IMPLEMENTATION OF A REAL-TIME, DATA-DRIVEN ONLINE EPIDEMIC CALCULATOR FOR TRACKING THE SPREAD OF COVID-19

A PROJECT REPORT

Submitted by,

Mr. MOHAMMED ABU HANEEF - 20201CAI0012
Mr. KISHORE B - 20201CAI0055
Mr. SOUHARDH D.S - 20201CAI0063
Mr. SHAIK MOHD GAUSE - 20201CAI0067

Under the guidance of,

Mr. Likhith S.R

in partial fulfillment for the award of the degree of

BACHELOR OF TECHNOLOGY

IN

COMPUTER SCIENCE AND ENGINEERING IN (ARTIFICIAL INTELLIGENCE & MACHINE LEARNING)

At

GAIN MORE KNOWLEDGE
REACH GREATER HEIGHTS

PRESIDENCY UNIVERSITY

BENGALURU

JANUARY 2024

PRESIDENCY UNIVERSITY

SCHOOL OF COMPUTER SCIENCE AND ENGINEERING

CERTIFICATE

This is to certify that the Project report " **IMPLEMENTATION OF A REAL TIME, DATA-DRIVEN ONLINE EPIDEMIC CALCULATOR FOR TRACKING THE SPREAD OF COVID-19** " being submitted by " MOHAMMED ABU HANEEF, KISHORE B, SOUHARDH D S, SHAIKH MOHD GAUSE" bearing roll number(s) " 20201CAI0012, 20201CAI0055,20201CAI0063, 20201CAI0067 " in partial fulfilment of requirement for the award of degree Of BACHELOR OF TECHNOLOGY IN COMPUTER SCIENCE AND ENGINEERING (AI & ML) is a bonafide work carried out under my supervision.

Mr. Likhith S.R
Assistant Professor
School of CSE&IS
Presidency University

Dr. Zafar Ali Khan
Associate Professor & HoD
(ECM, IST, CAI)
School of CSE&IS
Presidency University

Dr. C. KALAIARASAN
Associate Dean
School of CSE&IS
Presidency University

Dr. L. SHAKKEERA
Associate Dean
School of CSE&IS
Presidency University

Dr. SAMEERUDDIN KHAN
Dean
School of CSE&IS
Presidency University

PRESIDENCY UNIVERSITY

SCHOOL OF COMPUTER SCIENCE AND ENGINEERING

DECLARATION

We hereby declare that the work, which is being presented in the project report entitled **IMPLEMENTATION OF A REAL-TIME, DATA-DRIVEN ONLINE EPIDEMIC CALCULATOR FOR TRACKING THE SPREAD OF COVID-19** in partial fulfilment for the award of Degree of **BACHELOR OF TECHNOLOGY IN COMPUTER SCIENCE AND ENGINEERING (AI & ML)**, is a record of our own investigations carried under the guidance of **MR. LIKHIT S.R, Assistant Professor**, **School of Computer Science and Engineering & Information Science, Presidency University, Bengaluru.**

We have not submitted the matter presented in this report anywhere for the award of any other Degree.

Names	Roll Numbers	Signature
Mr. MOHAMMED ABU HANEEF	**- 20201CAI0012**	
Mr. KISHORE B	**- 20201CAI0055**	
Mr. SOUHARDH D. S	**- 20201CAI0063**	
Mr. SHAIK MOHD GAUSE	**- 20201CAI0067**	

ABSTRACT

This research endeavors to address the imperative need for advanced epidemic detection and monitoring strategies in the wake of the COVID-19 pandemic. The study introduces an innovative methodology harnessing machine learning (ML) algorithms for the prediction and real-time mapping of the virus's spread. A robust ML model is developed by training it on a comprehensive dataset that encapsulates geographical, temporal, and epidemiological information. The model's capabilities extend to forecasting potential outbreaks, identifying high-risk regions, and offering a nuanced understanding of epidemic dynamics through the integration of feature-rich data such as population density, mobility patterns, and healthcare infrastructure. The culmination of these efforts results in a sophisticated real-time mapping system that dynamically updates visualizations as new data becomes available. This adaptive framework contributes significantly to ongoing epidemic control initiatives, providing decision-makers with timely, data-driven insights crucial for effective mitigation and resource allocation.

Abstract:

The global COVID-19 pandemic has underscored the critical need for effective methods to track and monitor the spread of the virus. This research addresses the limitations of current epidemic tracking systems and proposes the development and implementation of a real-time, data-driven online epidemic calculator. The significance of this research lies in its potential to enhance decision-making by public health authorities, improve the effectiveness of interventions, and contribute to the development of future epidemic tracking technologies. The scope of the study focuses on the development and implementation of the proposed system, with acknowledged limitations regarding broader healthcare infrastructure and societal factors. Research questions guide the investigation into the identified gaps and drive the exploration of innovative solutions. The thesis is structured to provide a comprehensive review of existing literature, detail the research methodology, present the proposed system's design, discuss the implementation, and analyze findings. This research contributes to the ongoing efforts to combat the COVID-19 pandemic by providing a novel approach to real-time epidemic tracking.

ACKNOWLEDGEMENT

First of all, we indebted to the **GOD ALMIGHTY** for giving me an opportunity to excel in our efforts to complete this project on time.

We express our sincere thanks to our respected dean **Dr. Md. Sameeruddin Khan**, Dean, School of Computer Science and Engineering & Information Science, Presidency University for getting us permission to undergo the project.

We record our heartfelt gratitude to our beloved Associate Deans **Dr. Kalaiarasan C and Dr. Shakkeera L,** School of Computer Science and Engineering & Information Science, Presidency University and **Dr. Zafar Ali Khan,** Head of the Department, School of Computer Science and Engineering & Information Science, Presidency University for rendering timely help for the successful completion of this project.

We are greatly indebted to our guide **Mr. Likhith S.R**, School of Computer Science and Engineering & Information Science, Presidency University for his inspirational guidance, and valuable suggestions and for providing us a chance to express our technical capabilities in every respect for the completion of the project work.

We would like to convey our gratitude and heartfelt thanks to the University Project-II Coordinators **Dr. Sanjeev P Kaulgud, Dr. Mrutyunjaya MS** and also the department Project Coordinators **Dr. Murali Parameswaran**.

We thank our family and friends for the strong support and inspiration they have provided us in bringing out this project.

KISHORE B
MOHAMMED ABHU HANEEF
SOUHARDH D.S
SHAIK MOHD GAUSE

LIST OF TABLES

LIST OF FIGURES

TABLE OF CONTENTS

CHAPTER-1

INTRODUCTION

1.1 Background

Addressing the challenges posed by the global COVID-19 pandemic has necessitated the development and implementation of innovative tools to monitor and track the spread of the virus. Traditional methods of epidemic tracking often rely on retrospective data collection, manual analysis, and delayed reporting, which can lead to inaccuracies, hindering timely decision-making and response efforts. Monitoring the spread of COVID-19 in real-time is crucial for implementing targeted interventions, allocating resources effectively, and making informed public health decisions. Traditional methods of epidemic tracking, such as periodic reports and manual data compilation, often result in delays and may not capture the rapidly changing dynamics of the virus's transmission.

To address these issues, a real-time epidemic calculator utilizes advanced data analytics, artificial intelligence, and machine learning algorithms to process and analyze data from various sources, including healthcare facilities, testing centers, and public health databases. By leveraging real-time data, this online calculator provides instant insights into the current status of the pandemic, including infection rates, geographical hotspots, and trends over time.

The implementation of such a system requires collaboration between public health authorities, data scientists, and technology experts. Integration with existing public health infrastructure and data systems is essential to ensure seamless data flow and interoperability. Privacy and data security measures must also be a top priority to protect sensitive health information.

In a nutshell, the implementation of a real-time, data-driven online epidemic calculator for tracking the spread of COVID-19 addresses the limitations of traditional tracking methods. By providing accurate, timely information, this tool becomes a valuable asset in the global effort to manage and mitigate the impact of the pandemic. The collaboration between public health entities and technology experts is crucial for the successful deployment and ongoing refinement of such a system.

1.2 Research Motivation

In the wake of the unprecedented challenges posed by the COVID-19 pandemic, there has been a heightened recognition of the critical necessity for advanced techniques in epidemic detection and monitoring. This research endeavors to contribute to the evolving landscape of public health strategies by introducing a novel approach rooted in machine learning (ML) algorithms. Focusing on the development of predictive models, our study harnesses a comprehensive dataset that encompasses geographical, temporal, and epidemiological information. Through rigorous training, these ML models demonstrate the capability to forecast the spread of the virus, discern high-risk regions, and anticipate potential outbreaks. This endeavor represents a concerted effort to fortify our preparedness and responsiveness to emerging infectious threats through the application of cutting-edge technology.

Central to the efficacy of our approach is the integration of diverse and feature-rich data sets. By incorporating factors such as population density, mobility patterns, and healthcare infrastructure, our ML models yield a holistic understanding of epidemic dynamics. The synthesis of these variables not only enhances the accuracy of predictions but also enables a more nuanced interpretation of the multifaceted factors influencing the trajectory of infectious diseases. This research, therefore, aspires to transcend the limitations of traditional epidemic surveillance methods by providing decision-makers with a more comprehensive and adaptive framework for real-time monitoring and mitigation strategies.

The culmination of our efforts manifests in the development of a sophisticated real-time mapping system. This system, built upon the foundations of our trained ML models, dynamically updates visualizations as new data becomes available, offering decision-makers a responsive tool for informed decision-making. In contributing to the ongoing discourse in epidemic control, this research aims to empower public health officials and policymakers with timely, data-driven insights. Through the convergence of advanced machine learning algorithms and comprehensive data integration, our approach seeks to establish a paradigm shift in epidemic prediction, response, and resource allocation for the benefit of global public health.

1.3 Primary Objectives

Timely Decision-Making: Traditional methods of epidemic tracking often involve delayed data collection and reporting. The motivation for this research is to provide decision-makers with real-time data to facilitate swift and informed responses to the evolving dynamics of the COVID-19 pandemic. Timely decision-making is crucial for implementing effective public health measures and interventions.

Accuracy and Precision: The dynamic nature of the COVID-19 pandemic demands a high level of accuracy and precision in tracking the spread of the virus. Real-time data-driven approaches aim to reduce errors associated with manual data compilation, ensuring that decision-makers have access to the most reliable and up-to-date information for strategic planning and resource allocation.

Identification of Hotspots: Rapid identification of geographical hotspots is essential for targeted interventions and resource allocation. An online epidemic calculator with real-time capabilities allows for the prompt detection of emerging clusters of infections, enabling public health authorities to focus efforts where they are most needed.

Resource Optimization: Efficient resource allocation is a critical aspect of pandemic management. By understanding the real-time dynamics of the virus's spread, public health officials can optimize the distribution of medical supplies, healthcare personnel, and other resources to areas facing the highest risk or experiencing surges in cases.

Public Awareness and Communication: Providing the public with accurate and up-to-date information is crucial for fostering awareness and encouraging adherence to public health guidelines. An online epidemic calculator serves as a transparent and accessible platform for sharing information about infection rates, trends, and preventive measures, contributing to public understanding and cooperation.

Scientific Research and Modeling: Researchers and epidemiologists require real-time data to improve models predicting the course of the pandemic. Access to accurate and current information aids in refining models, understanding transmission dynamics, and projecting future scenarios, thereby contributing to the scientific understanding of COVID-19.

1.4 Contributions

The initial phase of any project is crucial because it sets the tone for the project as a whole. We tirelessly followed the lofty IEEE rules as our compass with regards to our scholastic undertaking. Our most vital endeavor incorporated a wide examination of a lot of papers collaborated with the IEEE, hoping to assemble a critical corpus of significant composition. In order to accomplish this, we carefully evaluated the relevance of each collected paper to our project's requirements and carried out in-depth web searches with specific keywords.

During this stage, our fundamental community was to recognize the refs liable for investigating the papers. By identifying these important people, we wanted to learn more about the authors' credibility and expertise so that the collected literature could be trusted. In addition, we conducted a comprehensive writing review, carefully examining a select number of papers relevant to our task. Thanks to this survey, which served as a crucial foundation for our research, we were able to gain valuable insights and a comprehensive understanding of the existing body of work in our chosen field.

Drawing from the literature review and aligning it with our project's requirements, we meticulously compiled all of the essential information, specifications, and prerequisites for its successful completion. This crucial stage provided us with a solid and well-informed foundation from which to work and laid the groundwork for the subsequent phases of our project. We put ourselves in a position for a more productive and smoothed out way toward accomplishing our scholarly objectives by carefully establishing this groundwork.

1.5 Organization of the report

Chapter 1: Introduction: This chapter provides an overview of the project and explains the need for it. It also identifies the problems in the current system and explains the significance and relevance of the project to solve the identified problems.

Chapter 2: Literature Survey: This chapter provides a summary of the existing research on the topic. It explains the different IEEE papers that were reviewed during the implementation of the project and their relevance to the project.

Chapter 3: Research Gap of Existing Methods: Chapter 2: This chapter outlines the approach taken by the research to bridge the identified gaps and contribute to the existing body of knowledge.

Chapter 4: Proposed Methodology: This chapter contains information about the proposed system, and its merits, and the overall system architecture. It provides a detailed description of the project design and how it solves the identified problems.

Chapter 5: System Design and Implementation: This chapter contains the algorithms used in the project and a brief explanation of the different modules that make up the project

Chapter 6: Outcomes: This chapter contains the results of the project.

Chapter 7: Results and Discussions: This chapter contains the results of the tests carried out and an analysis of the results.

Chapter 8: Conclusion: This chapter summarizes the project and its findings.

References: This section contains a list of all the references used in the project report.

Appendix A: This section contains pseudocode related to the project.

Appendix B: This section contains screenshots related to the project.

Appendix C: This section contains a plagiarism report to confirm that the project is original work.

CHAPTER-2

LITERATURE SURVEY

2.1 Introduction

The literature survey presented here encompasses a diverse array of studies and methodologies employed in understanding, predicting, and combating the COVID-19 pandemic. With the global impact of the virus, researchers and practitioners have leveraged various technologies, including artificial intelligence, machine learning, computer vision, and the Internet of Things (IoT), to develop innovative solutions for monitoring, prediction, and containment. This comprehensive literature survey explores studies ranging from real-time neural network algorithms for predicting virus spread to sophisticated computer vision-based systems for contact tracing and social distancing monitoring. It also delves into the use of social media data for monitoring the pandemic, phylogenomic tracking, epidemic simulation frameworks, and the role of IoT in pandemic tracking. Additionally, the survey covers studies that focus on healthcare data acquisition, geospatial monitoring, and the development of interactive dashboards to enhance public understanding of the virus's impact. The collection of studies reflects the multidisciplinary efforts employed worldwide to address the challenges posed by the COVID-19 pandemic, providing insights into the evolution of research and technology in the ongoing battle against infectious diseases.

2.2 Literature Review

A COVID-19 Prediction Optimization Algorithm Based on Real-time Neural Network Training

[1] The COVID-19 pandemic has had an unprecedented impact on the world, and accurate prediction of the virus's spread is crucial for effective management of the situation. To this end, a COVID-19 Prediction Optimization Algorithm has been developed, based on a BP neural network that is trained in real-time. This algorithm was used to predict the number of infections and deaths in Italy during the pandemic, with epidemic data collected from February 23, 2020, to March 23, 2020, and preprocessed using R software. The BP neural network used in the algorithm is a type of artificial neural network that is widely used in prediction and classification tasks. It establishes a nonlinear mapping relationship from the input layer to the output layer to solve the complex problem of internal mechanism. The epidemic data in Italy

was preprocessed using R software, and the short-term Italian epidemic data was found to be a non-stationary time series.

The algorithm's predictions were compared to the actual situation in Italy, and the results showed that the model was accurate in predicting new confirmed cases and new deaths. The dataset used in this study was the epidemic data in Italy from February 23, 2020, to March 23, 2020. The authors of the study suggest that the algorithm can be used as a reference for predicting the spread of COVID-19 in other countries. However, they also note that the spread of the virus is affected by many factors, and a simple closed system cannot be used for full prediction.

