PREDICTION AND ANALYSIS OF LIVER PATIENT DATA USING IBM MACHINE LEARNING SERVICE

An Industrial/Practical Training Report

Submitted to the Faculty of Engineering of

JAWAHARLAL NEHRU TECHNOLOGICAL UNIVERSITY KAKINADA,

KAKINADA

In partial fulfillment of the requirements for the award of the Degree of

BACHELOR OF TECHNOLOGY

IN

INFORMATION TECHNOLOGY

BY

GADE AKASH REDY (19481A1230)

DEPARTMENT OF INFORMATION TECHNOLOGY

SESHADRI RAO GUDLAVALLERU ENGINEERING COLLEGE

(An Autonomous Institute with Permanent Affiliation to JNTUK, Kakinada)

SESHADRI RAO KNOWLEDGE VILLAGE

GUDLAVALLERU – 521356

ANDHRA PRADESH

2022-2023

SMARTBRIDGE
Let's Bridge the Gap

Date: 30/04/2022

TO WHOMSOEVER IT MAY CONCERN

This is to certify that **Mr./Ms. Akash Reddy** pursuing B.Tech, Information Technology from Seshadri Rao Gudlavalleru Engineering College, Gudlavalleru has successfully completed his/her Virtual Internship Program from **27/02/2022** to **30/04/2022.**

During this period he/she had learned the concepts of **Machine Learning & Deep Learning** and successfully completed a project **"Prediction And Analysis Of Liver Patient Data Using IBM Machine Learning Service".**

Refer the enclosed **Certificate of Merit** for his/her performance during the tenure of Virtual Internship Program.

We wish him/her all the best for his/her future endeavours.

For **SmartBridge Educational Services Pvt. Ltd.,**

SmartBridge Educational Services Pvt. Ltd. Hyderabad

Jayaprakash.ch,
Program Manager

SMARTBRIDGE
Let's Bridge the Gap

Certificate of Merit

This is to certify that **Mr./Ms. Akash Reddy** has completed his/her Virtual Internship Program from **27/02/2022** to **30/04/2022** with **Machine Learning & Deep Learning** as the specialization and secured a SKILL INDEX **9** of 10.

Career Readiness Factor (CRF)

Evaluation Metrics: (on a scale of 1 to 4)

1- Rarely/poorly displays characteristic
2- Occasionally displays characteristic
3- Frequently displays characteristic
4- Always displays characteristic

NA – Not Applicable

Motivation/Enthusiasm	3
Leadership Qualities	3
Flexibility towards work	3
Professionalism/Work Ethics	4
Self-Confidence	3
Ability to work independently	4
Oral/written communication	3
Problem solving skills	3
Over All Score	26

Date: 30/04/2022

Jayaprakash. ch
Program Manager

This is to certify that

Akash Reddy

successfully completed and received a passing grade in

Design Thinking

(CEDTH1IN, provided by IBMCE)

A course on **smartbridge.skillsnetwork.site**
Powered by **IBM Developer Skills Network.**

Issued by
SmartBridge

Jagadisha Bhat
Country Manager - Software Services
IBM India Pvt Ltd

November 24, 2022

Authenticity of this certificate can be validated by going to:
https://courses.smartbridge.skillsnetwork.site/certificates/5e9d2087632140d894bcabdb6a92251e

12/4/22, 10:06 AM IBMCE CEDTH1IN Certificate | SmartBridge

Certificate ID Number: **5e9d2087632140d894bcabdb6a92251e** **November 24, 2022**

This is to certify that

Akash Reddy

successfully completed and received a passing grade in

DevOps Fundamentals

(CEDEV1IN, provided by IBMCE)

A course on smartbridge.skillsnetwork.site
Powered by IBM Developer Skills Network.

Issued by
SmartBridge

Jagadisha Bhat
Country Manager - Software Services
IBM India Pvt Ltd

November 24, 2022

Authenticity of this certificate can be validated by going to:
https://courses.smartbridge.skillsnetwork.site/certificates/1d19b154c2ec4be8b78e0eb67afeac69

12/4/22, 10:07 AM IBMCE CEDEV1IN Certificate | SmartBridge

Certificate ID Number: 1d19b154c2ec4be8b78e0eb67afeac69 November 24, 2022

This is to certify that

Akash Reddy

successfully completed and received a passing grade in

Agile Training Module

(CEAGL1IN, provided by IBMCE)

A course on smartbridge.skillsnetwork.site
Powered by IBM Developer Skills Network.

Issued by
SmartBridge

Jagadisha Bhat
Country Manager - Software Services
IBM India Pvt Ltd

November 24, 2022

Authenticity of this certificate can be validated by going to:
https://courses.smartbridge.skillsnetwork.site/certificates/d8d3ca608d3a428c9e586484af06d9ad

Certificate ID Number: **d8d3ca608d3a428c9e586484af06d9ad** **November 24, 2022**

SESHADRI RAO GUDLAVALLERU ENGINEERING COLLEGE

(An Autonomous Institute with Permanent Affiliation to JNTUK, Kakinada)

SESHADRI RAO KNOWLEDGE VILLAGE, GUDLAVALLERU

DEPARTMENT OF INFORMATION TECHNOLOGY

CERTIFICATE

This is to certify that the project report entitled **"Prediction And Analysis Of Liver Patient Data Using IBM Machine Learning Service"** is a bonafide record of work carried out by **G.Akash Reddy(19481A1230)** as a part of internship in the partial fulfillment of the requirements for the award of the Degree of **Bachelor of Technology** in **Information Technology** of **Jawaharlal Nehru Technological University Kakinada,** Kakinada during the academic year 2022-23.

Head of the Department

External Examiner **(DR.CH.KAVITHA)**

ACKNOWLEDGEMENT

I am glad to express my deep sense of gratitude to **Smart Bridge Mentors** for their guidance and cooperation in completing this externship project. Through this, I want to convey my sincere thanks to them for inspiring assistance during my project.

