## PROJECT REPORT FORMAT

#### 1.INTRODUCTION

- 1.1 Project Overview
- 1.2 Purpose

#### 2. LITERATURE SURVEY

- 2.1 Existing Problem
- 2.2 References
- 2.3 Problem Statement Definition:

#### 3.IDEATION &PROPOSED SOLUTION

- 3.1 Emphathy Map Canvas
- 3.2 Ideation & Brainstorming
- 3.3 ProposedSolution
- 3.4 Problem Solution Fit

## 4. REQUIREDMENT ANALYSIS

- **4.1 Functional Requirements**
- **4.2 Non-Functional Requirements**

## **5. PROJECT DESIGN**

**5.1 Data Flow Diagram** 

## **5.2 Solution and Technical Architecture**

#### 6. PROJECT PLANNING & SCHEDULING

- **6.1 Sprint Plannning & Estimation**
- **6.2 Reports from JIRA**

## 7. CODING AND SOLUTIONING

- **7.1 Feature 1**
- **7.2 Feature 2**

### 8. TESTING

- 8.1 Test Case
- **8.2 User Acceptance Testing**

#### 9.RESULTS

**9.1 Performance Metrics** 

## **10.ADVANTAGES AND DISADVANTAGES**

- 11. CONCLUSION
- 12. FEATURE ENHANCEMENTS
- 13.APPENDIX

#### 1.INTRODUCTION

## 1.1 Project Overview:

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. Diagnosis of liver infection at 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 chronic liver disease diagnosis by prediction. In this work, we used five algorithms Logistic Regression, Decision Tree, Support Vector Machine, Naive Bayes, and Random Forest. The performance of different classification techniques was evaluated on different measurement techniques such as accuracy, precision, recall, and specificity. Moreover, our present study mainly focused on the use of clinical data for liver disease prediction and explores different ways of representing such data through our analysis.

### 1.2 Purpose:

Early prediction of liver disease is very important to save human life and take proper steps to control the disease. Decision Tree algorithms have been successfully applied in various fields especially in medical science. This research work explores the early prediction of liver disease using various decision tree techniques. The liver disease dataset which is select for this study is consisting of attributes like total bilirubin, direct bilirubin, age, gender, total proteins, albumin and globulin ratio. The main purpose of this work is to calculate the performance of various decision tree techniques and compare their performance. The analysis proves that Decision Stump provides the highest accuracy than other techniques.

#### 2. LITERATURE SURVEY

### 2.1 Existing Problem:

Only two systems exist in the same domain, according to a thorough investigation into the subject. First, the system is entirely manual. It has the capacity to store patient information and medical records. The initial system's key characteristics are as follows. The second system is more effective than the first. It was discovered from a related research study that the system is constructed utilizing the KNN method. The Limitation are the entire system was manual, it fails to accurately predict a value using the KNN algorithm and This system takes long time to provide the user with an output.

#### 2.2 References:

A.K.M Sazzadur Rahman, F. M. Javed Mehedi Shamrat, Zarrin Tasnim, Joy Roy, Syed Akhter Hossain

ResearchGate - 2019

Chronic Liver Disease is the leading cause of global death that impacts the massive quantity of humans around the world. This disease diagnosis is very costly and complicated. Therefore, this paper evaluates the performance of different Machine Learning algorithms in order to reduce the high cost of chronic liver disease diagnosis by prediction. Six machine learning techniques have been applied including Logistic Regression, K Nearest Neighbors, Decision Tree, Support Vector Machine, Naïve Bayes, and Random Forest. The performance was evaluated on different measurement techniques such as accuracy, precision, recall, f-1 score, and specificity and the result was that LR achieved the highest accuracy.

Rayyan AzamKhan, Yigang Luo, Fang Xiang Wu

ScienceDirect – 2022

This paper mainly focuses on the computer-aided diagnosis of hepatic lesions in view of diffuse- and focal liver disorders. This is based on three image acquisition modalities: ultrasonography, computed tomography, and magnetic resonance imaging. Insightful analysis is presented for each preliminary step, particularly preprocessing, attribute analysis, and

classification techniques to accomplish clinical diagnostic tasks. In preprocessing denoising, deblurring, and segmentation methods are used. Denoising is mainly performed with nonlinear models.