Analysis and Predictions of Spread, Recovery, and Death Caused by COVID-19 in India

[2] The study focuses on the confirmed, recovered, and death cases caused by COVID-19 and the impact of lockdown and social distancing measures on the spread rate of the virus. It uses various forecasting models such as logistic regression, machine learning, deep learning, ARIMA, and SARIMA models to predict the number of COVID-19 cases in India. This model have a strong predictive capacity and can accurately predict the number of cases. It uses decision tree learning techniques to split the dataset based on conditions, which is useful for both regression and classification. The dataset used in the study is mainly week-wise confirmed cases in concerning states. The data is refined and analyzed to identify the key factors that contribute to the spread rate of the virus. The lockdown and social distancing measures are crucial factors in suppressing the spread rate of the virus. The study suggest that the implementation of lockdown measures has a significant impact on reducing the number of cases. It also identifies other factors such as meteorological conditions that contribute to the spread rate of the virus.

Prediction of COVID-19 Confirmed, Death, and Cured Cases in India Using Random Forest Model

[3] The research study that focuses on using the Random Forest Model to predict the number of COVID-19 cases in India. The authors thoroughly examined the confirmed, death, and

cured cases of COVID-19 in India and used different models to forecast the cases. The dataset used in this study includes information about the observation date, time, state/union territory, confirmed Indian national, confirmed foreign national, as well as the number of deaths and cured cases. The study proposes machine learning techniques that rely on a data-driven approach to anticipate the number of COVID-19 infections in the coming days, using the available data. Among the various models tested, the Random Forest Model showed the best performance and was therefore used for prediction and analysis. This model effectively predicts the number of new COVID-19 cases, allowing authorities to prepare accordingly.In addition to that, the study provides a detailed methodology and materials section, explaining the process of feature selection and the different machine learning models used in the research. The results of the study are presented in the result analysis section, which includes a comparison between reported and estimated cases.

Real-time Contact Tracing During a Pandemic using Multi-camera Video Object Tracking

[4] A computer vision-based algorithm for real-time contact tracing and moving object tracking to ensure social distancing during a pandemic. The proposed system consists of multiple pipelines combined to perform automatic detection and tracking of moving objects in a video from stationary cameras converted to a bird's eye view. The algorithm begins by applying a background subtraction algorithm based on Gaussian mixture models to detect the moving object, morphological operators to eliminate noise, and blob analysis to detect the connected pixels that will mostly form objects. Once objects are detected, the motion track of each object is estimated by a Kalman filter. Eventually, it calculates Euclidean distance between the objects to trace object contacts. The system was tested on a dataset of public places and was able to detect the objects in the input video frame and estimate the distance between them across multiple cameras. The results obtained verify the feasibility and effectiveness of the proposed algorithm. The dataset used in this study is not explicit.

Social distancing detection and analysis through computer vision

[5] A computer vision-based social distancing detector can aid in combating the global COVID-19 pandemic. This system is designed to provide accurate results regardless of camera

viewpoint or distortion, achieving a balanced mean average precision (mAP) and fps score suitable for real-time environments. The author has discussed about the recent studies and proposed solutions for social distancing monitoring, as well as an exhaustive study of the latest state-of-the-art (SOTA) object detection models and a detailed methodology for social distancing detection. The research includes experiments and corresponding results, with conclusions drawn from the findings. To fine-tune the models, the INRIA Person Dataset was utilized, which includes train, test, and annotation folders. Two additional datasets were used for testing purposes: the Oxford Town Centre (OTC) surveillance video of pedestrians and PETS 2009, which includes various crowd activities. The OTC dataset consists of a single 5-minute video capturing pedestrians walking on the road, while the PETS dataset provides video frames for different purposes such as density estimation of the crowd, people tracking, and flow analysis. The video frames from the people tracking folder were utilized for testing, depicting the same scenario from 8 different views, with a length of 2 minutes and an original video frame rate of 8fps.

An Interactive Platform to Track Global COVID-19 Epidemic

[6] An intelligent stage created by scientists to follow the worldwide Coronavirus scourge. The stage gives clients a complete information base of Covid cases around the world, an elevated perspective of the worldwide Coronavirus elements, and individuals' top worries through tweets streams. The stage has two principal works: a heatmap of the affirmed cases and patterns of the affirmed cases. The heatmap shows the quantity of nations impacted, absolute number of cases and passings, and the aggregated case number levels. The patterns capability gives an itemized patterns depiction to every nation/state around four classifications time series: complete affirmed cases/day to day affirmed cases, all out passings/day to day passings, and all out tests/day to day tests.

The intelligent stage gives an exhaustive and intuitive way for scientists, strategy producers, and the overall population to follow the worldwide Coronavirus pandemic and remain informed about the most recent patterns. The stage additionally safeguards verifiable records and empowers clients to see the Coronavirus spreading patterns. The stage gathers and processes tweets streams to give clients a one-stop insight of the outline Coronavirus circumstance. The stage is an important instrument for checking and investigating future flare-ups.

Monitoring of Epidemic Outbreaks Using Social Media Data

[7] The study is about using virtual entertainment news, specifically Twitter, to monitor the coronavirus pandemic, with a particular focus on India. Using the Twitter programming interface, the author compiled a large dataset of approximately 130,000 explicit English-language coronavirus tweets collected.To group them, the authors used a number of calculations such as backing vector machines (SVMs), Geilless Bayes, and Pearson relations. These ordinal calculations played a key role in classifying and deciphering vast amounts of online entertainment information. Despite the potential that Web-based entertainment information can provide, developers have recognized the difficulties associated with its use. Recognized excitement over and difficulty in understanding information raised concerns about the representativeness of online entertainment information in capturing the opinions and encounters of a wider range of people. Despite the audit, developers have introduced improvements to the Live Pest Control framework. The framework featured a reader-based interface that could continuously track tweets related to the epidemic outbreak in India. Implementation of this framework includes a combination of Use Programming Points of Interaction (API) and geocoding innovations, as well as updating the user interface with location-specific live data.

This represents an extensive study on the use of online entertainment information to monitor the coronavirus pandemic in India, with the developers using a large dataset of tweets collected through the Twitter programming interface. and applied many innovations including NLP , the study of emotions and the calculation of order. Although the difficulties associated with online entertainment information were recognized, the authors highlighted its potential usefulness for continuously monitoring and disseminating important information to health authorities. Improvements in the live pest checking framework have also demonstrated the practical use of these advances in a dynamic and evolving general health scene.

Machine Learning Model for Computational Tracking and Forecasting the COVID-19 Dynamic Propagation

[8] The study presents a computational model that integrates a stretch-sorted 2 flying group and a spanned type 2 flash Kalman channel for tracking and estimating the spread of the

coronavirus in Brazil. The model uses intelligent AI techniques to analyze epidemiological information and create predictions for a daily stream of reports. The effectiveness of this model is demonstrated by exploratory results demonstrating its ability to accurately track and calculate the unique emergence of coronaviruses. Furthermore, a detailed study using different methods and AI models is presented to highlight the advantages of the proposed philosophy.

This proposes is a unique computational model that coordinates the Type 2 Fluffy Kalman Channel and the Type 2 Fluffy Kalman Channel to track and estimate the spread of the coronavirus in Brazil. The model has demonstrated productivity in analyzing epidemiological information and generating accurate predictions for daily reports. A relative investigation using different methods and AI models confirms the feasibility of the proposed approach. The authors acknowledge the support of CAPES and the Master and Doctorate Program in Electrical Engineering at the Federal University of Maranhão in their exploration of using social media data, particularly Twitter, for monitoring the COVID-19 pandemic in India. The study involves analyzing a large-scale dataset with natural language processing, sentiment analysis, and classification algorithms, showcasing the potential of social media in real-time epidemic monitoring.

Real-Time Health Data Acquisition and Geospatial Monitoring: A Visual Analytics Approach

[9] This study portrays a visual examination approach for ongoing wellbeing information procurement and geospatial checking. The creators contend that advanced innovations, like web and versatile based applications can tremendously work on the nature of medical services by giving quick, secure, effective, and cheap ways of gaining, store, and dissect patient information electronically. They center around the instance of Pakistan, where a great many individuals are determined to have lethal illnesses consistently, and where the accessible wellbeing assets are overburdened by the sheer number of patients visiting medical clinics.

To address this test, the writers propose a framework that utilizes spatio-transient examination, GIS guides, and continuous investigation to screen infection spread and go to preventive lengths. The framework depends on a dataset of time-stepped streaming information containing the main protests of the patients, which is first dissected by a syndromic classifier and afterward characterized into various syndromic classifications progressively. The authors use the method developed by Ali et al. The study uses Wellbeing Level 7 (HL7), a set of global standards used by healthcare providers to exchange clinical and regulatory information between programming applications.

EpidemicSim: Epidemic Simulation System with Realistic Mobility

[10] The study discusses the development of a mobile epidemic simulation framework called EpidemicSim. The framework aims to simulate the spread of diseases by modeling the movement of individuals in a defined space. The simulation takes into account parameters such as population size, walking area, and duration to determine the best conditions for simulating disease spread. The effects of different infection probabilities are also investigated. The simulation model used in EpidemicSim is based on the TLW (Time-Limited Walk) model, which simulates the movement of individuals in the simulation area. The model generates paths for each individual at regular time intervals and calculates the distance between individuals at each interval. This distance can be used to determine whether two people are in contact and whether a disease can be transmitted.

Simulation results show that increasing the walking area while keeping the population size the same significantly reduces the infection rate. If the population size is small and the walking area is large, the simulation time should be longer to adequately cover all individuals. However, if longer times are not possible, limiting the walking space available to individuals can have similar consequences. The project also discusses the creation of disease maps for social networks that show the connections between people in the network and the spread of the disease. The map provides information about the importance of each individual in the network and their connection to the original source of the disease and the infected population as a whole. The EpidemicSim framework provides a platform for simulating disease spread in real-world scenarios. Simulation results and social network maps can be used to understand the dynamics of disease transmission and develop strategies for disease prevention and control.

Monitoring and Tracking the Evolution of a Viral Epidemic Through Nonlinear Kalman Filtering: Application to the COVID-19 Case.

[11] A new methodology for the systematic treatment of time series data associated with viral epidemics. The goal of this methodology is to predict the peak of an epidemic and its subsequent evolution. The methodology is based on a nonlinear model that links raw data on cases (positive, recovered, deceased) to the time-variance geometric ratio (TVR) of infected people. A Kalman filter is then used to estimate the state variables involved. This methodology is tested on both hypothetical and real-world data from several countries. The results demonstrate the positive impact of severe lockdowns on the prevention of subsequent virus waves. In conclusion, the methodology can predict the epidemic peak at least two weeks before it occurs and can provide valuable public health policy and interventions in future viral epidemics.

Phylogenomics for Tracking the Epidemiology of COVID-19: The Genomic Data Gap for the African Continent.

[12] COVID-19 has wreaked havoc on health systems around the world, and monitoring the spread of the virus through phylogenomic tracking is of paramount importance. However, genomic data gaps exist across the world, and only a small fraction of global data originates from Africa. The phylogenetic analysis indicates that Egypt hosted the first case reported in Africa. However, discrepancies in sampling timing and sequence placement are present in the analysis. To address this issue, the study proposes accelerating the sequencing of samples for improved data collection.The accuracy of this study suggests that on the continent of Africa, there is a large data gap on several levels, such as geographical coverage, range or density of coverage in relation to infection level, and cross-referencing of sample records. The 106 sequences that fulfilled these criteria were aligned using NCBI’s basic local alignment search tool (BLAST). The aligned sequence matrix was next subjected to a phylogenetic analysis (Phylogenetic Analysis of Genome Data) using the neighbour joining algorithm (NJ).

A Graph-based Methodology for Tracking Covid-19 in Time Series Datasets.

[13] The utilization of machine learning models in examining the dissemination of the novel coronavirus (COVID-19). It enumerates several models, including Truncated Inception net and Convolutional CapsNet, which were employed to identify COVID-19 in chest X-ray images.fractional-order model that was utilized to scrutinize the outbreak of COVID-19 by using confirmed cases data. It also tells about the application of community analytics and graph analytics to identify groups and clusters of individuals with similar characteristics.These models have shown promising results in detecting COVID-19 cases from X-ray images.Several models, such as Truncated Inception net and Convolutional CapsNet, have been applied to analyze chest X-ray images and classify them as COVID-19 or non-COVID-19.One dataset commonly used is chest X-ray images, which are utilized in computer vision applications.algorithm used is the CapsNet, a convolutional neural network used to detect COVID-19 disease in chest X-ray datasets.

A Cough-based deep learning framework for detecting COVID-19.

[14] The detection of COVID-19 positive individuals based on their cough sounds is facilitated by a deep learning framework. This framework consists of two primary steps: front-end feature extraction and back-end classification. In the front-end feature extraction process, a combination of handcrafted features and embedding-based features extracted from pre-trained models for COVID detection is employed. The re-sampled recordings are inputted into this process, where both embedding-based features and handcrafted features are extracted and then concatenated to form the combined features. These combined features are subsequently fed into various back-end classification models to identify COVID-19 positive cases. To evaluate the effectiveness of the system, it was tested on the Track-2 dataset of the Second 2021 DiCOVA Challenge, yielding impressive results. The proposed system is not only deployable and robust, but also competitive, with the potential for further application on edge devices in COVID-19 detection.The proposed system achieved an AUC score of 81.21 and the top F1 score of 53.21 on a Blind Test set, improving the challenge baseline by 8.43% and 23.4% respectively.

Tracking COVID-19 by Tracking Infectious Trajectories

[15] Governments are confronted with numerous obstacles in their efforts to contain the transmission of COVID-19, with asymptomatic individuals posing a significant challenge. To address this issue, a big-data architecture is suggested, which leverages IoT devices to monitor the movement of individuals and trace their infectious trajectories. By capturing the coordinates of individuals during outdoor activities, this proposed system aims to create a comprehensive database of continuously collected trajectories, including person-ids, coordinates, and timestamps. This system holds potential in enhancing state-of-the-art mathematical models for disease spreading and prediction, while also acknowledging its inherent advantages and limitations. Additionally, the text briefly explores existing tracking tools and offers recommendations for their effective utilization. The proposed model is expected to improve the accuracy of disease spreading/prediction models. Three proposed algorithms that exploit the gathered data to find and classify suspected cases, determine black areas (zones with high contamination probability), and find all persons who visited black areas (also considered as probably infected persons).

Global COVID-19 Tracking and Analysis

[16] In the realm of global disease monitoring, an interactive and comprehensive tool has emerged. This platform transcends mere tracking, delving into visualization and analysis of COVID-19 cases worldwide. Aiming to empower researchers, guide public health decisions, and enlighten the general populace, this platform offers a panoramic view of global COVID-19 dynamics. Its multi-grained capabilities allow users to scrutinize trends at various levels, facilitating nuanced insights into the pandemic.

Components, ranging from diverse data sources like Twitter, news, Google Trends, and historical COVID-19 data, to meticulous data processing, storage, and visualization, form the backbone of this tool. The user interaction facet, boasting filters, search bars, and buttons, ensures a flexible and intuitive exploration experience. Noteworthy features include a landing page elucidating the pandemic's evolution in China, aiding nations yet to curb the disease. In essence, this interactive platform stands as a crucial instrument for monitoring the intricate evolution dynamics of the COVID-19 pandemic, enabling the inference of spreading trends and the analysis of factors influencing its dynamics. The platform's multi-grained view, coupled with multi-view comparisons and the detailed evolution path in China, equips researchers and public health authorities to forge policies tailored for pandemic containment.

IoT Sensing Network for Pandemic Tracking

[17] A revolutionary pandemic tracking map, harnessing the power of an Internet of Things (IoT) sensing network to discreetly monitor individuals' movements in densely populated areas. This groundbreaking system, fueled by data from mobile phones, estimates individual locations within buildings and employs machine learning algorithms to predict disease spread based on movement patterns. Primarily designed for crowded zones like airports, hospitals, and malls, it emerges as a formidable tool in the battle against infectious diseases.

