**A PROJECT REPORT (IB345-G51)**

**ON**

**DISEASE PREDICTION USING MACHINE LEARNING**

*A report submitted in partial fulfilment of the requirement for the award of*

*The degree of*

**BACHELOR OF TECHNOLOGY**

**In**

**Computer Science And Engineering**



**Submitted to:**

Mr. Aditya Dev Mishra

Assistant Professor

Computer Science And Engineering.

**Submitted by:**

Harsh Dhonchak

Harsh Kumar

Kartikey Dixit

Kanak Agrawal

**SCHOOL OF COMPUTING**

**DIT UNIVERSITY, DEHRADUN**

(State Private University through State Legislature Act No. 10 of 2013 of Uttarakhand and approved by UGC)

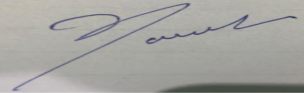
**Mussoorie Diversion Road, Dehradun, Uttarakhand - 248009, India.**

**2022**

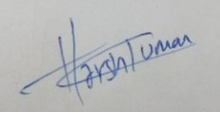
CANDIDATES DECLARATION

I hereby certify that the work, which is being presented in the Report, entitled **Disease Prediction Using Machine Leaning**, in partial fulfilment of the requirement for the award of the Degree of **Bachelor of Technology** and submitted to the DIT University is an authentic record of my work carried out during the period from ***17 August 2021*** to ***30 November 2021*** under the guidance of Mr. Aditya Dev Mishra.

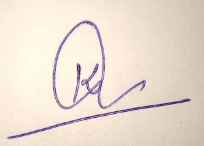
**DATE:**  **19/11/2021**



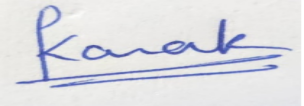
**Harsh Dhonchak**



**Harsh Kumar**



**Kartikey Dixit**



**Kanak Agrawal**

**ACKNOWLEDGEMENT**

We have taken efforts in this project. However, it would not have been possible without the kind support and help of many individuals. We would like to extend our sincere thanks to all of them.

First of all we are thankful to our project guide Mr. Aditya Dev Mishra under whose guideline we were able to progress in our project. We are wholeheartedly thankful to him for giving us his valuable time & attention & for providing us a systematic way for progressing our project in time.

Then we are thankful to all members of this project, without their support and handwork’s this could never be possible to make any progress in this project.

**Harsh Dhonchak**

**Harsh Kumar**

**Kartikey Dixit**

**Kanak Agrawal**

**ABSTRACT**

The system “Disease Prediction Using Machine Learning(ML)” is developed for accurate prediction of the users disease using ML(Machine Leaning). If. the. patient. is. not. seriously. ill. and. the. user. simply. wants. to. know. what. type. of. ailment. he. or. she. has. suffered. from. It. is. a. .system. that. gives. the. user. with. tips. and. techniques. for. maintaining. the. user's. health. system., as. well. as. a. way. to. forecast. disease. using. this. prediction. This. is. a. great. tool. for. the. health. sector. to. inform. the. user., as. well. as. for. the. user. if. he. or. she. does. not. wish. to. visit. a. hospital. Or. other. clinic. The customer can search out what illness he or she has by simply putting the symptoms and a few other details into the software. Python and Machine Learning is used to implement the whole project and also the user interface is created using Node.js ,with the help of previously available dataset.

**Table Of Content**

* Candidate’s Declaration
* Acknowledgement
* Abstract
* **Chapter 1-:**
  1. **Disease Prediction**
  2. **Problem Definition**
  3. **Project Features**
* **Chapter 2-:**
  1. **Purpose of the project**
  2. **System Flow chart Diagram**
  3. **State Chart Diagram**
* **Chapter 3-: Literature Survey**

**3.1 Machine Leaning**

**3.2 Python**

**3.3 Node.js**

* **Chapter 4-:**

**4.1 Tools and Technology required (Software & Hardware)**

* **Chapter 5-:**

**5**.1 **Implementation Screen shots**

* **Chapter 6-:**

**6**.1 **Conclusion And Future Enhancement**

Chapter 1

Disease Prediction :

The system “Disease Prediction Using Machine Learning(ML)” is developed for accurate prediction of the users disease using ML(Machine Leaning). If. the. patient. is. not. seriously. ill. and. the. user. simply. wants. to. know. what. type. of. ailment. he. or. she. has. suffered. from. It. is. a. .system. that. gives. the. user. with. tips. and. techniques. for. maintaining. the. user's. health. system., as. well. as. a. way. to. forecast. disease. using. this. prediction. This. is. a. great. tool. for. the. health. sector. to. inform. the. user., as. well. as. for. the. user. if. he. or. she. does. not. wish. to. visit. a. hospital. Or. other. clinic. The customer can search out what illness he or she has by simply putting the symptoms and a few other details into the software. Python and Machine Learning is used to implement the whole project and also the user interface is created using Node.js ,with the help of previously available dataset.

