A

PROJECT REPORT

ON

"Evaluation based Approaches for Liver Disease Prediction using Machine Learning Algorithms"

Submitted for the degree of

Bachelors of Computer Engineering

BY

Patil Devesh Pandurang	ROLL NO. 41
Bangar Ajay Shivaji	ROLL NO. 08
Gawali Subodh Manik	ROLL NO. 23

UNDER THE GUIDANCE OF

PROF. KANCHAN UMAVANE

DEPARTMENT OF COMPUTER ENGINEERING

SHIVAJIRAO S. JONDHLE COLLEGE OF ENGINEERING AND TECHNOLOGY,

ASANGAON.

UNDER UNIVERSITY OF MUMBAI

2023-24

DECLARATION

This is to certify that the project report entitled **"EVALUATION BASED APPROACHES FOR LIVER DISEASE PREDICTION USING MACHINE LEARNING ALOGRITHMS"** which is submitted by the project members in partial fulfilment of the requirement for the award of degree B.E in computer engineering to Mumbai University, Mumbai comprises only the original work and due acknowledgement has been made in the text to all other material used.

Date:

Name of the Students:

DEVESH PANDURANG PATIL

BANGAR AJAY SHIVAJI

GAWALI SUBODH MANIK

PROJECT CO-ORDINATOR

PROF.VISHAL SHINDE

CERTIFICATE

This is to certify that Project Report entitled **"EVALUATION BASED APPROACHES FOR LIVER DISEASE PREDICTION USING MACHINE LEARNING ALOGRITHMS"** which is submitted by project members in partial fulfilment of the requirement for the award of degree B.E in computer engineering to Mumbai University is a record of the candidates own work carried out by him/her under my supervision. The matter embodied in this thesis is original and has not been submitted for the award of any other degree.

Date:

Name of the Students:

DEVESH PANDURANG PATIL

AJAY SHIVAJI BANGAR

SUBODH MANIK GAWALI

Name of the Guide: PROF. KANCHAN UMAVANE

Sign and Date:

PROF. KANCHAN UMAVANE
(Internal Examiner)

PROF.
(External Examiner)

PROF. VISHAL SHINDE
HOD
Department of Computer Engg.

DR.MRS.GEETHA K. JAYARAJ
Principal,
SSJCET, Asangaon

ACKNOWLEDGEMENT

This is great pleasure and immense satisfaction to express our deepest sense of gratitude and thanks to everyone who has helped us in completing our work successfully. We are presenting this Project report on **"EVALUATION BASED APPROACHES FOR LIVER DISEASE PREDICTION USING MACHINE LEARNING ALGORITHMS"** as part of the curriculum of B.E. Computer Engineering. Inspiration and guidance are invaluable in every aspect of life especially in the field of academics, which we have received from our respected **Project Guide:** Prof. Kanchan Umavane and **Head of Computer Department:** Prof. Vishal R. Shinde. Besides, we take this opportunity to express our sincere gratitude to the **Principal:** Dr. Mrs. Geetha K. Jayaraj, SSJCET for providing a good environment and facilities to complete this project. We would also like to thank all our staff and friends who have directly or indirectly guided and helped us in the preparation of this report and also for giving us an unending support right from the stage this idea was conceived.

DEVESH PANDURANG PATIL
AJAY SHIVAJI BANGAR
SUBODH MANIK GAWALI

LIST OF PUBLICATIONS

ABSTRACT

The life of humans living without liver tumors is one of the fundamental care of human livelihood. Therefore, for better care, detection of liver disease at a primitive phase is necessary. For medical experts, predicting the illness in the early stages due to subtle signsis a very difficult task. Many, when it is too late, the signs become evident. The current work aims to augment the perceive nature of liver disease by means of machine learning methods to solve this epidemic. The key purpose of the present work focused on algorithmsfor classification of healthy people from liver datasets. Centered on their success variables,this research also aims to compare the classification algorithms and to provide predictionaccuracy results.[1]

Key Word- Support vector machines, Machine learning algorithms, Liver diseases, Prediction algorithms, Classification algorithms, Task analysis, Tumors.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABBREVATIONS

Abbrevation	Description
SVM	Support Vector Machines
LR	Logistic Regression
RF	Random Forest
DFD	Data Flow Diagram
UML	Unified Modelling Language
ML	Machine Learning

INTRODUCTION

AND

MOTIVATION

1. INTRODUCTION & MOTIVATION

1.1 INTRODUCTION

The scale of patient medical records increases day by day in the health care sector. Data mining is the method of using a computer-based information system (CBIS), using modern tactics, to uncover insights from data. Ensemble learning is a well-known method used for prediction by integrating multiple ensemble models of machine learning[5]. Aggregations of various classifiers are J48, C4.5 and Naive Bayes, etc. . Ensembles search for better outcomes than all of the simple classifiers[4]. Here is a high probability of liver failure among Indians. India is expected to become the World Capital of Liver Diseases by 2025. Because of the deskbound lifestyle, increased alcohol intake and smoking, the pervasive prevalence of infection inside liver in India is contributing around 100 forms of liver infections are present. It would also be of great value in the medical field to build a computer that will increase the diagnosis of the disease. systems that can assist physicians in making correct treatment choices, and the patient queue will also be minimized by liver specialists such as endocrinologists assisted by Automated categorization Methods for Disorders in Liver part. In medical diagnosis and disease prediction, classification techniques are widely common.[1].

1.2 AIM & OBJECTIVE

AIM

The aim of evaluation-based approaches for liver disease prediction using machine learning algorithms is to assess and improve the performance and reliability of machine learning models in predicting liver diseases. These approaches are crucial for developing accurate, clinically relevant, and practical tools for diagnosing liver conditions. Here are some specific goals and objectives of such approaches:

OBJECTIVE

- Develop machine learning models that can accurately identify the presence of liver disease at an early stage, allowing for timely medical intervention and treatment.
- Create predictive models that can assess the risk of liver disease for individuals based on their medical history, demographics, and other relevant factors, aiding in proactive healthcare management.
- Educate patients about their liver disease risk factors and potential interventions, promoting healthier lifestyles and increased awareness.
- Consider the broader public health impact of accurate liver disease prediction, including potential reductions in disease prevalence and improved population health outcomes.

1.3 LIMITATION ON EXISTING SYSTEM

The scale of patient medical records increases day by day in the health care sector.Major issues deliberated on patients with liver disease are not readily detected at starting phase since that can usually operate even though it is partly impaired. An early detection of liver disorders will improve the survival rate of the patient. There is a high probability of liver failure among Indians.it is very difficult to detect in early stages of the disease with high accuracy recovery of the disease.[1]

DISADVANTAGES OF EXISTING SYSTEM:

- we cannot predict proper accuracy results.
- we can not predict liver disease in early stages.

Algorithm: KNN ,Random forest

1.5 PROPOSED SYSTEM

The proposed work aims to enhance the predictive and classification quality of healthcare data by developing a hybrid predictive classifier model using the classifier ensemble. This project can help doctors make correct treatment choices, and the patient queue will also be minimized by liver specialists such as endocrinologists assisted by Automated categorization Methods for Disorders in Liver part. In medical diagnosis and disease prediction, classification techniques are widely common. Michael J Sorich described on chemical datasets, the classification (SVM)and logistic regression provides better prediction results.[1]

PROJECT OVERVIEW

2. PROJECT OVERVIEW

2.1 LITREATURE SURVEY

Paper 1

Paper Name: Liver Cancer Prediction in a Viral Hepatitis Cohort: A Deep Learning Approach

Authors Name : Dinh-Van Phan.Chien-Lung Chan,Ai-Hsien Adams Li

Explanation :

The main cause of liver illnesses, including liver cancer, which is the main cause of cancer-related mortality, is viral hepatitis. Unfortunately, therapy for this disease is typically challenging or downright impractical because it is discovered in its latter stages. Using a cohort of patients with hepatitis, this study used (DL) deep learning models to identify liver cancer early. In order to examine viral hepatitis patients from 2002 to 2010, it is polled one million randomly selected samples from the National Health Insurance Research Database (NHIRD). Then, based on the hepatitis cohort's medical history, this employed DL models to forecast instances of liver diseased In a cohort of people with hepatitis, the CNN model can accurately predict the development of liver cancer.[2]

Paper 2

Paper Name : A hybrid intelligent system for medical data classification

Authors Name : Manjeevan Seera

Explanation :

