GUJARAT TECHNOLOGICAL UNIVERSITY AHMEDABAD

Government Engineering College, Gandhinagar

A Project on

"COVID-19 Prediction System"

Under the subject of

PROJECT-I

B.E Semester-VII

(Information Technology)

Submitted by:

Group: - 99544

Sr No.	Name Of Student	Enrollment No.
1	Malaviya Darshan S.	170130116022
2	Patel Herry D.	170130116035
3	Prajapati Akshay N.	170130116043

Guided By

Prof.Anamika Mittal

Prof. Dr.Dhaval Parikh

Head Of Department

Academic Year

2020-2021

GUJARAT TECHNOLOGICAL UNIVERSITY AHMEDABAD

Government Engineering College, Gandhinagar

CERTIFICATE

This is to certify that the project entitled "**COVID-19 Prediction System**" is submitted by **Team No. 99544.** In partial fulfillment of the requirement for the award of degree Bachelor of Engineering in Department of Information Technology from Gujarat Technological University of the academic year 2020-2021.

ACKNOWLEDGEMENT

We sincerely feel that credit of the project work could not be narrowed down to only for on individual. The development of this project involves many valuable contributions.

We express our sincere gratitude towards our guide **Prof.Anamika Mittal** Project Supervisor. Government Engineering College, Gandhinagar for her valuable guidance and encouragement during our project.

We would like also thankful to our Head of the Department **Prof. D.A.Parikh** for providing all the facilities and support.

Finally, we would like to thank our parents, friends and Faculty Members for being with us support directly or indirectly.

Malaviya Darshan S. [170130116022]
Patel Herry D. [170130116035]
Prajapati Akshay N. [170130116043]

ABSTRACT

The successful application of data mining in highly visible fields like e-business, commerce and trade has led to its application in other industries. The medical environment is still information-rich but knowledge weak. There is a wealth of data possible within the medical systems. However, there is a lack of powerful analysis tools to identify hidden relationships and trends in data. The disease is a term that assigns to a large number of health care conditions related to the body. These medical conditions describe the unexpected health conditions that directly control all the body parts. Medical data mining techniques like association rule mining, classification, clustering is implemented to analyze the different kinds of general body-based problems. Classification is an essential problem in data mining. Several of popular classifiers construct decision trees to generate class models.

"Covid-19 Prediction System" is a web application for those who need tocheck for that they have or not symptoms of covid-19 infection. Here, patient has must upload an X-Ray of their chest and after uploading, our ML Model check that the patient has or not the covid-19 disease. Here, ML model work on our train dataset, which was trained by data of positive and negative symptoms of the covid-19 patient.

Initially, using front-end technology like HTML5, CSS3 and Bootstrap, we build a web application, wherein the company, as a user, will create their account to access this facility. Python-based web application will be build using the Flask framework and model based on Machine Learning and Database based on the MySQL database. After uploading x-ray by the patient the ML model check using their dataset, the patient has or not covid-19 infection.

TABLE OF CONTENTS

1.Introduction

1.1 Project Summary

- This project is based on machine learning and in which we are trying to make a model which get the user's x-ray and based on that data it will predict user has or not a covid-19 infection and also provide that user have or not a risk of covid-19. In other words, there is such algorithm behind the scene that gets the user's x-ray and through the x-ray mining process and some calculation, it will provide the result.
- First, we will create a database that will store the x-ray of user's in the end machine. Then through the backend of the model, it will process the x-ray and gives the result of the test. We will use Python and ML for backend process.

1.2 Aim of the project

- The purpose of doing this project is that it will give the result of that user have or not a covid-19 infection or about risk. Another purpose for making this is that it will give new wings to machine learning technology.

1.3 Technology and Literature Review

Introduction Of Python:

Python is a widely-used general-purpose, high-level programming language. It was created by Guido van Rossum in 1991 and further developed by the Python Software Foundation. It designed with an emphasis on code readability, and its syntax allows programmers to express their concepts in fewer lines of code.

An overview of Python:

Python is a high-level, interpreted, interactive, and object-oriented scripting language. Python is designed to be highly readable. It uses English keywords frequently whereas other languages use punctuation, and it has fewer syntactical constructions than other languages.

Python is Interpreted − Python is processed at runtime by the interpreter. You do not need to compile your program before executing it. This is similar to PERL and PHP.

Python is Interactive − You can sit at a Python prompt and interact with the interpreter directly to write your programs.

Python is Object-Oriented − Python supports an Object-Oriented style or technique of programming that encapsulates code within an object.

Python is a Beginner's Language − Python is an excellent language for beginner-level programmers and supports developing a wide range of applications from simple text processing to WWW browsers to games.

- **Overview of MySQL database:**

MySQL is an in-process library that implements a self-contained, serverless, zero-configuration, transactional SQL database engine. It is a database, which is zero-configured, which means like other databases you do not need to configure it in your system.

MySQL engine is not a standalone process like other databases, you can link it statically or dynamically as per your requirement with your application. MySQLaccesses its storage files directly.

- **Overview of Jupyter Notebook:**

The Jupyter Notebook is an open source web application that you can use to create and share documents that contain live code, equations, visualizations, and text. Jupyter Notebooks are a spin-off project from the IPython project, which used to have an IPython Notebook project itself.

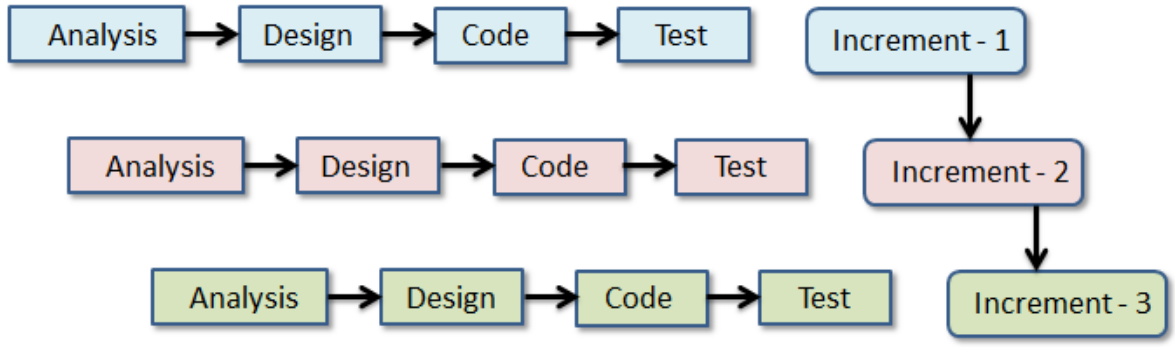
1.4 Plan of Project

- The Project plan should be considered the most vital step prior to project development. First of all, we did market researched of the application serving the very purpose of Resume Parsing. We analyzed them and planned to add some more modules to become more useful and productive.