Many other institutions had also done DPUML before, but our goal is to make it innovative and valuable to the users of this system. This Disease Prognosis Using Machine Learning was built solely with the Python programming language and Machine Learning, as well as a previously available dataset. Doctors nowadays use a variety of scientific methods and methodologies for identifying and diagnosing not just common ailments, but also numerous fatal disorders.

Therapeutic efficacy is always credited to a correct and complete diagnosis. Doctors may occasionally make incorrect judgements while diagnosing a patient's disease; As a result, disease prediction. systems that use ML algorithms can assist in obtaining credible results in such circumstances. The research illness prediction using machine learning was established to tackle general disease at an earlier stage. As we all know, mankind has grown so engrossed in the competitive environment of economic progress that he or she is unconcerned about health. According to studies, 40% of people overlook general sickness, which eventually leads to dangerous disease.

The main motivation for creating this project is to allow users to receive a health check-up at their leisure.

Problem Definition :

In today's health-care industry, there are a number of issues with technologies or devices that produce inaccurate or undesirable outcomes. As a result, we are constructing a software or project that will deliver accurate predictions based on information provided by the user as well as datasets available in that machine in order to prevent undesirable outcomes and acquire the correct and desired outcomes. The health business lacks information and knowledge, yet it is a large industry with a lot of work to perform. So, with the aid of all those algorithms, strategies, and approaches, we completed this project that will assist individuals in need. Many people go to clinics and hospitals for health check-ups but the problem here is that sometimes the patient may not get the result due to various problems like the doctor being on leave or may be busy in any other work or may the unavailability of the machine etc. and also the patient has to travel in order to reach the hospital or clinic which can be time consuming and costly. If a person has been experiencing a few symptoms but is unaware what illness he or she is suffering from, it could convert into a big disease in future. This disease prediction will be incredibly valuable to a wide range of people, including children, teenagers, adults, and senior citizens, in order to avoid this and learn about the ailment in its early stages.

##### PROJECT FEATURES:

The following are the characteristics of Disease Prediction Using Machine Learning :

* Using datasets, this project will forecast the diseases of patients based on their symptoms and other basic information.
* This is based on the hospitals' past datasets, therefore after comparing, it can produce almost exact findings.
* It can anticipate the patient's disease, solve various difficulties, and prevent many aspects with the use of Disease prediction.
* It secures the system, ensuring that no one can access it or make modifications to it. The user must select symptoms from a drop-down menu, and the illness is predicted using algorithms . The user must choose all symptoms in order to get accurate findings.
* Here, we can swiftly prepare data and transform it into an algorithm, minimising the overall workload of the project.
* It provides all of the necessary options for choosing kinds and characteristics.
* The user first registers and than uses the login id and password to enter into the system and further use the applications features.
* When a user first accesses the system, they must register by clicking the register/signup button.
* After which user needs to provide some basic details of signup and then the details of user are saved in system

Chapter 2

**Project Purpose :**

The goal of this research, "Disease Prediction Using Machine Learning," is to accurately forecast a patient's disease using all of their general information as well as their symptoms. Using that information and after comparing from previous patient datasets we will predict the disease that the patient has been through. If we find out in the early stages of diseases with the help of this project the problem can be handled, and this prognostication system can be proved very helpful in the medical field in general.

If the health sector embraces this idea, doctors' workloads will be decreased and they will be able to anticipate the patient's disease more accurately. The overall goal of this Disease prediction is to make predictions for numerous and commonly occurring diseases that, if left untreated and sometimes ignored, can turn in life threatening disease which can further spread among other people. If the health sector embraces this idea, doctors' workloads will be decreased and they will be able to anticipate the patient's disease more accurately. The overall goal of this Disease prediction is to make predictions for numerous and commonly occurring diseases that, if left untreated and sometimes ignored, can turn in life threatening disease which can further spread among other people. Based on the symptoms, this method will estimate the most likely disease. The health business lacks information and knowledge, yet it is a large industry with a lot of work to perform. So, with the aid of all those algorithms, strategies, and approaches, we completed this project that will assist individuals in need.

