

A Review of Liver Patient Analysis using Machine Learning

1.Introduction

1.1 OVERVIEW:

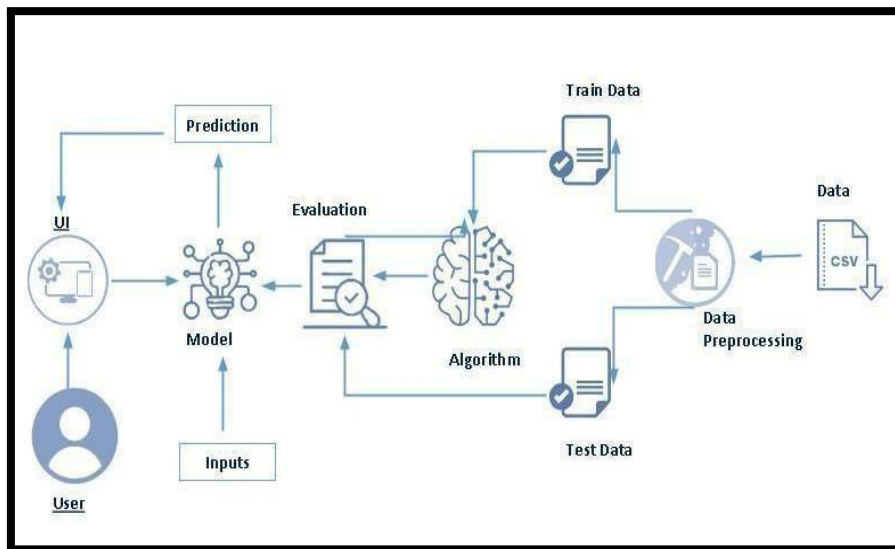
Liver diseases averts the normal function of the liver. This disease is caused by an assortment of elements that harm the liver. Diagnosis of liver infection at the preliminary stage is important for better treatment. In today's scenario devices like sensors are used for detection of infections.

Accurate classification techniques are required for automatic identification of disease samples. This disease diagnosis is very costly and complicated. Therefore, the goal of this work is to evaluate the performance of different Machine Learning algorithms in order to reduce the high cost of liver disease diagnosis.

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. In this project we will analyse the parameters of various classification algorithms and compare their predictive accuracies so as to find out the best classifier for determining the liver disease.

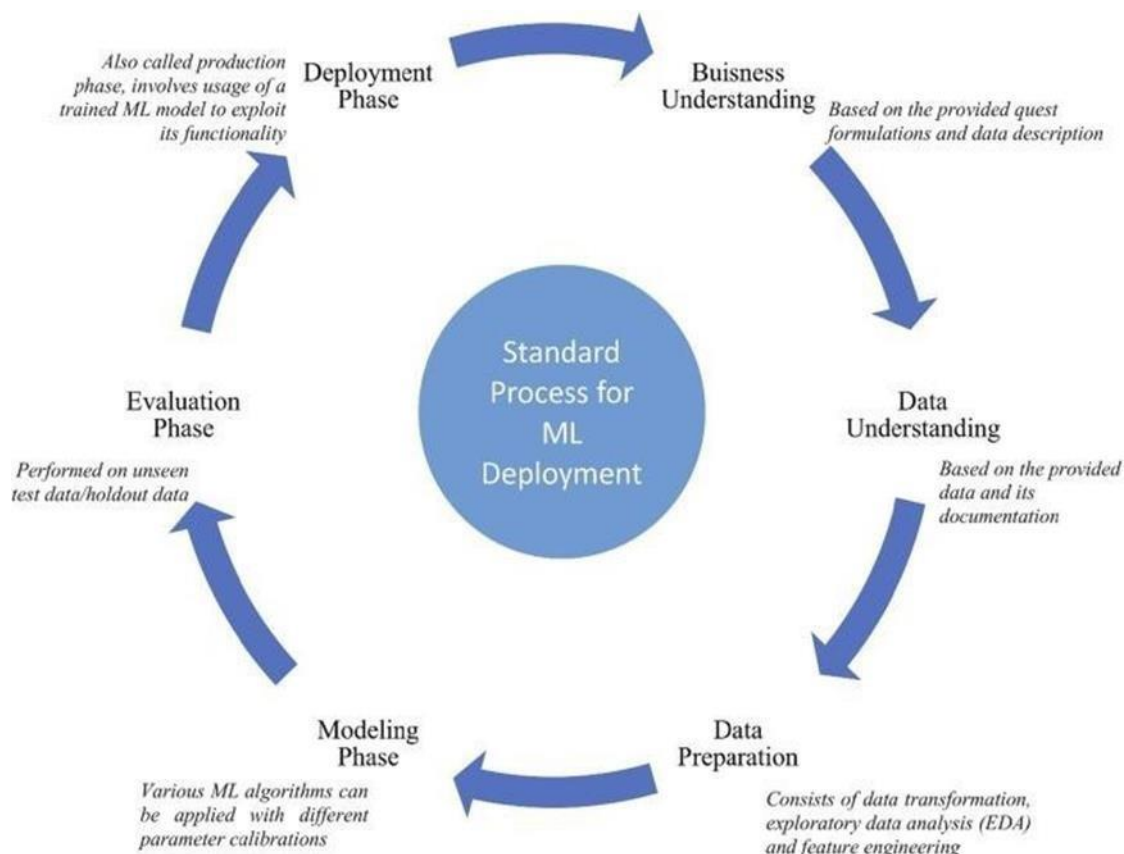
This project compares various classification algorithms such as Random Forest, Logistic Regression, KNN and ANN Algorithm with an aim to identify the best technique. Based on this study, Random Forest with the highest accuracy outperformed the other algorithms and can be further utilised JNH in the prediction of liver disease and can be recommended to the user.

Technical Architecture:



Technical about the Project:

Machine learning is mainly based on predictions made based on experience. It enables machines to make data-driven decisions, which is more efficient than explicitly programming to carry out certain tasks. These algorithms are designed in a fashion that gives exposure to new data that can help organisations learn and improve these strategies.



A PROJECT DESCRIPTION:

The liver is an organ about the size of a football. It sits just under your rib cage on the right side of your abdomen. The liver is essential for digesting food and ridding your body of toxic substances. Liver disease can be inherited (genetic). Liver problems can also be caused by a variety of factors that damage the liver, such as

viruses, alcohol use and obesity. Over time, conditions that damage the liver can lead to scarring (cirrhosis), which can lead to liver failure, a life-threatening condition. But early treatment may give the liver time to heal. Liver disease doesn't always cause noticeable signs and symptoms. If signs and symptoms of liver disease do occur, they may include: Skin and eyes that appear yellowish (jaundice) Abdominal pain and swelling 3Swelling in the legs and ankles Itchy skin Dark urine color Pale stool color Chronic fatigue Nausea or vomiting Loss of appetite Tendency to bruise easily

Project Flow:

- User interacts with the UI to enter the input
- Entered input is analysed by the model which is integrated.
 - Once model analyses the input the prediction is

showcased on the UI To accomplish this, we have to

complete all the activities listed below.

