

A MAJOR PROJECT REPORT
ON
**MULTI DISEASE PREDICTION USING MACHINE
LEARNING**

A dissertation submitted in partial fulfilment of the
Requirements for the award of the degree of

BACHELOR OF TECHNOLOGY

in

INFORMATION TECHNOLOGY

Submitted by

T.Arya (19B81A12C2)

K.Manasvini (19B81A12D4)

Ch.Mukesh(19B81A12E2)

Under the esteemed guidance of

Mrs.Swathi Agarwal

Associate Professor, IT Department



DEPARTMENT OF INFORMATION TECHNOLOGY

CVR COLLEGE OF ENGINEERING

ACCREDITED BY NBA, AICTE & Affiliated to JNTU-H Vastunagar, Mangalpally

(V), Ibrahimpatnam (M), R.R. District, PIN-501 510

2022-2023



Cherabuddi Education Society's
CVR COLLEGE OF ENGINEERING

(An Autonomous Institution)

ACCREDITED BY NATIONAL BOARD OF ACCREDITATION, AICTE

(Approved by AICTE & Govt. of Telangana and Affiliated to JNT University)

Vastunagar, Mangalpalli (V), Ibrahimpatan (M), R.R. District, PIN - 501 510

Web : <http://cvr.ac.in>, email : info@cvr.ac.in

Ph : 08414 - 252222, 252369, Office Telefax : 252396, Principal : 252396 (O)

DEPARTMENT OF INFORMATION TECHNOLOGY

CERTIFICATE

This is to certify that the Major Project Report entitled “**Multi Disease Prediction Using Machine Learning** ” is a bonafide work done and submitted **T.Arya(19B81A12C2)** , **K.Manasvini(19B81A12D4)**, **Ch.Mukesh(19B81A12E2)** during the academic year 2022-2023, in partial fulfilment of requirement for the award of Bachelor of Technology degree in Information Technology from Jawaharlal Nehru Technological University Hyderabad, is a bonafide record of work carried out by them under my guidance and supervision.

Certified further that to the best of my knowledge, the work in this dissertation has not been submitted to any other institution for the award of any degree or diploma.

INTERNAL GUIDE

Mrs. Swathi Agarwal

Associate Professor, IT Department

HEAD OF THE DEPARTMENT

Dr. Bipin Bihari Jayasingh

HOD, IT Department

MAJOR-PROJECT COORDINATOR

Dr. Rakesh Kumar Godi

Professor, IT Department

EXTERNAL EXAMINER

ACKNOWLEDGEMENT

The satisfaction of completing this major project would be incomplete without mentioning our gratitude towards all the people who have supported us. Constant guidance and encouragement have been instrumental in the completion of this project.

We offer our sincere gratitude to our internal guide **Mrs. Swathi Agarwal**, Associate Professor, IT Department, CVR College of Engineering for her immense support, timely co-operation and valuable advice throughout the course of our project work.

We are thankful to **Dr. Rakesh Kumar Godi**, Major Project coordinator, Professor, IT Department, CVR College of Engineering for his supportive guidelines and for having provided the necessary help for carrying forward this project without any obstacles and hindrances.

We would like to thank the Head of Department, Professor **Dr. Bipin Bihari Jayasingh**, for his meticulous care and cooperation throughout the project work.

We would like to thank Chairman, Principal and Vice Principal for availing infrastructural facilities to complete the project work in time.

We also thank the **Project Review Committee Members** for their valuable suggestions.

DECLARATION

We hereby declare that the project report entitled “**Multi Disease Prediction Using Machine Learning**” is an original work done and submitted to IT Department, CVR College of Engineering, affiliated to Jawaharlal Nehru Technological University Hyderabad, Hyderabad in partial fulfilment of the requirement for the award of Bachelor of Technology in **Information Technology** and it is a record of bonafide project work carried out by us under the guidance of **Dr.Rakesh Kumar Godi ,Professor, Department of Information Technology.**

We further declare that the work reported in this project has not been submitted, either in part or in full, for the award of any other degree or diploma in this institute or any other Institute or University.

Signature of the Student

(T. Arya)

(19B81A12C2)

Signature of the Student

(K. Manasvini)

(19B81A12D4)

Signature of the Student

(Ch. Mukesh)

(19B81A12E2)

ABSTRACT

Many models for health care analysis are concentrating on one disease per analysis. Like one analysis is for diabetes analysis, one for cancer analysis, one for skin diseases and so on. There is no common system where one analysis can perform more than one disease prediction. In this model we propose a system which is used to predict multiple diseases by using Flask API. In this model we perform Diabetes analysis, heart disease and breast cancer analysis. We try to implement multiple disease analysis used machine learning algorithms, TensorFlow and Flask API. Python pickling is used to save the model behavior and python unpickling is used to load the pickle file whenever required. The importance of this analysis is that while analyzing the diseases all the parameters which causes the disease is included so it possible to detect the maximum effects which the disease will cause. For example, for diabetes few parameters like age, sex, BMI , insulin, glucose, blood pressure, diabetes pedigree function, pregnancies, considered in addition to age, sex, BMI, insulin, glucose, blood pressure, diabetes pedigree function, pregnancies included serum creatinine, potassium, Glasgow Coma Scale, heart rate/pulse Rate, respiration rate, body temperature, low density lipoprotein (LDL), high density lipoprotein (HDL), TG (Triglycerides). Model's behavior will be saved as python pickle file. Flask API will invoke the corresponding model and returns the status of the patient. The importance of this analysis to analyze the maximum diseases, so that to monitor the patient's condition and warn the patients in advance to decrease mortality ratio.

