000	35.000000	42.00000
75%	58.000000	2.600000	1.300000	298.000000	60.500000	87.00000
max	90.000000	75.000000	19.700000	2110.000000	2000.000000	4929.00000

```
# 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
```

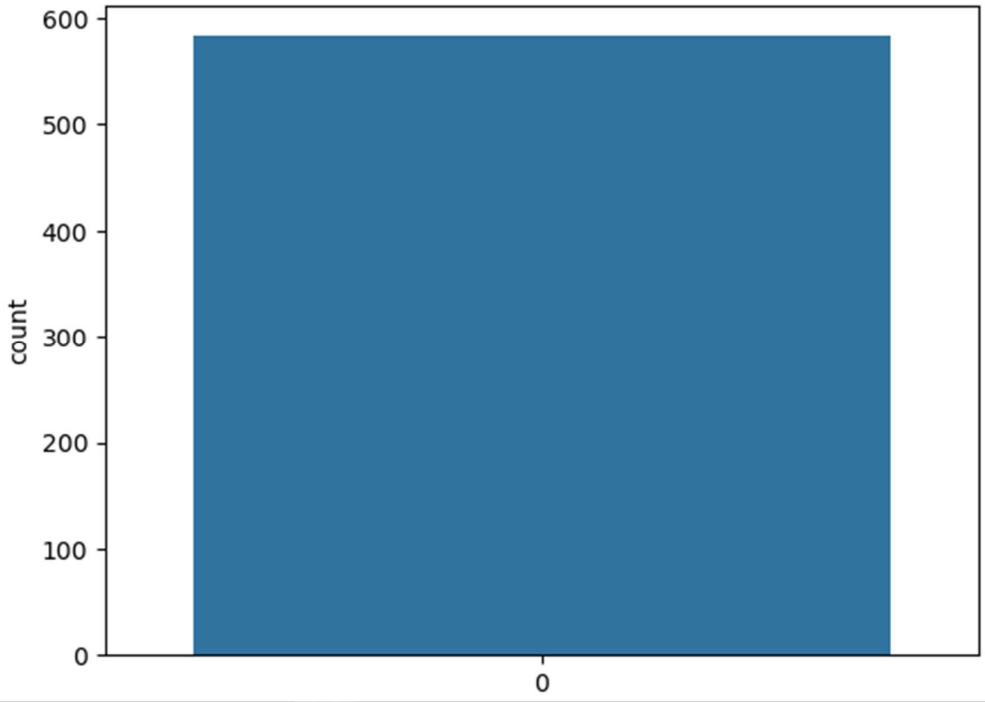


In [12]:
```
# Histrogram of Age:
plt.figure(figsize=(8,5))
sns.histplot(dataset['Age'], kde=True)
plt.title('Age', fontsize=20)
plt.show()
```



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

# Visualization:
dataset['Gender'] = dataset['Gender'].map({'Male': 0, 'Female': 1})
sns.countplot(dataset['Gender'])
plt.show()
```

```
Total Male   : 441
Total Female : 142
```




```
: # 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.

In [19]:
```
dataset.columns
```

Out[19]:
```
Index(['Age', 'Gender', 'Total_Bilirubin', 'Direct_Bilirubin',
       'Alkaline_Phosphotase', 'Alamine_Aminotransferase',
       'Aspartate_Aminotransferase', 'Total_Protiens', 'Albumin',
       'Albumin_and_Globulin_Ratio', 'Dataset'],
      dtype='object')
```

