Parshvanath Charitable Trust's
A. P. SHAH INSTITUTE OF TECHNOLOGY, THANE
(All Programs Accredited by NBA)
Department of Information Technology

Hepatitis B Mortality prediction

1.Aagya Singh (20104076)
2. Aditya Waingade (20104099)
3. Devesh Zope (20104086)

Project Guide

Prof .Roshna Sangle

.

Contents

- Introduction
- Objectives
- Scope
- Literature Survey
- Proposed System
- Project Outcomes
- Block Diagram
- Use Case/DFD
- Technology Stack

1.Introduction:

Problem Identified:

- Predicting mortality risk in patients with hepatitis B a viral infection that can cause liver damage and cancer so it's necessary to predict the future severity prior so that the healthcare professionals can determine appropriate treatment plans and improve patient outcome.

Solution proposed:

- This model can potentially help clinicians identify high-risk patients early on and tailor treatment plans accordingly improving overall patient outcomes.
- This model is trained using a dataset of clinical and demographic variables from patients with HepatitisB, including age , gender , liver function tests, etc

2.Objectives:

1. To identify patients who are at an increased risk of mortality or death from complications associated with the infection.

2. To find out the probability of death based on various clinical and demographic factors such age, sex, ascities, bilirubin content etc.

3. To improve patient care at an early stage by prior prediction.

4. To provide healthcare providers with information to guide their clinical-decision making to improve patient survival rate.

3. Scope

1. Identifying high-risk individuals who may benefit from closer monitoring, more frequent testing, and earlier intervention.
2. Estimates the burden of hepatitis B-related mortality and potential impact of the disease .
3. Early stage identification and management of individuals at risk of liver disease progression.
4. Reducing the overall burden of the disease’s mortality.

4. Literature Survey

Title	Author	Algorithm's Used	Result
Development of a system for Predicting Mortality in Hepatitis B related liver failure	Xu et al. (2020)	Logistic Regression, Decision Tree	Identified Serum bilirubin, prothrombin time as the most important predictors of mortality.
Predicting the survival of Hepatitis B virus-related liver failure patients	Zhang et al.	Decision tree	Identified bilirubin content and age as the most important predictors of mortality.

5. Proposed System

1. Data Collection ,cleaning and preprocessing: Collecting data from multiple resources including electronic medical record, laboratory results, demographics, and clinical notes.

2. Feature selection: It is an important step in mortality prediction as it helps to identify the most relevant variables or features that can accurately predict the outcome.

3. Model Selection and model training: In our system we have used logistic regression, KNN, decision tree to build a model and compare the accuracies of each model.

4. Deployment: Model deployment has been carried out through streamlit and can be used by healthcare professionals to predict the mortality risk in Hepatitis B patients.

6. Algorithm used

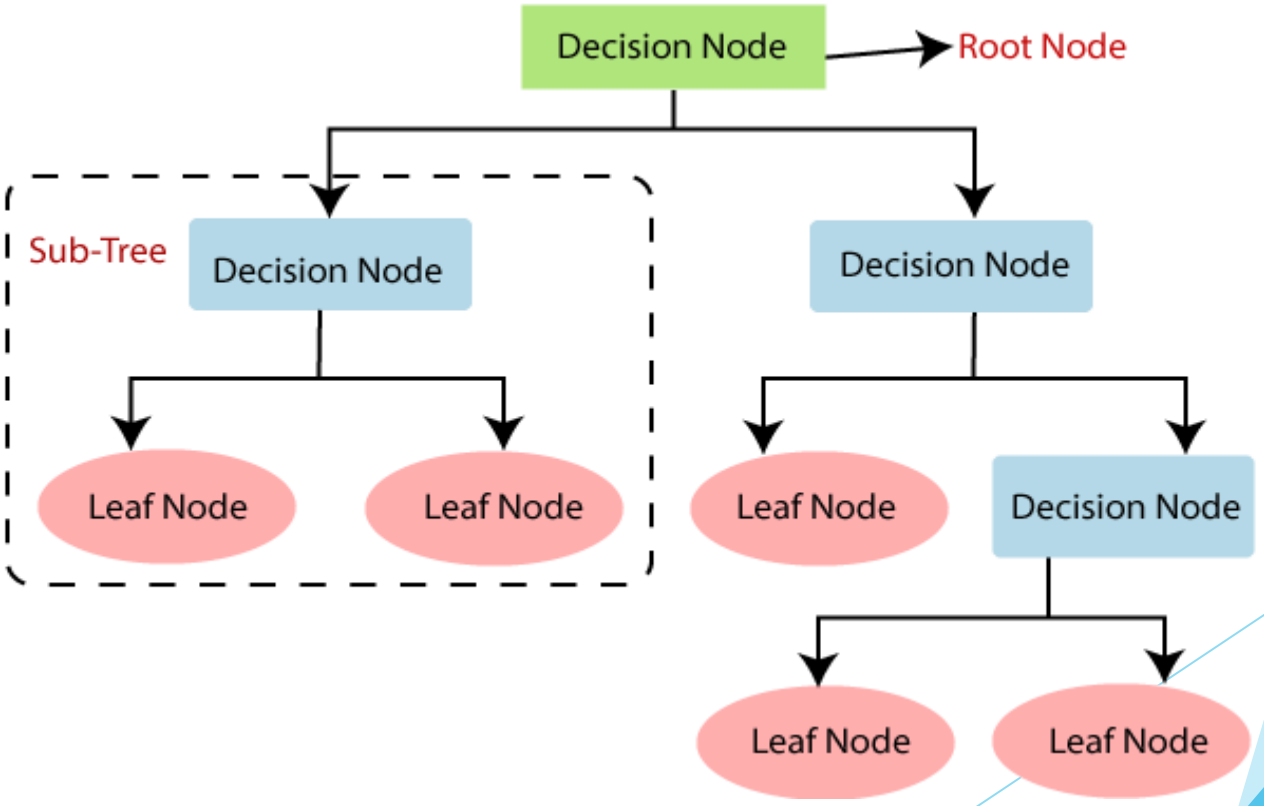
1. Logistic Regression: It is a supervised learning technique. It is used for classification problems. It gives the probabilistic values which lie between 0 and 1.

