



PRESIDENCY COLLEGE

(AUTONOMOUS)

AFFILIATED TO BENGALURU CITY UNIVERSITY, APPROVED BY AICTE, DELHI & RECOGNISED BY THE GOVT. OF KARNATAKA
RE-ACCREDITED BY NAAC WITH 'A+' GRADE

Department of Computer Applications

MCA – A SECTION (Batch- 2021-2023)

Session: May 2023 to August 2023

PROJECT WORK – SYNOPSIS

**TITLE OF THE PROJECT : DiagnoseMe : A Predictive Health
Analysis for Early Disease Detection**

STUDENT NAME & SECTION : Nakul S Anand T - MCA-A

REGISTRATION NO. : 21P01068

GUIDE NAME : Mr. N Karthik

OBJECTIVE:

The primary objective of this project is to build a disease prediction model using machine learning algorithms. The model will be trained on a dataset containing information on the symptoms and medical histories of patients with multiple diseases. The project aims to develop a reliable and accurate model that can predict the occurrence of multiple diseases based on the symptoms reported by patients.

INTRODUCTION:

The field of medical diagnosis and treatment has been revolutionized by machine learning and artificial intelligence. Using large datasets and sophisticated algorithms, it is now possible to predict the likelihood of a patient having a particular disease based on their symptoms and other clinical data. This project aims to use machine learning algorithms to predict the occurrence of multiple diseases based on the symptoms of patients.

MAJOR MODULES:

- Module 1 - Data Acquisition
- Module 2 - Exploratory Data analysis
- Module 3 - Data Preprocessing and cleaning



- Module 4 -Implementation of Machine learning algorithm
- Module 5 - Valuation of the model
- Module 6 - Deployment

Functionalities of Modules:

Data Acquisition:

- Acquired dataset from Machine Learning Repository
- Making sure that the dataset collected is reliable and trustworthy.
- Storage of the collected dataset will be done in Sqlite3 for a Secure and efficient way of storing.

Exploratory Data Analysis:

- Analyzing the dataset using statistical techniques to identify hidden patterns.
- Finding relationships between features.
- Feature selection is done accordingly.
- Visualization will be done to showcase the relationship between features.

Data preprocessing and cleaning:

- Collected raw dataset will be preprocessed and cleansed before analyzing.
- Finding outliers and handling them will be the top priority.
- Filling missing values and handling, Reformatting will be done for a better dataset analysis.
- Data transformation will be done for a better model valuation.

Implementation of Machine learning algorithm:

- Will implement machine learning algorithms accordingly.
- Different machine learning algorithms will be tested to find the best performing algorithm.

Valuation of the model:

- Valuation of the model will be done using different accuracy metrics measurements to check the model's performance and reliability.

Deployment:

- The model will be deployed as a Web Application using streamlit/flask.
- Then the End user can predict whether they have a certain disease in an easy, free of cost and efficient manner.



PRESIDENCY COLLEGE

(AUTONOMOUS)

AFFILIATED TO BENGALURU CITY UNIVERSITY, APPROVED BY AICTE, DELHI & RECOGNISED BY THE GOVT. OF KARNATAKA
RE-ACCREDITED BY NAAC WITH 'A+' GRADE

HARDWARE AND SOFTWARE REQUIREMENTS:

HARDWARE	
PROCESSOR	Intel Pentium P6200
RAM	2 GB
HDD	256 GB
SOFTWARE	
CLIENT SIDE TECHNOLOGIES(FRONTEND)	Streamlit/Flask (Python)
SERVER SIDE TECHNOLOGIES(BACKEND)	Python
IDE (IF ANY)	Google Colab, Spyder
CONNECTION TECHNOLOGY (DRIVER SOFTWARE OR ANY INTERFACE)	Sqlite3
OPERATING SYSTEM	Windows 8+
TOOLS(FRONT END AND BACK END)	Python, Sqlite3, Git

Gantt Chart:

SRS /SRA	System Design	Coding	Testing	Deployment and Report
27.05.2023	24.06.2023	22.07.2023	29.07.2023	18.08.2023

NOTE: Project hard copy submission must be done on or before 26.08.2023.

Guide Signature

Student Signature