TABLE OF FIGURES

S. No.	Name	Pg. No.
1.	Flow Chart.....	20
2.	System Architecture.....	22
3.	Use Case Diagram.....	23
4.	Class Diagram.....	24
5.	Sequence Diagram.....	25
6.	Activity Diagram.....	26
7.	Component Diagram.....	27
8.	Deployment Diagram.....	28

TABLE OF CONTENTS

S.No.	Name	Pg. No.
1.	Introduction.....	8
	1.1 Motivation.....	9
	1.2 Problem Statement.....	10
2.	Literature Review.....	11
3.	Software And Hardware Specifications.....	14
	3.1 System Requirements.....	14
	3.2 Functional Requirements.....	15
	3.3 Non-Functional Requirements.....	15
	3.4 Hardware Requirements.....	17
	3.5 Software Requirements.....	17
	3.6 Software Description.....	18
4.	Flowchart.....	20
5.	Design.....	21
	5.1 System Architecture.....	22
	5.2 Use Case Diagram.....	23
	5.3 Class Diagram.....	24
	5.4 Sequence Diagram.....	25
	5.5 Activity Diagram.....	26
	5.6 Component Diagram.....	27
	5.7 Deployment Diagram.....	28

CHAPTER – 1

INTRODUCTION

Multiple Disease Prediction using Machine Learning is a system which predicts the disease based on the information provided by the user. It also predicts the disease of the patient or the user based on the information or the symptoms he/she enter into the system and provides the accurate results based on that information.

If the patient is not much serious and the user just wants to know the type of disease, he/she has been through. It is a system which provides the user the tips and tricks to maintain the health system of the user and it provides a way to find out the disease using this prediction.

Now-a-days health industry plays major role in curing the diseases of the patients so this is also some kind of help for the health industry to tell the user and also it is useful for the user in case he/she doesn't want to go to the hospital or any other clinics, so just by entering the symptoms and all other useful information the user can get to know the disease he/she is suffering from and the health industry can also get benefit from this system by just asking the symptoms from the user and entering in the system and in just few seconds they can tell the exact and upto some extent the accurate diseases. This DPUML is previously done by many other organizations but our intention is to make it different and beneficial for the users who are using this system.

This Multiple Disease Prediction Using Machine Learning is completely done with the help of Machine Learning and Python Programming language with Tkinter Interface for it and also using the dataset that is available previously by the hospitals using that we will predict the disease.

Now a day's doctors are adopting many scientific technologies and methodology for both identification and diagnosing not only common disease, but also many fatal diseases. The successful treatment is always attributed by right and accurate diagnosis. Doctors may sometimes fail to take accurate decisions while diagnosing the disease of a patient, therefore Multiple Disease Prediction systems which use machine learning algorithms assist in such cases to get accurate results.

The project Multiple Disease Prediction using machine learning is developed to overcome general disease in earlier stages as we all know in competitive environment of economic development the mankind has involved so much that he/she is not concerned about health according to research there are 40% peoples how ignores about general disease which leads to harmful disease later.

The main reason of ignorance is laziness to consult a doctor and time concern the peoples have involved themselves so much that they have no time to take an appointment and consult the doctor which later results into fatal disease. According to research there are 70% peoples in India suffers from general disease and 25% of peoples face death due to early ignorance the main motive to develop this project is that a user can sit at their convenient place and have a check-up of their health the UI is designed in such a simple way that everyone can easily operate on it and can have a check-up.

1.1 MOTIVATION:

The purpose of making this project called “Multiple Disease Prediction Using Machine Learning” is to predict the accurate disease of the patient using all their general information’s and also the symptoms. Using this information, there we will compare with our previous datasets of the patients and predicts the disease of the patient he/she is been through. If this Prediction is done at the early stages of the disease with the help of this project and all other necessary measure the disease can be cured and in general this prediction system can also be very useful in health industry. If health industry adopts this project, then the work of the doctors can be reduced and they can easily predict the disease of the patient. The general purpose of this Multiple Disease Prediction is to provide prediction for the various and generally occurring diseases that when unchecked and sometimes ignored can turns into fatal disease and cause lot of problem to the patient and as well as their family members. This system will predict the most possible disease based on the symptoms. The health industry in information yet and knowledge poor and this industry is very vast industry which has lot of work to be done. So, with the help of all those algorithms, techniques and methodologies we have done this project which will help the peoples who are in the need.

1.2 PROBLEM STATEMENT:

Now-a-days in Health Industry there are various problems related to machines or devices which will give wrong or unaccepted results,so to avoid those results and get the correct and desired results we are building a program or project which will give the accurate predictions based on information provided by the user and also based on the datasets that are available in that machine.The health industry in information yet and knowledge poor and this industry is very vast industry which has lot of work to be done. So, with the help of all those algorithms, techniques and methodologies we have done this project which will help the peoples who are in the need.So the problem here is that many people goes to hospitals or clinic to know how is their health and how much they are improving in the given days, but they have to travel to get to know there answers and sometimes the patients may or may not get the results based on various factors such as doctor might be on leave or some whether problem so he might not have come to the hospital and many more reasons will be there so to avoid all those reasons and confusion we are making a project which will help all those person's and all the patients who are in need to know the condition of their health, and at sometimes if the person has been observing few symptoms and he/she is not sure about the disease he/she is encountered with so this will lead to various diseases in future.So,to avoid that and get to know the disease in early stages of the symptoms this Multiple Disease Prediction will help a lot to the various people's ranging from children to teenagers to adults and also the senior citizens.

CHAPTER – 2

LITERATURE SURVEY

Classification and Diagnosis of Diabetes: Standards of Medical Care in Diabetes—2018 American Diabetes Association Diabetes Care 2018; 41(Supplement 1): S13–S27. <https://doi.org/10.2337/dc18-S002>.

In this paper, we have proposed a diabetes prediction model for better classification of diabetes which includes few external factors responsible for diabetes along with regular factors like Glucose, BMI, Age, Insulin, etc.

Chaitrali S. Dangare, Sulabha S. Apte, "Improved Study of Heart Disease Prediction System using Data Mining Classification Techniques", International Journal of Computer Applications (0975 888)Volume 47No.10, June 2012.