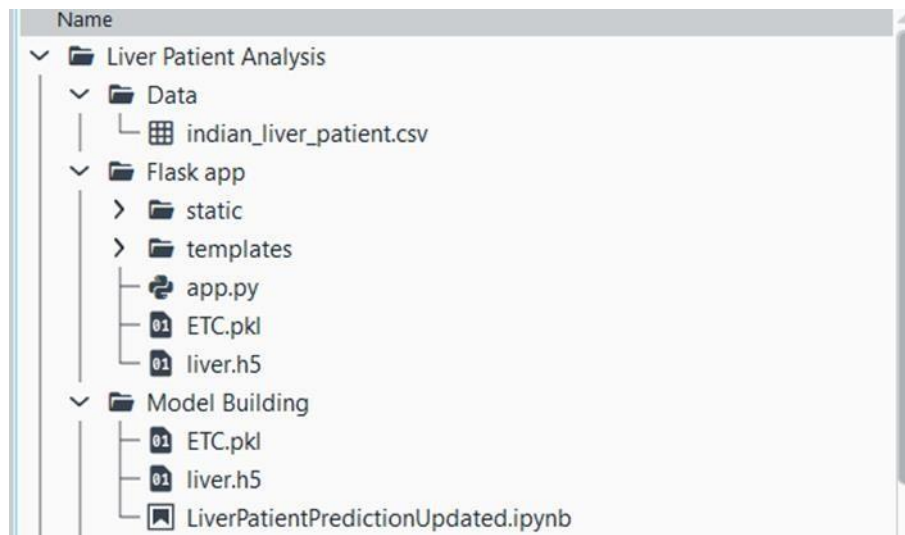
Define Problem / Problem Understanding

- Specify the business problem
- Business requirements
- Literature Survey
- Social or Business Impact.
- Data Collection & Preparation
 - Collect the dataset
 - Data Preparation
- Exploratory Data Analysis
 - Descriptive statistical
 - Visual Analysis
- Model Building
 - Training the model in multiple algorithms
 - Testing the model
- Performance Testing & Hyperparameter Tuning
 - Testing model with multiple evaluation metrics
 - Comparing model accuracy before & after applying hyperparameter tuning
- Model Deployment
 - Save the best model
 - Integrate with Web Framework
- Project Demonstration & Documentation
 - Record explanation Video for project end to end solution

- Project Documentation-Step by step project development procedure

Project Structure:

Create the Project folder which contains files as shown below



- Training folder contains a model training file.
- We are building a flask application which needs HTML pages stored in the templates folder and a python script app.py for scripting.
- ETC.pkl is our saved model. Further we will use this model for flask integration.

1.2 Purpose

The liver is essential for digesting food and ridding your body of toxic substances. Liver disease can be inherited (genetic). Liver problems can also be caused by a variety of factors that damage the liver, such as viruses, alcohol use and obesity.

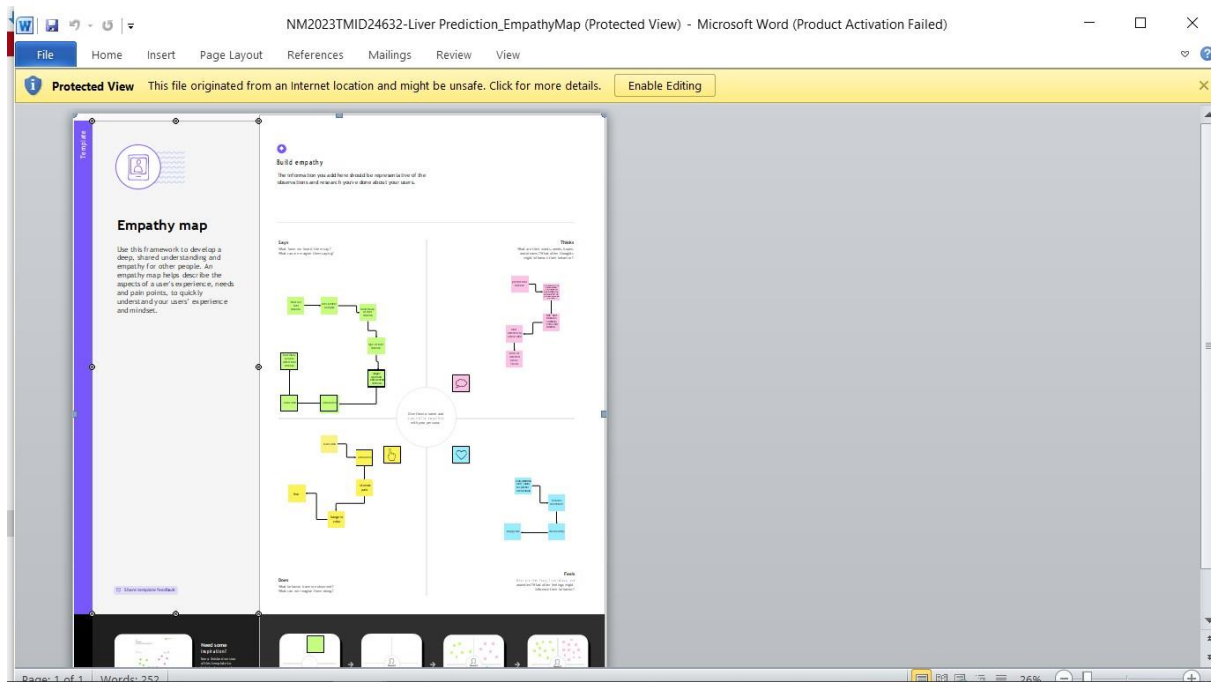
The liver helps your body maintain a healthy level of blood sugar. Your liver supplies glucose to your blood when it's needed. It also removes glucose from your blood when there's too much.

2.Problem Definition and Thinking:

