Project Based Learning-II
(Guidelines and Work Book)
Course Code: 210258
(2019 Course)

Second Year Engineering

Year: 2021 - 2022

Group ID: | 1 | 1 |

Team Members: 1.   SAI&D02 Prathamesh Patil

2.   SAI&D07 Ganesh Jagzap

3.   SAI&D08 Jay Sawant

4.   SAI&D09 Atharv Wankhede

Project Title   :   **Heart Disease Prediction System**

Name of Mentor :   Mrs. Shubhangi Vairagar

DPU

DEPARTMENT OF ARTIFICIAL INTELLINENCE AND DATA SCIENCE

DR. D. Y. PATIL INSTITUTE OF TECHNOLOGY, PIMPRI, PUNE
SAVITRIBAI PHULE PUNE UNIVERSITY
2021 - 2022

Preamble

For better learning experience, along with traditional classroom teaching and laboratory learning; project-based learning has been introduced with an objective to motivate students to learn by working in group cooperatively to solve a problem, Project-based Learning (PBL) is a student centric pedagogy that involves a dynamic classroom approach in which it is believed that students acquire a deeper knowledge through active exploration of real-world challenges and problems. Students learn about a subject by working for an extended period of time to investigate and respond to a complex question, challenge or a problem. It is a style of active learning and inquiry-based learning. (Reference: Wikipedia). Problem based learning will also redefine the role of teacher as mentor in learning process. Along with communicating knowledge to students, often in a lecture setting, the teacher will also to act as an initiator and facilitator in the collaborative process of knowledge transfer and development.

This is a recommended workbook for PBL that will serve the purpose and facilitate the job of students, mentor and coordinator. This workbook will reflect accountability, punctuality, technical writing ability and work flow of the work undertaken.

CERTIFICATE

This is to certify that Mr./ Ms. **1.SAI&D02 Prathamesh Patil**

                                  2.SAI&D07 Ganesh Jagzap

                                  3.SAI&D08 Jay Sawant

                                  4.SAI&D09 Atharv Wankhede

Group No. **11** Branch **Artificial Intelligence and Data Science** has successfully completed the work associated with **Project Based Learning II (210258**) titled as **Heart Disease Prediction System** and has submitted the work book associated under my supervision, in the partial fulfillment of Second Year Bachelor of Engineering (Choice Based Credit System) (2019 course) of Savitribai Phule Pune University.

Date:

Place: Pimpri

  Guide   Head   Principal

(Name & Sign)   (Name & Sign)   (Name & Sign)

Table of Contents

Sr. No.	Description	Page No.
1.	Preamble	2.
2.	Certificate	3.
3.	Project Based Learning Syllabus	5.
4.	Recommended Guidelines and Phases	7.
5.	Evaluation and Continuous Assessment Sheet	8.
6.	Project Information Sheet	9.
7.	Assessment Remarks Sheet	10
8.	Project Monitoring (1 sheet per week)	12 onwards

1. Project Based Learning Syllabus:

Course Objectives:
1. To emphasize learning activities that is long-term, interdisciplinary and student-centric.
2. To inculcate independent learning by problem solving with social context.
3. To engages students in rich and authentic learning experiences.
4. To provide every student the opportunity to get involved either individually or as a group so as to develop team skills and learn professionalism

Course Outcomes:
CO1: Project based learning will increase their capacity and learning through shared cognition.
CO2: Students able to draw on lessons from several disciplines and apply them in practical way.
CO3: Learning by doing approach in PBL will promote long-term retention of material and replicable skill, as well as improve teachers' and students' attitudes towards learning

Group Structure:
Working in supervisor/mentor – monitored groups. The students plan, manage and complete a task/project/activity which addresses the stated problem.
- There should be team/group of 5 -6students
- A supervisor/mentor teacher assigned to individual groups

Selection of Project/Problem:
The problem-based project-oriented model for learning is recommended. The model begins with the identifying of a problem, often growing out of a question or "wondering". This formulated problem then stands as the starting point for learning. Students design and analyze the problem within an articulated interdisciplinary or subject frame. A problem can be theoretical, practical, social, technical, symbolic, cultural and/or scientific and grows out of students' wondering within different disciplines and professional environments. A chosen problem has to be exemplary. The problem may involve an interdisciplinary approach in both the analysis and solving phases. By exemplarity, a problem needs to refer back to a particular practical, scientific, social and/or technical domain. The problem should stand as one specific example or manifestation of more general learning outcomes related to knowledge and/or modes of inquiry. There are no commonly shared criteria for what constitutes an acceptable project. Projects vary greatly in the depth of the questions explored, the clarity of the learning goals, the content and structure of the activity.
- A few hands-on activities that may or may not be multidisciplinary
- Use of technology in meaningful ways to help them investigate, collaborate, analyze, synthesize and present their learning.
- Activities may include-Solving real life problem, investigation /study and Writing reports of in depth study, field work.

Assessment:
The institution/head/mentor is committed to assessing and evaluating both student performance and program effectiveness. Progress of PBL is monitored regularly on weekly basis. Weekly review of the work is necessary. During process of monitoring and continuous assessment AND evaluation the individual and team performance is to be measured. PBL is monitored and continuous assessment is done by supervisor/mentor and authorities. Students must maintain an institutional culture of authentic collaboration, self-motivation, peer-learning and personal responsibility. The institution/department should support students in this regard through guidance/orientation programs and the provision of appropriate resources and services. Supervisor/mentor and Students must actively participate in assessment and evaluation processes.
Group may demonstrate their knowledge and skills by developing a public product and/or report and/or presentation.

- Individual assessment for each student (Understanding individual capacity, role and involvement in the project)
- Group assessment (roles defined, distribution of work, intra-team communication and togetherness)
- Documentation and presentation

Evaluation and Continuous Assessment:

It is recommended that the all activities are to be record and regularly, regular assessment of work to be done and proper documents are to be maintained at college end by both students as well as mentor (you may call it PBL work book).