I express my heartful gratitude and deep indebtedness to our beloved Head of the Department **Dr.Ch.Kavitha** for her great help and encouragement in doing my project successfully.

I also express my gratitude to our principal **Dr.G.V.S.N.R.V Prasad** for his encouragement and facilities provided during the course of project.

I express my heartful gratitude to all Faculty Members, all Lab Technicians, who helped me in all aspects of lab work. I thank one and all who have rendered help to me directly or indirectly in the completion of this project.

Project Associate

G.AKASH REDDY (19481A1230)

ABSTRACT

Chronic Liver Disease is the leading cause of global death that impacts the massive quantity of humans around the world. This disease is caused by an assortment of elements that harm the liver. For example, obesity, an undiagnosed hepatitis infection, alcohol misuse. Which is responsible for abnormal nerve function, coughing up or vomiting blood, kidney failure, liver failure, jaundice, liver encephalopathy and there are many more. 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 chronic liver disease diagnosis by prediction. In this work, we used 3 algorithms K Nearest Neighbors, Support Vector Machine and Random Forest. The performance of different classification techniques was evaluated on different measurement techniques such as accuracy, precision, recall, f-1 score, and specificity. We found the accuracy 65%, 67%, 62% for RF, SVM and KNN. The analysis result shown the SVM achieved the highest accuracy. Moreover, our present study mainly focused on the use of clinical data for liver disease prediction and explore different ways of representing such data through our analysis.

Keywords: Machine Learning, Liver Disease, Classification, Supervised learning, Computational Intelligence, Regression,RandomForest,SupportVectorMachine,K-NearestNeighbors.

INDEX

TITLE	PAGE NO
CHAPTER 1: INTRODUCTION	**1-3**
1.1 OVERVIEW	1
1.2 PURPOSE	1
1.3 PRE-REQUISITES	2
1.4 PROJECT FLOW	3
CHAPTER 2: LITERATURE SURVEY	**4**
2.1 EXISTING PROBLEM	4
2.2 PROPOSED SOLUTION	4
CHAPTER 3: THEORTICAL ANALYSIS	**5-7**
3.1 BLOCK DIAGRAM	5
3.2 SOFTWARE SPECIFICATIONS	5-6
3.3 HARDWARE SPECIFICATIONS	7
CHAPTER 4: EXPERIMENTAL INVESTIGATIONS	**8-11**
4.1 IBM CLOUD	8
4.2 DATASET COLLECTION	8-9
4.3 ALGORITHMS	10-12
CHAPTER 5: FLOW CHART	13
CHAPTER 6: RESULTS	14
CHAPTER 7: ADVANTAGES AND DISADVANTAGES	15
CHAPTER 8: APLICATIONS	16
CHAPTER 9: CONCLUSION	17
CHAPTER 10: FUTURE SCOPE	18
CHAPTER 11: BIBILOGRAPHY	19
CHAPTER 12: APPENDIX	20-22
CHAPTER 13: PROJECT WORK MAPPING WITH PROGRAMME OUTCOMES	23-26

LIST OF FIGURES

FIG NO.	FIGURE NAME	PAGE NO
Fig 3.1.1	Block diagram of Machine Learning	5
Fig 3.2.1	Anaconda Navigator	6
Fig 3.2.2	Jupyter Notebook	6
Fig 4.1	Previewing the dataset	9
Fig 4.1.2	Test and trained datasets	9
Fig 4.1.1	Random Forest model	10
Fig 4.2.1	Support Vector Machine model	11
Fig 4.3.1	K Nearest Neighbors model	11
Fig 4.4	Model stored as pickle	12
Fig 5.1	Flow diagram of our system	13
Fig 6.1	Home Page	14
Fig 6.2	Prediction Page	14
Fig 6.3	Result Page	14
Fig 12.1.1	Installation step 1	20
Fig 12.1.2	Installation step 2	21
Fig 12.1.3	Installation step 3	21
Fig 12.1.4	Installation step 4	22
Fig 12.2.1	Project Structure	*22*

CHAPTER-1

INTRODUCTION

Machine learning is a growing technology which enables computers to learn automatically from past data. Machine learning uses various algorithms for building mathematical models and making predictions using historical data or information.Currently, it is being used for various tasks such as image recognition, speech recognition, email filtering, Facebook auto-tagging, and many more.

A Machine Learning system learns from historical data, builds the prediction models, and whenever it receives new data, predicts the output for it. The accuracy of predicted output depends upon the amount of data, as the huge amount of data helps to build a better model which predicts the output more accurately.

OVERVIEW

The liver is the largest organ of the body and it is essential for digesting food and releasing the toxic element of the body. Theviruses and alcohol use lead the liver towards liver damage and lead a human to a life-threatening condition. According to the Global Burden of Disease (GBD) project, published in BMC Medicine, one million peoples are died in 2010 because of cirrhosis and millions are suffering from liver cancer. Machine learning has made a significant impact on the biomedical field for liver disease prediction and diagnosis. Machine learning offers a guarantee for improving the detection and prediction of disease that has been made an interest in the biomedical field and they also increase the objectivity of the decision-making process. By using machine learning techniques medical problems can be easily solved and the cost of diagnosis will be reduced. In this study, the main aspect is to predict the results more efficiently and reduce the cost of diagnosis in the medical sector. Therefore, we used different classification techniques for the classification of patients has liver disease or not.

PURPOSE

By Prediction And Analysis Of Liver Patient Data Using IBM Machine Learning Service We Will:

- Know fundamental concepts and can work on IBM Machine Learning Service.
- Gain a broad understanding of Classification Techniques.

PRE-REQUISITES:

In order to develop this Project we need to install following software/packages:

Anaconda Navigator :

Anaconda Navigator is a free and open-source distribution of the Python and R programming languages for data science and machine learning related applications. It can be installed on Windows, Linux, and macOS. Anaconda is an open-source, cross-platform, package management system. Anaconda comes with so very nice tools like Jupyter Lab, Jupyter Notebook,QtConsole, Spyder, Glueviz, Orange, Rstudio, Visual Studio Code. For this project, we will be using Jupyter notebook and Spyder

To install Anaconda navigator and to know how to use Jupyter Notebook & Spyder using Anaconda watch the video.