In this paper, a hybrid intelligent system that consists of the Fuzzy Min-Max neural network, the Classification and Regression Tree, and the Random Forest model is proposed, and its efficacy as a decision support tool for medical data classification is examined. The hybrid intelligent system aims to exploit the advantages of the constituent models and, at the same time, alleviate their limitations. It is able to learn incrementally from data samples (owing to Fuzzy Min-Max neural network), explain its predicted outputs (owing to the Classification and Regression Tree),and achieve high classification performances (owing to Random Forest). To evaluate the effectiveness of the hybrid intelligent system, three benchmark medical data sets, viz., Breast Cancer Wisconsin, Pima Indians Diabetes, and Liver Disorders from the UCI Repository of Machine Learning, are used for evaluation. A number of useful performance metrics in medicalapplications which include accuracy, sensitivity, specificity, as well as the area under the Receiver Operating Characteristic curve are computed. The results are analyzed and compared with those from other methods published in the literature. The experimental outcomes positivelydemonstrate that the hybrid intelligent system is effective in undertaking medical data classification tasks. More importantly, the hybrid intelligent system not only is able to produce good results but also to elucidate its knowledge base with a decision tree. As a result, domain users (i.e., medical practitioners) are able to comprehend the prediction given by the hybrid intelligent system; hence accepting its role as a useful medical decision support tool.[4]

Paper 3

Paper Name : An intelligent model for liver disease diagnosis

Authors Name : Rong-HoLin

Explanation :

Liver disease, the most common disease in Taiwan, is not easily discovered in its initial stage; early diagnosis of this leading cause of mortality is therefore highly important. The design of aneffective diagnosis model is therefore an important issue in liver disease treatment. This study accordingly employs classification and regression tree (CART) and case-based reasoning (CBR)techniques to structure an intelligent diagnosis model aiming to provide a comprehensive analytic framework to raise the accuracy of liver disease diagnosis .Methods: Based on theadvice and assistance of doctors and medical specialists of liver conditions, 510 outpatient visitors using ICD-9 (International Classification of Diseases, 9th Revision)codes at a medical centerin Taiwan from 2005 to 2006 were selected as the cases in the data set for liver disease diagnosis. Data on 340 patients was utilized for the development of the model and on 170 patients utilized to perform comparative analysis of the models. This paper accordinglysuggests an intelligent model for the diagnosis of liver diseases which integrates CART andCBR. The major steps in applying the model include: adopting CART to diagnose whether a patient suffers from liver disease; for patients diagnosed with liver disease in the first step, employing CBR to diagnose the types of liver diseases. Results: In the first phase, CART isused to extract rules from health examination data to show whether the patient suffers fromliver disease. The results indicate that the CART rate of accuracy is 92.94%.In the second phase, CBR is developed to diagnose the type of liver disease, and the new case triggers theCBR system to retrieve the most similar case from the case base in order to support the treatment of liver disease. The new case is supported by a similarity ratio, and the CBR diagnostic accuracy rate is 90.00%. Actual implementation shows that the intelligent diagnosis model is capable of integrating CART and CBR techniques to examine liver diseases with considerable accuracy. The model can be used as a supporting system in making decisionsregarding liver disease diagnosis and treatment. The rules extracted from CART are helpful to physicians in diagnosing liver diseases.[7]

2.1.1 COMPARATIVE ANALYSIS

Table .2.1.1 Comparative Analysis of Existing System

Paper Title	Author	Year	Publication	Description
Evaluation based Approaches for Liver Disease Prediction Machine Learning Algorithms	C.Geetha Dr.AR.Arunachalam	**2021**	IEEE	Model for predicting liver illness using logical regression and a support vector machine
Liver Cancer Prediction in a Viral Hepatitis Cohort: A Deep Learning Approach	Dinh-Van Phan.Chien-Lung Chan,Ai-Hsien Adams Li , Ting-Ying Chien,Van-Chuc Nguyen	**2020**	IJC	The results suggested that the Convolution Neural Networks model may accurately identify liver cancer.
Hybrid intelligent system for medical data classification	Manjeevan, Seera	**2014**	International Journal	Tasks involving the categorization of medical data are successfully completed by the hybrid system.
An intelligent model for liver disease diagnosis	Rong- HoLin	**2009**	PubMed publication	When making judgments about the diagnosis and treatment of liver illness, the model can be a useful tool.

2.1.2 MATHEMATICAL MODEL

Logistic Regression

Logistic regression is an algorithm that is used in solving classification problems. It is a predictive analysis that describes data and explains the relationship between variables. Logistic regression is applied to an input variable (X) where the output variable (y) is a discrete value which ranges between 1 (yes) and 0 (no).

It uses logistic (sigmoid) function to find the relationship between variables. The sigmoid function is an S-shaped curve that can take any real-valued number and map it to a value between0 and 1, but never exactly at those limits.

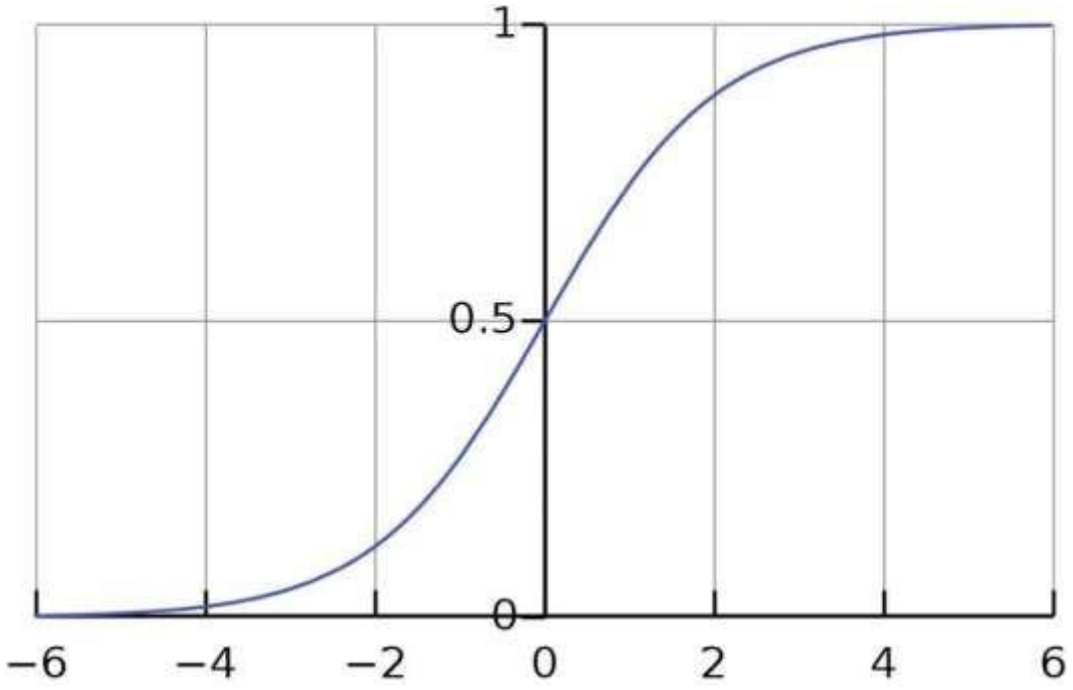


Fig 2.1.1 Illustration of a Sigmoidcurve

Support Vector Machine

The main objective of SVM is to find the optimal hyperplane which linearly separates the data points in two component by maximizing the margin . doted line is hyperplane, separating blue and pink classes balls.

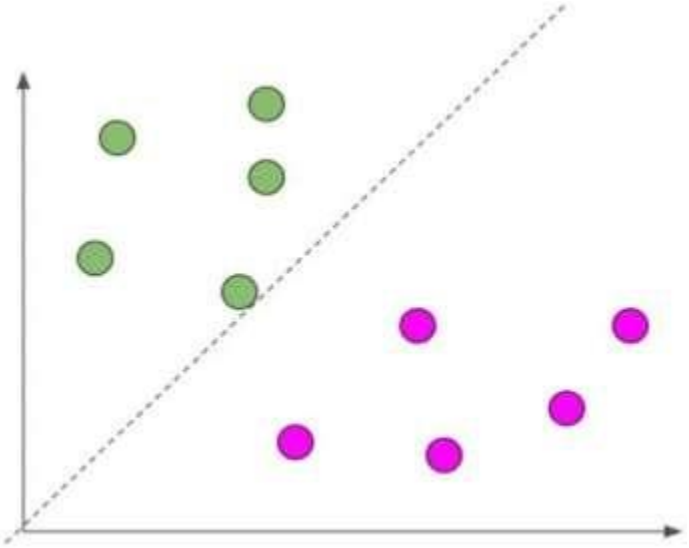


Fig.2.1.2 Support Vector Machine

Basic Linear Algebra

Vectors

Vectors are mathematical quantity which has both magnitude and direction. A point in the 2D plane can be represented as a vector between origin and the point.