Continuous Assessment Sheet (CAS) is to be maintained by all mentors/department and institutes. Recommended parameters for assessment, evaluation and weightage:

- Idea Inception (5%)
- Outcomes of PBL/ Problem Solving Skills/ Solution provided/ Final product (50%) (Individual assessment and team assessment)
- Documentation (Gathering requirements, design & modeling, implementation/execution, use of technology and final report, other documents) (25%)
- Demonstration (Presentation, User Interface, Usability etc) (10%)
- Contest Participation/ publication (5%)
- Awareness /Consideration of -Environment/ Social /Ethics/ Safety measures/Legal aspects (5%)

PBL workbook will serve the purpose and facilitate the job of students, mentor and project coordinator. This workbook will reflect accountability, punctuality, technical writing ability and work flow of the work undertaken.

References:

- Project-Based Learning, Edutopia, March 14, 2016.
- What is PBL? Buck Institutes for Education
- www.schoology.com
- www.wikipedia.org
- www.howstuffworks.com

2. Recommended Guidelines and Phases:

PBL is learning through activity. One of the teachers can be appointed as coordinator for PBL. Following are the recommended guidelines that will work as an initiator and facilitator in process of completion of PBL.

1. In first week of commencement of 2nd semester or preferably at the end of first semester let the coordinator create awareness about PBL (what, why, and how) among the students. Convey students expected outcomes, assessment process and evaluation criteria.

2. Get groups of students registered preferably 4-6 students per group.

3. Assign mentor to each group.

4. Provide guidelines for title identification (Problem can be some real-life situation that needs technology solutions. This situation can be identified by meeting people around, visiting various industries, society, and institutes. The solution can be prototype, model, convertible solutions, survey and analysis, simulation, and similar).

5. Let students submit the problem identified in prescribed format (Title, Problem statement, details of a problem undertaken, and what is need of solution to the problem)

6. Coordinator and mentor can approve the problem statements based on feasibility and learning outcomes expected for first year engineering students

7. Mentor is to monitor progress of the task during phases of project work. Broadly phases may include-requirements gathering, preparing a solution, technology design for the solution. (Optional phases-implementation and testing)

8. Weekly monitoring and continuous assessment record is to be maintained by mentor.

9. Get the report submitted at the end of semester.

3. Evaluation and Assessment Sheet (To be filled in my mentor)

Sr. No.	Details	Maximum Marks	Marks Obtained
1.	Problem Identification (Idea Inception)	10	
2.	Problem Analysis (Requirement Gathering)	15	
3.	Proposed Solution Model/Design/ Process / prototype	20	
4.	Technology Solution Model	15	
5.	Expected Outcomes	05	
6.	Implementation and Testing	10	
7.	Regularity (Attendance + Weekly Progress Reporting)	10	
8.	Awareness /Consideration of - Environment/ Social /Ethics/ Safety measures/Legal aspects	05	
9.	Contest Participation/ publication	05	
10.	Report	05	
	Total Marks	100	

Date:

Name & Sign of Mentor

4. Project Information Sheet

Project ID	11				
Title	**Heart Disease Prediction System**				
Problem Statement	Create a heart disease prediction system using machine learning where diseases are predicted by the machine learning algorithm on the basis of the symptoms provided by the user.				
Name of Mentor	Mrs. Shubhangi Vairagar				
Group Members	Division	Roll No.	Name	Mobile Number	Email ID
		SAI&D02	Prathamesh Patil	7083331851	20510441.dypit@dypvp.edu.in
		SAI&D07	Ganesh Jagzap	9075279739	20510148.dypit@dypvp.edu.in
		SAI&D08	Jay Sawant	9834730241	20510928.dypit@dypvp.edu.in
		SAI&D09	Atharv Wankhede	7385340991	20510790.dypit@dypvp.edu.in

5. **Continuous Assessment and Remarks Sheet**

Problem Identification (Idea Inception) –

- A study shows that from 1990 to 2016 the death rate due to heart diseases have increased around 34 per cent from 155.7 to 209.1 deaths per one lakh population in India.
- Thus, preventing heart diseases has become more than necessary. Good data-driven systems for predicting heart diseases can improve the entire research and prevention process, making sure that more people can live healthy lives. This is where Machine Learning comes into play. Machine Learning helps in predicting the heart diseases, and the predictions made are quite accurate.
- A dataset is formed by taking into consideration some of the information of some individuals. The problem is based on the given information about each individual we have to calculate that whether that individual will suffer from heart disease or not.
- Disease prediction using machine learning is a system where diseases are predicted by the machine learning algorithm on the basis of the symptoms provided by the user.

Problem Analysis (Requirement Gathering) –

- Machine learning is that domain of Data science that helps in predicting things based on the past data.
- In this project we will be collecting datasets, manipulate them and use machine learning algorithm (Random Forest) to create system that will efficiently predict the disease the user is suffering from on the basis of symptoms he/she possesses.

The system comprises of 2 major modules as follows:

❖ **Admin Module**

1. Add Training Data
2. Add Doctor Details
3. View User Details
4. View Feedback
5. View Doc Details
6. View Training Data

❖ **User Module**

1. Register (With Details like Age, Sex, etc.)
2. Check Heart (By providing Details like:-
    - o Age in Year
    - o Gender
    - o Chest Pain Type
    - o Fasting Blood Sugar
    - o Resting Electrographic Results(Restecg)
    - o Exercise Induced Angina(Exang)
    - o The slope of the peak exercise ST segment
    - o CA – Number of major vessels colored by fluoroscopy
    - o Thal
    - o Trest Blood Pressure
    - o Serum Cholesterol
    - o Maximum heart rate achieved(Thalach)
    - o ST depression induced by exercise(Oldpeak)

3. System will accordingly view Doctor to consult.

4.   Give Feedback
5.   View Doctor

❖ **User Classes and Characteristics**:

User A:
      Signifies a novice system customer. This user has no prior experience about using of this system. User A will find the product easy to use due to simple explanatory interface of the GUI (Graphical User Interface). The User A will have to enter his/her credentials (login details) which are new to the system and that will be saved permanently to the system database.

