A REVIEW OF LIVER PATIENT ANALYSIS METHODS USING MACHINE LEARNING

Submitted in partial fulfillment of requirement for the award of the Degree

Bachelor of Computer Science

In the faculty of Computer Science of Bharathiar University, Coimbatore

Submitted by

TEAM ID: NM2023TMID17436

SANGEETHA D

LOGESHWARI M

VAISHNAVI N

SIVASHANKARI U

Under the guidance of

Dr.V.PUNITHAVATHI MCA,M.Phil,DOA.,Ph.D.

Lecturer , Department of Computer Science

DEPARTMENT OF COMPUTER SCIENCE

L.R.G GOVERNMENT ARTS COLLEGE FOR WOMEN

(Affiliated To Bharathiar University)

TIRUPUR-4

APRIL-2023

LRG GOVERNMENT ARTS COLLEGE

NAAN MUDHALVAN PROJECT WORK

(AFFILIATED TO BHARATHIAR UNIVERSITY)

TIRUPUR-641602

TITLE :A REVIEW OF LIVER PATIENT ANALYSIS METHODS USING MACHINE LEARNING

This is to certify that this is a bonafide record of work done by the above students of III B.Sc Computer Science Degree **NAAN MUDHALVAN PROJECT** during the year 2022-23

Submitted for the Naan Mudhalvan project work held on April 2023

CLASS MENTOR **HEAD OF DEPARTMENT**

[Dr.V.PUNITHAVATHI] **[Dr.R.Parimala]**

INDEX

1.1 INTRODUCTION

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 complicate 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 in the prediction of liver disease and can be recommended to the user.

2. PROBLEM SELECTION

Selecting the right problem for liver patient analysis using machine learning can have a significant impact on the success of your project. Here are a few potential problem statements for this type of analysis:

1. **Predicting liver disease:** Given a dataset of liver patient records, build a machine learning model that can accurately predict whether a patient has liver disease or not. This can be framed as a binary classification problem, where the goal is to distinguish between patients with liver disease and those without.

2. **Severity prediction:** Given a dataset of liver patient records, build a machine learning model that can predict the severity of the disease. This can be framed as a regression problem, where the goal is to predict a continuous value, such as the patient's liver function score.

3. **Treatment recommendation:** Given a dataset of liver patient records and their treatment outcomes, build a machine learning model that can recommend the most effective treatment for a new patient. This can be framed as a recommendation problem, where the goal is to suggest a treatment that maximizes the patient's chances of recovery.

4. **Risk factor identification:** Given a dataset of liver patient records, build a machine learning model that can identify the most significant risk factors for liver disease.

3. IDEATION

Analyzing liver disease patient data using machine learning can provide valuable insights into the factors that contribute to liver disease and help healthcare professionals identify patients at high risk for developing liver disease or its complications. Here are some ideas for analyzing liver patient data using machine learning:

- **Predicting liver disease**: You can build a predictive model using machine learning algorithms that analyze patient data, such as age, gender, family history, blood test results, and lifestyle factors, to predict the likelihood of a patient developing liver disease. This can help healthcare professionals identify patients at high risk for liver disease and provide early intervention.

- **Identifying risk factors:** You can use machine learning to identify the risk factors that contribute to liver disease. This can be done by analyzing patient data to identify common patterns and trends that may be associated with liver disease, such as alcohol consumption, obesity, or viral infections.

- **Early detection of liver disease:** You can develop a machine learning model that analyzes patient data to detect early signs of liver disease. This can be done by analyzing blood test results, imaging scans, and

other diagnostic tests to identify changes that may indicate the presence of liver disease.

- **Treatment optimization:** You can use machine learning to optimize treatment for patients with liver disease. This can be done by analyzing patient data to identify which treatments are most effective for specific types of liver disease and for patients with specific risk factors.
- **Patient outcome prediction**: You can develop a machine learning model that predicts the likely outcome of liver disease for patients based on their medical history, lifestyle factors, and other clinical data. This can help healthcare professionals determine the most appropriate treatment plan for individual patients and monitor their progress over time.

4.REQUIREMENT ANANLYSIS

Requirement analysis in liver patient analysis using machine learning involves understanding the problem at hand, defining the objectives and constraints of the analysis, and identifying the data requirements and machine learning techniques that can be applied to address the problem. Here are some key steps in requirement analysis for liver patient analysis using machine learning:

- **Problem Understanding:** The first step is to understand the problem of liver patient analysis, which involves predicting whether a patient has liver disease based on various clinical and demographic features. It is important to understand the context of the problem, such as the purpose of the analysis, the target audience, and the expected outcomes.

