MAJOR PROJECT REPORT

On

IDENTIFICATION OF COVID-19 SPREADERS USING MULTIPLEX NETWORKS APPROACH

Submitted in partial fulfillment of the requirements for the award of the degree of

BACHELOR OF TECHNOLOGY

In

INFORMATION TECHNOLOGY

By

Sabrina Shaik 17261A1251

Under the guidance of

Mrs. J.Aruna Shanthi

Assistant Professor, IT.

DEPARTMENT OF INFORMATION TECHNOLOGY

MAHATMA GANDHI INSTITUTE OF TECHNOLOGY

(Affiliated to JNTUH, Hyderabad; Six UG Programs Accredited by NBA 3times; Accredited by NAAC with 'A' Grade)

Kokapet (vill), Gandipet (Mandal), Ranga Reddy (Dist.), Chaitanya Bharathi P.O., Hyderabad-500 075.

2020-2021

CERTIFICATE

This is to certify that the major project work entitled **IDENTIFICATION OF COVID-19 SPREADERS USING MULTIPLEX NETWORKS APPROACH s**ubmitted by **Sabrina Shaik (17261A1251)** and in partial fulfilment of requirements for the award of degree of Bachelor of Technology in Information Technology as specialization is a record of the bonafide work carried out under the supervision of Mrs. **J.Aruna Shanthi**, and this has not been submitted to any other university or institute for award of degree or diploma.

Project Guide	Project Coordinators
Mrs. J. Aruna Shanthi	**Mrs. J. Aruna Shanthi**
Assistant Professor	Assistant Professor
Dept. of IT	Dept. of IT

EXTERNAL EXAMINER	**DR. D. VIJAYA LAKSHMI**
	Professor and HOD, Dept. of IT

DECLARATION

We hereby declare that the research work entitled **IDENTIFICATION OF COVID-19 SPREADERS USING MULTIPLEX NETWORKS APPROACH** is original and bonafide work carried out by us as a part of fulfilment for Bachelor of Technology in Information Technology, Mahatma Gandhi Institute of Technology, Gandipet, Hyderabad, under the guidance of **Mrs. J. Aruna Shanthi**, Asst.Prof., Department of IT, MGIT.

No part of the project work is copied from books/journals/internet and wherever the partition is taken, the same has been duly referred in the text. The reported is based on the project work done entirely by us and not copied from any other source.

Sabrina Shaik **17261A1251**

ACKNOWLEDGEMENT

The satisfaction that accompanies the successful completion of any task would be incomplete without introducing the people who made it possible and whose constant guidance and encouragement crowns all efforts with success. They have been a guiding light and source of inspiration towards the completion of the project.

We would like to express our sincere gratitude and indebtedness to our project guide **Mrs. J. Aruna Shanthi,** Assistant Professor, Dept. of IT**,** who has supported us throughout our project with immense patience and expertise.

We are also thankful to the honorable Principal of MGIT **Prof. K. JAYA SANKAR** and **DR. D. VIJAYA LAKSHMI,** HOD and professor, Department of IT, for providing excellent infrastructure and a conducive atmosphere for completing this project successfully.

We are also extremely thankful to our Project Coordinator **Mrs. J. Aruna Shanthi**, Assistant Professor, Dept. of IT, for her valuable suggestions and interest throughout the course of this project.

We convey our heartfelt thanks to the lab staff for allowing us to use the required equipment whenever needed.

Finally, we would like to take this opportunity to thank our families for their support all through the work. We sincerely acknowledge and thank all those who gave directly or indirectly their support in completion of this work.

Sabrina Shaik 17261A1251

ABSTRACT

In this work, we present a system to recognize COVID-19 spreaders utilizing the examination of the connection between socio-socialand connection characteristics, such as GDP, life expectancy, number of air travel, and budget for health, among others including financial attributes with the quantity of contaminations and passing's brought about by the COVID-19 infection in various nations.

For this, we break down the data of each nation utilizing the mind-boggling systems approach, explicitly by examining the spreaders nations dependent on the separator set in 5-layer multiplex systems.

This methodology can help identify the spreader countries and obtains a classification of the countries based on their characteristics, where, in the spreaders set, the countries have high, medium or low values in the different socio-cultural and economic aspects; however, the characteristic that everyone shares are the high value in air connections. To mitigate a second outbreak of COVID-19 in the world, the countries that are in the union of both separator sets must reinforce their sanitary measures; in contrast, the countries that are at the intersection of the two separate sets, in addition to improving their sanitary measures, must regulate airflow to contain the spread of the disease.

Based on the information collected and modeled (until May 15, 2020), we can affirm that, by changing the relationships of the air flow, the risk of a second outbreak of COVID 19 can be minimized.

The outcomes show that, we acquire a characterization of the nations considering their numerical qualities in economics, populace, Gross Domestic Product (GDP), wellbeing and air associations; where, in the spreader set there are those nations that have high, medium or low qualities in the various attributes; nonetheless, the viewpoint that all the nations having a place with the separator set offer is a high incentive in air associations.

LIST OF FIGURES

TABLE OF CONTENTS

INTRODUCTION

In the current year (2020), the world has faced a disease caused by SARS-CoV2, known as COVID-19. This virus began to spread in Wuhan China on December 31, 2019 and, because the virus has spread rapidly throughout the world, it has been determined as a pandemic by the World Health Organization (WHO) in January 2020. COVID-19 is classified as a virus composed of single stranded RNA strands, and the symptoms of the disease caused by COVID-19 are headache, dry cough, malaise, fever, and respiratory failure. However, cases of asymptomatic people against the virus have been identified, which implies a real challenge for health institutions.

As of March 31, 2020, the United States has become the epicenter of the pandemic, followed by Italy, Spain, China, and Germany with 186,265, 105,792, 95,923, 82,278, and 71,690 confirmed cases, respectively. These five countries represent 63.4% of the total confirmed cases worldwide. The recovery percentages of this new pandemic are led by China, Spain, Germany, Italy and Iran with 42.8, 10.8, 9.1, 8.8, and 8.2% of cases. It is essential to mention that in Europe and America, the spread of COVID-19 throughout the world can be attributed to differences in health infrastructure, air travel, human development, and other socio-cultural and economic factors.

Therefore, in this work, we analyze the effect of sociocultural and economic factors, air travels, human development on both spreading and growth of COVID-19 in each country, using the Vertex Separator Problem (VSP) in multiplex complex networks. Here, it is essential to mention that the data for each topic was obtained from the websites of the European Union (EU), the World Health Organization (WHO), the World Bank (WB), the International Monetary Foundation (IMF) and Transparency International (TI) until May 15, 2020.

On the other hand, a complex network is a network with non-trivial topological characteristics that do not occur in simple networks, such as degree distributions, hierarchical structures, community structures, and high local cohesiveness (measured through the clustering coefficients).

Currently, there are several measures to identify influential or spreaders nodes in complex networks, where the most classic and important are: Closeness Centrality (CC) it Classifies the importance based on the inverse sum of the shortest distances to all other nodes from a central node, measuring the global structure. Degree centrality (DC) it Classifies each node based on its degree; therefore, the more connections a particular node has, the more important it is considered. Betweenness Centrality (BC) it Classifies the nodes by the number of the shortest paths between any other pair of nodes that cross through it.

In recent years, the study of complex networks and specifically of multilayer networks has been emphasized; this is thanks to the fact that most real systems have structures with multiple types of links or interactions between nodes; for example: Multimodal transport systems, biological systems, social networks and numerous modes of communication. Multiplex networks are a particular class of multilayer networks, which were introduced to better model complex real-world systems. The main characteristic of multiplex networks is that all the nodes in each layer are replicated in other layers, and there is a direct link between each replica node to denote the relationship.

The information of COVID-19 (infections and deaths), includes the period from 12/31/19 to 05/15/20 and the indicators used to define the socio-cultural-economic characteristics, are: projected real Gross Domestic Product (GDP) (2020), projected consumer prices (2020), special drawing rights (millions), quota (millions), human development index (HDI) (2009-2018), corruption perception index (2018), Gross national income (GNI) per capita, GNI per capita rank minus HDI rank, country population (population/km2), real population density (hectares by person), Gini Coefficient, current health expenditure (% of GDP 2000-2016) and air travels (20192020).

1.1 OBJECTIVE

we present a methodology to identify COVID-19 spreaders using the analysis of the relationship between socio-cultural and economic characteristics with the number of infections and deaths caused by the COVID-19 virus in different countries.

For this, we analyze the information of each country using the complex networks approach, specifically by analyzing the spreaders countries based on the separator set in 5-layer multiplex networks.

The results show that, we obtain a classification of the countries based on their numerical values in socioeconomics, population, Gross Domestic Product (GDP), health and air connections; where, in the spreader set there are those countries that have high, medium or low values in the different characteristics. the aspect that all the countries belonging to the separator set share is a high value in air connections.

1.2 PROBLEM DEFINITION

Obtain a classification of the countries based on their numerical values in socioeconomics, population, Gross Domestic Product (GDP), health and air connections. Where, in the spreader set there are those countries that have high, medium or low values in the different characteristics.

1.3 EXISTING SYSTEM

The existing system uses **Time Series Analysis** for the prediction of the possible covid cases for the next 15 days.

ARIMA is a model is used which is used for predicting future trends on a time series data.

1.4 PROPOSED SYSTEM

We present a methodology to identify COVID-19 spreaders using the analysis of the relationship between socio-cultural and economic characteristics with the number of infections and deaths caused by the COVID-19 virus in different countries. Using 5-layer multiplex network GDP, life expectancy, number of air travel, and budget for health.

We tend to integrate both ARIMA and PROPHET models which together provide an accurate result for non-linear trends.