User B:
      An experienced customer. This user has a prior experience about using of this system. User B will find the product easy to use due to simple explanatory interface of the GUI (Graphical User Interface). The User B will have to enter his/her credentials (login details) if he/she wants or else it is already saved to the system.

Admin:
      The admin is totally familiar with the functioning of the system. The administrative person is responsible for storing and organising the credentials of the user A and user B and specially to resolve the system bugs or errors if any.

❖ **Operating Environment**:

- Windows 10 OS.
- Active Internet Connection.
- Desktop.
- Python and its libraries (NumPy, Pandas, Matplotlib, scikitlearn)
- Tkinkter for GUI application.

❖ **Constraints**:
      If the information given by the end user is incorrect i.e., if the values for the particular symptoms are entered incorrect i.e., out of consideration limit then it may create invalid result for the same case of the user.

External interface Requirements

1. **Users Interface**: -

The user will have a desktop application on a pc. The user will have to input his personal details, data related to predict disease on the desktop app and the result will be also shown on this app too.

2.**Hardware Interfaces**: -

The user will need to have a desktop in order to access the application and a valid internet connection.

3.**Software Interfaces**: -

The support libraries will include NumPy, pandas, Matplotlib, Scikit-learn and the algorithms will be one of the following logistic regression, random forest, naive bayes and a desktop application for user interface.

Functional Requirements: -

The details of the functionalities are as follows: -

1.Master file maintenance.

2. Login/signup of user

3.Saving data from user.

4.Input of data from prediction.

5.Comparison of data with the datasets.

6.Calculation for prediction.

7.Revaluation

8. Revaluation Journal

9. Linkage with other modules

10. Issue Division

11.Personnel and Administration Division

12.Research and Statistical Division

13. General Reports

14.MIS/Time Series Reports

5. **Detailed non -functional requirements**: -

(i) Functionality: - At a particular time huge number of people can access the application.

(ii) Usability: The desktop user interface shall be Windows 95/98/2000 XP/Windows 7/ 7 Ultimate ,8,9,10,11.

(iii) Performance: The heart disease prediction system can support many customers at a time. The speed and accurate transaction decide the performance factor. The screen must be clearly visible to the user. The system must record the input of the symptoms.

(iv) Security: The login credentials of the user will have to be properly stored so that someone cannot hack the system and get their hands on login credentials and datasets.

(v) Scope: The scope of this project is to allow the user to get prediction of the heart disease he/she may have through the symptoms provided by the user.

Proposed Solution Model/Design/ Process / prototype –
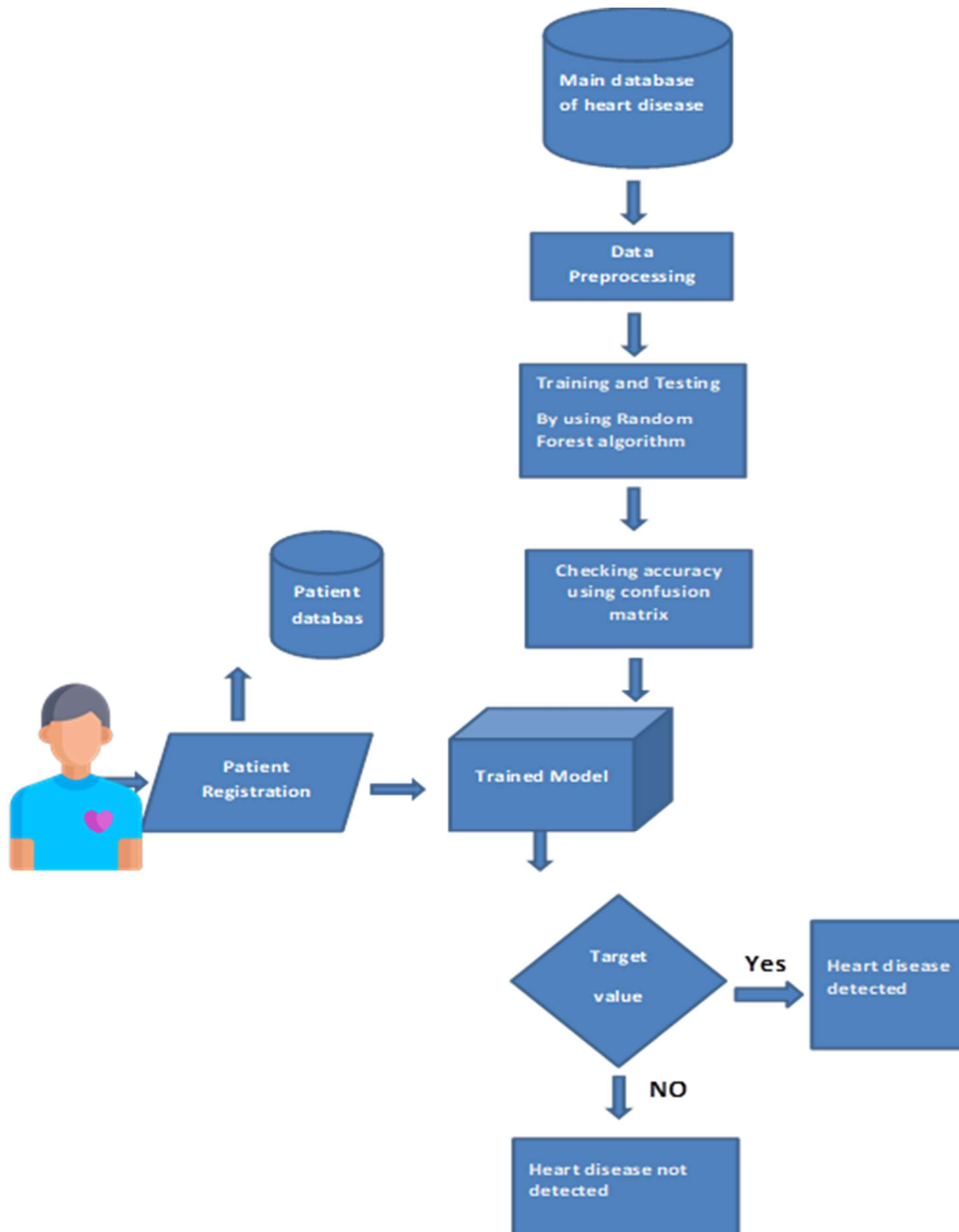
Technology Solution Model-