#### Chokka Anuradha, D Swapna, Balamuralikrishnan Thati

#### IEEE

In the human body one of the most important organs is the liver. If the regular functionality of the liver is disturbed then this condition is called disease-affected liver. Therefore, an early stage of disease detection is more important which helps in disease prevention at starting stage with small medications. But, it is too difficult to identify Liver disease at the early stages because symptoms are very less at the starting stage. Lab results with physical examination are involved in the Traditional methods. This paper aims to represent a Diagnosing for Liver disease prediction in Patients using Combined Machine Learning Models. Optimized three machine learning algorithms are used for the accurate diagnosis of liver disease and they are Artificial Neural Networks (ANN), Decision Trees, and K-Nearest Neighbors (KNN). With the help of these algorithms, given data is classified and results are produced. The future data is predicted with the help of past and present data. The accuracy results are produced by comparing three classification algorithms.

#### Fahad Mostafa, Easin Hasan, Morgan Williamson, Hafiz Khan

#### MDPI - 2021

ML algorithms are trained to detect the possibility of liver disease to assist healthcare workers. Correlation of chosen variable with the risk of liver disease is performed to train the model. ML methods were able to identify the liver disease with high accuracy. The PCA results showed five important factors for liver disease diagnosis: AST, ALT, GGT, BIL, and ALP. In a real situation, a clinician can strongly suspect liver disease using only these five variables, as they are very descriptive for liver function. The ratio of ALT and AS denotes the cause of a liver injury. GGT and ALP increase in circulation with the severity of a liver injury. Additionally, the injury proximity to the bile duct is determined by the concentration of ALP. This study shows several machine learning approaches with PCA, which outperformed the classification. Among three ML classification methods, the performance of SVM and RF is better than ANN.

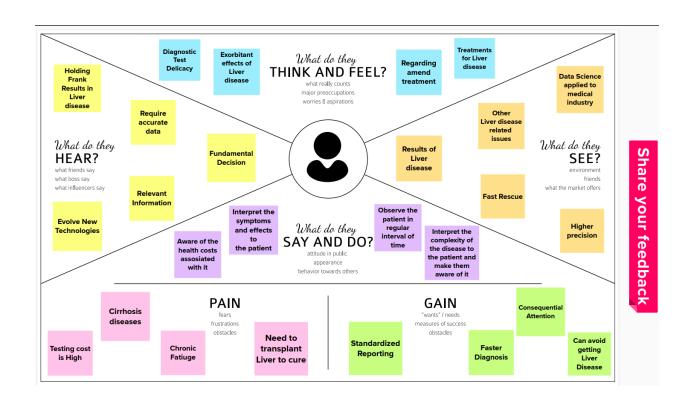
#### 2.3 Problem Statement Definition:

Given a dataset containing biological and diagnostic data of 583 Indian patients,

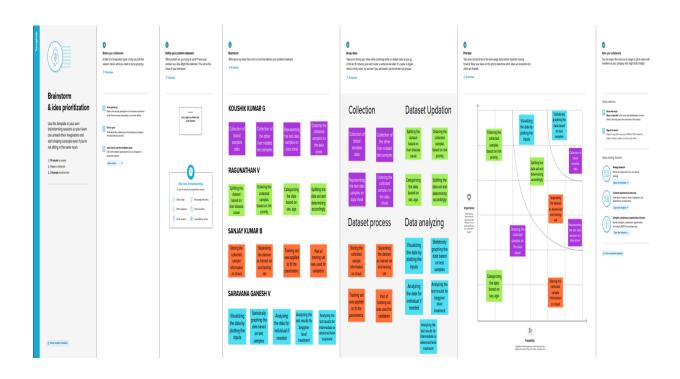
this project aims to identify a suitable machine learning algorithm which is capable of identifying whether a person has liver disease or not. This is a binary classification problem to be solved using supervised learning. We have ten features for each point and a label which identifies whether the patient is suffering from liver disease or not. In order to arrive at the solution, our aim should be train various supervised learning models on this dataset so that we have a well performing model which is able to classify any new data point as positive or negative with a reasonable degree of accuracy and perform better than the benchmarks.