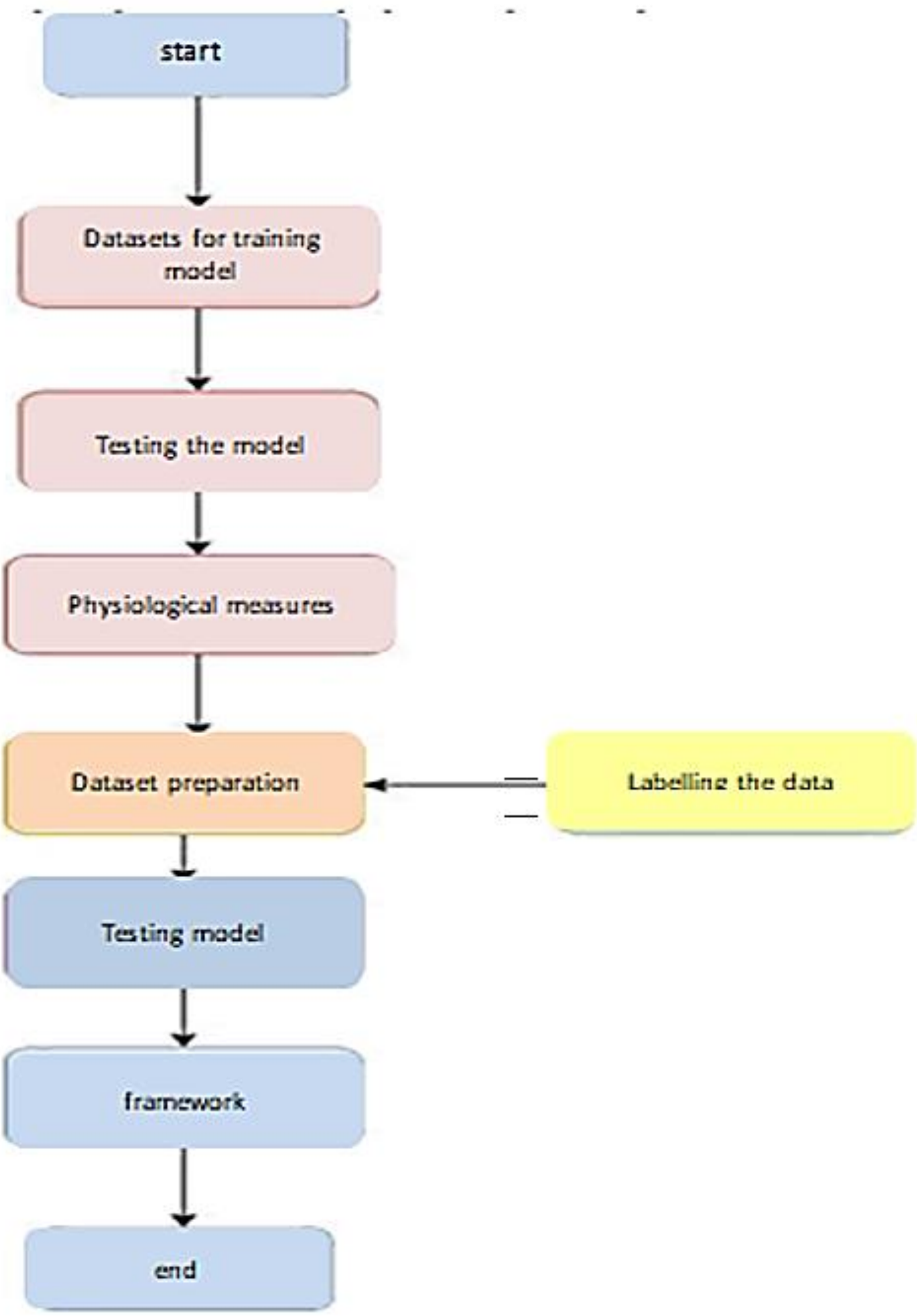

Figure no 1.4 Proposed model

LITERATURE SURVEY

1. C. Sohrabi et al., “World health organization declares global emergency:

A review of the 2019 novel coronavirus (covid-19),” International Journal of Surgery, 2020.

An unprecedented outbreak of pneumonia of unknown aetiology in Wuhan City, Hubei province in China emerged in December 2019. A novel coronavirus was identified as the causative agent and was subsequently termed COVID-19 by the World Health Organization (WHO). Considered a relative of severe acute respiratory syndrome (SARS) and Middle East respiratory syndrome (MERS), COVID-19 is caused by a beta coronavirus named SARS-CoV-2 that affects the lower respiratory tract and manifests as pneumonia in humans. Despite rigorous global containment and quarantine efforts, the incidence of COVID-19 continues to rise, with 90,870 laboratory-confirmed cases and over 3,000 deaths worldwide. In response to this global outbreak, we summarize the current state of knowledge surrounding COVID-19.

The onset of the novel Coronavirus Disease 2019 (COVID-19) outbreak in Wuhan, China, suggests animal-to-person spread and later person-to-person spread. The complete clinical picture following COVID-19 infection is not yet fully understood. A recent report on over 72,000 COVID-19 cases by the Chinese Center for Disease Control and Prevention showed the case fatality rate was overall 2.3%. The mortality rises to 8% in patients between 70 and 79-years-old, and spikes to 14.8% in those aged 80 and above [1]. Sorhabi et al. give an informative and comprehensive account of the timeline, etiology, symptoms, supportive treatment, and transmission prevention of COVID-19 [2].

The WHO's declaration of COVID-19 to be a Public Health Emergency of International Concern is attributable to the high case fatality rates in China and the global economic effect of COVID-19, which may compound the current ongoing influenza epidemic[3]. Furthermore, there is the potential for higher death rates in countries with vulnerable health systems in resource-limited regions. The ability to control local transmission depends on the application of the principles of rapid identification, prevention, and control, followed by patient isolation, rapid diagnosis, and contact tracing. Some countries remain ill-equipped with limited diagnostic capacity, resulting in delays from suspected case identification to vector confirmation and patient isolation, which increases the risk of disease transmission. [4] Though, 74% of countries in Africa have an influenza pandemic preparedness plan; however, most are outdated and inadequate to deal with a global pandemic such as COVID-19 [5].

Recent epidemics and pandemics (e.g., severe acute respiratory syndrome (SARS), H1N1 pandemic, Middle East respiratory syndrome (MERS), and Ebola) have highlighted the need to reinforce national public health capabilities, including diseases surveillance systems and health care workforce. Pandemic preparedness requires specific training in surveillance, epidemic response, and diagnostic testing. Strengthening health care systems worldwide, particularly in resource-poor settings, is imperative. Recurrent novel pandemics is our new global reality

The Internet plays a large role in disseminating anti-vaccination information. This paper builds upon previous research by analyzing the arguments proffered on antivaccination websites, determining the extent of misinformation present, and examining discourses used to support vaccine objections. Arguments around the themes of safety and effectiveness, alternative medicine, civil liberties, conspiracy theories, and morality were found on the majority of websites analyzed; misinformation was also prevalent. The most commonly proposed method of combating this misinformation is through better education, although this has proven ineffective. Education does not consider the discourses supporting vaccine rejection, such as those involving alternative explanatory models of health, interpretations of parental responsibility, and distrust of expertise. Anti-vaccination protestors make postmodern arguments that reject biomedical and scientific "facts" in favors of their own interpretations. Pro-vaccination advocates who focus on correcting misinformation reduce the controversy to merely an "educational" problem; rather, these postmodern discourses must be acknowledged in order to begin a dialogue. With morbidity and mortality from vaccine-preventable diseases [VPDs] having reached record lows [1], vaccines are one of the most successful tools for biomedical science and public health. Yet paradoxically, the effectiveness of vaccination has led to the reemergence of anti-vaccination sentiments. Vaccines may be seen as unnecessary or dangerous because incidence rates of VPDs in developed countries have plummeted. Vaccine "reactions" – negative health events following vaccination, attributed to the vaccine – then appear to be more common than the diseases themselves [2]. In this way, vaccines can be considered victims of their own success.

The media plays a large role in disseminating and sensationalizing vaccine objections. Such objections are part of what has been called the “anti-vaccination movement”, which has had a demonstrable impact on vaccination policies, and individual and community health [3].

A common sequence to vaccination scares involves scientific debate about potential vaccine risks, which communication technology transmits via a rhetoric of doubt; parents incorporate this with personal experiences and spread their views to their social groups [4]. These social groups exert considerable pressure on vaccination decisions by creating a “local vaccination culture” [5]. With the prominence of the Internet in today’s world, ∗ Correspondence address: 110 Parkwood Cres., Hamilton, Ontario L8V 4Z7, Canada. Tel.: +1 905 387 3141. E-mail addresses: aniakata@gmail.com, kataa@mcmaster.ca. the attitudes, beliefs, and experiences of that local culture can quickly become global. Internet usage statistics show approximately 74% of Americans and 72% of Canadians are online [6].

An estimated 75–80% of users search for health information online [7]. Of these users, 70% say the information they encounter online influences their treatment decisions [8]. In 2006, 16% of users searched online for information on immunizations or vaccinations [9]. While online research is more convenient and accessible than reading medical literature or visiting health practitioners, too great a reliance on Internet-based information can be problematic. Over half (52%) of users believe “almost all” or “most” information on health websites is credible [8]; yet the availability of inaccurate and deceptive information online has labelled the Internet a “modern Pandora’s box” The nature of the Internet allows any and all opinions to spread widely and instantaneously.

Individuals and groups gain exposure online without being filtered or reviewed – and anti-vaccination advocates have taken advantage of this fact. Anti-vaccination messages are more common on the Internet than in other forms of media, increasing the likelihood that vaccination decisions may be based on misleading information. Indeed, parents who exempt children from vaccination are more likely to have obtained information from the Internet than parents who have their children vaccinated; they are also more likely to have used certain antivaccination websites.

This demonstrates the importance of understanding what messages are presented online and why they may be accepted. The body of research examining online anti vaccinations is not large, nor has there been a recent update. Only on reanalysis examined misinformation and deception on such sites, but was not quantitative. Prior research also acknowledged the need to understand discourses underlying antivaccination arguments but did not elaborate upon them.

This analysis aims to address these issues by answering two main questions. First, what information is proffered on anti-vaccination websites, and what is its accuracy? Second, what discourses make these vaccine objections appealing? Data collection Web searches were conducted on May 21, 2009 using the terms "vaccine", "vaccination", and "immunization OR immunization" input into Google.com (the American version of the search engine) and Google.ca (the Canadian version). Google was chosen as it is the most popular search engine, accounting for 73% of all Internet searches. Results were classified as anti-vaccination and included for content analysis if they opposed childhood vaccination for any reason, without meeting any of the following exclusion criteria: listserv or newsgroup pages; pages solely containing brief notices about other website content; news results, medical journals or library sites; video results; book previews non-English sites; sites exclusively about adult immunization; (8) sites exclusively about veterinary vaccination and inactive links. Criteria were applied to the anti-vaccination websites and coded as present or absent. Criteria were adapted from previous online antivaccination studies, as well as created by the author.

Online health information seekers examine the first 10 search results 97.2% of the time; therefore, only the first 10 results retrieved per term were examined. Of 30 total Google.com results, 5 of 21 immunization sites (24%) were classified as antivaccination. Of 30 total Google.ca results, 2 of 16 immunization sites (13%) were classified as anti-vaccination. To amass additional websites for a more meaningful study, the Canadian searches were extended to 50 results per term. Of 150 total results, 5 of 86 immunization sites (6%) were classified as anti-vaccination (two were duplicates of American results). Combining the American and Canadian results, 8 anti-vaccination websites were subjected to content analysis. Overall, American searches returned more anti-vaccination results (24%) than Canadian searches (6%), indicating American parents are more likely to encounter anti-vaccination sites via Google than are Canadian parents. Neither search engine returned any anti-vaccination results for "immunization OR immunization"; this was expected based on research that found anti-vaccination groups avoid using the term “immunization” as they tend not to believe that vaccines confer immunity [16]. Although prior studies returned more search results [11], this does not necessarily mean the number of anti-vaccination websites has decreased, but rather that their search rankings may have shifted. Nevertheless, the proportion of sites retrieved for some search terms is notable – 71% of results from the Google.com “vaccination” search were classified as antivaccination.

Safety themes were present on all anti-vaccination websites analyzed. Every site claimed vaccines are poisonous and cause idiopathic illnesses. Sites stressed that vaccines contain substances poisonous to humans, including anti-freeze, ether, formaldehyde, mercury, and nanobacteria. Pertinent information was not elaborated upon – for instance, that the amount of potentially harmful substances in vaccines is not enough to produce toxic effects in humans, or that ether does not refer to the anesthetic but to a chemical compound. Illnesses attributed to vaccines included:

SYSTEM REQUIREMENTS SPECIFICATION

INTRODUCTION

- A condition or capability needed by a user to solve a problem or achieve an objective.
- A condition or capability that must be met or possessed by a system or system component to satisfy a contract, standard, specification or other formally imposed documents.