- **Objective Definition:** Once the problem is understood, the next step is to define the objectives of the analysis. This includes identifying the metrics that will be used to measure the performance of the machine learning model, such as accuracy, precision, recall, or F1-score. Additionally, it may be necessary to define the constraints of the analysis, such as data privacy or regulatory requirements.

- **Data Collection and Preparation:** The success of machine learning models depends on the quality of data used to train and validate them. Hence, it is important to identify the data sources, collect relevant data, and prepare it for analysis. This may involve data cleaning, normalization, and

transformation to ensure that the data is in a suitable format for machine learning algorithms.

- **Feature Selection:** Feature selection is the process of identifying the most relevant features that are predictive of liver disease. This can be done using various statistical and machine learning techniques, such as correlation analysis, principal component analysis, or recursive feature elimination.

- **Model Selection and Evaluation**: The next step is to select suitable machine learning models for the analysis, such as logistic regression, decision trees, or support vector machines. These models should be evaluated using appropriate performance metrics and cross-validation techniques to ensure that they are robust and generalize well to new data.

- **Deployment and Monitoring:** Once the machine learning model is developed and tested, it should be deployed in a production environment.
- It is important to monitor the performance of the model over time, and update it as necessary to ensure that it continues to provide accurate predictions. This may involve re-training the model with new data or fine-tuning its parameters to optimize its performance.

5.PROJECT DESIGN

Designing a project to analyze liver patient data using machine learning can be a valuable contribution to healthcare. Here is a general outline of the project design:

- **Data Collection:** Collect relevant data related to liver patients from multiple sources. This may include hospital records, clinical trials, or publicly available datasets. It is essential to ensure that the data is comprehensive and includes various aspects such as demographic information, medical history, lab results, and imaging reports.

- **Data Preprocessing:** Cleaning and pre-processing the data is a critical step in the machine learning pipeline. This may include data cleaning, data normalization, missing data handling, and outlier detection.

- **Feature Engineering:** After pre-processing the data, we need to extract relevant features that could help the machine learning model to learn the underlying patterns in the data. Feature engineering can include domain-specific knowledge, such as medical expertise, to extract meaningful features.

- **Model Selection:** Next, we need to select an appropriate machine learning algorithm that best suits the dataset and the problem statement.

This may include supervised or unsupervised learning algorithms. Some of the commonly used algorithms for healthcare analysis include logistic regression, decision trees, random forests, and neural networks.

- **Model Training and Validation:** After selecting the model, we need to train the model using the training dataset and validate its performance using the validation dataset. We need to ensure that the model does not overfit or underfit the data.

- **Model Evaluation:** Once the model is trained and validated, we need to evaluate its performance on the test dataset. We may use metrics such as accuracy, precision, recall, and F1 score to evaluate the model's performance.

- **Deployment:** Finally, we need to deploy the model in a real-world scenario. This may include integrating the model with a web application or healthcare information system.

6. PROJECT PLANNING PHASE

Developing a project plan is an essential step to ensure that your liver patient analysis project using machine learning is successful. Here are some steps that can help you in creating a project plan:

1. **Define the project goal:** The first step is to define the project's goal, which in your case, would be to develop a machine learning model that can accurately predict the risk of liver disease in patients.

2. **Gather and analyze data:** Once you have defined the project goal, the next step is to gather the necessary data. This could include medical records, lab reports, and imaging data. You will also need to analyze the data to identify any patterns or correlations that can be used to train the machine learning model.

3. **Identify machine learning algorithms:** After analyzing the data, you will need to identify the appropriate machine learning algorithms that can be used to train the model. This could include supervised learning algorithms such as logistic regression, decision trees, and neural networks.

4. **Train and test the model:** With the algorithms selected, you can proceed to train the machine learning model using the gathered data. It is essential to split the data into training and testing sets to evaluate the model's accuracy.

5. **Evaluate and optimize the model:** After training the model, you need to evaluate its performance against the testing dataset. If the model's accuracy is not satisfactory, you may need to tweak the algorithms or the model's parameters to improve its accuracy.

6. **Deploy the model:** Once you have optimized the model, you can deploy it into production, making it available for use by healthcare professionals.

7. **Monitor and maintain the model:** As the model is deployed, you will need to monitor its performance regularly and make necessary adjustments to maintain its accuracy.

8. **Communicate the results:** Finally, you need to communicate the results of your analysis to healthcare professionals, ensuring that they understand the model's limitations and its potential applications.

These are some essential steps to consider while developing a project plan for liver patient analysis using machine learning.