#### 3.IDEATION & PROPOSED SOLUTION

## 3.1 Emphathy Map Canvas:



# 3.2 Ideation & Brainstorming:



# 3.3 ProposedSolution:

| S.No. | Parameter                                | Description   |  |  |
|-------|--|---|--|--|
| 1.    | Problem Statement (Problem to be solved) | To predict the liver disease of a patient inorder by building machine learning models with the help of past/historical datas of liver disease patients.   |  |  |
| 2.    | Idea / Solution description              | Our idea is to build an effective and efficient model to predict the liver disease of a patient in a starting stage and make them take required treatment and health instructions.  |  |  |
| 3.    | Novelty / Uniqueness                     | Our Model will give the precise result by analysing the patients blood samples etc,. Which will inform us before the disease is affected. We use new algorithms to make the decision precisely.   |  |  |
| 4.    | Social Impact / Customer Satisfaction    | The customer will get satisfied and it will create a greater impact among the social economy.  Because, it is going to play a major role in the medical industry. Also, it gives the exact result of the possibility for which a liver disease can occur. |  |  |
| 5.    | Business Model (Revenue Model)           | The model which is used is first trained and tested with lakhs of samples. Among different models we will use the efficient and effective model to predict the disease.   |  |  |
| 6.    | Scalability of the Solution              | The outcome of the solution will vary according to the machine learning model we are creating to predict liver disease. In this case we are providing a high precise solution providing a model.  |  |  |



## 4. REQUIREDMENT ANALYSIS

## 4.1 Functional Requirements:

Following are the functional requirements of the proposed solution.

| FR No. | Functional Requirement (Epic)           | Sub Requirement (Story / Sub-Task)  |
|--------|---|---|
| FR-1   | Patient with symptoms of liver disease  | Patient dataset such as Total Bilirubin, Direct Bilirubin, Total Proteins, Albumin etc.   |
| FR-2   | Predicting the disease using algorithms | Machine learning  |
| FR-3   | Pre-processing the Data set of patient. | MPCA  |
| FR-4   | Classification of algorithm             | KNN ,SVM, Navis bayes   |
| FR-5   | Building and training the system        | In this phase, we split the dataset into training and test dataset, and then trained the models using training dataset  |
| FR- 6  | Testing the model                       | In this phase, we tested the accuracy of the models with<br>the test dataset that was formed in previous phase and<br>the most accurate model is figured out. |

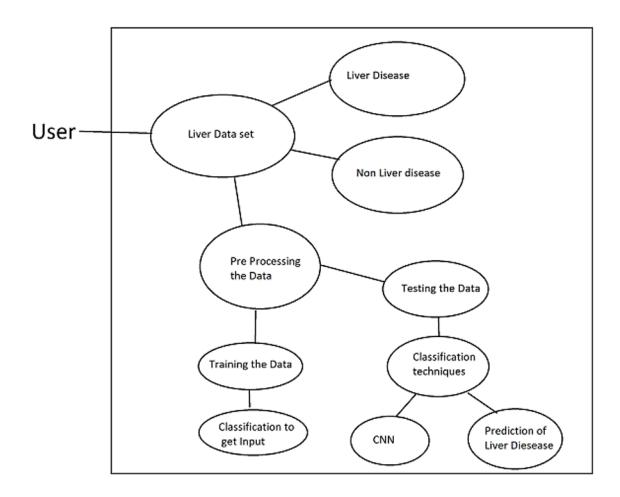
# **4.2 Non-Functional Requirements:**

| FR No. | Non-Functional Requirement | Description   |
|--------|----------------------------|---|
| NFR-1  | Usability                  | It is simpler and possible to predict liver disease at<br>an earlier stage. Because that it benefits all kinds of<br>people, it is a cost-effective option. |
| NFR-2  | Security                   | Early diagnosis of liver illness allows patients to receive treatment before the disease progresses and saves lives.  |
| NFR-3  | Reliability                | This approach offers excellent performance and scalability, making it more dependable.  |
| NFR-4  | Performance                | It provides accuracy of over 90%. Thus, it has a high performance rate.   |
| NFR-5  | Availability               | By having few basic data set of people we can predict the disease.  |
| NFR-6  | Scalability                | It has more efficiency in detecting liver disease prediction than any other models.   |