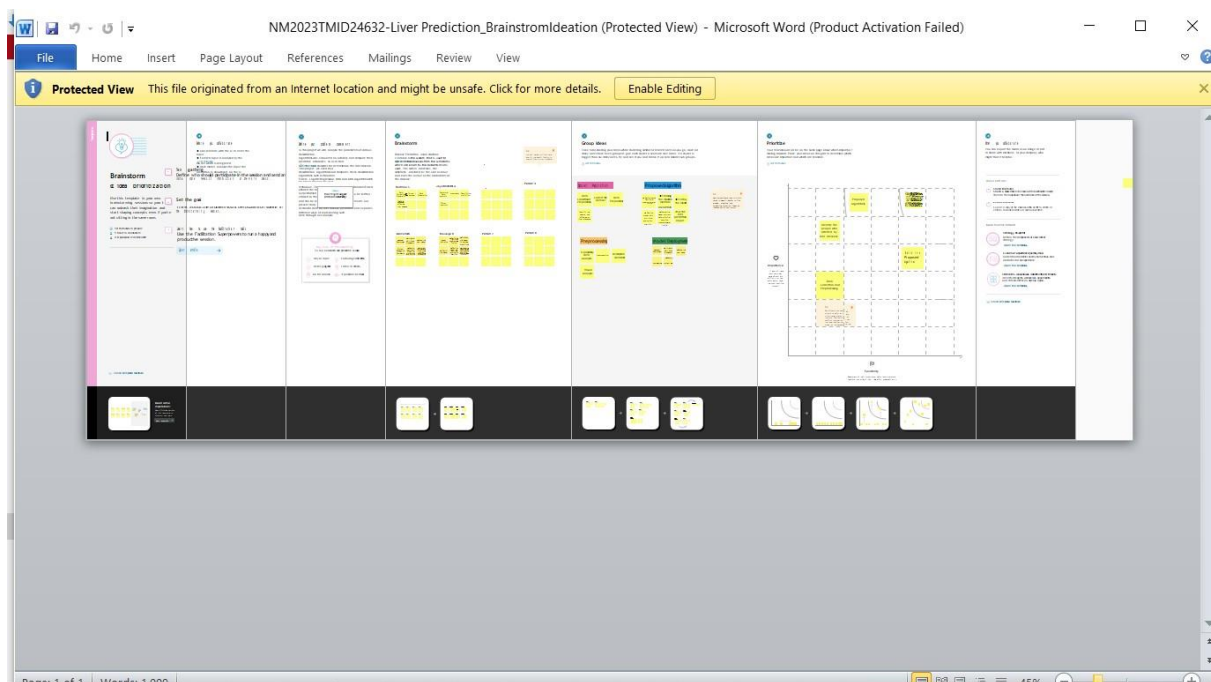
Machine learning has become an increasingly popular tool in recent years ,given its ability to automatically detect patterns in data and make predictions about future events .this can be extremely useful for asking decisions in a wide range of

domains, from financial trading to medical diagnoses Empathy Map.

2.1 Empathy Map



2.2 Ideation & Brainstorming Map:



3. RESULT

Final findings(output) of the project along with screenshots

/C:/Users/ELCOT/Downloads/goto% X

+

←

→

↺

file:///C:/Users/ELCOT/Downloads/goto predict.html

☆

🔒

📄

☰

Liver Patient Prediction

Age:

Gender:

Total_Bilirubin:

Alkaline_phosphatase:

Aspartate_Aminotransferase:

Albumin:

Direct_Bilirubin:

Alamine_Aminotransferase:

/C:/Users/ELCOT/Downloads/goto% X

+

←

→

↺

file:///C:/Users/ELCOT/Downloads/goto predict.html

☆

🔒

📄

☰

Total_Bilirubin:

Alkaline_phosphatase:

Aspartate_Aminotransferase:

Albumin:

Direct_Bilirubin:

Alamine_Aminotransferase:

Total_Protiens:

Albumin_and_Globulin_Ratio:

predict



Liver Patient Prediction

You have a liver disease problem, You must and should consult a doctor. Take care

4. ADVANTAGES AND DISADVANTAGES

List of advantages and disadvantages of the proposed solution

Advantages:

1. It's nutrient dense:

Rich in protein, low in calories and packed with essential vitamins and minerals, liver is one of the most nutrient-dense foods available. What makes it even more of a 'superfood' is that the nutrients are easily accessible by the body.

2. Rich source of vitamin A:

Liver and liver products such as pâté are a rich source of vitamin A in the form of retinol, which is the 'active' form. Muscle meat is not such an impressive source and plant sources have to be converted to the active form in the gut, so that the body can use them.

3. Valuable dietary source of vitamin D:

There are a limited number of food sources of vitamin D, with liver being one. Vitamin D, known as the sunshine vitamin, is typically formed by the action of sunlight on the skin and is difficult to obtain in adequate amounts from our diets. With low levels of sun light during the autumn and winter months, it's common for levels to be low.

The form of vitamin D supplied by liver is D3, the more 'active' form, which once again is found in much lower levels in muscle meat.

4. Useful source of B vitamins:

Liver is an impressive source of B vitamins including folate, choline and vitamin B12. These nutrients are needed for a number of functions in the body and especially for metabolism.

5. Valuable source of iron:

Rich in iron, liver supplies the form known as haem, which is especially helpful for supporting adequate iron stores. Maintaining these levels is particularly relevant for menstruating women who are commonly low in this important mineral.

Disadvantages:

The same vitamins and minerals that make liver so nutritionally potent can create complications for people with certain medical conditions. That's why you should check with your doctor before you start eating a lot more of it.

Here are some possible drawbacks of eating large amounts of liver:

High cholesterol: Liver is high in dietary cholesterol. While many people can eat high-cholesterol foods without a problem, people trying to lower their

cholesterol or who take cholesterol medications should watch their intake. Getting too much cholesterol can increase your risk of heart disease.

Vitamin A toxicity: It's possible, and dangerous, to get too much vitamin A. Eating large amounts of liver can lead to symptoms of vitamin A toxicity, which happens when your own liver can't process the excess vitamin A quickly enough. Most doctors recommend that people without vitamin deficiencies eat just one serving of liver per week.

Medication interactions: Some medications are known to interact with vitamin A. Since liver is rich in vitamin A, anyone on medications such as Orlistat or certain psoriasis medications should talk to their doctor before adding liver to their diet.

5. APPLICATIONS

The areas where this solution can be applied

The liver filters all of the blood in the body and breaks down poisonous substances, such as alcohol and drugs. The liver also produces bile, a fluid that helps digest fats and carry away waste.

Functions of the Liver

The liver is an essential organ of the body that performs over 500 vital functions. These include removing waste products and foreign substances from the bloodstream, regulating blood sugar levels, and creating essential nutrients. Here are some of its most important functions:

Albumin Production: Albumin is a protein that keeps fluids in the bloodstream from leaking into surrounding tissue. It also carries hormones, vitamins, and enzymes through the body.

Bile Production: Bile is a fluid that is critical to the digestion and absorption of fats in the small intestine.

Filters Blood: All the blood leaving the stomach and intestines passes through the liver, which removes toxins, byproducts, and other harmful substances.

Regulates Amino Acids: The production of proteins depend on amino acids. The liver makes sure amino acid levels in the bloodstream remain healthy.

Regulates Blood Clotting: Blood clotting coagulants are created using vitamin K, which can only be absorbed with the help of bile, a fluid the liver produces.

Resists Infections: As part of the filtering process, the liver also removes bacteria from the bloodstream.

Stores Vitamins and Minerals: The liver stores significant amounts of vitamins A, D, E, K, and B12, as well as iron and copper.

Processes Glucose: The liver removes excess glucose (sugar) from the bloodstream and stores it as glycogen. As needed, it can convert glycogen back into glucose.

6.CONCLUSION

Conclusion summarizing the entire work and findings

The main roles of the liver include removing toxins, processing food nutrients and regulating body metabolism. Important causes of liver disorders are fatty liver, hepatitis virus infections and alcohol. Cirrhosis (liver scarring), the end-result of many liver disorders, can lead to liver failure. On this page Functions of the liver Symptoms of liver disease Causes of liver disease Disorders of the liver Complications of liver disease Diagnosis of liver disease Treatment for liver disease Where to get help The liver is on the upper right side of the abdomen, just below the diaphragm. It is the largest internal organ of the human body and weighs around 1.5kg in the average adult. Blood from the digestive system must first filter through the liver before it travels anywhere else in the body.

