

Covid analysis

anonymous marking enabled

Submission date: 12-Jun-2022 06:35AM (UTC-0400)

Submission ID: 1855210302

File name: 1RV20MC028_20MCA41_Ishita_4.docx (5.02M)

Word count: 10556

Character count: 60863

CHAPTER 1

INTRODUCTION

Introduction

The Coronavirus-2 and its associated illness, coronavirus disease COVID which was first detected in China in late 2019, pose the most significant public health threat in the past 100 years. This sickness has swept the globe like wildfire. The worldwide reaction to the outbreak has been patchy. Some governments reacted quickly and efficiently, while others failed terribly. Strict lockdowns are one means of restricting viral propagation; however, this comes with a high economic cost due to company closures and job losses, and there is a chance that the spread will restart once the lockdowns are lifted.

1.1 Project Description

To combat COVID, most governments have focused on developing smart lockdown tactics that use a number of technological solutions. Overall, early disease prevention initiatives and a mix of high and low-tech solutions enabled several nations (such as China and South Korea) to limit disease spread. These countries rely on passengers being monitored and quarantined. Human movement was severely hampered in the early days of the outbreak. People began to withdraw within their houses. However, it resulted in the closure of the transportation industry. Individuals with minimal risk were also asked to adhere to quarantine, and their physical presence was either monitored by state operatives or tracked via cell tower location data. Identified patients were kept in covid wards for about a fortnight, determined by the availability of health care facilities.

Tracking, screening, and forecasting existing and future infected cases, early detection and diagnosis of infection, medicine and vaccine research, and lowering the burden of healthcare professionals are all essential uses of AI in the battle against the covid pandemic. The propagation may potentially be handled as a data science problem. Non-traditional strategies for processing covid

data, such as data collection and interpretation, modelling, prediction, and data visualization and communication, are required to manipulate covid infection dynamics.

If a person is diagnosed with Corona virus and the diagnosis is validated, the next crucial step is to trace the person's contacts in order to prevent the According to the World Health Organization, the virus is spread mostly by saliva, droplets, or nasal secretions.⁵¹ Contact tracing is an vital public health method for preventing the transmission of SARS-Cov-2. Contact tracing is a way of recognizing and monitoring people who have recently been in contact with a Covid patient in order to keep the spread of the virus further. In general, the method detects the ill person 14 days after the exposure.

If implemented completely, this approach has the potential to break the chain of transference of the present new coronavirus and suppress the epidemic by increasing options for effective controls and assisting in the reduction of the contemporary pandemic's scope. In this regard, a method of digital contact tracing with a mobile application is available in many affected countries, leveraging a variety of technologies such as GPS, Bluetooth, and smartphone tracking data. The digital contact tracing method is substantially faster than the non-digital technique and may be completed virtually in real time. All these digital apps are intended to capture an individual's personal details, which will be processed by Machine Learning algorithms and techniques to identify people who are at risk to the new virus owing to their recent string of interactions.

1.2 Company Profile

Insurance, healthcare, banking and financial services, media, retail, and other industries are served by EXL Service, a worldwide analytics and digital solutions firm.
The company's headquarters are situated in New York, and it employs over 30,000+ people across various countries and continents.

Vikram Talwar and Rohit Kapoor co-founded EXL in the year 1999. Vikram Talwar served the company as CEO as well as a vice-chairman after the company was founded in year 1999. Rohit Kapoor is the company's CEO as well as the vice-chairman of the company at the moment. Kapoor has previously held the post of president of the corporation. Conseco bought EXL in August 2001 and ran it as a completely owned subsidiary. Conseco sold the firm to Oak Hill Capital Partners and FTVentures in 2002. On October 20, 2006, the company's shares began trading on the NASDAQ market under the ticker EXLS.

With the purchase of peer business Inductis, EXL embarked along the route of expansion through acquisitions shortly before its NASDAQ debut. This acquisition was made to diversify EXL's income stream; previous to the acquisition, business process outsourcing (BPO) accounted for 90% of EXL's revenue; after the acquisition, BPO accounted for 70% of EXL's revenue, with the rest coming from research and analytics and advisory services. Through the acquisition of OPI in 2011, EXL increased their distribution reach in Eastern Europe and Southeast Asia. EXL was said to have planned a massive purchase spree in the US healthcare business in 2013, in response to the Affordable Care Act's inception and demand that all Americans have health insurance. The healthcare market accounted for around 10% of EXL's revenue at the time of this study in 2013.

EXL set foot in Latin America in 2014 by the way of a strategic joint venture with CT&S, which paired EXL's technical prowess with CT&S's local understanding and brand image. EXL's escalating client base for Spanish-language services was also amplified by this joint venture. EXL expanded its healthcare analytics capabilities by acquiring SCIOInspire Holdings in 2018. EXL purchased Clairvoyant in 2021, gaining scale in data, artificial intelligence, and cloud engineering.

1.3 Dissertation Organization

The remaining dissertation is arranged as follows:

Chapter 1: Introduction

In first chapter the part includes the description of the project i.e., what is the need of this project, what are the tools required to develop this project and what will be the outcome of this project. Second part includes brief introduction about the company and domain in which this project is based on. Third part includes the technical features of this project.

Chapter 2: Literature Survey

It consists of four segments where the first part discusses about any literature survey done with respect to content related to the project. The second part is about the existing & proposed systems. The tools & technologies used for project will be third and the requirements of hardware & software for running the project is the fourth segment

Chapter 3: Software Requirement Specification

In third chapter the first part includes the brief introduction of the project, abbreviations used while writing the document and overview of the project. The second part includes functional requirements i.e., the modules used in the project and what will be the input, processing and output of each module. It also includes non-functional requirements, design constraints i.e., hardware limitations and software compliance.

47

Chapter 4: System Design

Chapter 4 is all about system design i.e., a bigger view of the project on a large scale by explaining the problem and specifying the modules

Chapter 5: Detailed Design Fifth chapter includes the description of the detailed design. It includes object modelling which contain class diagram for the project, dynamic modelling which contains use case diagram, sequence diagram and activity diagram, functional modelling which contain data flow diagrams. These diagrams are included to describe the flow of the project.

Chapter 6: Implementation

This chapter provides code snippets related to the project & an explanation to implementation of the modules along with screenshots is given.

Chapter 7: Testing

Seventh chapter includes testing. It includes test cases for each module. This chapter include one failed test case.

Chapter 8: Conclusion

This chapter tells the whole summary of the project, comparison of existing system and proposed system and what all new things have been implemented.

Chapter 9: Future Enhancements

This chapter gives any future works/improvements that can be implemented in the future to improve and enhance the overall efficiency of the project.

**52
CHAPTER 2****LITERATURE REVIEW****2.1 Literature Survey**

Individuals' Rustam, F., Reshi, A., and Mehmood, "COVID-19 Future Forecasting Using Supervised Machine Learning Models", Machine learning (ML)-based forecasting techniques have demonstrated their use in predicting perioperative outcomes and improving decision-making about future actions. Many application fields that required the detection and prioritization of adverse aspects for a threat have long used machine learning models. The research shows that machine learning algorithms can predict the number of patients who would be afflicted by COVID-19, which is now regarded as a possible threat to humanity. Each of the models makes three types of predictions: the number of newly infected cases, the number of fatalities, and the number of recoveries the next ten days. The study's findings show that using these strategies in the current COVID-19 pandemic scenario is a promising mechanism [1].

"Regression Analysis of COVID-19 Using Machine Learning Algorithms," with the guide of utilizing Ekta Gambhir, Ritika Jain, Alankrit Gupta, and Uma Tomer (2020). The original Coronavirus keeps on being an approaching danger to general wellbeing, and dropping down in records as one of the worldwide's most terrible pandemics is going. The expectation is to offer donors an additional skill of the manner in which unique Machine Learning designs is most likely used in genuine worldwide situations. This concentrates moreover looks at the overarching design or test of Covid-19 transmission in India, further to the overall numbers. This analyze utilizes records from the Indian Ministry of H(health)F(Family)Welfare to give several advancements and styles situated in different components of the worldwide. The records for this inspect become accumulated over a span of 154 days. The kind of dataset is the reason relapse assessment become chosen for the current day inconvenience proclamation. Since the dataset is relentless, relapse assessment is top of the line ideal for foreseeing a constant laid out factor from a fixed of fair-minded factors [2]. M. A. Sarwar, N. Kamal, W. Hamid, and M. A. Shah (2018, September). "In Healthcare, Prediction of Diabetes Using Machine Learning

Algorithms." In a variety of industries, machine learning techniques are utilised to do predictive analytics over massive data.¹⁶ Predictive analytics in healthcare is a difficult task, but it can ultimately assist practitioners in making timely decisions regarding a patient's health and treatment based on huge data. This study examines predictive analytics in healthcare, and it employs six different machine learning algorithms.¹⁸ A dataset of a patient's medical record is gathered for the experiment, and six different machine learning algorithms are applied to the dataset.³⁸ The performance and accuracy of the algorithms used are compared and reviewed. The various machine learning approaches employed in this study are compared. A new study indicates which algorithm is superior for diabetes prediction. The goal is to use machine learning techniques to assist doctors and practitioners in the early detection of diabetes [3].

D. Dahiwade, G. Patle, and E. Meshram (2019, March). "Utilizing a Machine Learning Approach to Design a Disease Prediction Model." People are inclined to a whole parcel of infirmities due to their living way of behaving and the US of the climate. Thus, foreseeing contamination at an early recognition will turn into a fundamental errand. Nonetheless, logical clinical specialists figure out it hard to make specific expectations dependent generally upon side effects and side effects. The most troublesome undertaking is productively foreseeing contamination. To treatment this issue, measurements mining plays a fundamental component in disease forecast. Clinical innovative information produces a magnificent measure of insights every year. The appropriate appraisal of clinical insights has been profited from early impacted character care because of the sped up measure of insights development withinside the clinical and medical services fields. Information mining uncovers stowed away examples the use of sickness measurements. measurements covered in a heap of clinical measurements We predominant a major sickness forecast dependent totally upon the impacted person's side effects and side effects. For compelling ailment expectation, we rent the K-Nearest Neighbor (KNN) and Convolutional Neural Network (CNN) framework concentrating on calculations. An assortment of disease side effects and side effects is needed for sickness expectation. For an exact visualization, the person's dwelling conduct and exam measurements are viewed as in this exquisite ailment expectation. The CNN set of guidelines has an exactness of 84.5 percent in exquisite

sickness forecast, this is higher than the KNN approach. Furthermore, KNN makes some

higher memories and memory need than CNN. Subsequent to anticipating exquisite disease, this instrument can compute the gamble of rich sickness, showing whether or presently as of now not the gamble is lower or higher [4].

S. Subudhi, A. Verma, A. B. Patel, C. C. Hardin, M. J. Khandekar, H. Lee, et al (2021). In COVID-19, researchers compared machine learning methods for predicting ICU admission and death. Machine learning methods could help clinicians identify high-risk individuals because forecasting COVID-19's course is difficult. The results of 17 18 machine learning algorithms for predicting ICU admission and mortality in COVID-19 patients are compared in this study. We generated and internally validated models 11 using COVID-19 patient data from the Mass General Brigham (MGB) Healthcare 10 database, and we further evaluated them using temporally diverse individuals who presented to the ED between May and August 2020 (n = 1711). Ensemble-based models 11 outperform traditional models, according to our findings. COVID-19 predicts 5-day 10 ICU admission and 28-day death better than other model types. CRP, LDH, and O2 saturation were important for ICU admission models, but the most relevant factors for predicting mortality were eGFR 60 ml/min/1.73 m² and neutrophil and lymphocyte percentages. Such models could aid clinical decision-making in the event of future infectious disease epidemics, such as COVID-19 [5].

"A Survey of Mathematical, Machine Learning, and Deep Learning Models for COVID-19 Transmission and Diagnosis," by 40 C. C. John, V. Ponnusamy, S. K. Chandrasekaran, and R. Nandakumar, 2021. COVID-19 is a life-threatening virus with 30 worldwide ramifications. Drugs and vaccines have yet to be found because the disease 6 is caused by a new coronavirus with unknown gene information. For the time being, 30 illness spread analysis and prediction using a mathematical and data-driven model will be extremely useful in initiating prevention and control measures such as lockdown and quarantine. Various mathematical and machine-learning methods have been suggested 6 for assessing and predicting the spread. For a given circumstance, each model has its own set of constraints and benefits. To provide greater insight, this paper covers state-of-the-art mathematical models for COVID-19, including compartment models,

Covid Patient Tracing and Diagnosis System

JUNE 2022

statistical models, and machine learning models, in order to find an acceptable model for disease spread analysis. Furthermore, a precise COVID-19 diagnosis is another critical step in identifying the infected person and preventing further dissemination. Because the virus is spreading quickly, there is a need for an automated diagnosis system that can manage a huge population. Diagnostic mechanisms based on deep learning and machine learning will be more appropriate for this purpose [6].