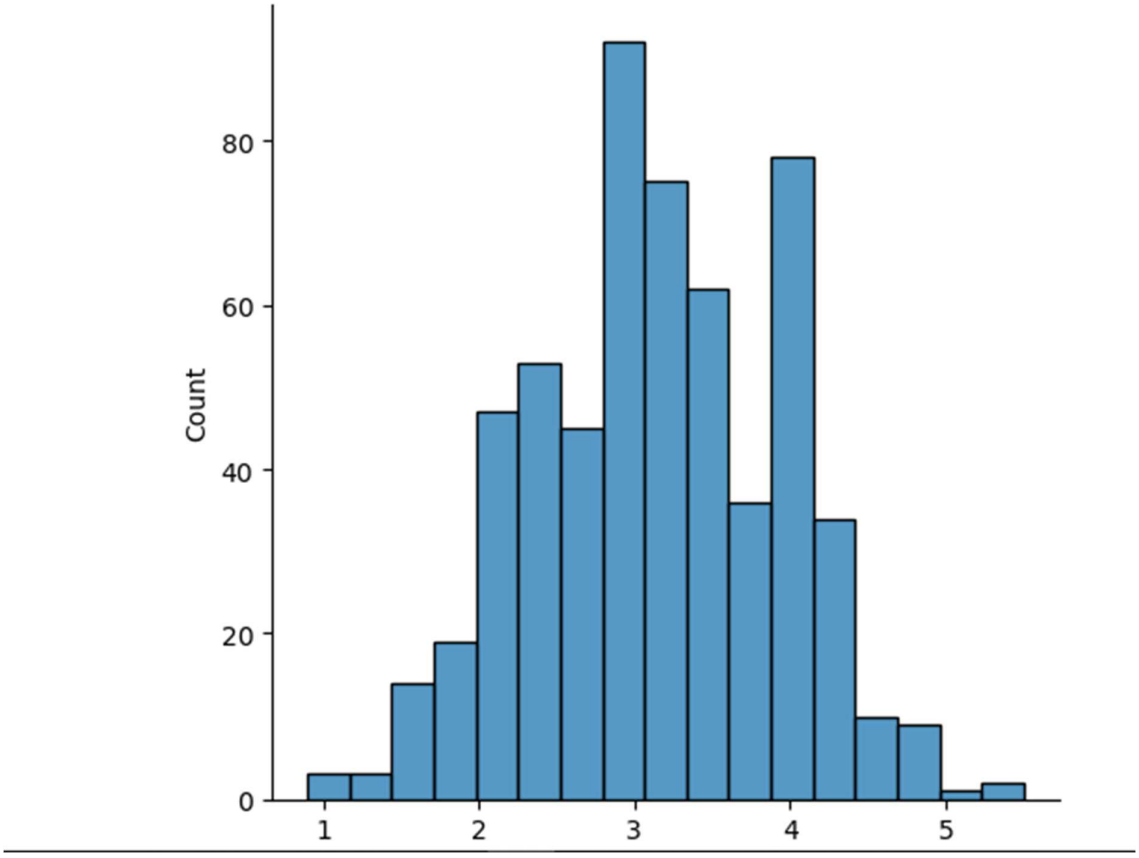
In [20]:
```
# Droping 'Direct_Bilirubin' feature:
dataset = dataset.drop('Direct_Bilirubin', axis=1)
```

In [21]:
```
dataset.columns
```

Out[21]:
```
Index(['Age', 'Gender', 'Total_Bilirubin', 'Alkaline_Phosphotase',
       'Alamine_Aminotransferase', 'Aspartate_Aminotransferase',
       'Total_Protiens', 'Albumin', 'Albumin_and_Globulin_Ratio', 'Dataset'],
      dtype='object')
```