3.1 FUNCTIONAL REQUIREMENTS

1.Data Collection

2.Data Preprocessing

3.Training and Testing

4.Modeling

5.Predicting

3.2 NON-FUNCTIONAL REQUIREMENTS

NON-FUNCTIONAL REQUIREMENT (NFR) specifies the quality attribute of a software system. They judge the software system based on Responsiveness, Usability, Security, Portability and other non-functional standards that are critical to the success of the software system. Example of nonfunctional requirement, “how fast does the website load?” Failing to meet non-functional requirements can result in systems that fail to satisfy user needs. Non- functional Requirements allows you to impose constraints or restrictions on the design of the system across the various agile backlogs. Example, the site should load in 3 seconds when the number of simultaneous users are> 10000. Description of non-functional requirements is just as critical as a functional requirement.

• Usability requirement

• Serviceability requirement

- Manageability requirement
- Recoverability requirement
- Security requirement
- Data Integrity requirement
- Capacity requirement
- Availability requirement
- Scalability requirement
- Interoperability requirement
- Reliability requirement
- Maintainability requirement
- Regulatory requirement
- Environmental requirement

3.3 HARDWARE REQUIREMENTS

Minimum hardware requirements are very dependent on the particular software being developed by a given Enthought Python / Canopy / VS Code user. Applications that need to store large arrays/objects in memory will require more RAM, whereas applications that need to perform numerous calculations or tasks more quickly will require a faster processor.

- Operating system : windows, Linux
- Processor : minimum intel i3
- Ram : minimum 4gb
- Hard disk : minimum 250gb

3.4 SOFTWARE REQUIREMENTS

The functional requirements or the overall description documents include the product perspective and features, operating system and operating environment, graphics requirements, design constraints and user documentation.

The appropriation of requirements and implementation constraints gives the general overview of the project in regards to what the areas of strength and deficit are and how to tackle them.

- Python idle 3.7 version (or)
- Anaconda 3.7 (or)
- Jupyter (or)
- Google colab

SYSTEM ANALYSIS

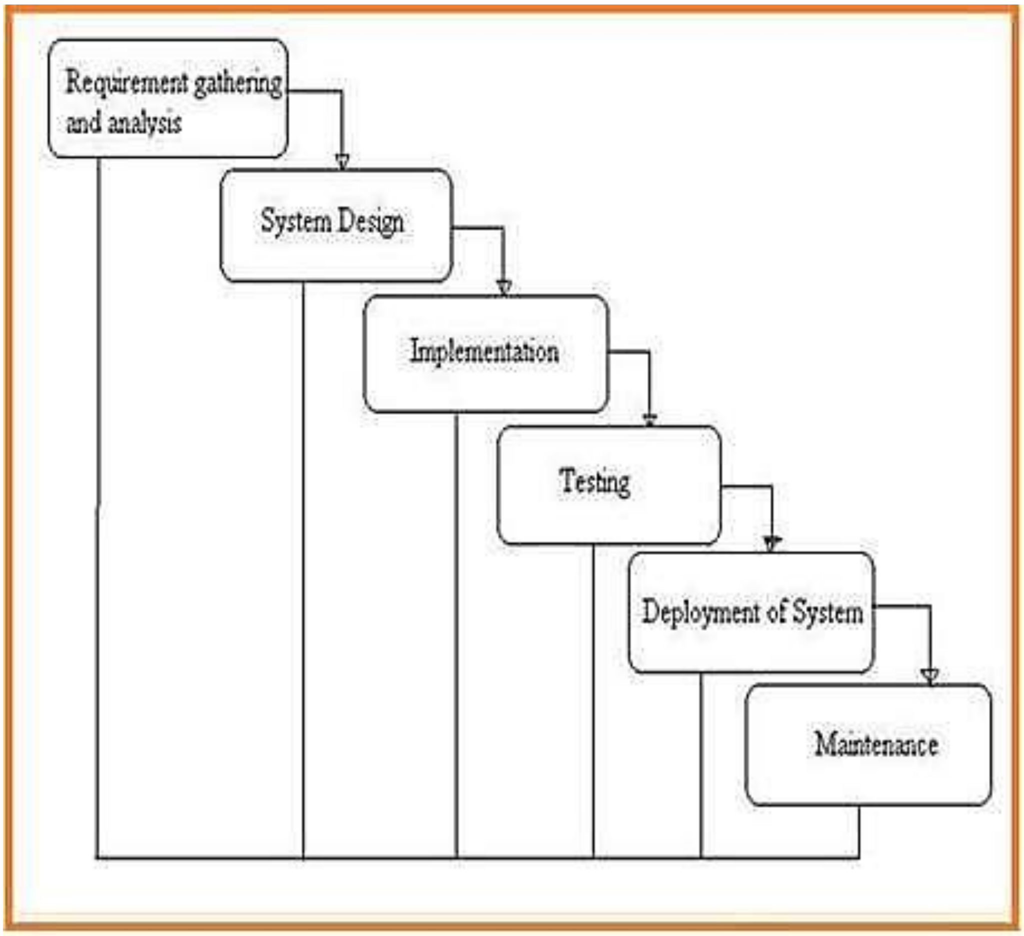

Fig.no. 4.0 Structure of Project

4.1 MODULES DESCRIPTION

OrdinalEncoder : To maximize the use of dataset

Seaborn: To plot worldwide covid 19 cases
Matplotlib: To plot total Confirmed Cases in million, Comparison With other Countries

ARIMA & PROPHET: To plot predicting cases for next 15 days
RandomForestRegressor: To calculate mse,rmse

Mean_squared_error,rmse: Metrices used for accuracy

Flask: To design UI

4.2 PROPOSED ARCHITECTURE/ BLOCK DIAGRAM

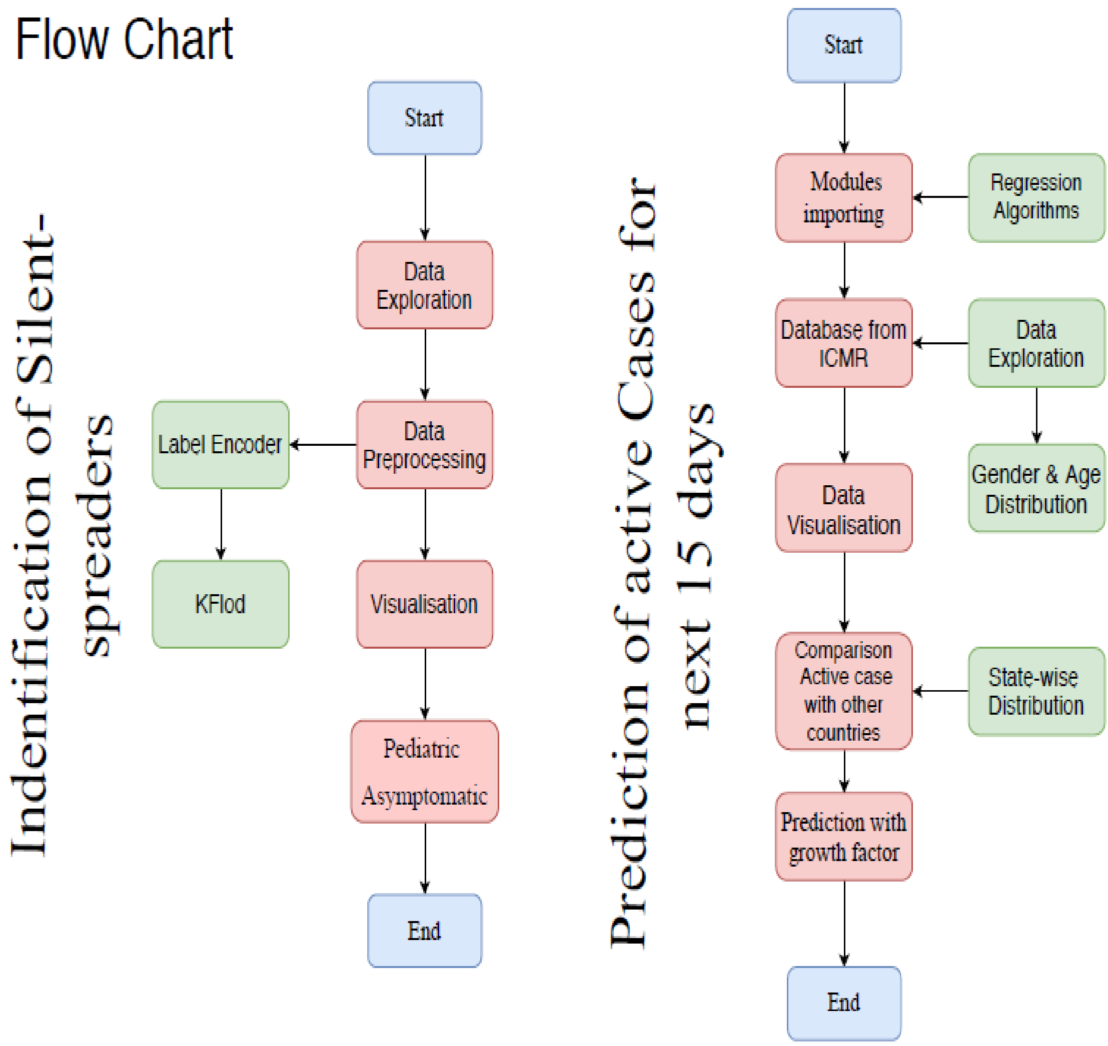

Fig.no. 4.1 Proposed architecture

SYSTEM DESIGN

INTRODUCTION

- System design is divided into three types like GUI designing.
- UML designing with avails in development of project in facile way with different actor and its utilizer case by utilizer case diagram, flow of the project utilizing sequence.
- Class diagram gives information about different class in the project with methods that have to be utilized in the project.
- The third and most import for the project in system design is Data base design where we endeavor to design data base predicated on the number of modules in our project

5.1 SYSTEM COMPONENTS

- **Input**-The input is live data feed from server
- **Output**-The output is the accuracy in identifying Covid-19 spreaders
- **Key Technology**-Machine learning, Deep learning, Python packages
- **Implementation**-
  - Gathering the Covid database
  - Pre-processing the dataset and analysis dataset
  - Training the models to analysis the documents
  - Obtain the accuracy in prediction
- **Data Sets**- The datasets is download from kaggle

5.2 UML DIAGRAMS

The System Design Document describes the system requirements, operating environment, system and subsystem architecture, files and database design, input formats, output layouts, human-machine interfaces, detailed design, processing logic, and external interfaces.

Global Use Case Diagrams:

Identification of actors:

Actor: Actor represents the role a user plays with respect to the system. An actor interacts with, but has no control over the use cases.

The actors identified in this system are:

- System Administrator
- Customer
- Customer Care
- Identification of use cases:

Use case: A use case can be described as a specific way of using the system from a user's (actor's) perspective.

Graphical representation:

- A more detailed description might characterize a use case as:
- Pattern of behavior the system exhibits
- A sequence of related transactions performed by an actor and the system
- Delivering something of value to the actor Use cases provide a means to:
- capture system requirements
- communicate with the end users and domain experts
- test the system
- Guide lines for identifying use cases:
- For each actor, find the tasks and functions that the actor should be able to perform or that the system needs the actor to perform. The use case should represent a course of events that leads to clear goal ❍Name the use cases.
- Describe the use cases briefly by applying terms with which the user is familiar. This makes the description less ambiguous

Flow of Events

A flow of events is a sequence of transactions (or events) performed by the system. They typically contain very detailed information, written in terms of what the system should do, not how the system accomplishes the task. Flow of events are created as separate files or documents in your favorite text editor and then attached or linked to a use case using the Files tab of a model element.