R. Sujath, J. M. Chatterjee, and A. E. Hassanien (2020). "A COVID-19 pandemic estimating form the use of device acquiring information on in India." With realities examination and realities mining, records and report time helps within the dynamic system principally founded absolutely on notable realities. The amount of realities to be had is colossal, making gathering records and removing an intriguing example from it a hard endeavor. The cutting edge realities on affirmed, recuperated, and downfall in India throughout a drawn out time span supports foreseeing and estimating the near predetermination. The rendition's exactness will be progressed through incorporate additional abilities including numerous clinics, the impacted individual's resistant framework, the patient's age, orientation, and the endeavors taken to struggle the unfurl of the ailment of the infection, etc, so one can offer total records. Starting today, it is imperative that yards to hold be severe and cautious in nature so one can address this fundamental condition through friendly partition, lockdown, time limit, quarantine, and seclusion. We can likewise furthermore complete that the MLP strategy produces higher expectation results than the LR and VAR procedures using WEKA and Orange through looking through on the expected qualities and assessing them to models from John Hopkins UniversityFootnote11 realities. We can utilize profound acquiring information on cycles to gauge time assortment realities within the predetermination to create higher forecasts [7].

³⁴ Gupta, Rajan, Manan Bedi, Prashi Goyal, Srishti Wadhera, and Vaishnavi Verma are the people of the (2020). "Examination of COVID-19 Tracking Tool in India: Case Study of Aarogya Setu Mobile Application" Different states of their different countries are quickly developing COVID-19 checking stuff or contact-following applications. The Aarogya Setu instrument, progressed with the guide of utilizing the Indian specialists, is the worry of this analyze. It is a telephone programming progressed with the guide of involving the Ministry of Health as a piece of the E-Governance

endeavor to tune and show Indian residents within the battle towards the unfurl of COVID-19. The look's at will probably learn about the apparatus' severa abilities and to give assorted measurements innovation ideas utilized within the product, notwithstanding the device's value in taking care of the proceeding with pandemic. At the point when an individual is near a COVID-19 excited individual, the application utilizes Bluetooth and GPS period to tell them. To investigate COVID-19 unfurl in India, the product utilizes various Data Science ideas comprising of order, alliance rule mining, and bunching. The look at moreover shows possible moves up to the program, comprising of utilizing AI and PC creative and judicious to go over COVID-19 patients. Versatile period subject matter experts, measurements innovation experts, logical specialists, wellbeing related forefront representatives, public overseers, and specialists officials could all partake in the examination [8].

J. Berglund, J. Berglund, J. Berglund, J. Berglund "There's an App for That: COVID-19 Tracking" Vice President Mike Pence advised the snap in March 2020, sooner than COVID-19 laid proclaim to the United States, at the Covid pestilence, which on the time turned into a threat to travelers on board the Grand Princess journey convey secured off the bank of California [1]. Cases had ascended in Iran, Italy, and Spain on the time, but the disorder turned out to be regardless a spine chiller to most extreme Americans. At the point when mentioned roughly a possible episode in Washington, D.C., Vice President Mike Pence alluded the subject to Dr. Anthony Fauci, top of the National Institute for Allergy and Infectious Disease and an individual from AmericaA Task Force. Fauci expressed that each and every individual who inspected great for the affliction could be exposed to contact following, which he characterized as "the overall population wellness weapon" for halting episodes. The World Health Organization proclaimed a global pandemic now at this point not extended later, and the infection began to unfurl all through the country [9].

C. Sandeepa, C. Moremada, N. Dissanayaka, T. Gamage, and M. Liyanage. Sandeepa, C., Moremada, C., Dissanayaka, N., Gamage, T., and Liyanage, M. (2020, September). "A System for Tracking Social Interactions and Predicting Potential COVID-19 Patients" Allowing for proper social distancing and isolation or confinement to identify possible patients is the greatest strategy to prevent the virus from spreading quickly. Because the virus has a 14-day incubation period, it's critical to track all social

interactions during that time and place such potential patients in social isolation. For the time being, however, proper social interaction tracking systems and patient prediction algorithms based on this data are lacking. This study looks at how to track users' social interactions and forecast infection risk based on such interactions. We began by developing a social contact tracking system based on BLE (Bluetooth Low Energy) and GPS. Then, using the obtained data, we created an algorithm to estimate the likelihood of becoming infected with COVID-19. Finally, a system prototype with a mobile app and a web monitoring tool is implemented. In addition, we used a graph-based model to simulate the system and assess the behaviour of the proposed algorithm, which confirms that self-isolation is significant in slowing illness progression [10].

R. Mallik, D. Sing, and R. Bandyopadhyay (2020). "GPS Tracking App for Police to Track COVID-19 Patients for Safe Distancing." Traffic congestion is unavoidable in places with high population density, such as Bengaluru and Mumbai. According to the TomTom traffic index for 2019, Bengaluru is the world's most congested city, with Mumbai in fourth place. Transporting a COVID-19 patient between two locations quickly during a traffic jam could be difficult. The position information of an ambulance transporting COVID-19 patients may be valuable to traffic cops in isolating the ambulance from crowded areas and directing them to clear roads, allowing the patients to be relocated more quickly [11].

S. Saha, S. Maity, T. Bhattacharyya, T. Bhattacharyy (2021, February). "An Application Based on GPS Tracking and QR Code Scanner to Assist Health Workers and Ordinary Citizens During Covid-19" The pollution expense unleashed obliteration over the world, first influencing European global areas lastly the United States. Thus, on this work, we suggest a product program utility that gathers individual records to process the risk principally based absolutely at the merits of a re-did overview, provoking the individual to confirm assuming they might be encountering any medicinally settled side effects, after which sends it to the area civil center. Thus, the requirement for logical workers to move house to house on a customary premise, gambling with tainting, can be diminished. This product program comprises of GPS following, a danger stage forecast, a contamination diagram, and the ability to compute logical earnestness, among various things. Basically, our gadget is intended to direct

consistently representatives and go to places in high-peril circumstances. This proposed period is utilized to confine capacity suggestive and asymptomatic clients from getting into public spots comprehensive of division shops and markets in area of interest zones. The product program will create a totally extraordinary QR code for each individual, for you to show their all around certainly worth basically founded absolutely on a mandatory study to conclude whether or presently no longer they might be impacted. The QR code might be examined to benefit get right of section to the site [12].

2.1 Existing and Proposed System

2.1.1 Problem statement

The health sector is always in need of new technology, whether it's in the shape of gear or software that can do important jobs. When Covid struck, there was no means to fight the disease or even tell if someone was sick or not, which was a critical period for the health domain. During the Covid-19 epidemic, the number of persons infected worldwide increased dramatically, with the majority of those infected coming into touch with people who were already infected or showing symptoms.

2.1.2 Scope of Project

Machine learning is a cutting-edge method with numerous applications in prediction. For the COVID pandemic, this technique must be used to detect high-risk patients, persons who have come into touch with positives, and other abnormalities. Rather than a standard overtly calculation-based method, it gives accurate and valuable features. Furthermore, this technique is useful in predicting the danger of COVID in healthcare during this crisis. Machine learning considers risk factors such as age, location, and climate.

2.1 Tools and Technologies

Python

33

Python's design philosophy stresses code readability, as seen by the extensive usage of whitespace. All programming paradigms are supported, including structured programming (particularly procedural), object-oriented programming, and functional programming.

Jupyter

8

JupyterLab is the latest web-based interactive development environment for notebooks, code, and data. Its flexible interface allows users to configure and arrange workflows in data science, scientific computing, computational journalism, and machine learning. A modular design invites extensions to expand and enrich functionality. Jupyter Notebook can hook up with several kernels to permit programming in numerous languages. A Jupyter kernel is a program chargeable for handling varied forms of requests (code execution, code completions, inspection), and providing a reply. Kernels consult with the opposite components of Jupyter victimization ZeroMQ, and therefore will be on a similar or remote machines. not like many different Notebook-like interfaces, in Jupyter, kernels aren't aware that they're connected to a particular document, and might be connected to several shoppers at once. typically kernels permit execution of solely one language, however there are one or two of exceptions.[citation needed] By default Jupyter Notebook ships with the IPython kernel. As of the 2.3 unharness (October 2014), there are forty nine Jupyter-compatible kernels for several programming languages, together with Python, R, Julia and Haskell. A Jupyter Notebook can be regenerate to variety of open commonplace output formats (HTML, presentation slides, LaTeX, PDF, ReStructuredText, Markdown, Python) through "Download As" within the internet interface, via the nbconvert library or "jupyter nbconvert" instruction interface in a very shell. To alter image of Jupyter notebook documents on the web, the nbconvert library is provided as a service through

NbViewer which may take a URL to any publically out there notebook document, convert it to hypertext markup language on the fly and show it to the user. The notebook interface was additional to IPython in the 0.12 unharness (December 2011), renamed to Jupyter notebook in 2015 (IPython 4.0 is Jupyter 1.0). Jupyter Notebook is analogous to the notebook interface of different programs love Maple, Mathematica, and SageMath, a procedure interface vogue that originated with Mathematica within the 1980s.[19] Jupyter interest overtook the recognition of the Mathematica notebook interface in early 2018. JupyterLab may be a newer computer program for Project Jupyter. It offers the building blocks of the classic Jupyter Notebook (notebook, terminal, text editor, file browser, wealthy outputs, etc.) in a very versatile user interface. the primary stable unharness was proclaimed on February 20, 2018.

CSV

15

A CSV is a comma-separated values record, which lets in records to be stored in a tabular format. CSVs seem like a garden-range spreadsheet however with a .csv extension. CSV documents may be used with maximum any spreadsheet program, together with Microsoft Excel or Google Spreadsheets. They fluctuate from different spreadsheet record kinds due to the fact you may best have a unmarried sheet in a record, they cannot keep cell, column, or row. Also, you can't now no longer keep formulation on this format.

Pygal

Pygal is an open-supply information visualization library in Python. It is one of the exceptional python libraries to create distinctly interactive plots and charts for unique datasets. Also, it lets in you to down load your visualizations in SVG (Scalable Vector Graphics) or PNG (Portable Graphics Format) for more than one programs and personalize it accordingly.

DBSCAN

Grouping evaluation or unquestionably Clustering is essentially an Unsupervised perusing technique that partitions the measurements components into various one-of-a-kind clumps or associations, to such an extent that the measurements components withinside similar associations have comparative homes and measurements

components in unmistakable associations have particular homes in some sense. It comprises of numerous unmistakable techniques put together absolutely for the most part with respect to differential advancement. For example K-Means (distance among components), Affinity proliferation (chart distance), Mean-shift (distance among components), DBSCAN (distance among closest components), Gaussian blends (Mahalanobis distance to focuses), Spectral grouping (diagram distance) and so on. Essentially, all bunching procedures utilize a similar strategy for example first we compute similitudes and afterward we use it to group the insights components into associations or clusters. Here we're equipped for cognizance on Density-based absolutely totally spatial grouping of utilizations with commotion (DBSCAN) bunching technique. Bunches are thick locales withinside the measurements space, isolated by means of districts of the lower thickness of components. The DBSCAN set of strategies is put together absolutely totally with respect to this natural idea of "bunches" and "commotion". The key thought is that for every component of a group, the organization of a given range needs to envelop basically a base measure of components.

Pandas

Pandas is a software program library written for the Python programming language for statistics manipulation and evaluation. In particular, it gives statistics systems and operations for manipulating numerical tables and time series. It is unfastened software program launched below the three-clause BSD license. The call is derived from the time period "panel statistics", an econometrics time period for statistics units that consist of observations over more than one time durations for the equal individuals. Its call is a play at the phrase "Python statistics evaluation" itself. Wes McKinney began out constructing what might grow to be pandas at AQR Capital whilst he changed into a researcher there from 2007 to 2010. Pandas is especially used for statistics evaluation

and related manipulation of tabular statistics in Dataframes. Pandas lets in uploading statistics from numerous report codecs which include comma-separated values, JSON, Parquet, SQL database tables or queries, and Microsoft Excel. Pandas lets in numerous statistics manipulation operations which include merging, reshaping, selecting, in addition to statistics cleaning, and statistics wrangling functions. The improvement of

pandas brought into Python many similar functions of operating with Dataframes that have been installed within the R programming language. The pandas library is constructed upon some other library NumPy, that is orientated to effectively operating with arrays in preference to the functions of operating on Dataframes.