2. KNN: It is a supervised learning technique. It assumes the similarity between the new case data and available cases and put the new case into the category that is most similar to the available categories.

3. Decision Tree: It is a tree-structured classifier, where internal nodes represent the features of a dataset, branches represent the decision rules and each leaf node represents the outcome.

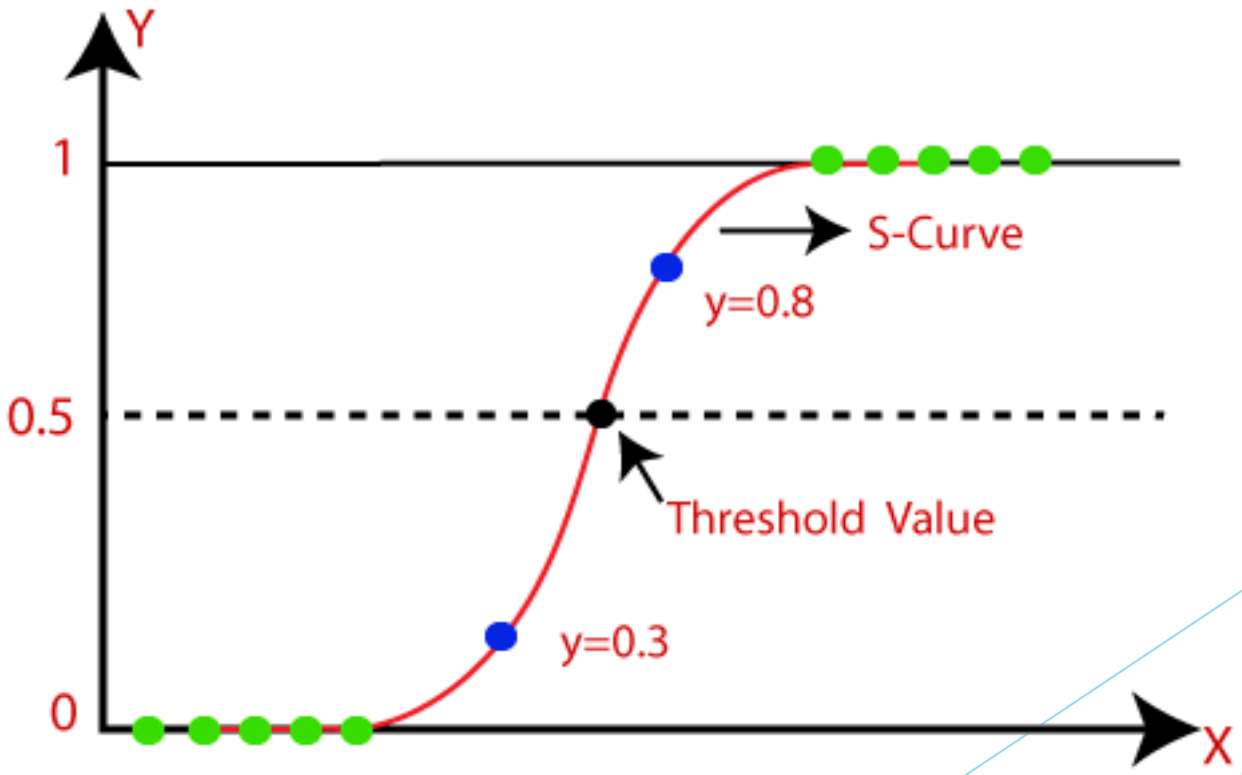
6.1 Decision Tree

It is a tree-structured classifier, where internal nodes represent the features of a dataset, branches represent the decision rules and each leaf node represents the outcome.