A flow of events should include:

- When and how the use case starts and ends
- Use case/actor interactions
- Data needed by the use case
- Normal sequence of events for the use case
- Alternate or exceptional flows Construction of Use case diagrams:

Use-case diagrams graphically depict system behavior (use cases). These diagrams present a high level view of how the system is used as viewed from an outsider's (actor's) perspective. A use-case diagram may depict all or some of the use cases of a system.

A use-case diagram can contain:

- actors ("things" outside the system)

- use cases (system boundaries identifying what the system should do)
- Interactions or relationships between actors and use cases in the system including the associations, dependencies, and generalizations.

Relationships in use cases:

1. **Communication:**

The communication relationship of an actor in a use case is shown by connecting the actor symbol to the use case symbol with a solid path. The actor is said to communicate with the use case.

2. **Uses:**

A Uses relationship between the use cases is shown by generalization arrow from the use case.

3.Extends:

The extend relationship is used when we have one use case that is similar to another use case but does a bit more. In essence it is like subclass.

SEQUENCE DIAGRAMS

A sequence diagram is a graphical view of a scenario that shows object interaction in a time- based sequence what happens first, what happens next. Sequence diagrams establish the roles of objects and help provide essential information to determine class responsibilities and interfaces. There are two main differences between sequence and collaboration diagrams: sequence diagrams show time-based object interaction while collaboration diagrams show how objects associate with each other. A sequence diagram has two dimensions: typically, vertical placement represents time and horizontal placement represents different objects.

Object:

An object has state, behavior, and identity. The structure and behavior of similar objects are defined in their common class. Each object in a diagram indicates some instance of a class. An object that is not named is referred to as a class instance.

The object icon is similar to a class icon except that the name is underlined: An object's concurrency is defined by the concurrency of its class.

Message:

A message is the communication carried between two objects that trigger an event. A message carries information from the source focus of control to the destination focus of control.

Link:

A link should exist between two objects, including class utilities, only if there is a relationship between their corresponding classes. The existence of a relationship between two classes symbolizes a path of communication between instances of the classes: one object may send messages to another. The link is depicted as a straight line between objects or objects and class instances in a collaboration diagram. If an object links to itself, use the loop version of the icon.

CLASS DIAGRAM:

Identification of analysis classes:

A class is a set of objects that share a common structure and common behavior (the same attributes, operations, relationships and semantics). A class is an abstraction of real-world items. There are 4 approaches for identifying classes:

a. Noun phrase approach:
b. Common class pattern approach.

c. Use case Driven Sequence or Collaboration approach.

d. Classes , Responsibilities and collaborators Approach

1. Noun Phrase Approach:

The guidelines for identifying the classes:

- Look for nouns and noun phrases in the use cases.
- Some classes are implicit or taken from general knowledge.
- All classes must make sense in the application domain; Avoid computer implementation classes – defer them to the design stage.
- Carefully choose and define the class names After identifying the classes we have to eliminate the following types of classes:
- Adjective classes.

2. Common class pattern approach:

The following are the patterns for finding the candidate classes:

- Concept class.
- Events class.
- Organization class
- Peoples class
- Places class
- Tangible things and devices class.

3. Use case driven approach:

We have to draw the sequence diagram or collaboration diagram. If there is need for some classes to represent some functionality then add new classes which perform those functionalities.

4. CRC approach:

The process consists of the following steps:

- Identify classes' responsibilities (and identify the classes)
- Assign the responsibilities
- Identify the collaborators. Identification of responsibilities of each class:

The questions that should be answered to identify the attributes and methods of a class respectively are:

- What information about an object should we keep track of?
- What services must a class provide? Identification of relationships among

Three types of relationships among the objects are:

Association: How objects are associated?

Super-sub structure: How are objects organized into super classes and sub classes?

Aggregation: What is the composition of the complex classes?

Association:

The **questions** that will help us to identify the associations are:

a. Is the class capable of fulfilling the required task by itself?

b. If not, what does it need?

c. From what other classes can it acquire what it needs? Guidelines for identifying the tentative associations:

- A dependency between two or more classes may be an association. Association often corresponds to a verb or prepositional phrase.
- A reference from one class to another is an association. Some associations are implicit or taken from general knowledge.

Some common association patterns are:

Location association like part of, next to, contained in….. Communication association like talk to, order to ……

We have to eliminate the unnecessary association like implementation associations, ternary or n- ary associations and derived associations.

Super-sub class relationships:

Super-sub class hierarchy is a relationship between classes where one class is the parent class of another class (derived class).This is based on inheritance.

Guidelines for identifying the super-sub relationship, a generalization are

1. Top-down:

Look for noun phrases composed of various adjectives in a class name. Avoid excessive refinement. Specialize only when the sub classes have significant behavior.

2. Bottom-up:

Look for classes with similar attributes or methods. Group them by moving the common attributes and methods to an abstract class. You may have to alter the definitions a bit.

3. Reusability:

Move the attributes and methods as high as possible in the hierarchy.

4. Multiple inheritances:

Avoid excessive use of multiple inheritances. One way of getting benefits of multiple inheritances is to inherit from the most appropriate class and add an object of another class as an attribute.

Aggregation or a-part-of relationship:

It represents the situation where a class consists of several component classes. A class that is composed of other classes doesn't behave like its parts. It behaves very difficultly. The major properties of this relationship are transitivity and anti symmetry.

The questionswhose answers will determine the distinction between the part and whole relationships are:

- Does the part class belong to the problem domain?
- Is the part class within the system's responsibilities?
- Does the part class capture more than a single value?(If not then simply include it as an attribute of the whole class)
- Does it provide a useful abstraction in dealing with the problem domain? There are three types of aggregation relationships. They are:
- Assembly:
- It is constructed from its parts and an assembly-part situation physically exists.
- Container:
- A physical whole encompasses but is not constructed from physical parts.
- Collection member:
- A conceptual whole encompasses parts that may be physical or conceptual. The container and collection are represented by hollow diamonds but composition is represented by solid diamond.

5.3.1 Use Case Diagram

A use case diagram in the Unified Modeling Language (UML) is a type of behavioral diagram defined by and created from a Use-case analysis. Its purpose is to present a graphical overview of the functionality provided by a system in terms of actors, their goals (represented as use cases), and any dependencies between those use cases. The main purpose of a use case diagram is to show what system functions are performed for which actor. Roles of the actors in the system can be depicted.

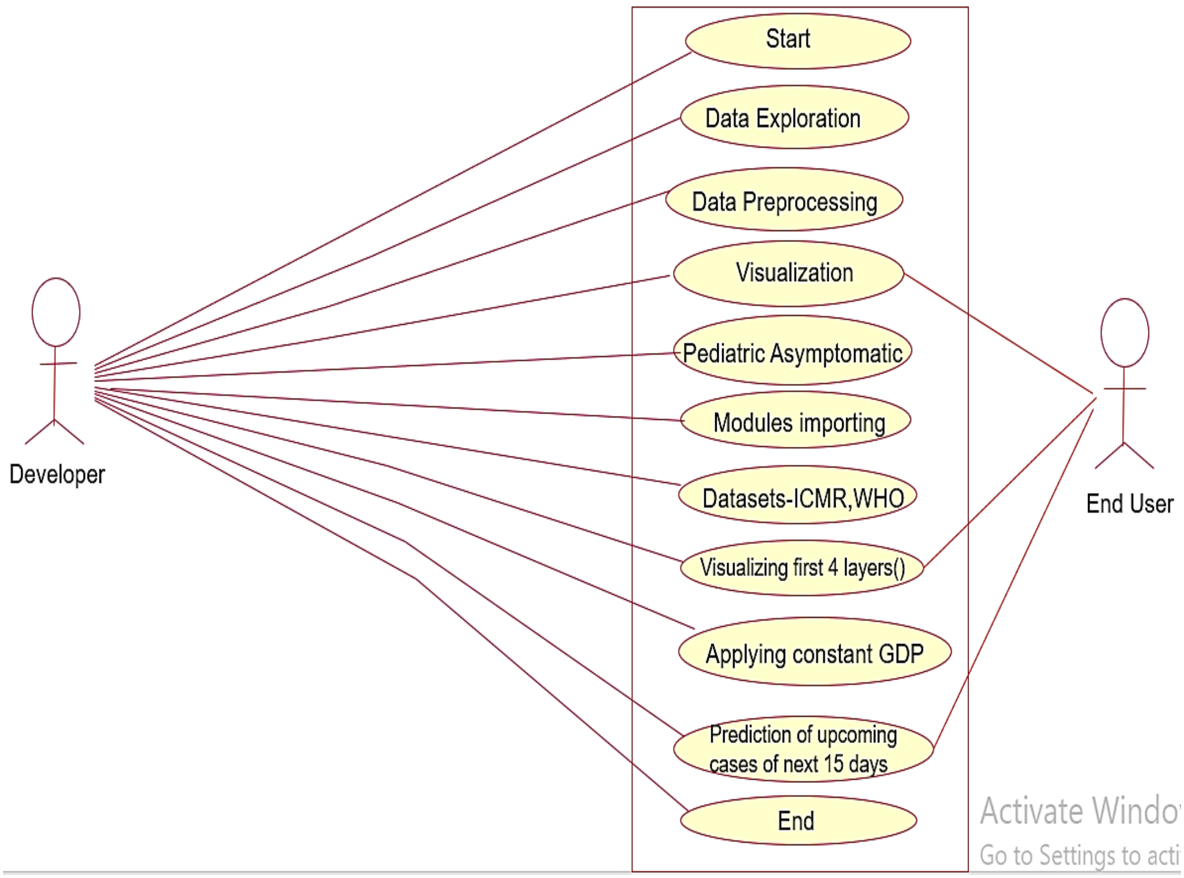

Fig.no. 5.2.1 Use Case Diagram

5.3.2 Class Diagram

In software engineering, a class diagram in the Unified Modeling Language (UML) is a type of static structure diagram that describes the structure of a system by showing the system's classes, their attributes, operations (or methods), and the relationships among the classes. It explains which class contains information.