SYSTEM FLOW CHART DIAGRAM

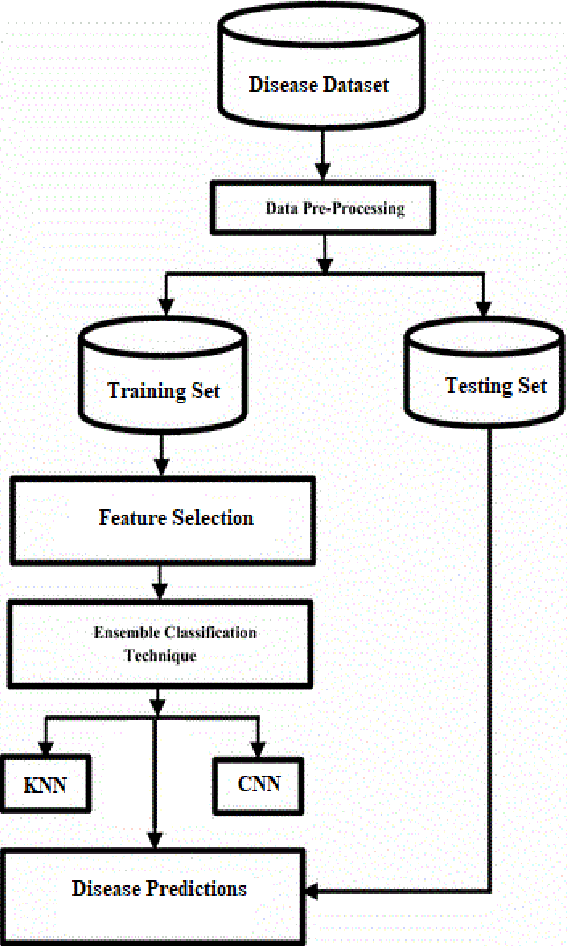


Fig:2.1 System flow Diagram (source : https://d3i71xaburhd42.cloudfront.net/65c2d88c893d69a62ca7a5b38d0ae5491292d8f4/3-Figure1-1.png)

State Chart Diagram :

Diagram

Description automatically generated

Fig:2.2 State Chart Diagram

Chapter 3

**LITERATURE SURVEY**

**MACHINE LEARNING :**

The study of computer systems which will improve themselves automatically supported experience and data is understood as machine learning (ML). so as to get predictions or judgements without being explicitly programmed, machine learning algorithms build a model using sample data, mentioned as "training data." Machine learning algorithms are utilized in a spread of applications, including medicine, email filtering, speech recognition, and computer vision, where traditional algorithms would be difficult or impossible to style.

Every machine learning is not statistical learning, whereas it is related to some mathematical statistics for making predictions with computers. It is assumed that methods, algorithms, and thesis that have worked well in the past are hopefully to function well in the future. These inferences can be clear, such as "since the sun has risen every morning over the last 10,000 days, it will most likely rise again tomorrow morning." It is assumed that methods, algorithms, and thesis that have worked well in the past are hopefully to function well in the future. These inferences can be clear, such as "since the sun has risen every morning over the last 10,000 days, it will most likely rise again tomorrow morning." Mathematical optimization research provides tools, theory, and application domains to the subject of machine learning. A related topic of research is data mining, which focuses on exploratory data analysis via unsupervised learning. Some machine learning implementations use data and neural networks to replicate the operation of a biological brain. Machine learning is sometimes known as predictive analytics when applied to commercial concerns.

It is assumed that methods, algorithms, and thesis that have worked well in the past are hopefully to function well in the future. These inferences can be clear, such as "since the sun has risen every morning over the last 10,000 days, it will most likely rise again tomorrow morning." They can be subtle, such as "X% of families have geographically distinct species with colour variants, hence there is a Y% possibility that unknown black swans exist."

Machine learning algorithms can do tasks that aren't explicitly encoded into them. It comprises computers learning from data in order to complete specified jobs. It is feasible to design algorithms that teach the machine on how to perform all steps necessary to solve the problem at hand for basic jobs entrusted to computers; no learning is required on the computer's part. It might be difficult for a human to manually construct the algorithms required for more complex jobs. In fact, rather than having human programmers explain each essential step, it may be more beneficial to aid the computer in constructing its own algorithm.

Machine learning is a field of study that use a number of techniques to train computers to accomplish tasks for which no perfect answer exists. When there are many alternative answers, one option is to flag some of the right responses as legitimate. We utilise this data as training data to develop the algorithm(s), which it then uses to find valid responses. The MNIST collection of handwritten digits, for example, has frequently been used to train a digital character recognition system.

Tasks for machine learning Typically, machine learning tasks are divided into numerous categories:

Learning under supervision:

The objective is to acquire a generic rule that maps inputs to outputs by presenting the computer with sample inputs and their desired results. The input signal may only be available in limited numbers or be limited to a single response in some instances.

**Active learning:**

The computer can only obtain training labels for a limited number of instances (because to budget constraints), and it must also optimize the objects for which labels are obtained. When used interactively, they might be provided to the user for labelling.

**Unsupervised learning:**

The learning process isn't labelled, so it's up to it to figure out how to find structure in the data. Unsupervised learning can be a goal in and of itself (identifying hidden patterns in data), or it can be a means to an end (finding hidden patterns in data) (feature learning).

**Semi-supervised learning:**

Only a portion of the training signal is sent to the computer: a training set that is missing some (usually many) of the objective outputs.

**Reinforcement learning**:

Only data (in the form of rewards and penalties) is delivered in response to the program's behaviour in a dynamic environment, such as while driving a car or playing a game against an opponent.

**PYTHON :**

Python is a high-level interpreted programming language which is used for various of purposes. Functional programming, aspect-oriented programming (including metaprogramming and metaobjects), object-oriented programming, and structured programming are all possible with its capabilities. Widening allows for a wider variety of specimens, including contract design and logic programming. Python uses dynamic typing, reference counting, and a garbage collector that detects cycles to keep memory under check. Late binding (dynamic name resolution) is also available, which binds method and variable names during programme execution.

Python's developers strive to avoid premature optimization by rejecting non-critical Python changes that may provide minor performance gains at the expense of clarity. A Python programmer can use PyPy, a just-in-time compiler, or relocate time-sensitive routines to C extension modules when performance is necessary. Python’s developers made it a top priority to make the language fun to use.