1.5 Project Development Approach

- A Software Process model is a simplified abstract representation of a software process, which is presented from a particular perspective. Planning prepares a framework that makes a reasonable estimate of the project. To accomplish it, software development models are used. Incremental model is used to satisfy this purpose.

- **Incremental Model:**



- The system uses this method of software development where the model is analyzed, designed, tested, and implemented. In the incremental model, the system is developed in various units. These units entail the requirement planning, development, and test phases. This module is useful for big and small products. It is divided into teams. It satisfies the customer's needs.

- The Incremental Model combines elements of the linear sequential model with the iterative philosophy of prototyping. The incremental model applies linear sequences in a staged fashion as calendar time progresses. Each linear sequence produces a deliverable "increment" of the software. In incremental model first increment is called core product. In core product basic requirements are added, but some new supplementary features remain undelivered. This core product is used by the customers to evolutes the system and the next increment is planned to develop.

- During first increment analysis phase, customer and developers specifies as many requirements as possible and prepare documentation. Now the first version of the product with the minimal and essential features is launched to market. Based on the feedback and experience with this version, a list of additional features is added. This process is repeated

following the delivery of each increment, until the complete product is produced.

- With this approach the first model may be available within a few weeks or months. In this model, less cost and time are required to develop the first increment called the core product. Less risk has occurred to develop the smaller systems represented by the increments. Incremental funding is allowed, means only one or two increments might be funded when the program starts. It can results in better testing, because testing each increment is likely to be more comfortable than the testing the entire system. The feedback providing at each increment is useful for determining the final requirement of system.

2. System Analysis and Study

2.1 System Requirement specification

2.1.1 User Characteristics

- Two types of user interaction with the system:
  1. User
     Users interact with the system through the user interface and cannot perform any read/write operation on the database.

  2. Data Administrator
     Data administrators manage the database and can perform read and write both operations.

2.1.2 Hardware and Software requirement specification

Hardware : Android , IOS Device & Laptop / Desktop

Software : Web browser (i.e. Google Chrome)

2.1.3 Constraints

- Reliability Requirements :
  The main reliability requirement is the validation used. Without proper system did not allow entering that value into the database. All the required validation controls are kept to keep the system secure.

- The following are some of the reliability requirements:
  - In the email ID the user can not enter any dummy value, the validation checks that whether there is a '@' or '_' symbol in that.
  - Any null value is not allowed in place of compulsory fields.

- The numeric field, users cannot enter any character value.
- The date of birth, user cannot enter date and time other than given format.
- Entered password and confirm password must match to each other.
- User can not re-register an account on his primary email.

- Safety and Security :

  - Considerations Safety :
    The source of this software will be kept at more than one place with user-ID, password and also in CD ROM in case of server failure.

  - Security :
    Security in this software provides to a different user in different ways by giving other user id. If the user is admin, he has all the privileges and constraints. He can access the entire database. He can change or delete the databases from other user's accounts. HR departments has limited access according to their role. Because of limited privileges, one of them cannot update other details of the candidate.

2.1.4 Assumption and dependencies :

- Only admin has all privileges.
- User allow access to add x-ray.
- User allow using cookies.
- Algorithm has its own decision.

2.1.5 The Requirement of new System :

- The system should record all details of the user.
- The System should allow the user to change profile.
- The System should allow the user to give feedback on system.
- The system should have the facility of administrative help for the user to solve his important problems.
- The system should have full authority to access user account.
- The system should allow admin to send a new notification.

- Username and password will send to registered email or mobile.
- Password recovery system should there in case of user forgot their password.

2.2 Feasible Study

- A Feasible study gives an overview of the system. First, it gathers information about user and system. Check for system requirements and compare them with the local machines. If it is not compatible, then it will show a warning about can't be compatible with your device.
- A feasible study, it first asks for the user to allow to access of media and use of cookies then store available information in a database like cookies and x-ray.
- The gathered information will further divide into the classical label and will normalize according to requirement. The Feasible study also gives a report on covid-19 symptoms and recommendations of further treatment. The waiting state of algorithm initialization, feasible analysis of data ensures that data is proper, and there is no inconsistency.

  - Technical feasibility :
    This assessment ensures that all required technical support from the user as well as from admin and data administrator will provide as sufficient.
    To meet system requirement at end-user, it is required to check and for that purpose, technical feasibility study assessment is there.

  - Economic feasibility :
    The system is open-source and free to use for any user.
    This is required for system admin that they have enough cost ready to use for maintaining server as well as to the central database.
    Some cost of the technical support team and Help desk member is there to make running system effectively.

  - Operational feasibility :
    Operational feasibility is a measure of how well a proposed system solves the problem. It is required to solve problematic characteristics of the system i.e. improper data, undesirable system and incorrect input.

It is simplified in a manner that once the user allows to getting access to his media, it is the system duty to fetch data(x-ray) from device and evaluate it adequately.
Operational feasibility ensures that all process like input and data processing does in a way the same as in a flow diagram. It is mandatory to maintain server and database so that it doesn't cause any data inconsistency.

- Schedule feasibility :
  As per the flow diagram and use case scenario, it is mandatory that all transactions and interconnection dose in a scheduled time.
  Once the user enters into the system, it starts collecting data from the local device when the user uploads the x-ray. If it is failed to get data, there might be an improper data format or incompatible device. In such a case, the system should do the whole rollback procedure and provide an accurate information regarding the error occurred.
  It is assumed that all interconnection of data and processing of data is doing the work on a time.

2.3 Features

- In this project we are making web application so it can be run on any type of devices without much issue . It can take patient chest X-Ray image as an input and pass the data to the pre-trained machine learning model based on past patient data .After that using machine learning it can predict whether the patient having chance of Covid-19 positive or not based on past patient data .We are using VGG16 and Resnet50 algorithm for the classification task in our machine learning model.