7. PROJECT DEVELOPMENT PHASE

The development phase for a liver patient analysis project can be broken down into several key steps:

1. **Defining the problem:** The first step is to clearly define the problem you are trying to solve. For a liver patient analysis project, this may involve identifying the key factors that contribute to liver disease and developing a model to predict the likelihood of a patient developing

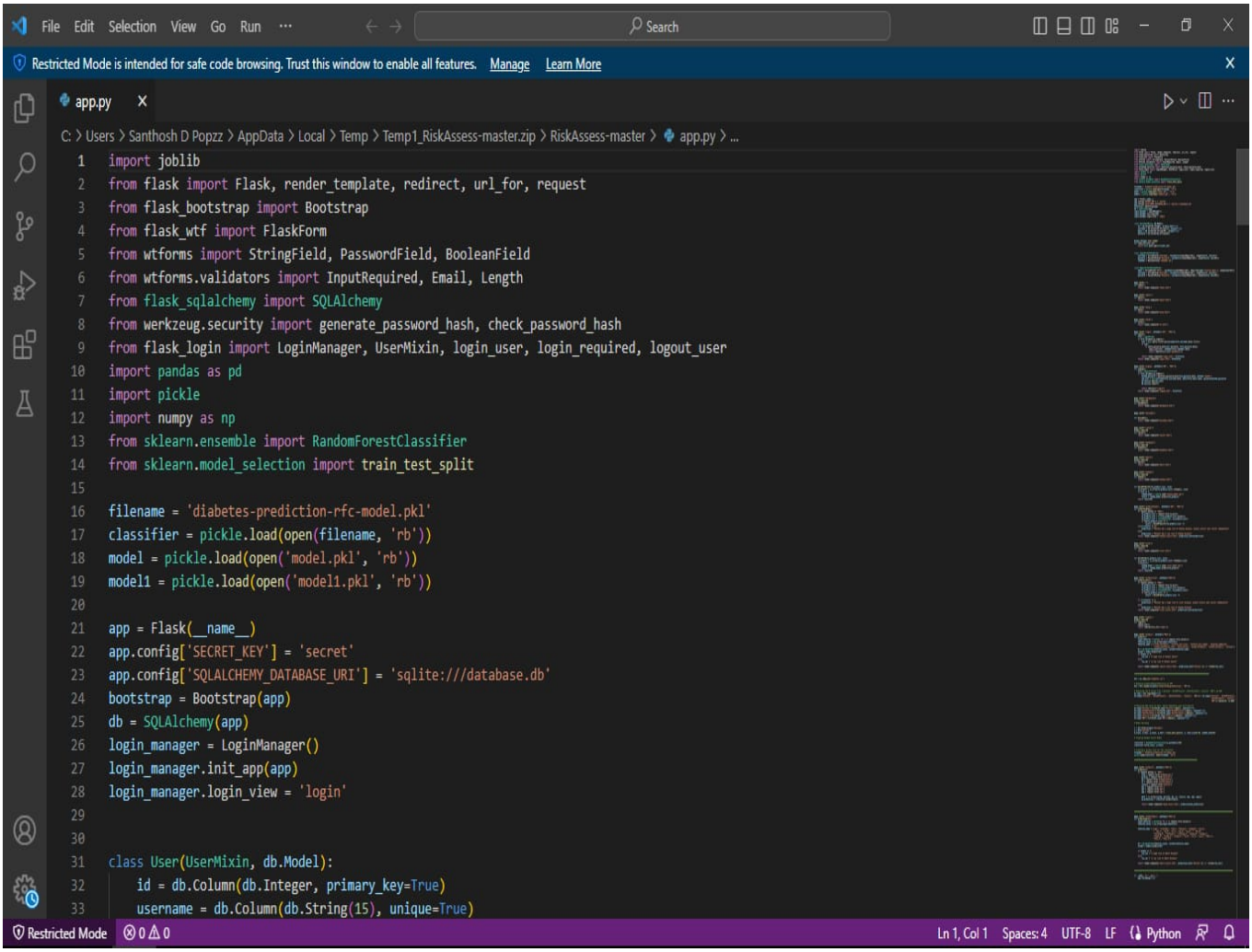
2. **Gathering data:** The next step is to gather relevant data for your project. This may include medical records, lab test results, patient demographics, and other relevant information.

3. **Data preprocessing:** Once you have gathered the data, you will need to preprocess it to make it suitable for analysis. This may involve cleaning the data, removing missing values, and transforming the data into a format that can be used by your analysis tools.