Link:https://youtu.be/5mDYijMfSzs

To make a responsive python script you must require the following packages

Requests: Allows you to send HTTP requests using Python.

Flask**:** Web framework used for building Web applications.

Link: https://youtu.be/akj3_wTploU

If you are using **anaconda navigator**, follow below steps to download required packages:

- Open the anaconda prompt.
- Type "pip install requests" and click enter.
- Type "pip install Flask" and click enter.

If you are using Pycharm IDE, you can install the packages through the command prompt and follow the same syntax as above.

Prior Knowledge : One should have knowledge on the following Concepts:

Requests: Allows you to send HTTP requests using Python.

Flask: Web framework used for building Web applications.

Link: https://youtu.be/lj4I_CvBnt0

PROJECT FLOW

1. Install Required Libraries.

2. Data Collection.
    - Collect the dataset or Create the dataset

3. Data Pre- processing.
    - Import the Libraries.
    - Importing the dataset.
    - Understanding Data Type and Summary of features.
    - Take care of missing data & create columns.
    - Data Visualization.
    - Drop the column from data frame, merge the data frames.
    - Observing Target, Numerical and Categorical Columns
    - Label Encoding & Splitting the Dataset into Dependent and Independent variables
    - Splitting Data into Train and Test.

4 .Model Building
    - Training and testing the model
    - Evaluation of Model
    - Saving the Model

5. Application Building
    - Create an HTML file
    - Build a Python Code

6. Final UI
    - Dashboard Of the flask app.

CHAPTER-2
LITERATURE SURVEY

EXISTING PROBLEM

liver disease is one of the major health problems in the world. Every year, around 2 million people died worldwide because of liver disease .we could not even predict the presence of liver disease in our body. It leads to a person's death if he could not identify the symptoms of liver disease in his body.

PROPOSED SOLUTION

Artificial Intelligence being the trending technology can end up with the best solutions for every typical problem. Here we are using supervised Machine Learning algorithms. Our proposed model says weather a person is having liver disease or not and instructs him to consult a doctor in case of its presence. As time progresses and more data are collected, the supervised learning willproduce more accurate results and will be helpful in determining fare optimizer and dynamic availability of adjustments and continuously improve future recommendations.

CHAPTER-3

THEORITICAL ANALYSIS

BLOCK DIAGRAM

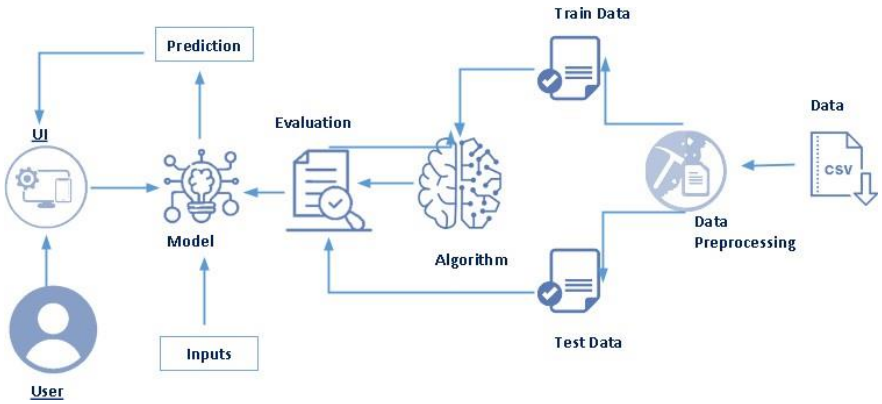

Fig 3.1.1 Block diagram of Machine Learning

SOFTWARE REQUIREMENTS

- To complete the project, you need the following packages and libraries.
  1. Anaconda Navigator
  2. Jupyter
  3. Numpy
  4. Pandas
  5. Matplotlib
  6. Scikit-Learn

Install Anaconda software

To install Anaconda on your local system, go through the below links according to your system requirements. After Anaconda is installed, run the .exe folder.

Install Required Libraries

Search Anaconda Navigator and open a Jupyter notebook.