3. Design

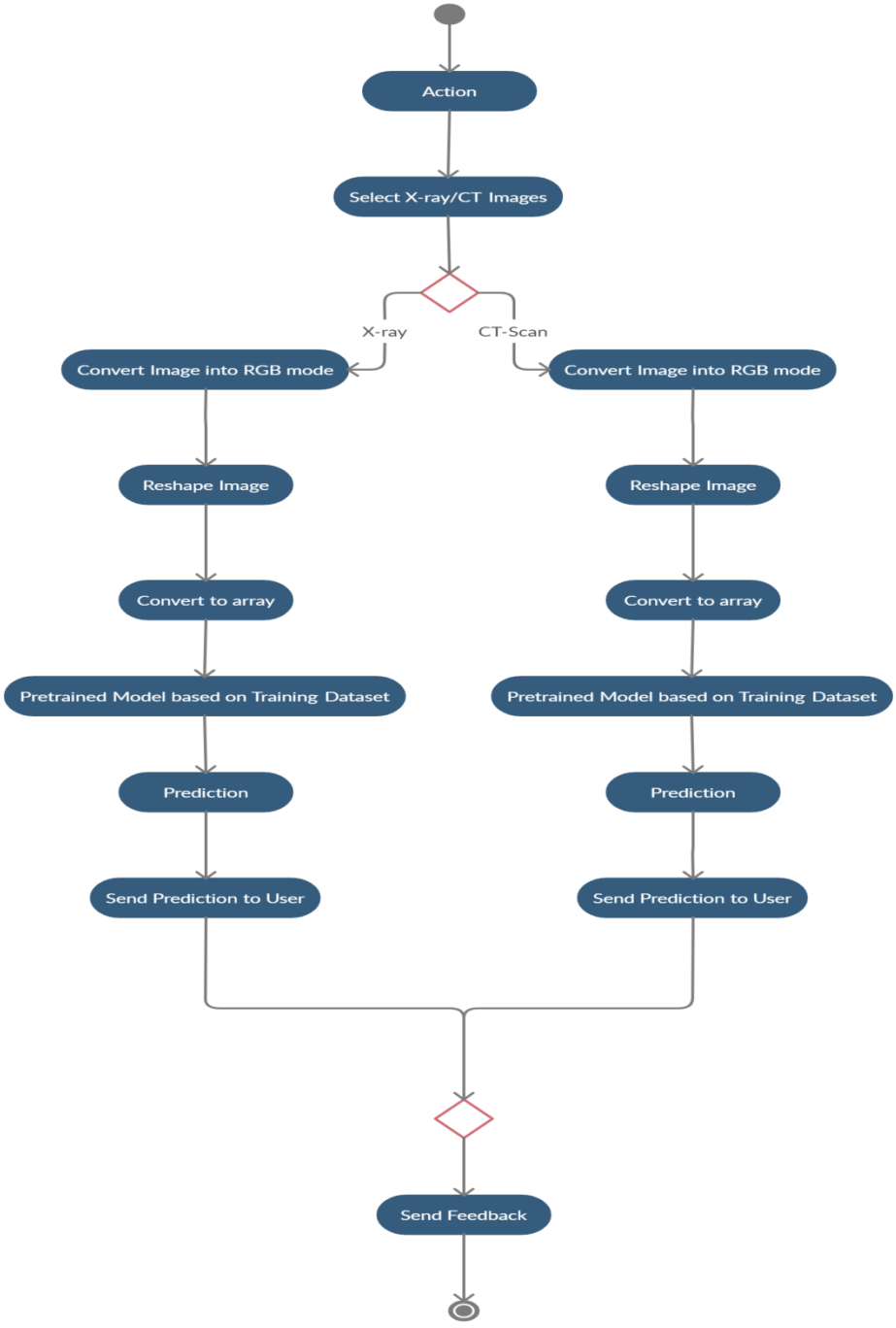
3.1 Diagrams

3.1.1 Use Case Diagram

A use case is a set of scenarios that describe an interaction between a user and a system. A use case diagram displays the relationship among actors and uses cases. The two main components of a use case diagram are use cases and actors.

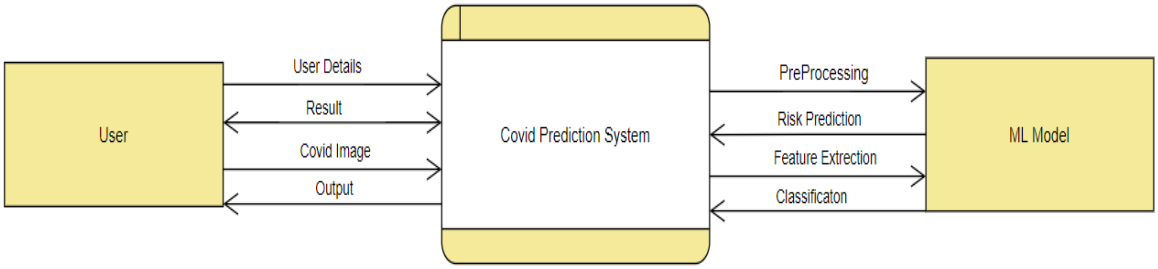
3.1.2 Activity Diagram

An Activity diagram is 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. This flow can be sequential, branched, or concurrent. Activity diagrams deal with all types of flow control by using different elements.