In [22]:
```
import seaborn as sns

sns.displot(dataset['Albumin'])
```

Out[22]:
```
<seaborn.axisgrid.FacetGrid at 0x17b57561570>
```

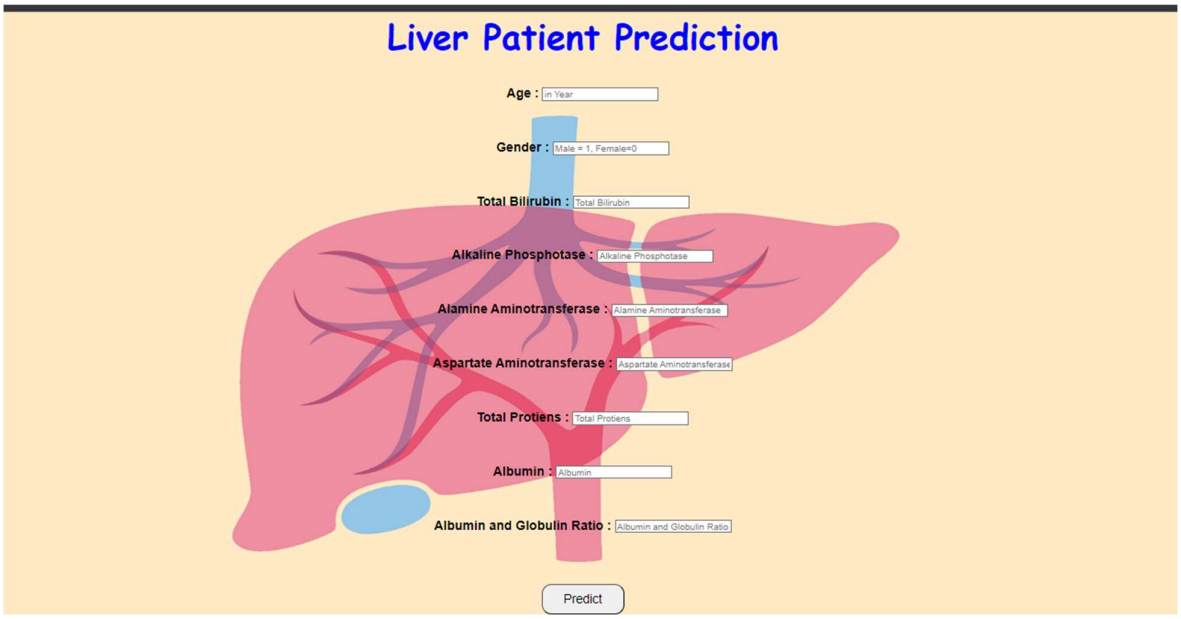


Activity : Flask app

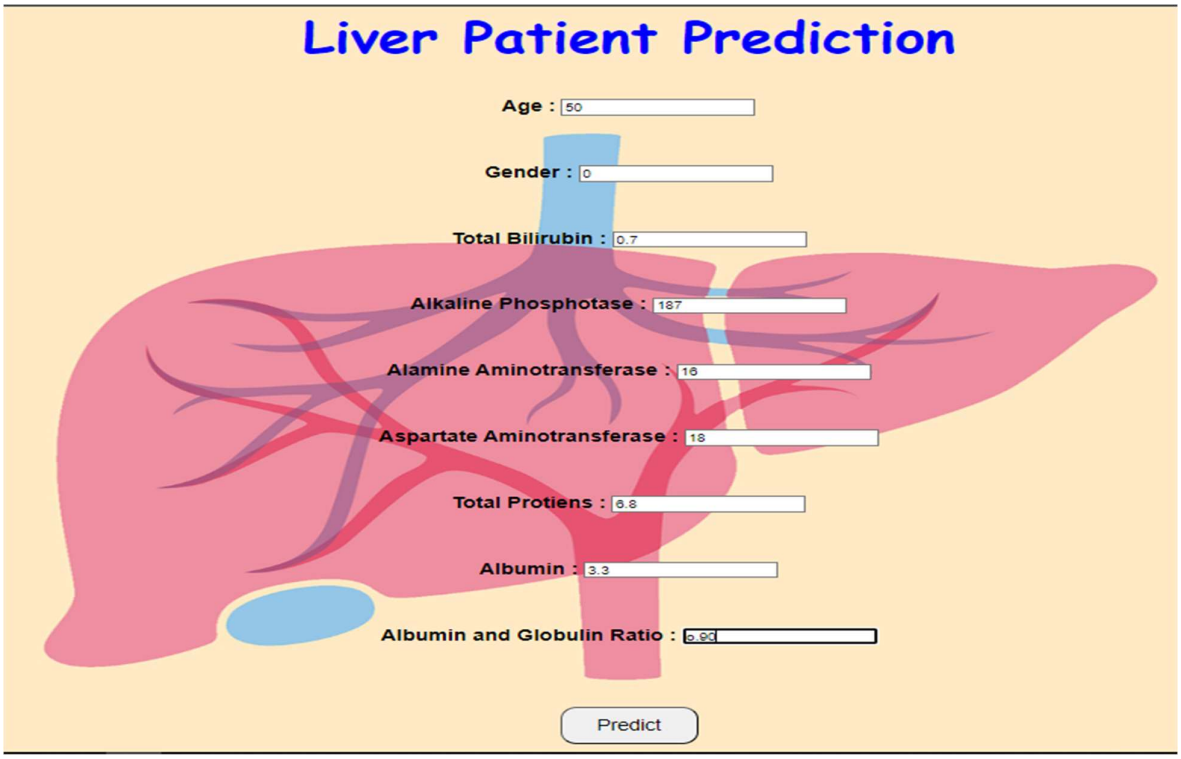
```
File Edit Search Source Run Debug Consoles Projects Tools View Help
C:\Users\ELCOT\Documents\Liver-Disease-Prediction-Project-main\app.py
app.py
```

```python
from flask import Flask, render_template, request
import numpy as np
import pickle


app = Flask(__name__)
model = pickle.load(open('Liver2.pkl', 'rb'))

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

@app.route("/predict", methods=['POST'])
def predict():
    if request.method == 'POST':
        Age = int(request.form['Age'])
        Gender = int(request.form['Gender'])
        Total_Bilirubin = float(request.form['Total_Bilirubin'])
        Alkaline_Phosphotase = int(request.form['Alkaline_Phosphotase'])
        Alamine_Aminotransferase = int(request.form['Alamine_Aminotransferase'])
        Aspartate_Aminotransferase = int(request.form['Aspartate_Aminotransferase'])
        Total_Protiens = float(request.form['Total_Protiens'])
        Albumin = float(request.form['Albumin'])
        Albumin_and_Globulin_Ratio = float(request.form['Albumin_and_Globulin_Ratio'])


        values = np.array([[Age,Gender,Total_Bilirubin,Alkaline_Phosphotase,Alamine_Aminotransferase,Aspartate_Aminotransferase,Total_Protiens,Albumin,
        prediction = model.predict(values)

        return render_template('result.html', prediction=prediction)


if __name__ == "__main__":
    app.run(debug=True)
```