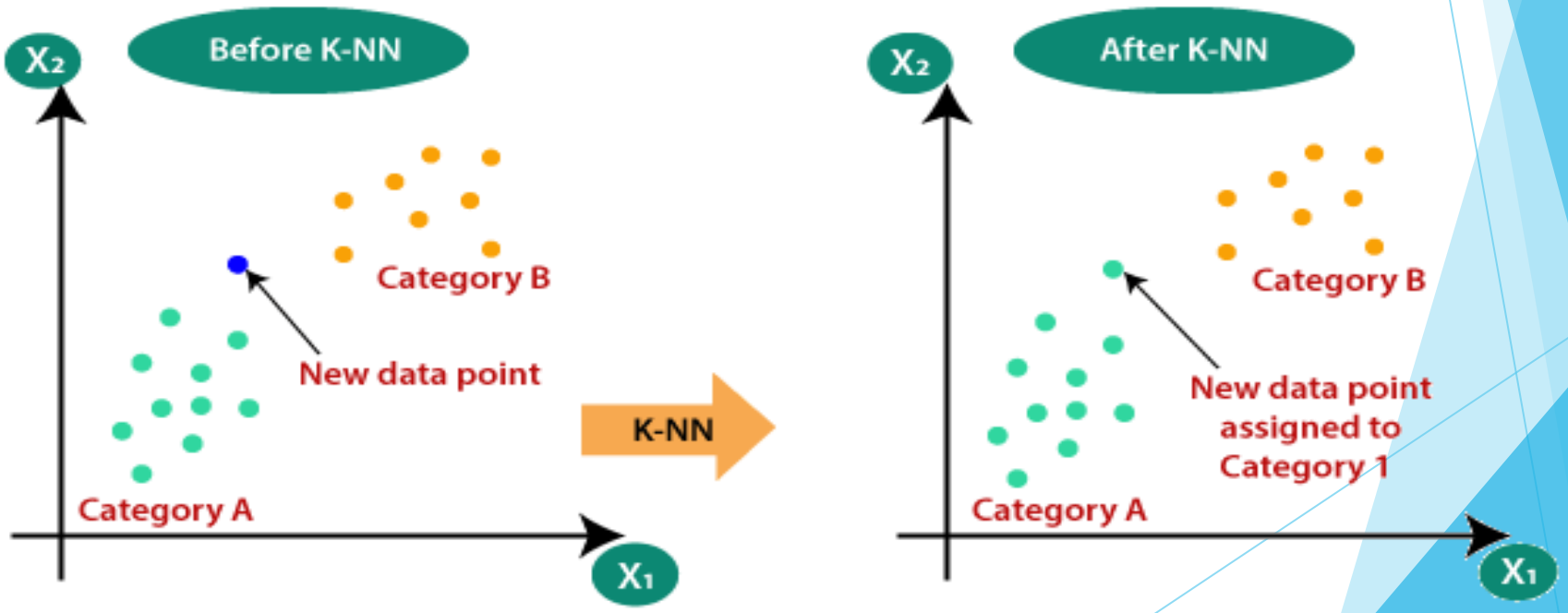
6.2 Logistic Regression

Logistic Regression is a supervised learning technique. It is used for classification problems. It gives the probabilistic values which lie between 0 and 1.



6.3 KNN

KNN: It is a supervised learning technique. It assumes the similarity between the new case data and available cases and put the new case into the category that is most similar to the available categories.



7. Project Outcomes

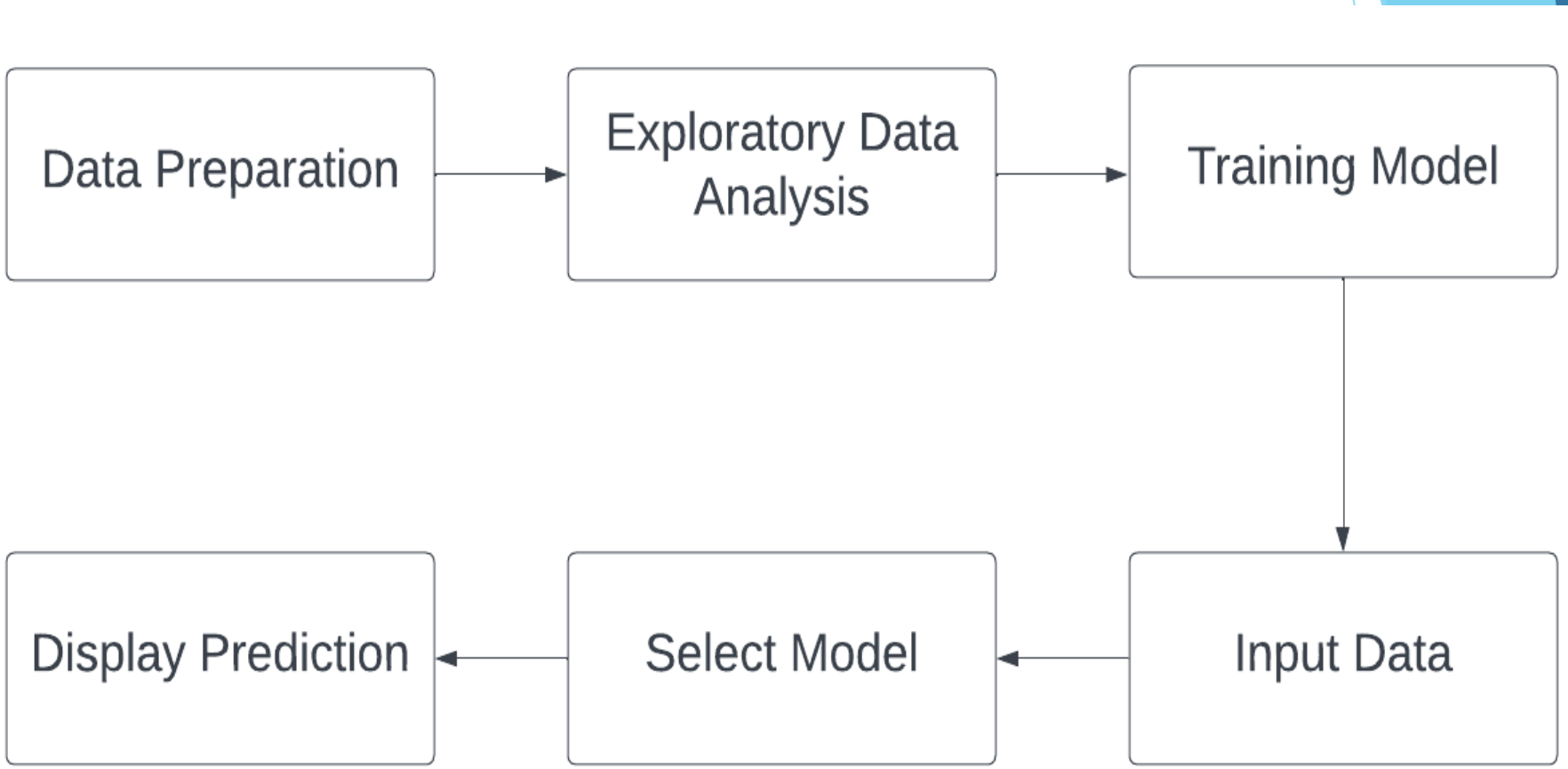
1. Users will have the ability to detect the disease at an early stage where treatment is more likely to be effective.

