**CHAPTER 1-INTRODUCTION**

There are many tools related to disease analysis. But particularly heart related diseases, tuberculosis, malaria have been analysed and risk level is generated. But generally there are no such tools that are used for prediction of general diseases. So Disease Predictor helps for the prediction of the general diseases analysis and diagnosing of disease become a challenging factor faced by doctors and hospitals both in India and abroad. In order to reduce the large scale of deaths from disease, a quick and efficient Prediction technique is to be discovered. The researchers accelerating their research works to develop a software with the help machine learning algorithm which can help doctors to take decision regarding both prediction and diagnosing of disease. The main objective of this project is predicting the disease of a patient using machine learning algorithms. Comparative study of the various performance of machine learning algorithms is done through graphical representation of the results. The highest mortality of both India and abroad is due to disease. So it is vital time to check this death toll by correctly identifying the disease before time. The matter become a headache for all doctors both in India and abroad. Now a days 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 disease prediction systems which use machine learning algorithms assist in such cases to get accurate results.

# TECHNOLOGIES USED

* Python
* Machine learning
* Tkinter

# CHAPTER 2-RATIONALE

Disease Prediction determines whether person can have disease in near future or having it right now. I used various features for determining the disease using the dataset for evaluation. The dataset in the atlas is sectioned, based on various features. A machine learning algorithm, designed using python language libraries, is used. The outputs generated are reliable and can be readily used by the doctors.

# CHAPTER 3-OBJECTIVES

The objectives of the project DISEASE ANALYSIS are:

1. To find the rate of disease in human beings.
2. To measure the probability of a user for having disease.
3. To implement rule based algorithm as analysing technique into a system.
4. To predict that whether a person will have disease or not.
5. To represent and measure the variation of each factor responsible for disease.

**CHAPTER 4-FEASIBILITY STUDY**

A feasibility study is an analysis used in measuring the ability and likelihood to complete a project successfully including all relevant factors. It must account for factors that affect it such as economic, technological, legal and scheduling factors. Project managers use feasibility studies to determine potential positive and negative outcomes of a project before investing a considerable amount of time and money into it.

* **Economic & Ecological feasibility**

Economics is the study of value, costs, resources, and their relationship in a given context or situation. In the discipline of software engineering, activities have costs, but the resulting software itself has economic attributes as well. Software engineering economics provides a way to study the attributes of software and software processes in a systematic way that relates them to economic measures.

* **Legal/Ethical Feasibility:**

The project undertaken is a legally feasible as it does not violate any rights of our users as well as use authorized trademark.

* **Marketing feasibility**

This project will be all time trending and highly feasible. This type of application is having a great demand by faculty. The existing system is time consuming and manual which is not viable.

* **Financial Feasibility**

Financially, it is also very feasible it made up in very low cost and will be free for all customers.

**CHAPTER 5-METHODOLOGY/PLANNING OF WORK**

**Step 1**: Data collection and dataset preparation This will involve collection of medical information from various sources like hospitals, then pre-processing is applied on dataset which will remove all the unnecessary data and extract important features from data.

**Step 2**: Developing a probabilistic model and deep learning approach (RNN) for Disease Prediction in this step probabilistic model and deep learning approach based on RNN is to be developed it will run effectively on extensive databases of healthcare. And generate decision tree also it can deal with a huge number of information variables without variable deletion.

**Step 3**: Training and experimentation on datasets The Disease Prediction model will be trained on the dataset of diseases to do the prediction accurately and produce Confusion matrix. In this project 3 different algorithms were used -