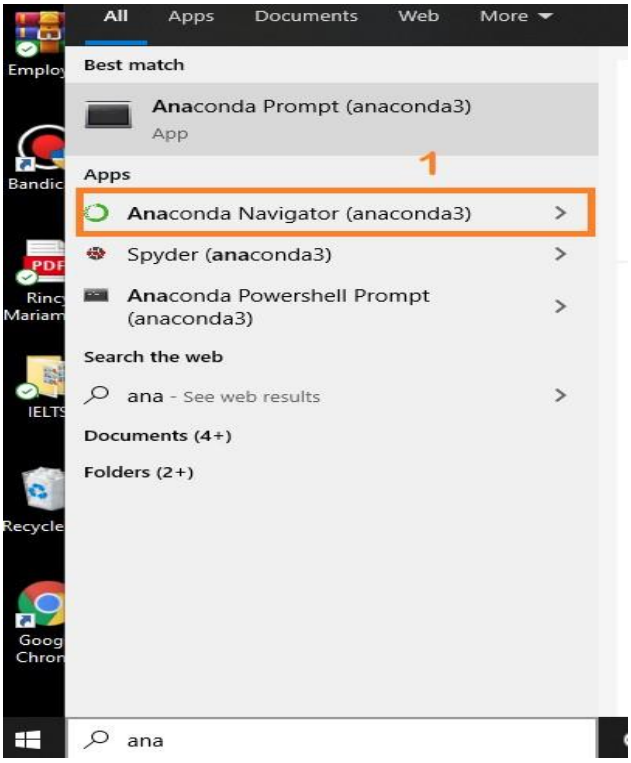

Fig 3.2.1 Anaconda Navigator

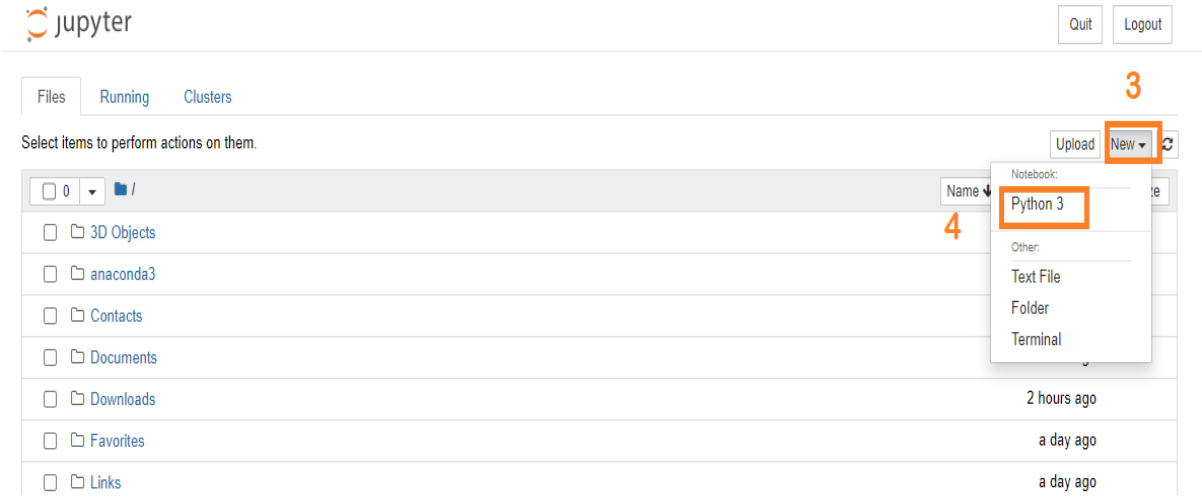

Fig 3.2.2 Anaconda Navigator Jupyter Notebook

HARDWARE SPECIFICATIONS

REQUIREMENT	SPECIFICATIONS
Operating System	Microsoft Windows Unix Linux
Processing	Minimum: 4 CPU cores for one user. For each deployment, a sizing exercise is highly recommended.
RAM	Minimum 8 GB
Operating system specifications	File descriptor limit set to 8192 on UNIX and Linux
Disk Space	A Minimum of 7 GB of free space is required to install the software.

CHAPTER-4
EXPERIMENTAL INVESTIGATIONS

Analysis or the investigation made while working on the solution:

While working on the solution we investigated on what is 3D Printing, IBM cloud, IBM Watson studio, Machine Learning service, Cloud Object Storage. The key role on investigation is collection of dataset.

IBM CLOUD ACCOUNT

IBM Acquired soft layer, a public cloud platform, to serve as the foundation for its IaaS offering.In October 2016, IBM rolled the soft layer brand under its Blue mix brand of PaaS offerings, giving users to access both IaaS and PaaS resources from a single console. IBM cloud provides a full-stack, public cloud platform with various products in the Catalog, including options for compute,storage,networking, end to end developer solutions for app development, testing and deployment,security databases, and cloud native services.

Creating the IBM cloud account by going to the IBM cloud login page and click create on IBM cloud account. Enter our IBM id and an ID is created based on the email that we enter.Completing the remaining fields with our information and click create account by this the account is created.

DATA SET COLLECTION

- ML depends heavily on data, without data, it is impossible for an “AI” to learn. It is the most crucial aspect that makes algorithm training possible. In Machine Learning projects, we need a training data set. It is the actual data set used to train the model for performing various actions.
- There are many features which are responsible for occurrence of Liver Disesease , e.g. Age,Gender,Total Bilirubin,Alamne_Aminotransferase etc. For better prediction of the Liver Cancer, we should consider as many essential features as possible.
- You can collect dataset from different open sources like kaggle.com, data.gov, UCI machine learning repository etc.
- Here we are using a data set which you can find in the below link and you can download it from the link.

Link: https://www.kaggle.com/datasets/uciml/indian-liver-patient-records

```
In [8]: import pandas as pd
        import numpy as np
        import matplotlib.pyplot as plt
        import seaborn as sns
        import pickle

In [11]: data = pd.read_csv(r'C:\Users\NAME\OneDrive\Desktop\Internship Project\Notebook\indian_liver_patient.csv')

In [12]: data.head()
```

Out[12]:

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

4.1 previewing the dataset

- There are several Machine learning algorithms to be used depending on the data you are going to process such as images, sound, text, and numerical values. The algorithms that you can choose according to the objective that you might have it may be Classification algorithms are Regression algorithms.

Example:

1.Random Forest Regression / Classification.

2. Support Vector Machine.

2. K-Nearest Neighbors / Classification.

- You will need to train the datasets to run smoothly and see an incremental improvement in the prediction rate.

```
In [15]: x=data.iloc[:,0:-1]
         y=data.iloc[:,-1]
         print(x)
         print(y)
```

```
      Age  Gender  Total_Bilirubin  ...  Total_Protiens  Albumin  Albumin_and_Globulin_Ratio
0      65       0              0.7  ...             6.8      3.3                        0.90
1      62       1             10.9  ...             7.5      3.2                        0.74
2      62       1              7.3  ...             7.0      3.3                        0.89
3      58       1              1.0  ...             6.8      3.4                        1.00
4      72       1              3.9  ...             7.3      2.4                        0.40
..    ...     ...              ...  ...             ...      ...                         ...
578    60       1              0.5  ...             5.9      1.6                        0.37
579    40       1              0.6  ...             6.0      3.2                        1.10
580    52       1              0.8  ...             6.4      3.2                        1.00
581    31       1              1.3  ...             6.8      3.4                        1.00
582    38       1              1.0  ...             7.3      4.4                        1.50

[583 rows x 10 columns]
0      1
1      1
2      1
3      1
4      1
      ..
578    2
579    1
580    1
581    1
582    2
Name: Dataset, Length: 583, dtype: int64
```

```
In [16]: from sklearn.model_selection import train_test_split
         xtrain,xtest,ytrain,ytest=train_test_split(x,y,test_size=0.2)

In [17]: xtrain.shape
Out[17]: (466, 10)
```