Random forest is a Supervised Machine Learning Algorithm that is used widely in Classification and Regression problems. It builds decision trees on different samples and takes their majority vote for classification and average in case of regression.

One of the most important features of the Random Forest Algorithm is that it can handle the data set containing continuous variables as in the case of regression and categorical variables as in the case of classification. It performs better results for classification problems.

Expected Outcomes-

- This system signifies its scope towards the early prediction of the heart disease which will be very helpful for the patient to identify whether he/she is suffering from heart disease or not.
- Moreover, it is very efficient for detecting the rare disease at early stage which will help the medical researchers to specify the treatment to the patient.
- Our desktop application aims for the same goal to predict the occurrence of heart related disease based on the various datasets acquired from Kaggle. We expect to find various hidden relationships and patterns in the data by applying machine learning algorithms on this data, to find presence of heart related disease.
- Clearly accurate diagnosis requires huge amount of complex data which is too big to analyze using traditional techniques. Our application will have a good interface where both doctor and patient can interact with the system, and input all the health parameters of the patient and system in the backend using Random Forest algorithm will predict whether the patient has heart related issue or not.

Implementation and Testing-

This system is implemented using python programming language a simple Graphic User Interface(GUI) is created using the Tkinter library of python. In this system user inputs various heart parameters in the required textboxes, which is given to the machine learning algorithm (Random Forest Classifier) and that machine learning algorithm predicts the output value based on the parameters.

The testing of accuracy of the model is done using the accuracy_score method of scikit-learn library. True positive, True Negative, False Positive and False Negative of the model can be obtained by the confusion_matrix. In this way accuracy_score returns the accuracy of the model, which is displayed on the screen as a note for precautionary purposes.

Regularity (Attendance + Weekly Progress Reporting) – (To be filled by Teacher)

Awareness /Consideration of -Environment/ Social /Ethics/ Safety measures/Legal aspects-(To be filled by teacher)
Contest Participation/ publication-(To be filled by teacher)
Report –(To be filled by Teacher)

6. Project Monitoring/ Progress Information Sheets (one sheet per week, 3-4 weeks sheet should be added)

Week 1
Date: 10/02/2022
Current Work phase of project-   **Domain Selection**   • A task has been given by the teacher to select different domains on the basis of the interest of the students.   • The teacher instructed to select any one domain which students want to work upon their PBL project.   • And also instructed to prepare a presentation based on the domain selection.
Discussions Held   • Our team leader has taken the First inter group discussion meeting based on the task given by the teacher.   • We discussed lots of domain topics such as machine learning, artificial intelligence and data science, computer vision, IOT (Internet of Things), Block-chain-Technology, etc.   • We found that our interest is towards the domain of Data science & machine learning.   • Data science and machine learning gives knowledge and insights obtained from the data & used to solve various complex business problems.   • Hence, we decided to prepare our PBL project on the domain of Machine learning and data science.   • Each group has given his task to work upon the presentation.   • Lastly, we prepared a presentation based on machine learning and data science domain and taken the rehearsal of the presentation one day prior.
Progress till Date   • We have to present the Presentation of Domain selection on 11th February.   • All group members have explained the topics perfectly as decided in the meeting.   • Finally, we successfully discussed the domain selection till the date of task completion.
Remark   **Sign of Mentor**

Week 2

Date: 14/02/2022

Current Work phase of project-

Topic Selection:

- A task has been given by the teacher to select 4 different topics on the basis of the interest of the students.
- The teacher instructed to select any one topic which students want to work upon their PBL project.
- And also instructed to prepare a presentation based on the topic selection.

Discussions Held

- Our team leader has taken the second inter group discussion meeting based on the task given by the teacher.
- We discussed lots of topics, among those topics we selected the 4 main topics which are-
  - **Disease Prediction**
  - **Credit card fraud detection**
  - **Sales Recommendation**
  - **Movies Recommendation**
- We found that our interest is towards the topic of Heart Disease Prediction System.
- Disease prediction using machine learning is a system where diseases are predicted by the machine learning algorithm on the basis of the symptoms provided by the user.
- Hence, we decided to prepare our PBL project on the Heart Disease Prediction System
- Each group has given his task to work upon the presentation.
- Lastly, we prepared a presentation based on the Heart Disease Prediction system and taken the rehearsal of the presentation one day prior.

Progress till Date

- We have to present the Presentation of Topic selection on 15th February.
- All group members have explained the topics perfectly as decided in the meeting.
  Finally, we successfully discussed the Topic selection till the date of task completion, and selected the topic of **Heart Disease Prediction System.**

Remark

Sign of Mentor

Week 3

Date: 17/02/2022

Current Work phase of project-

Literature Survey:

- A task has been given by the teacher to take the Literature survey based on the Heart Disease Prediction System.
- The teacher instructed to visit IEEE website and The Google Scholar website to search and find out whether someone has implemented or worked upon such type of project.
- And also instructed to prepare a google sheet which should be edited by all the group members.