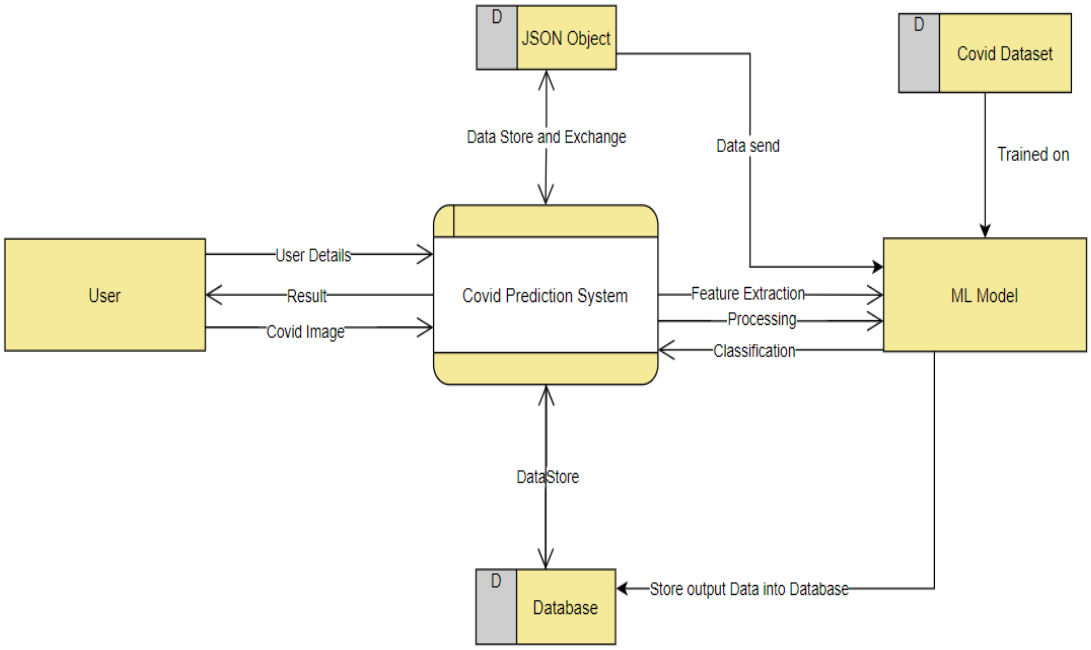


3.1.3 Data Flow Diagram

A Data Flow Diagram (DFD) is a graphical representation of the "flow" of data through an information system modelling, its process aspects. Often it is a preliminary step used to create an overview of the system that can later be elaborated.