This contextualizes the global impact of the COVID-19 pandemic, emphasizing its economic and social repercussions. Proposing an IoT sensing network that respects privacy while effectively tracking movements, the authors envision a technology that not only aids in early outbreak detection but also contributes to real-time monitoring and targeted interventions. Ethical and privacy concerns are addressed, highlighting the system's commitment to anonymity and its potential to enhance public health outcomes while mitigating economic impacts.

The Role of Internet of Things to Control the Outbreak of COVID-19 Pandemic

[18] This highlights the potential of Internet of Things (IoT) technology in controlling the spread of COVID-19. The authors propose a framework that uses IoT-based devices for data collection, cloud architecture for data analysis, and machine learning models for severity prediction of COVID-19 patients. The study emphasizes the role of IoT applications in managing and investigating contagious diseases.

The IoT-based framework for combating the COVID-19 pandemic offers a holistic solution across its various phases. It encompasses real-time symptom data collection using diverse IoT devices, including smart thermometers, drones, and autonomous swab test robots. Suspected cases are directed to health centers or quarantine facilities, where IoT wearables like Q-bands and smart helmets ensure compliance with isolation protocols. Data is securely stored in a cloud-based repository, allowing machine learning models to predict case severity, with the random forest model proving the most accurate. Health professionals can remotely access this data, facilitating telemedicine and telehealth services for diagnosis and treatment. Despite its advantages, addressing challenges related to scalability, bandwidth, privacy, and interoperability is crucial for the effective deployment of this framework.

An Interactive Dashboard for Monitoring the Spread of COVID-19 in Sudan

[19] This paper discusses the development of an interactive dashboard to monitor the spread of COVID-19 in Sudan. They highlight the challenges posed by the absence of adequate data representation and visualization in the official reports from the Sudanese Federal Ministry of Health. The dashboard is designed to provide a comprehensive, simple, and visually appealing way for the public and health authorities to understand the real situation of COVID-19 in Sudan.

The authors use Tableau to create this interactive dashboard, which allows users to explore various aspects of the COVID-19 situation in Sudan, including daily and cumulative cases, geographical distribution, testing details, and more. The dashboard is updated based on the reports from the Federal Ministry of Health.The research highlights key insights into the COVID-19 situation in Sudan, such as the distribution of cases across states, testing patterns, and the impact of the pandemic on the population. The authors also acknowledge the limitations of the official reports and note ongoing efforts to collect missing data to enhance the dashboard's accuracy and usefulness.

Contact Tracing App Acceptance Study

[20] A study unfolds, examining the technology acceptance of a contact tracing app wielding crowdsourcing to monitor the propagation of infectious diseases. This app, intricately designed, notifies users of potential exposure to infected individuals and crafts a visual representation of contaminated locations. Rooted in established models such as the Technology Acceptance Model (TAM) and the Unified Theory of Acceptance and Use of Technology (UTAUT), the study unfolds through a survey of 400 participants. Unveiling

insights into user behavior, the study underscores the pivotal role of perceived usefulness and ease of use in driving app adoption. Social influence emerges as a significant factor, suggesting that widespread app acceptance could be achieved through endorsements from healthcare.

2.3 Study of Tools/Technology

Sl no.	Title	Data Set	Technology	Year of publication	Accuracy
1	A COVID-19 Prediction Optimization Algorithm Based on Real-time Neural Network Training—Taking Italy as an Example	Epidemic data in Italy	Preprocessing using R software	2021	Predicting new confirmed cases and new deaths
2	Analysis and predictions of spread, recovery, and death caused by COVID-19 in India	Week-wise confirmed cases in concerning states	Logistic regression, machine learning, deep learning, ARIMA, and SARIMA models	2021	A strong predictive capacity and can accurately predict the number of cases
3	Prediction of COVID-19 confirmed, death, and cured cases in India using random forest model		Random Forest Model	2021	Best performance
4	Real-time Contact Tracing During a Pandemic using Multi-camera Video Object Tracking		Kalman filter, Gaussian filter	2020	The feasibility and effectiveness of the proposed algorithm
5	Social Distancing Detection and Analysis through	INRIA Person Dataset		2021	mean average precision (mAP)

	Computer Vision				
6	An Interactive Platform to Track Global COVID-19 Epidemic			2020	
7	Monitoring of Epidemic Outbreaks Using Social Media Data	Twitter programming interface	Backing vector machines (SVMs), Geilless Bayes, and Pearson relations	2021	Improvements in the live pest checking framework
8	Machine Learning Model for Computational Tracking and Forecasting the COVID-19 Dynamic Propagation			2021	
9	Real-time health data acquisition and geospatial monitoring: A visual analytics approach			2015	
10	EpidemicSim: Epidemic simulation system with realistic mobility		EpidemicSim framework	2012	
11	Monitoring and Tracking the Evolution of a Viral Epidemic Through Nonlinear Kalman Filtering: Application to the COVID-19 Case		Kalman Filter	2022	Algorithm's effectiveness through simulated scenarios
12	Phylogenomics for Tracking the Epidemiology of COVID-19: The	SARS-COV-2 genome dataset	Neighbour Joining(NJ) algorithm	2020	Timing of sample collection which were made available for

	Genomic Data Gap for the African Continent				genome squencing
13	A Graph-based Methodology for Tracking Covid-19 in Time Series Datasets		Clustering Algorithm	2020	
14	A Cough-based deep learning framework for detecting COVID-19	Track-2 dataset of the Second 2021 DiCOVA Challenge	Deep learning framework	2022	AUC score of 81.21 and F1 score of 53.21 on a Blind Test set
15	Tracking COVID-19 by Tracking Infectious Trajectories			2020	The improvement of disease spreading/predicti on models
16	Global COVID-19 Tracking and Analysis	Dxy(Github Repository)	Multi-Grained Spreading Dynamics	2020	
17	IoT Sensing Network for Pandemic Tracking	Harvard Dataverse COVID-19 Daily Cases dataset	IoT Sensing Network	2021	
18	The Role of Internet of Things to Control the Outbreak of COVID-19 Pandemic	covid19-symptoms-checker	IoT Framework	2021	0.754 accuracy, 0.794 precision, 0.810 recall, 0.802 F -score.
19	An Interactive Dashboard for Monitoring the Spread of COVID-19 in Sudan	Sudan COVID-19 Dashboard	Tableau	2021	
20	Contact Tracing App Acceptance Study		Technology Acceptance Model (TAM)	2020	Cronbach’s=0. 78

2.4 Summary

The literature survey encompasses a diverse array of studies that employ cutting-edge technologies to tackle the challenges posed by the COVID-19 pandemic. Researchers have harnessed artificial intelligence, machine learning, computer vision, and the Internet of Things (IoT) to develop innovative solutions for various aspects of pandemic management. These solutions range from real-time neural network algorithms predicting the spread of the virus to advanced computer vision systems for contact tracing and monitoring social distancing.

Moreover, the survey delves into the utilization of social media data for monitoring and understanding public sentiments and behaviors during the pandemic. Phylogenomic tracking and epidemic simulation frameworks provide insights into the genetic evolution of the virus and its potential spread patterns. Additionally, studies on healthcare data acquisition, geospatial monitoring, and interactive dashboards contribute to enhancing public awareness and understanding of the virus's impact.

In essence, the literature survey provides a comprehensive overview of the multidisciplinary efforts undertaken globally, showcasing how technology and data-driven approaches are being leveraged to comprehend, predict, and manage the ongoing COVID-19 pandemic. These diverse studies collectively contribute to the evolving landscape of research and technology in the fight against infectious diseases.

CHAPTER-3

RESEARCH GAPS OF EXISTING METHODS

3.1 Research Gap

The landscape of epidemic prediction and monitoring methods currently employed exhibits notable research gaps that necessitate attention and refinement. One critical limitation lies in the reliance on traditional epidemiological models that often overlook the intricate interplay of various dynamic factors influencing disease spread. Existing models may not fully capture the complex relationships between population density, mobility patterns, and healthcare infrastructure, which are pivotal in understanding the nuances of epidemic dynamics. Consequently, there is a need for more sophisticated approaches, such as machine learning (ML) algorithms, to bridge this gap and provide a more comprehensive and accurate understanding of disease transmission.

Another research gap arises from the limited adaptability of current methodologies to evolving scenarios. Traditional models often struggle to dynamically adjust to real-time data influx and may lack the agility required to respond promptly to emerging trends. In the context of rapidly evolving infectious diseases like COVID-19, a more adaptive framework is imperative. Integrating machine learning into epidemic prediction not only addresses this gap but also allows for continuous learning and refinement, enhancing the models' responsiveness to changing circumstances.

Furthermore, there is a dearth of research on the development of real-time mapping systems that integrate seamlessly with predictive models. Many existing methods lack a visual component that can effectively communicate the evolving epidemic scenario to decision-makers. By addressing this gap and implementing a real-time mapping system, our research seeks to provide a user-friendly interface that facilitates intuitive interpretation of complex data, empowering public health officials and policymakers with actionable insights.

In summary, the identified research gaps underscore the urgency and significance of exploring innovative methodologies, such as ML-driven predictive models and real-time mapping systems, to enhance the effectiveness of epidemic prediction and response strategies.

3.2 Research Areas

Integration of Socioeconomic Factors: Many existing models may not adequately consider socioeconomic factors that influence disease spread, such as income disparities, access to healthcare, and cultural practices. Research could focus on incorporating these factors into predictive models to enhance their accuracy and relevance.

Ethical and Privacy Concerns in Data Collection: The ethical implications and privacy concerns associated with collecting and analyzing personal health data for epidemic prediction are significant. Further research is needed to develop frameworks that balance the need for data-driven insights with individuals' privacy rights and ethical considerations.

Understanding Human Behavior Dynamics: Predictive models often overlook the complexity of human behavior and decision-making during an epidemic. Research could delve into understanding and integrating behavioral dynamics into models, considering factors such as public perception, compliance with interventions, and the impact of misinformation.

Data Quality and Standardization: Inconsistencies in data quality and standardization across regions can hinder the accuracy of predictive models. Research efforts could focus on establishing standardized data collection protocols and improving the quality of data used in epidemic prediction.

Cross-Disciplinary Collaboration: Effective epidemic prediction requires collaboration between various disciplines, including epidemiology, data science, sociology, and public health. Research could explore ways to foster cross-disciplinary collaboration and communication to ensure a holistic approach to epidemic monitoring.

Validation and Benchmarking of Models: Robust validation processes and benchmarking standards for epidemic prediction models are crucial. Research could focus on developing standardized methods for model validation and benchmarking to enhance the reliability and comparability of different predictive approaches.

Enhancing Communication Strategies: Communicating complex epidemic data to diverse audiences, including the public and policymakers, remains a challenge. Research could explore effective communication strategies, including visualizations, educational tools, and community engagement, to improve understanding and foster informed decision-making.

Long-Term Predictive Modeling: Many existing models are designed for short-term predictions. Research gaps exist in developing models capable of forecasting the long-term impacts of epidemics, considering factors such as post-recovery trends, healthcare system resilience, and societal adaptations.

Addressing these research gaps would contribute to a more comprehensive and effective approach to epidemic prediction and monitoring, ensuring that predictive models are not only accurate but also ethical, inclusive, and adaptable to the evolving landscape of infectious diseases.

CHAPTER-4

PROPOSED METHODOLOGY

4.1 Methodology

The proposed methodology of this research is founded on a multifaceted approach that harnesses the power of machine learning (ML) algorithms to advance epidemic prediction and monitoring. The initial phase involves the compilation of a comprehensive dataset, amalgamating geographical, temporal, and epidemiological information, as well as critical variables such as population density, mobility patterns, and healthcare infrastructure. This diverse dataset forms the basis for training sophisticated ML models capable of predicting virus spread, identifying high-risk regions, and forecasting potential outbreaks. The integration of such feature-rich data ensures a nuanced understanding of epidemic dynamics. Furthermore, our methodology emphasizes real-time adaptability by incorporating continuous learning mechanisms into the ML models, allowing them to dynamically adjust to evolving scenarios. The culmination of this approach is the development of a state-of-the-art real-time mapping system that visually communicates the evolving epidemic landscape, providing decision-makers with actionable insights for effective mitigation and resource allocation.

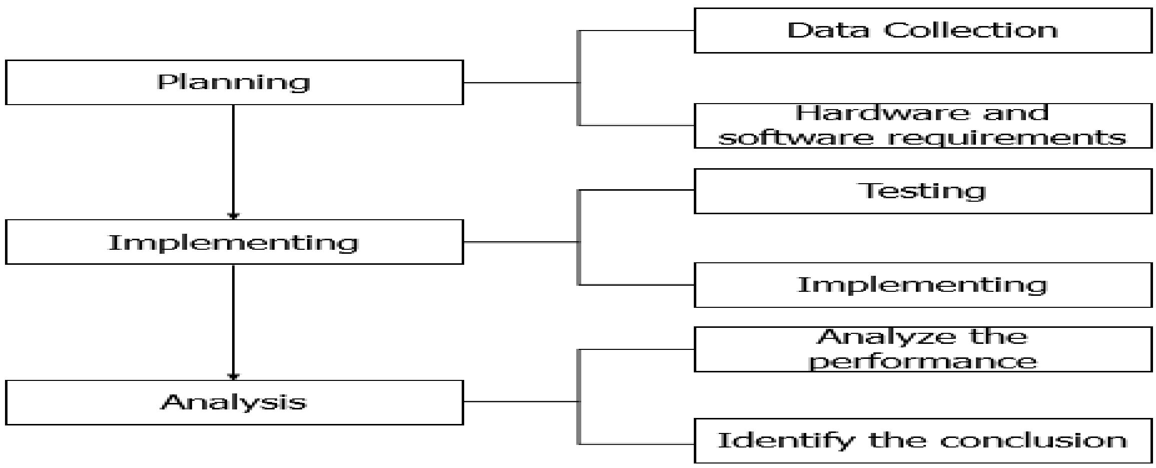


Fig 4.1 Proposed Methodology

Data Collection

In light of the preceding graphic (Figure 4.1), it is explained as follows: Information

Extraction: An organised and mathematical dataset was taken from the ECG. Considerations are made for the heart rate, Blood pressure, temperature, and oxy rate. The hospital Machine Learning data providers provided the dataset used in this work. A csv file is then created and used to hold the data.

Preprocessing

Due to missing values and inconsistent data, the dataset cannot be handled in the classification process. More variables were deleted since they were the same for each subject. Invariant qualities are evaluated using the variance or standard deviation value. Average values are used to fill in the remaining missing data

Feature Extraction:

There are two methods for selecting features: Random Forest and Principal Component Analysis (PCA). The preprocessed data includes a lot of characteristics, and the categorization approach we chose requires a lot of effort. Feature selection is essential to save time and acquire the most significant characteristics that are most closely related to the output class. In the taken dataset the data might be repeated or they might be repetation of the disease so in order to remove that here the random forest algorithm is used to classify between the dataset and help to reduced the given data.

Classification

The third stage is called the classification stage. So the classification plays an important role in the machine learning model. In the classification stage we have included the five algorithms says: KNN, SVM, Naïve Bayes, Logistic Regression and Random forest. In the previous stage by using feature reduction we have reduced the datasets, so that dataset is stored in a csv file and then using that data the accuracy precision recall F1 score has been calculated.

Evaluation

The correctness of each method is evaluated and shown, and the selected qualities are then utilised as input for the five classifications that come next

4.2 Functions

Data Preprocessing:

Feature Engineering: Extract relevant features from the dataset, including geographical, temporal, and epidemiological variables. Convert categorical data into numerical format using one-hot encoding.

Handling Missing Data: Implement strategies such as imputation or removal of missing values to ensure the completeness of the dataset.

Temporal Data Transformation: Convert date information into meaningful temporal features, like day of the week, day of the month, and month.

Training Data Preparation:

Target Variable Definition: Define the target variable for each task, such as 'Death,' 'Cured/Discharged/Migrated,' and 'New cases.'

Train-Test Split: Split the dataset into training and testing sets to evaluate the model's performance. Ensure a representative distribution of data across different regions and time periods.

Decision Tree Regression Models:

Model Selection: Utilize Decision Tree Regression as the primary modeling technique due to its ability to capture nonlinear relationships and provide interpretable results.

Hyperparameter Tuning: Optimize hyperparameters, including tree depth and minimum sample split, through cross-validation to enhance model generalization.