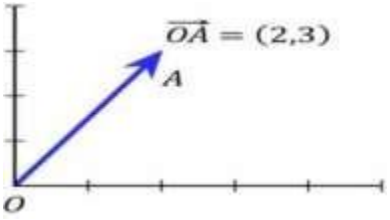


Fig 2.1.3 $\overrightarrow{OA}$ is a vector & length between O & A is its magnitude

Length of Vectors

Length of vectors are also called as norms. It tells how far vectors are from the origin.

Length of vector $x(x_1, x_2, x_3)$ is calculated as : $||x|| + \sqrt{(x^2 + x_2^2 + x_3^2)}$

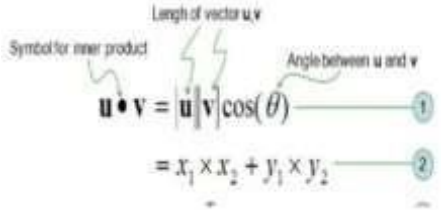
Direction of vector

Length of vector $x(x_1, x_2, x_3)$ is calculated as : $\left\{ \frac{x_1}{\|x\|}, \frac{x_2}{\|x\|}, \frac{x_3}{\|x\|} \right\}$

Dot Product

Dot product between two vectors is a scalar quantity . It tells how to vectors are related.

Two vect u and v are calculated as:



Hyper-plane

It is plane that linearly divide the n-dimensional data points in two component. In case of 2D, hyperplane is line, in case of 3D it is plane.It is also called as *n-dimensional line. Fig.3* shows, a blue line(hyperplane) linearly separates the data point in two components.

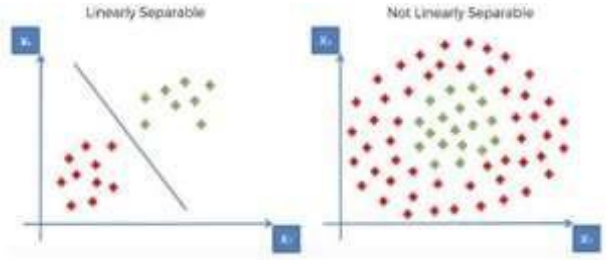


***Fig.2.1.4, hyperplane* is line divides data point into two classes(red & green)**

These *written*

*a*x+b-y=0Let vectors X=(x,y) and W=(a,-1)then on vectors form hyperplane is :W.X+b+0*

2.2 PROBLEM STATEMENT

The primary objective of this project is to develop a machine learning model that can predict the presence or risk of liver disease in individuals based on relevant clinical and demographic factors using SVM and Logical regression Algorithm. The model should achieve high accuracy, precision, recall, and be capable of generalizing to diverse patient populations.

Liver disease is a significant global health concern, affecting millions of individuals worldwide. Early detection and accurate diagnosis of liver diseases are crucial for effective treatment and improving patient outcomes. Machine learning offers the potential to assist healthcare professionals in identifying individuals at risk of liver disease and providing timely interventions.[1]

2.3 SYSTEM OVERVIEW

User:

The User can register the first. While registering he required a valid user email and mobile for further communications. Once the user register then admin can activate the user. Once admin activated the user then user can login into our system. User can upload the dataset based on our dataset column matched. For algorithm execution data must be in float format. Here we took liver disease dataset for testing purpose. User can also add the new data for existing dataset based on our Django application. User can click the Classification in the web page so that the data calculated Accuracy, precision, sensitivity and specificity based on the algorithms. User can click Prediction in the web page so that user can write the review after predict the review That will display results depends upon review like positive,negative or neutral.

Admin:

Admin can login with his login details. Admin can activate the registered users. Once he activate then only the user can login into our system. Admin can view the overall data in the browser. Admin can click the Results in the web page so calculated Accuracy, precision, sensitivity and specificity based on the algorithms is displayed. All algorithms execution complete then admin can see the overall accuracy in web page.

Data Preprocessing:

A dataset can be viewed as a collection of data objects, which are often also called as a records, points, vectors, patterns, events, cases, samples, observations, or entities. Data objects are described by a number of features that capture the basic characteristics of an object, such as the mass of a physical object or the time at which an event occurred, etc. The data preprocessing in this forecast uses techniques like removal of noise in the data, the expulsion of missing information, modifying default values if relevant and grouping of attributes for prediction at various levels.

2.4 PROJECT TIMELINE CHART

Fig.2.4 Project Timeline Chart

2.5 TASK DISTRIBUTION

2.5.1 DESIGN PHASE

Table No 2.5.1: Design Phase

Name of Student	Task Performed	Result
Devesh P. Patil	Project Design	Completed Successfully
Ajay S. Bangar	Suggested Appropriate contents	Completed Successfully
Subodh M. Gawali	Documentation	Completed Successfully

2.5.2 IMPLEMENTATION PHASE

Table No 2.5.2: Implementation Phase

Name of Student	Implementation Task	Result
Devesh P. Patil	Project Implementation	Completed Successfully
Ajay S. Bangar	Implemented Project Algorithm.	Completed Successfully
Subodh M. Gawali	Documentation	Completed Successfully

SOFTWARE REQUIRMENT SPECIFICATION

3. SOFTWARE REQUIREMENT SPECIFICATION

3.1 HARDWARE REQUIREMENTS

System	:	Intel core i3.
Hard Disk	:	256 GB.
Monitor	:	14’ Color Monitor.
Mouse	:	Optical Mouse.
Ram	:	4 GB.
Keyboard	:	101 Keyboard Keys

3.2SOFTWARE REQUIREMENTS

Operating system	:	Windows 10 and above.
Coding Language	:	Python
Software’s used	:	eclipse IDE, Python 3.6.3 .
Languages Used	:	JavaScript, HTML,CSS,python
Technology Used	:	Machine Learning
Algorithms	:	SVM, Logistic Regression

SYSTEM DESIGN

4. SYSTEM DESIGN

4.1 DESIGN SPECIFICATION

4.1.1 ALGORITHM

Algorithm for Logistic regression-

```
def start_logistic_regression():
    tuned_params = {'C': [0.0001, 0.001, 0.01, 0.1, 1, 10, 100, 1000, 10000], 'penalty': ['l1', 'l2']}
    model = GridSearchCV(LogisticRegression(), tuned_params, scoring='roc_auc', n_jobs=
    1) model.fit(X_train, y_train)
    model.best_estimator_
    ## Predict Train set results
     y_train_pred=model.predict(X_train)
     ## Predict Test set results
    y_pred = model.predict(X_test)
    # Get just the prediction for
     the positive class (1) y_pred_proba = model.predict_proba(X_test)[:,1]
    # Display first 10 predictions
    y_pred_proba[:10]
    i = 28  ## Change the value of i to get the details of any point (56, 213, etc.)
    print('For test point {}, actual class = {}, precited class = {}, predicted probability =
    {}'. format(i, y_test.iloc[i], y_pred[i], y_pred_proba[i]))
    confusion_matrix(y_test, y_pred).
    plt.title('Receiver Operating Characteristic')

    def start_svm():
    from sklearn import svm
    def svc_param_selection(X, y, folds):
    Cs = [0.001, 0.01, 0.1, 1, 10]
    gammas = [0.001, 0.01, 0.1, 1]
    param_grid = {'C': Cs, 'gamma': gammas}
    grid_search = GridSearchCV(svm.SVC(kernel='rbf'),
```

```
 param_grid, cv=nfolds) grid_search.fit(X_train, y_train)
grid_search.best_params_
return
grid_search.best_params_
# svClassifier = SVC(kernel='rbf')
# svClassifier.fit(X_train, y_train)
```

4.1.2 WORKING OF ALGORITHM

Logistic Regression Algorithm

1. Initialize Logistic Regression Model:

- Create an object `lgr` for the logistic regression model.

```
python
from sklearn.linear_model import LogisticRegression
lgr = LogisticRegression()
```

2. Fit the Model with Oversampled Data:

- Use the oversampled data `X_over` (features) and `y_over` (target labels) to train the logistic regression model.

```
python
lgr.fit(X_over, y_over)
```

3. Make Predictions:

- Use the trained logistic regression model to predict labels for the test data `X_test`.

```
python
y_pred = lgr.predict(X_test)
```

4. Generate Classification Report:

- Calculate the classification report to evaluate the model's performance. The classification report provides metrics like precision

5. Prediction:

- After training, for a new input \(x \), the logistic regression model calculates the probability \(h_{\theta}(x) \). If the probability is greater than 0.5, the model predicts `y = 1`; otherwise, it predicts `y = 0`.