This paper has analysed prediction systems for Heart disease using a greater number of input attributes. The system uses medical terms such as sex, blood pressure, cholesterol like 13 attributes to predict the likelihood of patient getting a heart disease.

M.Mehdy, E.E.Shair and P.Y.Ng, "Decision Tree algorithm for Earlier Detection of Breast Cancer", Hindawi, Computational and Mathematical Methods in Medicine, Volume 2017, Article ID 2610628.

The authors used a 10-fold cross-validation technique from the UCI machine learning repository named Wisconsin breast cancer diagnosis dataset, the techniques achieved the accuracy of 96% for Decision Tree algorithm.

DhfarHamed, Jwan K. Alwan, Mohamed Ibrahim, Mohammad B. Naeem "The Utilisation of Machine Learning Approaches for Medical Data Classification" in Annual Conference on New Trends in Information & Communications Technology Applications - march2017 A. J. Jenkins, M. V. Joglekar, A. A. Hardikar, A. C. Keech, D. N. O'Neal, and A. S. Januszewski, "Biomarkers in diabetic retinopathy," The Review of Diabetic Studies, vol. 12, no. 1-2, pp. 159–195, 2015.

In Machine Learning there are various types of algorithms such as Regression, Linear Regression, Logistic Regression, Naive Bayes Classifier, Bayes theorem, KNN (K-Nearest Neighbour Classifier), Decision Tress, Entropy, ID3, SVM (Support Vector Machines), K-means Algorithm, Random Forest and etc.,

Machine learning explores the study and construction of algorithms that can learn from and make predictions on data Machine learning is closely related to computational statistics, which also focuses on prediction-making through the use of computers. It has strong ties to mathematical optimization, which delivers methods, theory and application domains to the field.

Diez Roux AV, Merkin SS, Arnett D, Chambless L, Massing M, et al. 2019. Neighbourhood of residence and incidence of coronary heart disease. *N. Engl. J. Med.* 345:99–106 The first study to report that neighbourhood socioeconomic disadvantage was independently associated with the incidence of coronary heart disease.

A variety of characteristics of neighbourhoods, including the availability of resources and services to promote or maintain healthy lifestyles as well as the physical and social environment, may be related to cardiovascular risk. Although studies have suggested that neighbourhood characteristics may be related to the prevalence of, risk factors for, and mortality due to coronary heart disease, 8,9,13-15 the extent to which neighbourhood characteristics are related to the incidence of coronary heart disease has not been established.

T hota R.N.Acharya S.H.Garg M.L.Curcumin and/or omega -3 polyunsaturated fatty acids supplementation reduces insulin resistance and blood lipids in individuals with high risk of type 2 diabetes: a randomised controlled trial.Lipids Health Dis.2019;18: 31

Lowering insulin resistance and dyslipidaemia may not only enhance glycaemic control but also preserve the β -cell function, reducing the overall risk of developing type 2 diabetes (T2D). The current study was aimed to evaluate the effects of curcumin and/or long-chain omega-3 polyunsaturated fatty acids (LCn-3PUFA) supplementation on glycaemic control and blood lipid levels in individuals at high risk of developing T2D.

Cosentino F, Grant PJ, Aboyans V, Bailey CJ, Ceriello A, Delgado V, et al. ESC Scientific Document Group. 2019 ESC Guidelines on diabetes, pre-diabetes, and cardiovascular diseases developed in collaboration with the EASD. Eur Heart J 2020 41:255 -323. doi: 10.1093/eurheartj/ehz486.

The Emerging Risk Factor Collaboration, a meta-analysis of 102 prospective studies, showed that DM in general (data on DM type were unavailable) confers a two-fold excess risk of vascular outcomes (coronary heart disease, ischaemic stroke, and vascular deaths), independent of other risk factors. The excess relative risk of vascular events with DM was greater in women and at younger ages.

Vogel, P., Klooster, T., Andrikopoulos, V., & Lungu, M. 2017, September. A low-effort analytics platform for visualizing evolving Flask-based Python web services. In 2017 IEEE Working Conference on Software Visualization (VISSOFT) (pp. 109-113). IEEE

Within the Python community, Flask² is a very popular web framework³. It provides simplicity and flexibility by implementing a bare-minimum web server, and thus advertises as a micro-framework. Despite their popularity, to the best of our knowledge, there is no simple solution for monitoring the evolving performance of Flask web applications. Thus, every one of the developers of these projects faces one of the following options when confronted with the need of gathering insight into the runtime behaviour of their implemented services.

CHAPTER – 3

SOTWARE AND HARDWARE SPECIFICATIONS

3.1 SYSTEM REQUIREMENTS:

A requirement is a feature that the system must have or a constraint that it must be accepted by the client. Requirement Engineering aims at defining the wants of the system under construction. Requirement Engineering include two main activities requirement elicitation which results in the specification of the system that the client understands and analysis which in analysis model that the developer can unambiguously interpret. A requirement may be a statement about what the proposed system will do.

Requirements can be divided into two major categories:

- Functional Requirements.
- Non-Functional Requirements.

3.2 FUNCTIONAL REQUIREMENTS:

A Functional requirement defines a function of a system or its component. A function is described as a set of inputs, the behaviour, and outputs. Functional requirements may be calculations, technical details, data manipulation and processing and other specific functionality that define what a system is supposed to accomplish. Behavioural requirements describing all cases where the system uses the functional requirements are captured in use cases. Functional requirements are supported by non-functional requirements (also known as quality requirements), which impose constraints on the design or implementation (such as performance requirements, security, or reliability).

As defined in requirements engineering, functional requirements specify particular results of a system. This should be contrasted with non-functional requirements which specify overall characteristics such as cost and reliability. Functional requirements drive the application architecture of a system, while non-functional requirements drive the technical architecture of a system.