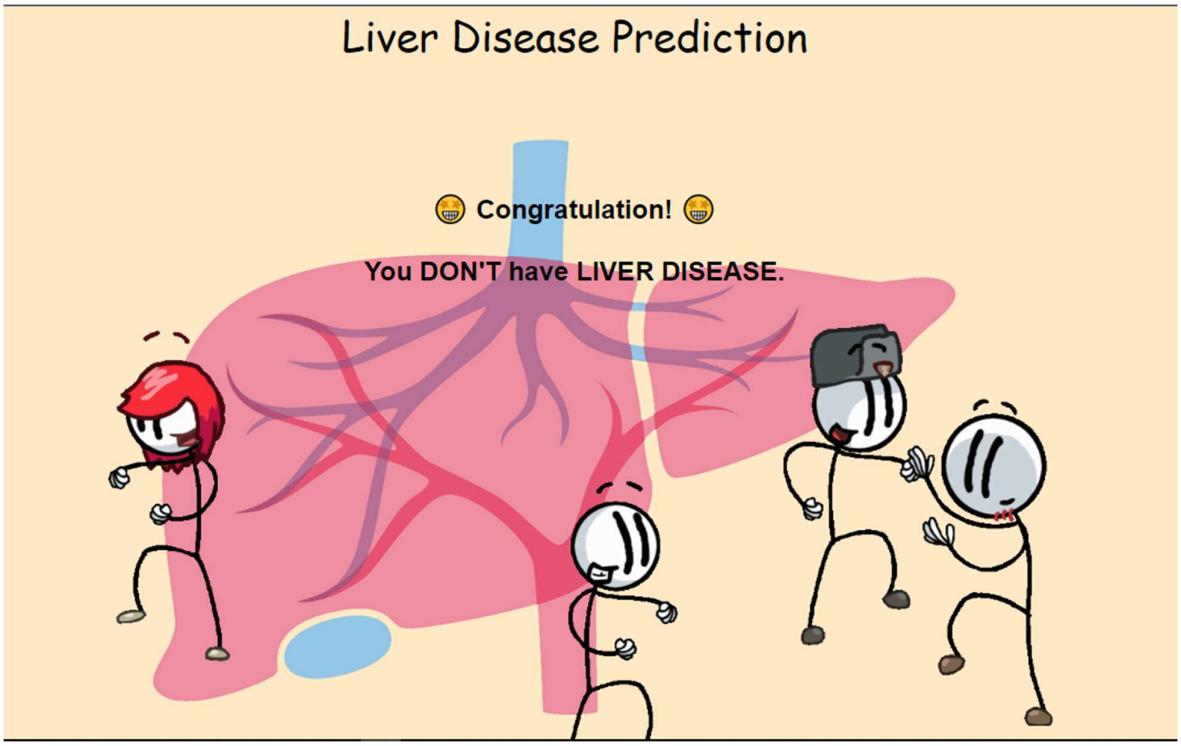
Result:



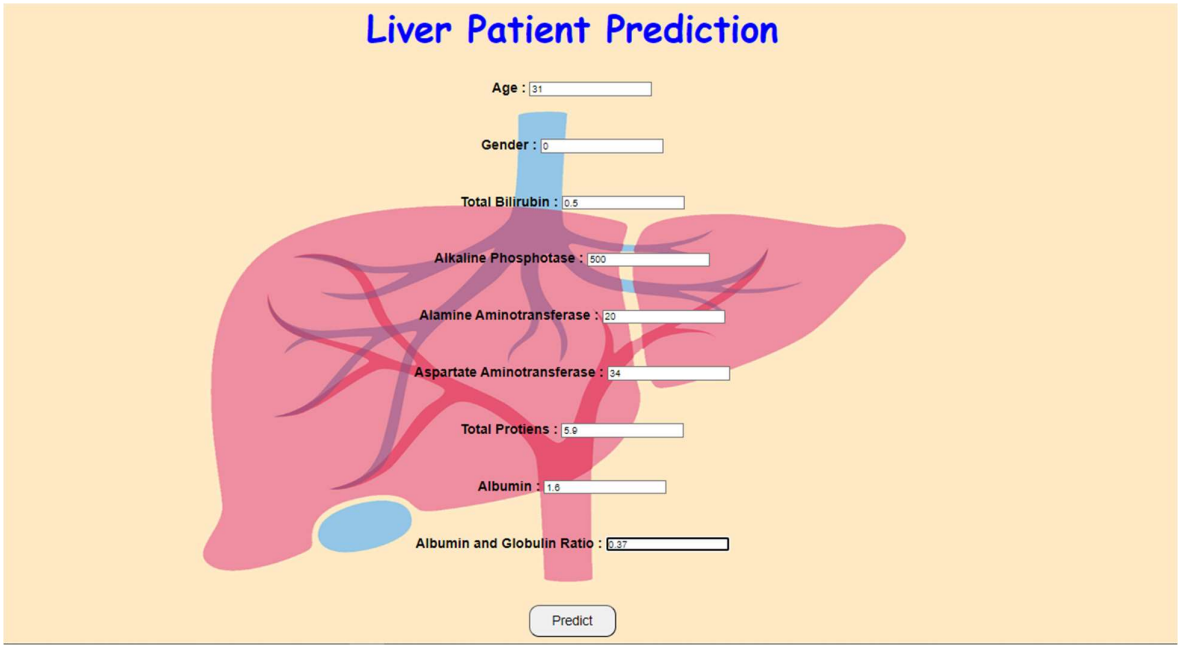
Enter the values in the box:



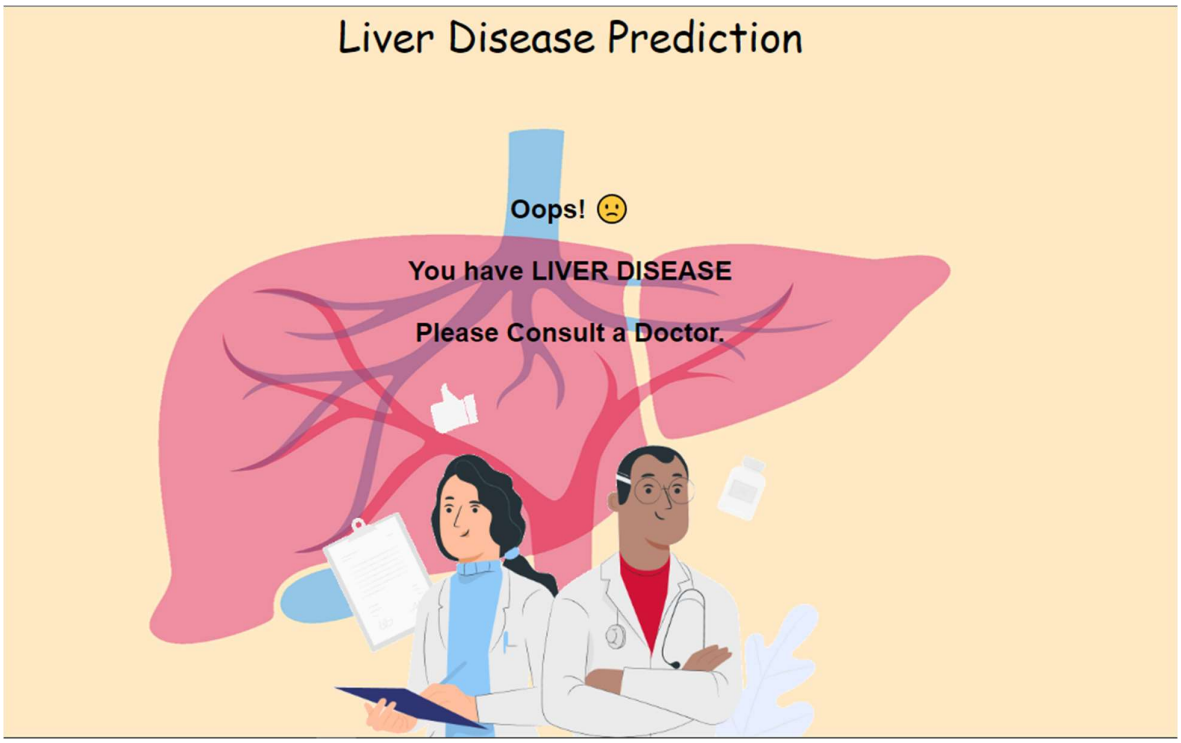
After the output:



Enter the value to fill the box:



After the output:



4.Trailhead Profile Public URL:

Team Lead - https://trailblazer.me/id/mdharan1

Team Member 1 - https://trailblazer.me/id/pjanor61

Team Member 2 - https://trailblazer.me/id/gselvamk

Team Member 3 - http://trailblazer.me/id/gkali22

5. Advantages & Disadvantages:

Advantages:

Early detection: Analysis of liver patients can help identify liver disease at an early stage, when it may be more treatable and may have a better prognosis.

Treatment planning: Analyzing liver patients can help doctors understand the severity of the disease, which can inform treatment planning and help them develop a personalized treatment plan for the patient.

Research: Analysis of liver patients can help researchers understand the underlying causes of liver disease, develop new treatments, and improve the overall understanding of liver function and disease.

Prevention: Analyzing liver patients can help identify risk factors for liver disease and inform prevention strategies to reduce the incidence of liver disease in the population.

Improved patient outcomes: By analyzing liver patients, doctors can monitor treatment effectiveness and adjust treatment plans as necessary to achieve better patient outcomes

Disadvantages:

False positive or false negative results: Liver patient analysis tests may give inaccurate results, which can lead to misdiagnosis or delayed treatment.

Limited information: Liver patient analysis can only provide information about the liver and its function, but it may not reveal the underlying cause of liver disease or other related health issues.

Invasive procedures: Some liver patient analysis tests, such as liver biopsy, can be invasive and may carry risks such as bleeding or infection.

Costly: Liver patient analysis tests can be expensive, especially if multiple tests are needed to diagnose and monitor liver disease.

Emotional distress: The process of undergoing liver patient analysis and waiting for results can be stressful and anxiety-provoking for patients.

Limitations in detecting early-stage liver disease: Some liver patient analysis tests may not be sensitive enough to detect early-stage liver disease, which can delay diagnosis and treatment.

6.Applications:

Diagnosis: Liver patient analysis can help diagnose liver diseases such as hepatitis, cirrhosis, and liver cancer. The analysis can detect abnormal liver function, the presence of liver enzymes, and the levels of bilirubin and albumin in the blood.

Monitoring: Patients with liver diseases need to be monitored closely to ensure that their condition does not worsen. Liver patient analysis can help doctors monitor a patient's liver function and identify any changes that may indicate a need for further treatment.

Treatment: Liver patient analysis can help doctors determine the best course of treatment for a patient. For example, a liver biopsy can be analyzed to determine the extent of liver damage and whether a patient is a candidate for a liver transplant.

Research: Liver patient analysis can help researchers study the causes and effects of liver diseases. By analyzing patient data, researchers can identify trends and patterns that can inform future research and treatment.

7.Feature Scope:

- **Patient demographics**
- **Medical history**
- **Liver function tests**
- **Imaging studies**
- **Lifestyle factors**
- **Environmental exposures**
- **Genetic factors**