The main roles of the liver include:

*Removing toxins from the body

- *Processing food nutrients

- *Helping to regulate body metabolism.

Conditions that can prevent the liver from performing its vital functions include:

- *Fat accumulation

- *Alcohol misuse

- *Viral infection

- *Iron or copper accumulation

- *Toxic damage

- *Cancer.

The most common cause of liver disease is non-alcoholic fatty liver disease ('fatty liver'). Cirrhosis is the end-result of many liver conditions. It involves severe scarring of the liver (with liver nodule formation). Cirrhosis is associated with a progressive decline in liver function resulting in liver failure.

7.FUTURE SCOPE

Enhancements that can be made in the future:

To diagnose you, your healthcare provider will ask about your health history. He or she will also give you a physical exam. You will also need some tests.

You may have one or more of these tests:

Cholecystography. This is also called oral cholecystography or a gallbladder series. A series of X-rays are taken of the gallbladder after you swallow a Endoscopic retrograde cholangiopancreatography (ERCP).

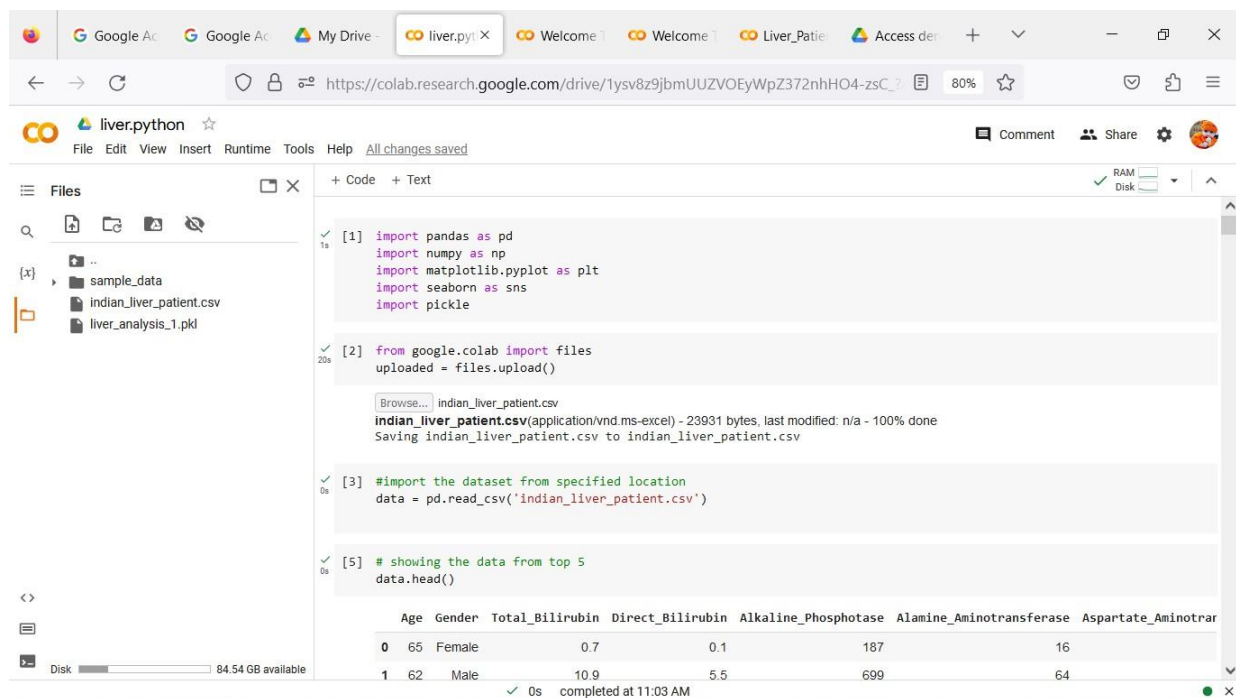
This is a procedure that helps diagnose and treat problems in the liver, gallbladder, bile ducts, and pancreas. It uses X-rays and a long, flexible, lighted tube (endoscope). The scope is put into your mouth and throat.

special contrast dye. This test can show gallstones, inflammation of the gallbladder (cholecystitis), and other problems. But this test isn't used often CT scan.

This is a test that uses X-rays and a computer to make detailed images of the body. A CT scan shows details of the bones, muscles, fat, and organs. CT scans are more detailed than general X-rays.

8.APPENDIX

Source Code



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

[2] from google.colab import files
uploaded = files.upload()

Browse... indian_liver_patient.csv
indian_liver_patient.csv(application/vnd.ms-excel) - 23931 bytes, last modified: n/a - 100% done
Saving indian_liver_patient.csv to indian_liver_patient.csv

[3] #import the dataset from specified location
data = pd.read_csv('indian_liver_patient.csv')

[5] # showing the data from top 5
data.head()
```

	Age	Gender	Total_Bilirubin	Direct_Bilirubin	Alkaline_Phosphotase	Alamine_Aminotransferase	Aspartate_Aminotransferase
0	65	Female	0.7	0.1	187		16
1	62	Male	10.9	5.5	699		64

completed at 11:03 AM

Google Chrome browser window showing a Google Colab notebook titled "liver.pyt". The address bar shows the URL: https://colab.research.google.com/drive/1ysv8z9jbmUUZVOEyWpZ372nhHO4-zsC_?. The notebook interface includes a menu bar (File, Edit, View, Insert, Runtime, Tools, Help) and a toolbar with icons for file operations and runtime status. The left sidebar shows the file explorer with a folder named "sample_data" containing files "indian_liver_patient.csv" and "liver_analysis_1.pkl". The main code area displays the following code cells:

```
[6]   
[7] data.describe()  
[8] data.info()
```

The output of cell [7] shows a summary statistics table:

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

The output of cell [8] shows the data structure:

```
Out[8]:   
Range: 0 to 582 (583 rows)  
Columns: 7 (Age, Total_Bilirubin, Direct_Bilirubin, Alkaline_Phosphotase, Alamine_Aminotransferase, Aspartate_Aminotransferase, Gender)  
dtypes: object(1), int64(6)
```

The notebook status bar indicates "completed at 11:03 AM".

Google Chrome browser window showing the same Google Colab notebook. The code cells are:

```
[5]   
[6] data.tail()
```

The output of cell [5] shows a subset of the data:

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

The output of cell [6] shows the last five rows of the data:

	Age	Gender	Total_Bilirubin	Direct_Bilirubin	Alkaline_Phosphotase	Alamine_Aminotransferase	Aspartate_Aminotransferase
578	60	Male	0.5	0.1	500	20	20
579	40	Male	0.6	0.1	98	35	35
580	52	Male	0.8	0.2	245	48	48
581	31	Male	1.3	0.5	184	29	29
582	38	Male	1.0	0.3	216	21	21

The notebook status bar indicates "completed at 11:03 AM".

liver.python

File Edit View Insert Runtime Tools Help All changes saved

Files

- sample_data
 - indian_liver_patient.csv
 - liver_analysis_1.pkl

Code