Numpy

NumPy can be a library for the Python programming language, as well as working with huge, multi-faceted clusters and grids, joined with an enormous arrangement of undeniable level numerical capacities to perform on such exhibits. . NumPy's connection, Numeric, was initially made with the assistance of Jim Hugunin's work with commitments from totally various engineers. In 2005, Travis Oliphant made NumPy utilizing the coordination of Numarray contender capacities into Numeric,²⁸ with significant changes. NumPy is an open source programming framework program and has various benefactors. NumPy is an undertaking upheld by NumFOCUS in charge matters. NumPy is focused on Python's CPython reference execution, for example an enhanced bytecode translator. Math calculations composed for this Python model frequently run a horrendously related stack more slow than incorporated identical calculations because of absence of compiler enhancement. NumPy takes care of the issue of gradualness with the assistance of four-layered cluster mining, abilities and administrators that function admirably on exhibits; utilizing them joins code altering, predominantly inward circles, utilizing NumPy. Mining NumPy in Python gives MATLAB happiness for however long they are deciphered, permitting anybody to compose quick applications as long as most tasks are performed on exhibits or grids and not on aimless. By examination, MATLAB has a ton of extra toolboxes, including Simulink, while NumPy itself is coordinated with Python, a stylish and extensive extra programming language. Moreover, extra Python programs are accessible; SciPy can be a library with extra MATLAB-like capacities, and Matplotlib is a plotting bundle that gives MATLAB-like plotting abilities. Inside, each MATLAB and NumPy depends on BLAS and LAPACK for unpracticed arithmetical

calculation. Python ties of the broadly utilized and required local scratch pad library OpenCV utilizes NumPy exhibits to store and perform measurements. Since pictures with more than one channel are in a real sense drawn as third-layered exhibits, ordering,

falling, or overlaying completely various clusters are seriously unpracticed ways to deal with creating extraordinary pixel consents for a photograph. NumPy outlines as day to day details in OpenCV for shots, mined execution scores, portion separating and more that significantly improve on programming and troubleshooting processes.

24

2.1 Hardware and Software Requirements

2.1.1 Hardware requirements

- Processor: Intel® Core™ i3 processor, Ryzen 3 or above
- Hard Disk: 500 GB or above
- Memory (RAM): minimum 2GB
- Ethernet connection (LAN) or wireless adapter (WiFi)
- Graphical Processing Unit (GPU) 1050 Ti or above

24

2.1.2 Software requirements

- Operating System – Window 8-10, Linux 18.0
- Browser – Chrome v101.0.4951.54, Microsoft Edge
v100.0.1185.39
- Programming languages – Python 3.6.0 or above
- Visualization tools – Seaborn 0.11.2, Matplotlib 0.11.0

25
CHAPTER 3

**SOFTWARE REQUIREMENT
SPECIFICATION**

3.1 Introduction

A software program necessities specification is a documentation that includes an in-depth description of the capability of the machine. The Software Requirement Specification (SRS) level of software program improvement is wherein the necessities of the machine beneath attention are documented with a purpose to lay the muse for the software program improvement activities. Consistency, completeness of all important necessities and their definitions, and correctness are all traits of a valid SRS document.

This chapter gives an in-depth description of the proposed machine in addition to its predicted results. The machine's definitions, widespread product description and functions, person characteristics, purposeful and non-purposeful requirements, and machine hassle constraints are all protected in the specification.

3.1.1 Definitions, Acronyms and Abbreviations

np – Numpy

pd – Pandas

DBSCAN – Density Based Clustering

CSV – Comma Separated Values

3.2 General Description

The work-flow of the system is presented in this section, which includes the product description, product functions, system user characteristics, general constraints, and assumptions.

3.2.1 Product Description

The product will be able to help medical professionals to find a particular covid patient and analyze the gravity of the situation by checking for other individuals who came in close contact with them, to further widen their search they will be able to check which one of them were already infected and who wasn't, thus in such cases they will be able to prevent the disease from spreading and also by taking the further necessary measures.

Let's check in-depth, the product will choose a random id of an infected individual particularly and plot it on a graph, further the id and date of the infected individual will be used to locate which cluster he/she belongs to which further helps to narrow down the search and one will get the list of people who came in close contact to them, which later will be used to check which of them were already infected and which ones aren't, lastly it will also predict the probability of them being infected after they directly or indirectly came in contact with the patient/patients.

3.2.2 Product Functions

Locating covid patients – How well can the developed system locate an individual.

Locate people who came in contact with the patient – The system should be able to locate the other people who came in contact with the patient within the given distance or radius.

Check if other people located were infected or not – To check for other patients among the clusters being formed.

Probability of person getting infected – To find the probability of the non-infected person of getting infected after coming in contact to the patient

3.2.3 User Characteristics

The system is used by the people who are medical professionals who provide medical and healthcare services. The user must be familiar with the patient details.

3.2.4 Assumptions and Dependencies

The user who will use this application should have the patient details with them along with an efficient device to smoothly run ML algorithms.

3.3 Functional Requirements

3.3.1 Locating covid patients

Purpose: To locate the infected patient

Input: Latitude and Longitude

Function: Search for a particular patient

Output: Maps the searched patient's coordinates either in map or on the cluster

3.3.2 Infected Patient Diagnosis

26

Purpose: To locate people who possibly came in contact with the infected person and inform them

Input: 2 Latitudes and Longitudes

Function: Search through the mentioned coordinates and finds other people

Output: List of the people

3.3.3 Identification of probably infected patients

Purpose: To check whether the people located were infected themselves or not for further processing

Input: Person's name and id

Function: Checks whether the person is healthy or infected

Output: Gives details of the patient

3.3.4. Probability of people being infected

Purpose: To find the probability or percentage of the non-infected getting Covid

Input: Person's id

Function: Checks the level of the symptoms

Output: Gives a percentage value

3.4 Non-Functional Requirements**Performance –**

Performance in terms of accuracy in finding/locating patients based on latitude and longitude.

Scalability –

Whether the system will be able to adapt to a different dataset with a different size.
Should be able to handle 5000 sets of data up to 5000000 sets of data.

Accuracy –

The probability of the system being accurate is 96%.

3.5 Design Constraints

The system should have anaconda/python installed in it.

The system should have Jupyter notebook to run the ML code.

25
CHAPTER 4

SYSTEM DESIGN

System design provides an overview of the system's architecture, including how the system is connected internally, how work flows within the system, and the concept of complete system components.

4.1 Architectural design

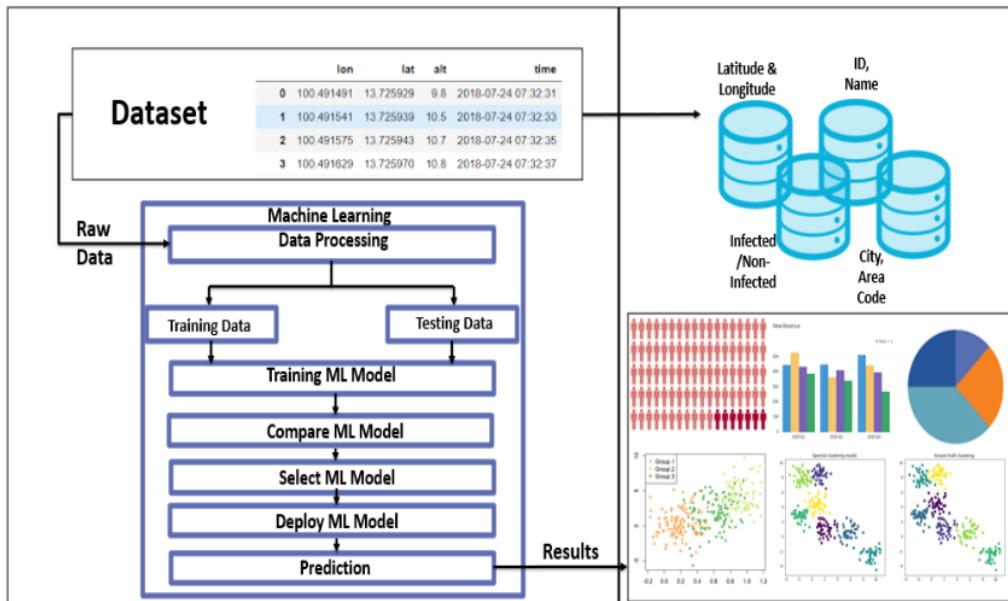


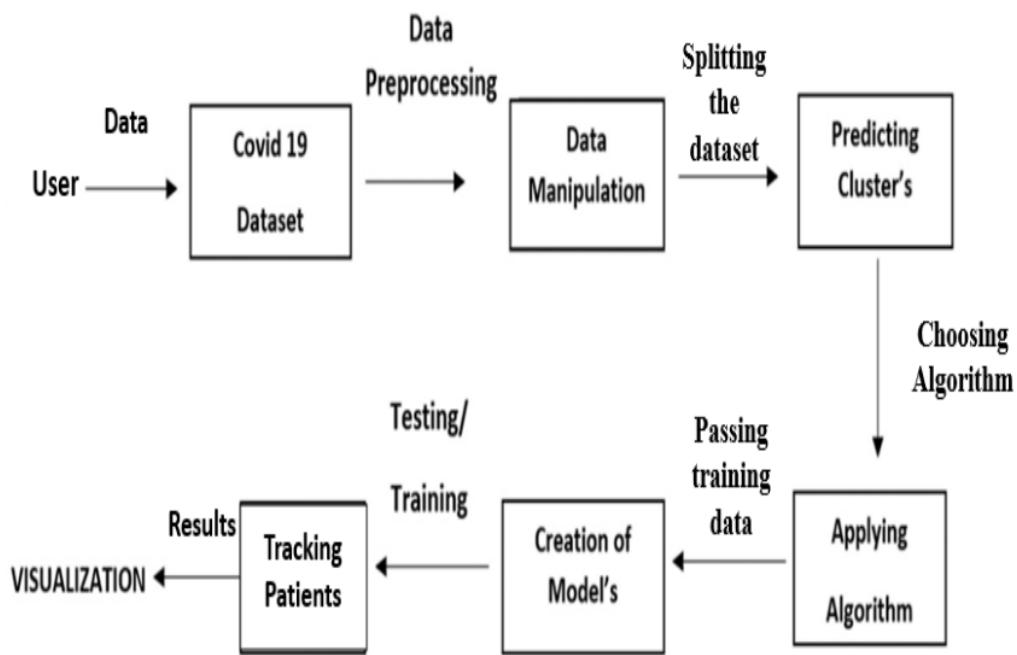
Fig 4.1 Architecture Diagram

Figure 4.1 shows the architecture of the proposed structure and how the system is going to be developed. It also shows the different components one to work with to completely built a Contact tracing model.

4.1.1 Problem Statement

The health sector is always in need of new technology, whether it's in the shape of gear or software that can do important jobs. When Covid struck, there was no means to fight the disease or even tell if someone was sick or not, which was a critical period for the health domain. During the Covid-19 epidemic, the number of persons infected worldwide increased dramatically, with the majority of those infected coming into touch with people who were already infected or showing symptoms.

4.1.2 Block Diagram



36
Fig 4.2 Block Diagram

Fig 4.2 shows the flow of the request from a user to the providing of patient details. When the service is requested by the user, it is accepted through the model built and the task is further performed to search who is nearest to the location of the patient.

4.1.3 Module Specification

Quality Assurance for Patient Tracking

Tracking Patient

Tracking Patient's whereabouts

Location patients using latitude and longitude

Quality Assurance for Infected Patient Diagnosis

Sorting individuals who came in contact with patient

Matching dates with that of patient

Checking whether other people are from the same cluster only

Quality Assurance for Identification of probably infected patients

List of patients who are infected

List of patients not infected

Probability of the non-infected patient of getting infected

Quality Assurance for Probability of person getting infected

Percentage value of the non-infected getting infected

The percentage value should be calculated depending on:

- Whether came in contact with patient or not
- The range of symptoms

4.2 Context Diagram

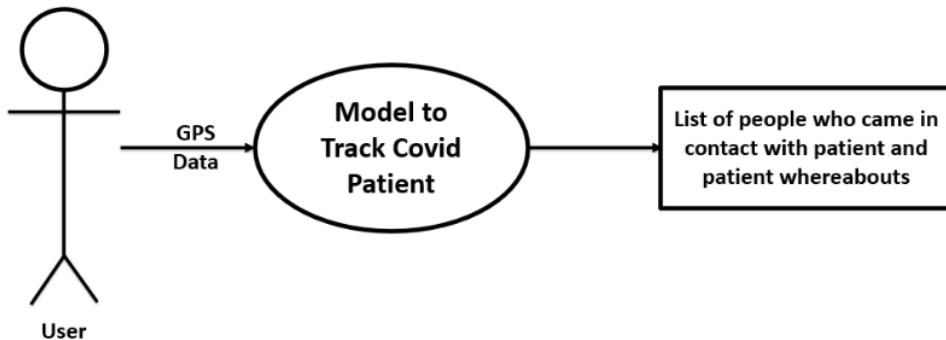


Fig 4.3 Context Diagram

Fig 4.3 shows the flow between the components of the system. This is a high-level diagram that shows the internal process that takes place when a user requests a service. The model provides service based on the GPS data passed. Once the data is available the clusters are formed and the model searches for the patient and other individuals.