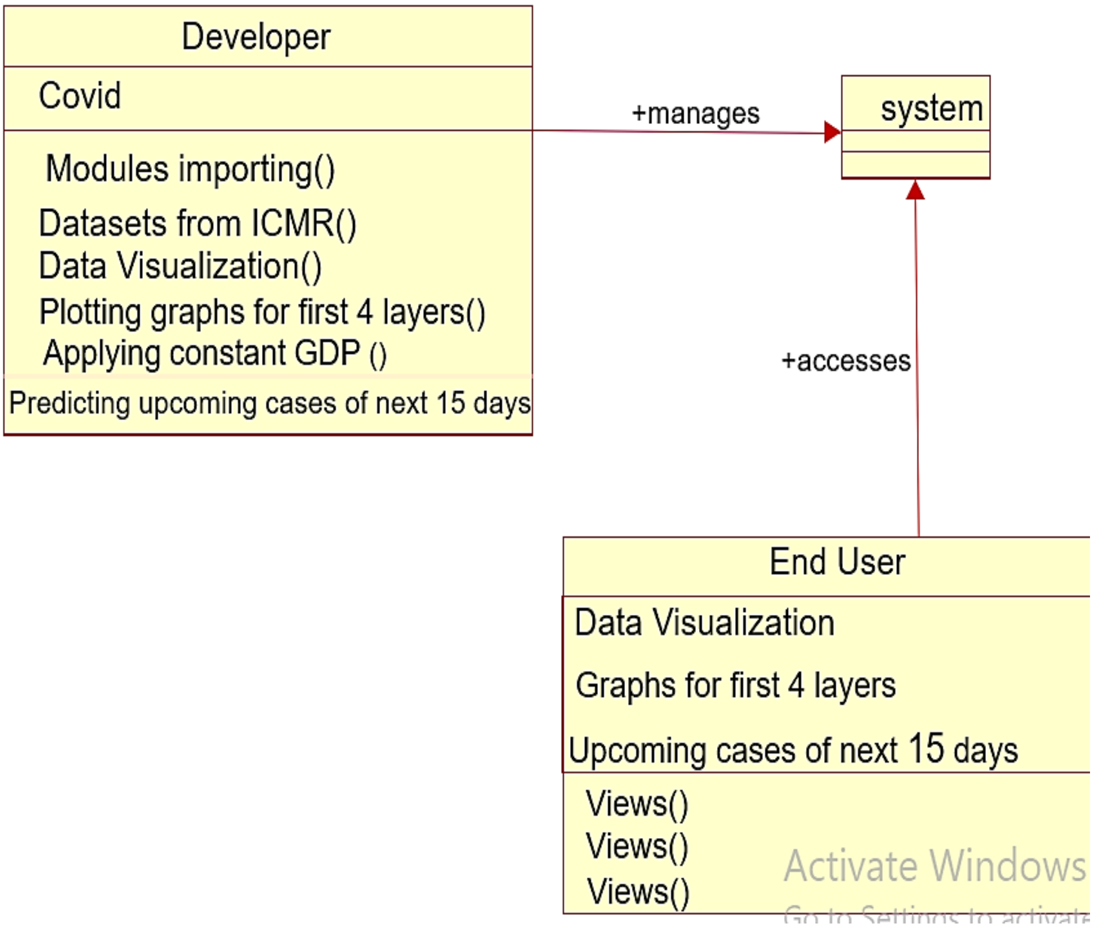

Fig.no. 5.2.2 Class Diagram

5.3.3 Sequence Diagram

A sequence diagram in Unified Modeling Language (UML) is a kind of interaction diagram that shows how processes operate with one another and in what order. It is a construct of a Message Sequence Chart. Sequence diagrams are sometimes called event diagrams, event scenarios, and timing diagrams.

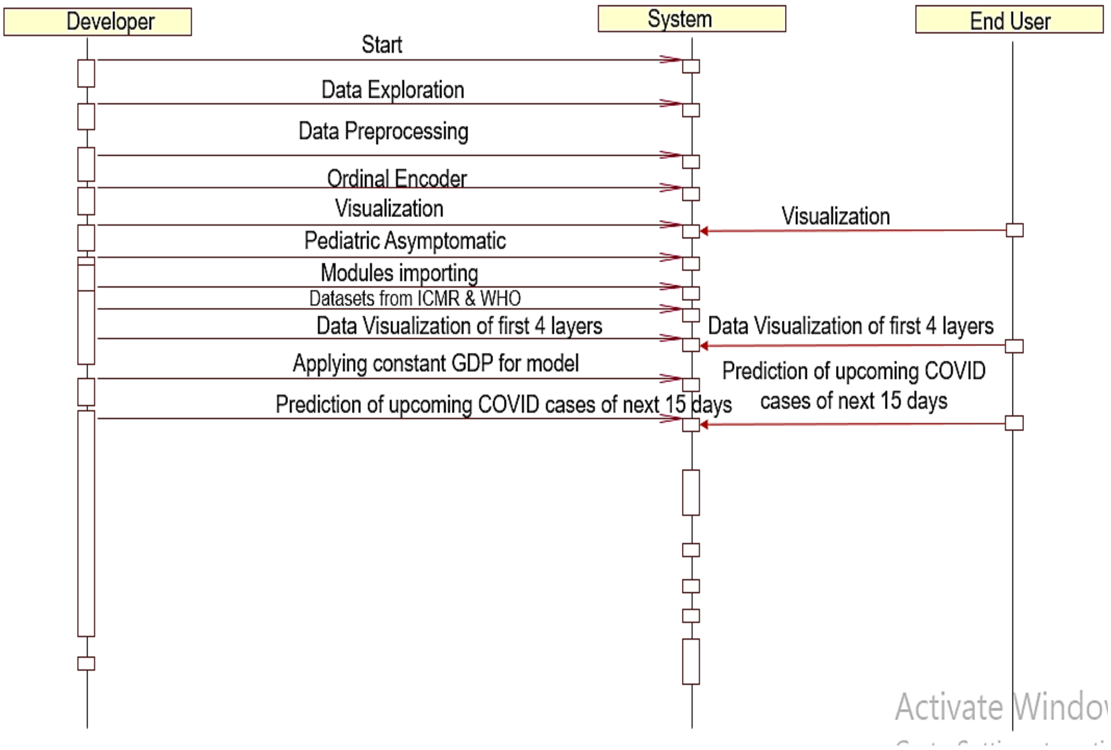

Fig.no. 5.2.3 Sequence Diagram

5.3.4 Activity Diagram

Activity diagram is basically a flowchart to represent the flow from one activity to another activity. The activity can be described as an operation of the system. The control flow is drawn from one operation to another.
It describe the dynamic aspects of the system.

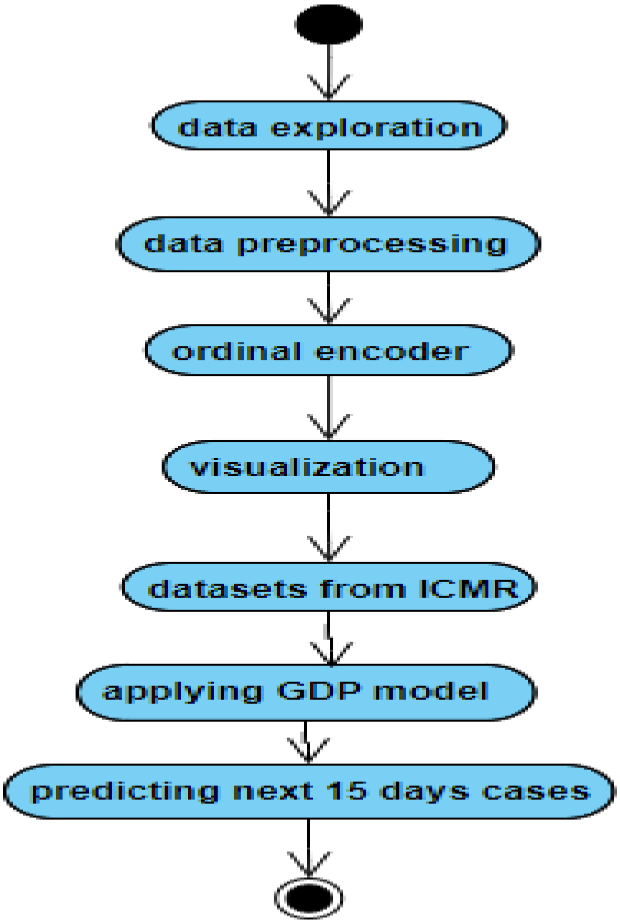

Fig.no.5.2.4 Activity diagram

5.3.5 . Component Diagram

The purpose of a component diagram is to show the relationship between different components in a system. the term "component" refers to a module of classes that represent independent systems or subsystems with the ability to interface with the rest of the system.

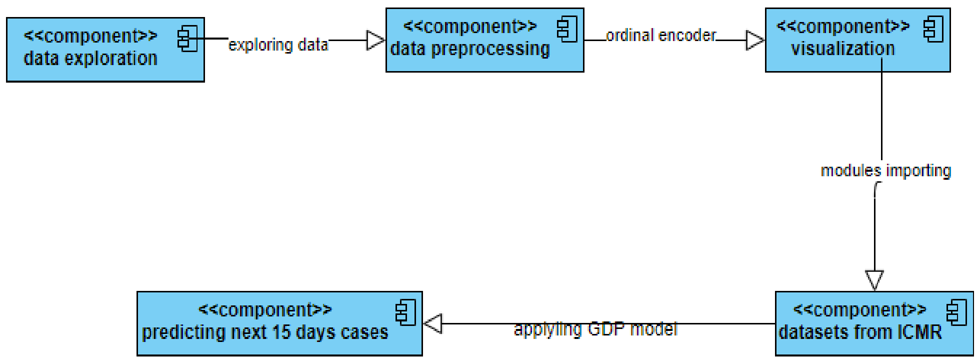

Figure no. 5.2.5 Component diagram

5.3.6 Collaboration Diagram

A collaboration diagram, also known as a communication diagram, is an illustration of the relationships and interactions among software objects in the Unified Modeling Language .These diagrams can be used to portray the dynamic behavior of a particular use case and define the role of each object.

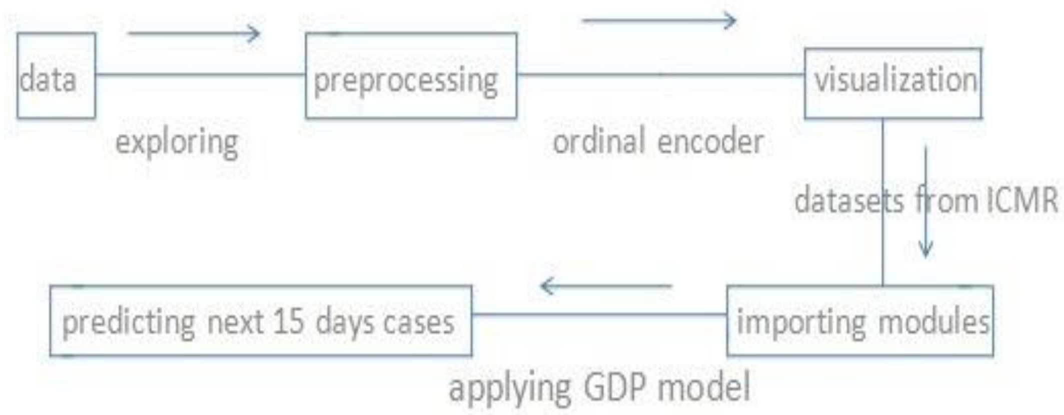

Figure no. 5.2.6 Collaboration

5.3.7 Deployment Diagram

A deployment diagram is a UML diagram type that shows the execution architecture of a system, including nodes such as hardware or software execution environments, and the middleware connecting them. Deployment diagrams are typically used to visualize the physical hardware and software of a system.