```
[8] data.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_Protiens        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
dtypes: float64(5), int64(5), object(1)
memory usage: 50.2+ KB
```

```
[9] data.isnull().any()
```

```
Age                False
Gender             False
Total_Bilirubin    False
Direct_Bilirubin   False
Alkaline_Phosphotase False
Alamine_Aminotransferase False
Aspartate_Aminotransferase False
Total_Protiens     False
Albumin            False
Albumin_and_Globulin_Ratio False
Dataset            False
dtype: bool
```

completed at 11:03 AM

liver.python

File Edit View Insert Runtime Tools Help All changes saved

Files

- sample_data
 - indian_liver_patient.csv
 - liver_analysis_1.pkl

Code

```
[9] data.isnull().any()
```

```
Alkaline_Phosphotase False
Alamine_Aminotransferase False
Aspartate_Aminotransferase False
Total_Protiens        False
Albumin               False
Albumin_and_Globulin_Ratio True
Dataset               False
dtype: bool
```

```
[10] data.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
```

```
[ ]
```

```
[11] data[data['Dataset']==1]
```

completed at 11:03 AM

Google Ac Google Ac My Drive - liver.pyt Welcome Welcome Liver_Patie Access der + -

← → ↺ https://colab.research.google.com/drive/1ysv8z9jbmUUZVOEyWpZ372nhHO4-zsC_? 80% ☆

liver.python ☆ File Edit View Insert Runtime Tools Help All changes saved Comment Share

Files + Code + Text RAM Disk

sample_data indian_liver_patient.csv liver_analysis_1.pkl

```
[11] data[data['Dataset']==1]
```

	Age	Gender	Total_Bilirubin	Direct_Bilirubin	Alkaline_Phosphotase	Alamine_Aminotransferase	Aspartate_Aminotr
0	65	Female	0.7	0.1	187		16
1	62	Male	10.9	5.5	699		64
2	62	Male	7.3	4.1	490		60
3	58	Male	1.0	0.4	182		14
4	72	Male	3.9	2.0	195		27
...
576	32	Male	15.0	8.2	289		58
577	32	Male	12.7	8.4	190		28
579	40	Male	0.6	0.1	98		35
580	52	Male	0.8	0.2	245		48
581	31	Male	1.3	0.5	184		29

416 rows × 11 columns

0s completed at 11:03 AM

Google Ac Google Ac My Drive - liver.pyt Welcome Welcome Liver_Patie Access der + -

← → ↺ https://colab.research.google.com/drive/1ysv8z9jbmUUZVOEyWpZ372nhHO4-zsC_? 80% ☆

liver.python ☆ File Edit View Insert Runtime Tools Help All changes saved Comment Share

Files + Code + Text RAM Disk

sample_data indian_liver_patient.csv liver_analysis_1.pkl

```
[12] data['Dataset'].unique()
array([1, 2])

[16] # checking for missing data
data.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

[ ] #mode imputation
#data['Albumin_and_Globulin_Ratio'].fillna(data['Albumin_and_Globulin_Ratio'].mode()[0])

[14] data_1 = data.dropna()

0s completed at 11:03 AM
```

Google Ac Google Ac My Drive - liver.pyt Welcome Welcome Liver_Patie Access der + -

https://colab.research.google.com/drive/1ysv8z9jbmUUZVOEyWpZ372nhHO4-zsC_? 80%

liver.python

File Edit View Insert Runtime Tools Help All changes saved

Files

- sample_data
 - indian_liver_patient.csv
 - liver_analysis_1.pkl

+ Code + Text

```
[15] # checking for the missing data after cleaning data
#data['Albumin_and_Globulin_Ratio'].fillna(data['Albumin_and_Globulin_Ratio'].mode()[0])
data.isnull().sum()

Age                0
Gender              0
Total_Bilirubin     0
Direct_Bilirubin    0
Alkaline_Phosphatase 0
Alamine_Aminotransferase 0
Aspartate_Aminotransferase 0
Total_Protiens      0
Albumin             0
Albumin_and_Globulin_Ratio 4
Dataset             0
dtype: int64

[ ]

[17] plt.figure(figsize=(15,10))
plt.subplot(3,3,1)
plt.scatter(data_1['Age'], data_1['Dataset'],)
plt.ylabel('Dataset')
plt.xlabel('Age')

plt.subplot(3,3,2)
```

0s completed at 11:03 AM

Google Ac Google Ac My Drive - liver.pyt Welcome Welcome Liver_Patie Access der + -

https://colab.research.google.com/drive/1ysv8z9jbmUUZVOEyWpZ372nhHO4-zsC_? 80%

liver.python

File Edit View Insert Runtime Tools Help All changes saved

Files

- sample_data
 - indian_liver_patient.csv
 - liver_analysis_1.pkl

+ Code + Text

```
[17] plt.xlabel('Age')

plt.subplot(3,3,2)
plt.scatter(data_1['Gender'], data_1['Dataset'],)
plt.ylabel('Dataset')
plt.xlabel('Gender')