CHAPTER 5

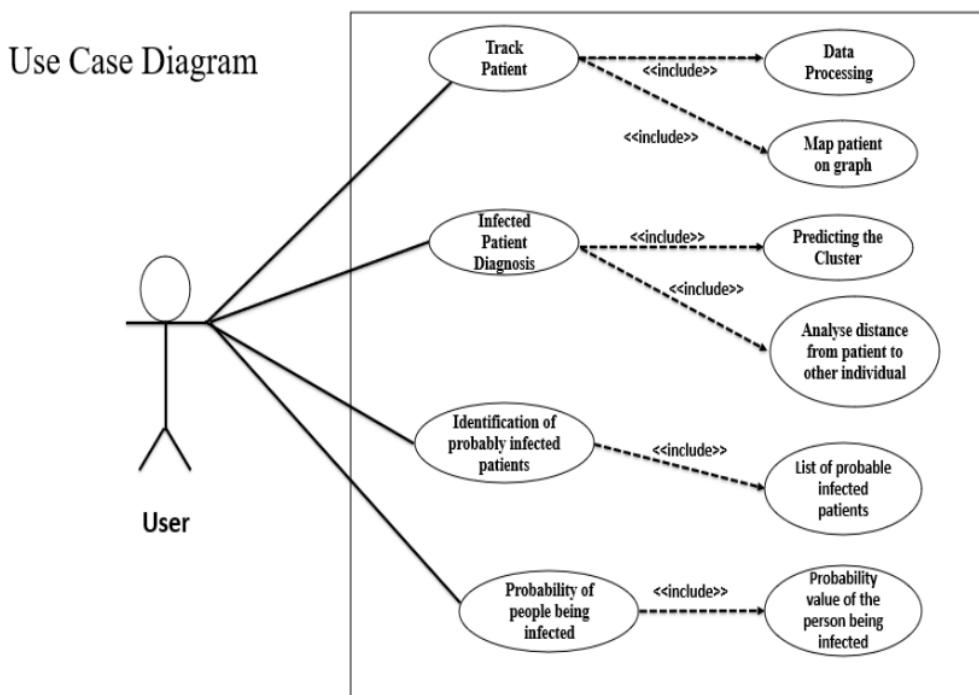
DETAILED DESIGNS

5.1 System Design

System design administer a sketch and analysis of the system's architecture, including how the system is connected internally, how work flows within the system, and the concept of complete system components.

5.1.2 Dynamic Modelling

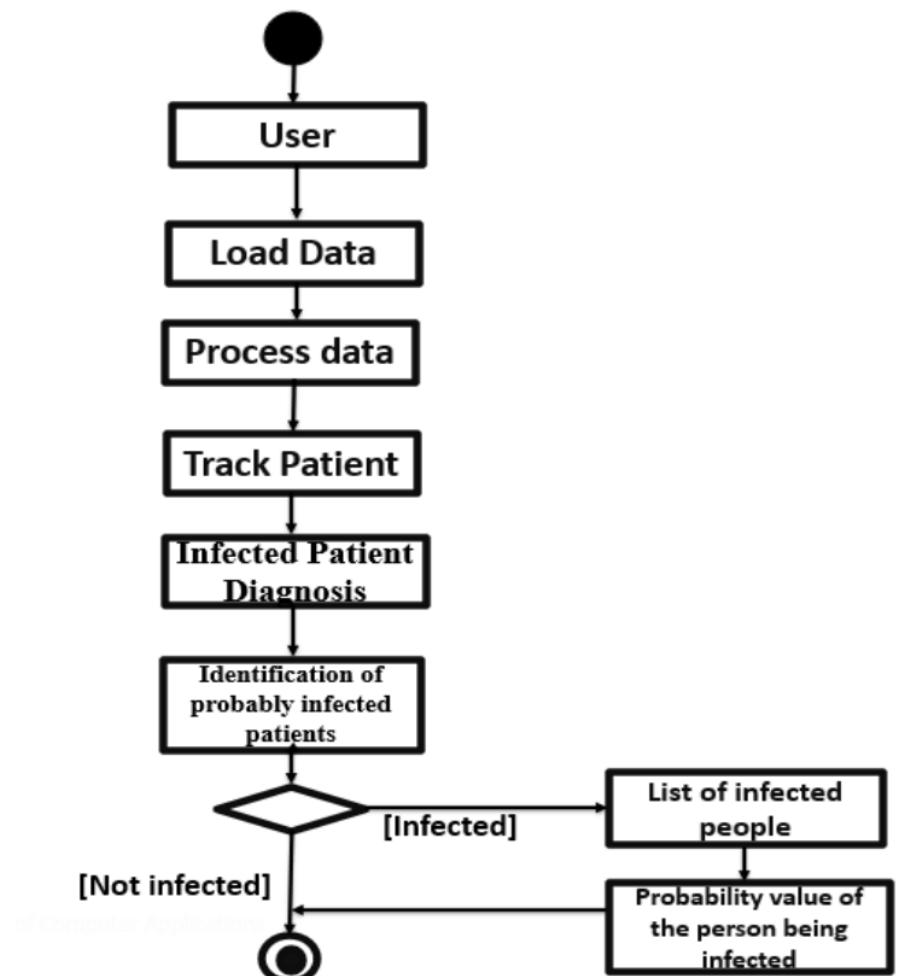
5.1.2.1 Use Case Diagram



46
5.1 Use-case Diagram

Fig 5.1 shows the actor, User, present in the system. The actor has specific functions to perform. The user tracks patients using the GPS data . They further diagnose the infected patient to check for people who came in contact with them. After which it checks among the other individuals who are infected and who aren't.

5.1.2.2 Activity Diagram



5.2 Activity Diagram

Fig 5.2 describes the activities that take place in the system. The activity starts when the User loads the data. After which they form clusters using ML algorithm, once the clusters are formed a random id is chosen and further used for processing.

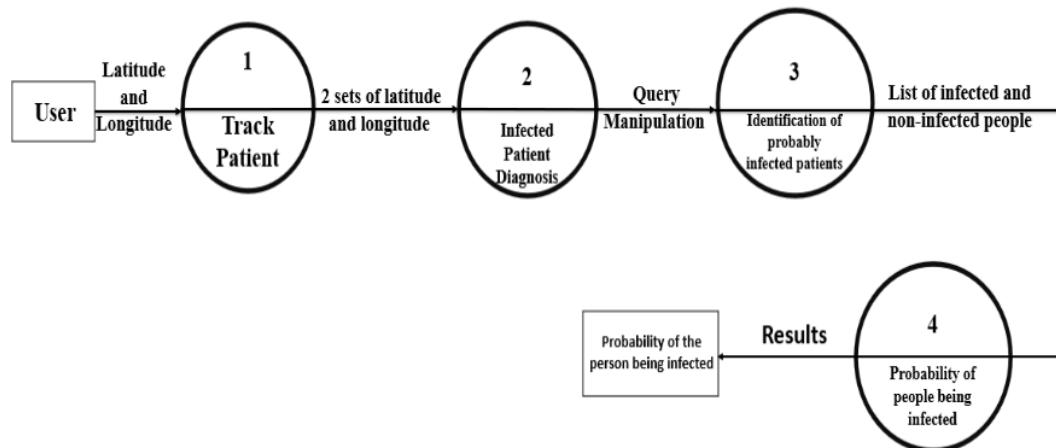
5.1.2 Functional Modelling

5.1.2.1 Data Flow Diagram

21
Fig 5.3 DFD Level-0



The fig. 5.3 shows the highest level of the DFD. The process involved when a user requests a patient to be tracked is shown.



21
Fig 5.4 DFD Level-1

The fig. 5.4 shows the data flow between user, track patient, infected patient diagnosis and identification of probable infected patient. When the data is loaded by the user, the details of the patient is taken and the location is tracked. The patient information is transferred to the infected patient diagnosis. Once the patients are recognized the details are sent to the user.

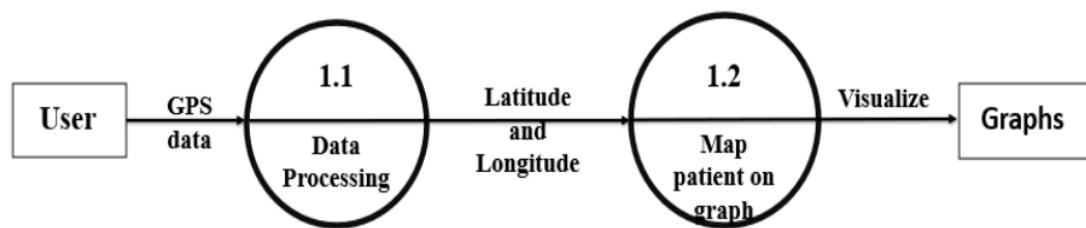


Fig 5.5 DFD Level-2 Patient Tracking

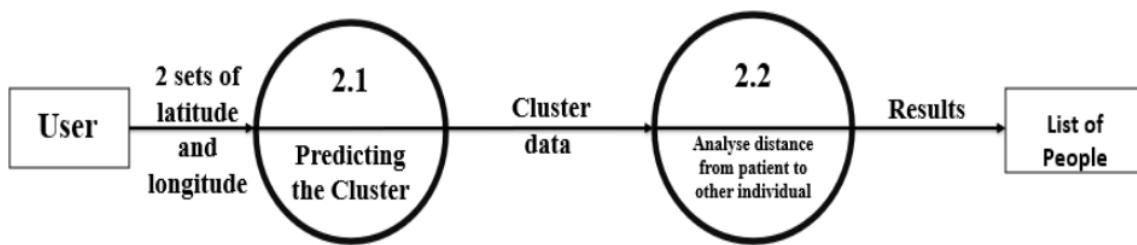


Fig 5.6 DFD Level-2 Infected Patient Diagnosis

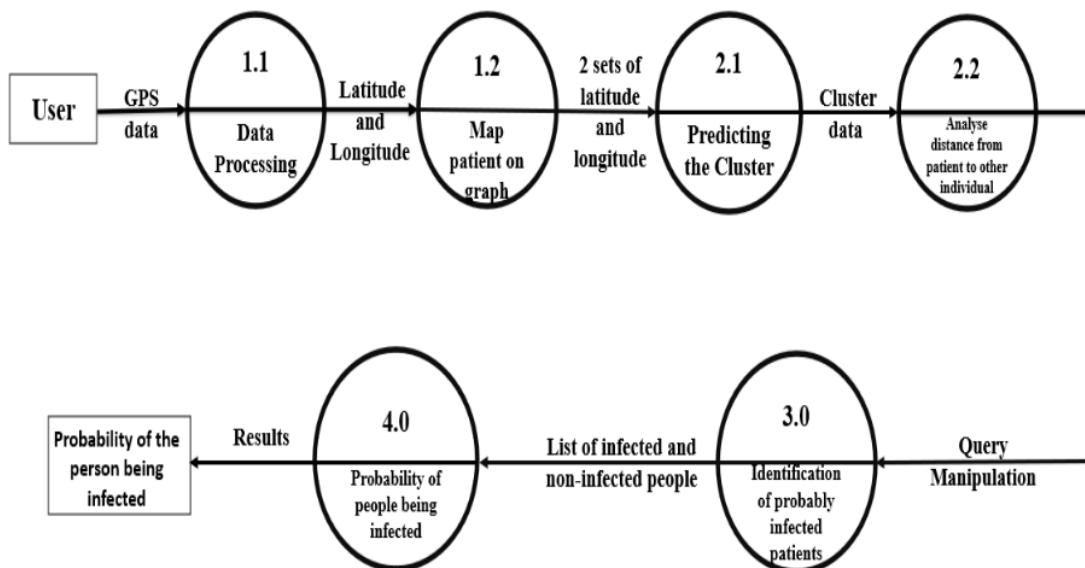
**Fig 5.5 Final DFD**

Fig 5.5 shows the next level of DFD that is the in-depth flow of data that takes place. The data that is flown from the user to the model, from the model to the patient diagnosis, from patient diagnosis to the identifying probable patients.