2. Reduce diagnostic errors that can occur due to human bias or errors in interpretation of medical data.

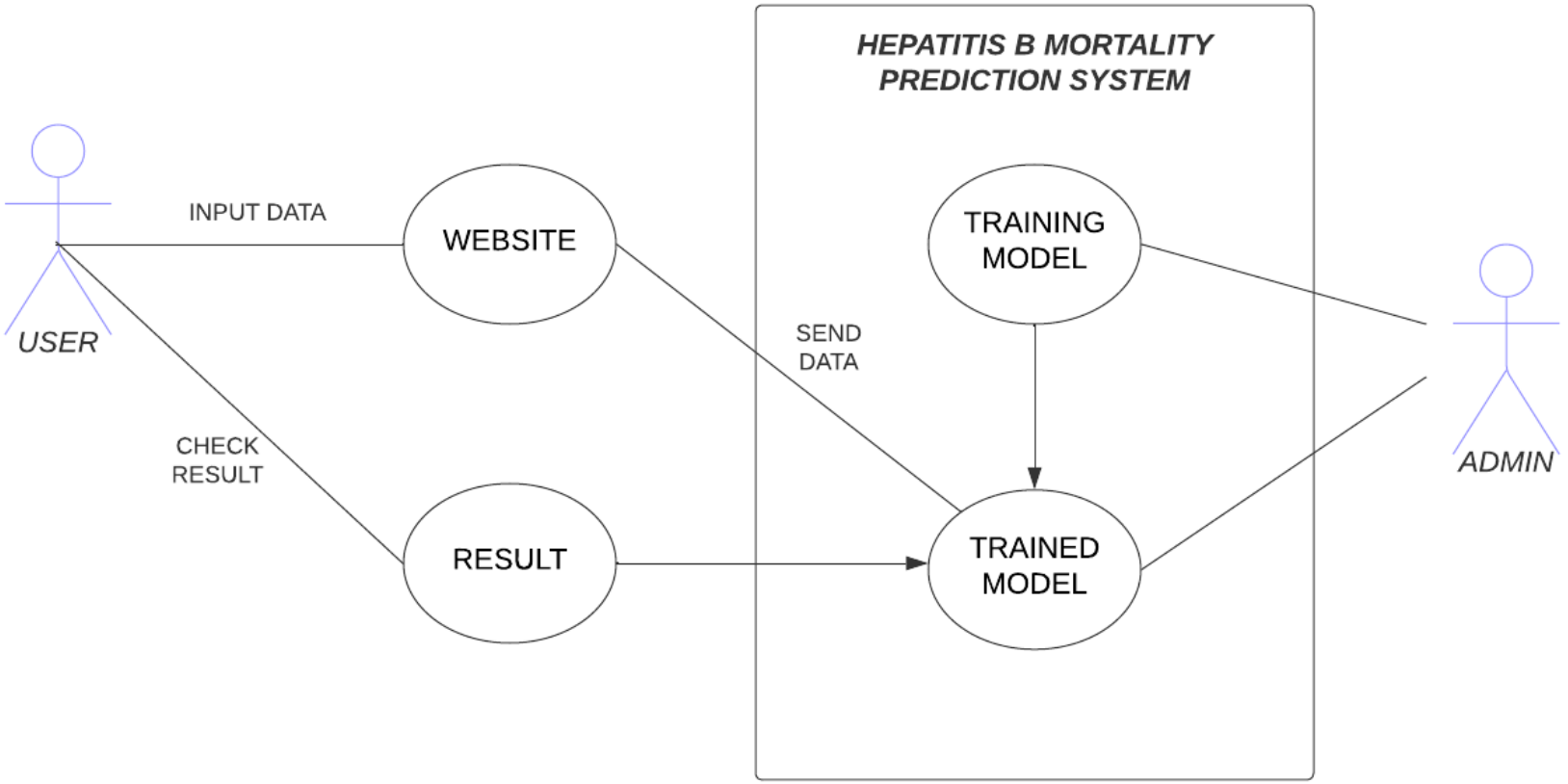
3. The project can lead to the development of a user-friendly platform or an application that can be integrated into existing healthcare and workflows.