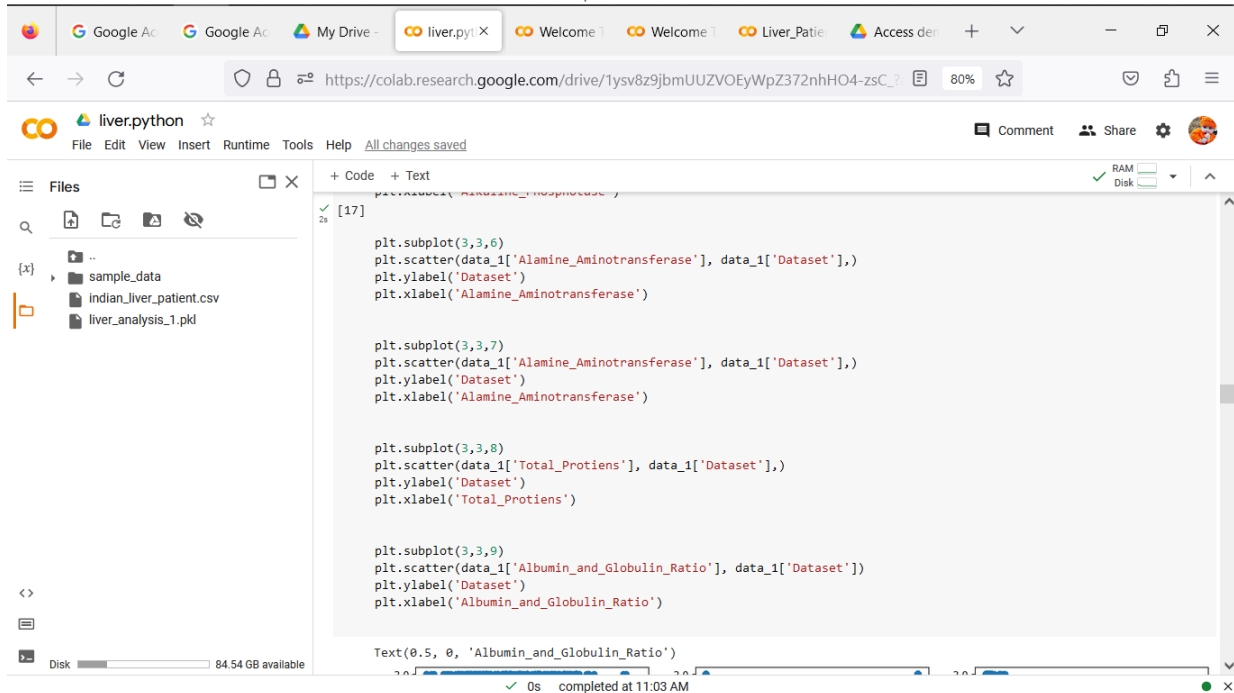
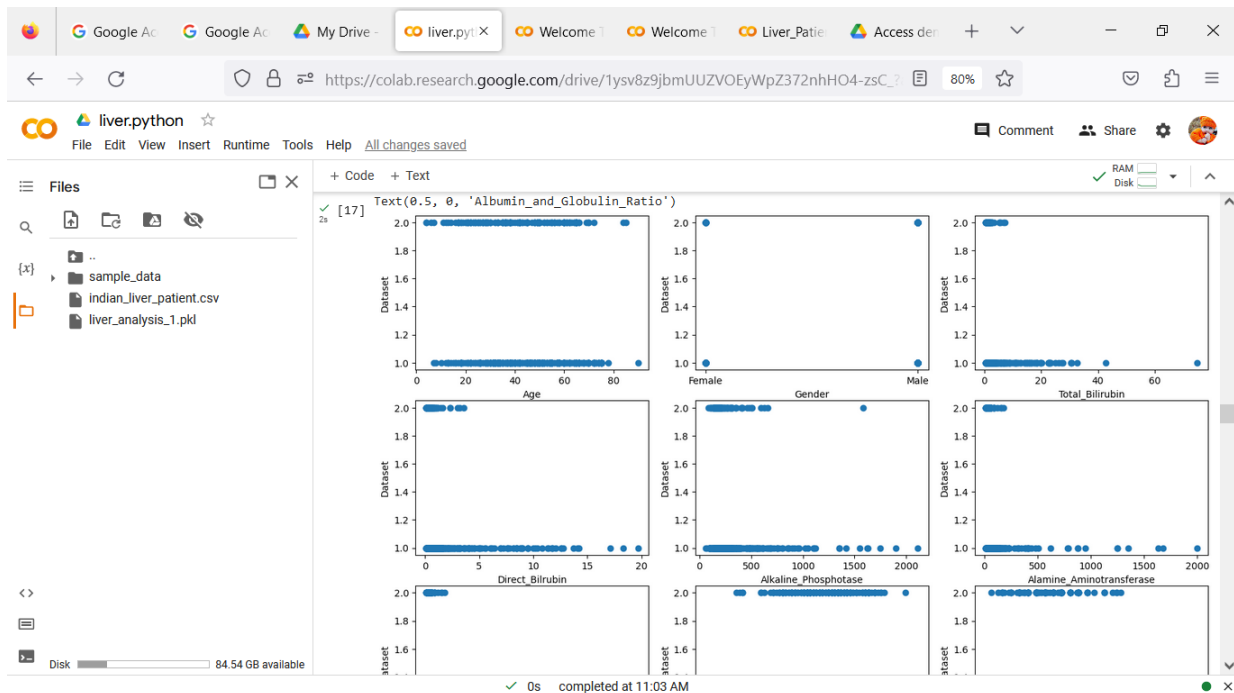
plt.subplot(3,3,3)
plt.scatter(data_1['Total_Bilirubin'], data_1['Dataset'],)
plt.ylabel('Dataset')
plt.xlabel('Total_Bilirubin')