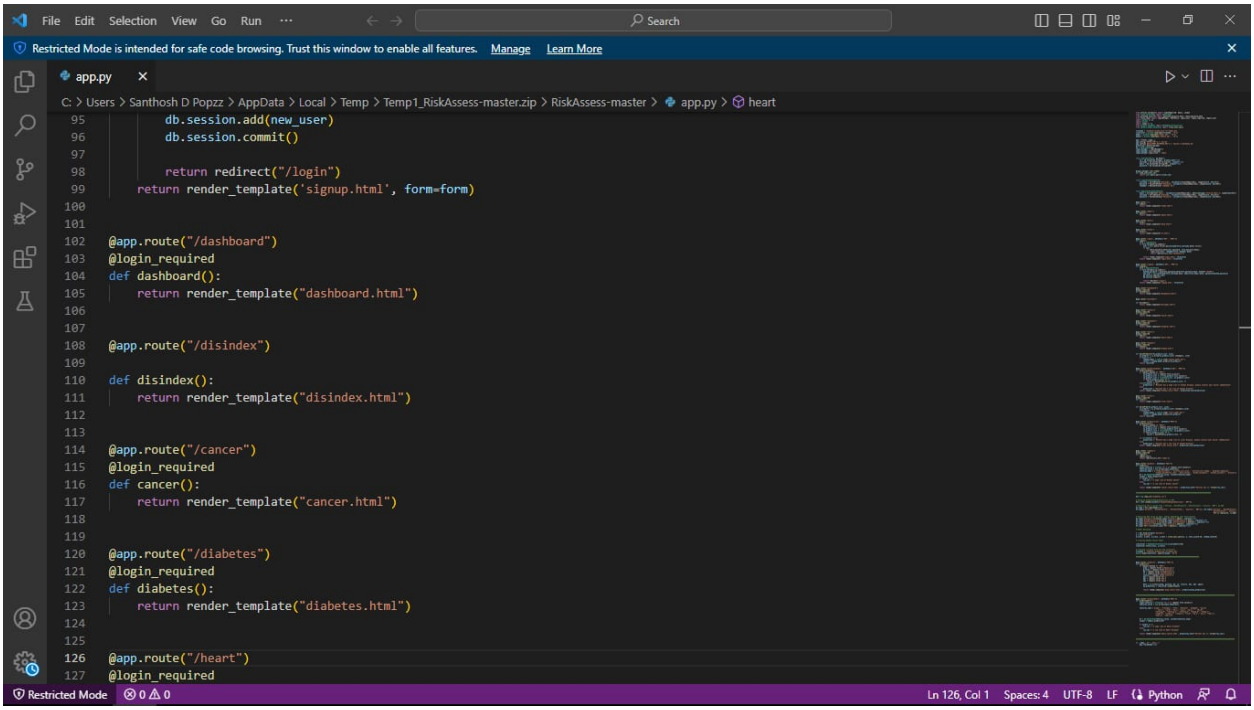
4. **Feature engineering:** Feature engineering involves selecting the most relevant features from your data and creating new features that may be useful for your analysis. In the context of a liver patient analysis project, this may involve identifying key risk factors for liver disease and creating new features that capture these risk factors.