Discussions Held

- Our team leader has taken the third-inter group discussion meeting based on the task given by the teacher.
- A common google sheet has been created by the group leader and its access is given to all the group members.
- Then, a short discussion meet has been taken by the group leader to discuss the tasks which are distributed among all.
- The google sheet is divided into the four compartments, such as author/publisher name, problem statement, advantages and disadvantages.
- All are instructed to perform this task during the lecture hours itself.

Progress till Date

- We have to present the Presentation on the literature survey.
- All group members have explained the topics perfectly as decided in the meeting.
- Finally, we successfully discussed the entire literature survey and also communicated with the teacher.

Remark

Sign of Mentor

Week 4

Date: 21/02/2022

Current Work phase of project-

SRS (Software Requirement Specification)

- A task has been given by the teacher to create the Software Requirement Specifications based on the heart disease prediction system.
- The teacher instructed to describe the SRS among the following headings-

o **Section :1.0**: - Discuses the scope and the purpose of the software.
o **Section :2.0**: - Describe the overall functionalities and the constraints of the software and the user characteristics.
o **Section :3.0**: - Explains features, external, functional and non-functional requirements of the system.

- And also instructed to demonstrate/present the SRS to them.

Discussions Held

- Our team leader has taken the fourth inter group discussion meeting based on the task given by the teacher.
- According to the guidance of the team leader, the SRS task is divided among all the four group members.
- As the SRS contains lots of points which are interdependent on one another, the team has discussed and taken collaborative efforts for the creation of the SRS.
- The team has communicated all the functionalities, modules and other non-functional requirements and wherever necessary team has taken discussion review with the teacher.

Progress till Date

- We have to present the SRS (Software Requirement Specification) to the teacher till the date decided.
- All group members have explained the topics perfectly as decided in the meeting.
- Finally, we successfully discussed the entire SRS (Software Requirement Specification) and also communicated with the teacher.

Remark

Sign of Mentor

Week 5

Date: 24/03/2022

Current Work phase of project-

- A task has been given by the teacher to create synopsis of the project.
- Teacher has asked us to write the synopsis in the following format :-
- Section 1.0 :- Discuss the introduction of the software.

Discussions Held
1.First we wrote about the introduction of the project that is on heart disease prediction.
2.Then we did a literature survey on 3 papers and understood what type of algorithms we have to use in the project and on how to use it in order to predict heart diseases.
3.After that we started making an architectural diagram of the project .
4.After this we move on how it can be modified and on how to use it in future such as how we can incorporate other disease predictions in the system.
5.Finally we move on to the conclusion on how this system will help to predict the medical results efficiently. In this predicting system, we will provide a user-friendly interface that can be used by the users to detect whether their medical test results are positive or normal, i.e., it will detect the disease.
6.For the above task we used the following websites :-
- Google.com
- Geeksforgeeks.com
- Kaggle.com
- Tutorialpoint.com
- Javatpoint.com

7.After writing the synopsis we presented it in front of mam and then we were told some flaws about the architectural diagram.
8.Later mam gave us an insight on how to make project and how to improve our architectural diagram.
9.Discussed about the project and on how to make synopsis.

Progress till Date

1. Domain Selection
2. Topic Selection
3. Literature survey
4. SRS
5. Synopsis

Remark

Sign of Mentor

Week 6

Date: 21/04/2022

Current Work phase of project-

- A task has been given by the teacher to create abstract of the project.
- In the abstract we have to talk about the problem we are solving and about how we are incorporating many technologies in the project and about the aim of the project.
- And finally told us to present the abstract.

Discussions Held

1. Discussed about the project and on how to make abstract.
2. First, we wrote about the introduction of heart disease and how it impacts the society.
3. Then we wrote how different technologies will help in medical science.
4. Finally, we wrote about what is our aim via our project and how we will implement it using various algorithms.

Progress till Date

1. Domain Selection
2. Topic Selection
3. Literature survey
4. SRS
5. Synopsis
6. Abstract

Remark

Sign of Mentor

Week 7

Date: 06/05/2022

Current Work phase of project-

- Teacher gave us a task to make a report on our project.
- In the report we have to write the following details -
- What is the domain of the project.
- How we have done the literature survey.
- How we decided our topic.
- What is the topic.
- What we have done in the semester to complete our project sequentially.
- Finally we presented it in front of teacher.

Discussions Held

- First, we wrote everything about our domain, topic Selection, what we did in literature survey.
- Then we discussed about how we implemented our project in a sequential manner.
- Finally, we wrote everything in a single pdf.
- And then we presented it in front of teacher.

Progress till Date

1.Domain Selection
2.Topic Selection
3.Literature survey
4.SRS
5.Synopsis
6.Abstract
7.Project
8.Report

Remark

Sign of Mentor

Add Your Project Code and Output print screen. You have copy entire project code here and all output screens.