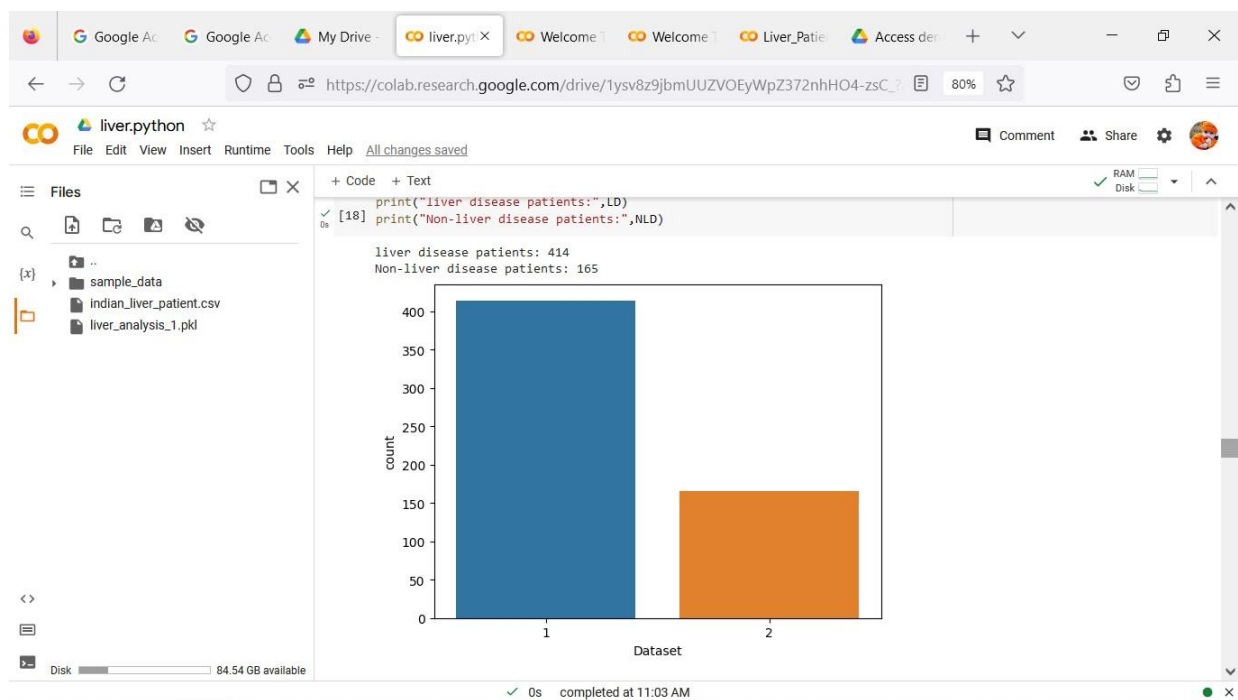
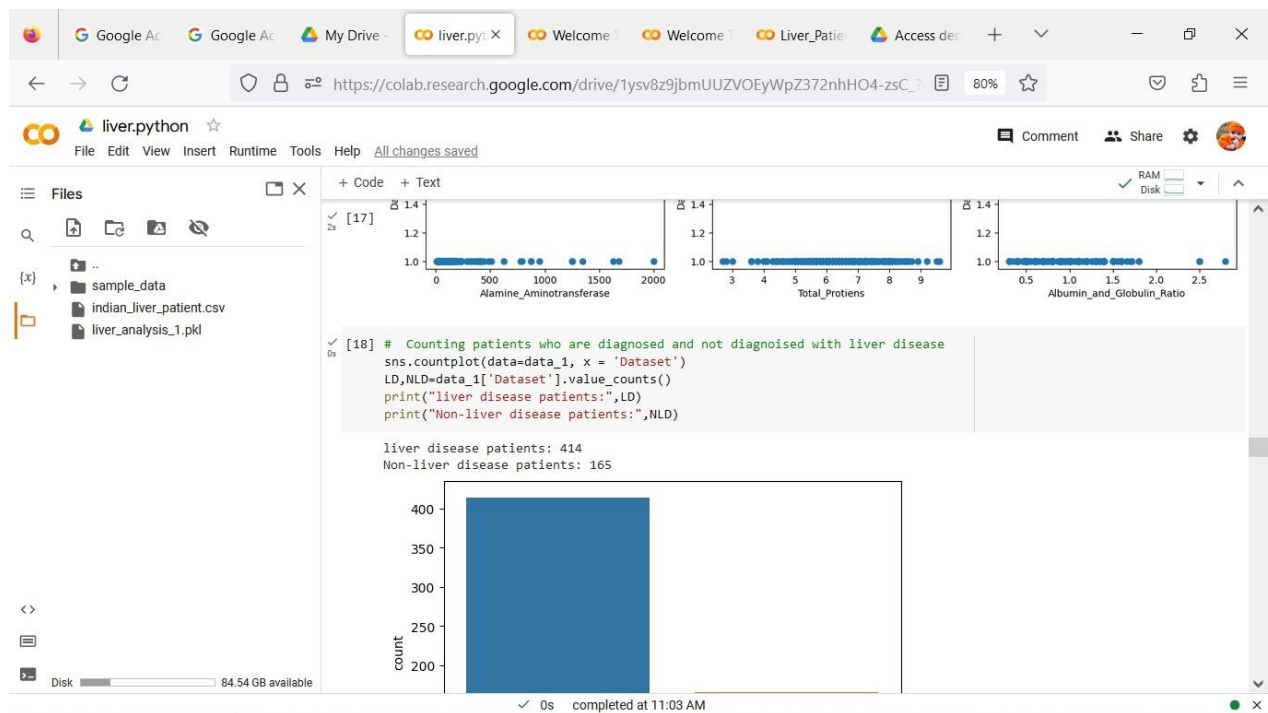
plt.subplot(3,3,4)
plt.scatter(data_1['Direct_Bilirubin'], data_1['Dataset'],)
plt.ylabel('Dataset')
plt.xlabel('Direct_Bilirubin')

plt.subplot(3,3,5)
plt.scatter(data_1['Alkaline_Phosphatase'], data_1['Dataset'],)
plt.ylabel('Dataset')
plt.xlabel('Alkaline_Phosphatase')

plt.subplot(3,3,6)
```

0s completed at 11:03 AM





Google Ac

Google Ac

My Drive

liver.pytX

Welcome

Welcome

Liver_Patie

Access den

←

→

↺

🔒

🔗

📄

https://colab.research.google.com/drive/1ysv8z9jbmUUZVOEyWpZ372nhHO4-zsC_?

80%

☆

🔍

📄

☰

liver.python

File Edit View Insert Runtime Tools Help All changes saved

Comment Share ⚙️ 👤

Files

🔍

📁 ..

📁 sample_data

📄 indian_liver_patient.csv

📄 liver_analysis_1.pkl

<>

📄

📁

Disk 84.54 GB available

+ Code + Text

✓ [19]

✓ [20]

```
# Importing the LabelEncoder library from sklearn=learn
from sklearn.preprocessing import LabelEncoder
le=LabelEncoder()
# Converting Textual data into numeric data
data_1['Gender'] = le.fit_transform(data_1['Gender'])
data_1.head()
```

<ipython-input-20-0e8da2ac9edc>:5: SettingWithCopyWarning:

A value is trying to be set on a copy of a slice from a DataFrame.

Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning

data_1['Gender'] = le.fit_transform(data_1['Gender'])

	Age	Gender	Total_Bilirubin	Direct_Bilirubin	Alkaline_Phosphotase	Alamine_Aminotransferase	Aspartate_Aminotransferase
0	65	0	0.7	0.1	187	16	
1	62	1	10.9	5.5	699	64	
2	62	1	7.3	4.1	490	60	
3	58	1	1.0	0.4	182	14	

✓ 0s

completed at 11:03 AM

Google Ac

Google Ac

My Drive

liver.pytX

Welcome

Welcome

Liver_Patie

Access den

←

→

↺

🔒

🔗

📄

https://colab.research.google.com/drive/1ysv8z9jbmUUZVOEyWpZ372nhHO4-zsC_?

80%

☆

🔍

📄

☰

liver.python

File Edit View Insert Runtime Tools Help All changes saved

Comment Share ⚙️ 👤

Files

🔍

📁 ..

📁 sample_data

📄 indian_liver_patient.csv

📄 liver_analysis_1.pkl

<>

📄

📁

Disk 84.54 GB available

+ Code + Text

✓ [19]

```
# Counting patients who are Male and who are Female
sns.countplot(data=data_1, x = 'Gender', label='Count')
m,f=data_1['Gender'].value_counts()
print("No of Males:",m)
print("No of Females:",f)
```

No of Males: 439

No of Females: 140

✓ 0s

completed at 11:03 AM

Google Chrome browser tabs: Google Ac, Google Ac, My Drive, liver.py, Welcome, Welcome, Liver_Patie, Access der.

Address bar: https://colab.research.google.com/drive/1ysv8z9jbmUUZVOEyWpZ372nhHO4-zsC_? 80%

liver.python

File Edit View Insert Runtime Tools Help All changes saved

Files

- sample_data
 - indian_liver_patient.csv
 - liver_analysis_1.pkl

Code

```
[20] 3 58 1 1.0 0.4 182 14
      4 72 1 3.9 2.0 195 27
```

```
[ ] # Converting Textual data into numeric data
data_1['Gender'] = le.fit_transform(data_1['Gender'])
```

<ipython-input-24-93691031c36e>:2: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead
See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning
data_1['Gender'] = le.fit_transform(data_1['Gender'])