5.1.2.2 Sequence Diagram

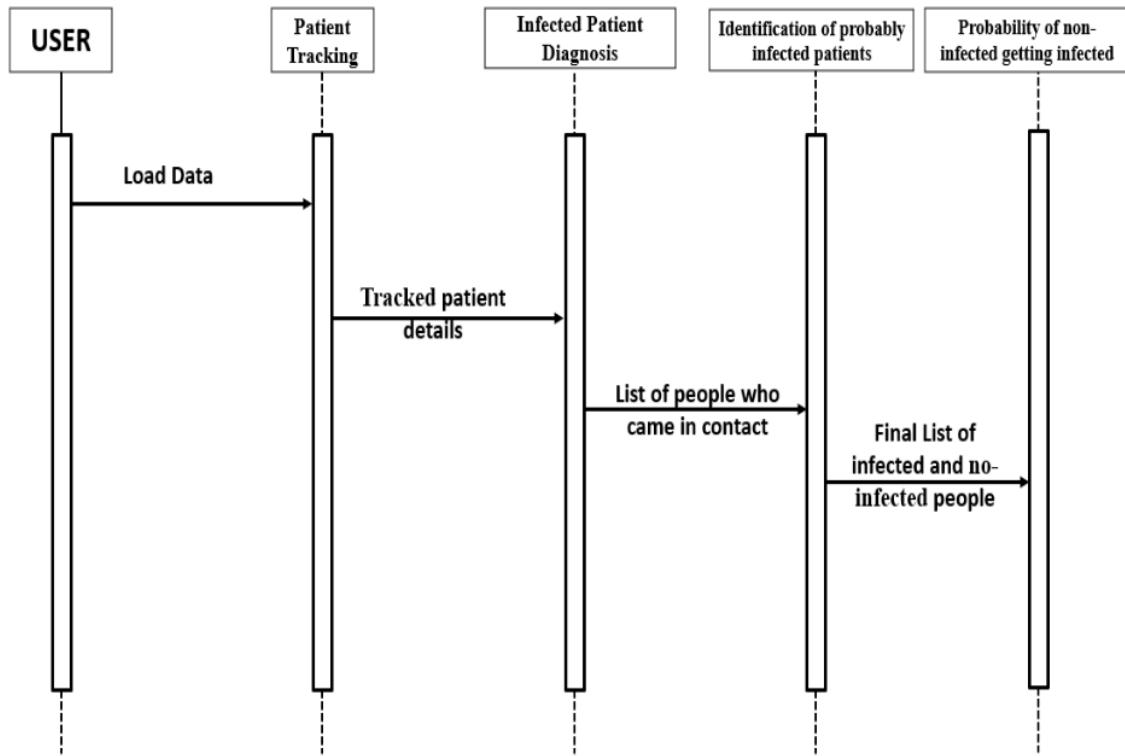


Fig 5.6 Sequence Diagram

Fig 5.6 gives a clear description of the data flow inside the Contact Tracing Mode

CHAPTER 6

IMPLEMENTATION

6.0 Code Snippets

DBSCAN

```
safedistance = (0.0018288/6371)
ls=[]
for date in data.Dates.unique():
    data1=data[data.Dates==date]
    print("Data shape : ",data1.shape[0])
    model = 17dbSCAN(eps=safedistance, min-samples=2, algorithm='ball-tree',
metric='haversine').fit(np.radians(data1[['lat', 'long']]))

    core-samples-mask = np.zeros_like(model.labels_, dtype=bool)
    core-samples-mask[model.core-sample-indices_] = True
    labels = model.labels_
    data1['Cluster'] = model.labels_.tolist()
    ls.append(data1)
data_final=pd.concat(ls, axis=0)
print("Final data shape : ", data_final.shape)
print (model.labels_)
```

Track Patient

```
# randomly find a number
import random

# random_id=random.randint(data_final1.id.min(),data_final1.id.max())
# random_id=random.choice(data_final1.id.tolist())
```

```
random_id=random.choice(data_final1.loc[data_final1.infected==True,'id'].tolist())

#find the cluster where he/she falls

random_id_cluster=data_final1[(data_final1.id==random_id)]
print(f"The random id is : {random_id} and belongs to cluster
{random_id_cluster['Cluster'].tolist()[0]}. Name :
{random_id_cluster['first_name'].tolist()[0]}. He was present on date
{random_id_cluster['Dates'].tolist()[0]}")
```

Infected Patient Diagnosis

```
cluster_req=data_final.loc[data_final.id==random_id,'Cluster'].tolist()[0]

infected_ids=data_final[(data_final.Dates==random_id_cluster['Dates'].tolist()[0]) &
(data_final.Cluster==cluster_req)]
infected_ids
```

```
22
disp_dict={}
lst=[]
for index, row in infected_ids.iterrows():
    if row['id'] !=random_id:
        lst.append((row['latitude'], row['longitude']))
```

```
22
xy_chart = pygal.XY(stroke=False)#,style=dark-lighten-style)

xy_chart.add(str(random_id_cluster['first_name'].tolist()[0]),
[(random_id_cluster['latitude'].tolist()[0], random_id_cluster['longitude'].tolist()[0])])
#other person
xy_chart.add("Infected",lst_infected)
```

```
xy_chart.add("Not Infected",lst_not_infected)
displays(HTML(base.html.format(rendered-chart=xy-chart.render(is-unicode=True)
)))

```

Identification of probably infected patients

```
infected_ids=data_final1[(data_final1.Dates==random_id_cluster['Dates'].tolist()[0])
& (data_final1.Cluster==random_id_cluster['Cluster'].tolist()[0])]
infected_ids
```

```
infected_ids=data_final1[(data_final1.Dates==random_id_cluster['Dates'].tolist()[0]) &
(data_final1.Cluster==random_id_cluster['Cluster'].tolist()[0]) &
(data_final1.infected==False) & (data_final1.id!=random_id)]
infected_ids
```

```
infected_ids=data_final1[(data_final1.Dates==random_id_cluster['Dates'].tolist()[0]) &
(data_final1.Cluster==random_id_cluster['Cluster'].tolist()[0]) & (data_final1.infected==True)
& (data_final1.id!=random_id)]
infected_ids
```

Probability of patient getting infected

```
cough=int(input("Cough Range : [Enter value between 1 and 5 , 1 being the
lowest and 5 being the highest] :"))
```

```
fever=int(input("Fever Range : [Enter value between 1 and 5 , 1 being the
lowest and 5 being the highest] :"))
```

```
breathe=int(input("Breathing Range : [Enter value between 1 and 5 , 1 being the
```

```
lowest and 5 being the highest] :"))
infected=input("In contact with infected? (y/n) :")
if infected.lower()=='y':
    infected=1
elif infected.lower()=='n':
    infected=0
else:
    print("Incorrect value entered. Please start again")

if infected==0 or infected==1:
    value=(cough+fever+breathe)/3*5
elif infected==1:
    value=((cough*2+fever2+breathe2)/(3*5)2)*0.5
print("The severity % is :",value)
[5:57 pm, 09/06/2022] Amar Chokher Bali: cough=int(input("Cough Range :
[Enter value between 1 and 5 , 1 being the lowest and 5 being the highest] :"))

fever=int(input("Fever Range : [Enter value between 1 and 5 , 1 being the lowest
and 5 being the highest] :"))

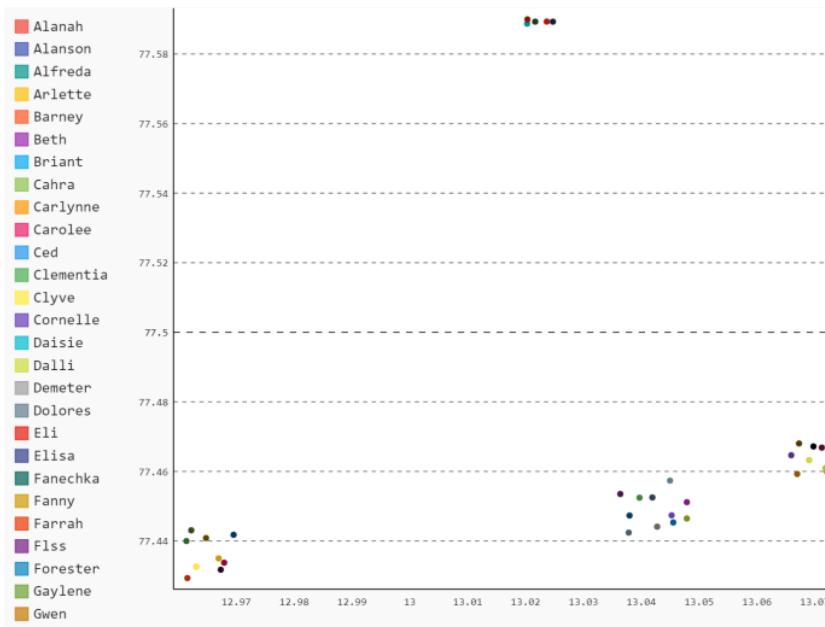
breathe=int(input("Breathing Range : [Enter value between 1 and 5 , 1 being the
lowest and 5 being the highest] :"))

infected=input("In contact with infected? (y/n) :")
if infected.lower()=='y':
    infected=1
elif infected.lower()=='n':
    infected=0
else:
```

```
print("Incorrect value entered. Please start again")
```

```
if infected==0 or infected==1:  
    value=(cough+fever+breathe)/3*5  
elif infected==1:  
    value=((cough * * 2+fever * * 2+breathe * * 2)/(3 * 5 ) * * 2) * * 0.5  
print("The severity % is :",value)
```

6.2 Implementation



49
Fig 6.1 Clusters

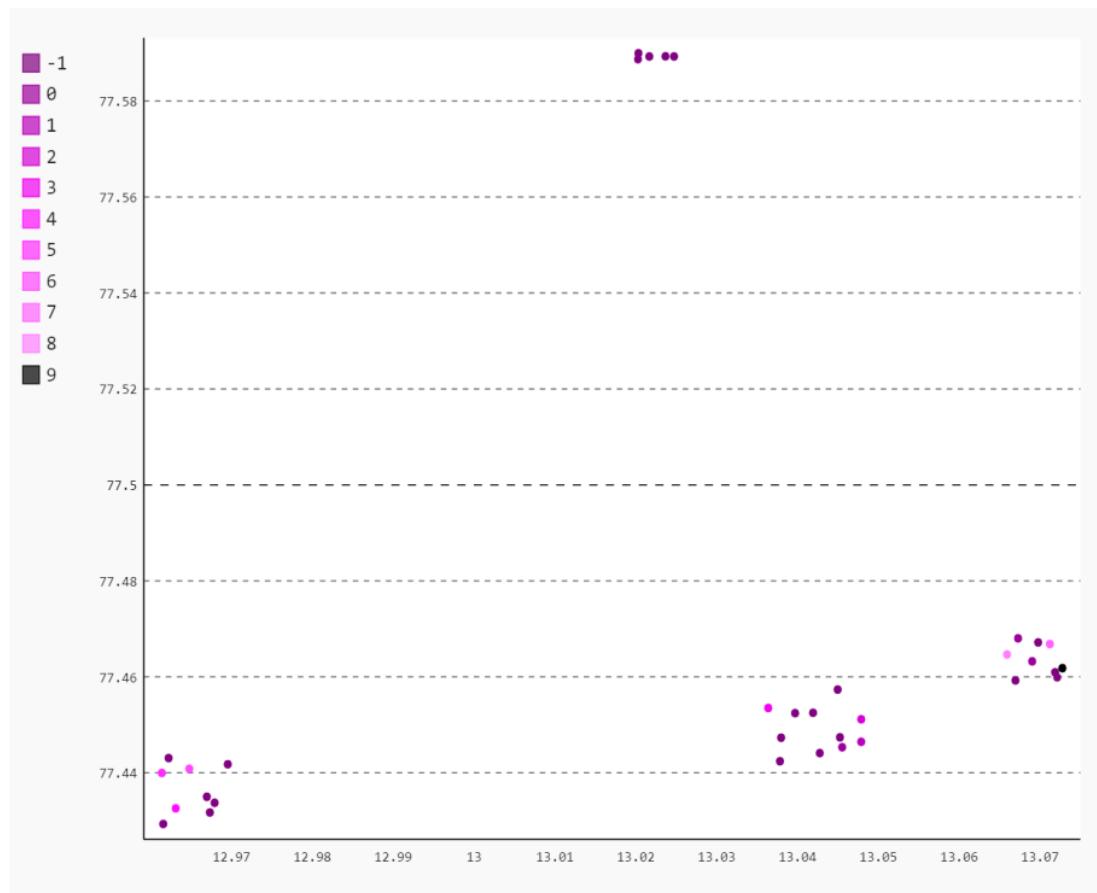
**Fig 6.2 Clusters with noise**

Fig 6.2 shows the clusters with a -1 cluster too which represents the outliers in the dataset and are needed to be removed, the numbers from 0 to 9 represents the clusters we actually need to work with for the contact tracing.