Fig 4.2 Test and trained dataset

1. RANDOM FOREST ALGORITHM

- A Random Forest is an ensemble technique capable of performing both regression and classification tasks with the use of multiple decision trees and a technique called Bootstrap and Aggregation, commonly known as bagging. The basic idea behind this is to combine multiple decision trees in determining the final output rather than relying on individual decision trees.
- A Random Forest is an ensemble technique capable of performing both regression and classification tasks with the use of multiple decision trees and a technique called Bootstrap and Aggregation, commonly known as bagging. The basic idea behind this is to combine multiple decision trees in determining the final output rather than relying on individual decision trees.

Build the model with the Random Forest Regressor:

- We're going to use x_train and y_train obtained above in train_test_split section to train our Random Forest regression model. We're using the fit method and passing the parameters as shown below

```
In [23]: from sklearn.ensemble import RandomForestClassifier
         RFmodel=RandomForestClassifier()

In [24]: RFmodel.fit(xtrain,ytrain)
Out[24]: RandomForestClassifier()

In [25]: RFpredict=RFmodel.predict(xtest)

In [26]: RFaccuracy=accuracy_score(RFpredict,ytest)
         RFaccuracy
Out[26]: 0.6581196581196581
```

Fig 4.1.1 Random forest model

2. SUPPORT VECTOR MACHINE ALGORITHM

- Support Vector Machine or SVM is one of the most popular Supervised Learning algorithms, which is used for Classification as well as Regression problems. However, primarily, it is used for Classification problems in Machine Learning.
- The goal of the SVM algorithm is to create the best line or decision boundary that can segregate n-dimensional space into classes so that we can easily put the new data point in the correct category in the future. This best decision boundary is called a hyperplane.
- SVM works by mapping data to a high-dimensional feature space so that data points can be categorized, even when the data are not otherwise linearly separable. A separator between the categories is found, then the data are transformed in such a way that the separator could be drawn as a hyperplane.

Build the model with the Support Vector Machine:

```
In [17]: from sklearn.svm import SVC
         from sklearn.ensemble import RandomForestClassifier
         from sklearn.neighbors import KNeighborsClassifier
         from sklearn.metrics import accuracy_score,confusion_matrix

In [18]: svm=SVC()
         RFmodel=RandomForestClassifier()
         KNNmodel=KNeighborsClassifier()

In [19]: svm.fit(xtrain,ytrain)
Out[19]: SVC()

In [20]: SVMpred=svm.predict(xtest)

In [21]: SVMaccuracy=accuracy_score(SVMpred,ytest)
         SVMaccuracy
Out[21]: 0.6752136752136753
```

Fig 4.2.1 Support Vector Machine model

3. K NEAREST NEIGHBORS

- K-NN algorithm can be used for Regression as well as for Classification but mostly it is used for the Classification problems.
- K-NN is a **non-parametric algorithm**, which means it does not make any assumption on underlying data.
- It is also called a **lazy learner algorithm** because it does not learn from the training set immediately instead it stores the dataset and at the time of classification, it performs an action on the dataset.
- KNN algorithm at the training phase just stores the dataset and when it gets new data, then it classifies that data into a category that is much similar to the new data.

Build the model with the K Nearest Neighbors:

```
In [28]: from sklearn.neighbors import KNeighborsClassifier
         KNN=KNeighborsClassifier()

In [29]: KNN.fit(xtrain,ytrain)
Out[29]: KNeighborsClassifier()

In [30]: KNNpred=KNN.predict(xtest)

In [31]: KNNaccuracy=accuracy_score(KNNpred,ytest)
         KNNaccuracy
Out[31]: 0.6239316239316239
```

Fig 4.3.1 K Nearest Neighbors model

- We have created 3 models SVM, KNN and Random Forest and predicted the result using three different models.
- Finally we displayed the accuracies of the 3 models individually.
- We observed that the accuracy of SVM model is more than the other 2 models .
- So we tend to finalize the SVM algorithm for our system usage with accuracy 67%.

Saving The Model

- Pickle is used for serializing and de-serializing Python object structures, also called marshalling or flattening. Serialization refers to the process of converting an object in memory to a byte stream that can be stored on disk or sent over a network. Later on, this character stream can then be retrieved and de-serialized back to a Python object.
- Save our model by importing pickle files.

```
In [34]: import pickle
         pickle.dump(svm,open('liver_analysis_svm1.pkl','wb'))

In [ ]:
```