- Functional Requirements concerns with the specific functions delivered by the system. So, Functional requirements are statements of the services that the system must provide.
- The functional requirements of the system should be both complete and consistent.
- Completeness means that all the services required by the user should be defined.
- Consistency means that requirements should not have any contradictory definitions.
- The requirements are usually described in a fairly abstract way. However, functional system requirements describe the system function in details, its inputs and outputs, exceptions and soon.
- Take user id and password match it with corresponding file entries. If a match is found then continue else raise an error message.

3.3 NON-FUNCTIONAL REQUIREMENTS:

- Non-functional Requirements refer to the constraints or restrictions on the system. They may relate to emergent system properties such as reliability, response time and store occupancy or the selection of language, platform, implementation techniques and tools.
- The non-functional requirements can be built on the basis of needs of the user, budget constraints, organization policies and etc.

1. Performance requirement:

All data entered shall be up to mark and no flaws shall be there for the performance to be 100%.

2. Platform constraints:

The main target is to generate an intelligent system to predict the adult height.

3. Accuracy and Precision:

Requirements are accuracy and precision of the data

4. Modifiability:

Requirements about the effort required to make changes in the software. Often, the measurement is personnel effort (person-months).

5. Portability:

Since mobile phone is handy so it is portable and can be carried and used whenever required.

6. Reliability:

Requirements about how often the software fails. The definition of a failure must be clear. Also, don't confuse liability with availability which is quite a different kind of requirement. Be sure to specify the consequences of software failure, how to protect from failure, a strategy for error Prediction, and a strategy for correction.

7. Security:

One or more requirements about protection of your system and its data.

8. Usability:

Requirements about how difficult it will be to learn and operate the system. The requirements are often expressed in learning time or similar metrics.

ACCESSIBILITY: Accessibility is a general term used to describe the degree to which a product, device, service, or environment is accessible by as many people as possible. In our project people who have registered with the cloud can access the cloud to store and retrieve their data with the help of a secret key sent to their email ids. User interface is simple and efficient and easy to use.

MAINTAINABILITY: In software engineering, maintainability is the ease with which a software product can be modified in order to include new functionalities can be added in the project based on the user requirements just by adding the appropriate files to existing project using net and programming languages. Since the programming is very simple, it is easier to find and correct the defects and to make the changes in the project.

SCALABILITY: System is capable of handling increase total throughput under an increased load when resources (typically hardware) are added. System can work normally under situations such as low bandwidth and large number of users.

PORTABILITY: Portability is one of the key concepts of high-level programming. Portability is the software code base feature to be able to reuse the existing code instead of creating new code when moving software from an environment to another. Project can be executed under different operation conditions provided it meets its minimum configurations. Only system files and dependant assemblies would have to be configured in such case.

VALIDATION: It is the process of checking that a software system meets specifications and that it fulfils its intended purpose. It may also be referred to as software quality control. It is normally the responsibility of software testers as part of the software development lifecycle. Software validation checks that the software product satisfies or fits the intended use (high-level checking), i.e., the software meets the user requirements, not as specification artefacts or as needs of those who will operate the software only; but, as the needs of all the stakeholders.

3.4 HARDWARE REQUIREMENTS:

- System :Pentium4, Intel Core i3,i5,i7 and 2GHz min
- RAM : 512Mb or above
- HardDisk : 10 GB or above

3.5 SOFTWARE REQUIREMENTS:

- Python 3.7
- scikit-learn
- Python Flask
- Windows 8 or Above Operating System
- Python libraries:
 - NumPy
 - Pandas

3.6 SOFTWARE DESCRIPTION :

PYTHON:

Python is a multi-paradigm programming language. Object-oriented programming and structured programming are fully supported, and many of its features support functional programming and aspect-oriented programming (including by metaprogramming and metaobjects. Many other paradigms are supported via extensions, including design by contract and logic programming. Python uses dynamic typing and a combination of reference counting and a cycle-detecting garbage collector for memory management. It also features dynamic name resolution (late binding), which binds method and variable names during program execution.

Python's developers strive to avoid premature optimization , and reject patches to non-critical parts of Python that would offer marginal increases in speed at the cost of clarity. When speed is important, a Python programmer can move time-critical functions to extension modules written in languages such as C, or use PyPy, a just-in-time compiler. Python is also available, which translates a Python script into C and makes direct C-level API calls into the Python interpreter. An important goal of Python's developers is keeping it fun to use. Python's design offers some support for functional programming in the Lisp tradition. It has filter, map, and reduce functions, list comprehensions, dictionaries, sets, and generator expressions. The standard library has two modules (itertools and functools) that implement functional tools borrowed from Haskell and Standard ML.

BENEFITS OF PYTHON :

- Presence of Third-Party Modules
- Extensive Support Libraries
- Open Source and Community Development
- Learning Ease and Support Available
- User-friendly Data Structures
- Productivity and Speed
- Highly Extensible and Easily Readable Language

TKINTER INTERFACE:

Tkinter is a Python binding to the Tk GUI toolkit. It is the standard Python interface to the Tk GUI toolkit and is Python's de facto standard GUI. Tkinter is included with standard Linux, Microsoft Windows and Mac OS X installs of Python. The name Tkinter comes from Tk interface. Tkinter was written by Fredrik Lundh. Tkinter is free software released under a Python license. As with most other modern Tk bindings, Tkinter is implemented as a Python wrapper around a complete Tool Command Language (TCL) interpreter embedded in the Python interpreter. Tkinter calls are translated into Tcl commands which are fed to this embedded interpreter, thus making it possible to mix Python and TCL in a single application. In Tkinter, the Frame widget is the basic unit of organization for complex layouts. A frame is a rectangular area that can contain other widgets. When any widget is created, a parent-child relationship is created. For example, if you place a text label inside a frame, the frame is the parent of the label. Python offers multiple options for developing GUI (Graphical User Interface). Out of all the GUI methods, tkinter is most commonly used method. It is a standard Python interface to the Tk GUI toolkit shipped with Python. Python with tkinter outputs the fastest and easiest way to create the GUI applications.