```
[21] data_1.head()
```

	Age	Gender	Total_Bilirubin	Direct_Bilirubin	Alkaline_Phosphotase	Alamine_Aminotransferase	Aspartate_Aminotransferase
0	65	0	0.7	0.1	187	16	
1	62	1	10.9	5.5	699	64	

0s completed at 11:03 AM

Google Chrome browser tabs: Google Ac, Google Ac, My Drive, liver.py, Welcome, Welcome, Liver_Patie, Access der.

Address bar: https://colab.research.google.com/drive/1ysv8z9jbmUUZVOEyWpZ372nhHO4-zsC_? 80%

liver.python

File Edit View Insert Runtime Tools Help All changes saved

Files

- sample_data
 - indian_liver_patient.csv
 - liver_analysis_1.pkl

Code

```
[21] 1 62 1 10.9 5.5 699 64
      2 62 1 7.3 4.1 490 60
      3 58 1 1.0 0.4 182 14
      4 72 1 3.9 2.0 195 27
```

```
[22] x=data_1.iloc[:,0:-1]
      y=data_1.iloc[:,1]
```

```
[23] # dividing the data into input and output
x=data_1.iloc[:,0:-1]
y=data_1.iloc[:,1]
```

```
[24] # importing the train_test_split from scikit-learn
from sklearn.model_selection import train_test_split
xtrain,xtest,ytrain,ytest=train_test_split(x,y,test_size=0.3)
```

```
[25] # Returns size of xtrain
xtrain.shape
```

0s completed at 11:03 AM

Google Chrome browser window showing a Google Colab notebook titled "liver.py". The address bar shows the URL: https://colab.research.google.com/drive/1ysv8z9jbmUUZVOEyWpZ372nhHO4-zsC_?. The notebook interface includes a menu bar (File, Edit, View, Insert, Runtime, Tools, Help) and a toolbar with icons for RAM, Disk, and other settings. The left sidebar shows the file explorer with a folder named "sample_data" containing files "indian_liver_patient.csv" and "liver_analysis_1.pkl". The main code editor displays the following code:

```
[25] # Returns size of xtrain
xtrain.shape

(485, 10)

[26] # Returns size of xtest
ytrain.shape

(485,)

[27] from sklearn.metrics import accuracy_score
from sklearn.metrics import confusion_matrix

[28] # Importing the machine learning model
from sklearn.svm import SVC
from sklearn.ensemble import RandomForestClassifier
from sklearn.neighbors import KNeighborsClassifier

[ ]

[29] # Initializing the machine learning model
```

The status bar at the bottom indicates "0s completed at 11:03 AM".

Google Chrome browser window showing the same Google Colab notebook. The address bar shows the URL: https://colab.research.google.com/drive/1ysv8z9jbmUUZVOEyWpZ372nhHO4-zsC_?. The notebook interface is the same as the previous screenshot. The main code editor displays the following code:

```
[ ]

[29] # Initializing the machine learning model
svm=SVC()
RFmodel=RandomForestClassifier()
KNmodel=KNeighborsClassifier()

[ ]

[30] # Support Vector Machine Model
from sklearn.svm import SVC
svm=SVC()

[31] # train the data with SVM model
svm.fit(xtrain, ytrain)

[32] SVCpred=svm.predict(xtest)
```

The status bar at the bottom indicates "0s completed at 11:03 AM".

Google Chrome browser window showing a Google Colab notebook titled "liver.py". The address bar shows the URL: https://colab.research.google.com/drive/1ysv8z9jbmUUZVOEyWpZ372nhHO4-zsC_?. The notebook interface includes a menu bar (File, Edit, View, Insert, Runtime, Tools, Help) and a toolbar with icons for file operations and runtime settings. The left sidebar shows the file explorer with a folder named "sample_data" containing files "indian_liver_patient.csv" and "liver_analysis_1.pkl". The main code area displays the following code cells:

```
[33] from sklearn.metrics import accuracy_score, confusion_matrix
# Checking for accuracy score from actual data and predicted data
SVCaccuracy=accuracy_score(SVCpred, ytest)
SVCaccuracy

0.7126436781609196

[34] # showing the confusion matrix
SVCcm=confusion_matrix(SVCpred, ytest)
SVCcm

array([[124, 50],
       [ 0,  0]])

[35] #Random Forest Classifier model
from sklearn.ensemble import RandomForestClassifier
RFmodel=RandomForestClassifier()

[36] # train the data with Random Forest model
RFmodel.fit(xtrain,ytrain)

RandomForestClassifier
RandomForestClassifier()
```

The notebook status bar indicates "0s completed at 11:03 AM".

Google Chrome browser window showing the same Google Colab notebook. The code area displays the following code cells:

```
[37] RFpred=RFmodel.predict(xtest)

[38] # Checking for accuracy score from actual data and predicted data
RFaccuracy=accuracy_score(RFpred,ytest)
RFaccuracy

0.6436781609195402

[39] # showing the confusion matrix
RFCm=confusion_matrix(RFpred, ytest)
RFCm

array([[99, 37],
       [25, 13]])

[ ]

[ ]

[40] # K-Nearest Neighbors Model
from sklearn.neighbors import KNeighborsClassifier
KNN = KNeighborsClassifier()
```

The notebook status bar indicates "0s completed at 11:03 AM".

Google Chrome browser window showing a Google Colab notebook titled "liver.py". The address bar shows the URL: https://colab.research.google.com/drive/1ysv8z9jbmUUZVOEyWpZ372nhHO4-zsC_?. The notebook interface includes a file explorer on the left showing a folder named "sample_data" containing "indian_liver_patient.csv" and "liver_analysis_1.pkl". The main code area shows the following code cells:

```
[41] # train the data with K-Nearest Neighbors Model
      KNN.fit(xtrain, ytrain)
      KNeighborsClassifier
      KNeighborsClassifier()

[42] KNNpred=KNN.predict(xtest)

[43] # Checking for accuracy score from actual data and predicted data
      KNNaccuracy=accuracy_score(KNNpred,ytest)
      KNNaccuracy
      0.6436781609195402

[44] # showing the confusion matrix
      KNNcm=confusion_matrix(KNNpred, ytest)
      KNNcm
      array([[99, 37],
             [25, 13]])

[ ]
```

The status bar at the bottom indicates "0s completed at 11:03 AM".

Google Chrome browser window showing the same Google Colab notebook. The code cells are now:

```
[44] # showing the confusion matrix
      KNNcm=confusion_matrix(KNNpred, ytest)
      KNNcm
      array([[99, 37],
             [25, 13]])

[ ]

[45] print("Support vector Machine Algorithms accuracy score : {value:.2f} %".format(value=SVCaccuracy*100))
      print("Random Forest Algorithms accuracy score : {value:.2f} %".format(value=RFaccuracy*100))
      print("K-Nearest Neighbors Algorithms accuracy score : {value:.2f} %".format(value=KNNaccuracy*100))

      Support vector Machine Algorithms accuracy score : 71.26 %
      Random Forest Algorithms accuracy score : 64.37 %
      K-Nearest Neighbors Algorithms accuracy score : 64.37 %

[46] # saving the model
      import pickle
      pickle.dump(svm, open('liver_analysis_1.pkl','wb'))
```

The status bar at the bottom indicates "0s completed at 11:03 AM".