6. Evaluation:

The classification report provides a comprehensive evaluation of the model performance by calculating various metrics like precision, recall, F1-score, and support for both classes (usually **0** and **1**)

4.2 SYSTEM ARCHITECTURE

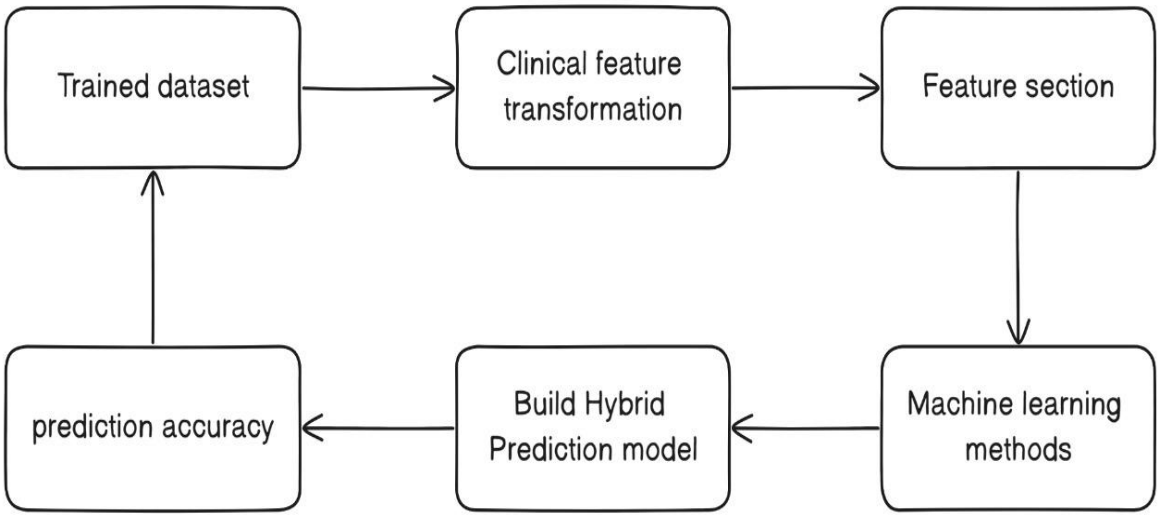


Fig.4.2 System Architecture Diagram

4.3 DATA FLOW DIAGRAM

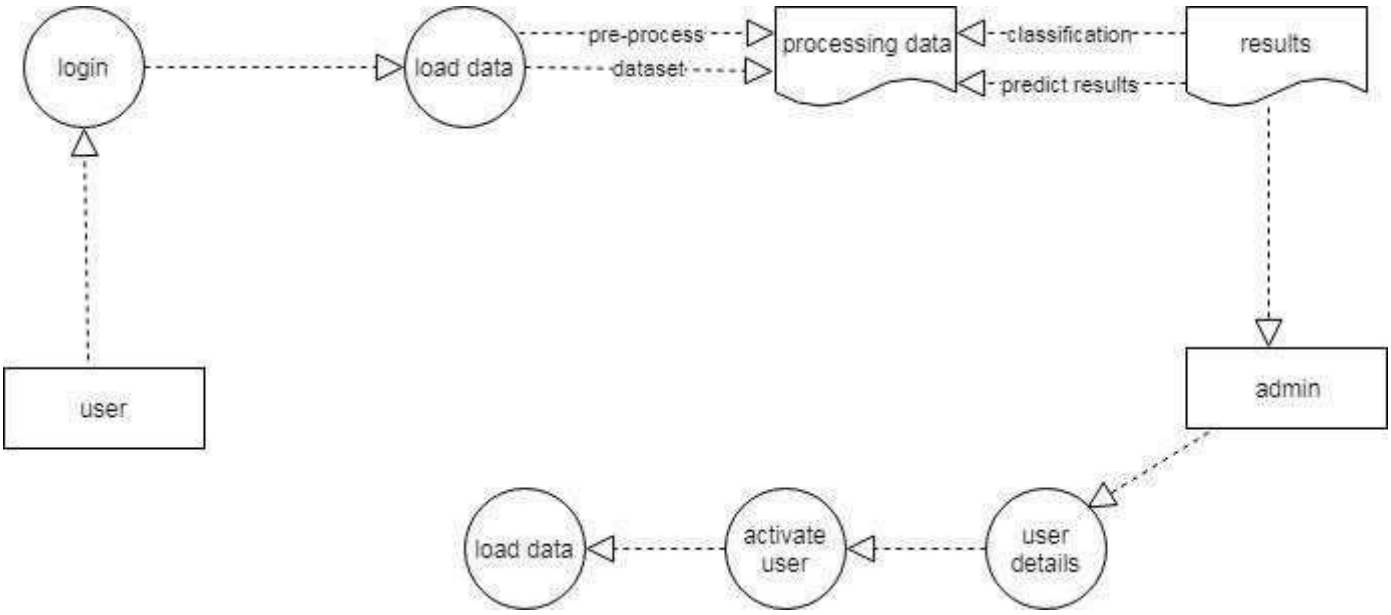


Fig.4.3.1 DFD Level 0 diagram

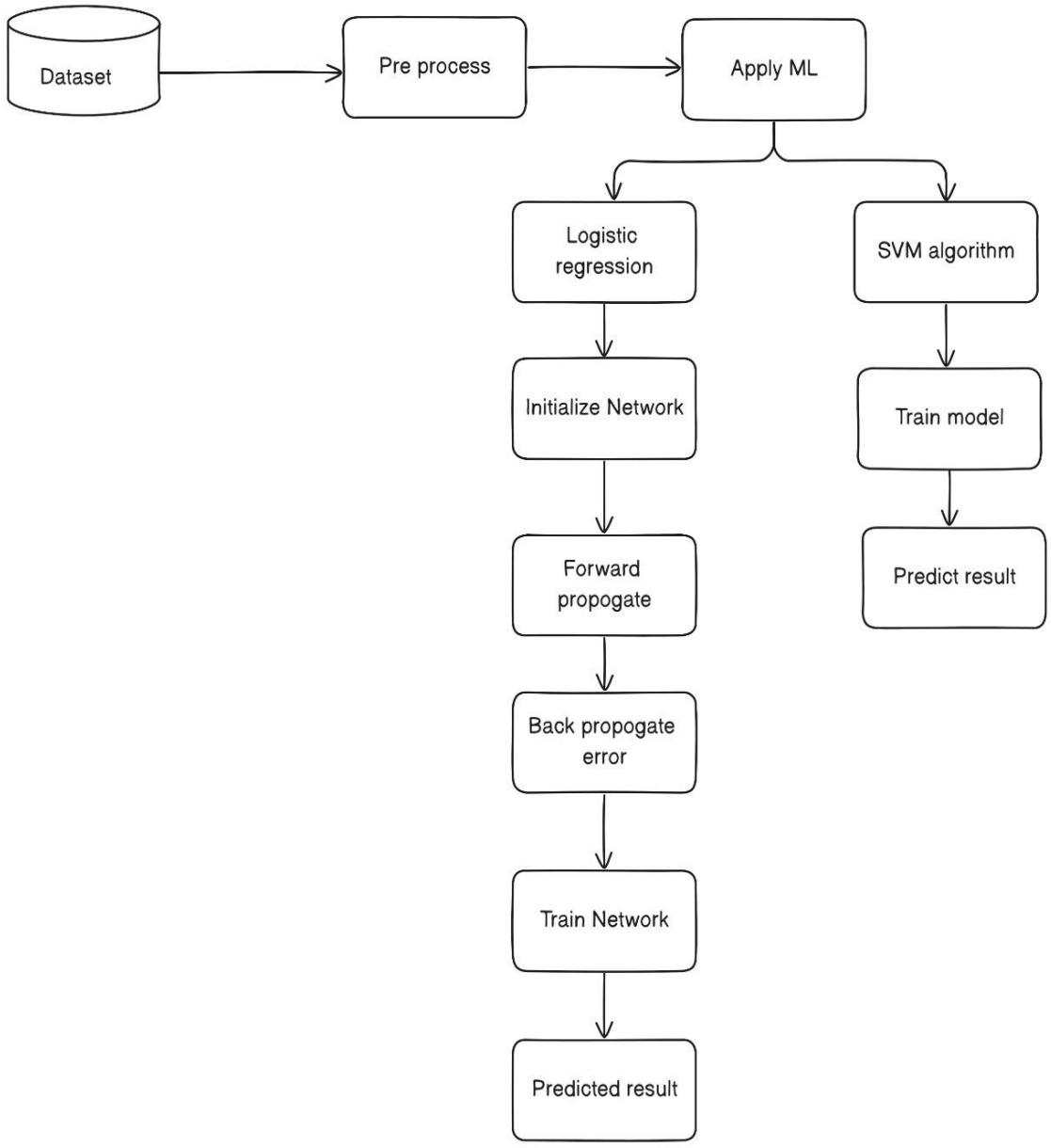


Fig.4.3.2 DFD Level 1 diagram

4.4 FLOW CHART

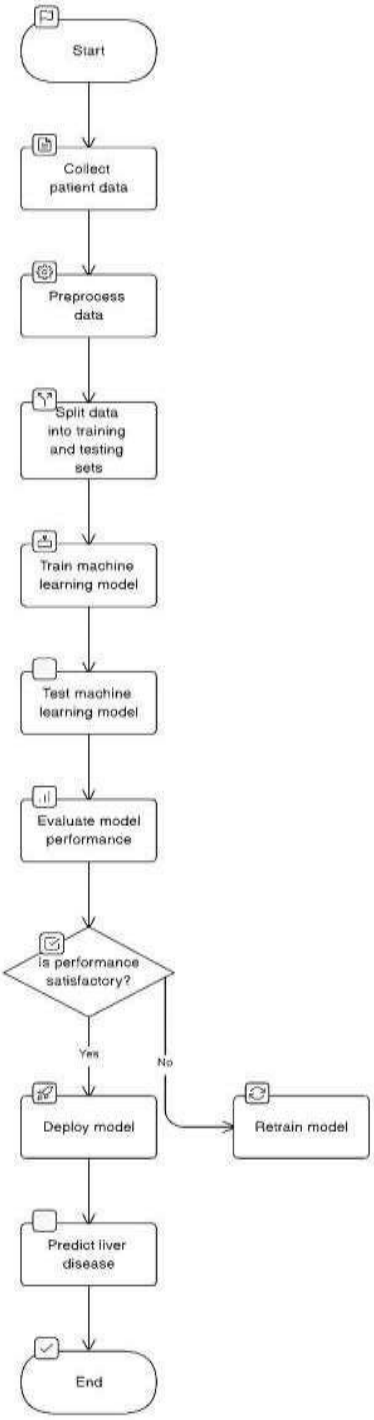


Fig.4.4 Flowchart

4.5 UML DIAGRAMS

4.5.1 CLASS DIAGRAM

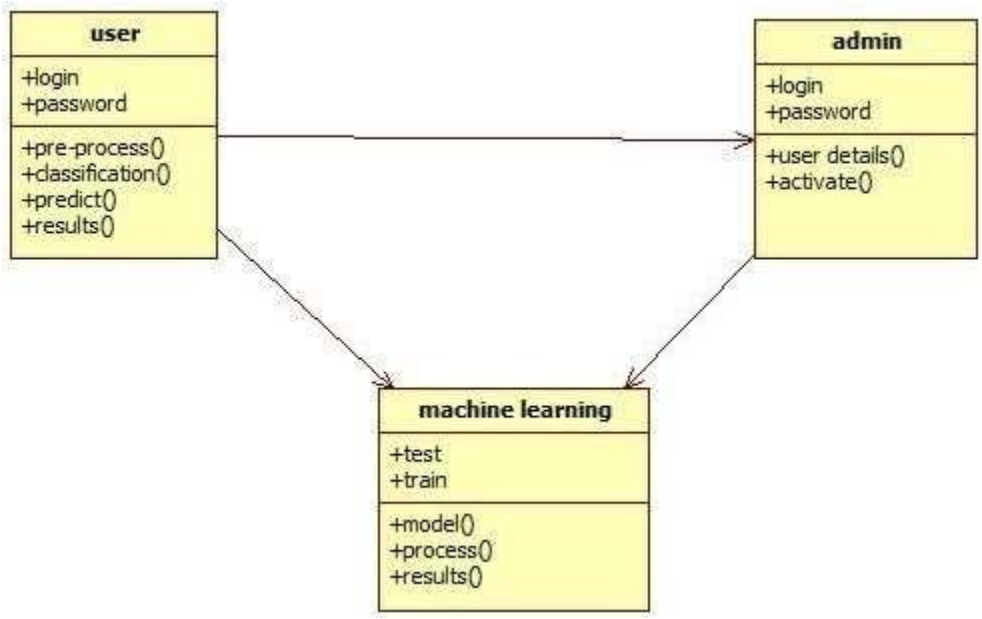


Fig.4.5.1 .Class Diagram

4.5.2 SEQUENCE DIAGRAM

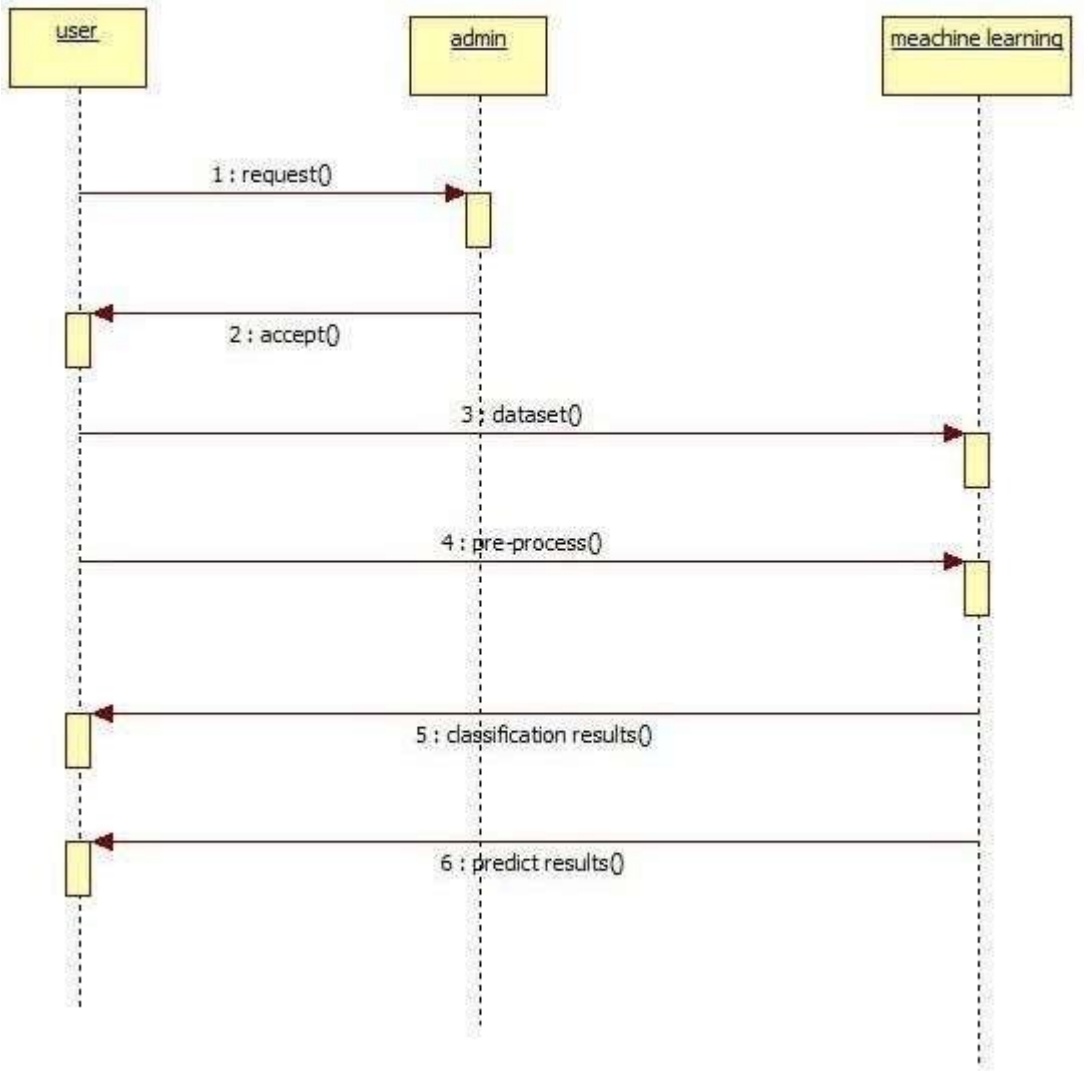


Fig.4.5.2 Sequence Diagram

4.5.3 USE CASE DIAGRAM