To create a tkinter:

- 1) Importing the module – tkinter.
- 2) Create the main window (container).
- 3) Add any number of widgets to the main window.
- 4) Apply the event Trigger on the widgets.
- 5) Importing tkinter is same as importing any other module in the python code.
- 6) Note that the name of the module in Python 2.x is 'Tkinter' and in Python 3.x is 'tkinter'.

CHAPTER 4

FLOWCHART

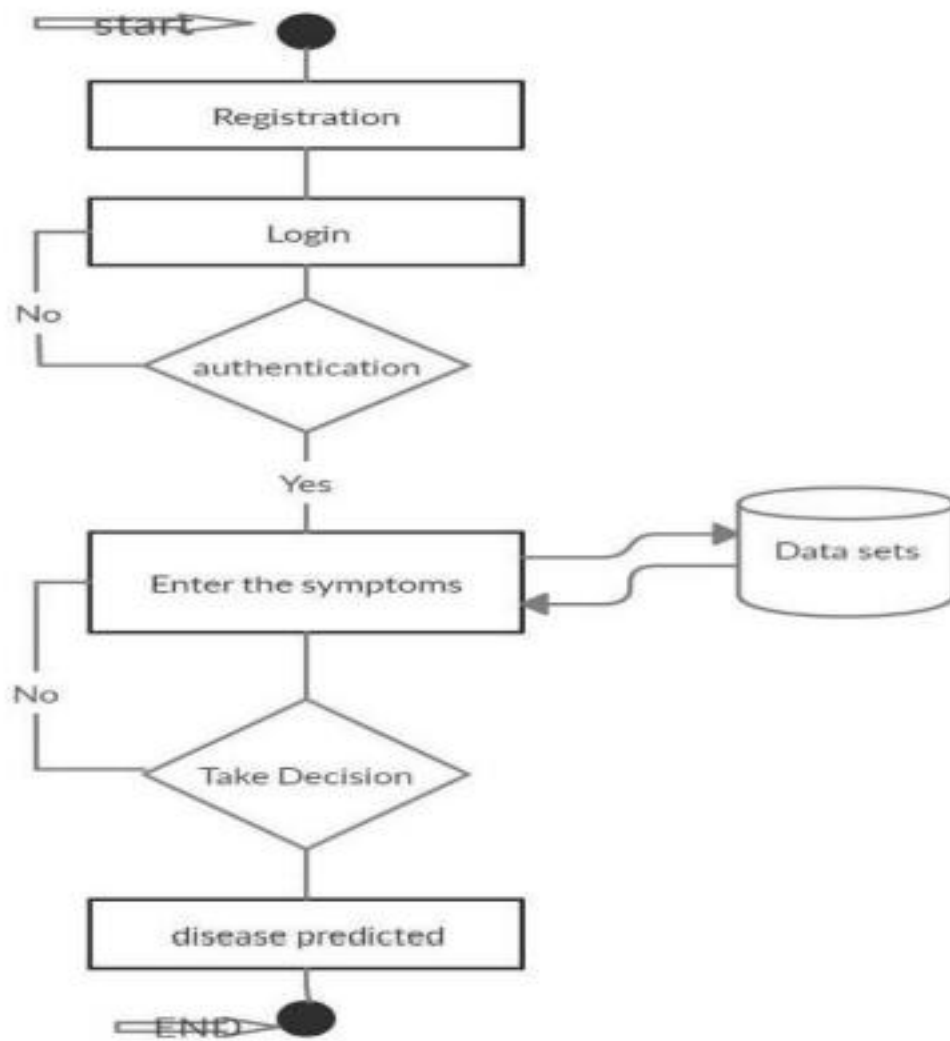


Fig 1: Flow chart

CHAPTER 5

DESIGN

The Design consist of various design which we have implemented in our system Multiple Disease Prediction using machine learning. This system has built with various designs such as data flow diagram, sequence diagram, class diagram, use case diagram, component diagram, activity diagram, state chart diagram, deployment diagram. After doing these various diagrams and based on these diagrams we have done our project. We have designed our system in such a way that whenever user login into the system, the user has to register to the system, and new user cannot use the system without registering in the system. After that for registration the user requires basic credentials such as username, age, email, phone, password. Then the user has to login to the system using the same username and password.

Here are the things that this system can perform.

- a. Entering Symptoms
- b. Disease Prediction

Entering Symptoms: Once user successfully logged in to the system then he/she has to select the symptoms from the given drop-down menu.

Disease prediction: The predictive model predicts the disease of a person he might have, based on the user entered symptoms.

5.1 SYSTEM ARCHITECTURE:

Multiple Disease Prediction using machine learning predicts the presence of the disease for the user based on various symptoms and the information the user gives such as sugar level, haemoglobin level and many more such general information through the symptoms. The architecture of the system Multiple Disease Prediction using machine learning consist of various datasets through which we will compare the symptoms of the user and predicts it, then the datasets are transformed into the smaller sets and from there it gets classified based on the classification algorithms lateron the classified data is then processed into the machine learning technologies through which the data gets processed and goes in to the Multiple Disease Prediction model using all the inputs from the user that is mentioned above. Then after user entering the above information and overall processed data combines and compares in the prediction model of the system and finally predicts the disease. An architecture diagram is a graphical representation of a set of concepts, that are part of an architecture, including their principles, elements and components. The diagram explains about the system software in perception of overview of the system.