DFD 0.0

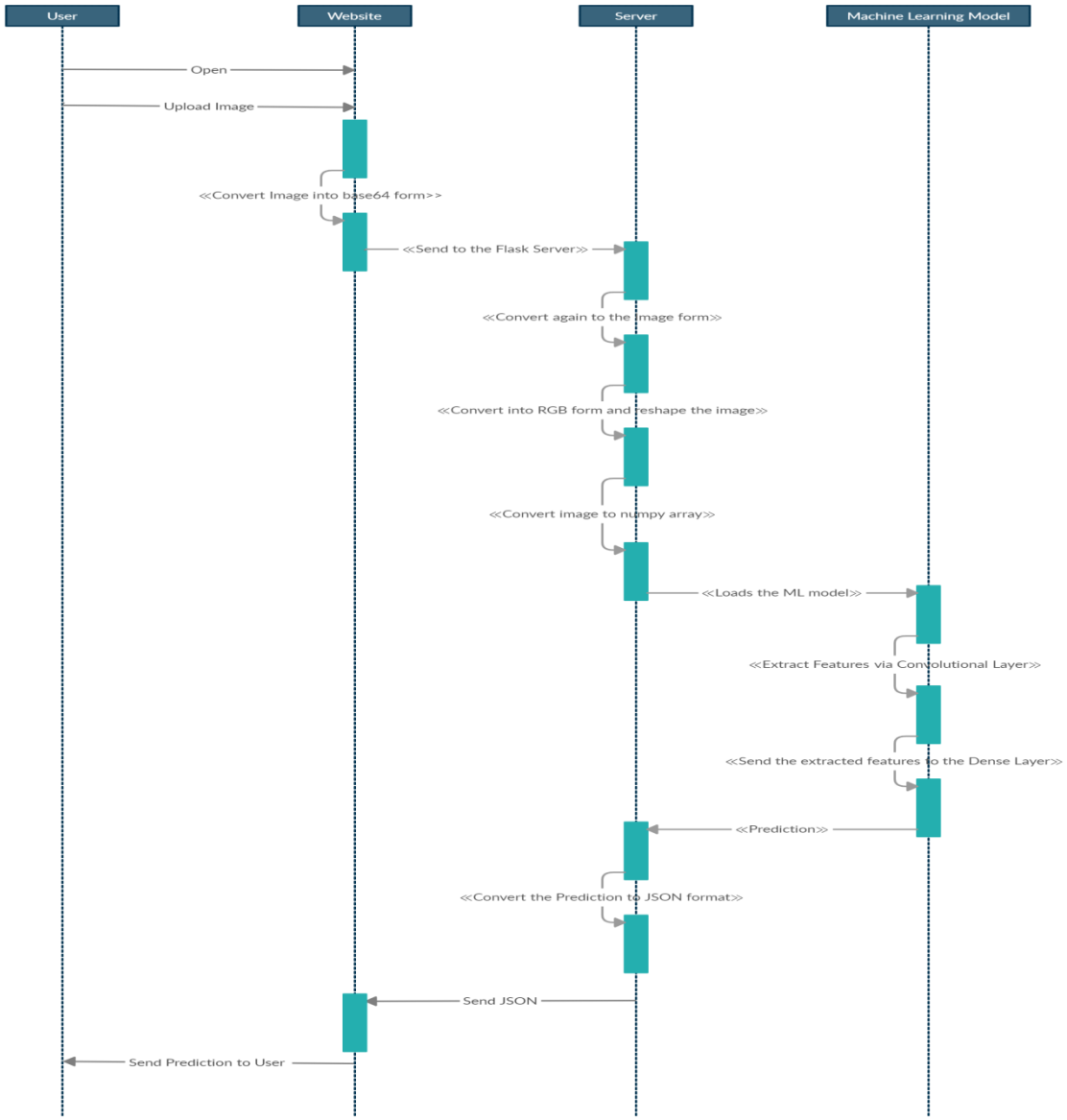


DFD 1.0

3.1.4 Sequence Diagram

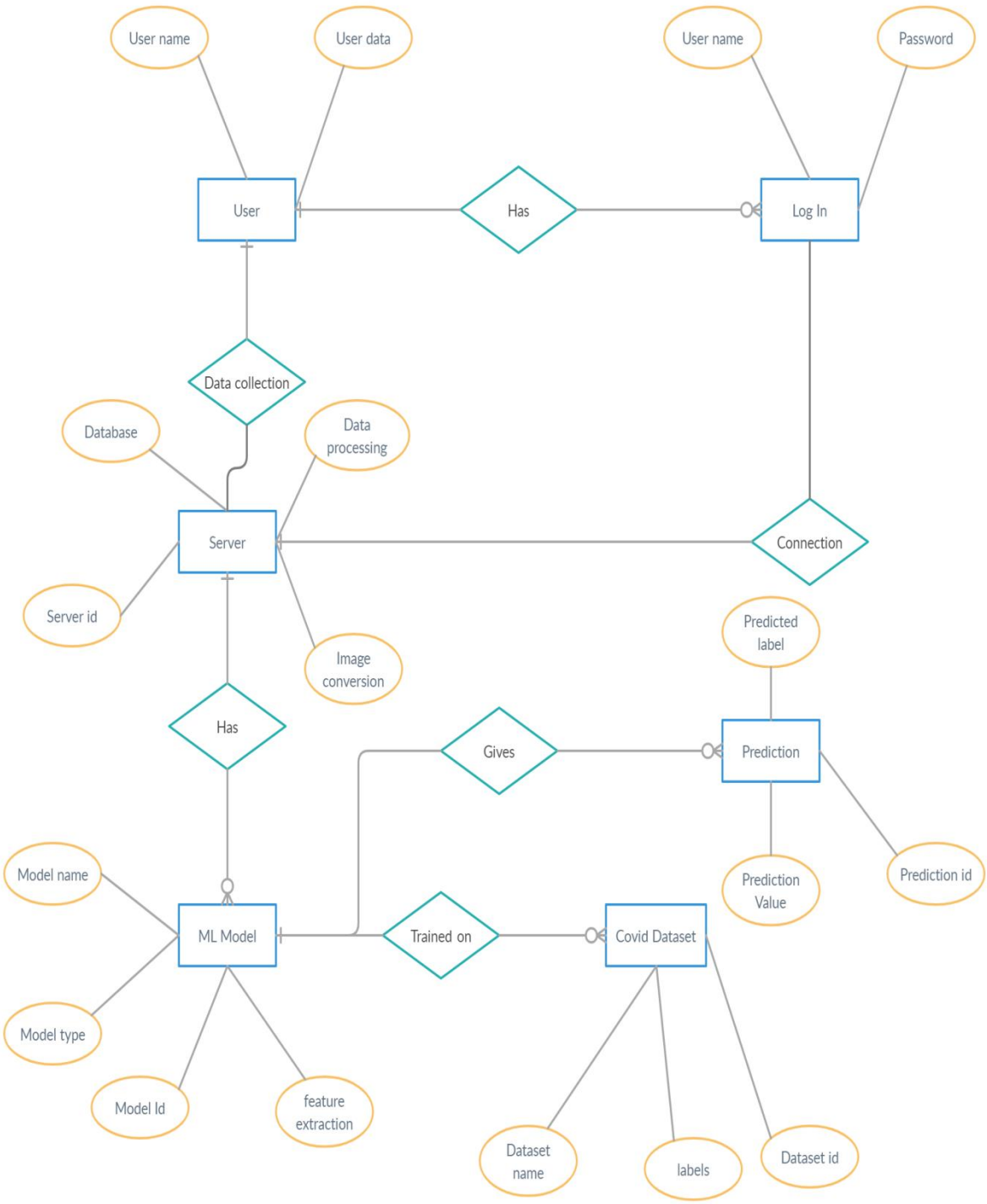
A sequence diagram is an interaction diagram that shows how objects operate with one another and in what order. It is a construct of a message sequence chart.

A sequence diagram shows object interactions arranged in time sequence. It depicts the objects and classes involved in the scenario and the sequence of messages exchanged between the objects needed to carry out the functionality of the scenario.



3.1.5 E-R Diagram

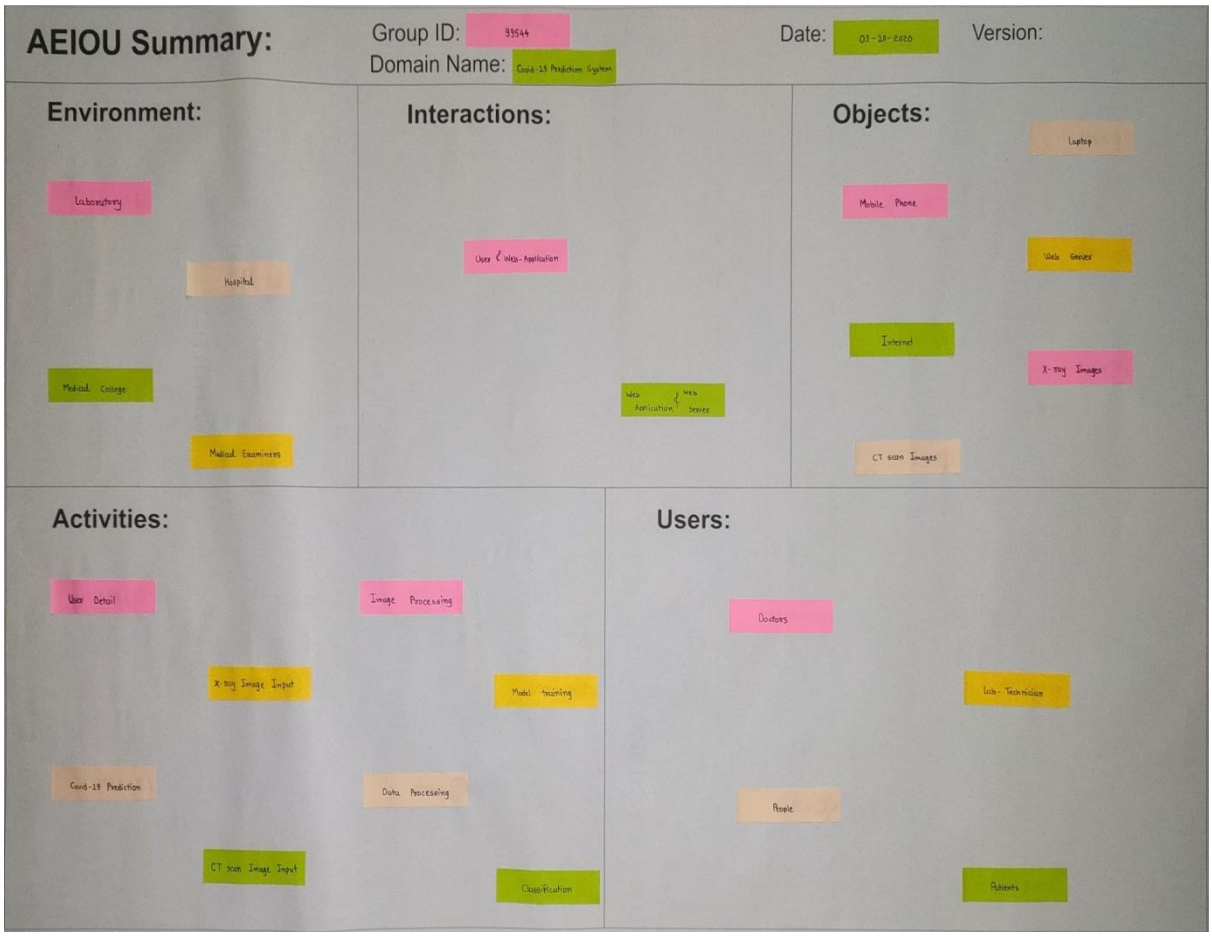
ER diagrams are used to analyze existing databases to find and resolve problems in logic or deployment. Drawing the map should reveal where it's going wrong. Business information systems: The charts are used to design or analyze relational databases used in business processes.