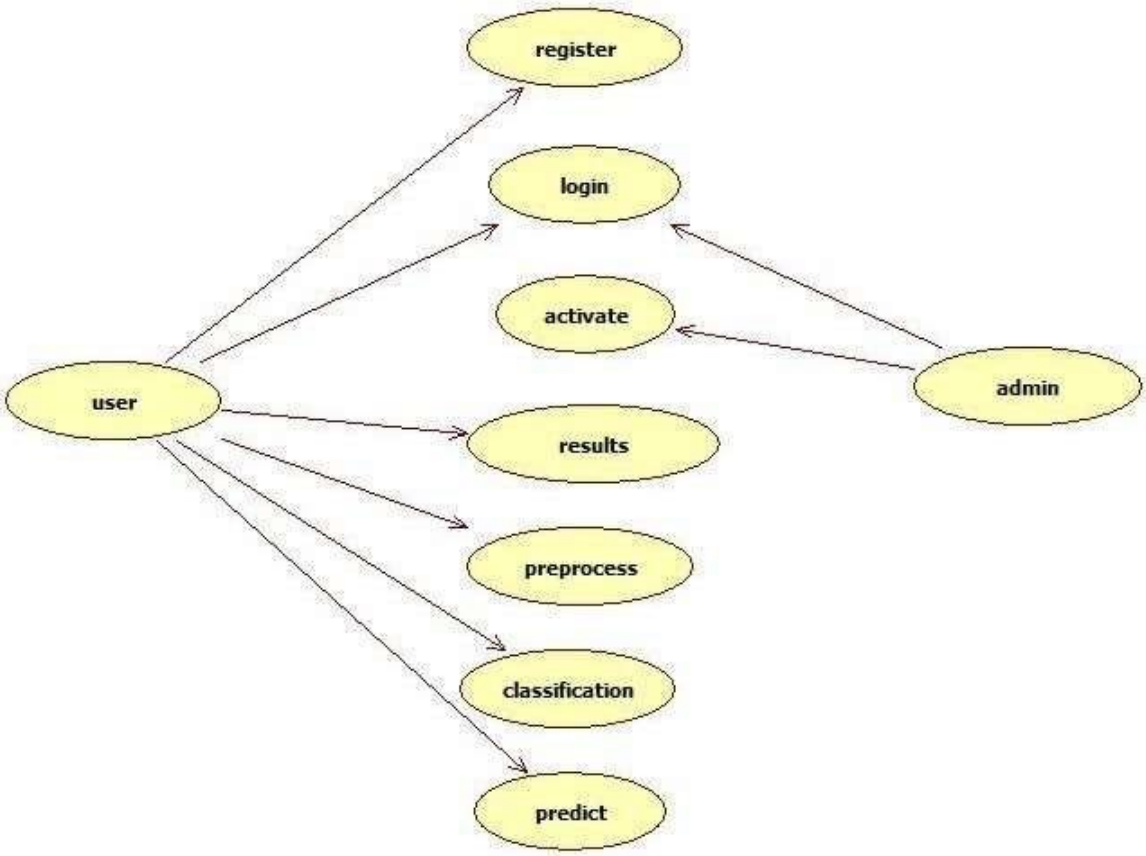


Fig.4.5.3 Use Case Diagram

4.5.4 ACTIVITY DIAGRAM

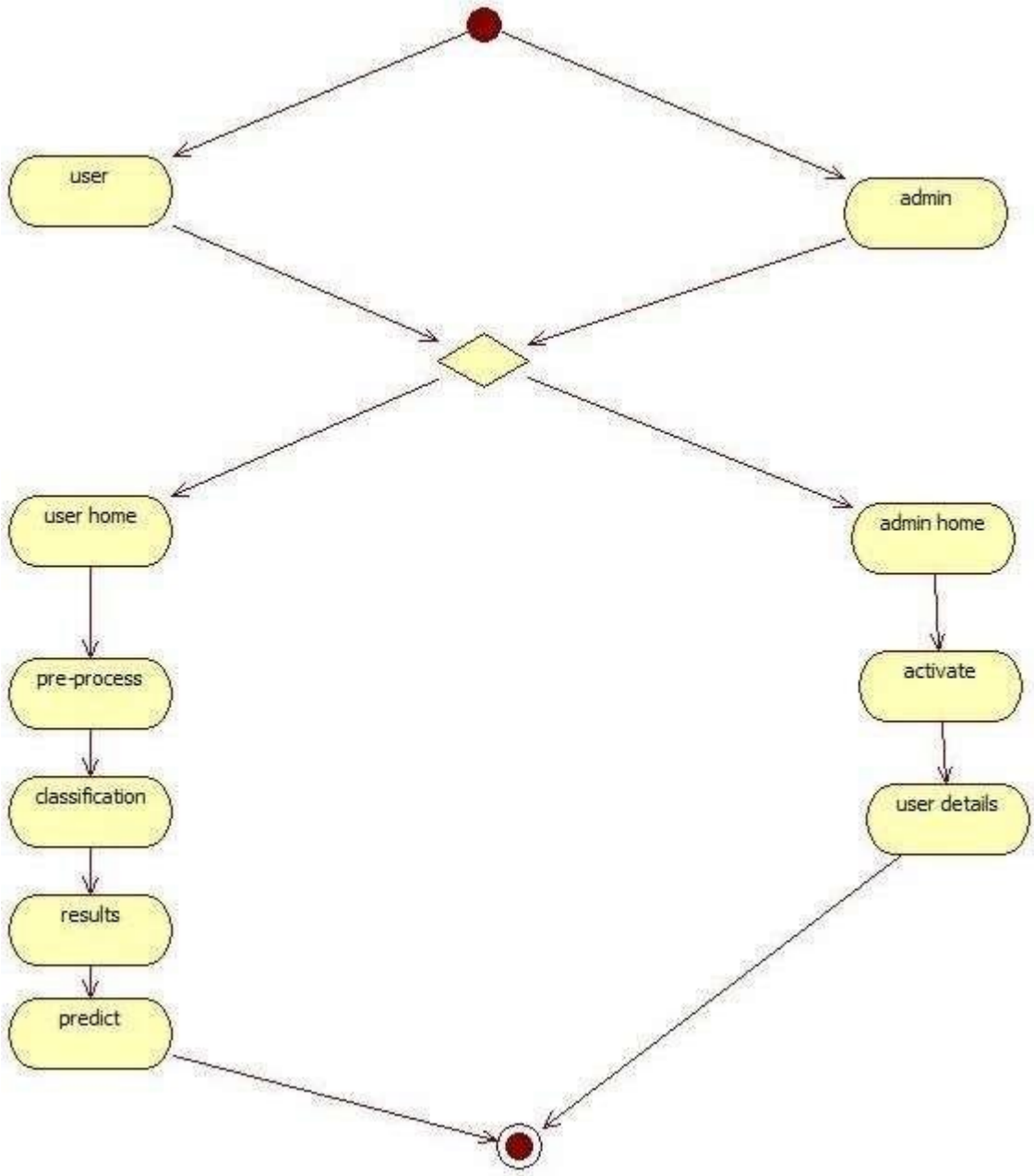


Fig. 4.5.4 Activity Diagram

PROJECT IMPLEMENTATION

5. PROJECT IMPLEMENTATION

5.1 TECHNOLOGY OVERVIEW

PYTHON

Python is a general-purpose interpreted, interactive, object-oriented, and high-level programming language. An interpreted language, Python has a design philosophy that emphasizes code readability (notably using whitespace indentation to delimit code blocks rather than curly brackets or keywords), and a syntax that allows programmers to express concepts in fewer lines of code than might be used in languages such as C++or Java. It provides constructs that enable clear programming on both small and large scales. Python interpreters are available for many operating systems. CPython, the reference implementation of Python, is open source software and has a community-based development model, as do nearly all of its variant implementations. C Python is managed by the non-profit Python Software Foundation. Python features a dynamic type system and automatic memory management. It supports multiple programming paradigms, including object-oriented, imperative, functional and procedural, and has a large and comprehensive standard library.

DJANGO

Django is a high-level Python Web framework that encourages rapid development and clean, pragmatic design. Built by experienced developers, it takes care of much of the hassle of Web development, so you can focus on writing your app without needing to reinvent the wheel. It’s free and open source.