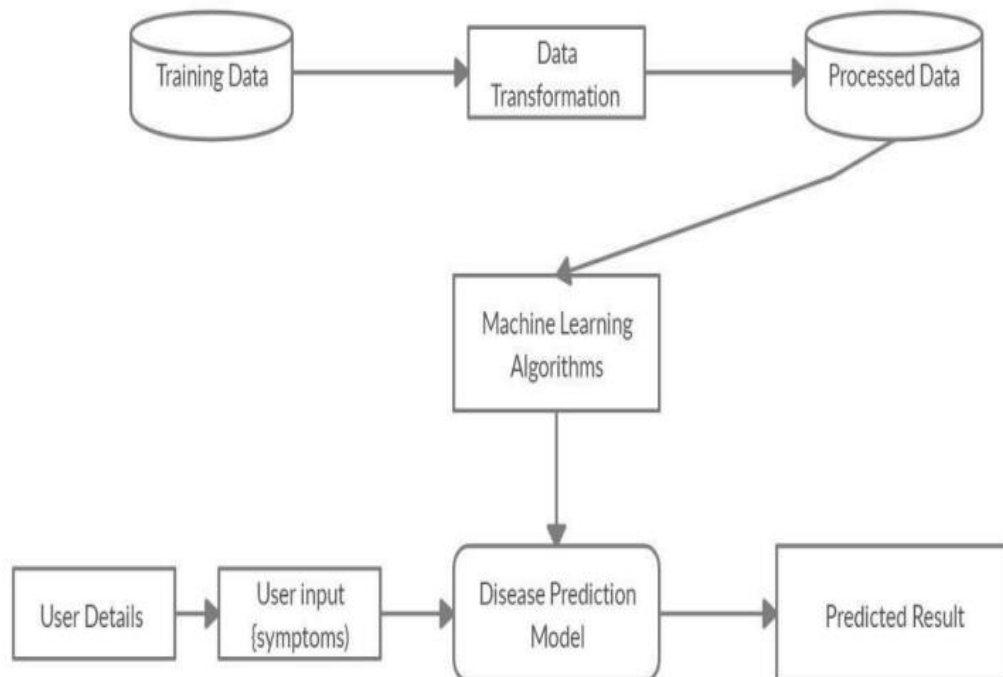


Fig 2: System Architecture

5.2 USE CASE DIAGRAM:

The Use Case diagram of the project Multiple Disease Prediction using machine learning consist of all the various aspects a normal use case diagram requires. This use case diagram shows how from starting the model flows from one step to another, like he enters into the system then enters all the information's and all other general information along with the symptoms that goes into the system, compares with the prediction model and if true is predicts the appropriate results otherwise it shows the details where the user if gone wrong while entering the information's and it also shows the appropriate precautionary measure for the user to follow. Here the use case diagram of all the entities is linked to each other where the user gets started with the system.

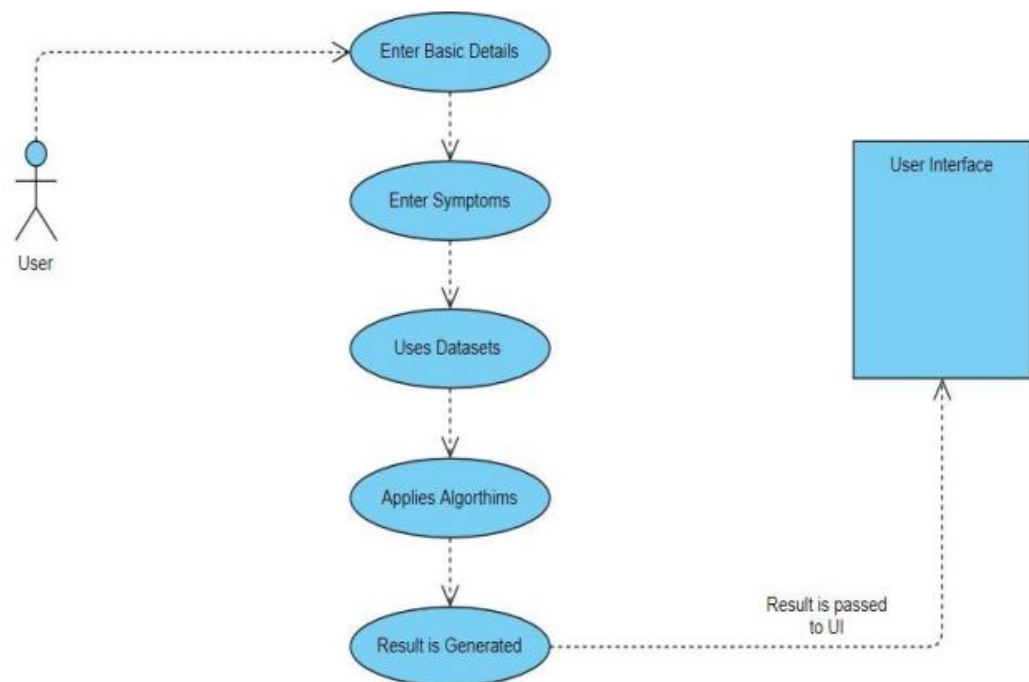


Fig 3: Use Case Diagram

5.3 CLASS DIAGRAM:

Multiple Disease Prediction using machine learning consist of class diagram that all the other application that consists the basic class diagram, here the class diagram is the basic entity that is required in order to carry on with the project. Class diagram consist information about all the classes that is used and all the related datasets, and all the other necessary attributes and their relationships with other entities, all these information is necessary in order to use the concept of the prediction, where the user will enter all necessary information such as user name, email, phone number, and many more attributes that is required in order to login into the system and using the files concept we will store the information of the users who are registering into the system and retrieves those information later while logging into the system.