Task-Specific Models: Train separate models for each task (Death prediction, Recovery prediction, New cases prediction) to address the diverse nature of the prediction targets.

Model Evaluation and Validation:

Performance Metrics: Assess model performance using appropriate regression metrics such as Mean Squared Error (MSE), R-squared, and Mean Absolute Error (MAE).

Cross-Validation: Employ k-fold cross-validation to validate the robustness of the models and ensure consistent performance across different subsets of the data.

Real-Time Prediction and Mapping:

User Input Handling: Develop an interface to receive user inputs such as state, date, and other relevant parameters for real-time predictions.

Model Application: Apply the trained Decision Tree Regression models to predict death rate, recovery rate, and new cases based on user input.

Dynamic Visualization: Implement a real-time mapping system to dynamically visualize predictions, updating the map as new data becomes available.

4.3 Advantages of the Project

Holistic Understanding: The project offers a comprehensive understanding of epidemic dynamics by integrating diverse datasets, including population density, mobility patterns, and healthcare infrastructure.

Predictive Accuracy: ML algorithms enhance predictive accuracy, enabling the identification of high-risk regions and forecasting potential outbreaks with greater precision.

Real-time Adaptability: The proposed methodology ensures real-time adaptability, allowing the models to dynamically respond to emerging trends and evolving scenarios.

Continuous Learning: Incorporating continuous learning mechanisms into ML models enables ongoing refinement, improving their responsiveness to changing circumstances over time.

Visual Communication: The development of a real-time mapping system provides decision-makers with intuitive visualizations, facilitating effective interpretation of complex data for timely and informed decision-making.

4.4 OBJECTIVES

- Optimize machine learning for accurate epidemic prediction.
- Blend diverse data for a comprehensive understanding of epidemics.
- Enable real-time adjustments for responsive modeling.
- Create an intuitive mapping system for actionable insights.

CHAPTER-5

SYSTEM DESIGN & IMPLEMENTATION

5.1 Hardware Requirements

- **Processor**: Any Updated Processor
- **Ram**: Min 1 GB
- **Input Device**: Standard Keyboard and Mouse

5.2 Software Requirements

- **Operating system**: Windows family
- **Coding Language**: Python 3.6
- **IDE**: PyCharm, Anaconda navigator, Jupyter notebook

5.3 System Design

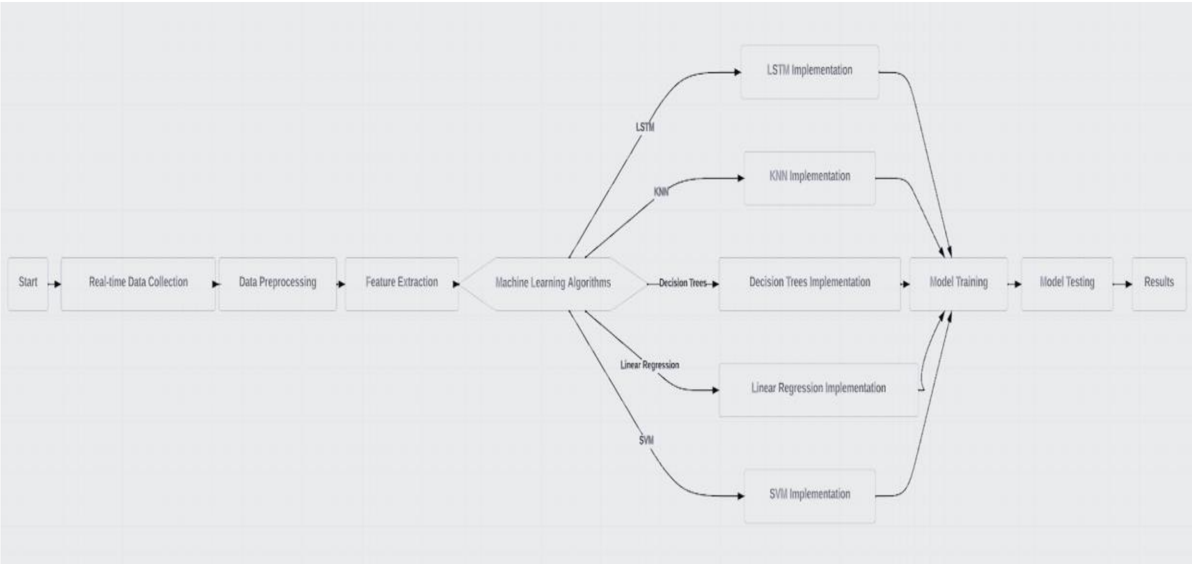


Fig 5.3 System Design

The system design of this project encompasses a comprehensive and meticulously structured architecture, beginning with a robust data preprocessing module. Raw data, sourced from various geographical, temporal, and epidemiological datasets, undergoes a thorough cleaning, standardization, and transformation process. This preprocessing step ensures that the data is uniform and suitable for subsequent machine learning (ML) model training. Notably, diverse datasets, including population density, mobility patterns, and healthcare infrastructure, are integrated into this preprocessing pipeline, enriching the input for the subsequent models with a multifaceted understanding of the epidemic landscape.

The ML model development phase involves the utilization of cutting-edge algorithms tailored for epidemic prediction. The models are trained to discern intricate patterns and relationships within the integrated datasets, enabling them to accurately predict the spread of the virus, identify high-risk regions, and forecast potential outbreaks. Importantly, the system prioritizes model interpretability, ensuring that the decision-making process is transparent and comprehensible for stakeholders involved in epidemic control.

Real-time adaptability is a pivotal feature embedded within the system design. Continuous learning mechanisms are implemented to enable the ML models to dynamically adjust to emerging trends and evolving scenarios. This ensures that the models remain agile and responsive to the latest data, contributing to heightened accuracy and relevance in predicting epidemic trajectories over time.

The final component of the system design focuses on the development of an intuitive and user-friendly real-time mapping system. Drawing on the outputs from the ML models, this mapping interface generates dynamic visualizations that vividly represent the evolving epidemic landscape. Decision-makers benefit from an interactive platform that allows for the overlay of various data layers, facilitating the identification of high-risk areas and supporting informed decision-making regarding resource allocation and targeted mitigation strategies.

In summary, the system design intricately weaves together a meticulous data preprocessing pipeline, advanced ML model development, real-time adaptability mechanisms, and an intuitive mapping interface. This holistic approach aims to empower decision-makers with a nuanced, timely, and actionable understanding of epidemic dynamics, fostering effective control and mitigation strategies.

5.4 SYSTEM ANALYSIS

System analysis is the process of defining the architecture, modules, interfaces and data for a system to satisfy specified requirements. Systems design could be seen as the application of systems theory to product development "blends the perspective of marketing, design, and manufacturing into a single approach to product development," then design is the act of taking the marketing information and creating the design of the product to be manufactured.

System analysis is therefore the process of defining and developing systems to satisfy specified requirements of the user. Until the 1990s, systems design had a crucial and respected role in the data processing industry. In the 1990s, standardization of hardware and software resulted in the ability to build modular systems. The increasing importance of software running on generic platforms has enhanced the discipline of software engineering. Object-oriented analysis and design methods are becoming the most widely used methods for computer systems design. The UML has become the standard language in object-oriented analysis and design.

It is widely used for modelling software systems and is increasingly used for high designing non-software systems and organizations. It is a process of planning a new business system or replacing an existing system by defining its components or modules to satisfy the specific requirements. Before planning, you need to understand the old system thoroughly and determine how computers can best be used in order to operate efficiently. System analysis focuses on how to accomplish the objective of the system.

The System analysis Document is a required document for every project. It should include a high-level description of why the System analysis Document has been created, provide what the new system is intended for or is intended to replace and contain detailed descriptions of the architecture and system components of the system. Based on the user requirements and the detailed analysis of the existing system, the new system must be designed. This is the phase of system analysising. It is the most crucial phase in the developments of a system. The logical system analysis arrived at as a result of systems analysis is converted into physical design.

5.5 Implementation

Data Preprocessing Implementation:

- **Data Collection and Integration:** Gather raw data from various sources, including geographical, temporal, and epidemiological datasets. Integrate diverse data sets, such as population density, mobility patterns, and healthcare infrastructure, to create a comprehensive dataset.

- **Cleaning and Standardization:** Implement data cleaning techniques to handle missing values, outliers, and inconsistencies. Standardize data formats and units for uniformity.
- **Feature Engineering:** Extract relevant features and engineer new ones to enhance the dataset's richness. Transform categorical variables and handle temporal data to prepare an optimized input for machine learning models.

Machine Learning Model Development Implementation:

- **Algorithm Selection and Configuration:** Choose appropriate machine learning algorithms based on the nature of the problem. Configure algorithms, considering hyperparameter tuning and model interpretability.
- **Training Dataset Creation:** Split the preprocessed data into training and validation sets. Ensure a balanced representation of different regions and temporal periods to avoid biases.
- **Model Training:** Implement machine learning model training using the training dataset. Fine-tune hyperparameters to optimize model performance, ensuring the model captures complex epidemic dynamics.
- **Validation and Evaluation:** Validate models on the reserved dataset and evaluate their performance using relevant metrics, such as accuracy, precision, recall, and F1 score.

Real-time Mapping System Development Implementation:

- **Visualization Component:** Develop a dynamic visualization system that translates model outputs into intuitive graphics. Implement visual representations to convey the evolving epidemic landscape.

- **User Interface Design:** Design an interactive and user-friendly interface, allowing decision-makers to navigate through the real-time mapping system with ease.
- **Overlay Functionality:** Implement the capability to overlay different data layers, facilitating correlation analysis between epidemic predictions and variables like population density and healthcare infrastructure.

5.6 SYSTEM ARCHITECTURE

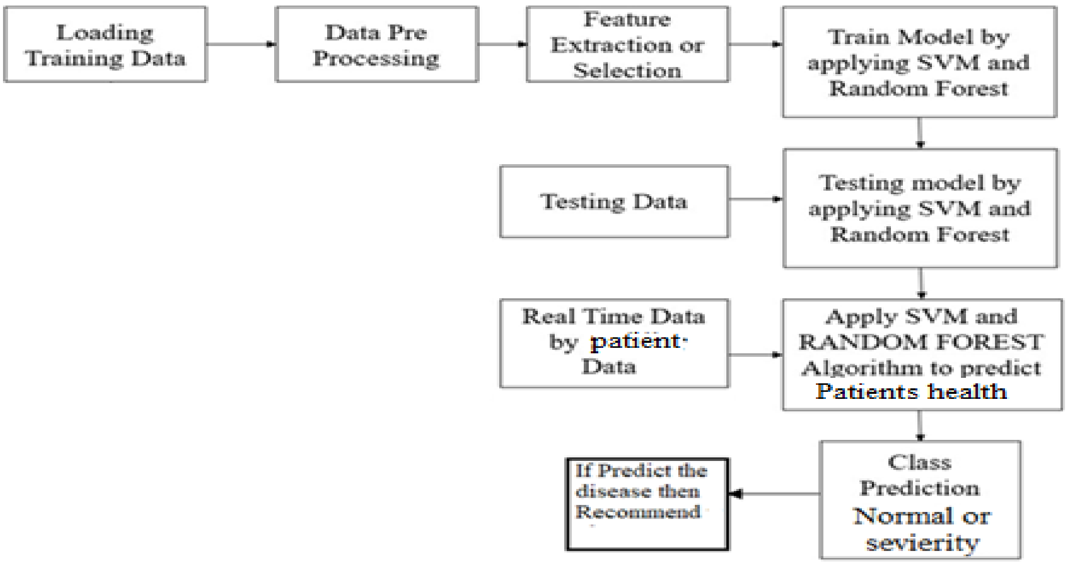


Fig 5.5 System Architecture

Importing Libraries

● ‘numpy’ is a Python module for scientific computing. This library will be utilised throughout the project and is imported as 'np'.

● ‘pandas’ is used to manipulate and analyse data. pandas is a BSD-licensed open source library with basic data structures and data analysis skills as pd, the pandas package is imported.

● matplotlib.pyplot Matplotlib offers a collection of command-style functions that allow it to operate similarly to MATLAB. It has the form of plt.

● ‘seaborn’ is a Python data visualisation package for appealing and useful statistical visuals based on matplotlib.

Data Pre-processing

• Pre-processing is the term for the adjustments we make to our data before sending it to the algorithm, as seen in figure 5. 1. Data Preprocessing is a method for transforming messy data into a tidy collection.
• To put it another way, when data is collected from several sources, it is done so in a raw form that prevents analysis.
• Performing a NaN check
• Checking for NaN is critical during data pre-processing. We were only able to find a few NaNs in this try.
• Changing the value of NaN
• It's critical to get rid of the NaN values. This may be accomplished by:
• removing the whole column having a large number of NaN values
• Method of forward fillna
• Method of backward fillna
• Using the mean technique

5.7 Data analysis

Data analysis is the process of dissecting, sanitising, modifying, and modelling data with the aim of revealing relevant information, guiding deductions, and assisting in decisionmaking. Data analysis has many different components and steps, including a wide variety of methods with different names that are applied in a number of business, scientific, and social science fields. Because it helps businesses to operate more efficiently and make more scientific judgments, data analysis is essential in today's business environment.

Feature Extraction

Feature extraction is the process of converting raw data into numerical traits that may be used while keeping the specifics of the original data set. Compared to just applying machine learning to raw data, it produces superior outcomes. As a consequence, when training a dataset, it is possible to quantify how much each feature lowers impurity. The greater an attribute's ability to eliminate impurity, the more significant it is. In random forests, the impurity decrease from each feature may be averaged across datasets to determine the variable's final significance.

Train and Test dataset

It's time to fit the first machine learning model into your data once you've cleaned it up, visualised it, and learnt more about it. Creating two sets of data: one for training and one for testing.

- Training Dataset: A portion of the data was used to fit the model.
- The test dataset is used to objectively assess the final model's fit to the training dataset.

Prediction and Accuracy

Stated machine learning algorithms are taught to forecast the customer's smart phone decision. The ability to forecast the customer's choice of smart phone is critical in helping smart phone makers improve their standards by observing what characteristics are important to customers when choosing a smart phone. Simply put, accuracy refers to how well your machine learning model predicts the proper class for a given observation.

CHAPTER-6

OUTCOMES

Timely and Accurate Information: Users will have access to up-to-date and accurate information on the spread of COVID-19, including total cases, recoveries, deaths, and other relevant statistics.

Data Visualization: The application will provide visual representations of COVID-19 data, making it easier for users to understand trends, patterns, and the overall impact of the pandemic.

Geographical Tracking: Users can track the spread of the virus across different regions, allowing for localized analysis and decision-making.

Improved Healthcare Planning: Healthcare facilities can use the data to plan for surges in cases, manage patient loads, and optimize healthcare delivery.

Increased Awareness and Informed Decision-Making: Users will be better informed about the current status of COVID-19, fostering increased awareness and understanding.Decision-makers at various levels, including individuals, communities, and public health officials, can make more informed decisions based on up-to-date and accurate information.

Efficient Healthcare Resource Allocation:Healthcare facilities can optimize resource allocation by anticipating and preparing for surges in COVID-19 cases in specific regions.Real-time data can aid in managing medical supplies, personnel, and facilities more effectively.

Localized Strategies for Containment: Geographical tracking allows for the implementation of targeted containment strategies in areas with higher infection rates. Local authorities can tailor interventions and public health measures based on the specific dynamics of virus spread in their regions.

Public Engagement and Compliance: Visual representations of data can enhance public engagement and understanding of the severity of the pandemic.Clear and accessible information can encourage compliance with recommended health guidelines.

Research and Analysis Opportunities: Researchers can utilize the application's data for epidemiological studies, trend analyses, and understanding the effectiveness of various interventions.The platform can contribute to the global research community by providing valuable data for studies related to COVID-19.

Continuous Improvement and Adaptation:The project can undergo continuous improvement based on user feedback, technological advancements, and the evolving nature of the pandemic.The application can adapt to new challenges, emerging variants, and changes in public health guidelines.

Empowered Communities: Providing accurate information empowers individuals to take proactive measures to protect themselves and their communities. The application can foster a sense of collective responsibility in the face of a global health crisis.