Fig 4.3.1 model stored as pickle

CHAPTER-5

FLOW CHART

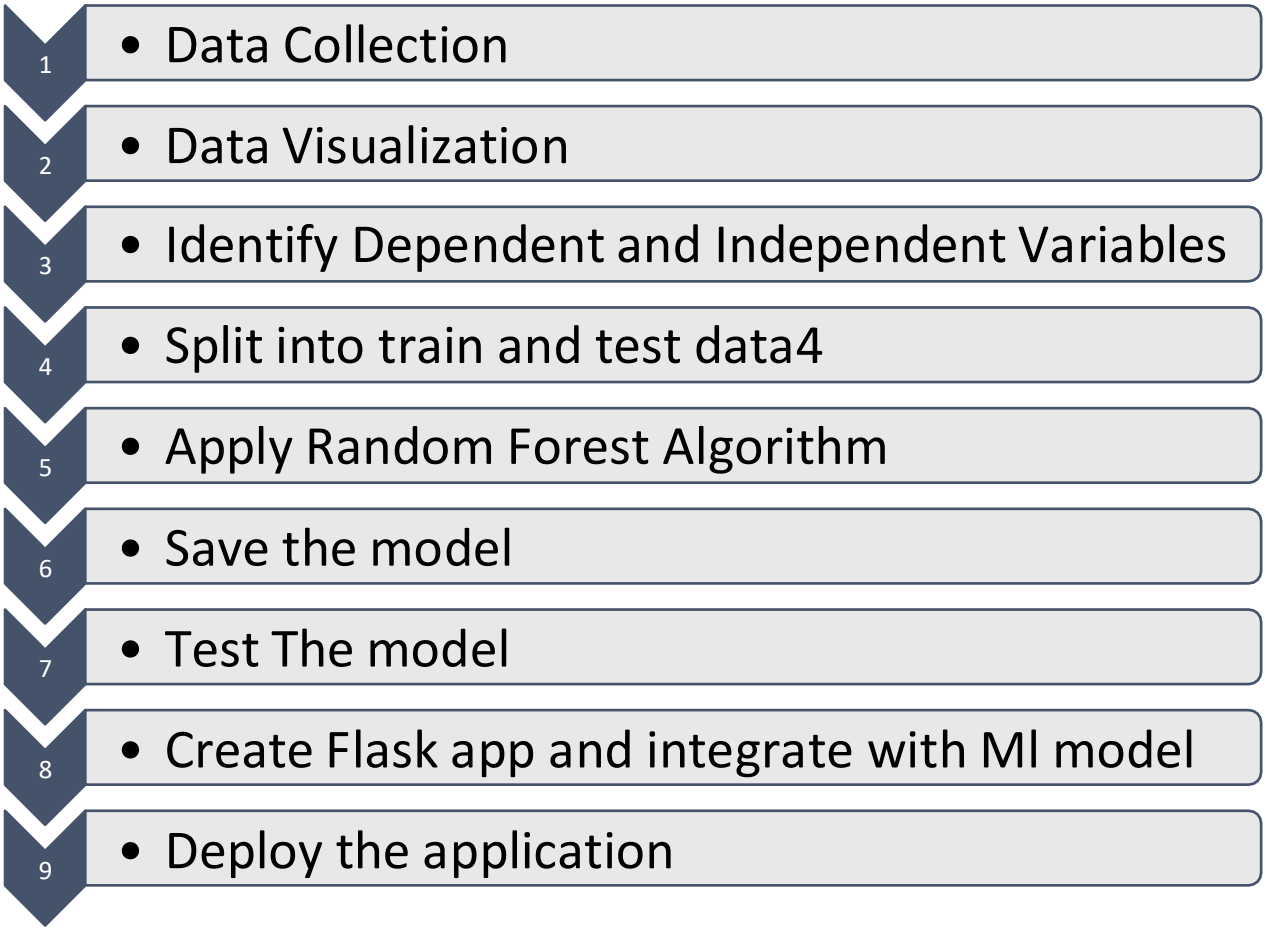

Fig 5.1 Flow diagram of our system

CHAPTER-6

RESULTS

Final Output of the Project

- This is the main page of Liver Disease Prediction. Where you may know about the project and also from this page user can click on the GO PREDICT button and they will redirect onto the prediction page for providing the inputs.

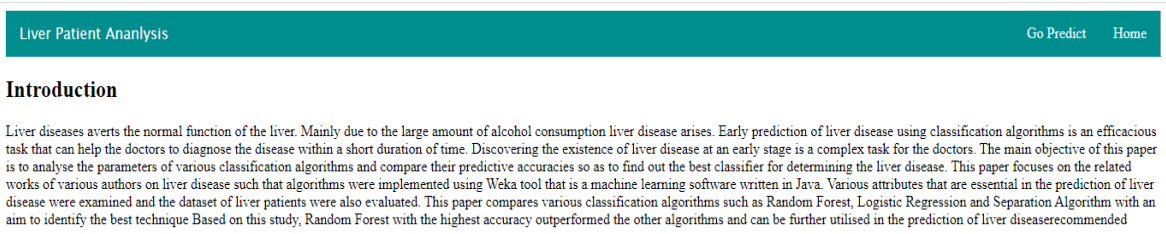

Fig 6.1 Home Page

- The prediction page user gives the input for predicting the output where they can choose Age,gender,Total_Bilirubin, Direct_Bilirubin, Alkaline_Phosphotase, Total_Protiens etc.. then click to predict the output.

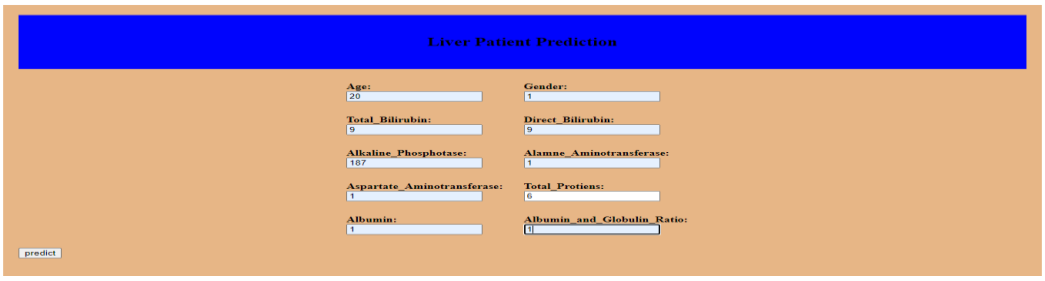

Fig 6.2 Prediction Page

- In the predict page user will get the output based on the inputs they given in the prediction page.

Fig 6.3 Result Page

CHAPTER-7

ADVANTAGES AND DISADVANTAGES

Advantages

- By this Project the Liver Disease prediction can be done.
- As we used Support Vector Machine Model the accuracy will be more and gives accurate values.

Disadvantages

- Users tend to loose their control on viewing their health status.
- Doest consider a lot more attributes that contribute to the Liver Disease.
- Can't be used for a purpose where the deficiency of health is high.

CHAPTER-8

APPLICATIONS

- It is used a software Application in Hospitals for Liver Disease Prediction.
- Used as general purpose checkup of Liver Disease in Homes.
- Identifying the Immediate Change in the Health functionality .