## **5. PROJECT DESIGN**

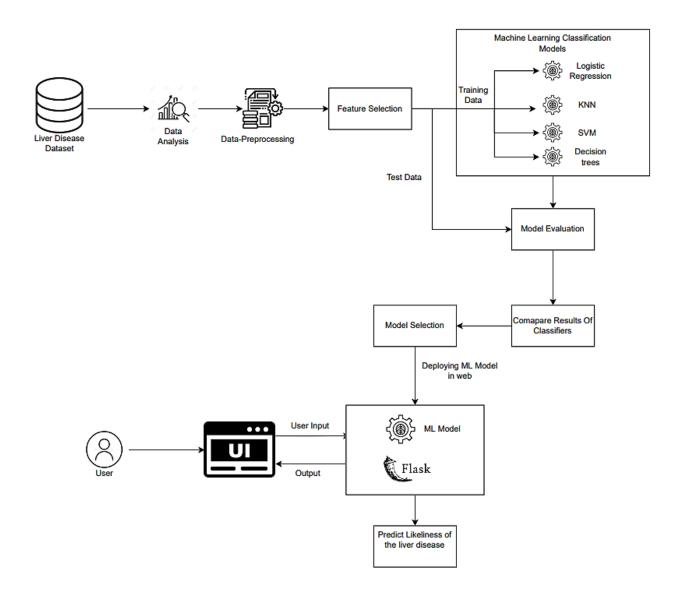
# **5.1 Data Flow Diagram:**

A Data Flow Diagram (DFD) is a traditional visual representation of the information flows within a system. A neat and clear DFD can depict the right amount of the system requirement graphically. It shows how data enters and leaves the system, what changes the information, and where data is stored.



## **5.2 Solution and Technical Architecture:**

## **SOLUTION ARCHITECTURE**



## 6.PROJECT PLANNING &SCHEDULING

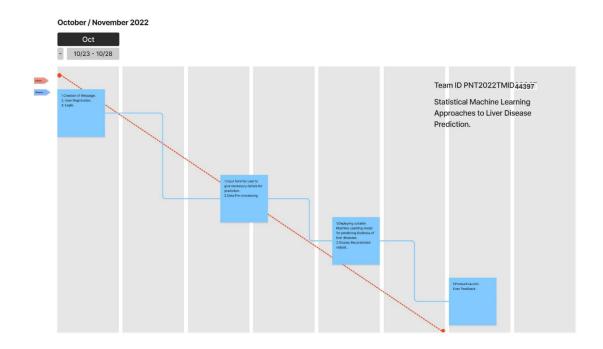
## **6.1 Sprint Plannning & Estimation:**

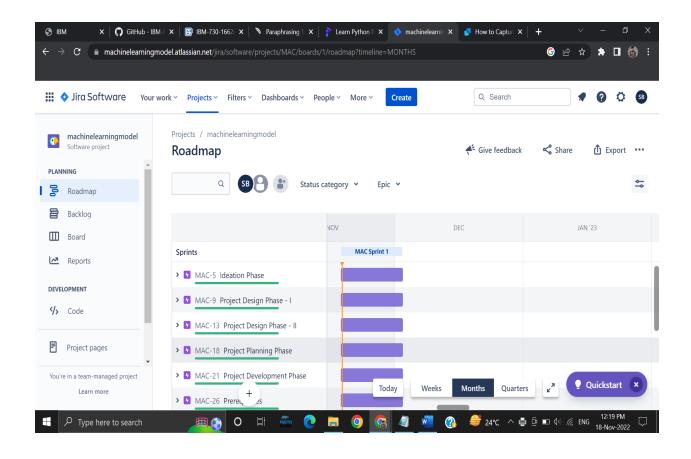
## **Sprint 1,2,3 and 4:**

- \* In the first sprint, we used a Jupyter notebook to develop a machine learning model, which we then saved in a pickle file. The dataset for patients with liver disease is included in this. This uses a variety of techniques to build the model.
- \* We developed a home page and a prediction page for Sprint 2. where users may enter their medical information, such as albumin, proteins, globulin, age, and gender, and the information is utilised to determine whether or not the user has liver disease.
- \* The prediction that users input on the prediction page is displayed on the Result page that we created for Sprint 3.
- \* Using the model that is stored in the pickle file, we created the python flask file in Sprint 4 to predict the model.

## **6.2 Reports from JIRA:**

## **Sprint Burndown**





## 7. CODING AND SOLUTIONING

#### **7.1 Feature 1:**

In general LD(Liver Disease) regular testing takes longer time for the result so, the treatement gets late. In our project the model predicts the disease in earlier.

## **7.2 Feature 2:**

As the model is deployed in the IBM Cloud wide range of peoples and hospitals get benefited.