Django's primary goal is to ease the creation of complex, database-driven websites. Django emphasizes reusability and "pluggability" of components, rapid development, and the principle of don't repeat yourself. Python is used throughout, even for settings files and data models

5.2 CODING

User side views.py

```
from django.shortcuts import render


# Create your views here.
from django.shortcuts import render, HttpResponse
from django.contrib import messages
from sklearn.tree import DecisionTreeClassifier
from .forms import UserRegistrationForm
from .models import UserRegistrationModel
# Create your views here.
def UserRegisterActions(request):
      if request.method == 'POST':
       form = UserRegistrationForm(request.POST)
      if form.is_valid():
          print('Data is Valid')
  formsave()
  messages.success(request, 'You have been successfully registered')
  form = UserRegistrationForm()
  return render(request, 'UserRegistrations.html', {'form': form})
      else:
   messages.success(request, 'Email or Mobile Already Existed')
   print("Invalid form")
     else:
      form = UserRegistrationForm()
  return render(request, 'UserRegistrations.html', {'form': form})


    def UserLoginCheck(request):
    if request.method == "POST":
      loginid = request.POST.get('loginid')
```

```
    pswd = request.POST.get('pswd')
    print("Login ID = ", loginid, ' Password = ', pswd)
    try:

        status = check.status
        print('Status is = ', status)
        if status == "activated":
            request.session['id'] = check.id
        request.session['loggeduser'] = check.name
        request.session['loginid'] = loginid
        request.session['email'] = check.email
        print("User id At", check.id, status)
        return render(request, 'users/UserHomePage.html', {})
        else:
            messages.success(request, 'Your Account Not at activated')
            return render(request, 'UserLogin.html')
       except Exception as e:
        print('Exception is ', str(e))
        pass
    messages.success(request, 'Invalid Login id and password')
   return render(request, 'UserLogin.html', {})

  def UserHome(request):
   return render(request, 'users/UserHomePage.html', {})

  def PreProcess(request):
 import matplotlib.pyplot as plt
 import seaborn as sns
# inline
import pandas as pd
from django.conf import settings
import os
```

```
path = os.path.join(settings.MEDIA_ROOT, 'indian_liver_patient.csv')
data = pd.read_csv(path)
# checking the stats
# given in the website 416 liver disease patients and 167 non liver disease patients
# need to remap the classes liver disease:=1 and no liver disease:=0 (normal convention to be
followed)
data['Dataset'] = data['Dataset'].map(
    {2: 0, 1: 1})
  count_classes = pd.value_counts(data['Dataset'], sort=True).sort_index()
  count_classes.plot(kind='bar')
 # plt.title("Liver disease classes bar graph")
 # plt.xlabel("Dataset")
 # plt.ylabel("Frequency")
 # plt.savefig('classlabels.png')
 # data['Dataset'] = data['Dataset'].map({2: 0, 1: 1})
 data['Albumin_and_Globulin_Ratio'].fillna(value=0, inplace=True)
 data_features = data.drop(['Dataset'], axis=1)
 data_num_features = data.drop(['Gender', 'Dataset'], axis=1)
 data_num_features.head()
 data_num_features.describe()  # check to whether feature scaling has to be performed or not
 from sklearn.preprocessing import StandardScaler
 scaler = StandardScaler()
 cols = list(data_num_features.columns)
  data_features_scaled = pd.DataFrame(data=data_features)
  data_features_scaled[cols] = scaler.fit_transform(data_features[cols])
 data_features_scaled.head()
 data_exp = pd.get_dummies(data_features_scaled)
data_exp.head()
# Set up the matplotlib figure
# f, ax = plt.subplots(figsize=(12, 10))
# plt.title('Pearson Correlation of liver disease Features')
# # Draw the heatmap using seaborn
```

```
    # sns.heatmap(data_num_features.astype(float).corr(), linewidths=0.25, vmax=1.0, square=True,
cmap="YlGnBu",
    #             linecolor='black', annot=True)
    # plt.savefig('corr.png')
    # plt.show()
    return render(request, 'PreProcess.html', {"data": data_num_features.to_html
```

TESTING

6. TESTING

6.1 SYSTEM TESTING

The purpose of testing is to discover errors. Testing is the process of trying to discover every conceivable fault or weakness in a work product. It provides a way to check the functionality of components, sub assemblies, assemblies and/or a finished product It is the process of exercising software with the intent of ensuring that the Software system meets its requirements and user expectations and does not fail in an unacceptable manner. There are various types of test. Each test type addresses a specific testing requirement.

TYPES OF TESTS

Unit testing

Unit testing involves the design of test cases that validate that the internal program logic is functioning properly, and that program inputs produce valid outputs. All decision branches and internal code flow should be validated. It is the testing of individual software units of the application .it is done after the completion of an individual unit before integration. This is a structural testing, that relies on knowledge of its construction and is invasive. Unit tests perform basic tests at component level and test a specific business process, application, and/or system configuration. Unit tests ensure that each unique path of a business process performs accurately to the documented specifications and contains clearly defined inputs and expected results.

Integration testing

Integration tests are designed to test integrated software components to determine if they actually run as one program. Testing is event driven and is more concerned with the basic outcome of screens or fields. Integration tests demonstrate that although the components were individually satisfaction, as shown by successfully unit testing, the combination of components is correct and consistent. Integration testing is specifically aimed at exposing the problems that arise from the combination of components.

System Test

Testing ensures that the entire integrated software system meets requirements. It tests a configuration to ensure known and predictable results. An example of system testing is the configuration oriented system integration test. System testing is based on process descriptions and flows, emphasizing pre-driven process links and integration points.

White Box Testing

White Box Testing is a testing in which in which the software tester has knowledge of the inner workings, structure and language of the software, or at least its purpose. It is purpose. It is used to test areas that cannot be reached from a black box level.

RESULTS AND DISCUSSIONS

7. RESULTS AND DISCUSSIONS

7.1 RESULT SETS

Table no. 7.1.1: Test Cases amd results

S.no	Test Case	Excepted Result	Result	Remarks(IF Fails)
1.	User Register	If User registration successfully.	Pass	If already user email exist then it fails.
2.	User Login	If Username and password is correct then it will getting valid page.	Pass	Un Register Users will not logged in.
3.	User View User	Show our dataset	Pass	If Data set Not Available fail.
4.	User Machine Learning	For our two Models the accuracy and F1 score will calculated.	Pass	If Accuracy and F1 Score Not Displaying fail
5.	User Predict	Display Review with true results	Pass	Results not True Fail
6.	View classification Report	Display Classification Results	Pass	Results not true Fail
7.	View Results	Display Results	Pass	results not Displayed Fail.
8.	Admin login	Admin can login with his login credential. If success he get his home page	Pass	Invalid login details will not allowed here
9.	Admin can activate the register users	Admin can activate the register user id	Pass	If user id not found then it won't login
10.	Results	For our Two models the accuracy and F1 Score	Pass	If Accuracy And F1 Score Not Displayed fail