###### BENEFITS OF PYTHON :

-> Third-party modules are present, as are extensive support libraries.

-> Easy readability

-> Ease syntax

-> User-Friendly

-> High-Extensible and Easy-to-Read Language Productivity and Speed

**Node.js :**

Node.js is an open-source programming language that runs on several platforms. Back-end. Runtime for JavaScript. Outside of a web browser, this environment employs the V8 engine to run JavaScript code. Node.js is used by developers to create command-line tools and server-side scripting, which entails running scripts on the server before sending the page to the browser. Node.js only supports JavaScript as a native language, however there are several compile-to-JS languages available. Node.js apps may now be created in a variety of languages, including Coffee Script, Dart, TypeScript, Clojure Script, and others.

As a result, Node.js serves as a java script model everywhere, combining server-side and client-side script development rather than using separate languages.

Despite the fact that.js is the standard filename suffix for JavaScript code, "Node.js" refers to the entire product rather than a single file. Asynchronous I/O is possible because to Node.js' event-driven design. These design decisions are directly aimed at increasing throughput and scalability in online applications with a lot of I/O activities, as well as real-time Web applications (e.g., Netflix, Uber, PayPal etc.).

Node.js only supports JavaScript as a native language, however there are several compile-to-JS languages available. Node.js apps may now be created in a variety of languages, including Coffee Script, Dart, TypeScript, Clojure Script, and others.

Using JavaScript and a library of "modules" that perform many fundamental operations, Node.js makes it easier to create Web servers and networking tools. Modules provide support for file system I/O, networking (DNS, HTTP, TCP, TLS/SSL, or UDP), binary data (buffers), cryptographic processes, data streams, and other critical services. Node.js modules make advantage of an API that makes creating attended apps more easier.

Node.js only supports JavaScript as a native language, however there are several compile-to-JS languages available. Node.js apps may now be created in a variety of languages, including Coffee Script, Dart, TypeScript, Clojure Script, and others.

Using JavaScript and a library of "modules" that perform many fundamental operations, Node.js makes it easier to create Web servers and networking tools. Binary data (buffers), networking (DNS, HTTP, TCP, TLS/SSL, or UDP), and cryptographic data.

Chapter **4**

Requirement Analysis

**Functional Requirements :**

A functional requirement specifies a system's or component's function. Inputs, outputs and behaviors make up a function. Technical details, calculation and data manipulation, some other functionalities are the basic requirements. Behavioral requirements are indeed the situations wherein the software performs the required functionality. Limiting the design or implementation helps to meet the fundamental needs.

Functional requirements, as developed by requirements engineering, provide specific system outcomes. Cost, reliability, and technical infrastructure are all determined through non-functional constraints. Functional Requirements look after the system's and application's unique functions. The system's functional requirements should be consistent and comprehensive. All required services should be there. The lack of definitions that are incompatible with one another is referred to as "consistency." A functional requirement specifies a system's or component's function . Inputs, outputs and behaviors make up a function. Technical details, calculation and data manipulation, some other functionalities are the basic requirements. Behavioral requirements are indeed the situations wherein the software performs the required functionality. Limiting the design or implementation helps to meet the fundamental needs.

Criteria is presented in generic manner. Functional system requirements, on the other hand, explain the system's function in full, including its inputs and outputs, as well as any exceptions. Let the user enter their id and password, then the system matches them to the appropriate file entries and if the details entered by the user are present then they can proceed else there will be a error message displayed.

**Non-Functional Requirements :**

• Non-functional: “Requirements” is defined as the restrictions/constraints on the system. They could be tied to new system characteristics, such as reliability, response time, or store occupancy, as well as language, platform, implementation methodology, and tool selection.

• Based on the user's needs, budget limits, organizational policies, and many more things the Non-functional requirements are created.

1. Performance requirement: All data entered must be accurate and free of errors in order for the performance to be perfect.

2. Platform constraints: The major goal is to create a smart system that can anticipate.

3. Data Correctness and Precision: Data accuracy and precision are requirements.

4. Modifiability: It should be easy and fast to change(that is add or remove features/problems) from the software. Personnel effort is frequently used as a metric (person- months).

5. Portability: A mobile phone is convenient because it is portable and everyone has a smartphone today.

6. Reliability: A clear definition of failure should be there. Also, don't mix up availability with reliability, which is a very distinct criterion. Make sure to include the implications of software failure, how to prevent failure, an error prediction technique, and a corrective strategy.