## **8.TESTING**

## 8.1 Test Case

| Test Scenario                         | Expected Result   |
|---------------------------------------|---|
| Verify the UI elements                | UI works fine   |
| Verify the button elements working    | Home page loads   |
| Verify the button elements working    | Predict page loads  |
| Verify the UI elements                | Shows the description about the project and the details of team |
|                                       | members   |
| Verify the predict button working     | Predict page loads  |
| Verify the UI elements                | Show the input boxes and the user to                            |
| verify the of elements                | enter their input   |
| Verify the working of prediction page | Predict page works fine   |
| Verify the UI elements                | Result page works fine  |

## **8.2 User Acceptance Testing**

## 1. **Purpose of Document**

The purpose of this document is to briefly explain the test coverage and open issues of the Liver Disease Prediction project at the time of the release to User Acceptance Testing (UAT).

## 2. **Defect Analysis**

This report showsthe number of resolved or closed bugs at each severity level, and how they were resolved

| Resolution     | Severity 1    | Severity 2 | Severity 3 | Severity 4 | Subtotal |  |
|----------------|---------------|------------|------------|------------|----------|--|
| By Design      | 7             | 0          | 0          | 0          | 7        |  |
| Duplicate      | 0             | 0          | 3          | 0          | 3        |  |
| External       | 0             | 3          | 0          | 1          | 4        |  |
| Fixed          | 2             | 4          | 2          | 4          | 12       |  |
| Not Reproduced | 0             | 0          | 1          | 0          | 1        |  |
| Skipped        | 0             | 0          | 1          | 1          | 2        |  |
| Won't Fix      | t Fix 0 5 0 0 |            | 5          |            |          |  |
| Totals         | 9             | 12         | 7          | 6          | 34       |  |

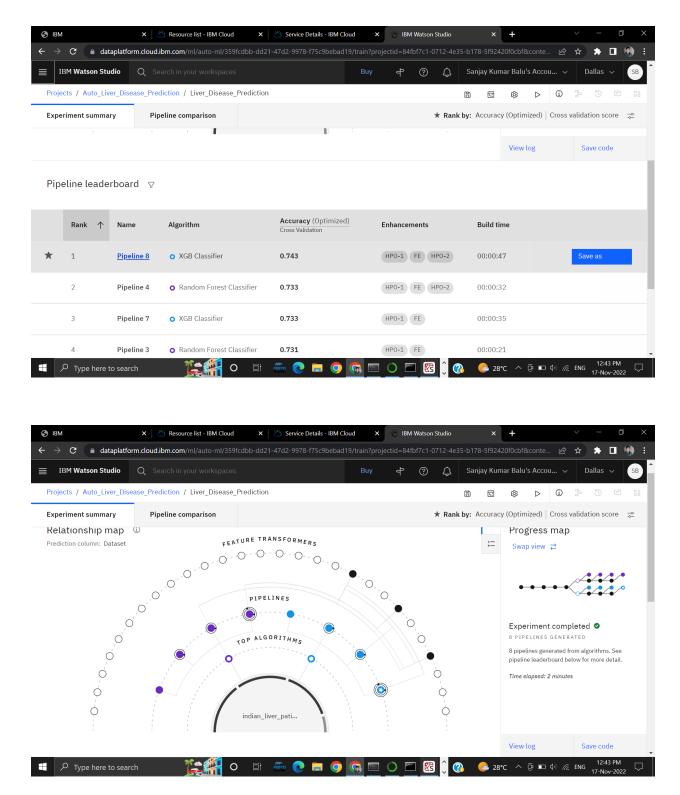
# 3.Test case Analysis

This report shows the number of test cases that have passed, failed, and untested

| Section             | Total Cases | Not Tested | Fail | Pass |
|---------------------|-------------|------------|------|------|
| Print Engine        | 7           | 0          | 0    | 7    |
| Client Application  | 51          | 0          | 0    | 51   |
| Security            | 2           | 0          | 0    | 2    |
| Outsource Shipping  | 3           | 0          | 0    | 3    |
| Exception Reporting | 9           | 0          | 0    | 9    |
| Final Report Output | 4           | 0          | 0    | 4    |
| Version Control     | 2           | 0          | 0    | 2    |