4. Users or patients will be able to reduce health-care costs associated with the disease.

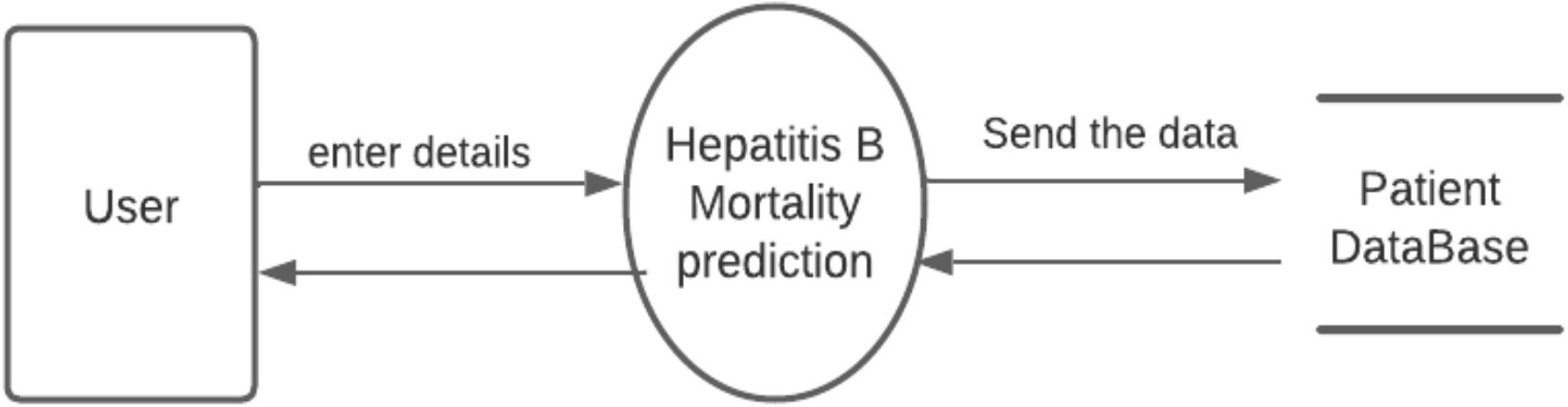
8. Block Diagram



9. Use Case

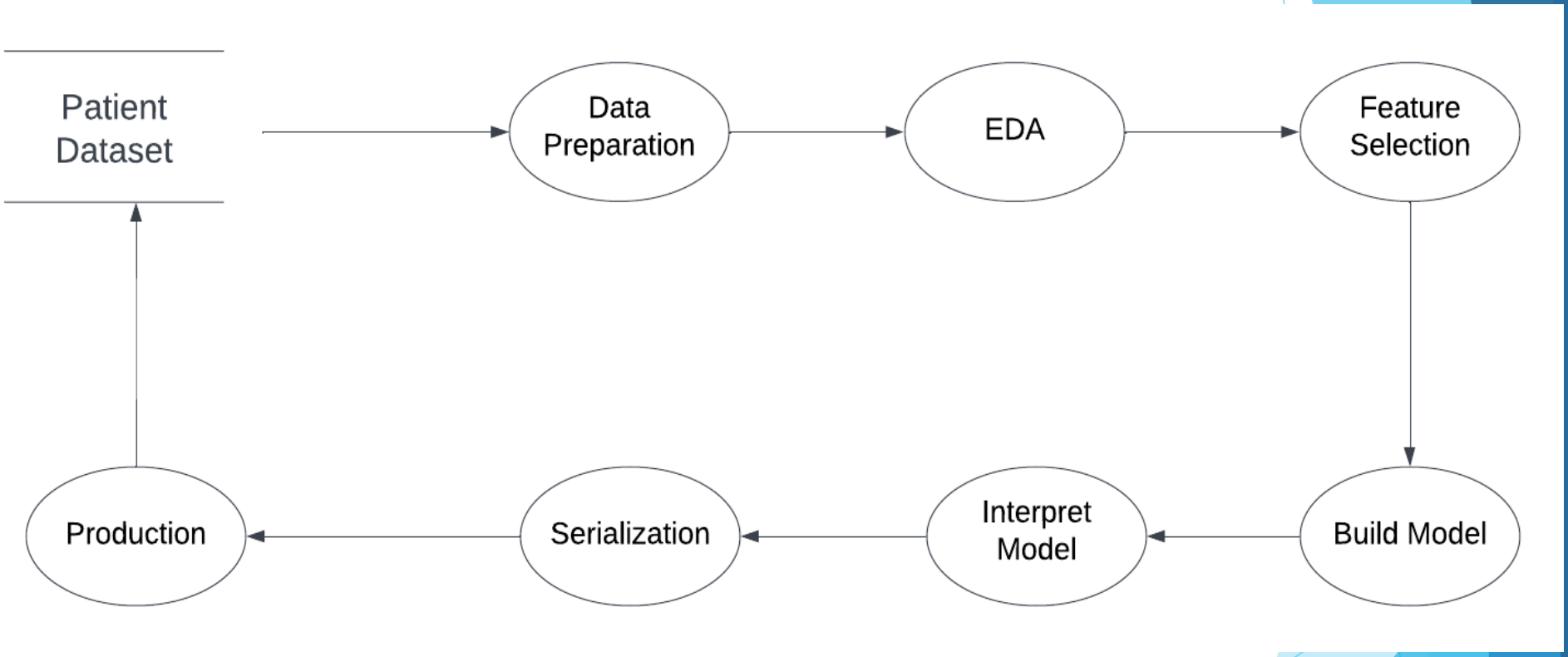

HEPATITIS B MORTALITY PREDICTION SYSTEM
USER
INPUT DATA
WEBSITE
CHECK RESULT
RESULT
SEND DATA
TRAINING MODEL
TRAINED MODEL
ADMIN

10. DFD



Level 0 - DFD