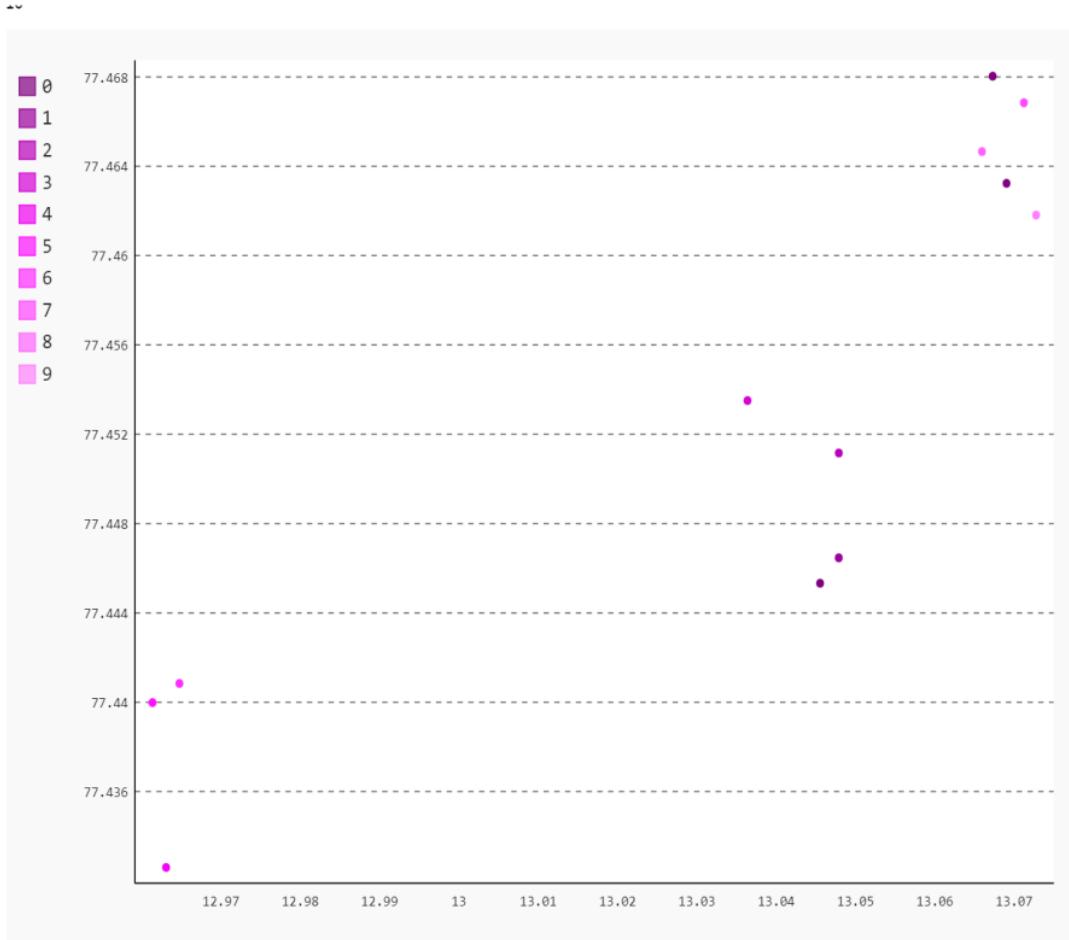
**Fig 6.3 Clusters without noise**

Fig 6.3. shows the final data and clusters without the outliers, thus after removing the -1 cluster the final cluster value is 10.

randomly find an id of the infected person

```
In [15]: # randomly find a number
import random
# random_id=random.randint(data_final1.id.min(),data_final1.id.max())
# random_id=random.choice(data_final1.id.tolist())
random_id=random.choice(data_final1.loc[data_final1.infected==True,'id'].tolist())
#find the cluster where he/she falls
random_id_cluster=data_final1[(data_final1.id==random_id)]
print(f"The random id is : {random_id} and belongs to cluster {random_id_cluster['Cluster'].tolist()[0]}. Name : {random_id_c
The random id is : 364 and belongs to cluster 7. Name : Lexie. He was present on date 25-04-2022
```

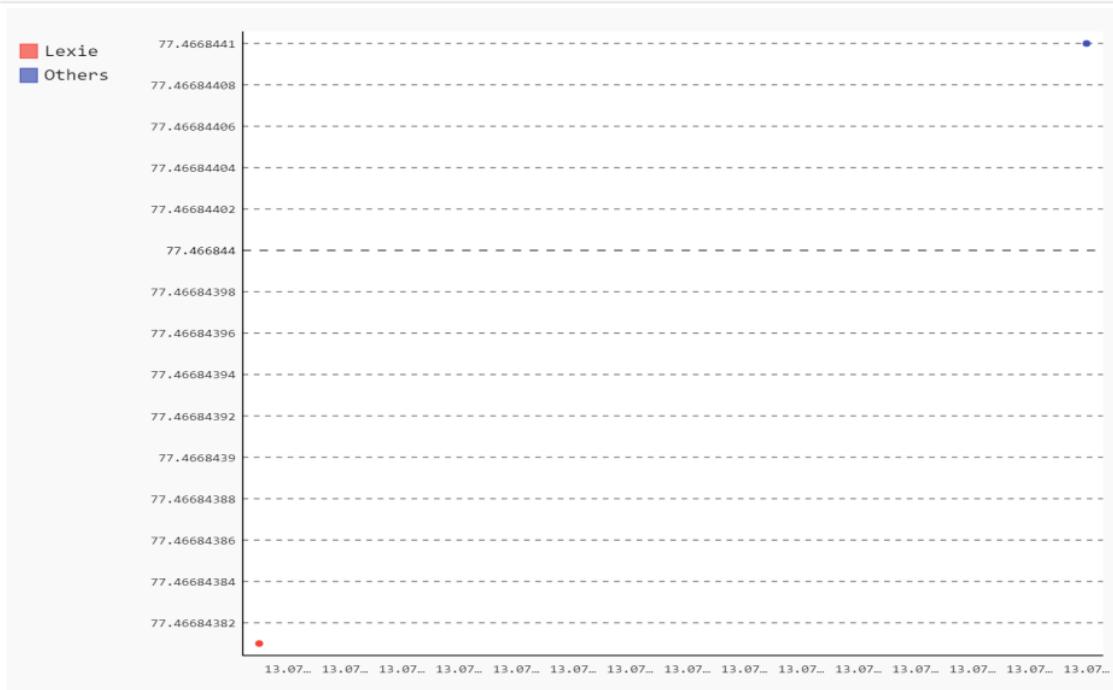
Fig. 6.4 Choose Random id**Fig. 6.5 Plotting single Patient**

Fig 6.5 and 6.4 shows the process of choosing a random patient and plotting that random patient on a graph using latitude and longitude.

```
In [111]: M infected_ids=data_final[(data_final.Dates==random_id_cluster['Dates'].tolist()[0]) & (data_final.Cluster==cluster_req)]  
infected_ids
```

Out[111]:

	id	first_name	last_name	gender	Area	latitude	longitude	Dates	infected	full name	Cluster
21	213	Mordy	Wherit	Male	Kadarannahalli	13.047887	77.446465	25-04-2022	True	Mordy Wherit	0
22	214	Randa	Kynsey	Female	Kadarannahalli	13.047887	77.446465	25-04-2022	True	Randa Kynsey	0

Fig. 6.6 List of people who came in contact**Fig. 6.7 Plotting newly identified individuals**

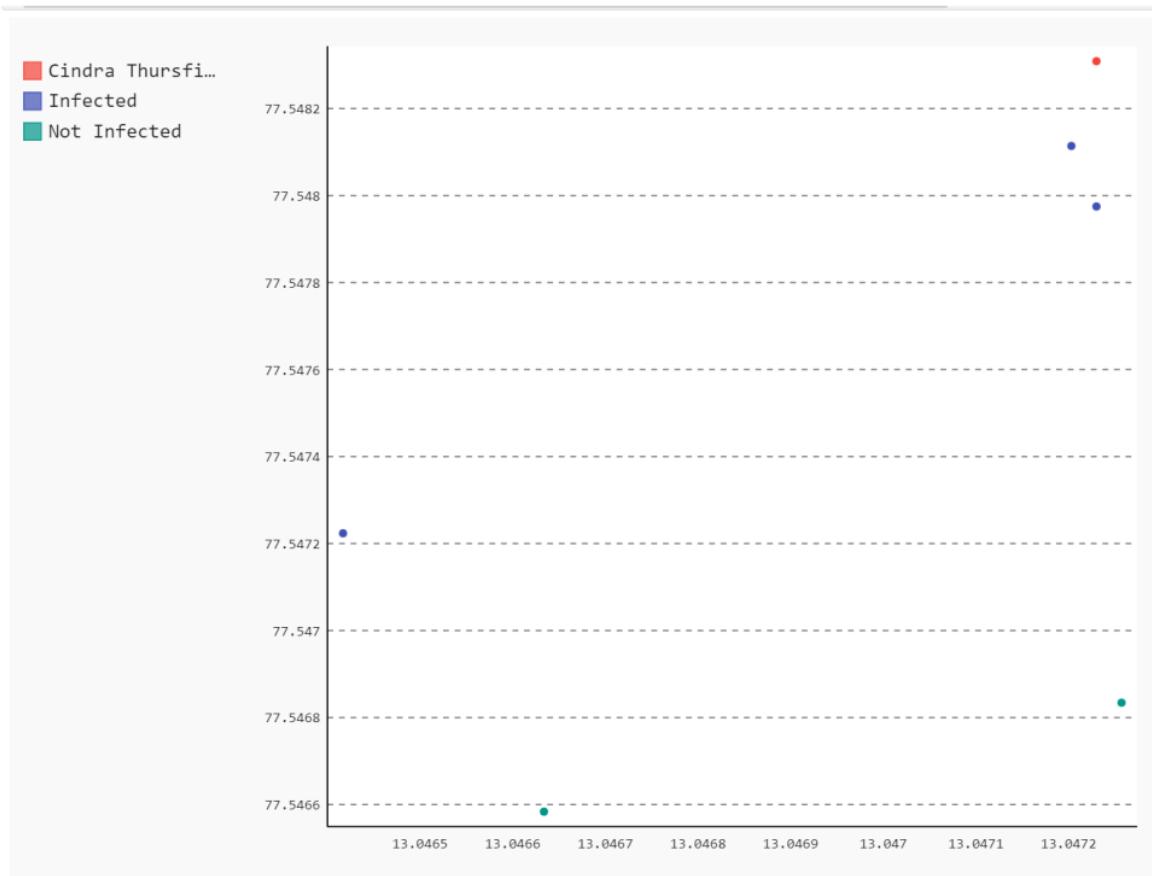


Fig. 6.8 Plotting infected and non-infected people

```
cough=int(input("Cough Range : [Enter value between 1 and 5 , 1 being the lowest and 5 being the highest] :"))

fever=int(input("Fever Range : [Enter value between 1 and 5 , 1 being the lowest and 5 being the highest] :"))

breathe=int(input("Breathing Range : [Enter value between 1 and 5 , 1 being the lowest and 5 being the highest] :"))

infected=input("In contact with infected? (y/n) :")

if infected.lower()=='y':
    infected=1
elif infected.lower()=='n':
    infected=0
else:
    print("Incorrect value entered. Please start again")

if infected==0 or infected==1:
    value=((cough+fever+breathe+infected)*100)/16
elif infected==1:
    value=((cough +fever +breathe+infected )*100 )/16
print("The severity % is :",value)
```

```
Cough Range : [Enter value between 1 and 5 , 1 being the lowest and 5 being the highest] :3
Fever Range : [Enter value between 1 and 5 , 1 being the lowest and 5 being the highest] :2
Breathing Range : [Enter value between 1 and 5 , 1 being the lowest and 5 being the highest] :2
In contact with infected? (y/n) :y
The severity % is : 50.0
```

Fig. 6.8 Probability of a person getting infected

CHAPTER 7

SOFTWARE TESTING

Every business enterprise need to provide number one cognizance to checking out whilst constructing any project. Testing is an crucial part of improvement , specially to make certain quality. Many smaller and mid- sized corporations don't provide checking out lots interest albeit it's important for handing over a strong product. Just like it's crucial to own checks present, it additionally topics how the company run those checks. Every business enterprise growing a product need to have checks present. Testing is an crucial part of improvement , specially to make certain quality. Many smaller and mid- sized corporations don't provide checking out lots interest albeit it's important for handing over a strong product. Just like it's crucial to own checks within the specific situation, it additionally topics how the company run those checks. Some corporations prefer to do guide checking out, even though that's now no longer the handiest approach. The subsequent logical step is automating your checking out technique the business enterprise have checks present. Manual checking out need to be minimized in any respect costs. And take a look at automation will increase average software program improvement performance and lets in for extra strong equipment to be built.The following are examples of the types of testing carried out as part of this project.

Unit Testing

Testcases are prepared for each and every feature and the entire flow is tested in such a manner that each and every testcase is passed in each of the cycle so that there is no bugs remaining due to code changes after each cycle,which enables to have a very bug free product after the end of the unit testing.

Load Testing

Load testing emphasizes testing of the behaviour of system during the course of load being put on it or computing device can handle high loads given a high demand of end-users.

7.1 Test Cases

The testing phase of the project consists of unit test cases. Since the modules are independent of each other completely as intermediary outputs are being fed for the next module as input, there is no connection or integration done at a later point of time

7.1: Unit test cases for Extraction of data

Test Case ID/No.	Test Case Description	Input/Inputs	Expected Outputs	Actual Outputs	Results
T01	Format of dataset dumped by client	Dataset at location	Appropriate format for import is available (.txt, .csv, .xlsx etc.,)	Appropriate format for import is available	Pass
T02	Format of dataset dumped by client	Dataset at location	Appropriate format for import is available (.txt, .csv, .xlsx etc.,)	Dataset in .flv format	Fail
T03	Validation of date format	Excel Dataset	Date format in “ddmmmyyy”	Date format in “ddmmmyy”	Pass

The test cases talk about validating the format of the data sets and location of

file which is fetched in notebook. The average number of patient details is generated using mackroo website where user has to enter the number of required columns and add the parameters as needed then it generates the number of rows with particular dummy data which can be used. This results in generating huge data in hundreds of files based on the time period for which the model was getting tested.