Backend Code: -

```python
import numpy as np #create numerical lists and arrays
import pandas as pd #creating datafraames fro m csv files
from sklearn.model_selection import train_test_split #splitting our datasets for training and testing
from sklearn.ensemble import RandomForestClassifier #ML algorithm
from sklearn.metrics import accuracy_score #testing our model accuracy
from sklearn.metrics import confusion_matrix#confusion matrix
import pickle

heartdataset1=pd.read_csv("D:\Learning only\PBL\data.csv")
heartdataset2=pd.read_csv("D:\Learning only\PBL\heart.csv")
heartdataset3=pd.read_csv("D:\Learning only\PBL\heart_cleveland_upload.csv")

heartdataset1.loc[ heartdataset1["target"] == 1, "target"] = 0
heartdataset1.loc[ heartdataset1["target"] == 2, "target"] = 1

heartdataset3.rename(columns = {'condition':'target'}, inplace = True)

vertical_concat = pd.concat([heartdataset1, heartdataset2], axis=0)
vertical_concat1=pd.concat([vertical_concat, heartdataset3], axis=0)

x=vertical_concat1.drop(columns='target',axis=1) #dropping column- axis=1 row- axis=0
y=vertical_concat1['target']

x_train,x_test,y_train,y_test=train_test_split(x,y,test_size=0.30,stratify=y,random_state=6)

print(x.shape,x_train.shape,x_test.shape)

model=RandomForestClassifier()

model.fit(x_train,y_train)

x_train_prediction=model.predict(x_train)
trainingdataaccuracy_x_train=accuracy_score(x_train_prediction,y_train)

print("Training data accuracy is ",trainingdataaccuracy_x_train)

x_test_prediction=model.predict(x_test)
testdataaccuracy=accuracy_score(x_test_prediction,y_test)

print("Testing data accuracy is ",testdataaccuracy)

# x_train_prediction=model.predict(x_train)
trainingdataaccuracy_x_train_conf_matrix=confusion_matrix(x_train_prediction,y_train)

print("Training data accuracy is ",trainingdataaccuracy_x_train_conf_matrix)
```

```
# x_test_prediction=model.predict(x_test)
testdataaccuracy_conf_matrix=confusion_matrix(x_test_prediction,y_test)

print("Testing data accuracy is ",testdataaccuracy_conf_matrix)

input_data=(43,0,0,132,341,1,0,136,1,3,1,0,3)
#change input data to a numpy array
input_data_numpy_arr=np.asarray(input_data)
#reshape the numpy array as we are predicting for only one instance
input_data_reshape=input_data_numpy_arr.reshape(1,-1)
print(input_data_reshape)
prediction=model.predict(input_data_reshape)
print(prediction)

if(prediction[0]==0):
   print("Heart disease not detected")
elif(prediction[0]==1):
   print("Heart disease detected")


with open('mlmodel','wb') as f:
   Model=pickle.dump(model,f)

with open('mlmodel','rb') as f:
   loaded_model=pickle.load(f)
```