7. Security: One or more security criteria for your system and data.

**Hardware and Software Requirements :**

* System : Pentium 4, Intel Core i3, i5, i7 and 2 GHz Minimum
* RAM : 512Mb or above
* Hard Disk : 10 GB or above
* Input Device : Keyboard and Mouse
* Output Device : Monitor or PC
* Operating System : Windows 7, 10 or Higher Versions
* Front End : Handlebars
* Back End : Node.js
* Programming Lang : Python, Java

**CHAPTER 5**

**IMPLEMENTATION SCREENSHOTS :**

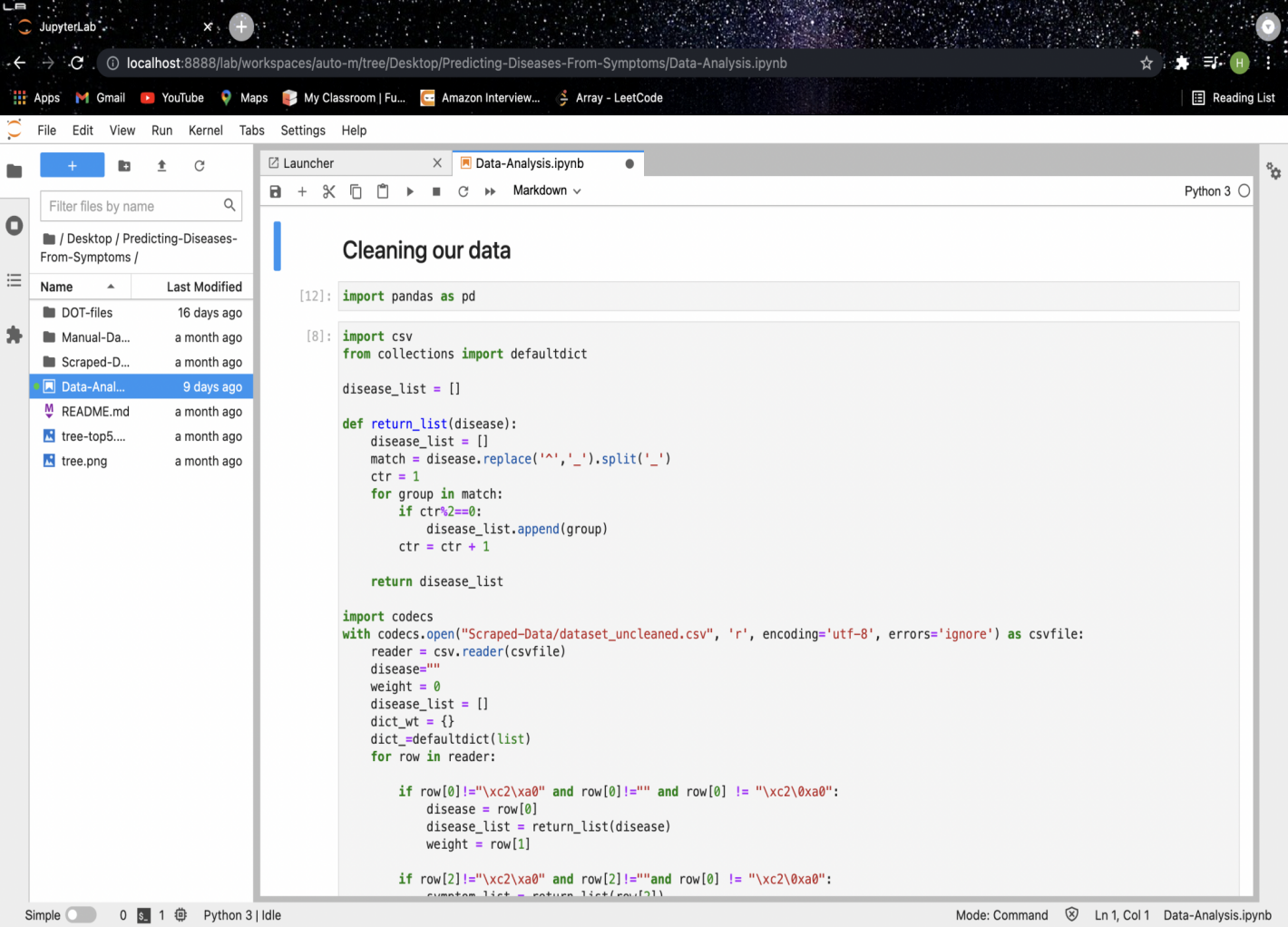


Fig:5.1

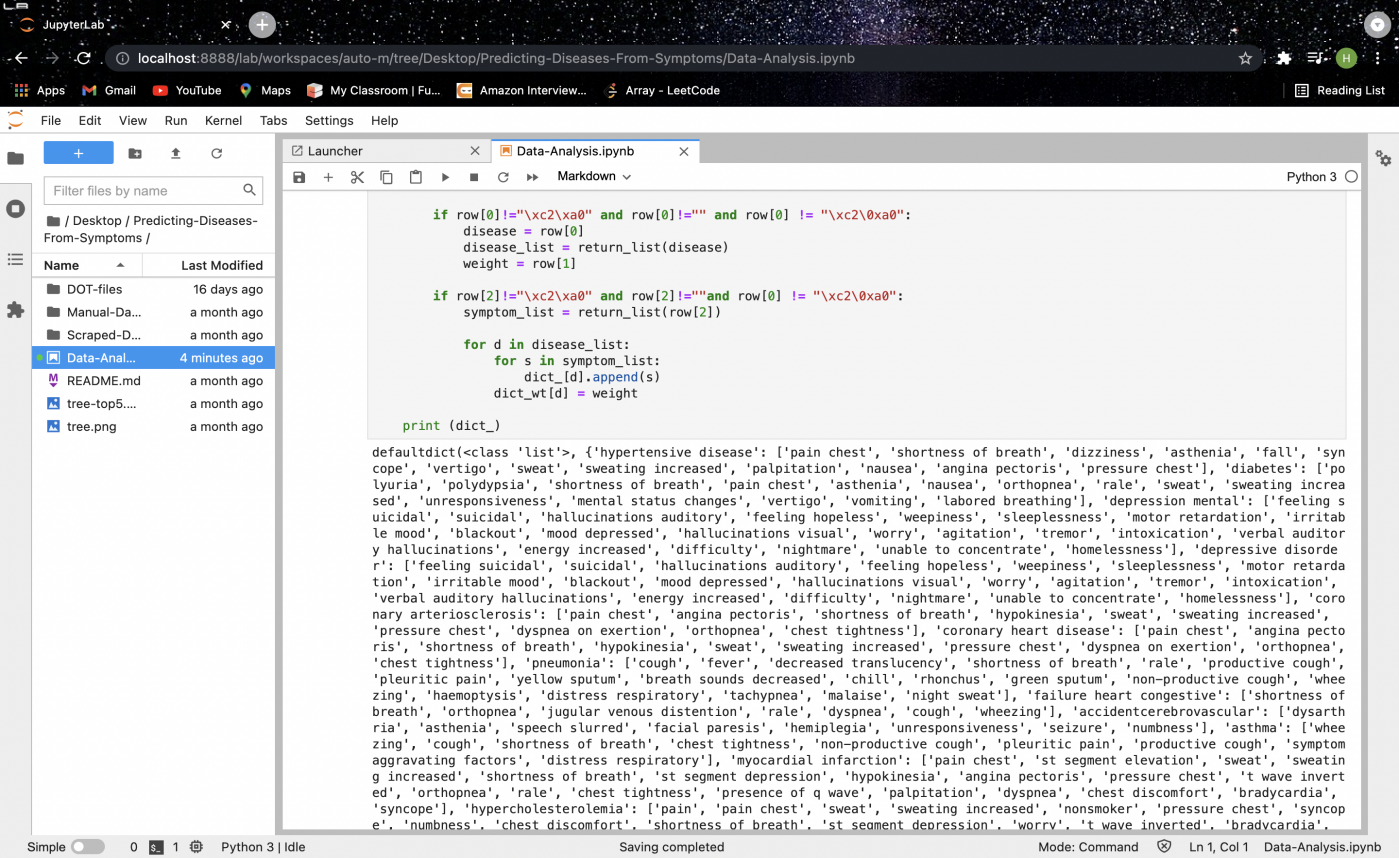


Fig:5.2

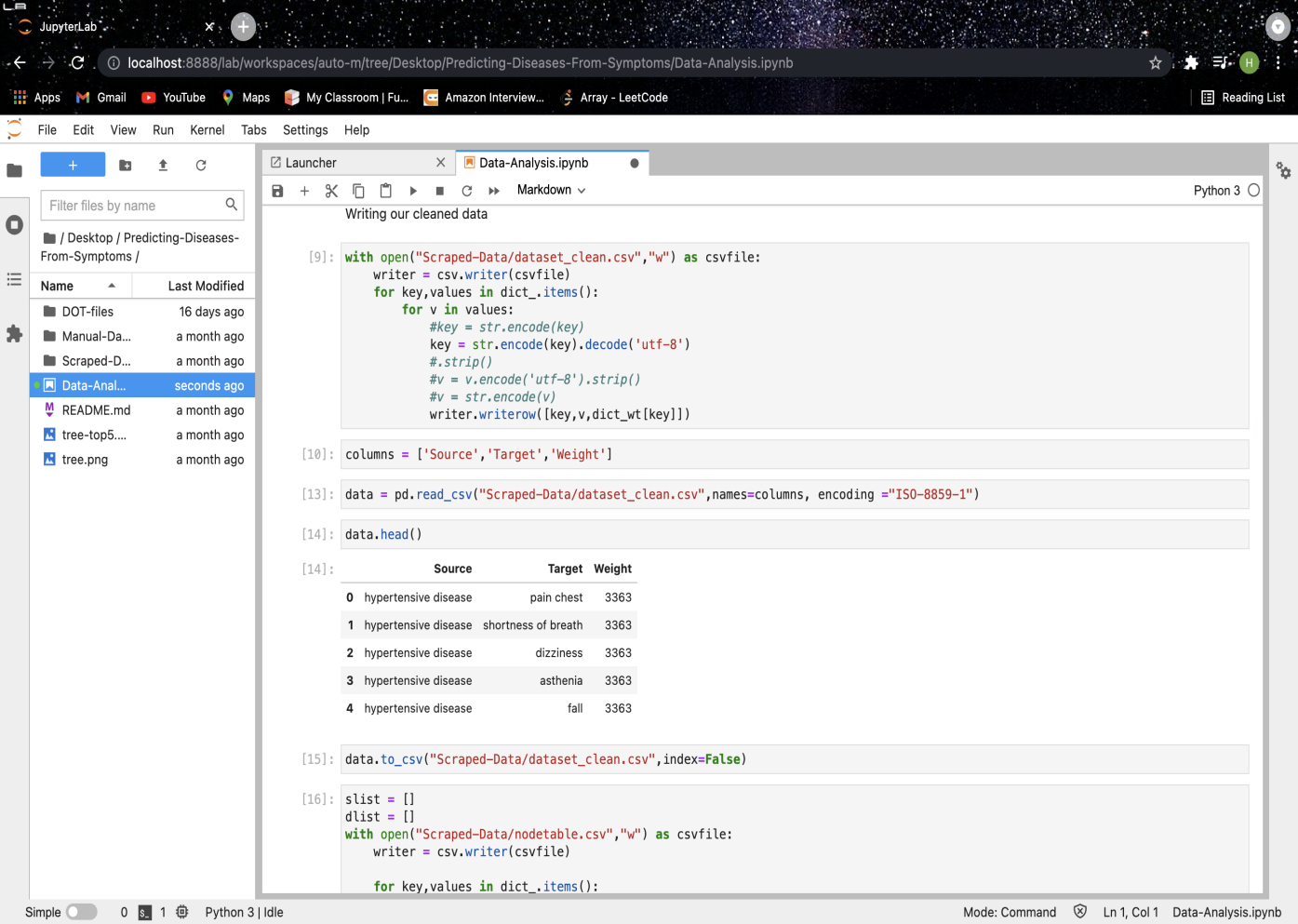


Fig:5.3

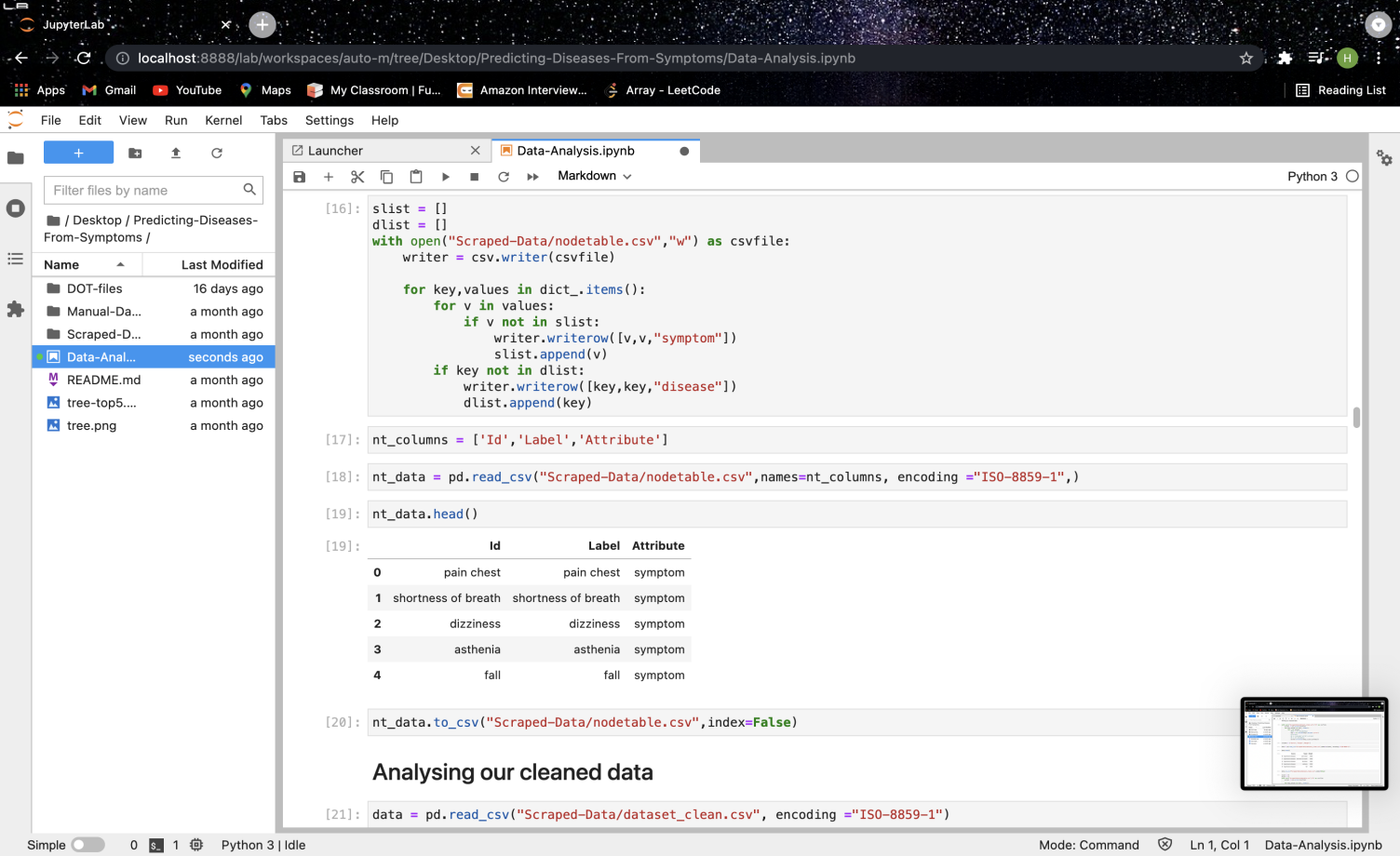


Fig:5.4

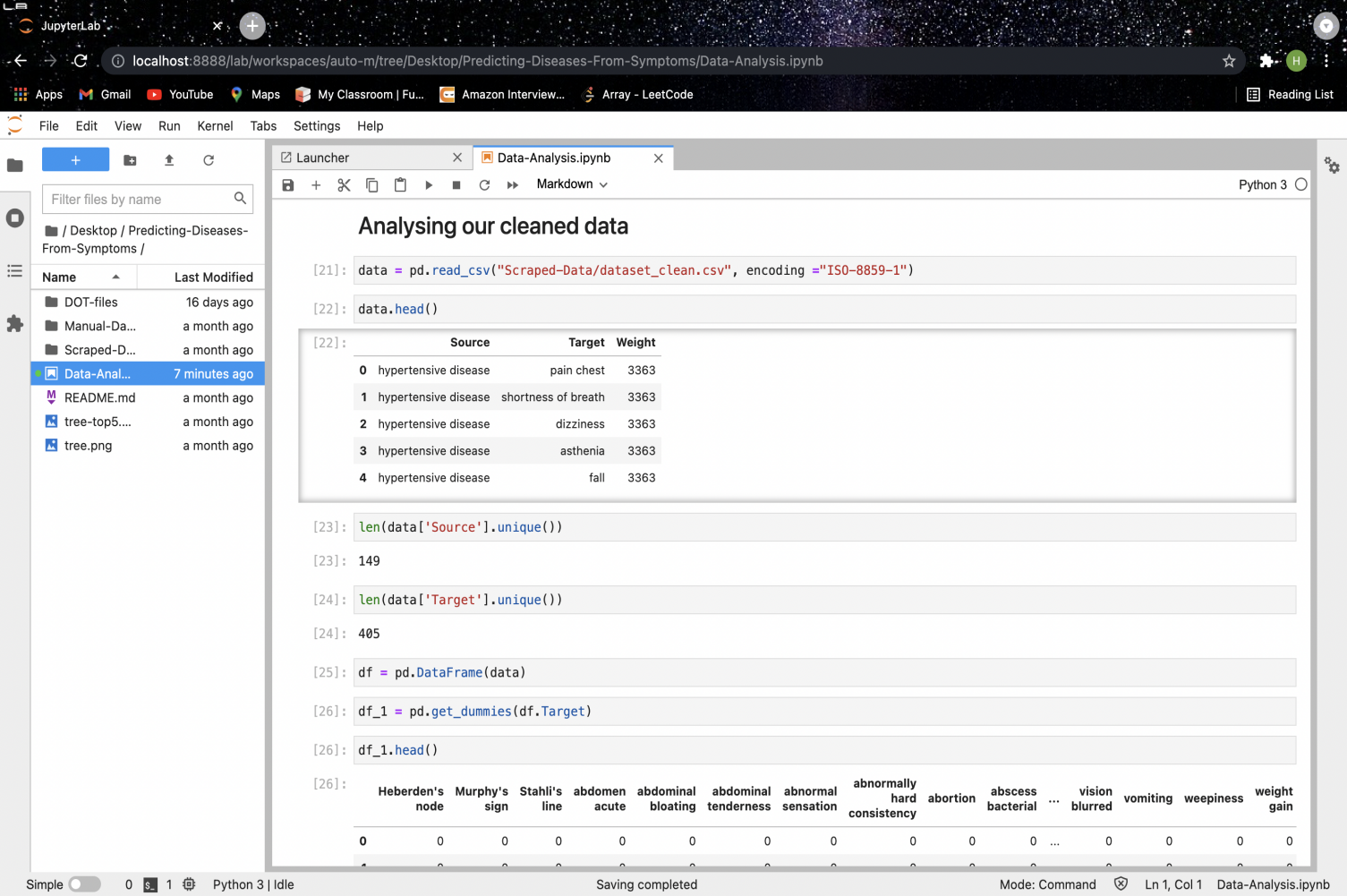


Fig:5.5

**CHAPTER 6**

**Conclusion :**

We may say that the initiative will aid in the alleviation of both the user's and the medic's concerns. The project will save the Health Industry a lot of time because users will only need to enter a few symptoms and will receive quick results that will be valuable to them. For higher accuracy we not only use the text data of the patient, but also structured data based on the proposed k-mean algorithm. We can get as high as 95% accuracy by combining the user entered data and the structured one with the help of k-mean algorithm.

Adopting this project will reduce the stress on doctors and people can be treated if the doctor is tension free. This project is made to provide a prediction on the disease a person is going through, which if ignored can be fatal and can even harm the whole family let alone the infected person.

**Future Enhancement :**

* The ability to change a user's personal information.
* More interactive user interface.
* Facilities for creating backups.
* It's possible to do this as a Mobile app.
* More information about the most recent diseases are available.