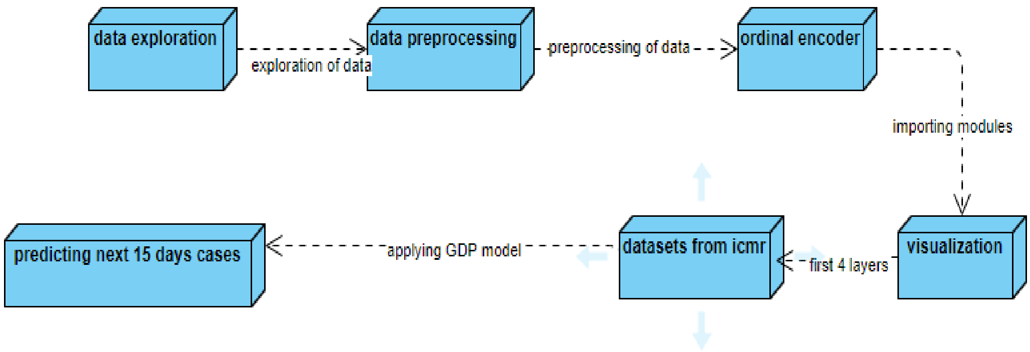

Figure no. 5.2.6 Deployment diagram

5.3 DATA DICTIONARY

S.no.	Acronym	Abbreviation
1	GDP	Gross Domestic Product
2	WHO	World Health Organization
3	VSP	Vertex Separator Problem
4	EU	European Union
5	WB	World Bank
6	IMF	International Monetary Foundation
7	CC	Closeness Centrality
8	DC	Degree Centrality
9	BC	Betweenness Centrality

10	NFR	Non-Functional Requirement
11	SARS	Severe Acute Respiratory Syndrome
12	MERS	Middle East Respiratory Syndrome
13	COVID-19	Corona Virus Disease 2019
14	VPDS	Vaccine-Preventable Diseases
15	CDC	Centers for Disease Control
16	MMR	Measles, Mumps, Rubella
17	SIDS	Sudden Infant Death Syndrome
18	UML	Unified Modeling Language

CODING & IMPLEMENTATION

6.1 SAMPLE CODE

```
import pandas as pd
import numpy as np
import datetime
import requests import
warnings

import matplotlib.pyplot as plt
import matplotlib import
matplotlib.dates as mdates import
seaborn as sns import squarify
import plotly.offline as py import
plotly_express as px

#Tensorflow for Multipplex Layer

import tensorflow
from keras.models import Sequential from
keras.layers import Dense

from xgboost import XGBRegressor from
lightgbm import LGBMRegressor
from sklearn.ensemble import RandomForestRegressor

from sklearn.metrics import mean_squared_error, mean_absolute_error from
sklearn.preprocessing import OrdinalEncoder

from sklearn.model_selection import train_test_split
from  statsmodels.tsa.arima_model  import  ARIMA
from fbprophet import Prophet
from fbprophet.plot import plot_plotly, add_changepoints_to_plot

from IPython.display import Image warnings.filterwarnings('ignore')
%matplotlib inline

age_details = pd.read_csv('/content/drive/My
Drive/covid_project/AgeGroupDetails.csv')
india_covid_19 = pd.read_csv('/content/drive/My
Drive/covid_project/covid_19_india.csv') hospital_beds
= pd.read_csv('/content/drive/My
Drive/covid_project/HospitalBedsIndia.csv')
individual_details = pd.read_csv('/content/drive/My
Drive/covid_project/IndividualDetails.csv')
```

```
ICMR_details = pd.read_csv('/content/drive/My Drive/covid_project/ICMRTestingDetails.csv')
ICMR_labs = pd.read_csv('/content/drive/My Drive/covid_project/ICMRTestingLabs.csv')
state_testing = pd.read_csv('/content/drive/My Drive/covid_project/StatewiseTestingDetails.csv')
population = pd.read_csv('/content/drive/My Drive/covid_project/population_india_census2011.csv')

world_population = pd.read_csv('/content/drive/My Drive/covid_project/population_by_country_2020.csv') confirmed_df
=
pd.read_csv('https://raw.githubusercontent.com/CSSEGISandData/COVID19/master/csse_covid_19_data/csse_covid_19_time_series/time_series_covid19_confirmed_global.csv') deaths_df
=
pd.read_csv('https://raw.githubusercontent.com/CSSEGISandData/COVID19/master/csse_covid_19_data/csse_covid_19_time_series/time_series_covid19_deaths_global.csv') recovered_df
=
pd.read_csv('https://raw.githubusercontent.com/CSSEGISandData/COVID19/master/csse_covid_19_data/csse_covid_19_time_series/time_series_covid19_recovered_global.csv')
latest_data =
pd.read_csv('https://raw.githubusercontent.com/CSSEGISandData/COVID19/master/csse_covid_19_data/csse_covid_19_daily_reports/04-04-2020.csv')

india_covid_19['Date'] = pd.to_datetime(india_covid_19['Date'],dayfirst = True)
state_testing['Date'] = pd.to_datetime(state_testing['Date'])
ICMR_details['DateTime'] =
pd.to_datetime(ICMR_details['DateTime'],dayfirst = True) ICMR_details
= ICMR_details.dropna(subset=['TotalSamplesTested',
'TotalPositiveCases'])

world_confirmed = confirmed_df[confirmed_df.columns[-1:]].sum()
world_recovered = recovered_df[recovered_df.columns[-1:]].sum()
world_deaths = deaths_df[deaths_df.columns[-1:]].sum() world_active
= world_confirmed - (world_recovered - world_deaths)

labels = ['Active','Recovered','Deceased'] sizes =
[world_active,world_recovered,world_deaths]
color= ['#66b3ff','green','red']
explode = []

for i in labels:
explode.append(0.05)

plt.figure(figsize= (15,10))
```

```
plt.pie(sizes, labels=labels, autopct='%1.1f%%', startangle=9, explode
=explode,colors = color)
centre_circle = plt.Circle((0,0),0.70,fc='white')

fig = plt.gcf()
fig.gca().add_artist(centre_circle)
plt.title('World COVID-19 Cases',fontsize = 20)
plt.axis('equal')   plt.tight_layout()

hotspots = ['China','Germany','Iran','Italy','Spain','US','Korea,
South','France','Turkey','UnitedKingdom','India']
dates = list(confirmed_df.columns[4:]) dates =
list(pd.to_datetime(dates)) dates_india =
dates[8:]

df1 = confirmed_df.groupby('Country/Region').sum().reset_index()
df2 = deaths_df.groupby('Country/Region').sum().reset_index() df3
= recovered_df.groupby('Country/Region').sum().reset_index()

global_confirmed = {}
global_deaths = {} global_recovered
= {}
global_active= {}

for country in hotspots:     k =df1[df1['Country/Region'] ==
country].loc[:,'1/30/20':]         global_confirmed[country] =
k.values.tolist()[0]

   k =df2[df2['Country/Region'] == country].loc[:,'1/30/20':]
global_deaths[country] = k.values.tolist()[0]

   k =df3[df3['Country/Region'] == country].loc[:,'1/30/20':]
global_recovered[country] = k.values.tolist()[0]

for country in hotspots:
   k = list(map(int._sub_, global_confirmed[country],
global_deaths[country]))global_active[country] = list(map(int._sub_, k,
global_recovered[country]))

fig = plt.figure(figsize= (15,15))
plt.suptitle('Active, Recovered, Deaths in Hotspot Countries and India as of May
15',fontsize = 20,y=1.0)
```

6.2 IMPLEMENTATION PROCEDURE

- Gathering the covid database
- Pre-processing the dataset and analysis dataset
- Training the models to analysis the documents
- Obtain the accuracy in prediction