Frontend Code: -

```
from email.header import Header
from tkinter import *
#from turtle import Screen
import numpy as np
import pandas as pd
from tkinter import messagebox
import pickle


# define functions


def inputlist():


   Show()
   disp = Tk()
```

```python
    disp.title("output screen")
    disp.geometry("932x532")
    disp.maxsize(932, 532)
    disp.minsize(932, 532)
    disp.config(bg="grey")

    age = Ageval1.get()
    gender0 = sexVal1.get()
    gender1=1
    if (gender0 == 'Female'):
        gender1 = 1
    elif (gender0 == 'Male'):
        gender1 = 0
    l1 = [ age, gender1, Cpval1.get(), trestbpsval1.get(), cholval1.get(), FBSval1.get(), restecgval1.get(),
thalachval1.get(), exangval1.get(), oldpeakval1.get(), slopeval1.get(), caval1.get(), thalval1.get()]

    # emptylabel = Label(disp, fg="green", bg="pink", padx=1232, pady=732).grid(row=0, column=0)
    print("Submitting form")
    print(
        f"{age, gender1, Cpval1.get(), trestbpsval1.get(), cholval1.get(), FBSval1.get(), restecgval1.get(),
thalachval1.get(), exangval1.get(), oldpeakval1.get(), slopeval1.get(), caval1.get(), thalval1.get()}")

    global pickle_prediction

    x= StringVar()

    with open('mlmodel', 'rb') as f:
        loaded_model = pickle.load(f)
    predictionarray = np.asarray(l1)

    input_data_reshape = predictionarray.reshape(1, -1)
    pickle_prediction = loaded_model.predict(input_data_reshape)
    for i in pickle_prediction:
        x=i
    print(x)
    if (x==1):
        finalheartprediction="You have heart disease"
    elif(x==0):
        finalheartprediction="You don't have heart disease"
    # l3=Entry(disp,textvariable=pickle_prediction_val).place(x=400,y=400)

    def PrintOutput():

        emptylable = Label(disp, fg="green", bg="grey", font="time 25 italic", padx=200, pady=40)
        emptylable.config(text=f"Result : {finalheartprediction}")
        emptylable.place(x=0,y=100)


    pdata=Button(disp,text="Result",font="Arial 20
bold",fg="white",bg="black",padx=20,pady=20,command=PrintOutput).place(x=360,y=300)
    note = Label(disp, text="Note : This model is 84.19% accurate",
            font="time 20 bold italic", fg="blue",bg="grey").place(x=200, y=400)
```

```python
def Show():

    age = Ageval1.get()
    gender0 = sexVal1.get()
    gender1=1
    if (gender0 == 'Female'):
        gender1 = 1
    elif (gender0 == 'Male'):
        gender1 = 0
    l1 = [ age, gender1, Cpval1.get(), trestbpsval1.get(), cholval1.get(), FBSval1.get(), restecgval1.get(),
thalachval1.get(), exangval1.get(), oldpeakval1.get(), slopeval1.get(), caval1.get(), thalval1.get()
]
    name=[nameval.get()]
    age1=[age]
    gender2=[gender1]
    cp=[Cpval1.get()]
    restbps=[trestbpsval1.get()]
    cholestrol=[cholval1.get()]
    fbs=[FBSval1.get()]
    restecg=[restecgval1.get()]
    thalach=[thalachval1.get()]
    exangval=[exangval1.get()]
    oldpeak=[oldpeakval1.get()]
    slope=[slopeval1.get()]
    ca=[caval1.get()]
    thalach=[thalachval1.get()]




    # with open("ganesh45.txt","a") as f:
        # f.write(f"name :{name}, age: {age}, sex :{ gender2}, cp: {cp}, trestbps: {restbps}, chol:
{cholestrol}, fbs:{ fbs}, restecg: {restecg},thalac: {thalach}, exang:{ exangval}, oldpeak:{ oldpeak},
slope: {slope}, ca: {ca}, thal: {thalach}\n")


dict={"name":name,"age":age,"sex":gender2,"cp":cp,"trestbps":restbps,"chol":cholestrol,"fbs":fbs
,"restecg":restecg,"thalach":thalach,"exang":exangval,"oldpeak":oldpeak,"slope":slope,"ca":ca,"th
al":thalach}
    df=pd.DataFrame(dict)
    # df.to_csv("D:\\Learning only\\patientdatabase.csv")
    with open("patientdatabase.csv",'a') as f:
        df.to_csv(f,mode='a',index=False,header=False)



    # emptylabel = Label(disp, fg="green", bg="pink", padx=1232, pady=732).grid(row=0, column=0)
    print("Submitting form")
    print(
        f"{ age, gender1, Cpval1.get(), trestbpsval1.get(), cholval1.get(), FBSval1.get(), restecgval1.get(),
thalachval1.get(), exangval1.get(), oldpeakval1.get(), slopeval1.get(), caval1.get(), thalval1.get()}")
```

```python
    global pickle_prediction
    pickle_prediction=StringVar()



    with open('mlmodel', 'rb') as f:
        loaded_model = pickle.load(f)
    predictionarray = np.asarray(l1)

    input_data_reshape = predictionarray.reshape(1, -1)
    pickle_prediction = loaded_model.predict(input_data_reshape)


    # filename=PhotoImage(file="heart icon.png")
    # backlable=Label(image=filename).place(x=0,y=0,height=332,weidth=200)


    # screen=Screen


def predict1():
    root.destroy()
    global pred
    pred = Tk()

    pred.config(bg="burlywood")

    pred.geometry("780x532")
    pred.maxsize(780, 532)
    pred.minsize(780, 532)

    name = Label(pred, text="Name", font="time 10 bold", fg="black", bg="burlywood").grid(row=0,
column=0, padx=14,
                                                               pady=14)
    Age = Label(pred, text="age", font="time 10 bold", fg="black", bg="burlywood").grid(row=0,
column=2, padx=14,
                                                               pady=14)
    height = Label(pred, text="Height", font="time 10 bold", fg="black",
bg="burlywood").grid(row=1, column=2, padx=14,
                                                               pady=14)
    weight = Label(pred, text="Weight", font="time 10 bold", fg="black",
bg="burlywood").grid(row=2, column=2, padx=14,
                                                               pady=14)
    sex = Label(pred, text="Gender", font="time 10 bold", fg="black", bg="burlywood").grid(row=1,
column=0, padx=14,
                                                               pady=14)
    cp_type = Label(pred, text="Chest pain type", font="time 10 bold", fg="black",
bg="burlywood").grid(row=2, column=0,
                                                               padx=14,
                                                               pady=14)
```

```python
    Resting_BP = Label(pred, text="trestbps", font="time 10 bold", fg="black",
bg="burlywood").grid(row=3,
                                                column=0,
                                                padx=14,
                                                pady=14)
    SC = Label(pred, text="Serum Cholesterol", font="time 10 bold", fg="black",
bg="burlywood").grid(row=3, column=2,
                                                padx=14, pady=14)
    FBS = Label(pred, text="Fasting Blood Sugar", font="time 10 bold", fg="black",
bg="burlywood").grid(row=5, column=0,
                                                padx=14,
                                                pady=14)
    RER = Label(pred, text="Resting ECG Results", font="time 10 bold", fg="black",
bg="burlywood").grid(row=5, column=2,
                                                padx=14,