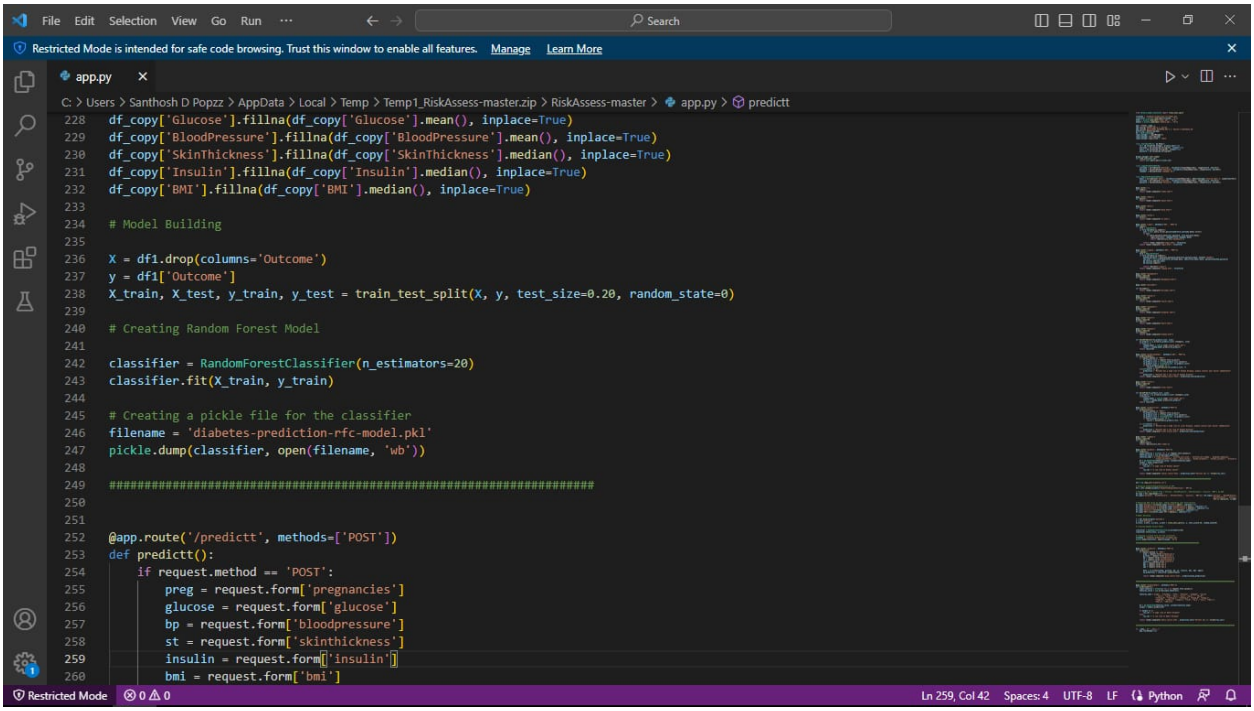
8. RESULTS AND DISCUSSION

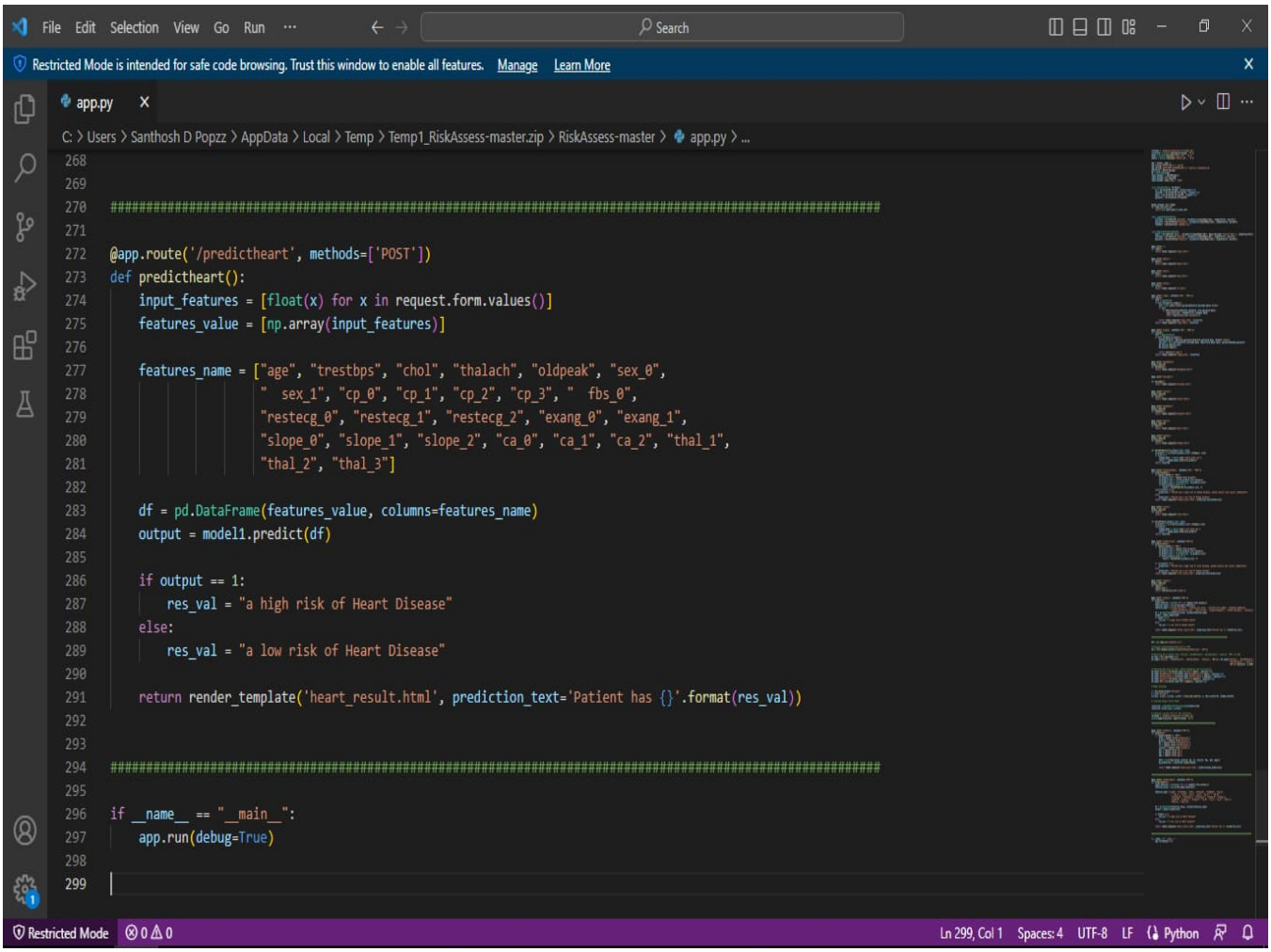
SAMPLE CODING:



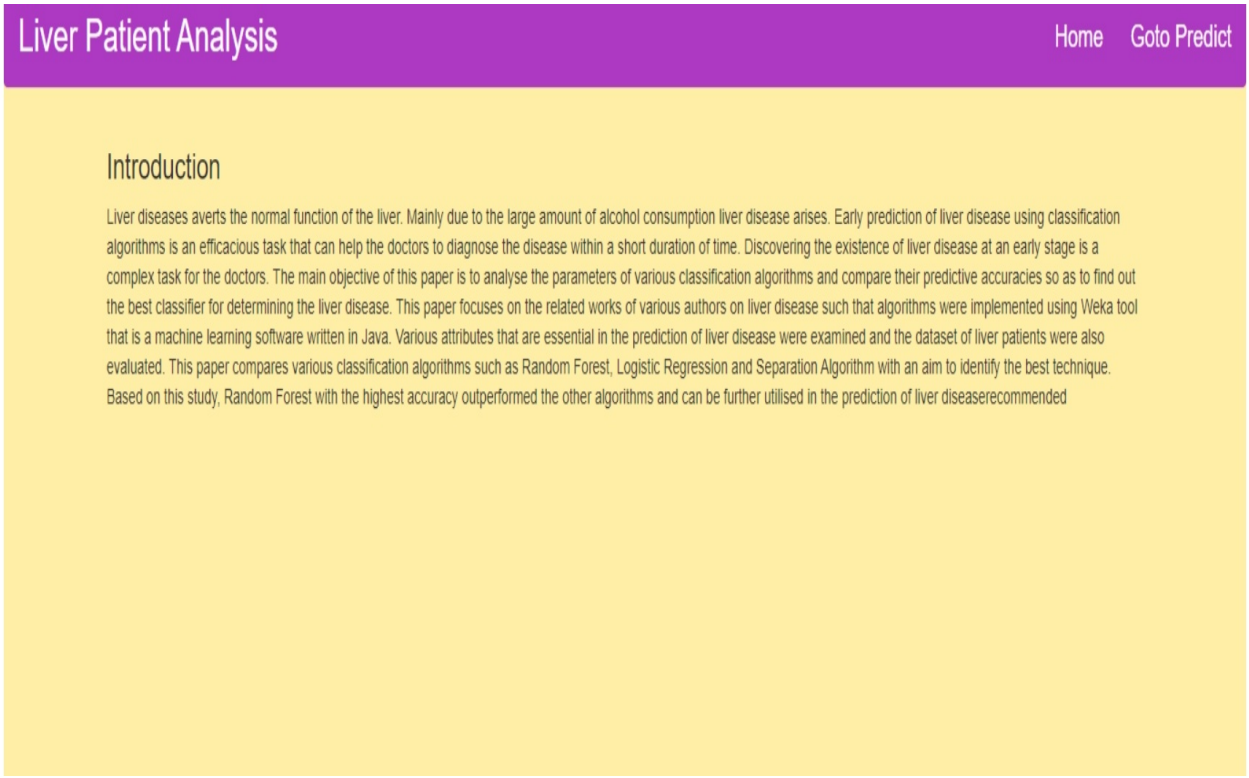
```
import joblib
from flask import Flask, render_template, redirect, url_for, request
from flask_bootstrap import Bootstrap
from flask_wtf import FlaskForm
from wtforms import StringField, PasswordField, BooleanField
from wtforms.validators import InputRequired, Email, Length
from flask_sqlalchemy import SQLAlchemy
from werkzeug.security import generate_password_hash, check_password_hash
from flask_login import LoginManager, UserMixin, login_user, login_required, logout_user
import pandas as pd
import pickle
import numpy as np
from sklearn.ensemble import RandomForestClassifier
from sklearn.model_selection import train_test_split

filename = 'diabetes-prediction-rfc-model.pkl'
classifier = pickle.load(open(filename, 'rb'))
model = pickle.load(open('model.pkl', 'rb'))
model1 = pickle.load(open('model1.pkl', 'rb'))

app = Flask(__name__)
app.config['SECRET_KEY'] = 'secret'
app.config['SQLALCHEMY_DATABASE_URI'] = 'sqlite:///database.db'
bootstrap = Bootstrap(app)
db = SQLAlchemy(app)
login_manager = LoginManager()
login_manager.init_app(app)
login_manager.login_view = 'login'


class User(UserMixin, db.Model):
    id = db.Column(db.Integer, primary_key=True)
    username = db.Column(db.String(15), unique=True)
```