RESULT

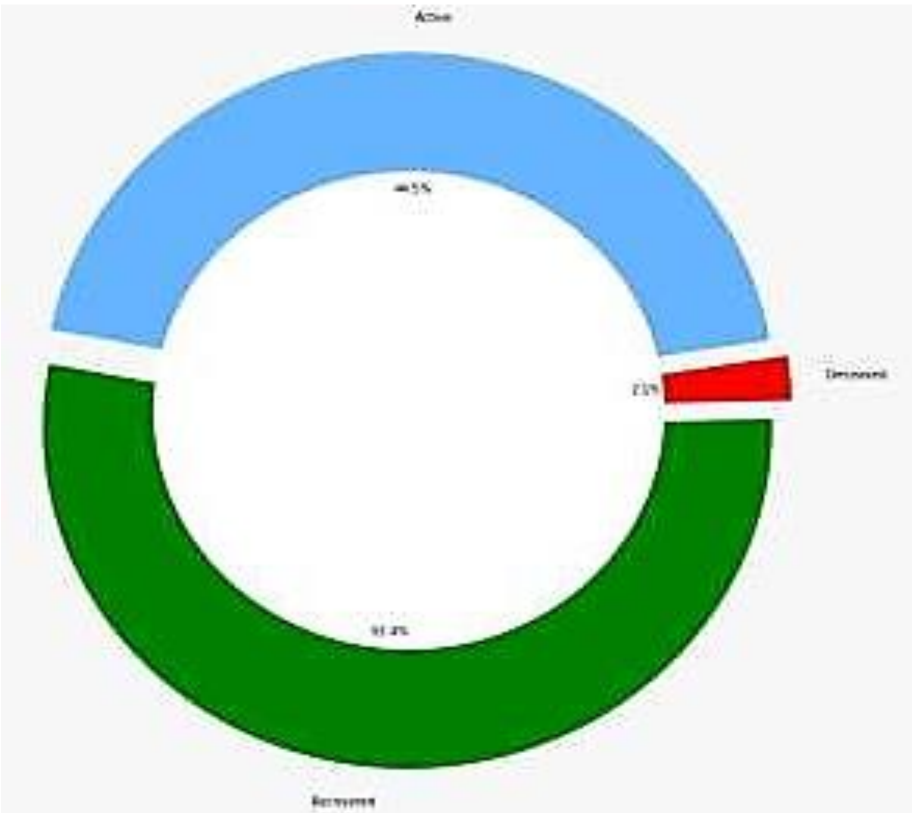

Fig. no. 7.1 World COVID-19 Cases

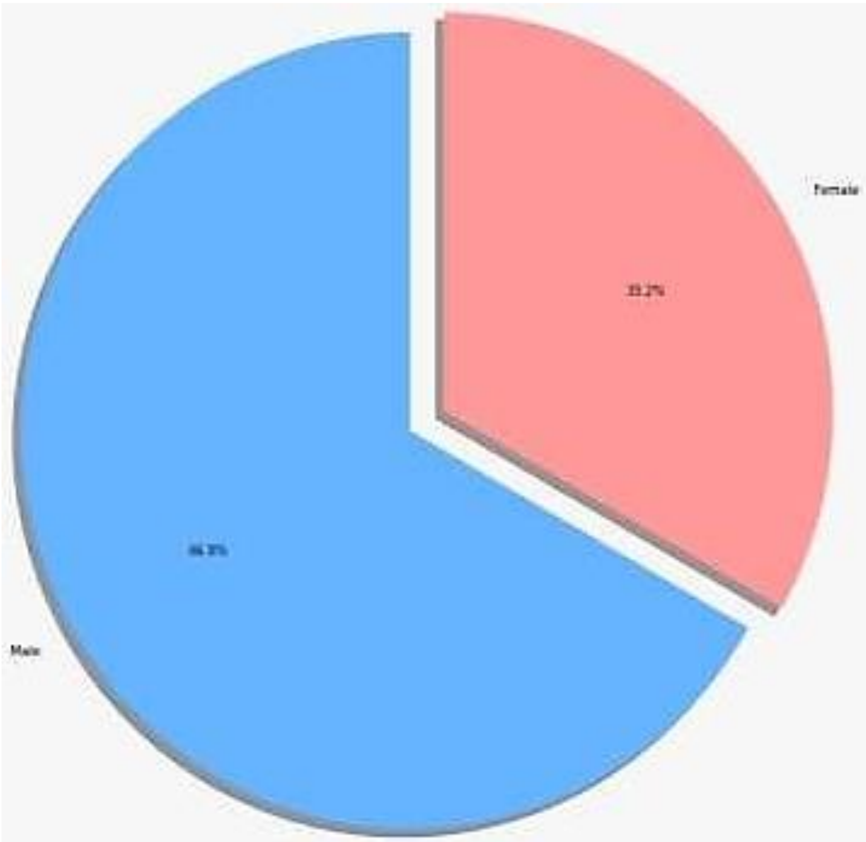

Fig. no. 7.2 Percentage of Gender

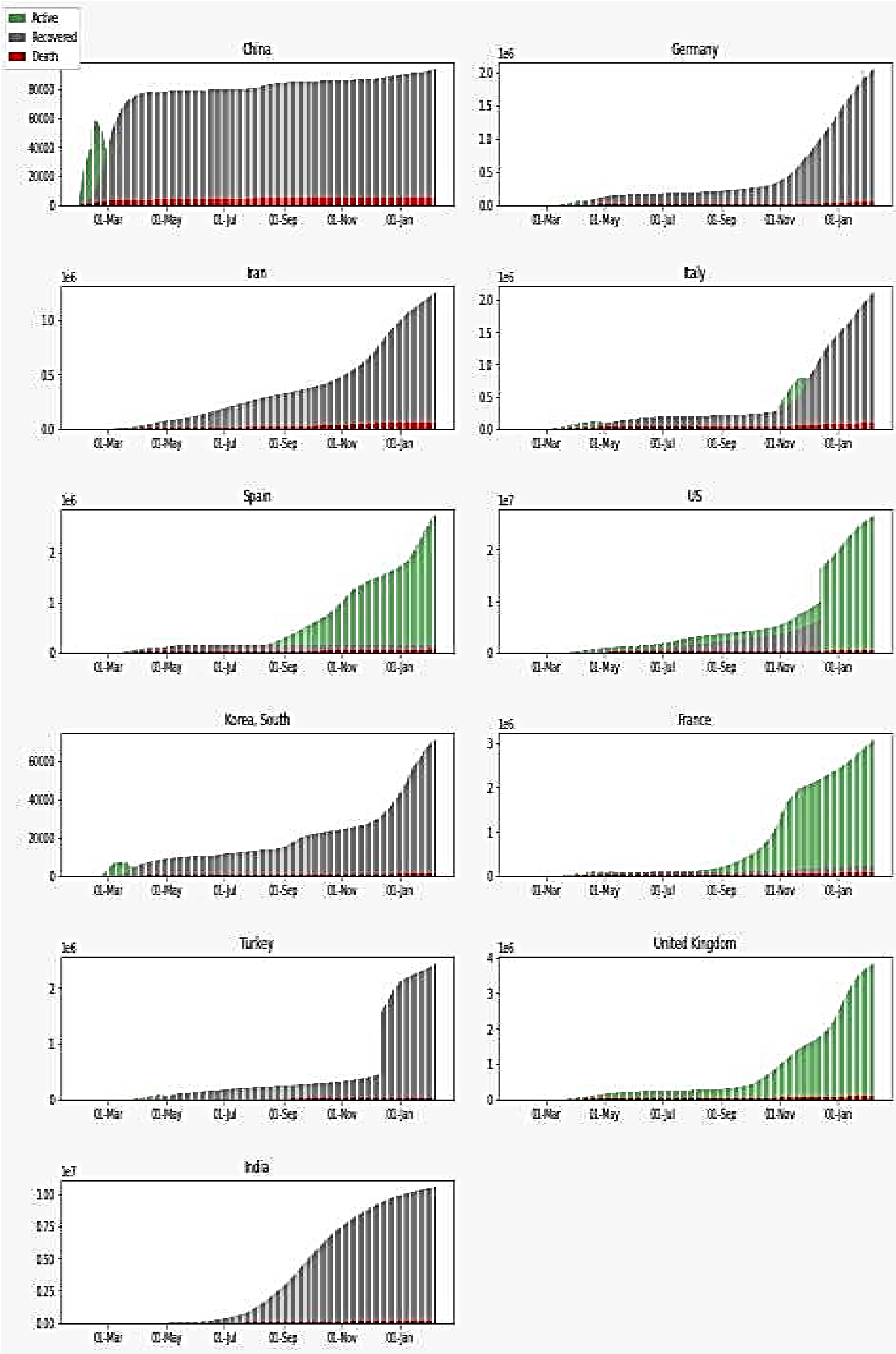

Fig. no. 7.3 Active, Recovered, Deaths in Hotspot Countries and India as of May 15

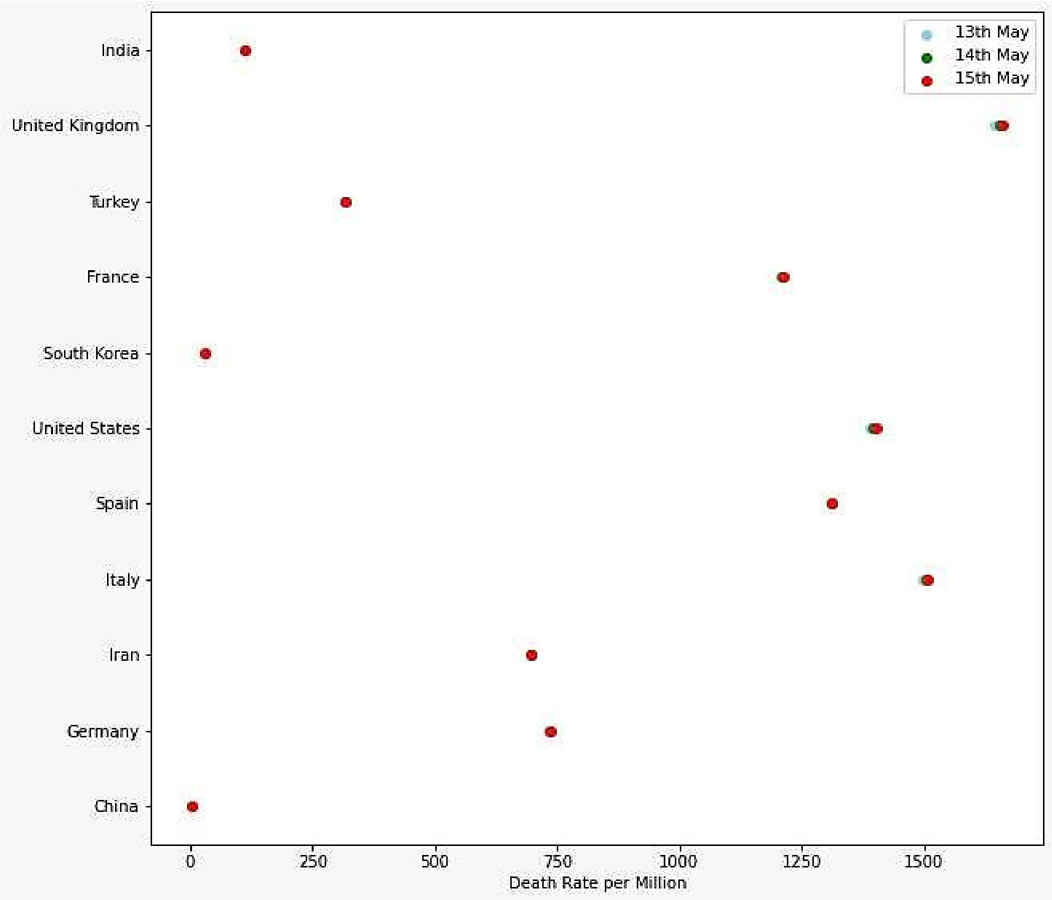

Fig. no. 7.4 Death Rate per Million in Hotspot Countries

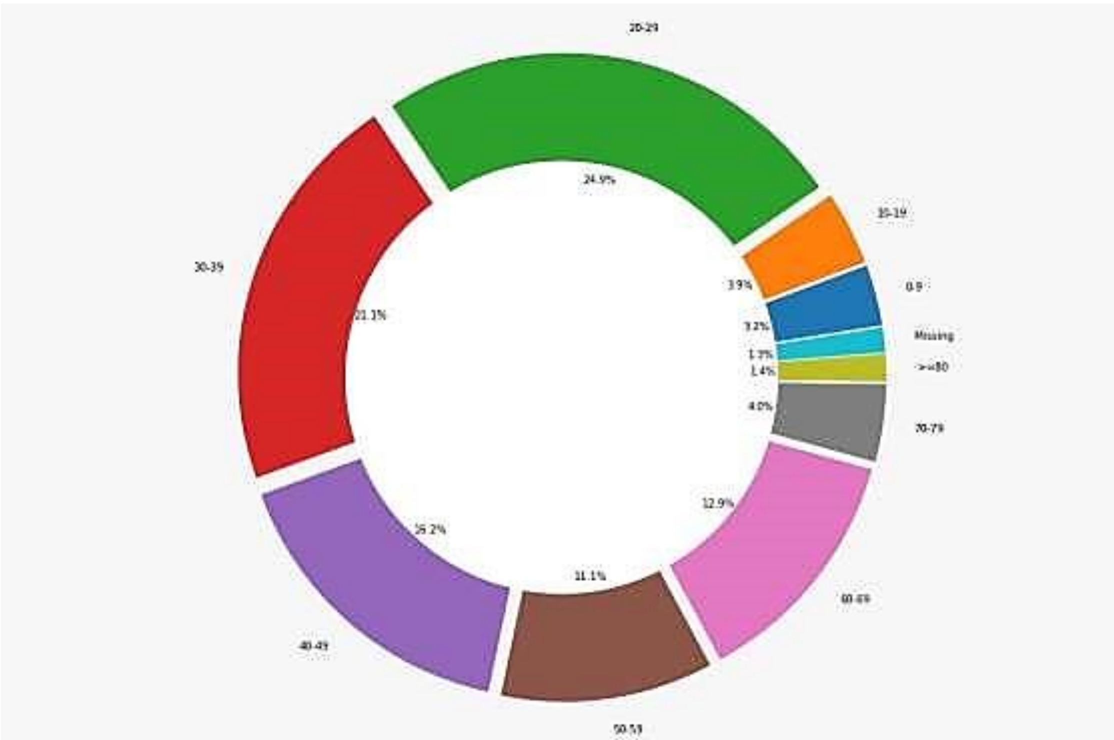

Fig. no. 7.5 India – Age Group wise Distribution

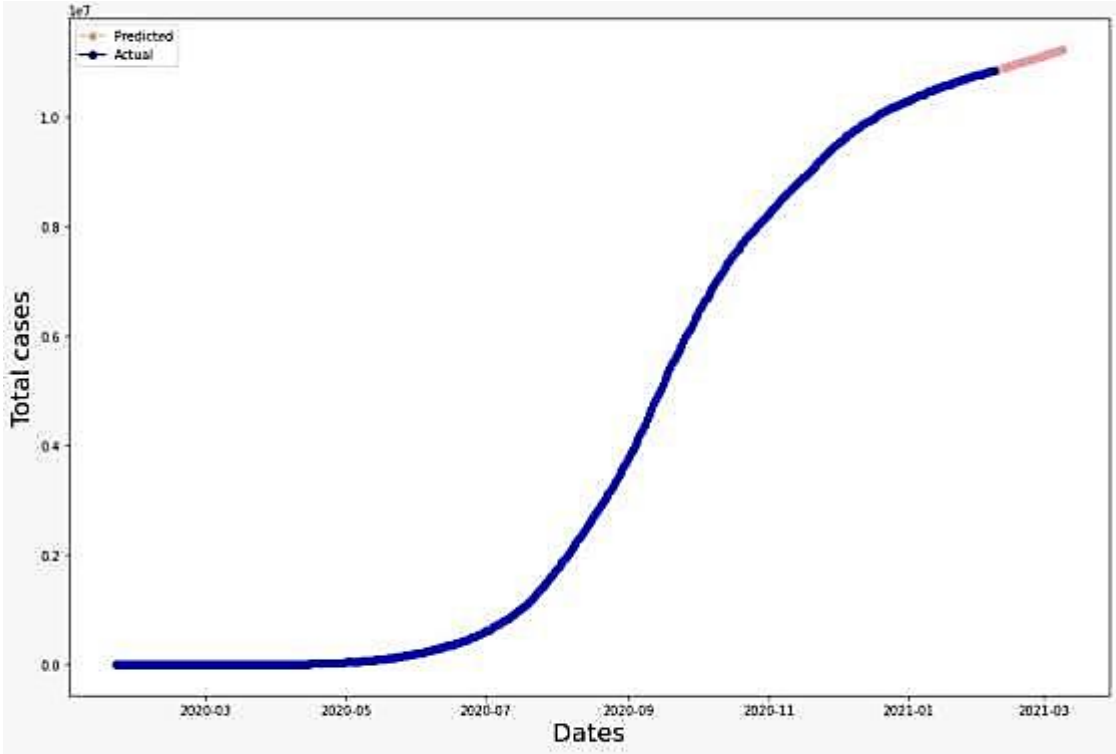

Fig. no. 7.6 Predicted Values for the next 15 Days

SYSTEM TESTING

8.1 TESTING

Testing is a process of executing a program with the aim of finding error. To make our software perform well it should be error free. If testing is done successfully it will remove all the errors from the software.