                                                pady=14)
    MHR = Label(pred, text="thalach", font="time 10 bold", fg="black",
bg="burlywood").grid(row=7, column=0,
                                                padx=14, pady=14)
    EIA = Label(pred, text="exang", font="time 10 bold", fg="black", bg="burlywood").grid(row=7,
                                                column=2,
                                                padx=14,
                                                pady=14)
    ST_d = Label(pred, text="oldpeak", font="time 10 bold", fg="black",
bg="burlywood").grid(row=9, column=0,
                                                padx=14, pady=14)
    SPE = Label(pred, text="Slop", font="time 10 bold", fg="black", bg="burlywood").grid(row=9,
column=2, padx=14,
                                                pady=14)
    NMV = Label(pred, text="ca", font="time 10 bold", fg="black", bg="burlywood").grid(row=11,
                                                column=0,
                                                padx=14,
                                                pady=14)
    Thalassemia = Label(pred, text="Thalassemia", font="time 10 bold", fg="black",
bg="burlywood").grid(row=11,
                                                column=2,
                                                padx=14,
                                                pady=14)

    global Ageval1, sexVal1, Cpval1, trestbpsval1, cholval1, FBSval1, restecgval1, thalachval1,
exangval1, oldpeakval1, slopeval1, caval1, thalval1, nameval

    nameval = StringVar()
    heightval = StringVar()
    weightval = StringVar()
    Ageval1 = IntVar()
    sexVal1 = StringVar()
    Cpval1 = IntVar()
    trestbpsval1 = IntVar()
    cholval1 = IntVar()
    FBSval1 = IntVar()
    restecgval1 =IntVar()
```

```python
    thalachval1 = IntVar()
    exangval1 = IntVar()
    oldpeakval1 = IntVar()
    slopeval1 = IntVar()
    caval1 = IntVar()
    thalval1 = IntVar()
    sextype = {"Male ", "Female"}
    ageentry = Entry(pred, textvariable=nameval).grid(row=0, column=1, padx=14, pady=14)
    entry1 = Entry(pred, textvariable=Ageval1).grid(row=0, column=3, padx=14, pady=14)
    entry2 = Radiobutton(pred, text="Female", variable=sexVal1, value="Female", font="time 8 bold", fg="black",
                bg="burlywood").grid(row=1, column=1, sticky=W, padx=10, pady=10)
    entry21 = Radiobutton(pred, text="Male", variable=sexVal1, value="Male", font="time 8 bold", fg="black",
                bg="burlywood").grid(row=1, column=1, padx=20, sticky=E, pady=10)
    heightentry = Entry(pred, textvariable=heightval, font="time 8 bold").grid(row=1, column=3, padx=14, pady=14)
    entry3 = Entry(pred, textvariable=Cpval1).grid(row=2, column=1, padx=14, pady=14)
    weightentry = Entry(pred, textvariable=weightval).grid(row=2, column=3, padx=14, pady=14)
    entry4 = Entry(pred, textvariable=trestbpsval1).grid(row=3, column=1, padx=14, pady=14)
    entry5 = Entry(pred, textvariable=cholval1).grid(row=3, column=3, padx=14, pady=14)
    entry6 = Entry(pred, textvariable=FBSval1).grid(row=5, column=1, padx=14, pady=14)
    entry7 = Entry(pred, textvariable=restecgval1).grid(row=5, column=3, padx=14, pady=14)
    entry8 = Entry(pred, textvariable=thalachval1).grid(row=7, column=1, padx=14, pady=14)
    entry9 = Entry(pred, textvariable=exangval1).grid(row=7, column=3, padx=14, pady=14)
    entry10 = Entry(pred, textvariable=oldpeakval1).grid(row=9, column=1, padx=14, pady=14)
    entry11 = Entry(pred, textvariable=slopeval1).grid(row=9, column=3, padx=14, pady=14)
    entry12 = Entry(pred, textvariable=caval1).grid(row=11, column=1, padx=14, pady=14)
    entry13 = Entry(pred, textvariable=thalval1).grid(row=11, column=3, padx=14, pady=14)

    # sbutton = Button(pred, text="Submit", font="Arial 20 bold", fg="red", bg="black", padx=20, pady=20,command=Show).grid(row=13, column=2, sticky=E, padx=20, pady=20)
    show = Button(pred, text="Predict", font="time 20 bold", padx=60, pady=20, fg="red", bg="black", command=inputlist).place(x=250,y=420)

    # scvalue = StringVar()
    # scvalue.set("")
    # screen = Entry(pred, textvar=scvalue, font="time 15 bold",borderwidth=10).place(x=600, y=200, width=600, height=50)
    # pre1 = Button(pred, text="Predict", font="time 20 bold ", borderwidth=6, padx=20, pady=20,bg="green",fg="black",command=getvals).place(x=800, y=300)
    # pred.mainloop()
    # f=Frame(pred,bg="white").grid(row=0,column=2)
    # emptylabe=Label(f,fg="green",font="Arial 20 bold",bg="yellow",padx=400,pady=20).grid(row=1,column=2)


root = Tk()

# define  window size
root.geometry("1232x732")
root.maxsize(1232, 732)
```

```python
root.minsize(1232, 732)

root.config(bg="olivedrab")
col = Label(
    text=" DR. D. Y. PATIL INSTITUTE OF TECHNOLOGY, PIMPRI, PUNE\n SAVITRIBAI
PHULE PUNE UNIVERSITY 2021 - 2022",
    fg="black", bg="olivedrab", font="Arial 15 bold").grid(row=0, column=1, pady=30)
branch = Label(text="DEPARTMENT OF ARTIFICIAL INTELLIGENCE AND DATA SCIENCE",
fg="black", bg="olivedrab",
            font="Arial 20 bold").grid(row=1, column=1)
root.title("Heart Disease prediction model")

# labeling

l1 = Label(root, text="Welcome to disease prediction model", fg="yellow", bg="olivedrab",
font="Arial 40 bold").grid(
    row=2, column=1, padx=160, pady=50)

b2 = Button(root, text="User registration", font="time 20 bold", fg="red", bg="black", padx=20,
pady=20,
        command=predict1, borderwidth=13).grid(row=3, column=1, pady=50)

members = Label(root,
        text=" Members:                    \n  Prathamesh patil\n  Ganesh jagzap   \n  Jay sawant
\n  Atharv wankhede",
        fg="black", bg="olivedrab", font="Arial 20 ").grid(row=4, column=1, sticky=E, padx=200)
mentor = Label(root, text="Mentor : Mrs .Shubhangi Vairagar ", fg="black", bg="olivedrab",
font="Arial 20").grid(row=4,
                                        column=1,
                                        sticky=W,
                                        padx=200,
                                        pady=20)


root.mainloop()
```

OUTPUT SCREENS

tk — □ ×

Name	Rahul	age	74
Gender	○ Female ● Male	Height	180
Chest pain type	0	Weight	80
trestbps	120	Serum Cholesterol	269
Fasting Blood Sugar	0	Resting ECG Results	2
thalach	121	exang	1
oldpeak	0.2	Slop	1
ca	1	Thalassemia	3

Predict

output screen — □ ×

Result : You don't have heart disease

Result

Note : This model is 84.19% accurate

tk

Name	PQR	**age**	56
Gender	⊙ Female ○ Male	**Height**	160
Chest pain type	3	**Weight**	75
trestbps	130	**Serum Cholesterol**	256
Fasting Blood Sugar	1	**Resting ECG Results**	2
thalach	142	**exang**	1
oldpeak	1.2	**Slop**	2
ca	1	**Thalassemia**	7

Predict

output screen

Result : You have heart disease

Result

Note : This model is 84.19% accurate