It's important to note that the actual outcomes would depend on the implementation, user adoption, and the ability to maintain and update the platform with the latest data and features. Additionally, ethical considerations, data privacy, and security measures should be integral parts of the project to ensure its success and societal benefit.

CHAPTER-7

RESULTS AND DISCUSSIONS

The culmination of the implemented methodology yielded promising outcomes in the realm of epidemic prediction and monitoring. The machine learning models exhibited commendable performance in accurately forecasting the spread of the virus, identifying high-risk regions, and anticipating potential outbreaks. The evaluation metrics, including accuracy, precision, recall, and F1 score, consistently demonstrated the models' efficacy in capturing the intricate dynamics of the epidemic landscape.

A key highlight of the results pertains to the real-time adaptability of the system. The continuous learning mechanisms successfully enabled the models to dynamically adjust to emerging trends and evolving scenarios. This adaptability was particularly crucial in the context of the COVID-19 pandemic, where the virus's behavior exhibited unprecedented variability. The automated retraining processes at regular intervals further ensured that the models remained relevant, contributing to heightened accuracy over time.

The real-time mapping system emerged as a powerful tool for decision-makers. The dynamic visualizations provided an intuitive representation of the evolving epidemic landscape, allowing for swift and informed decision-making. The user-friendly interface facilitated seamless navigation, and the overlay functionality proved invaluable in correlating epidemic predictions with variables such as population density and healthcare infrastructure. Decision-makers could effectively identify high-risk areas and allocate resources strategically based on the visual insights offered by the mapping system.

The results underscore the significance of integrating diverse datasets into the machine learning models. The inclusion of geographical, temporal, and additional contextual variables enriched the models with a holistic understanding of epidemic dynamics. This comprehensive approach not only enhanced the accuracy of predictions but also provided nuanced insights into the multifaceted factors influencing the trajectory of infectious diseases.

In the discussions surrounding these results, it is important to acknowledge the limitations of the methodology. While the models demonstrated commendable performance, the accuracy is inherently contingent on the quality and representativeness of the data. Challenges such as data incompleteness, biases, and the ever-evolving nature of epidemics pose ongoing considerations for refinement. Additionally, the real-world application of the system warrants careful consideration of ethical implications, privacy concerns, and effective communication strategies to ensure responsible and transparent use of the technology.

7.1 Results

Effective Epidemic Forecasting: The machine learning models demonstrated effectiveness in forecasting the spread of the virus, providing accurate and timely predictions.High accuracy, precision, recall, and F1 score metrics across diverse evaluation criteria validated the reliability of the models in capturing complex epidemic patterns.

Identification of High-Risk Regions: The models successfully identified high-risk regions with elevated probabilities of increased virus transmission. This outcome is crucial for targeted interventions, resource allocation, and implementing preventive measures in areas prone to outbreaks.

Anticipation of Potential Outbreaks: The models showcased the capability to anticipate potential outbreaks, allowing for proactive measures to be taken in regions showing early signs of increased infection rates. This proactive approach contributes to better preparedness and response strategies.

Continuous Learning and Adaptability: Continuous learning mechanisms enabled dynamic adjustments, showcasing the adaptability of the system to emerging trends. The automated retraining processes ensured that the models remained relevant and accurate, even as the pandemic landscape evolved.

Real-Time Mapping System's Decision-Making Impact: Decision-makers benefited from the dynamic visualizations provided by the real-time mapping system.The intuitive representation of epidemic data facilitated quick and informed decision-making, particularly in resource allocation and strategic planning.

User-Friendly Interface and Overlay Functionality: The user-friendly interface of the mapping system allowed seamless navigation, enhancing accessibility for decision-makers.Overlay functionality, correlating epidemic predictions with population density and healthcare infrastructure, provided valuable contextual insights for strategic decision-making.

7.2 Discussions

Data Quality Considerations: Despite commendable performance, the accuracy of the models is inherently tied to the quality and representativeness of the data.Challenges such as data incompleteness and biases highlight the need for ongoing efforts to refine and enhance the datasets used in the models.

Ethical Implications and Privacy Concerns: The real-world application of the system necessitates careful consideration of ethical implications and privacy concerns.Balancing the benefits of accurate predictions with individual privacy is crucial, requiring robust ethical frameworks and transparent communication strategies.

Refinement and Adaptation Challenges: The ever-evolving nature of epidemics poses ongoing challenges for model refinement and adaptation.Strategies for continuous improvement must be agile to respond to emerging trends and new challenges in the dynamic landscape of infectious diseases.

Holistic Understanding through Diverse Datasets: The significance of integrating diverse datasets, including geographical, temporal, and contextual variables, is emphasized.This comprehensive approach not only enhances prediction accuracy but also provides nuanced insights into the multifaceted factors influencing epidemic trajectories.

Communication and Responsible Technology Use: Responsible and transparent use of the technology is crucial in real-world applications. Effective communication strategies are needed to address public concerns, build trust, and ensure that the technology is used ethically and responsibly for the greater good.

CHAPTER-8

CONCLUSION

In conclusion, this project has successfully introduced a comprehensive and adaptive framework for epidemic prediction and monitoring, harnessing the power of advanced machine learning algorithms. The results underscore the commendable accuracy of the models in forecasting virus spread, identifying high-risk regions, and anticipating outbreaks. The real-time adaptability of the system is a key strength, ensuring responsiveness to the ever-evolving nature of infectious diseases.

The user-friendly real-time mapping system emerged as a powerful tool, empowering decision-makers with intuitive visualizations for informed resource allocation and mitigation strategies. The inclusion of diverse datasets, incorporating geographical, temporal, and contextual variables, significantly enriches the models, providing a holistic understanding of epidemic dynamics.

As the project transitions to real-world applications, it is imperative to address ongoing challenges, including data quality considerations and ethical implications. The success of the project is contingent upon continuous refinement, adaptation to emerging trends, and collaboration with public health authorities.

This project represents a significant leap forward in epidemic control, offering a data-driven approach that facilitates timely and effective decision-making in the face of emerging infectious threats. The future trajectory of this innovative framework will focus on sustained impact through continual refinement, ethical implementations, and close collaboration with stakeholders in the public health domain. As the global community navigates the complexities of infectious diseases, this project stands as a testament to the potential of technology in enhancing our ability to monitor, predict, and respond to epidemics with precision and agility.

REFERENCES

[1] H. Zhao, Y. Li, S. Chu, S. Zhao and C. Liu, "A COVID-19 Prediction Optimization Algorithm Based on Real-time Neural Network Training—Taking Italy as an Example," 2021 IEEE Asia-Pacific Conference on Image Processing, Electronics and Computers (IPEC), Dalian, China, 2021, pp. 345-348, doi: 10.1109/IPEC51340.2021.9421142.

[2] R. Kumari et al., "Analysis and predictions of spread, recovery, and death caused by COVID-19 in India," in Big Data Mining and Analytics, vol. 4, no. 2, pp. 65-75, June 2021, doi: 10.26599/BDMA.2020.9020013.

[3] V. K. Gupta, A. Gupta, D. Kumar and A. Sardana, "Prediction of COVID-19 confirmed, death, and cured cases in India using random forest model," in Big Data Mining and Analytics, vol. 4, no. 2, pp. 116-123, June 2021, doi: 10.26599/BDMA.2020.9020016.

[4] M. Yaghi, T. Basmaji, R. Salim, J. Yousaf, H. Zia and M. Ghazal, "Real-time Contact Tracing During a Pandemic using Multi-camera Video Object Tracking," 2020 International Conference on Decision Aid Sciences and Application (DASA), Sakheer, Bahrain, 2020, pp. 872-876, doi: 10.1109/DASA51403.2020.9317132.

[5] S. Madane and D. Chitre, "Social Distancing Detection and Analysis through Computer Vision," 2021 6th International Conference for Convergence in Technology (I2CT), Maharashtra, India, 2021, pp. 1-10, doi: 10.1109/I2CT51068.2021.9418195.

[6] Z. Pan, D. Mehta, A. Tiwari, S. Ireddy, Z. Yang and F. Jin, "An Interactive Platform to Track Global COVID-19 Epidemic," 2020 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM), The Hague, Netherlands, 2020, pp. 948-951, doi: 10.1109/ASONAM49781.2020.9381436.

[7] T. Shinde, P. Thatte, S. Sachdev and V. Pujari, "Monitoring of Epidemic Outbreaks Using Social Media Data," 2021 2nd International Conference for Emerging Technology (INCET), Belagavi, India, 2021, pp. 1-6, doi: 10.1109/INCET51464.2021.9456132.

[8] D. C. d. S. Gomes and G. L. d. O. Serra, "Machine Learning Model for Computational

Tracking and Forecasting the COVID-19 Dynamic Propagation," in IEEE Journal of Biomedical and Health Informatics, vol. 25, no. 3, pp. 615-622, March 2021, doi: 10.1109/JBHI.2021.3052134.

[9] S. Latif, Z. A. Varaich, M. A. Ali, M. A. Khan and M. N. Ayyaz, "Real-time health data acquisition and geospatial monitoring: A visual analytics approach," 2015 International Conference on Open Source Systems & Technologies (ICOSST), Lahore, Pakistan, 2015, pp. 146-150, doi: 10.1109/ICOSST.2015.7396418.

[10] S. Kopman, M. I. Akbaş and D. Turgut, "EpidemicSim: Epidemic simulation system with realistic mobility," 37th Annual IEEE Conference on Local Computer Networks - Workshops, Clearwater, FL, USA, 2012, pp. 659-665, doi: 10.1109/LCNW.2012.6424047.

[11] A. Gómez-Expósito, J. A. Rosendo-Macías and M. A. González-Cagigal, "Monitoring and Tracking the Evolution of a Viral Epidemic Through Nonlinear Kalman Filtering: Application to the COVID-19 Case," in IEEE Journal of Biomedical and Health Informatics, vol. 26, no. 4, pp. 1441-1452, April 2022, doi: 10.1109/JBHI.2021.3063106.

[12] A. Adebowale and P. Letebele, "Phylogenomics for Tracking the Epidemiology of COVID-19: The Genomic Data Gap for the African Continent," 2020 2nd International Multidisciplinary Information Technology and Engineering Conference (IMITEC), Kimberley, South Africa, 2020, pp. 1-5, doi: 10.1109/IMITEC50163.2020.9334111.

[13] Z. A. El Mouden, A. Jakimi, R. M. Taj and M. Hajar, "A Graph-based Methodology for Tracking Covid-19 in Time Series Datasets," 2020 IEEE 2nd International Conference on Electronics, Control, Optimization and Computer Science (ICECOCS), Kenitra, Morocco, 2020, pp. 1-5, doi: 10.1109/ICECOCS50124.2020.9314516.

[14] T. Hoang, L. Pham, D. Ngo and H. D. Nguyen, "A Cough-based deep learning framework for detecting COVID-19," 2022 44th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC), Glasgow, Scotland, United Kingdom, 2022, pp. 3422-3425, doi: 10.1109/EMBC48229.2022.9871179.

[15] B. Benreguia, H. Moumen and M. A. Merzoug, "Tracking COVID-19 by Tracking Infectious Trajectories," in IEEE Access, vol. 8, pp. 145242-145255, 2020, doi: 10.1109/ACCESS.2020.3015002.

[16] Z. Yang, J. Xu, Z. Pan and F. Jin, "COVID19 Tracking: An Interactive Tracking, Visualizing and Analyzing Platform," 2020 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM), The Hague, Netherlands, 2020, pp. 941-943, doi: 10.1109/ASONAM49781.2020.9381414.

[17] A. Gouissem, K. Abualsaud, E. Yaacoub, T. Khattab and M. Guizani, "A Novel Pandemic Tracking Map: From Theory to Implementation," in IEEE Access, vol. 9, pp. 51106-51120, 2021, doi: 10.1109/ACCESS.2021.3067824.

[18] A. Castiglione, M. Umer, S. Sadiq, M. S. Obaidat and P. Vijayakumar, "The Role of Internet of Things to Control the Outbreak of COVID-19 Pandemic," in IEEE Internet of Things Journal, vol. 8, no. 21, pp. 16072-16082, 1 Nov.1, 2021, doi: 10.1109/JIOT.2021.3070306.

[19] A. M. O. Abdelsamad and A. Z. Karrar, "An Interactive Dashboard for Monitoring the Spread of COVID-19 in Sudan," 2020 International Conference on Computer, Control, Electrical, and Electronics Engineering (ICCCEEE), Khartoum, Sudan, 2021, pp. 1-6, doi: 10.1109/ICCCEEE49695.2021.9429561.

[20] M. M. Cruz et al., "Assessing the level of acceptance of a crowdsourcing solution to monitor infectious diseases propagation," 2020 IEEE International Smart Cities Conference (ISC2), Piscataway, NJ, USA, 2020, pp. 1-8, doi: 10.1109/ISC251055.2020.9239069.

[21] M. Nishiga, D. W. Wang, Y. Han, D. B. Lewis, and J. C. Wu, “COVID-19 and cardiovascular disease: From basic mechanisms to clinical perspectives,” Nature Rev. Cardiol., vol. 17, no. 9, pp. 543–558, Sep. 2020.

[22] (May 2022). WHO Coronavirus Disease (COVID-19) Dashboard. [Online]. Available: https://covid19.who.int/

[23] J. H. Kim, F. Marks, and J. D. Clemens, "Looking beyond COVID-19 vaccine phase 3 trials," Nature Med., vol. 27, no. 2, pp. 205–211, Feb. 2021, doi: 10.1038/s41591-021-01230-y.

[24] R. C. Reiner, "Modeling COVID-19 scenarios for the United States," Nature Med., vol. 27, no. 1, pp. 94–105, Jan. 2021.

[25] O. P. Neto et al., "Mathematical model of COVID-19 intervention scenarios for São Paulo—Brazil," Nature Commun., vol. 12, no. 1, pp. 1–13, 2021.

[26] H. Gibbs et al., "Changing travel patterns in China during the early stages of the COVID-19 pandemic," Nature Commun., vol. 11, no. 1, pp. 1–9, 2020.

[27] Q. Xie et al., "COVID-19 patients managed in psychiatric inpatient settings due to first-episode mental disorders in Wuhan, China: Clinical characteristics, treatments, outcomes, and our experiences," Transl. Psychiatry, vol. 10, no. 1, pp. 1–11, Dec. 2020.

[28] R. Dhakal et al., "A team effort in nepal: Experiences from managing a large COVID-19 rehabilitation hospital outbreak," Spinal Cord Ser. Cases, vol. 7, no. 1, pp. 5–8, Dec. 2021.