## 9.RESULTS

## **9.1 Performance Metrics**



10.ADVANTAGES AND DISADVANTAGES

#### **ADVANTAGES**

- Costs for initial setup and upkeep are decreased.
- Enhances medical services
- Minimises the need for medical oversight
- Infrastructure, platform, and software are not issues
- Proactive planning for future trends in population health
- By analysing and assessing scores and analysing ongoing transactions, preventing ransom and other hacks
- Accelerating administrative tasks including insurance claim submission and discharge procedures
- Simplicity of maintaining electronic health records

#### **DISADVANTAGES**

- Internet cnnection is required
- User interface development is a little more difficult.
- additional procedures to keep in mind when creating various services

#### 11. CONCLUSION

This study developed an algorithm for predicting Liver Disease at earlier stage. The Dataset contains input parameters obtained from liver disease patients and the models are trained and validated using the valid parameters. To analyze the liver disease decision tree, support vector machine learning, random forest and etc are built.

#### 12. FEATURE ENHANCEMENTS

#### SMS/Email Module

In the proposed system, admin assigns ID and Password for doctors and receptionist and is intimated manually so we can add sms/email module as a feature enhancement where doctors and receptionist receive an sms or email regarding the ID and Password.

#### **QUERY MODULE**

We can add the query module as the feature enhancement to the application where doctors, receptionist and admin of the application can interact with each other.

#### 13.APPENDIX

#### **SOURCE CODE**

## webApp.py

#### home.html

```
essential in the prediction of liver disease
where examined and the data set of liver patients also evaluated. This paper compares various
classification algorithms such as random forest,
this study, show with the highest accuracy
outperformed the other algorithms and can be further utilised in the prediction of liver disease
recommended. (p)
(/div)
(/hody)
(/html)
```

#### home.css

```
box sizing: border-box;
}
boxy {
    margin: 0px;
    padding: 0px;
    padding: 0px;
    padding: 0px;
    padding: 0px;
    padding: 0px;
    font 'family: 'Raleway', sans-serif;
}

maxin {
    mini {
        boxeround.color: beige;
        buight: 50wh;
        position: relative;
}

nav {
        issplay: flee;
        justify-content: space-around;
        align-itens: center;
        top: 0;
        buight: 100%;
        buig
```

```
.menu li a {
    height: 48px;
    line height: 48px;
    padding: 90x 22px;
    display: flex;
    display: flex;
    destrowment upercase;
    font-weight: 900;
    letter-spacing: lps;
    color: black;
}

.content {
    Descground-color: 807fc5;
    display: flex;
    display: flex;
    idisplay: flex;
    idisplay: flex;
    idisplay: flex;
    postion: absolute;
    left: 58%;
    juitity-content space-around;
    postion: absolute;
    left: 58%;
    right: 90%;
    idisplay: flexible f
```

```
transition: all ease 0.2s;
}
```

## prediction.html

```
chtml>
chtmad>

ctitie>Liver Patient Analysis:/title>
cmeta name="viesport" content="withhedwice=width, initial-scale=1",
clink real="style=heet" her* "https://conis.cloudfianc.com/ana/libs/font_amescom/a,7.0/css/font_amescom.Bin.csa">
clink real="style=heet" https://fontamescom/as/libs/font_amescom/a,7.0/css/font_amescom.Bin.csa">
clink real="style=heet" here* "https://fontamestale.com" crossorigin>
clink real="style=heet" here* "fittos://fontamestale.com" crossorigin>
clink real="style=heet" here* "(un_fon("static ',filenames 'css/prediction.css')))">
clink real="style=heet" here* "(un_fon("static ',filenames 'css/prediction.css')))">
cheed*
cheedy:
cheed to come action='/logh' method='post">
cloud come action='/logh' meth
```

# prediction.css

```
body (
    background-color: rgbs(98, 245, 281, 0.753);
    font-family: "Raleway", same-serif;
}

11 {
    text-align: center;
    color: rgbs(255, 11, 11, 0.815);
} clobe tr td p {
    padding-sert: 405;
} .box {
    border-readus: 28pc;
    padding-sert: 28pc;
    sert: 28pc
```

```
outline: none;
color: white;
color: white;
transition: 0.25s;
}
.box isput[type='text]:focus;
box isput[type='text]:focus;
box isput[type='text]:focus;
border-color: #2ecc71;
}
.box isput[types tox]: focus;
border-color: #2ecc71;
}
.box isput[types tox]: focus;
text-align: center;
border-2x xoild #2ecc1;
padding: #4px #4px;
color: white;
border-radius: #2ex;
transition: 0.25s;
cursor: pointer;
}
.box p(
color: whitesmoke;
}
```