3.2 Design Engineering Canvas

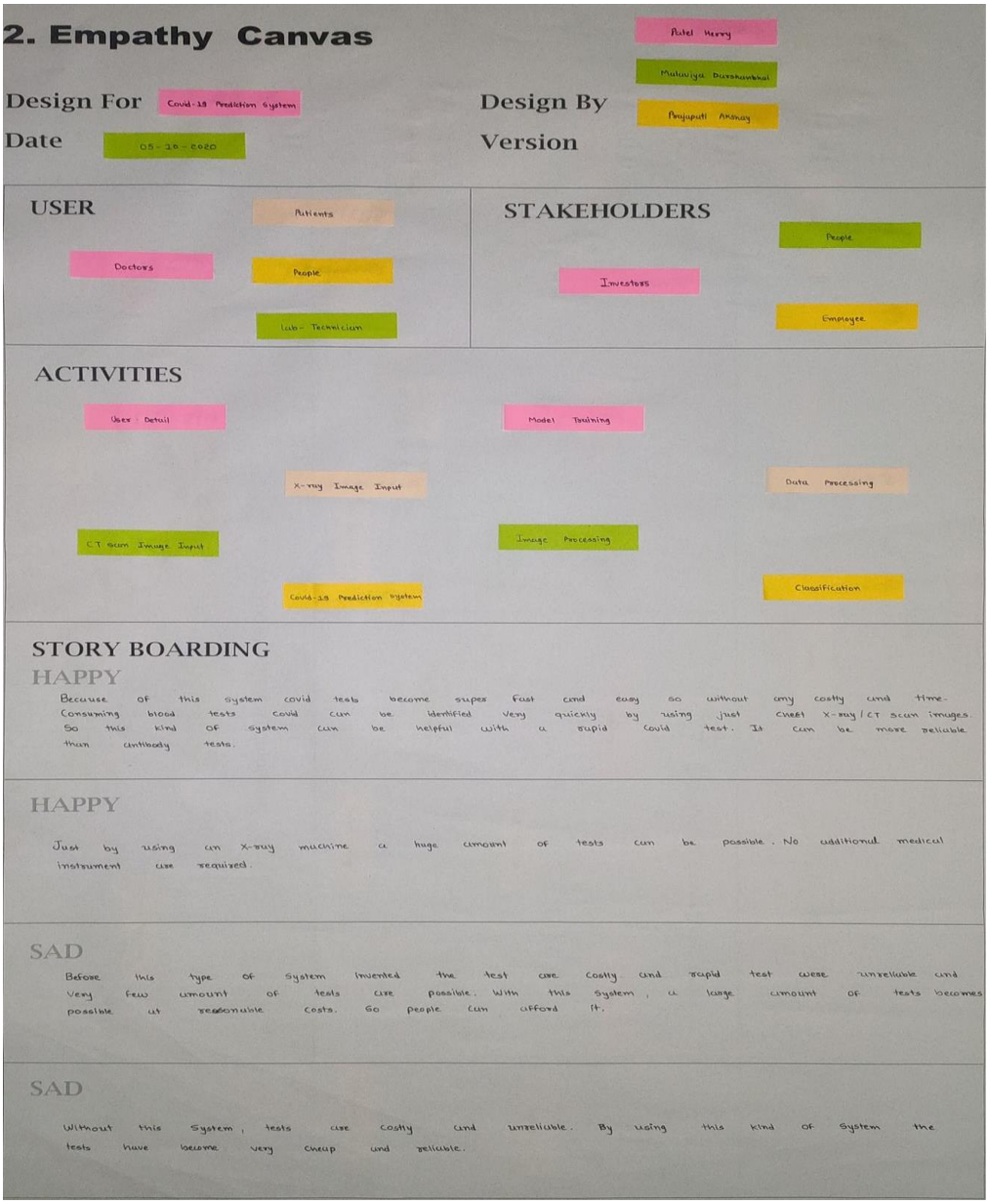
3.2.1 AEIOU Canvas

In AEIOU canvas, we have gone to many of the places and the people. We have gathered various new and essential data, bits of information, possibilities and functions. So, before we put any of our effort into making something; we thought that it would be great to discuss with the people first as what they are craving and what new they would like to see. Some of the information gathering, deeds are as shown in the below canvas:



3.2.2 Empathy Canvas

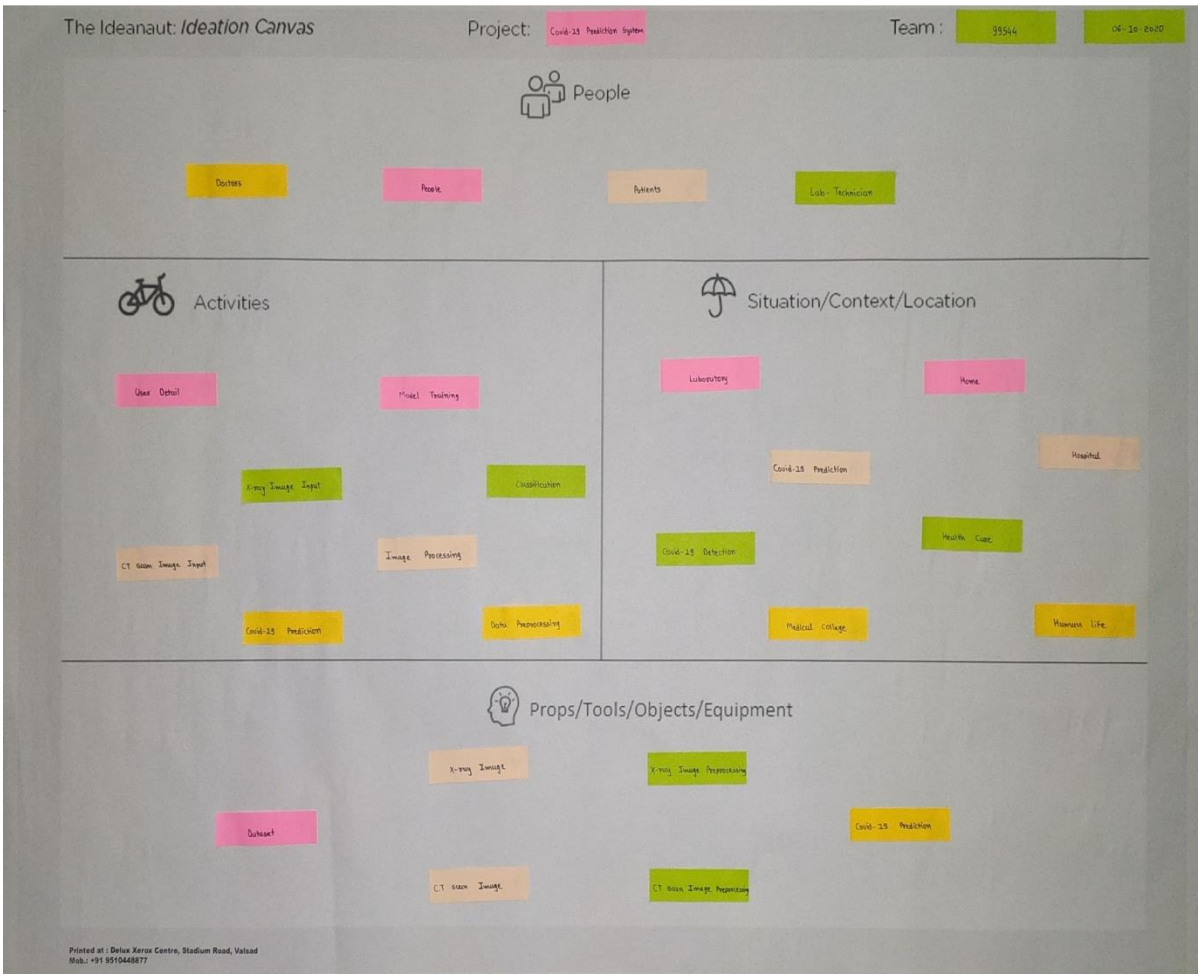
Empathy canvas includes factors like users, stakeholders, activities involved in a project and story boarding. In this canvas we have defined the sector for user of product and the stakeholders. We have also mentioned the activities that get involved in this project with a context storyboarding point describing the effect with and without the product. Such as, what people were missing, what they likely to have in a product, what changes this product might bring to the world.



3.2.3 Ideation Canvas

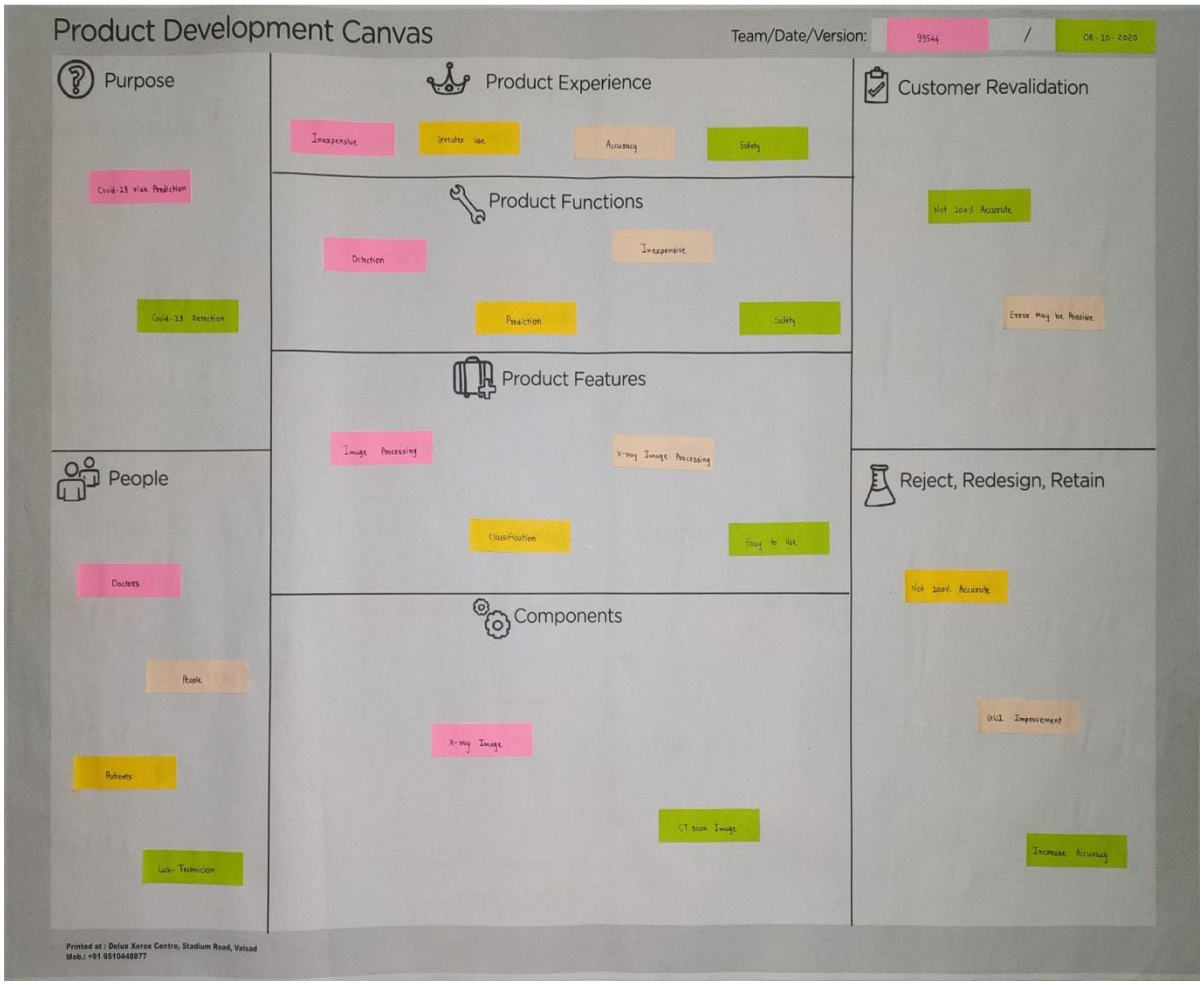
The ideation canvas covers the plot of the idea of the establishment and success of our project. It mainly contains the people who will be involved in the project or are related to the project; it may consist of users and stakeholders altogether.