10.1 DFD


Patient Dataset
Data Preparation
EDA
Feature Selection
Build Model
Interpret Model
Serialization
Production

11. Technology Stack

1.Programming Language : Python

2.IDE : Visual Studio Code, Jupyter Notebook

3.Libraries: Numpy, MatplotLib, pandas, etc

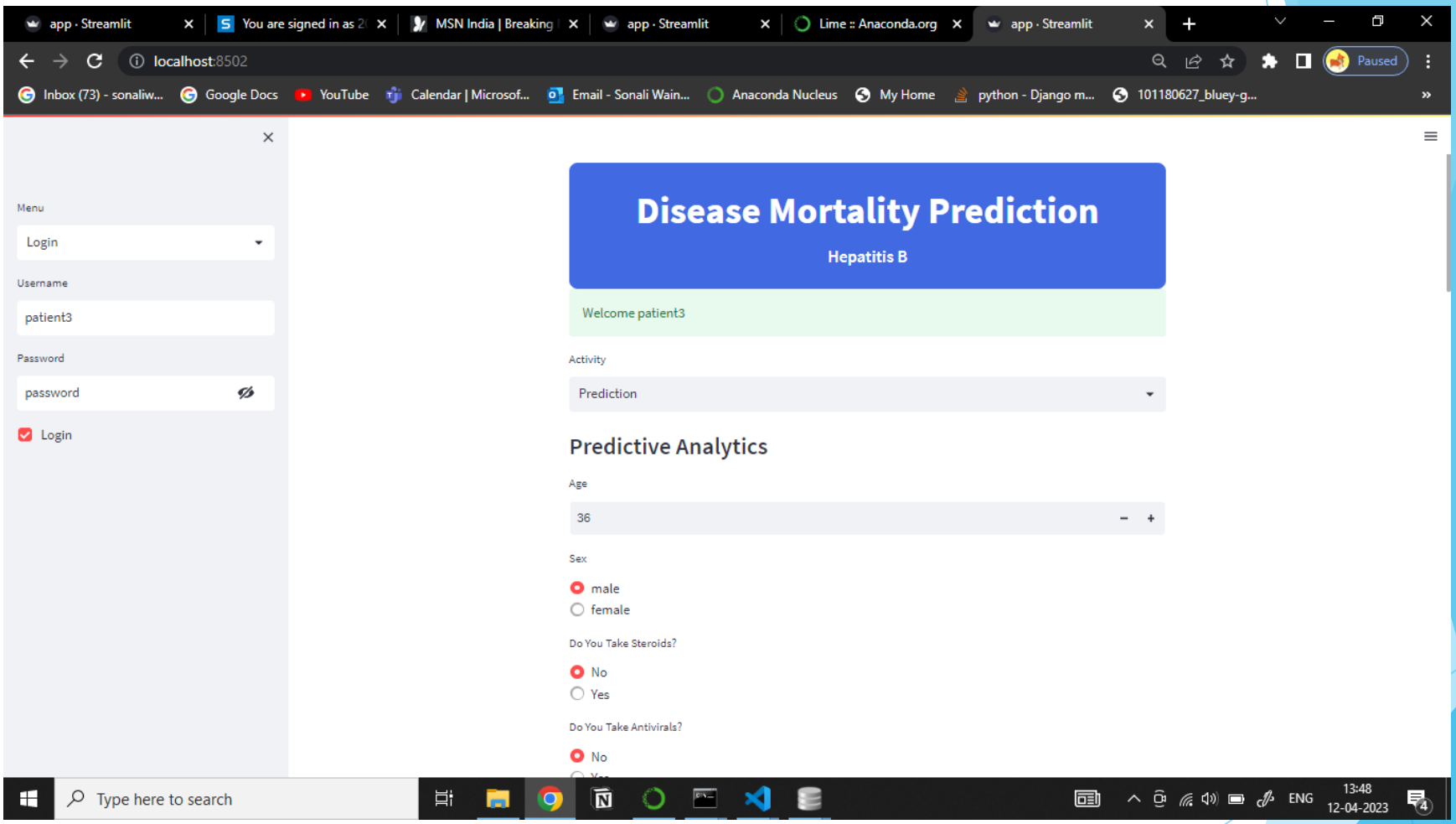
4.Algorithm: Logistic regression, KNN, Decision Tree