8.2 TESTING STRATEGIES

1. White Box Testing
2. Black Box Testing
3. Unit testing
4. Integration Testing
5. Alpha Testing
6. Beta Testing
7. Performance Testing and so on

White Box Testing

Testing technique based on knowledge of the internal logic of an application's code and includes tests like coverage of code statements, branches, paths, conditions. It is performed by software developers

Black Box Testing

A method of software testing that verifies the functionality of an application without having specific knowledge of the application's code/internal structure. Tests are based on requirements and functionality.

Unit Testing

Software verification and validation method in which a programmer tests if individual units of source code are fit for use. It is usually conducted by the development team.

Integration Testing

The phase in software testing in which individual software modules are combined and tested as a group. It is usually conducted by testing teams.

Alpha Testing

Type of testing a software product or system conducted at the developer's site. Usually, it is performed by the end users.

Beta Testing

Final testing before releasing application for commercial purpose. It is typically done by end- users or others.

Performance Testing

Functional testing conducted to evaluate the compliance of a system or component with specified performance requirements. It is usually conducted by the performance engineer.

Black Box Testing

Blackbox testing is testing the functionality of an application without knowing the details of its implementation including internal program structure, data structures etc. Test cases for black box testing are created based on the requirement specifications. Therefore, it is also called as specification-based testing. Fig.4.1 represents the black box testing:

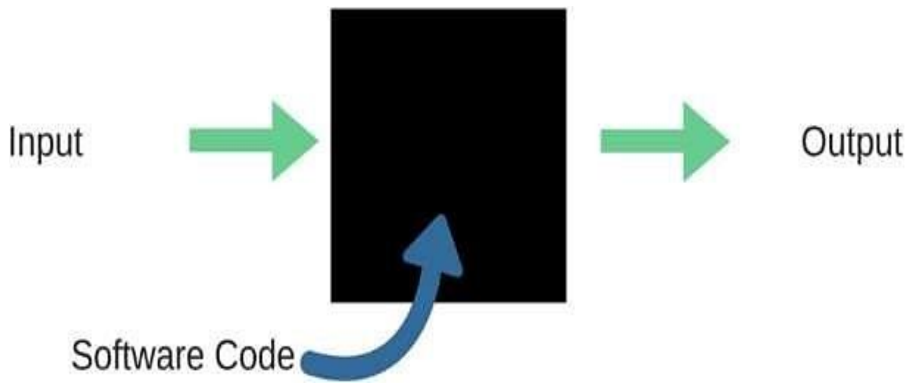

Fig.no. 8.1 Black Box Testing

When applied to machine learning models, black box testing would mean testing machine learning models without knowing the internal details such as features of the machine learningmodel, the algorithm used to create the model etc. The challenge, however, is to verify the test outcome against the expected values that are known beforehand.

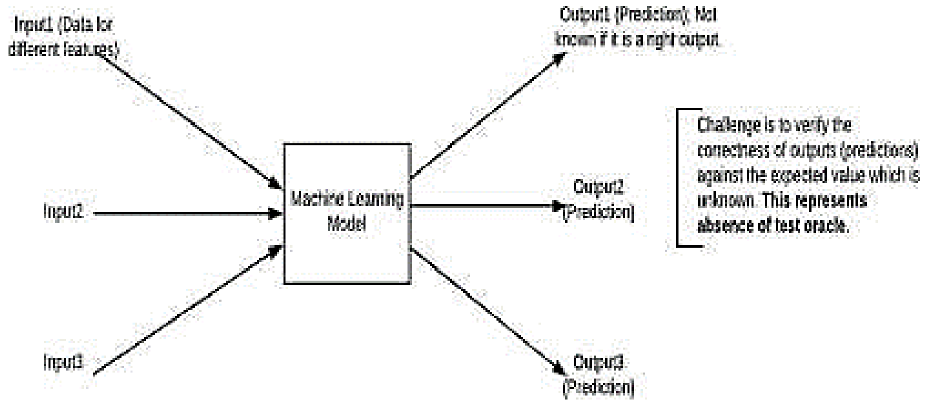

Fig.no. 8.2 Black Box Testing for Machine Learning algorithms

The above Fig.8.2 represents the black box testing procedure for machine learning algorithms.

8.3 TEST CASES

Table.7.2: Test Case Table

Test Case Id	Test Case Name	Test Case Description	Test Steps			Test Case Status	Test Priority
			Step	Expected	Actual		
01	Start the Application	Host the Application and test if it Starts making sure the required software is available	If it doesn't Start	We cannot run the applicatio n.	The application hosts success.	High	High
02	Home Page	Check the deployment environment for Properly loading the application.	If it doesn’t load.	We cannot access the applicatio n.	The application is running successfully.	High	High
03	User Mode	Verify the working of the application in freestyle mode	If it doesn’t Respond	We cannot use the Freestyle mode.	The application displays the Freestyle Page	High	High

04	Data Input	Verify if the application takes input and updates	If it fails to take the input or store in the Database	We cannot proceed further	The application updates the input to application	High	High

CONCLUSIONS AND FUTURE SCOPE

9.1 CONCLUSION

In this work, we present an analysis of the countries that are spreaders of COVID-19, based on the main socio-cultural, economic, and connection characteristics, such as GDP, life expectancy, number of air travel, and budget for health, among others. The results show that the methodology, can cause the rupture of the 5-layer multiplex network and help identify the spreader countries and obtains a classification of the countries based on their characteristics, where, in the spreaders set, the countries have high, medium or low values in the different sociocultural and economic aspects; however, the characteristic that everyone shares are the high value in air connections. On the other hand, we can affirm that to mitigate a second outbreak of COVID-19 in the world, the countries that are in the union of both separator sets must reinforce their sanitary measures; in contrast, the countries that are at the intersection of the two separate sets, in addition to improving their sanitary measures, must regulate airflow to contain the spread of the disease.

9.2 FUTURE SCOPE

Based on the information collected and modeled (until May 15, 2020), we can affirm that, by changing the relationships of the air flow, the risk of a second outbreak of COVID 19 can be minimized; however, we cannot quantify how much it can help.

REFERENCES

[1] WHO., “Timeline covid-19.URL:https://www.who.in t/news-room/detail/27-042020-who-timeline---COVID-19 -19,” 2020.

[2] C. Sohrabs et al., “World health organization declares global emergency: A review of the 2019 novel coronavirus (covid-19),” International Journal of Surgery, 2020. [3] W. H. Organization et al., “Coronavirus disease 2019 (covid-19): situation report,
72,” 2020.

[4] WHO., “Covid-19 dashboard.URL:https://www.who. int/,” 2020.

[5] M. Douglas et al., “Mitigating the wider health effects of covid-19 pandemic response,” Bmj, vol. 369, 2020.

[6] M. Gilbert et al., “Preparedness and vulnerability of African countries against importations of covid-19: a modelling study,” The Lancet, vol. 395, no. 10227, pp. 871– 877, 2020.

[7] A. R. Tuite et al., “Estimation of covid-19 outbreak size in Italy,” The Lancet. Infectious Diseases, vol. 20, no. 5, p. 537, 2020.

[8] R. Müller and D. Wagner, “α-vertex separator is np hard even for 3-regular graphs,” Computing, vol. 46, no. 4, pp. 343–353, 1991.

[9] E. Union., “Eu global response to covid-19. https://ec.e uropa.eu/internationalpartnerships/topics/eu-global-r esponse-covid-19_en,” 2020.

[10] W. Bank., “World bank group and covid-19 (coronavirus). https://www.worldbank.org/en/who-we-are/ news/coronavirus-covid19,” 2020.

[11] I. M. Foundation., “International monetary foundation data. https://www.imf.org/en/Data,” 2020.

[12] T.International.,“Transparency international (countries). https:// www. transparency.org/en/countries/,” 2020.

[13] G. D’Agostino and A. Scala, Networks of networks: the last frontier of complexity. Springer, 2014, vol. 340.

[14] W. Wang et al., “Epidemic spreading on complex networks with general degree and weight distributions,” Physical Review E, vol. 90, no. 4, p. 042803, 2014. [15] M. AL Zaabi et al., “Cursi: a crime investigation system using the relative

importance of information spreaders in networks depicting criminals' communications," IEEE Transactions on Information Forensics and Security, vol. 10, no. 10, pp. 2196–2211, 2015.

[16] C. C. Sarli and C. R. Carpenter, "Measuring academic productivity and changing definitions of scientific impact," Missouri medicine, vol. 111, no. 5, p. 399, 2014. [17] L.-l. Ma et al., "Identifying influential spreaders in complex networks based on gravity formula," Physical A: Statistical Mechanics and its Applications, vol. 451, pp. 205–212, 2016.

[18] S.-Y. Tan et al., "Efficient network disintegration under incomplete information: the comic effect of link prediction," Scientific reports, vol. 6, p. 22916, 2016.

[19] A. Wahid-Ul-Ashraf et al., "Newton's gravitational law for link prediction in social networks," in International Workshop on Complex Networks and their Applications. Springer, 2017, pp. 93–104.

[20] Y. Liu et al., "Core-like groups result in invalidation of identifying super-spreader by k-shell decomposition," Scientific reports, vol. 5, p. 9602, 2015.

[21] L. Lüet al., "Vital nodes identification in complex networks," Physics Reports, vol. 650, pp. 1–63, 2016.

[22] L. C. Freeman, "Centrality in social networks conceptual clarification," Social networks, vol. 1, no. 3, pp. 215–239, 1978.

[23] P. Bonacich and P. Lloyd, "Eigenvector-like measures of centrality for asymmetric relations," Social networks, vol. 23, no. 3, pp. 191–201, 2001.

[24] E. Cozzo et al., "Clustering coefficients in multiplex networks," arXiv preprint arXiv:1307.6780, 2013.

[25] P. J. Mucha et al., "Community structure in timedependent, multiscale, and multiplex networks," science, vol. 328, no. 5980, pp. 876–878, 2010.