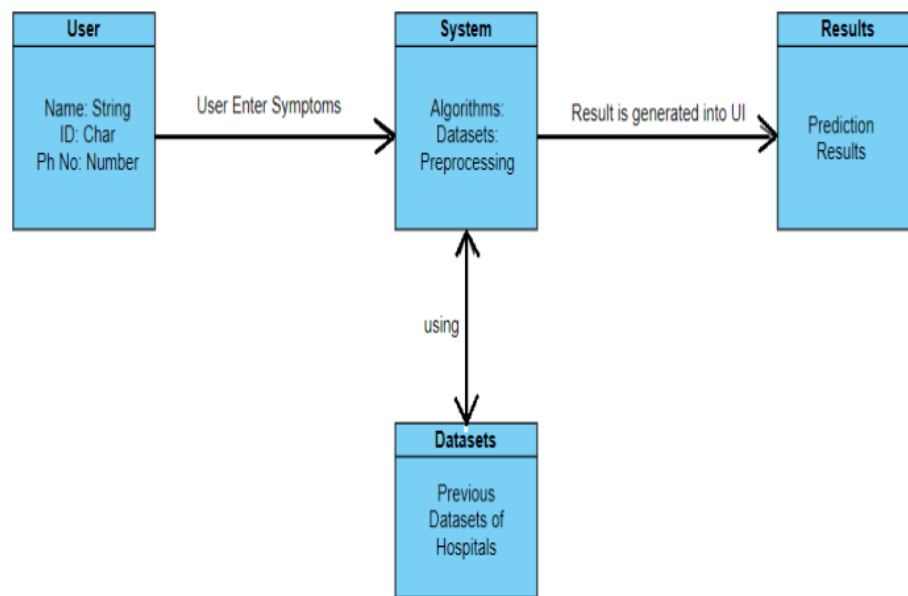


Fig 4: Class Diagram

5.4 SEQUENCE DIAGRAM:

The Sequence diagram of the project Multiple Disease Prediction using machine learning consist of all the various aspects a normal sequence diagram requires. This sequence diagram shows how from starting the model flows from one step to another,like he enter into the system then enters all the information's and all other general information along with the symptoms that goes into the system, compares with the prediction model and if true is predicts the appropriate results otherwise it shows the details where the user if gone wrong while entering the information's and it also shows the appropriate precautionary measure forthe user to follow.Here the sequence of all the entities are linked to each other where the user gets started with the system.

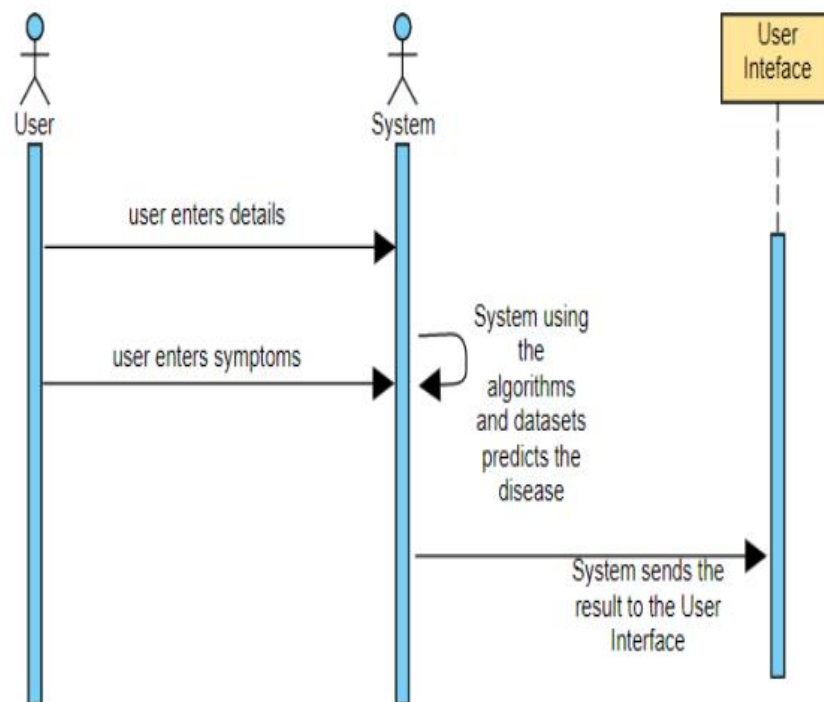


Fig 5: Sequence Diagram

5.5 ACTIVITY DIAGRAM:

Activity diagram is another important diagram in UML to describe the dynamic aspects of the system. Activity diagram is basically a flowchart to represent the flow from one activity to another activity. The activity can be described as an operation of the system. The control flow is drawn from one operation to another. Here in this diagram the activity starts from user where the user registers into the system then login using the credentials and then the credentials are matched in the system and if its true, then the user proceeds to the prediction phase where the prediction happens. Then finally after processing the data from datasets the analysis will happen then the correct result will be displayed that is nothing but the Output.