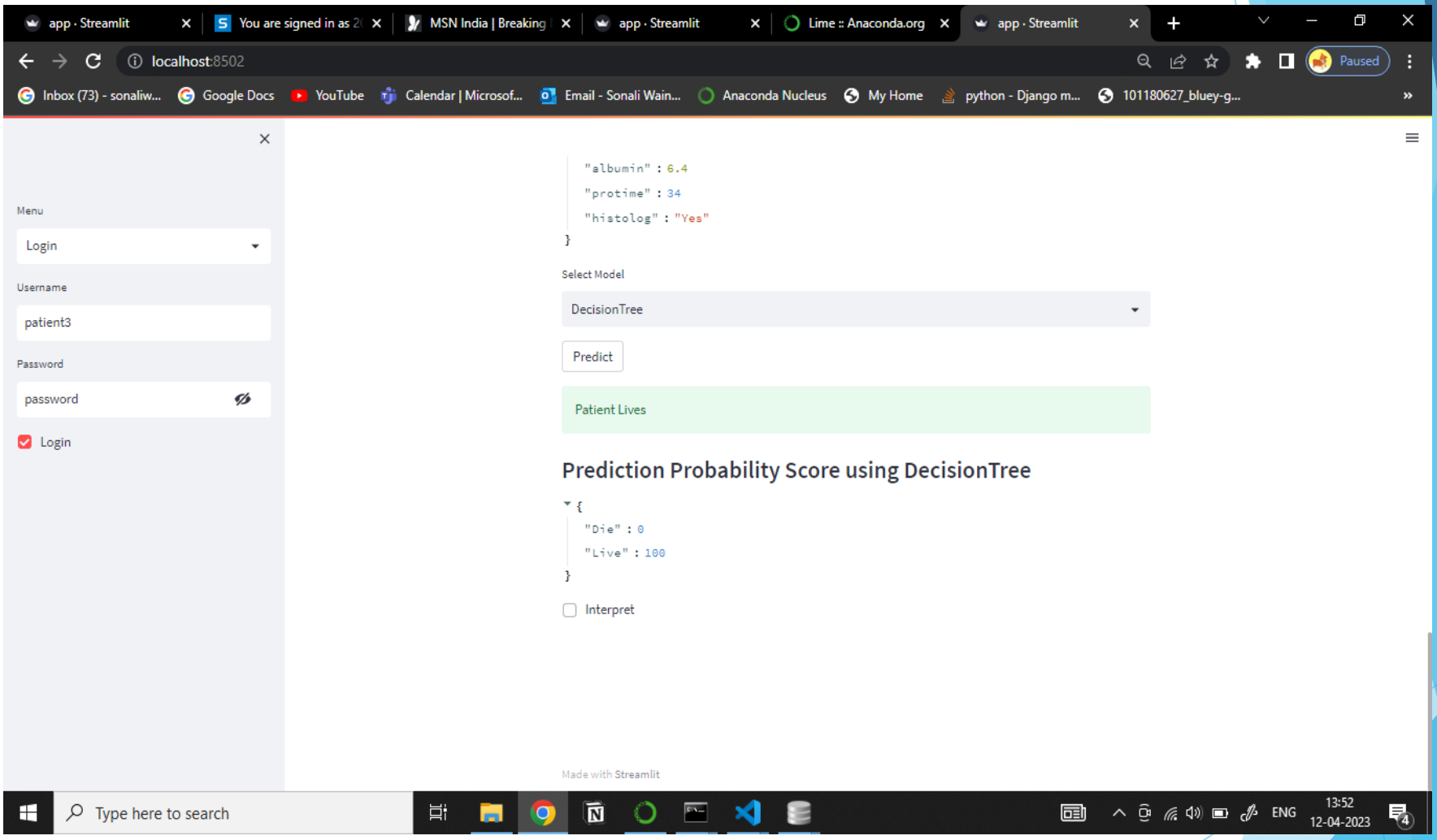
12. Suggestions in review 1

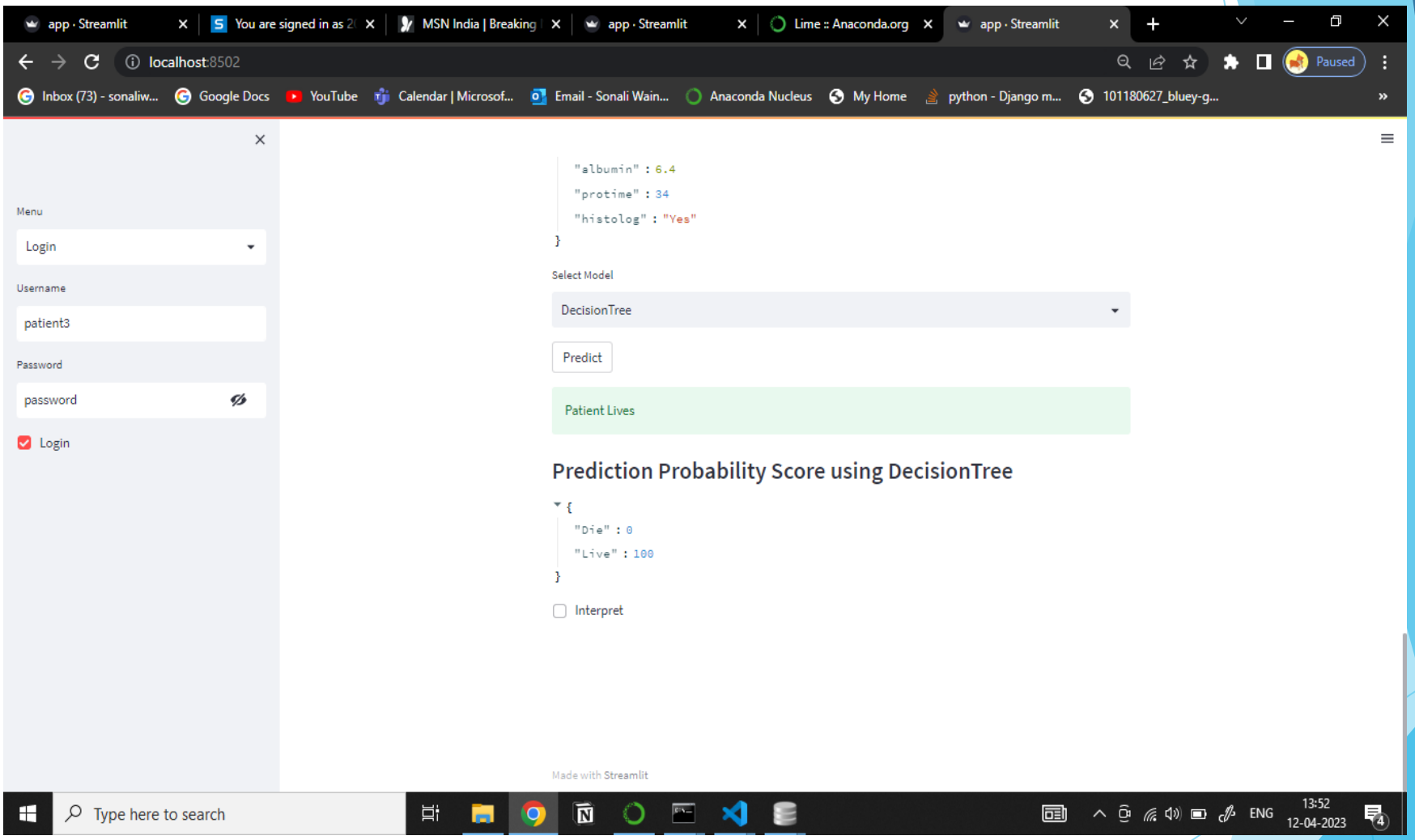
1.Suggestion1: Improvement in GUI

2.Suggestion2: Add validations

13. Result and Discussion




app · Streamlit
You are signed in as 2
MSN India | Breaking
app · Streamlit
Lime :: Anaconda.org
app · Streamlit
localhost:8502
Paused
Inbox (73) - sonaliw...
Google Docs
YouTube
Calendar | Microsof...
Email - Sonali Wain...
Anaconda Nucleus
My Home
python - Django m...
101180627_bluey-g...
Menu
Login
Username
patient3
Password
password
Login
"albumin" : 6.4
"protime" : 34
"histolog" : "Yes"
}
Select Model
DecisionTree
Predict
Patient Lives
Prediction Probability Score using DecisionTree
{
"Die" : 0
"Live" : 100
}
Interpret
Made with Streamlit
Type here to search
ENG
13:52
12-04-2023


app · Streamlit
You are signed in as 2
MSN India | Breaking
app · Streamlit
Lime :: Anaconda.org
app · Streamlit
localhost:8502
Paused
Inbox (73) - sonaliw...
Google Docs
YouTube
Calendar | Microsof...
Email - Sonali Wain...
Anaconda Nucleus
My Home
python - Django m...
101180627_bluey-g...
Menu
Login
Username
patient3
Password
password
Login
"albumin" : 6.4
"protime" : 34
"histolog" : "Yes"
}
Select Model
DecisionTree
Predict
Patient Lives
Prediction Probability Score using DecisionTree
{
"Die" : 0
"Live" : 100
}
Interpret
Made with Streamlit
Type here to search
ENG
13:52
12-04-2023

13. Conclusion and Future scope

Future Scope:

- For future scope, the project can be improved by incorporating more advanced machine learning and data pre-processing techniques, such as data normalization.

Conclusion:

The Hepatitis B mortality prediction is a significant application of machine learning and predictive analysis in the business world. Through this project, we can predict the mortality of patient using certain significant patient features.

Thank You.