Validation of dates is a common scenario where the model can find out the number of individuals who came in proximity with positive patient on the same day. Normalizing them would help the model recognize the data from all the time periods simultaneously that reduces time of execution.

7.2: Unit test cases for Locating covid patients

Test Case ID/No.	Test Case Description	Input/Inputs	Expected Outputs	Actual Outputs	Results
T01	Find Infected People	Person ID (56)	Details of Positive or Negative person	Details of Positive or Negative person	Pass
				Any other output	Fail
T02	Find Infected People	Person Name (“Ishita”)	Details of Positive or Negative person	Details of Positive or Negative person	Pass
				Any other output	Fail
T03	Find Infected People	Dates In format (“ddmmmyy”) (“4-9-2021”)	Detail's of infected person on mentioned date.	Detail's of infected person on mentioned date.	Fail

				Any other output	Fail
--	--	--	--	------------------	------

The Unit test cases for Locating covid patients talk about tracing the particular individual positive patient with inserting few unique values as patient id or name as a input. The details of covid positive patient will be displayed with location as in dataset it has mentioned about the coordinates

with place name also dataset contains the dates of the particular patient on which date he was present on which location, which gives us clear picture to figure out the covid infected and non-infected people.

7.3: Unit test cases for Locate people who came in contact with the infected patient

Test Case ID/No.	Test Case Description	Input/Inputs	Expected Outputs	Actual Outputs	Results
T01	checking People in radius with infected person	Dates (4-6-2022)	Details of people who is in mentioned radius	Details of people who is in mentioned radius	Pass
				Any other output	Fail
T02	Checking People in same clusters with infected person	Clusters Number (2)	Details of people who form in same clusters	Details of people who form in same clusters	Pass
				Any other output	Fail
T03	Finding People in same city	City Name (Bengaluru)	Display all the people within same city	Display all the people within same city	Fail

				Any other output	Fail
--	--	--	--	------------------	------

The Unit test cases for Locate people who came in contact with the infected patient talk about tracing the particular individual people who has been in contact

with positive patient within the given radius mentioned (**i.e., 0.0001876**) in eps value in algorithm where it defines the neighborhood around a data point i.e. if the distance between two points is lower or equal to 'eps' then they are considered as neighbors and **MinPts** which counts **Minimum number of neighbors (data points) within eps radius.** Larger the dataset, the larger value of **MinPts must be chosen.** The cluster is formed based on the DBSCAN algorithm where it basically focus on Density-based spatial clustering of applications with noise clustering method, also to find number of infected through clusters are one of the technique. Dataset contains the name of the cities through which we can detect the number of infected or non-infected people around the city.

Table 7.4: Unit test cases for check if other people located were covid infected or not

Test Case ID/No.	Test Case Description	Input/Inputs	Expected Outputs	Actual Outputs	Results
T01	Find infected people in contact	Clusters, Dates (3,2-42022)	Details of Infected people	Details of Infected people	Pass
				Any other output	Fail

Covid Patient Tracing and Diagnosis System

JUNE 2022

T02	Find non-infected people in contact	Clusters, City (4,Bengaluru)	Details of Non-infected People	Details of Non-infected People	Pass
				Any other output	Fail

The Unit test cases for check if other people located were covid infected or not and its results, tracing the particular individual people who has infected or not

followed by count of infected people and count of non-infected people based on the certain parameters. It can be based on city where the input will be city name and the list of people infected or non-infected will be displayed and second method is based on dates which can find out the infected or non-infected person with respective to city.

The other method followed will be through clusters where the count of infected and non-infected in same clusters can be displayed or respective other clusters based on safe distance mentioned, if the person is in under mentioned distance will be counted as infected vise-versa.

Table 7.5: Unit test cases to check if the probability value predicted is correct

Test Case ID/No.	Test Case Description	Input/Inputs	Expected Outputs	Actual Outputs	Results
T01	Find probability Of person being infected	Patient id (35)	Percentage value	Percentage value	Pass

				Any other output	Fail
T02	Check if infected value taken is correct or not	Came in contact with infected person ('n' or 'y')	Proceed further	Proceed further	Pass
				Any other output	Fail

The Unit test cases for check if the probability value predicted is correct and it's results. One can observe that if the values of person came in contact is anything else than 'n' or 'y' the model won't be able to predict the values correctly and will predict some incorrect value. Not only that the patient id should also be entered correctly otherwise the percentage value will not be correct.

CHAPTER 8

CONCLUSION

The COVID-19 epidemic continues to have an impact on everyone's way of life. Contact tracing is anticipated to be crucial in aiding health officials in promptly spotting those who may have been exposed to the virus. COVID had a terrible impact on human existence around the world, and as of this writing, the management of educational institutions and workplaces, as well as many other areas of daily life as we know it, remains uncertain. We can conclude from the preceding sections that a clustering technique like DBSCAN can cluster data points from the dataset without having previous knowledge of datasets. DBSCAN is effective in finding non-linearly separable clusters, on the other hand clustering algorithms like k-means are insufficient for this dataset. The DBSCAN function implemented in this project has three objective where it finds the infected person, the people who came in contact with infected patient followed with the list of infected people and non-infected people based on the same dates and clusters they fall into. Thus, giving us an effective way of tracking down potentially infected individuals. In the end, both digital contact tracing and manual approaches have their own limitations. However, employing machine learning to perform effective contact tracing can help us save on an array of human and financial resources.

CHAPTER 9

FUTURE ENHANCEMENTS

The current COVID-19 outbreak has sparked a widespread adoption of a virus tracking app. If these models and technologies had been in existence and extensively tested before to the commencement of the Covid-19 pandemic, we could have been better prepared for a pandemic of this magnitude.

1) Improvement in proximity accuracy:

In previous exchanges, we've discussed the current issues with proximity accuracy.

The Bluetooth technology was not created with location or proximity determination as a primary consideration. During the development phase of new "Bluetooth-like" protocols, a focus on location/proximity services should be prioritised. This approach should include antenna design, which allows for not only distance precision but also direction-finding. The Bluetooth 5.1 protocol does include direction-finding, albeit with limited precision. At the time of writing, no Covid19 tracing programme has such direction capabilities built into their proximity analysis.

2) A fully decentralised architecture for infection tracing:

One of the "takeaways" from the COVID19 disaster is that privacy concerns must be addressed in order for the public to use the technology more widely. None of the apps we've discussed so far have a fully decentralised design; they all rely on a central server to some extent, usually under the direction of a governing authority. It is necessary to conduct research on a fully decentralised system that uses some type of peer-to-peer network to permit privacy-preserving information sharing among user-devices.

3) Quantum communications:

At least in terms of deployment, is the most advanced of all the new quantum applications. Commercial quantum communications services are currently available, and proof-of-principle deployment in space has already taken place. The advanced communications security and enhanced privacy protections offered by this technology will most likely have the greatest impact on tracing applications. In theory, if properly implemented, future virus-tracking programmes will have unrestricted security and privacy, making hacking and unauthorised access to virus-tracking data obsolete.

BIBLIOGRAPHY

- [1] Rustam, Furqan, Aijaz Ahmad Reshi, Arif Mehmood, Saleem Ullah, Byung-Won On, Waqar Aslam, and Gyu Sang Choi. "COVID-19 future forecasting using supervised machine learning models." *IEEE access* 8 (2020): 101489-101499.
- [2] Gambhir, E., Jain, R., Gupta, A. and Tomer, U., 2020, September. Regression analysis of COVID-19 using machine learning algorithms. In 2020 International conference on smart electronics and communication (ICOSEC) (pp. 65-71). IEEE.
- Wala, T. & Sharma, A. K. (2014). *Improvised Software Testing Tool*, 3(9), 573–581.
- [3] Sarwar, M.A., Kamal, N., Hamid, W. and Shah, M.A., 2018, September. Prediction of diabetes using machine learning algorithms in healthcare. In 2018 24th international conference on automation and computing (ICAC) (pp. 1-6). IEEE.
- [4] Dahiwade, D., Patle, G., & Meshram, E. (2019, March). Designing disease prediction model using machine learning approach. In 2019 3rd International Conference on Computing Methodologies and Communication (ICCMC) (pp. 1211-1215). IEEE.
- [5] Subudhi, S., Verma, A., Patel, A. B., Hardin, C. C., Khandekar, M. J., Lee, H., ... & Jain, R. K. (2021). Comparing machine learning algorithms for predicting ICU admission and mortality in COVID-19. *NPJ digital medicine*, 4(1), 1-7.
- [6] John, Christopher Clement, Vijayakumar Ponnusamy, Sriharipriya Krishnan Chandrasekaran, and R. Nandakumar. "A survey on mathematical, machine learning and deep learning models for COVID-19 transmission and diagnosis." *IEEE reviews in biomedical engineering* (2021).
- [7] Sujath, R., Jyotir Moy Chatterjee, and Aboul Ella Hassanien. "A machine learning forecasting model for COVID-19 pandemic in India." *Stochastic Environmental Research and Risk Assessment* 34, no. 7 (2020): 959-972.
- [8] Gupta, Rajan, Manan Bedi, Prashi Goyal, Srishti Wadhera, and Vaishnavi Verma. "Analysis of COVID-19 tracking tool in India: case study of Aarogya Setu

- mobile application." Digital Government: Research and Practice 1, no. 4 (2020): 1-8.
- [9] Berglund, Jennifer. "Tracking COVID-19: there's an App for that." IEEE pulse 11, no. 4 (2020): 14-17.
- [10] Sandeepa, Chamara, Charuka Moremada, Nadeeka Dissanayaka, Tharindu Gamage, and Madusanka Liyanage. "Social interaction tracking and patient prediction system for potential COVID-19 patients." In 2020 IEEE 3rd 5G World Forum (5GWF), pp. 13-18. IEEE, 2020.
- [11] Mallik, Ranajoy, Dilip Sing, and Rajib Bandyopadhyay. "GPS tracking app for police to track ambulances carrying COVID-19 patients for ensuring safe distancing." Transactions of the Indian National Academy of Engineering 5, no. 2 (2020): 181-185.
- [12] Biswas, Suparna, et al. "GPS tracking, QR code scanner based application to help the health workers and common citizens during Covid-19." Journal of Physics: Conference Series. Vol. 1797. No. 1. IOP Publishing, 2021
- [13] "The Difference between Robotic Process Automation and Artificial Intelligence Available", pp. 293-296, June, 2018
- [14] J.G.Enríquez,A.Jiménez-Ramírez, F. J. Domínguez-Mayo, J. A. García-García, "Robotic Process Automation: A Scientific and Industrial Systematic Mapping Study", Access IEEE, vol. 8, pp. 39113-39129, 2020
- [15] S. Aguirre and A. Rodriguez, "Automation of a business process using robotic process automation (rpa): A case study", Workshop on Engineering Applications, pp.65-71, 2017.
- [16] C.Lamberton, D.Brigo and D.Hoy, "Impact of robotics rpa and ai on the insurance industry: challenges and opportunities", Journal of Financial Perspectives: Insurance edition, 2016.
- [17] Kaur, H. & Gupta, G. (2013). Comparative Study of Automated Testing Tools: Selenium, Quick Test Professional and Testcomplete, Int. Journal of Engineering Research and Applications, 3(5).
- [18] H. Ed-douibi, J.L.C. Izquierdo and J. Cabot, "Automatic Generation of Test Cases for REST APIs: A Specification-Based Approach", IEEE 22nd Intern. Enterprise Distributed Object Computing Conference, pp. 181-190, 2018.

- [19] T. Fertig and P. Braun, "Model-Driven Testing of RESTful APIs", Intern. Conferenceon World Wide Web, pp. 1497-1502, 2015.
- [20] S. Yatskiv, I. Voytyuk, N. Yatskiv, O. Kushnir, Y. Trufanova and V. Panasyuk, "Improved Method of Software Automation Testing Based on the Robotic Process Automation Technology", 2019 9th International Conference on Advanced Computer Information Technologies (ACIT), pp. 293-296, June, 2019
- [21] "The Difference between Robotic Process Automation and ArtificialIntelligence Available", pp. 293-296, June, 2018
- [22] J.G.Enríquez,A Jiménez-Ramírez, F. J. Domínguez-Mayo, J. A. García-García, "Robotic Process Automation: A Scientific and Industrial Systematic Mapping Study",Access IEEE, vol. 8, pp. 39113-39129, 2020
- [21] S. Aguirre and A. Rodriguez, "Automation of a business process using robotic process automation (rpa): A case study", Workshop on Engineering Applications, pp.65-71, 2017.
- [22] C.Lamberton, D.Brigo and D.Hoy, "Impact of robotics rpa and ai on the insurance industry: challenges and opportunities", Journal of Financial Perspectives: Insurance edition, 2016.
- [23] Kaur, H. & Gupta, G. (2013). Comparative Study of Automated Testing Tools: Selenium, Quick Test Professional and Testcomplete, Int. Journal of Engineering Research and Applications, 3(5).
- [24] Mustafa, K. M., Al-Qutaish, R. E., & Muhairat, M. I. (2009). Classification of software testing tools based on the software testing methods, in 2009 International Conference on Computer and Electrical Engineering, ICCEE 2009, 1, 229–233.
- [25] Kolli, R. (2016). An Empirical Study on Software Test Estimation.
- [26] Raulamo, P., Hosio, S. & Mäntylä, M. V (2019). Practitioner Evaluations on Software Testing Tools. In Proceedings of the Evaluation and Assessment on Software Engineering, 57–66.
- [27] Vos, T., Marín, B., Panach, I., Baars, A., Ayala, C., & Franch, X. (2011). Evaluating software testing techniques and tools. Actas de Las 16th Jornadas de Ingenieria Del Software y Bases de Datos, JISBD 2011, pp. 531–536.