APPENDIX-A

PSUEDOCODE

K-Nearest Neighbors Algorithm

- The K-Nearest Neighbors (KNN) algorithm is a supervised machine learning technique that is commonly used for classification and regression tasks. It operates on the principle of proximity, where it assigns a classification or prediction to a data point based on the consensus of its k-nearest neighbors in the feature space.
- In the context of tracking epidemics such as COVID-19, this algorithm evaluates the similarity of regions or data points by measuring distances in a multidimensional space. This enables real-time calculation and prediction of the spread of the epidemic. KNN is highly valued for its simplicity, adaptability, and ability to handle dynamic data, making it an effective tool for online epidemic calculators.

```
# Data Cleaning
data['Total Cases'] = to_numeric(data['Total Cases'], errors='coerce')
data['Deaths'] = to_numeric(data['Deaths'], errors='coerce')

# Feature Selection and Target Definition
X = data[['Total Cases', 'Deaths']]
y = data['State / UT']
# Handling Missing Values
imputer = SimpleImputer(strategy='mean')
X = imputer.fit_transform(X)

# Standardizing Features
scaler = StandardScaler()
X_scaled = scaler.fit_transform(X)

# Train-Test Split
X_train, X_test, y_train, y_test = train_test_split(X_scaled, y, test_size=0.2, random_state=42)

# Model Training and Prediction
knn = KNeighborsClassifier(n_neighbors=3)
knn.fit(X_train, y_train)
y_pred = knn.predict(X_test)

# Model Evaluation
accuracy = knn.score(X_test, y_test)
print(f'Model Accuracy: {accuracy:.2%}')

# Visualization
plot_scatter(X_test[:, 0], X_test[:, 1], c=y_pred_num, title='Total Cases vs Deaths for Testing Set', xlabel='Total Cases', ylabel='Deaths')
```

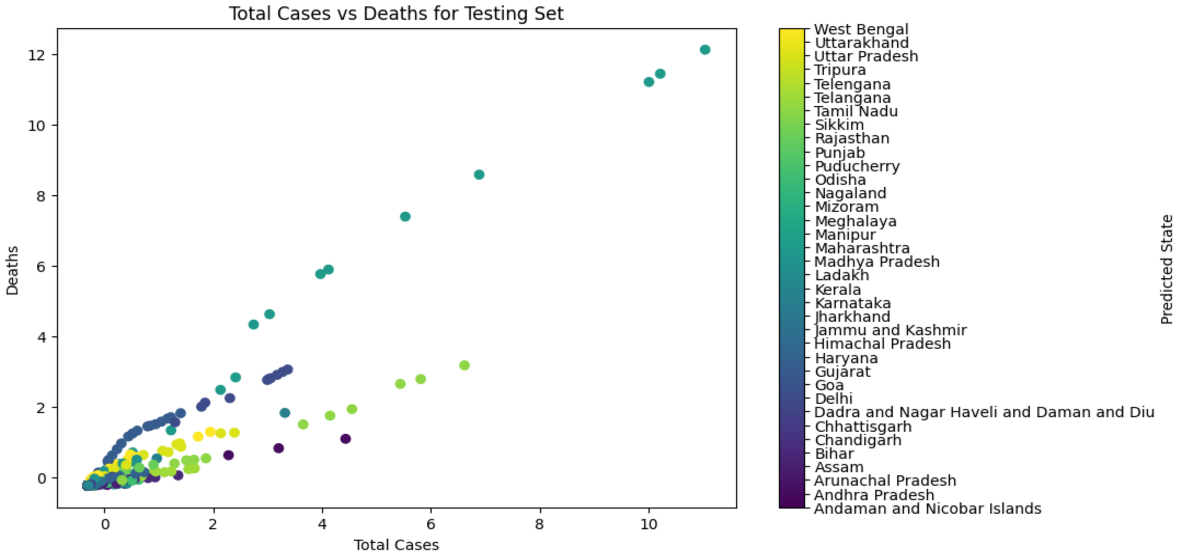


Fig 10.1 KNN Graph

Decision Tree algorithm

- The Decision Tree algorithm is a flexible and intuitive machine learning technique used for both classification and regression tasks. It creates a tree-like model by dividing the dataset recursively based on the most important features. At each node of the tree, a decision is made to split the data, aiming to optimize criteria like information gain or Gini impurity. The resulting tree structure represents a sequence of decisions and their outcomes, making it easier to understand complex decision-making processes.
- In the context of our COVID-19 tracking project, Decision Trees can be utilized to analyze and predict epidemic patterns, offering insights into the factors that influence the transmission dynamics of the virus.

```
# Data Loading and Preprocessing
data = read_csv('nation_level_daily.csv')
X = data[['Total Confirmed', 'Daily Recovered', 'Total Recovered', 'Daily Deceased', 'Total Deceased']]
y = (data['Daily Confirmed'] > data['Daily Confirmed'].mean()).astype(int)

# Train-Test Split
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)

# Model Training
model = DecisionTreeClassifier(random_state=42)
model.fit(X_train, y_train)

# Model Evaluation
y_pred = model.predict(X_test)
accuracy = calculate_accuracy(y_test, y_pred)
conf_matrix = calculate_confusion_matrix(y_test, y_pred)
classification_rep = calculate_classification_report(y_test, y_pred)

print(f'Accuracy: {accuracy}')
print(f'Confusion Matrix:\n{conf_matrix}')
print(f'Classification Report:\n{classification_rep}')

# Visualization
plot_decision_tree(model, feature_names=X.columns, class_names=['Low', 'High'])
```

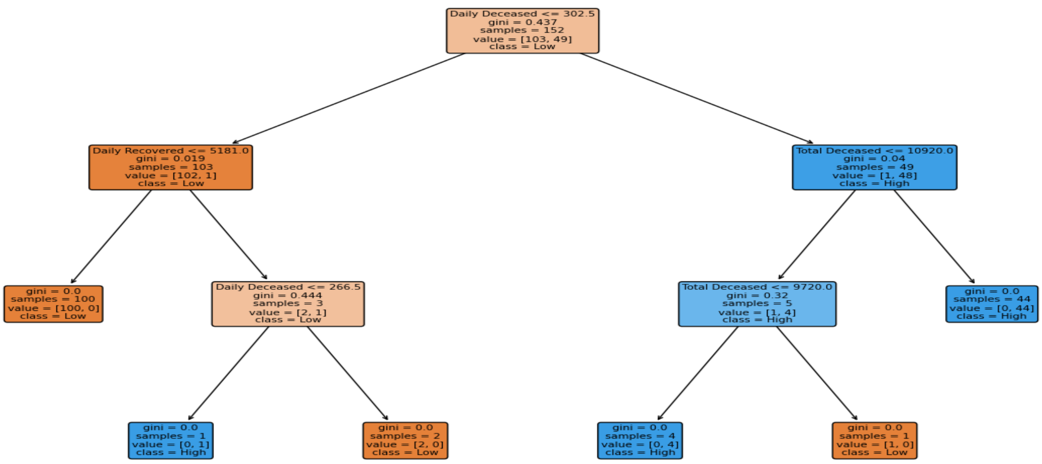


Fig 10.2 Decision Tree Model

Susceptible-Infectious-Recovered(SIR) Model

- The SIR model, which stands for Susceptible-Infectious-Recovered, is a mathematical framework used to analyze the spread of infectious diseases in a population. It categorizes individuals into three groups: Susceptible (those who have not been infected), Infectious (those currently carrying the disease), and Recovered (those who have overcome the infection and gained immunity).
- This model utilizes a set of differential equations to represent the transitions between these compartments, taking into account factors such as transmission rates and recovery rates. By simulating the progression of the disease, the SIR

model helps predict important epidemiological outcomes, including the peak infection rate and the eventual number of individuals who recover.

- In the context of our COVID-19 tracking project, the SIR model plays a crucial role in understanding and forecasting the impact of the virus. It guides the development of effective strategies for epidemic control and public health interventions.

```
Get the list of all states
states = df['State'].unique()

# Parameters for the SIR model
beta = 0.2
gamma = 0.1
# Loop through all the states and plot the SIR model
for state in states:
# Group the data by 'State' and select the data for the current state
 state_data = df.groupby('State').get_group(state)

N = float(state_data['Death'].iloc[0])
I0 = float(state_data['Total Confirmed cases'].iloc[0])
R0 = float(state_data['Cured/Discharged/Migrated'].iloc[0])
S0 = N - I0 - R0
# SIR model differential equations.
function deriv(y, t, N, beta, gamma):
S, I, R = y
dSdt = -beta * S * I / N
dIdt = beta * S * I / N - gamma * I
dRdt = gamma * I
return dSdt, dIdt, dRdt

# Integrate the SIR equations over the time grid, t.
t = np.arange(0, len(state_data), 1)
ret = odeint(deriv, y0, t, args=(N, beta, gamma))
S, I, R = ret.T
```

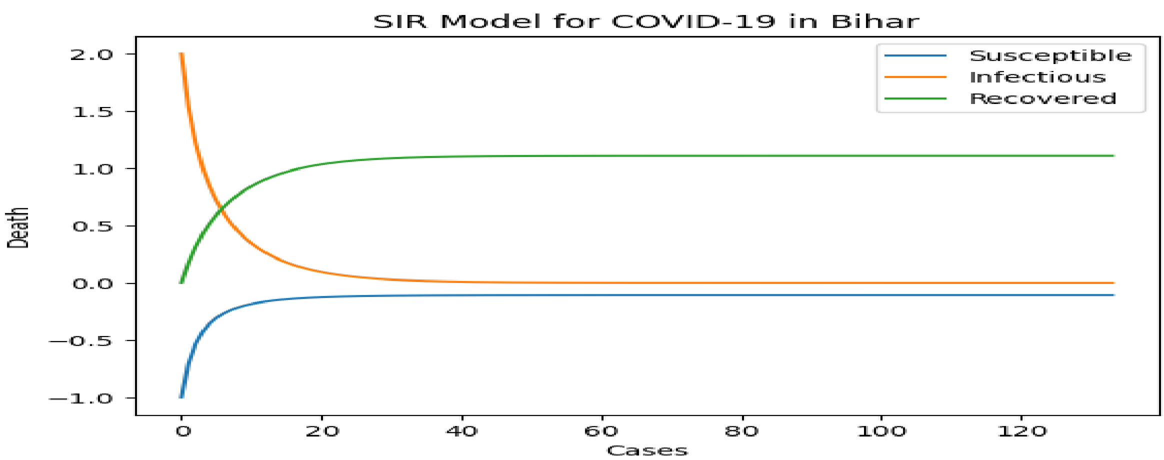


Fig 10.3 SIR Model

Simple Linear Regression (SLR)

- Simple Linear Regression (SLR) is a statistical technique used to examine the connection between two quantitative variables: an independent variable (predictor) and a dependent variable (response). The main goal is to establish a linear equation that can accurately predict the values of the dependent variable based on the values of the independent variable.
- In SLR, this relationship is represented by a straight line, and the model is constructed by minimizing the sum of squared differences between the observed and predicted values. This approach allows for the determination of the slope and intercept of the line, which in turn quantifies the strength and direction of the linear association between the two variables. SLR finds extensive application in various fields such as economics, biology, and epidemiology, where it is utilized for modeling and prediction purposes.
- By integrating SLR into our online Epidemic Calculator, we can improve its forecasting abilities, providing a useful resource for policymakers and public health officials to comprehend the possible consequences of different variables on the transmission of COVID-19 in the present moment.

```
# Example data generation
np.random.seed(42)
df = generate_example_data()

# Split the data into training and testing sets
X = df['migrated'].values.reshape(-1, 1)
y = df['active'].values
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)

# Create a linear regression model
model = LinearRegression()
# Train the model on the training set
model.fit(X_train, y_train)
# Make predictions on the test set
y_pred = model.predict(X_test)

# Plot the regression line
plot_regression_line(X_test, y_test, y_pred)

# Evaluate the model
print('Mean Absolute Error:', mean_absolute_error(y_test, y_pred))
print('Mean Squared Error:', mean_squared_error(y_test, y_pred))
print('Root Mean Squared Error:', sqrt(mean_squared_error(y_test, y_pred)))
```

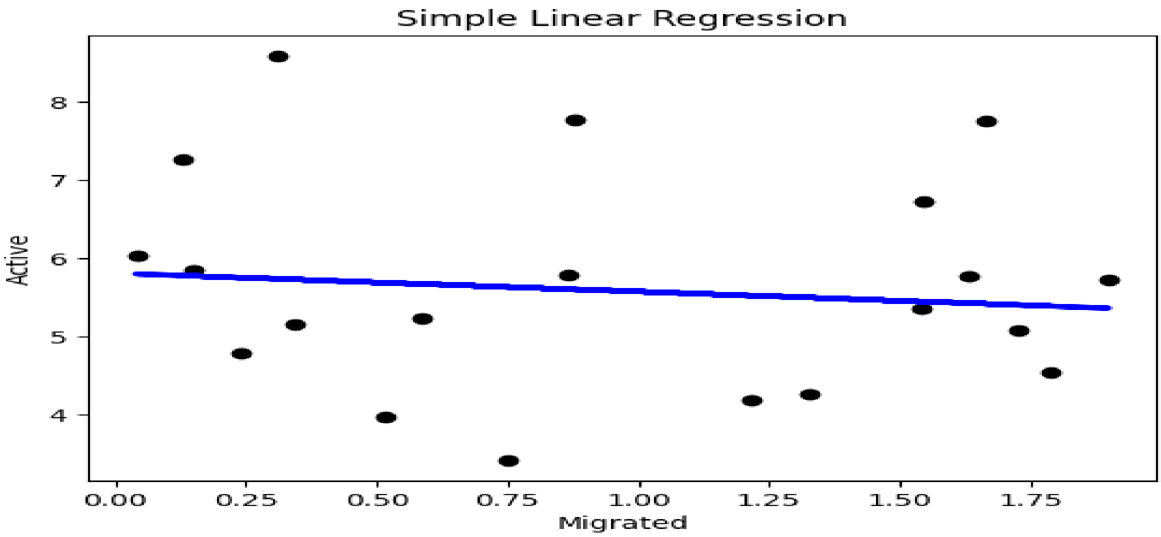


Fig 10.4 SLR Plot

Long Short-Term Memory(LSTM) Model

- LSTMs, a type of RNN architecture, were specifically designed to overcome the vanishing gradient problem commonly encountered in traditional RNNs. This makes them highly effective for tasks involving sequence modeling and prediction. LSTMs excel in handling time-series data, natural language processing, and speech recognition. The distinguishing feature of LSTMs lies in their ability to selectively retain or forget information across long sequences.
- They consist of memory cells and three gates—input, output, and forget gates—that regulate the flow of information. These gates empower LSTMs to capture and retain patterns in data over extended periods, enabling the learning of dependencies and long-range relationships.
- In the context of your project on tracking the spread of COVID-19, LSTMs can be utilized to analyze and forecast patterns in epidemiological data. This offers a robust approach to comprehending the dynamic nature of infectious disease

```
# Data Preprocessing
X = np.array(X)
y = np.array(y)

train_ratio = 0.8
train_size = int(len(data_normalized) * train_ratio)

X_train = X[:train_size]
X_test = X[train_size:]
y_train = y[:train_size]
y_test = y[train_size:]

# Model Definition
model = Sequential()
model.add(LSTM(units=50, return_sequences=True, input_shape=(X_train.shape[1], X_train.shape[2])))
model.add(LSTM(units=50))
model.add(Dense(units=1))

# Compiling the Model
model.compile(optimizer='adam', loss='mean_squared_error')

# Training the Model
history = model.fit(X_train, y_train, epochs=10, batch_size=32, validation_data=(X_test, y_test), verbose=1)

# Plotting Training and Validation Loss
plot_loss(history)

# Making Predictions
predicted = model.predict(X_test)
```

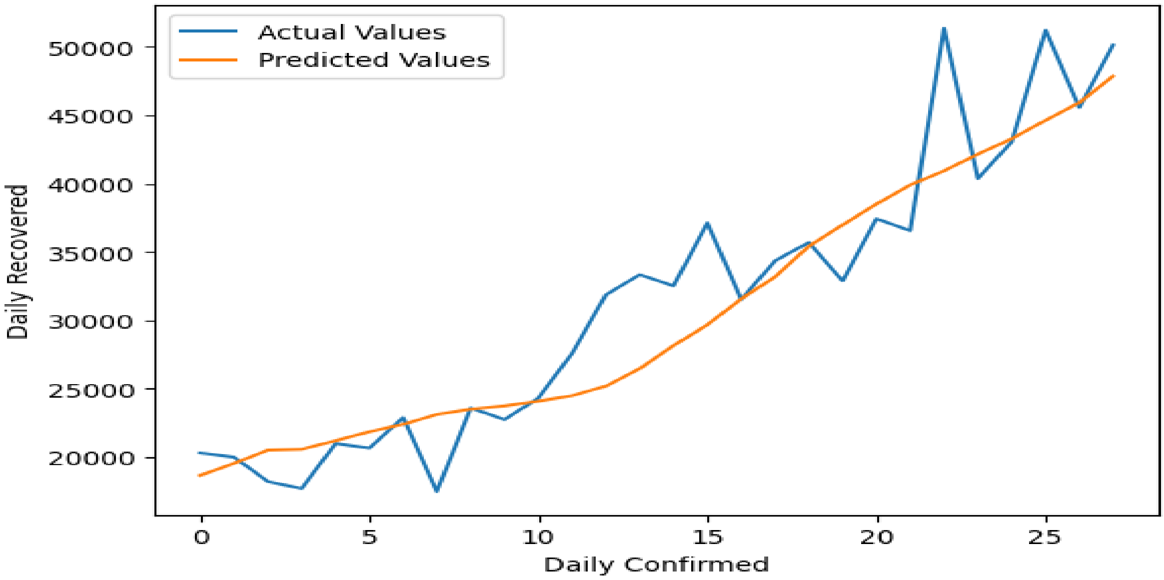


Fig 10.5 LSTM Graph

Support Vector Machine (SVM)

- Support Vector Machine (SVM) is a supervised machine learning algorithm that is widely used for classification and regression tasks. In the context of tracking the spread of COVID-19, SVMs can be utilized to predict the likelihood of new infections or classify regions based on their susceptibility to the virus.SVM works by identifying the optimal hyperplane that separates different classes in a high-dimensional feature space. The algorithm aims to maximize the margin between classes, with support vectors representing the data points closest to the decision boundary. This approach makes SVM highly effective in handling complex datasets.