File Edit Selection View Go Run ···
Search
Restricted Mode is intended for safe code browsing. Trust this window to enable all features. Manage Learn More
app.py
C: > Users > Santhosh D Popzz > AppData > Local > Temp > Temp1_RiskAssess-master.zip > RiskAssess-master > app.py > heart
db.session.add(new_user)
db.session.commit()
return redirect("/login")
return render_template('signup.html', form=form)
@app.route("/dashboard")
@login_required
def dashboard():
return render_template("dashboard.html")
@app.route("/disindex")
def disindex():
return render_template("disindex.html")
@app.route("/cancer")
@login_required
def cancer():
return render_template("cancer.html")
@app.route("/diabetes")
@login_required
def diabetes():
return render_template("diabetes.html")
@app.route("/heart")
@login_required
Restricted Mode
Ln 126, Col 1 Spaces: 4 UTF-8 LF Python


File Edit Selection View Go Run ···
Search
Restricted Mode is intended for safe code browsing. Trust this window to enable all features. Manage Learn More
app.py
C: > Users > Santhosh D Popzz > AppData > Local > Temp > Temp1_RiskAssess-master.zip > RiskAssess-master > app.py > predictt
df_copy['Glucose'].fillna(df_copy['Glucose'].mean(), inplace=True)
df_copy['BloodPressure'].fillna(df_copy['BloodPressure'].mean(), inplace=True)
df_copy['SkinThickness'].fillna(df_copy['SkinThickness'].median(), inplace=True)
df_copy['Insulin'].fillna(df_copy['Insulin'].median(), inplace=True)
df_copy['BMI'].fillna(df_copy['BMI'].median(), inplace=True)
# Model Building
X = df1.drop(columns='Outcome')
y = df1['Outcome']
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.20, random_state=0)
# Creating Random Forest Model
classifier = RandomForestClassifier(n_estimators=20)
classifier.fit(X_train, y_train)
# Creating a pickle file for the classifier
filename = 'diabetes-prediction-rfc-model.pkl'
pickle.dump(classifier, open(filename, 'wb'))
@app.route('/predictt', methods=['POST'])
def predictt():
if request.method == 'POST':
preg = request.form['pregnancies']
glucose = request.form['glucose']
bp = request.form['bloodpressure']
st = request.form['skinthickness']
insulin = request.form['insulin']
bmi = request.form['bmi']
Restricted Mode
Ln 259, Col 42 Spaces: 4 UTF-8 LF Python


File Edit Selection View Go Run ···
Search
Restricted Mode is intended for safe code browsing. Trust this window to enable all features. Manage Learn More
app.py
C: > Users > Santhosh D Popzz > AppData > Local > Temp > Temp1_RiskAssess-master.zip > RiskAssess-master > app.py > ...
268
269
270 ##
271
272 @app.route('/predictheart', methods=['POST'])
273 def predictheart():
274 input_features = [float(x) for x in request.form.values()]
275 features_value = [np.array(input_features)]
276
277 features_name = ["age", "trestbps", "chol", "thalach", "oldpeak", "sex_0",
278 " sex_1", "cp_0", "cp_1", "cp_2", "cp_3", " fbs_0",
279 "restecg_0", "restecg_1", "restecg_2", "exang_0", "exang_1",
280 "slope_0", "slope_1", "slope_2", "ca_0", "ca_1", "ca_2", "thal_1",
281 "thal_2", "thal_3"]
282
283 df = pd.DataFrame(features_value, columns=features_name)
284 output = model1.predict(df)
285
286 if output == 1:
287 res_val = "a high risk of Heart Disease"
288 else:
289 res_val = "a low risk of Heart Disease"
290
291 return render_template('heart_result.html', prediction_text='Patient has {}'.format(res_val))
292
293
294 ##
295
296 if __name__ == "__main__":
297 app.run(debug=True)
298
299
Restricted Mode ⊗ 0 ⚠ 0
Ln 299, Col 1 Spaces: 4 UTF-8 LF Python