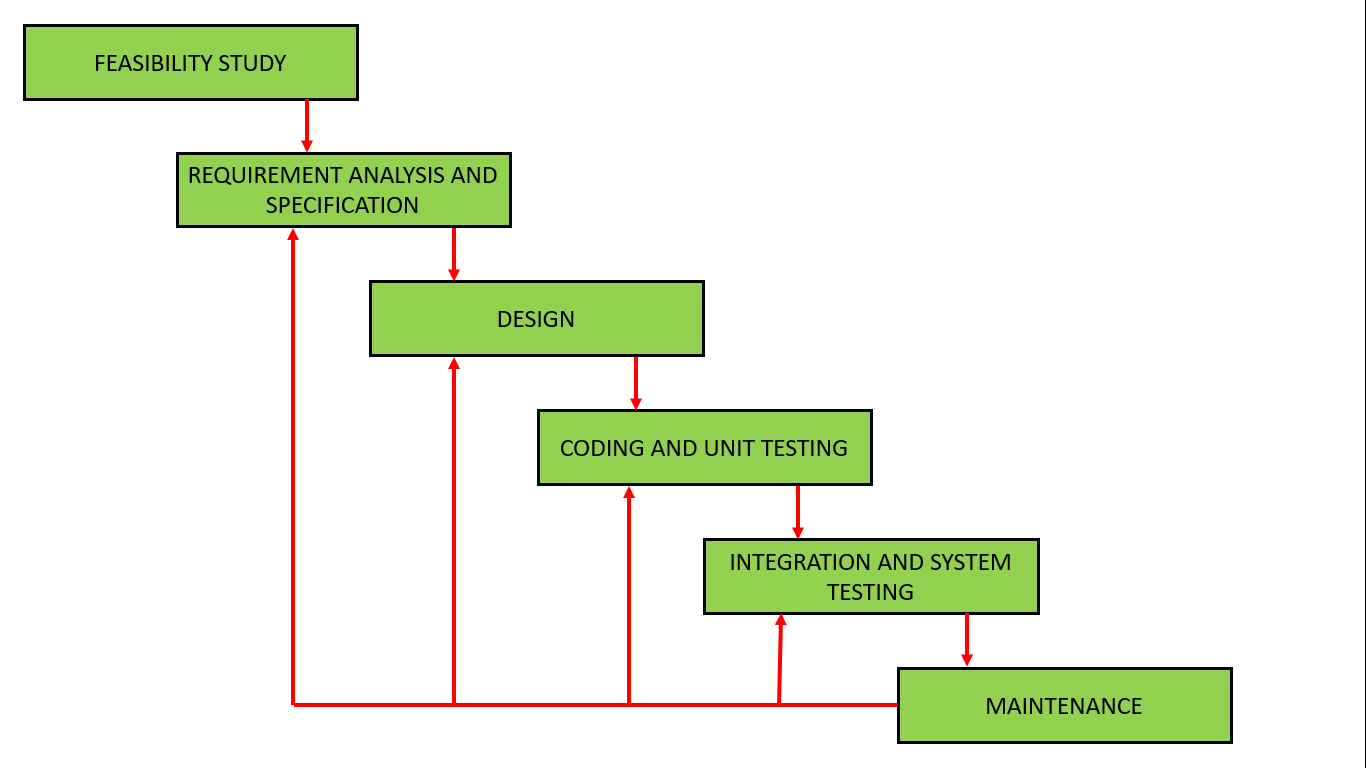
* Decision Tree
* Random Forest
* Naïve Bayes

**Step 4**: Deployment and analysis on real life scenario The trained and tested prediction model will be deployed in a real-life scenario made by the human experts & will be leveraged for further improvement in the methodology.

**Iterative Waterfall Method**

What is the most common way to plan out a project? Sequence the tasks that lead to a final deliverable and work on them in order. This is the Waterfall methodology — the traditional method for managing projects and the one that is simplest to understand. One task must be completed before the next one begins, in a connected sequence of items that add up to the overall deliverable. It’s an ideal method for projects that result in physical objects (buildings, computers), and project plans can be easily replicated for future use.

The power of this methodology is that every step is pre-planned and laid out in the proper sequence. While this may be the simplest method to implement initially, any changes in customers’ needs or priorities will disrupt the sequence of tasks, making it very difficult to manage.



**Fig: 5.1 Iterative waterfall model**

**CHAPTER 6-FACILITIES REQUIRED**

**Table No: 6.1 Hardware Specification**

|  |  |
| --- | --- |
| *Operating System* | Windows 7 and above |
| *Processor* | Intel Core Series or AMD Ryzen Series |
| *Speed* | 3.0 GHZ |
| *Memory* | 4 GB RAM |
| *Hard Disk Drive* | 200MB |

**Table No: 6.2 Software Specification**

|  |  |
| --- | --- |
| *Development Environment* | PyCharm IDE |
| *Language* | Python |
| *Front End* | Python(tkinter) |
| *Back End* | Python |

**CHAPTER 7-REFERENCES**

[1] Kaveeshwar, S.A., and Cornwall, J., 2014, “The current state of disease mellitus in India”. AMJ, 7(1), pp. 45-48.

[2] Dean, L., McEntyre, J., 2004, “The Genetic Landscape of Disease [Internet]. Bethesda (MD): National Center for Biotechnology Information (US); Chapter 1, Introduction to Disease. 2004 Jul 7.

[3] Mohammed, A.K., Sateesh, K. P., Dash G. N., 2013, “A Survey of Data Mining Techniques on Medical Data for Finding Locally Frequent Diseases” International Journal of Advanced Research in Computer Science and Software Engineering, 3(8), pp. 149-153.