CHAPTER-9

CONCLUSION

- From the entire findings we came to know fundamental concepts and can work on IBM Watson and machine learning.
  - → Gain a understanding of Random Forest algorithms
  - → Gain a understanding of SVM algorithms
  - → Gain a understanding of KNN algorithms
  - → Learn to build useful models on IBM cloud
  - → To create data visualizations for understanding.

CHAPTER 10

FUTURE SCOPE

❖ **Enhancements that can be made in the future:**

- This Liver Disease application can be further extended by adding more disease predictions and having provided with the remedy for it.
- We can develop the application by considering more attributes that has more impact on the disease and that could help in better prediction.

CHAPTER 11

BIBILOGRAPHY

- The Handbook of Liver Disease, 3rd Edition, by Drs. Lawrence S. Friedman and Emmet B. Keeffe.
- Machine Learning With Random Forests And Decision Trees: A Mostly Intuitive Guide, But Also Some Python by Scott Hartshorn (Goodreads Author)
- Gunn, Support Vector Machines for Classification and Regression, http://www.isis.ecs.soton.ac.uk/resources/svminfo/
- Random forest classifier for remote sensing classification

  M Pal - International journal of remote sensing, 2005 - Taylor & Francis

CHAPTER 12

APPENDIX

How to install the Anaconda Navigator:

1. Download the Anaconda installer.
2. RECOMMENDED: Verify data integrity with SHA-256. For more information hashes, see what about cryptographic hash verification?
3. Double click the installer to launch.
4. Click Next.
5. Read the licensing terms and click “I Agree”.
6. Select an install for “Just Me” unless you’re installing for all users (which require windows administrator privileges) and click next.
7. Select a destination folder to install Anaconda and click the Next button.

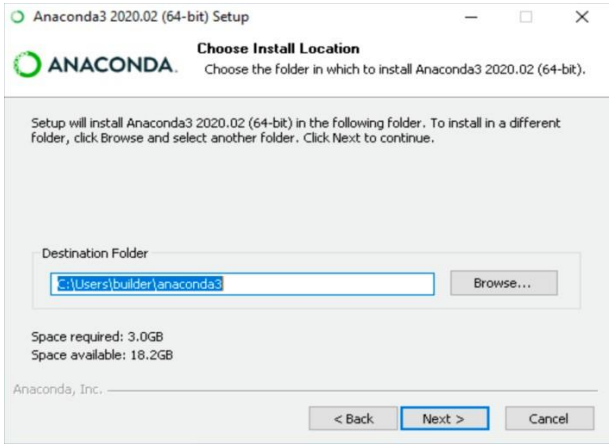

Fig 12.1.1 Installation step 1

8.Choose whether to add Anaconda to your PATH environment variable. We recommend not adding Anaconda to the PATH environment variable, since this can interfere with other software. Instead, use Anaconda software by opening Anaconda Navigator or the Anaconda Prompt from the Start Menu.

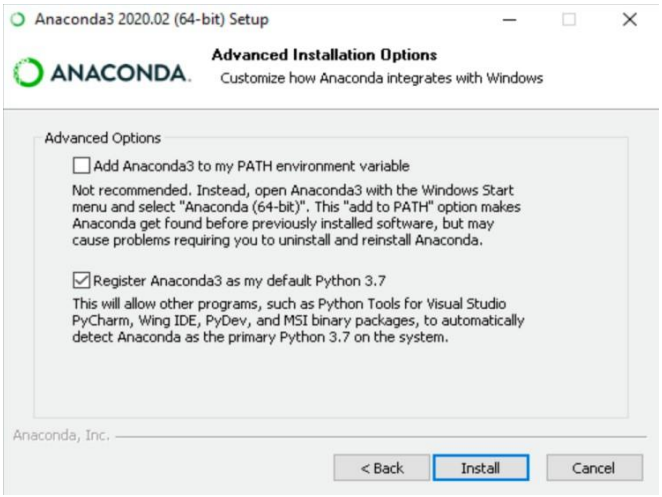

Fig 12.1.2 Installation step 2

9.Choose whether to register Anaconda as your default Python. Unless you plan on installing and running multiple versions of Anaconda or multiple versions of Python, accept the default and leave this box checked.