7.2 SCREENSHOTS

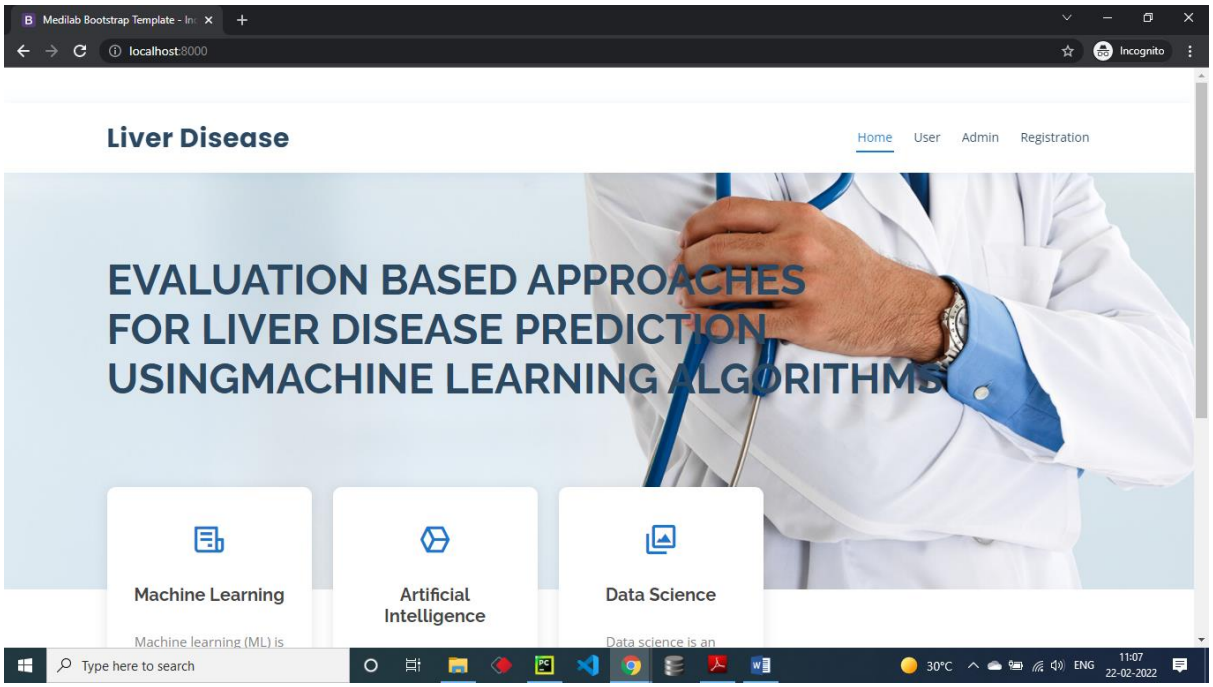


Fig.7.2.1: home page

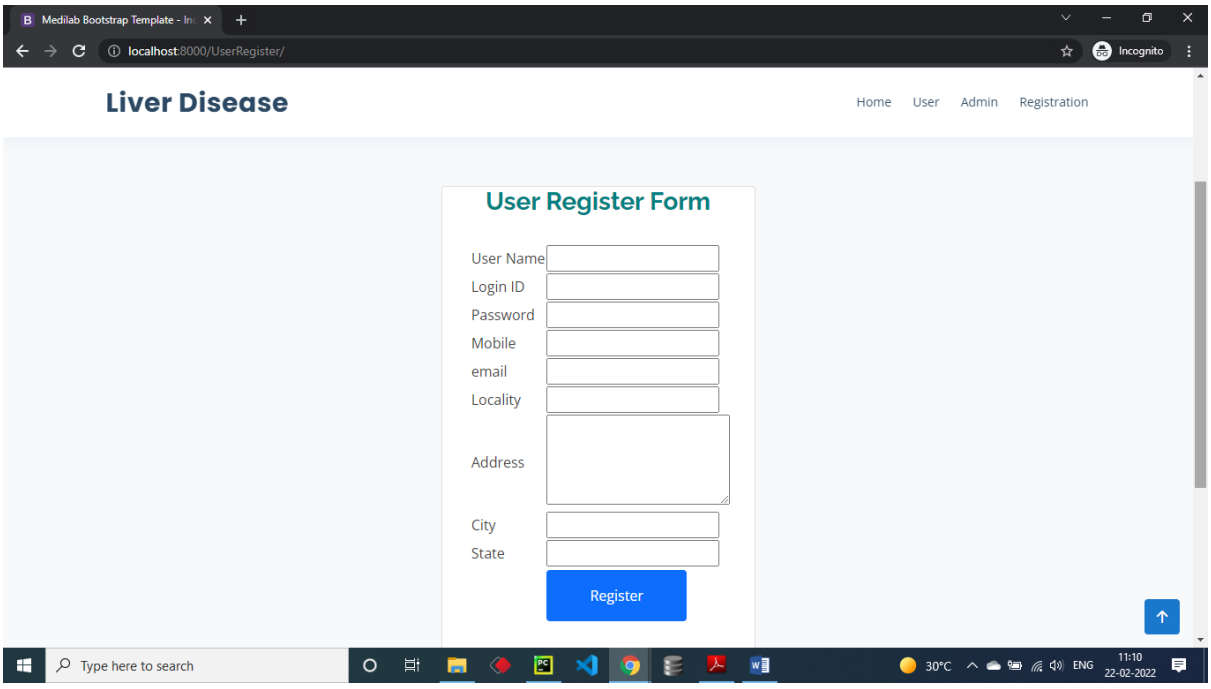


Fig.7.2.2: Registration form

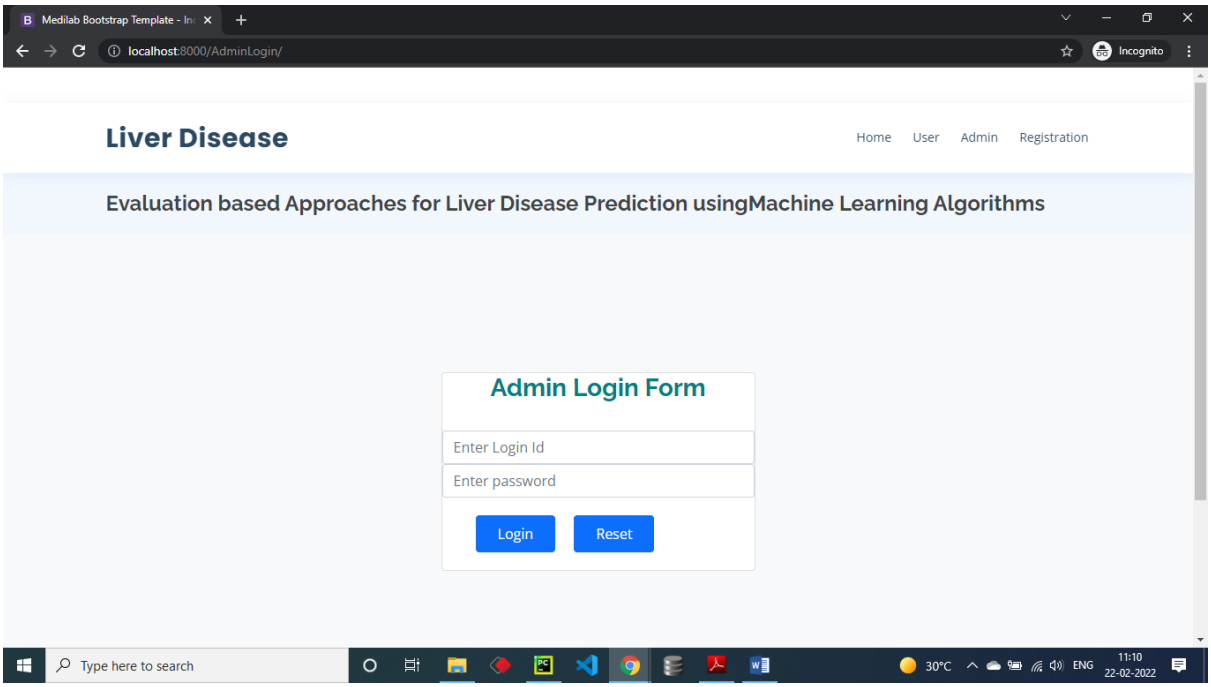


Fig.7.2.3: Admin Login

ADVANTAGES

8. ADVANTAGES

- The primary benefit of Algorithms for Machine Learning place of traditional predictive models is their ability to identify new patterns between variables and produce predictions by learning from preexisting data.
- It has been proved that MLAs raise the precision of identifying people who are at risk of illness.
- Learning methods with the assistance of a supervisor, teacher, or instructor are called as supervised learning. It consists of a training set of patterns linked to label data, which facilitates the learning and prediction of the algorithm from input to output.

CONCLUSION

CONCLUSION

Thus, we have tried to implement the paper "C.Geetha Dr.AR.Arunachalam" "Evaluation based Approaches for Liver Disease Prediction Machine Learning Algorithms" 2021 IEEE Access and conclusion is as follows methods for identifying and evaluating liver disease in patients have been proposed and assessed utilizing machine learning methods Logitistic regression, or SVM, is a machine learning technique that combines two main approaches. These models have been used in the prediction analysis.

FUTURE WORK

FUTURE WORK

Using all the models, the prediction analysis has been implemented and their performance has been assessed. The probability of liver disease prediction attained with high accuracy .In future , the present scenario can be compared with other techniques such as naive bayes classification, Random forest etc. Also this work can be further focused on implementation of parametric classifications by bio-inspired optimization algorithms.

REFERENCES

REFERENCES:-

[1] C. Geetha and A. Arunachalam, "Evaluation based Approaches for Liver Disease Prediction using Machine Learning Algorithms," 2021 International Conference on Computer Communication and Informatics (ICCCI), Coimbatore,India, 2021, pp. 1-4,

[2] Dinh-Van Phan.Chien-Lung Chan,Ai-Hsien A Adams Li , Ting-Ying Chien,Van-Chuc Nguyen," Liver Cancer Prediction in a Viral Hepatitis Cohort: A Deep Learning Approach"2020 International journal of cancer(IJC) 10.1002/ijc.33245

[3] Gullo, F.: From patterns in data to knowledge discovery: what data mining can do. Phys. Procedia 62, 18–22 (2015)

[4] Manjeevan Seera, Chee Peng Lim, A hybrid intelligent system for medical data classification, Expert Systems with Applications 41 (2014) 2239– 2249.

[5] Zhi-Hua Zhou, Ensemble Methods: Foundations and Algorithms, Machine Learning & Pattern Recognition Series, 2012.

[6] Ayodele, T.O.: Types of machine learning algorithms. In: New Advances in Machine Learning. In Tech (2010)

[7] Lin, Rong-Ho. "An intelligent model for liver disease diagnosis." Artificial intelligence in medicine vol. 47,1 (2009): 53-62. i:10.1016/j.artmed.2009.05.005

[8] Michael J Sorich. An intelligent model for liver disease diagnosis. Artificial Intelligence in Medicine 2009;47:53—62.

APPENDIX

APPENDIX

1) Open the Project path and run one by one

2) Python 1_Run Registration

3) Python 2_Run Admin(Activate the user)

4) Python 3_Run user login

5) Python 4_run preprocess dataset

6) Python 5_run logistic algorithms, SVM

7) Python 6_Test Prediction.