- [28] Kannan, S. & Pushparaj, T. (2014). A Study on Variations of Bottlenecks in Software Testing, International Journal of Computer Sciences and Engineering, 2(5), 8 - 14.
- [29] Malik, V. & Gahlan, M. (2016). Comparative Study of Automated Web Testing Tools, Int. J. Latest Trends Eng. Technol., 6(3), 1- 8.
Top 10 Testing Automation Tools for Software Testing, pp.
- [30]. Giuseppina Cretella and Beniamino Di Martino, "An Overview of Approaches for the Migration of Applications to the Cloud," in Lecture Notes in Information Systems and Organisation, Leonardo Caporarello,Beniamino Di Martino, and Marcello Martinez, Eds.: Springer, 2014, pp.
- [31]. Gos, Konrad & Zabierowski, Wojciech. (2020). The Comparison of Microservice and Monolithic Architecture 150-153.
10.1109/MEMSTECH49584.2020.9109514.
- [32]. C. M. Aderaldo, N. C. Mendonça, C. Pahl, and P. Jamshidi, "Benchmark requirements for microservices architecture research," in Proceedings of the 2017 IEEE/ACM 1st International Workshop on Establishing the Community-Wide Infrastructure for Architecture-Based Software Engineering, pp. 8–13, Buenos Aires, Argentina, May 2017
- [33]. C. Esposito, A. Castiglione, and K.-K. R. Choo, "Challenges in delivering software in the cloud as microservices," IEEE Cloud Computing, vol. 3, no. 5, pp. 10– 14, 2016.
- [34]. G. Mazlami, J. Cito, and P. Leitner, "Extraction of microservices from monolithic software architecture," IEEE International Conference on Web Services, vol. 40, no. 11, pp. 524–531, 2017
- [35]. Sachchidanand Sing, Nirmala Singh, "Containers & Docker: Emerging Roles
- [36]. N. Alshuqayran, N. Ali and R. Evans, "A Systematic Mapping Study in Microservice Architecture," 2016 IEEE 9th International Conference on Service-Oriented Computing and Applications (SOCA), 2016, pp. 44-51, doi:
10.1109/SOCA.2016.15.
- [37]. Chen Yang, "Checkpoint and Restoration of Micro-service in Docker Containers," School of Information Security Engineering, Shanghai Jiao Tong University, China
-

200240

[38]. D. N. Jha, S. Garg, P. P. Jayaraman, R. Buyya, Z. Li and R. Ranjan, "A Holistic Evaluation of Docker Containers for Interfering Microservices," 2018 IEEE International Conference on Services Computing (SCC), 2018, pp. 33-40, doi:10.1109/SCC.2018.00012

[39]. Fawaz Paraiso, Stéphanie Challita, Yahya Al-Dhuraibi, Philippe Merle, "Model-Driven Management of Docker Containers", University of Lille & Inria Lille - Nord Europe 2016

[40]. Alin Calinciu, Cristian Constantin Spoiala, Corneliu Octavian Turcu, Constantin Filote, "OpenStack and Docker: building a high-performance IaaS platform for interactive social media applications", May 19-21, 2016

[41]. R. R. Yadav, E. T. G. Sousa, and G. R. A. Callou, Docker Containers Versus Virtual Machine-Based Virtualization: Proceedings of IEMIS 2018

[42]. D. Liu, H. Zhu, C. Xu, I. Bayley, D. Lightfoot, M. Green, and P. Marshall. Cide: An integrated development environment for microservices. In 2016 IEEE International Conference on Services Computing (SCC), pages 808–812, June 2016

[43]. Bhawiyuga A, Data M and Warda A 2017 Architectural Design of Token based Authentication of MQTT Protocol in Constrained IoT Device, 2017 11th

International Conference on Telecommunication Systems Services and Applications (TSSA), 1–4 [44]. Zheng K and Jiang W 2014 A Token Authentication Solution for

Hadoop Based on Kerberos PreAuthentication, 2014 International Conference on Data Science and Advanced Analytics (DSAA) 354–60

[45]. Ethelbert O, Moghaddam F F , Wieder P and Yahyapour R A 2017 JSON Token-Based Authentication and Access ManagementSchema for Cloud SaaS Applications, in 2017 IEEE 5th International Conference on Future Internet of Things and Cloud (Cornel University Library) p 6

[46]. Mestre P, Madureira R, Melo-Pinto P and Serodio C 2017 Securing RESTful Web Services using Multiple JSON Web Tokens, in Proc. World Congress on Engineering 2017 418–23

[47]. Memeti A, Imeri F and Cico B 2017 REST vs. SOAP: Choosing the best web service while developing in-house web applications, Journal of Natural Sciences and Mathematics of UT 2(3) 63–8

[48]. Shofa R N, Aradea and Kurnia B B 2013 Penerapan Service Oriented Architecture (SOA)

Dalam Pembangunan Web Based Learning, Jurnal Penelitian SITROTIKA 9(2) 221–19

[49]. Pavlovski C, Warwar C, Paskin B and Chan G 2015 Unified Framework for Multifactor Authentication, 22nd International Conference on Telecommunications (ICT 2015) 209–13

[50]. Balaj Y 2017 Token-Based vs Session-Based Authentication: A survey 1–5

Covid analysis

ORIGINALITY REPORT

18%

SIMILARITY INDEX

11%

INTERNET SOURCES

10%

PUBLICATIONS

10%

STUDENT PAPERS

PRIMARY SOURCES

- | | | |
|---|---|-----|
| 1 | Dr. Jasmine K S*, Mr. Thejas T S. "IoT- A Technology Transfer in Manufacturing Future Products", International Journal of Recent Technology and Engineering (IJRTE), 2020 Publication | 3% |
| 2 | en.wikipedia.org Internet Source | 3% |
| 3 | Nadeem Ahmed, Regio A. Michelin, Wanli Xue, Sushmita Ruj et al. "A Survey of COVID-19 Contact Tracing Apps", IEEE Access, 2020 Publication | 1 % |
| 4 | Submitted to Gitam University Student Paper | 1 % |
| 5 | Submitted to International Management Institute Student Paper | 1 % |
| 6 | doctorpenguin.com Internet Source | 1 % |
| 7 | ijircce.com Internet Source | 1 % |

8	Submitted to Universitat Politècnica de València Student Paper	<1 %
9	jultika.oulu.fi Internet Source	<1 %
10	doaj.org Internet Source	<1 %
11	www.ncbi.nlm.nih.gov Internet Source	<1 %
12	Submitted to Jacobs University, Bremen Student Paper	<1 %
13	Submitted to Liverpool John Moores University Student Paper	<1 %
14	Submitted to SASTRA University Student Paper	<1 %
15	Submitted to National Institute of Technology, Sri Nagar Jammu & Kashmir Student Paper	<1 %
16	ieeexplore.ieee.org Internet Source	<1 %
17	Submitted to University College London Student Paper	<1 %
18	Sarra Samet, Mohamed Ridda Laouar, Issam Bendib. "Use of Machine Learning Techniques	<1 %

"How to Predict Diabetes at an Early Stage", 2021 International Conference on Networking and Advanced Systems (ICNAS), 2021

Publication

- | | | |
|----|--|------|
| 19 | Submitted to De Montfort University
Student Paper | <1 % |
| 20 | link.springer.com
Internet Source | <1 % |
| 21 | Submitted to Ramrao Adik Institute of Technology
Student Paper | <1 % |
| 22 | towardsdatascience.com
Internet Source | <1 % |
| 23 | www.irjmets.com
Internet Source | <1 % |
| 24 | www.jungo.com
Internet Source | <1 % |
| 25 | a4academics.com
Internet Source | <1 % |
| 26 | "Understanding COVID-19: The Role of Computational Intelligence", Springer Science and Business Media LLC, 2022
Publication | <1 % |
| 27 | Hemant Ghayvat, Muhammad Awais, Prosanta Gope, Sharnil Pandya, Subhankar Majumdar. "ReCognizing SUspect and | <1 % |

PredictiNg ThE SpRead of Contagion Based on Mobile Phone LoCation DaTa (COUNTERACT): A System of identifying COVID-19 infectious and hazardous sites, detecting disease outbreaks based on the internet of things, edge computing, and artificial intelligence", Sustainable Cities and Society, 2021

Publication

28

Submitted to Higher Education Commission Pakistan

<1 %

Student Paper

29

stackify.com

<1 %

Internet Source

30

Christopher Clement John, Vijayakumar Ponnusamy, Sriharipriya Krishnan Chandrasekaran, Nandakumar R. "A Survey on Mathematical, Machine Learning and Deep Learning Models for COVID-19 Transmission and Diagnosis", IEEE Reviews in Biomedical Engineering, 2021

<1 %

Publication

31

Gaurav Tripathi, Rakesh Kumar. "Early Prediction of Diabetes Mellitus Using Machine Learning", 2020 8th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO), 2020

<1 %

Publication

- 32 www.6gchannel.com Internet Source <1 %
- 33 Md. Riyazudin, Mousmi Ajay Chaurasia, Syed Ibrahim Ibaad, Alisha Lalani, Salva Fathima. "Chapter 11 Estimation of Quantity and Nutritional Information of Food Using Image Processing", Springer Science and Business Media LLC, 2022 <1 %
Publication
- 34 Raju Singh. "Cloud Computing and Covid-19", 2021 3rd International Conference on Signal Processing and Communication (ICPSC), 2021 <1 %
Publication
- 35 Saumya Gupta, Chander Prabha. "Voting Regression Model for Covid-19 Time Series Data Analysis", 2021 3rd International Conference on Advances in Computing, Communication Control and Networking (ICAC3N), 2021 <1 %
Publication
- 36 sersc.org Internet Source <1 %
- 37 www.geeksforgeeks.org Internet Source <1 %
- 38 Submitted to University of Hertfordshire Student Paper <1 %

39	dspace.library.uvic.ca:8080	<1 %
Internet Source		
40	pure.hw.ac.uk	<1 %
Internet Source		
41	Richvichanak Chrin, Sujing Wang. "Analysis and Prediction of COVID-19 Data using Machine Learning Models", 2021 10th International Conference on Computing and Pattern Recognition, 2021	<1 %
Publication		
42	Submitted to Southern University And A & M College	<1 %
Student Paper		
43	Submitted to University of Bedfordshire	<1 %
Student Paper		
44	assets.researchsquare.com	<1 %
Internet Source		
45	oaji.net	<1 %
Internet Source		
46	prasannanatarajan.github.io	<1 %
Internet Source		
47	studentsrepo.um.edu.my	<1 %
Internet Source		
48	uhra.herts.ac.uk	<1 %
Internet Source		

49	C.-H. Lin, L.-C. Fu. "Robot navigation through obstacles of general shapes using a center-line oriented algorithm", Conference Proceedings., IEEE International Conference on Systems, Man and Cybernetics, 1989 Publication	<1 %
50	medium.com Internet Source	<1 %
51	pubcovid19.pt Internet Source	<1 %
52	www.coursehero.com Internet Source	<1 %

Exclude quotes On
Exclude bibliography On

Exclude matches < 3 words

Covid analysis

PAGE 1

PAGE 2

PAGE 3

PAGE 4

PAGE 5

PAGE 6

PAGE 7

PAGE 8

PAGE 9

PAGE 10

PAGE 11

PAGE 12

PAGE 13

PAGE 14

PAGE 15

PAGE 16

PAGE 17

PAGE 18

PAGE 19

PAGE 20

PAGE 21

PAGE 22

PAGE 23

PAGE 24

PAGE 25

PAGE 26

PAGE 27

PAGE 28

PAGE 29

PAGE 30

PAGE 31

PAGE 32

PAGE 33

PAGE 34

PAGE 35

PAGE 36

PAGE 37

PAGE 38

PAGE 39

PAGE 40

PAGE 41

PAGE 42

PAGE 43

PAGE 44

PAGE 45

PAGE 46

PAGE 47

PAGE 48

PAGE 49

PAGE 50

PAGE 51

PAGE 52

PAGE 53

PAGE 54

PAGE 55

PAGE 56

PAGE 57

PAGE 58

PAGE 59