OUTPUT:


Liver Patient Analysis
Home
Goto Predict
Introduction
Liver diseases averts the normal function of the liver. Mainly due to the large amount of alcohol consumption liver disease arises. 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. Discovering the existence of liver disease at an early stage is a complex task for the doctors. The main objective of this paper is to 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 paper focuses on the related works of various authors on liver disease such that algorithms were implemented using Weka tool that is a machine learning software written in Java. Various attributes that are essential in the prediction of liver disease were examined and the dataset of liver patients were also evaluated. This paper compares various classification algorithms such as Random Forest, Logistic Regression and Separation 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 in the prediction of liver diseaserecommended

Liver Patient Prediction

Age:

Gender:

Enter 0 as male, 1 as female

Total_Bilirubin:

Direct_Bilirubin:

Alkaline_Phosphotase:

Alamine_Aminotransferase:

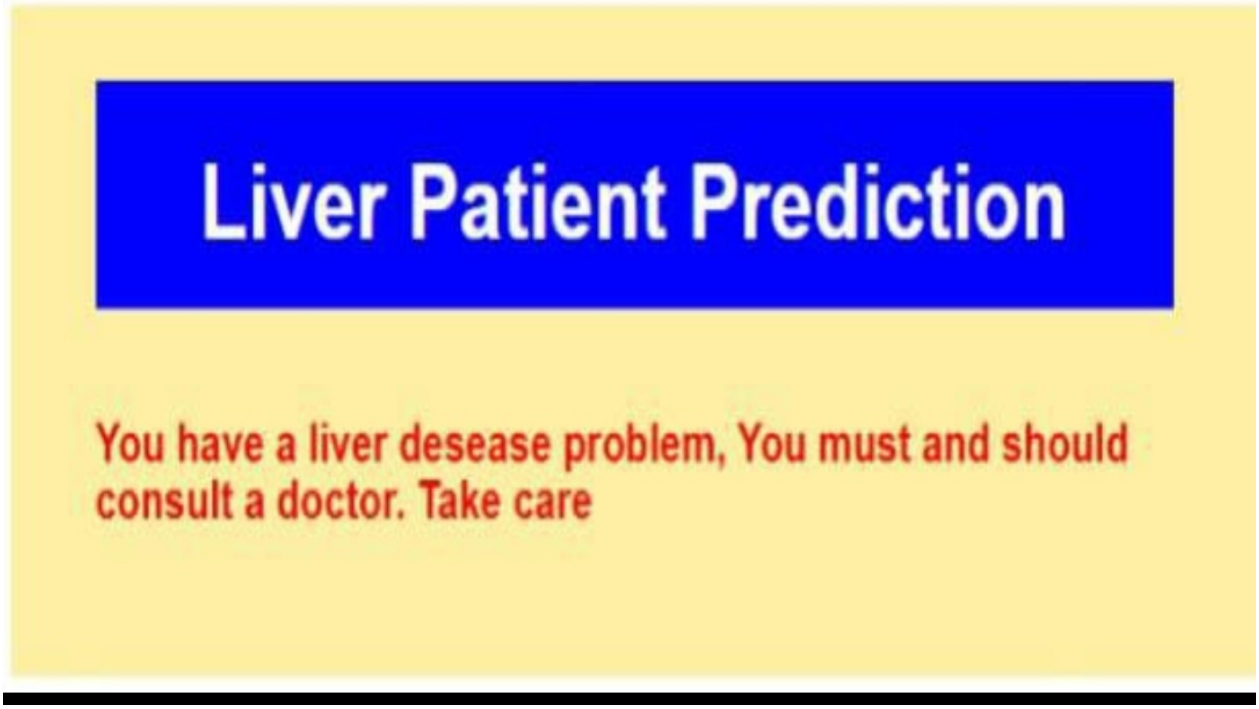
Aspartate_Aminotransferase:

Total_Protiens:

Albumin:

Albumin_and_Globulin_Ratio:

Predict


Liver Patient Prediction
You have a liver desease problem, You must and should consult a doctor. Take care

APPENDICES

Video Link: https://youtube.com/watch?v=S5J2CNV-bd8&feature=share

Project Link: https://github.com/vaishnavilrg/A-Review-of-Liver-Patient-Analysis-Methods-Using-Machine-Learning