10. Click the Install button. If you want to watch the packages Anaconda is installing, click Show Details.

11. Click the Next button.

12. Optional: To install PyCharm for Anaconda, click on the link to https://www.anaconda.com/pycharm.

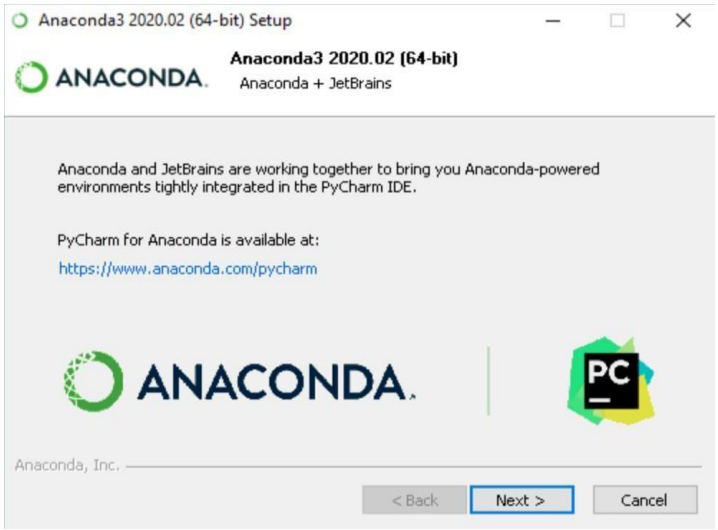

Fig 12.1.3 Installation step 3

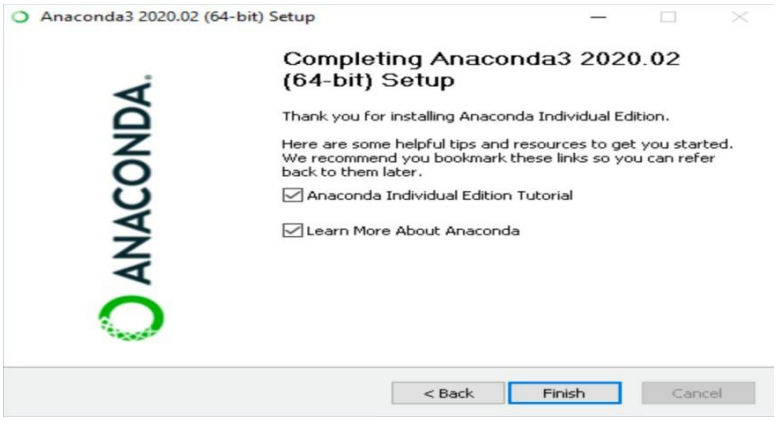

Fig 12.1.4 Installation step 4

Project Structure:

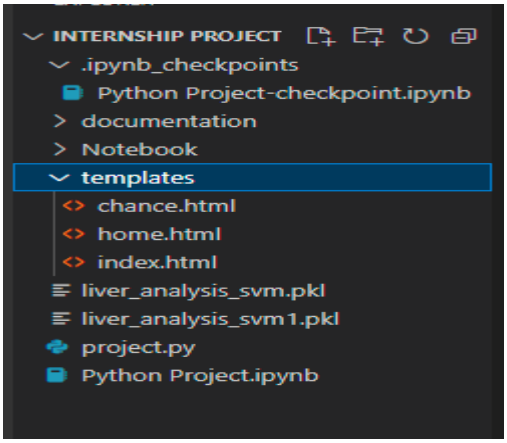

Fig 12.2.1 Project Structure

How to run project:

Step 1: Open anaconda prompt.

Step 2: Change the working directory to project directory.

Step 3: Move to the directory named flask.

Step 4: Launch the application app.py using **python app.py** command.

Step 5: Application runs at **http:localhost:5000** port

Step 6: Press Ctrl+ C to stop the execution.

CHAPTER 13

PROJECT WORK MAPPING WITH PROGRAMME OUTCOMES

Pos	1	2	3	4	5	6	7	8	9	10	11	12
Project	2	3	3	2	3	3	2	2	2	2	3	3

PROGRAMMEOUTCOMES(POs)

Engineering Graduates will be able to:

1. **Engineering knowledge:** Apply the knowledge of mathematics, science, engineering fundamentals, and an engineering specialization to the solution of complex engineering problems.
2. **Problem analysis:** Identify, formulate, review research literature, and analyze complex engineering problems reaching substantiated conclusions using first principles of mathematics, natural sciences, and engineering sciences.
3. **Design/development of solutions:** Design solutions for complex engineering problems and design system components or processes that meet the specified needs with appropriate consideration for the public health and safety, and the cultural, societal, and environmental considerations.
4. **Conduct investigations of complex problems:** Use research-based knowledge and research methods including design of experiments, analysis and interpretation of data, and synthesis of the information to provide valid conclusions.
5. **Modern tool usage:** Create, select, and apply appropriate techniques, resources, and modern engineering and IT tools including prediction and modeling to complex engineering activities with an understanding of the limitations.
6. **The engineer and society:** Apply reasoning informed by the contextual knowledge to assess societal, health, safety, legal and cultural issues and the consequent responsibilities relevant to the professional engineering practice.
7. **Environment and sustainability:** Understand the impact of the professional engineering solutions in societal and environmental contexts, and demonstrate the knowledge of, and need for sustainable development.

8. **Ethics:** Apply ethical principles and commit to professional ethics and responsibilities and norms of the engineering practice.
9. **Individual and team work:** Function effectively as an individual, and as a member or leader in diverse teams, and in multidisciplinary settings.
10. **Communication:** Communicate effectively on complex engineering activities with the engineering community and with society at large, such as, being able to comprehend and write effective reports and design documentation, make effective presentations, and give and receive clear instructions.
11. **Project management and finance:** Demonstrate knowledge and understanding of the engineering and management principles and apply these to one's own work, as a member and leader in a team, to manage projects and in multidisciplinary environments.
12. **Life-long learning:** Recognize the need for and have the preparation and ability to engage in independent and life-long learning in the broadest context of technological change.

PROGRAMME OUTCOMES	Mapping HIGH: 3/ MEDIUM: 2/LOW: 1	JUSTIFICATION
1	**2**	To apply the knowledge of mathematics.
2	**3**	By considering the problems of dynamic price prediction and analyze those problems and developing Algorithms.
3	**3**	This project meets the desired specification of the society.
4	**2**	We have created this user interface by using Flask.
5	**3**	By using the Python and Pandas we have created data-frames for storing data and user friendly interface.
6	**3**	In this developing process we were able to meet the local challenges as well as global challenges.
7	**2**	This interface does not provide benefits to all types of users which help for the society.
8	**2**	It will provide some ethical, social behavior in some aspects.
9	**2**	The work is done by team to function on multi-disciplinary team.
10	**2**	As our project is done in all aspects like communicating and documenting effectively.
11	**3**	Our project is developed by Python Programming language and it will engage in lifelong learning.
12	**3**	We find a solution to our problem by developing an application, which is effective for financial management.

PROGRAM SPECIFIC OUTCOMES

PSOs	1	2
PROJECT	3	2

Student will be able to

1. Organize, maintain and protect IT Infrastructural resources.
2. Design and Develop web, mobile, and smart apps based software solutions to the real world problems.

PROGRAM SPECIFIC OUTCOMES	Mapping HIGH:3/ MEDIUM: 2/LOW:1	JUSTIFICATION
1	**3**	Predicting the liver disease in a patient is a challenge. We made it easy by developing this project.
2	**2**	Software usage in liver disease prediction can reduce the time and increase efficiency.