#### result.html

```
.asobar albover, .dropdom:hover.dropbtn {
    background-color: reg(5, 15, 298);
    background-color: reg(5, 15, 298);
}

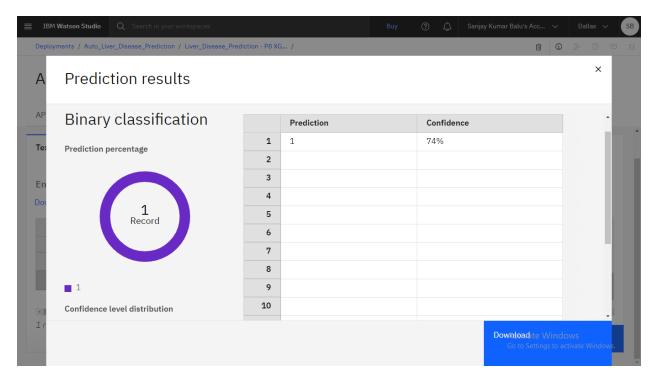
dropdom.content {
    display: come;
    postion: absolute;
    background-color: #89999;
    man-addit: 1800x;
    donodom-content a {
        finit: nome;
        color: black;
        padding: 120x 180x;
        display: block;
        text-align: left;
    }

    .dropdom-content albover {
        background-color: medd;
    }

    .dropdom-insort, .dropdom-content {
        display: block;
        insplay: insince block;
        insplay: insince block;
        insplay: insince block;
        box: sizing: border-box;
    }

        registerbor {
        background-color: medAFS0;
    }
```

```
culs clibranuseac/lib clibranuseac/libranuseac/libranuseac/libranuseac/libranuseac/libranuseac/libranuseac/libranuseac/libranuseac/libranuseac/libranuseac/libranuseac/libranuseac/libranuseac/libranuseac/libranuseac/libranuseac/libranuseac/libranuseac/libranuseac/libranuseac/libranuseac/libranuseac/libranuseac/libranuseac/libranuseac/libranuseac/libranuseac/libranuseac/libranuseac/libranuseac/libranuseac/libranuseac/libranuseac/libranuseac/libranuseac/libranuseac/libranuseac/libranuseac/libranuseac/libranuseac/libranuseac/libranuseac/libranuseac/libranuseac/libranuseac/libranuseac/libranuseac/libranuseac/libranuseac/libranuseac/libranuseac/libranuseac/libranuseac/libranuseac/libranuseac/libranuseac/libranuseac/libranuseac/libranuseac/libranuseac/libranuseac/libranuseac/libranuseac/libranuseac/libranuseac/libranuseac/libranuseac/libranuseac/libranuseac/libranuseac/libranuseac/libranuseac/libranuseac/libranuseac/libranuseac/libranuseac/libranuseac/libranuseac/libranuseac/libranuseac/libranuseac/libranuseac/libranuseac/libranuseac/libranuseac/libranuseac/libranuseac/libranuseac/libranuseac/libranuseac/libranuseac/libranuseac/libranuseac/libranuseac/libranuseac/libranuseac/libranuseac/libranuseac/libranuseac/libranuseac/libranuseac/libranuseac/libranuseac/libranuseac/libranuseac/libranuseac/libranuseac/libranuseac/libranuseac/libranuseac/libranuseac/libranuseac/libranuseac/libranuseac/libranuseac/libranuseac/libranuseac/libranuseac/libranuseac/libranuseac/libranuseac/libranuseac/libranuseac/libranuseac/libranuseac/libranuseac/libranuseac/libranuseac/libranuseac/libranuseac/libranuseac/libranuseac/libranuseac/libranuseac/libranuseac/libranuseac/libranuseac/libranuseac/libranuseac/libranuseac/libranuseac/libranuseac/libranuseac/libranuseac/libranuseac/libranuseac/libranuseac/libranuseac/libranuseac/libranuseac/libranuseac/libranuseac/libranuseac/libra
```



#### VIDEO LINK:

https://drive.google.com/file/d/174tVtKUJc8dth7CJSgAsY-Capo5GeiMd/view?usp=drivesdk