Secondly, it tells about the activities which are to be done or performed, along with the location, context and situations as what is to be done, where it would be performed, why it is being carried out.



3.2.4 Product Development Canvas

This is the final outcome canvas of all the previous canvas. It is generated once the final idea is established, talks, debates, and discussions when done. This canvas describes the actual product in detail, along with the problems arising with it. Also, clearing out our views on what people think about it, what new they have to share, what wrong things did they find out, what is the primary purpose, features and functions.



Conclusion and Future Work

4.1 Conclusion

- By making this kind of system, We can head one step towards the peak of machine learning. We conclude that this is based on data and data processing; we required a robust database and ML Model to meet the user's requirement. There is more than one step to make this system, and properly following them will lead us to make the final product.

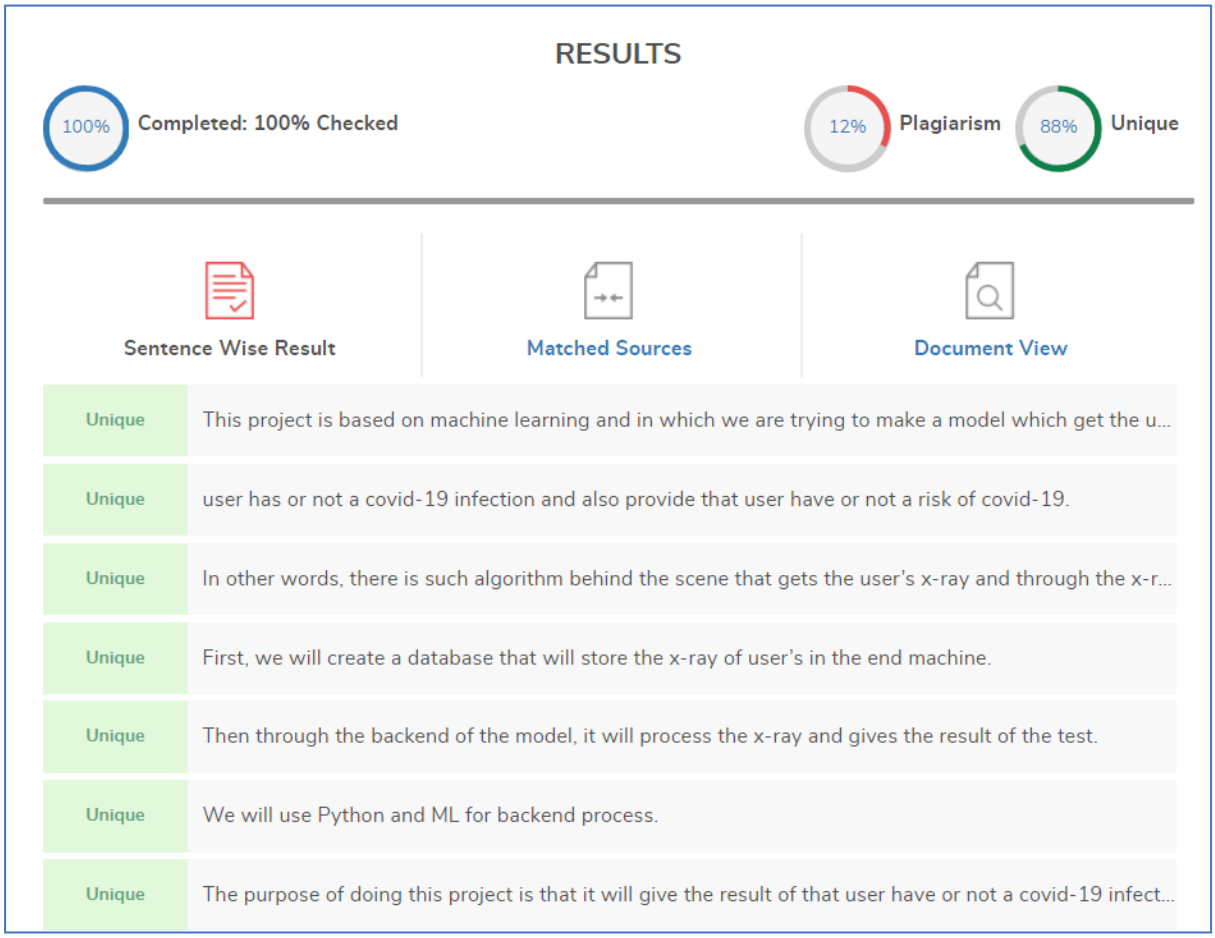
- We need user's permission to get access to their media as well as cookies. We need accurate classifier to classify the data and a more accurate algorithm to predict the result based on the input given.

4.2 Future Work

- The upcoming semester we will start implementing it on an actual term. There we require knowledge of Python and Machine learning. We started to learn machine learning and for the optimization of the database.

- In the next term, we will first make a base of the system and will decide to make it usable for users as much as possible. We will start it through the dataset and figure it out of the accuracy in different types of classification algorithms.

Plagiarism Report



References

- An update on COVID-19 for the radiologist - a British society of thoracic imaging statement,Clin, Radiol. (2020).
- Research Paper:COVID-19 Deep Learning Prediction Model Using Publicly Available Radiologist-Adjudicated Chest X-Ray Images as Training Data:Preliminary Findingds(Article ID 8828855).
- Kong W., Agarwal P.P. Chest imaging appearance of COVID-19 infection. Radiology: Cardiothorac. Imaging. 2020;2(1) doi: 10.1148/ryct.2020200028.
- Rodrigues J.C.L. An update on COVID-19 for the radiologist - a British society of thoracic imaging statement. Clin. Radiol 2020.
- Automatic Detection of Coronavirus Disease (COVID19) Using X-ray Images and Deep Convolutional Neural Networks. arXiv:2003.10849v1, 2020.
- Machine learning tutorials from youtube channel edureka.
- Python learning from youtube channel telusko.