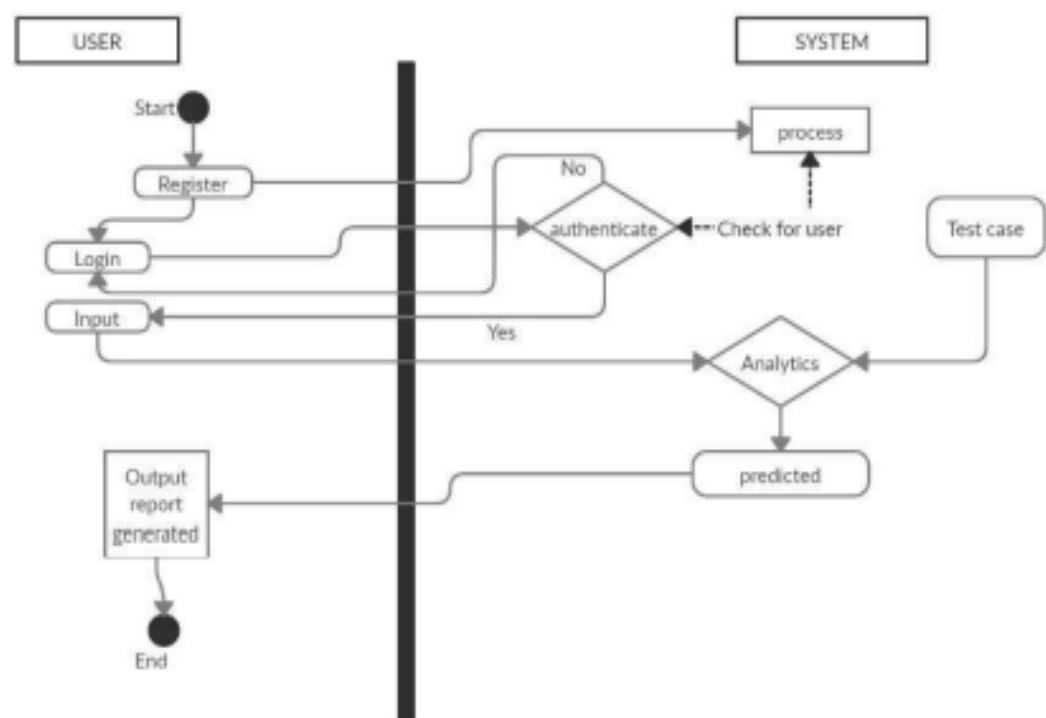


Fig 6: Activity Diagram

5.6 COMPONENT DIAGRAM:

A component diagram, also known as a UML component diagram, describes the organization and wiring of the physical components in a system. Component diagrams are often drawn to help model implementation details and double-check that every aspect of the system's required function is covered by planned development. Here component diagram consists of all major components that is used to built a system. So, Design, Algorithm, File System and Datasets all are linked to one another. Datasets are used to compare the results and algorithm is used to process those results and give a correct accuracy and design UI is used to show the result in an appropriate way in the system and file system is used to store the user data. So, like this all components are interlinked to each other.

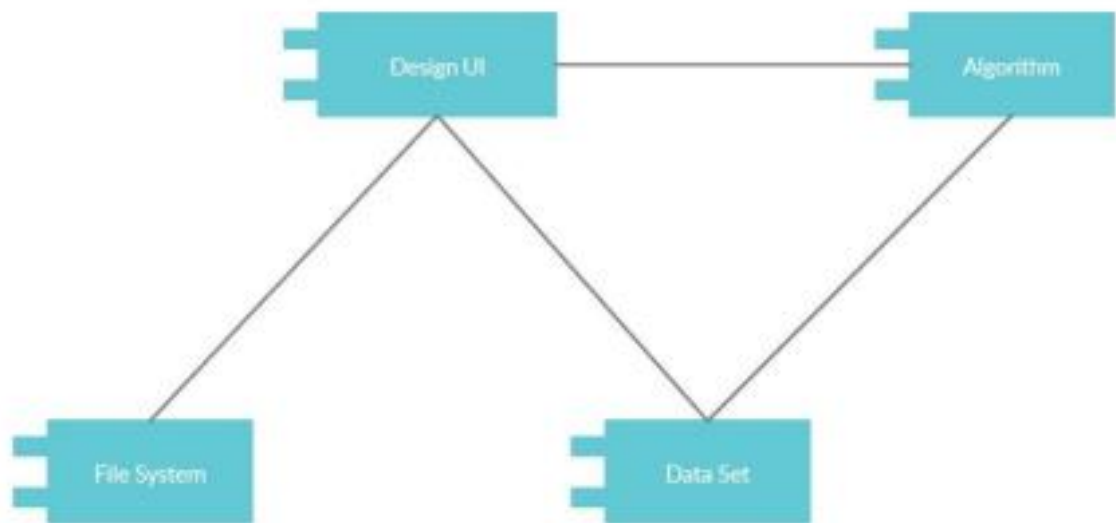


Fig 7: Component Diagram

5.7 DEPLOYMENT DIAGRAM:

A deployment diagram shows the configuration of run time processing nodes and the components that live on them. Deployment diagrams is a kind of structure diagram used in modelling the physical aspects of an object-oriented system. Here the deployment diagram show the final stage of the project and it also shows how the model looks like after doing all the processes and deploying in the machine. Starting from the system how it processes the user entered information and then comparing that information with the help of datasets, then training and testing those data using the algorithms such as decision tree, naïve Bayes, random forest. Then finally processing all those data and information the system gives the desired result in the interface.

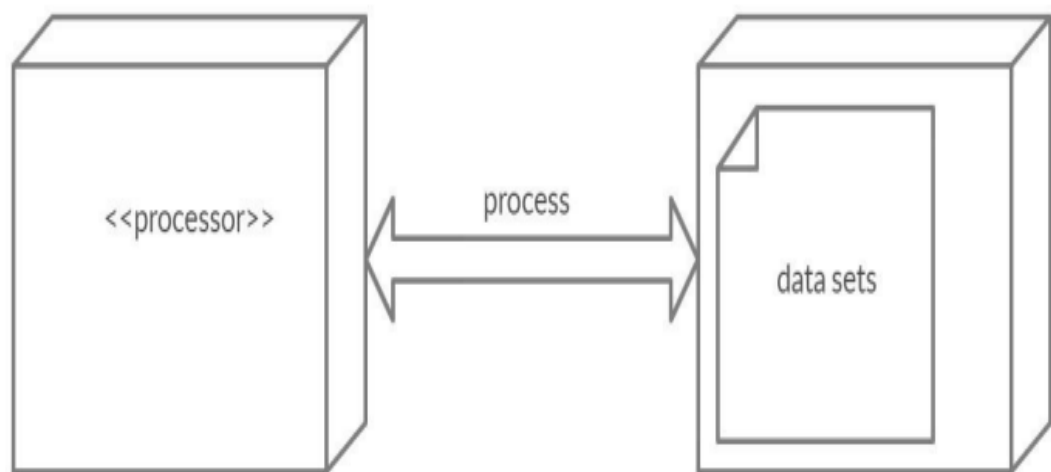


Fig 8: Deployment Diagram