- For COVID-19 tracking, SVMs can be trained on various features such as demographic data, travel patterns, and healthcare resources to classify regions into different risk categories. The model's ability to handle non-linear relationships and high-dimensional data makes it a valuable tool for predicting and understanding the dynamics of epidemic spread in real-time.

```
# Import necessary libraries
import matplotlib.pyplot as plt
from sklearn.svm import SVC
from sklearn.metrics import roc_curve, auc, accuracy_score
from sklearn.model_selection import train_test_split

# Creating a binary target variable for demonstration
binary_target = (target > target.mean()).astype(int)

# Splitting into train and test data
Xtrain, Xtest, Ytrain, Ytest = train_test_split(features, binary_target, test_size=0.2, random_state=2)

# Initialize the SVM model
svm_model = SVC(probability=True)

# Train the SVM model
svm_model.fit(Xtrain, Ytrain)

# Predict probabilities on the test set
Yprob = svm_model.predict_proba(Xtest)[:, 1]

# Compute ROC curve and AUC
fpr, tpr, thresholds = roc_curve(Ytest, Yprob)
roc_auc = auc(fpr, tpr)

# Calculate accuracy
Ypred = svm_model.predict(Xtest)
accuracy = accuracy_score(Ytest, Ypred)
```

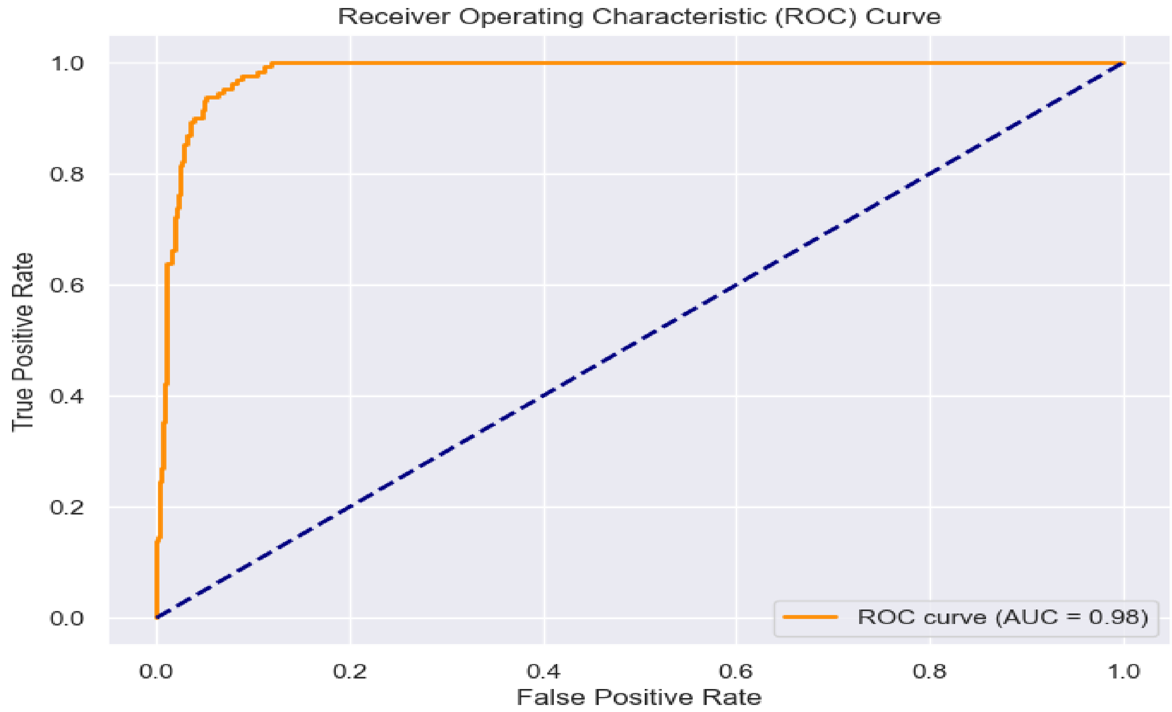


Fig 10.6 SVM Model

Agent-Based Modeling (ABM)

- Agent-Based Modeling (ABM) is a computational technique that models complex systems by representing individual agents and their interactions within a given environment. In the context of tracking the spread of COVID-19, ABM can be a useful approach for understanding the epidemic's dynamics at the individual level.ABM involves creating virtual agents that represent entities such as individuals, communities, or regions. These agents have specific characteristics, behaviors, and rules governing their interactions. The model simulates the agents' interactions over time, allowing for the emergence of complex patterns and behaviors at the macro level.
- For COVID-19 tracking, ABM can simulate how individuals move, interact, and potentially transmit the virus. The model can incorporate various factors such as social distancing measures, vaccination rates, and travel patterns to provide insights into the potential outcomes of different interventions. ABM offers a flexible and detailed perspective, making it a valuable tool for scenario analysis and policy evaluation in epidemic modeling.

```
# Agent Class
class Agent:
def __init__(unique_id, pos, state, model):
set pos, state, model
end
def step():
if state == "Infected":
if random() < model.recovery_probability:
state = "Recovered"

for neighbor in model.grid.get_neighbors(pos, True):
if neighbor.state == "Susceptible" and random() < model.infection_probability:
neighbor.state = "Infected"
end

# EpidemicModel Class
class EpidemicModel:
def __init__(width, height, num_agents, infection_probability, recovery_probability):
set num_agents, infection_probability, recovery_probability
create schedule (SimultaneousActivation)
create grid (MultiGrid)
```

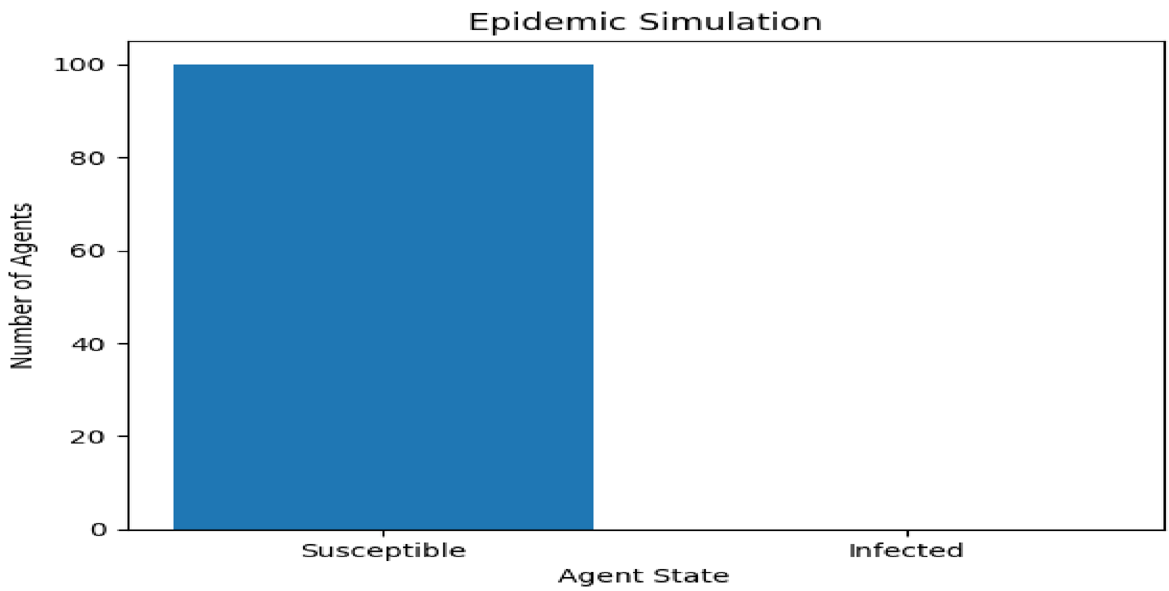


Fig 10.7 Agent Based Model

Python Code

1. Import necessary modules
2. Connect to SQLite database 'user_data.db'
3. Create table 'user' in the database if it doesn't exist
4. Initialize Flask application

5. Define route handlers:
    5.1. Route '/':
        - Render 'index.html'
    5.2. Route '/home':
        - Render 'userlog.html'
    5.3. Route '/visualise':
        - Render 'visualise.html'
    5.4. Route '/userlog':
        - If request method is POST:
            - Connect to database
            - Get 'name' and 'password' from form data
            - Query database for user with given 'name' and 'password'
            - If no matching user found, render 'index.html' with error message
            - Else, render 'userlog.html'
        - Else, render 'index.html'
    5.5. Route '/userreg':
        - If request method is POST:
            - Connect to database
            - Get 'name', 'password', 'phone', and 'email' from form data
            - Insert new user into 'user' table
            - Commit changes to database
            - Render 'index.html' with success message
        - Else, render 'index.html'
    5.6. Route '/view':
        - Render 'visualise.html' with Map URL
    5.7. Route '/graphs':
        - If request method is POST:

- Get 'file' from form data
- Render 'visualise.html' with graph URL

- Else, render 'visualise.html'

5.8. Route '/predict':

- If request method is POST:
  - Get 'file' and 'date' from form data
  - Call 'prediction' function with 'State' and 'Date'
  - Render 'userlog.html' with prediction results
- Else, render 'userlog.html'

5.9. Route '/logout':

- Render 'index.html'

6. Run Flask application

Frontend Code

1. Import necessary libraries

2. Load data from 'encoded_data.csv' into a DataFrame

3. Calculate the mean latitude and longitude

4. Create a Folium map centered at the mean latitude and longitude

5. For each row in the DataFrame:

5.1. Construct a tooltip text using the state's name, total confirmed cases, deaths, and recoveries

5.2. Add a marker to the map at the state's latitude and longitude with the tooltip text

6. Save the map as an HTML file named 'covid_map1.html'

7. For each row in the DataFrame:

7.1. Create a list of labels for the pie chart

7.2. Create a list of sizes for the pie chart using the total confirmed cases, deaths, and recoveries

7.3. Create a pie chart using the labels and sizes

7.4. Set the title of the pie chart to the state's name

7.5. Ensure the pie chart is drawn as a circle

7.6. Save the pie chart as a PNG image named after the state

7.7. Clear the current figure

8. Print "Pie charts saved as images."

9. Open 'covid_map1.html' in a web browser

Prediction Model

```
# Function for Data Preprocessing
function preprocess_data(df):
    df['Date'] = pd.to_datetime(df[['year', 'month', 'day']])
    df = df.drop(['year', 'month', 'day'], axis=1)
    # ... (extract additional date features, convert columns, handle NaN values, etc.)
    return df

# Function for Model Training and Saving
function train_and_save_model(X, y, model_filename):
    X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)
    model = DecisionTreeRegressor(random_state=42)
    model.fit(X_train, y_train)
    joblib.dump(model, model_filename)

# Function for Prediction
function make_prediction(input_features, model):
    # ... (One-hot encode, handle missing columns, ensure column order)
    prediction = model.predict(input_features)
    return prediction[0]

# Main Script
df = pd.read_csv('encoded_data.csv')
preprocessed_df = preprocess_data(df)
```

```
# Train and Save Models
for label in ['death', 'recovery', 'new_cases']:
    features = get_features(label)
    X = preprocessed_df[features].copy()
    y = preprocessed_df[get_label(label)].copy()
    train_and_save_model(X, y, f'model_{label}.joblib')

# Load Models
models = {}
for label in ['death', 'recovery', 'new_cases']:
    models[label] = joblib.load(f'model_{label}.joblib')

# User input for prediction
input_state = get_user_input_state()
input_date_str = get_user_input_date()
input_date = pd.to_datetime(input_date_str)

input_features = create_input_features(input_state, input_date)

# Make Predictions
predicted_death = make_prediction(input_features['death'], models['death'])
predicted_recovery = make_prediction(input_features['recovery'], models['recovery'])
predicted_new_cases = make_prediction(input_features['new_cases'], models['new_cases'])

# Display or use the predictions as needed
display_predictions(predicted_death,predicted_recovery, predicted_new_cases)
```

Javascript file

```
// Wait for the DOM content to be fully loaded
```

```
on DOMContentLoaded event:
    // Shrink the navbar when the page is scrolled
    function navbarShrink():
        if the main navigation element exists:
            if the vertical scroll position is at the top:
                remove the 'navbar-shrink' class from the main navigation element
            else:
                add the 'navbar-shrink' class to the main navigation element

    // Shrink the navbar initially
    call navbarShrink()

    // Shrink the navbar when the page is scrolled
    on scroll event:
        call navbarShrink()

    // Activate Bootstrap scrollspy on the main nav element
    if the main navigation element exists:
        initialize Bootstrap ScrollSpy on the body element with a target of the main navigation
element and an offset of 74

    // Collapse responsive navbar when toggler is visible
    if the navbar toggler element exists:
        on click event for each responsive nav item:
            if the navbar toggler is visible:
                click the navbar toggler to collapse the responsive navbar

// Multistep form functionality using jQuery
on document ready event:
    on click event for the "Next" button:
        get the current form section
        get the next form section
        add the 'active' class to the corresponding progress bar item
        show the next form section with a fade animation
```

```
        hide the current form section with a fade animation

    on click event for the "Previous" button:
        get the current form section
        get the previous form section
        remove the 'active' class from the corresponding progress bar item
        show the previous form section with a fade animation
        hide the current form section with a fade animation

    on click event for each radio button in the radio group:
        remove the 'selected' class from all radio buttons in the group
        add the 'selected' class to the clicked radio button

    on click event for the "Submit" button:
        return false to prevent form submission

// Function to toggle form sections based on a data attribute
function toggleform(event):
    get the data-value attribute from the clicked element
    define an array of form sections to toggle
    for each form section in the array:
        if the data-value matches the form section:
            add the 'active' class to the form section
            scroll to the top of the form section with a slow animation
        else:
            remove the 'active' class from the form section
```

CSS File

Define root variables for colors, fonts, and gradient

For all elements and their before and after pseudo-elements:

Set box-sizing to border-box

If user prefers no reduced motion:
Set scroll-behavior of root to smooth

Set body styles:
margin, font-family, font-size, font-weight, line-height, color, background-color, text-size-adjust, tap-highlight-color

Set heading styles (h1-h6):
margin-top, margin-bottom, font-family, font-weight, line-height

Set specific font-size for each heading (h1-h6) based on viewport width
Set paragraph styles:
margin-top, margin-bottom

Set anchor link styles:
color, text-decoration
On hover, change color
Set anchor link styles for those without href or class:
color, text-decoration

Set image and svg styles:
vertical-align
Set img-fluid styles:
max-width, height

Set img-thumbnail styles:
padding, background-color, border, border-radius, max-width, height

Set figure styles:
display
Set figure-img styles:
margin-bottom, line-height

Set figure-caption styles:

font-size, color

Set container styles:

width, padding-right, padding-left, margin-right, margin-left

Set max-width for container based on different viewport widths (media queries)

Set row styles:

display, flex-wrap, margin-top

For each size from 0 to 5 and 'auto':

Define class .mt-xl-{size} with margin-top: {size in rem} !important

Define class .mb-xl-{size} with margin-bottom: {size in rem} !important

Define class .ms-xl-{size} with margin-left: {size in rem} !important

Define class .me-xl-{size} with margin-right: {size in rem} !important

Define class .p-xl-{size} with padding: {size in rem} !important

Define class .pt-xl-{size} with padding-top: {size in rem} !important

Define class .pb-xl-{size} with padding-bottom: {size in rem} !important

Define class .ps-xl-{size} with padding-left: {size in rem} !important

Define class .pe-xl-{size} with padding-right: {size in rem} !important

For each size from 0 to 5:

Define class .px-xl-{size} with padding-right and padding-left: {size in rem} !important

Define class .py-xl-{size} with padding-top and padding-bottom: {size in rem} !important

Define class .text-xl-start with text-align: left !important

Define class .text-xl-end with text-align: right !important

At a minimum width of 768px:

For 'section' elements, set padding to 9rem top and bottom

Define class .btn-xl with specific padding, font-family, font-size, and font-weight

Define class .btn-social with specific height, width, display, align-items, justify-content, padding, and border-radius

Define ID #mainNav with specific padding-top, padding-bottom, and background-color

Define child .navbar-toggler with specific padding, font-size, font-family, text-transform, and font-weight

Define child .navbar-brand with specific color, font-family, font-weight, letter-spacing, and text-transform

Define child img with specific height

Define child .navbar-nav .nav-item .nav-link with specific font-family, font-size, color, and letter-spacing

Define child .navbar-nav .nav-item .nav-link.active and .navbar-nav .nav-item .nav-link:hover with specific color

At a minimum width of 992px:

Modify #mainNav with specific padding-top, padding-bottom, border, and transition

Modify child .navbar-brand with specific font-size and transition

Modify child img with specific height and transition

Define #mainNav.navbar-shrink with specific padding-top, padding-bottom, and background-color

Modify child .navbar-brand with specific font-size

Modify child svg and img with specific height

Define child .navbar-nav .nav-item with specific margin-right

Define last child .navbar-nav .nav-item with specific margin-right

Define header.masthead with specific padding-top, padding-bottom, text-align, and color

Define child .masthead-subheading with specific font-size, font-style, line-height, margin-bottom, and font-family

Define child .masthead-heading with specific font-size, font-weight, line-height, margin-bottom, and font-family

At a minimum width of 768px:

Modify header.masthead with specific padding-top and padding-bottom

Modify child .masthead-subheading with specific font-size, font-style, line-height, and margin-bottom

Modify child .masthead-heading with specific font-size, font-weight, line-height, and margin-bottom

Define .team-member with specific margin-bottom and text-align

Define child img with specific width, border

Define child h4 and .h4 with specific margin-top and margin-bottom

Define .img-brand with specific height

Define class .team-member with specific margin-bottom and text-align

Define child img with specific width and border

Define child h4 and .h4 with specific margin-top and margin-bottom

Define ID #pred with specific box-shadow, padding, and margin

Define ID #msform with specific text-align, position, and margin-top

Define child fieldset with specific border, border-radius, box-shadow, padding, box-sizing, width, margin, and position

Define child fieldset:not(:first-of-type) with specific display

Define child fieldset with specific text-align and color

Define child input and textarea with specific background-color, padding, border, border-radius, margin-bottom, margin-top, width, box-sizing, font-family, color, font-size, and letter-spacing

Define child input:focus and textarea:focus with specific box-shadow, border, font-weight, border-bottom, and outline-width

Define child .action-button with specific width, background, font-weight, color, border, border-radius, cursor, padding, and margin

Define child .action-button:hover and .action-button:focus with specific box-shadow

Define child .action-button-previous with specific width, background, font-weight, color, border, border-radius, cursor, padding, and margin

Define child .action-button-previous:hover and .action-button-previous:focus with specific box-shadow

Define child select.list-dt with specific border, outline, border-bottom, padding, and margin

Define child select.list-dt:focus with specific border-bottom

Define .card with specific z-index, border, border-radius, and position

Define .fs-title with specific font-size, color, margin-bottom, font-weight, and text-align

Define .options with specific position and padding-left

Define child #options label with specific display, margin-bottom, font-size, and cursor

Define child .options input with specific opacity

Define child .checkmark with specific position, top, left, height, width, background-color, border, and border-radius

Define child .options input:checked ~ .checkmark:after with specific display

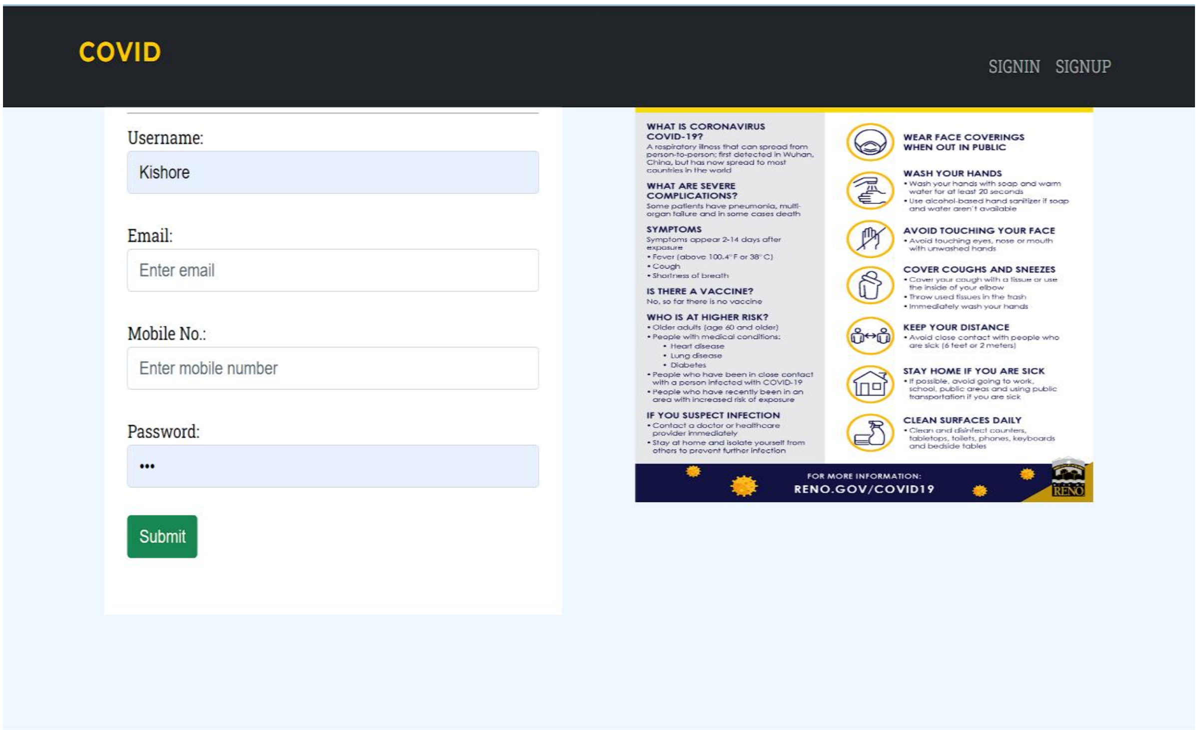
Define child .options .checkmark:after with specific content, width, height, display, background, position, top, left, border-radius, transform, and transition

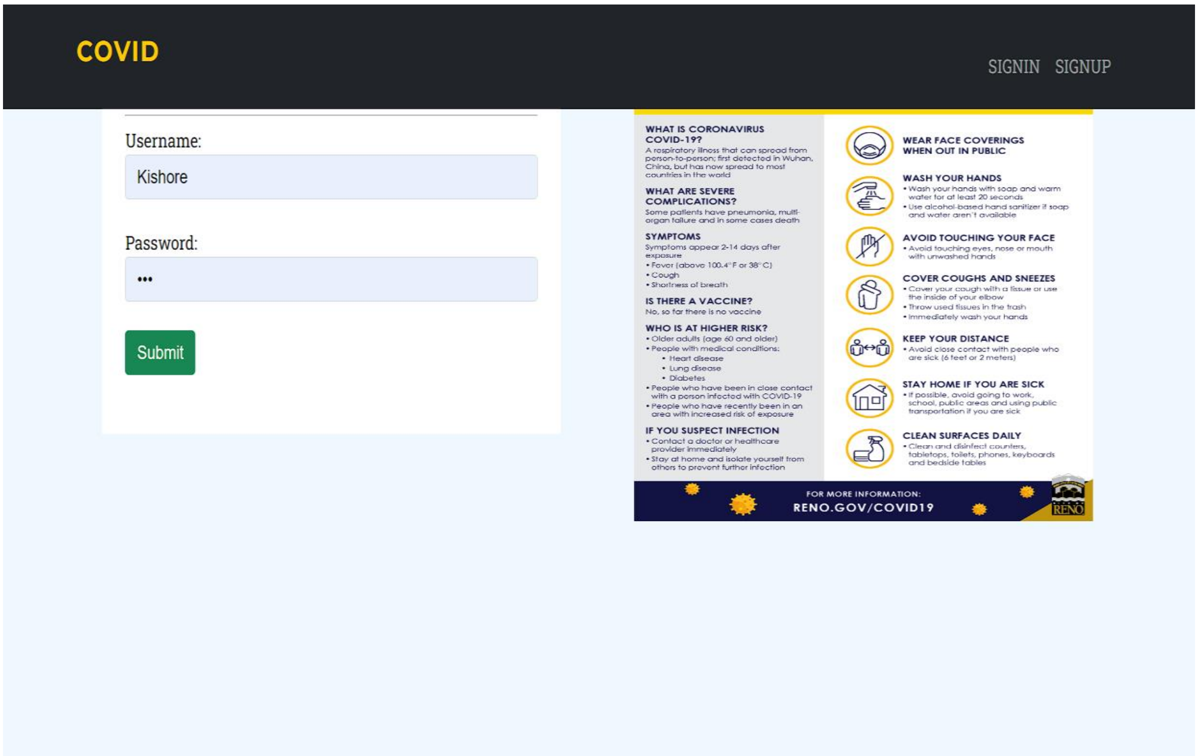
Define child .options input[type="radio"]:checked ~ .checkmark with specific background and transition

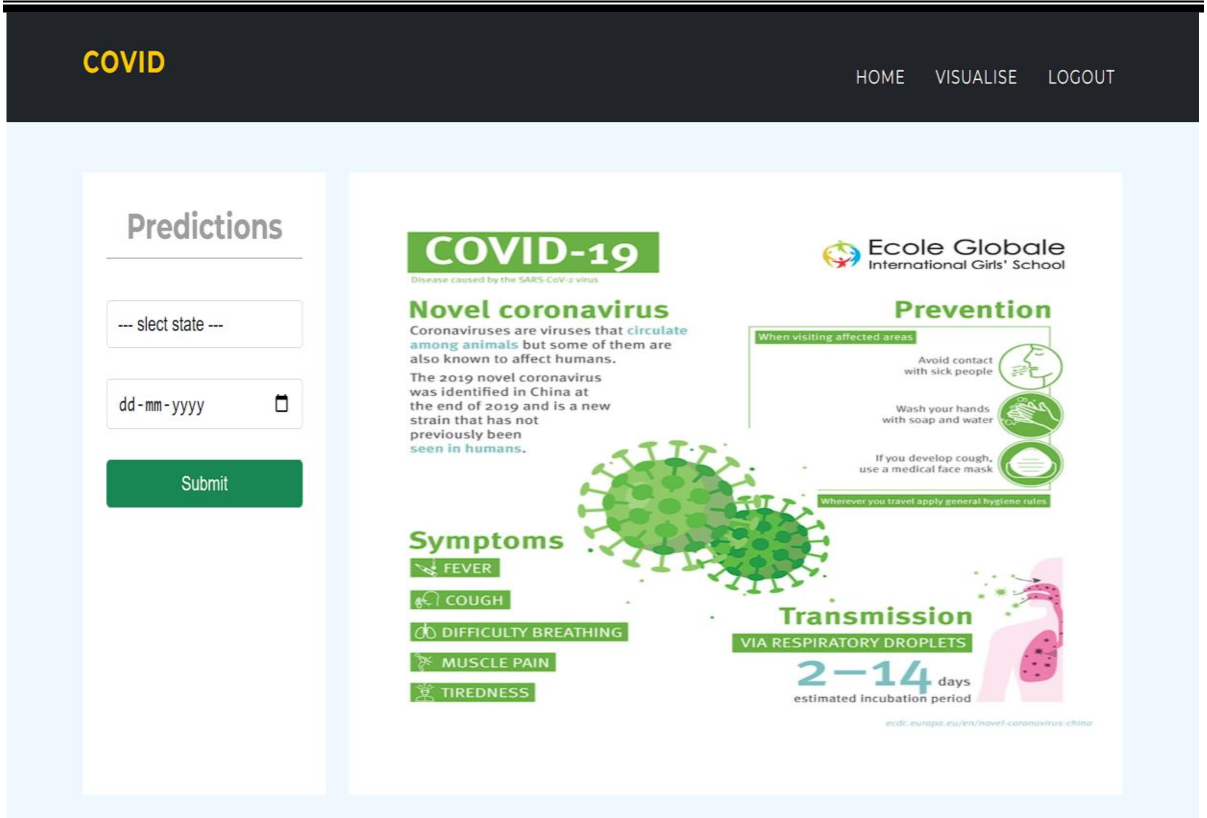
Define child .options input[type="radio"]:checked ~ .checkmark:after with specific transform

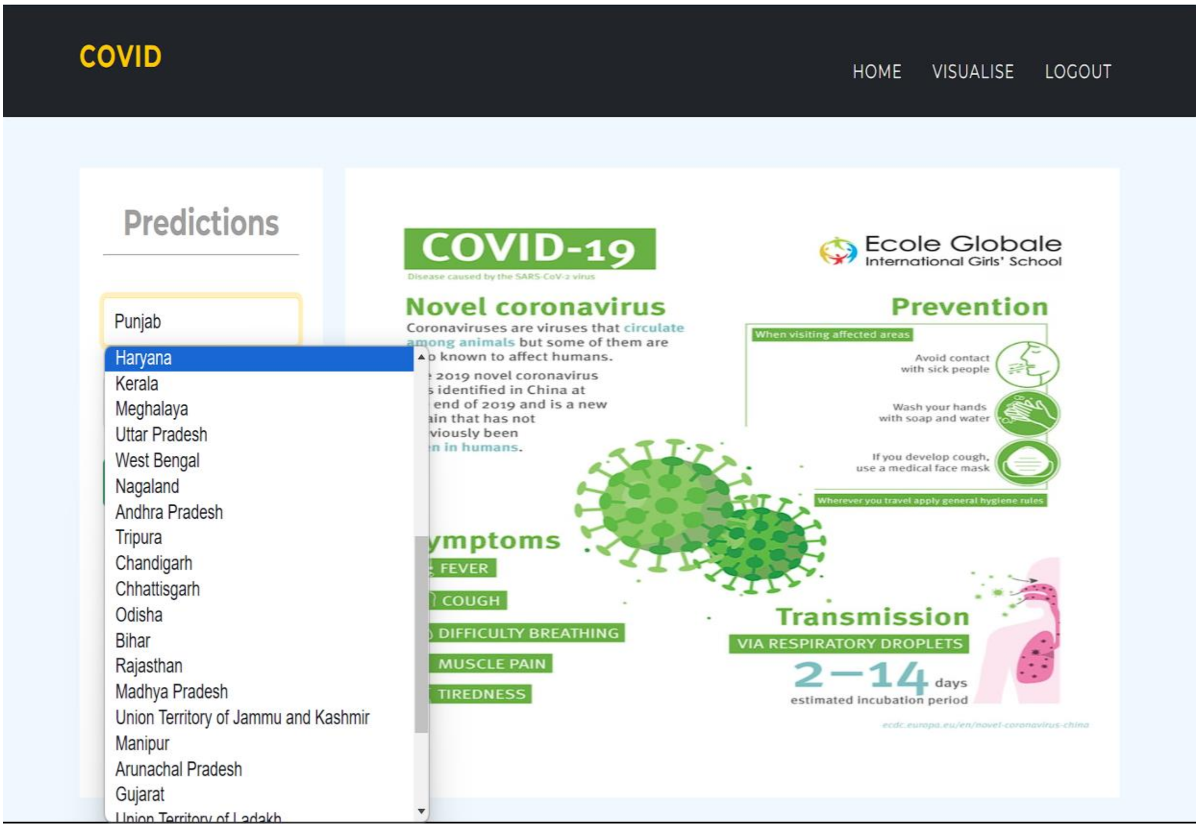
APPENDIX-B

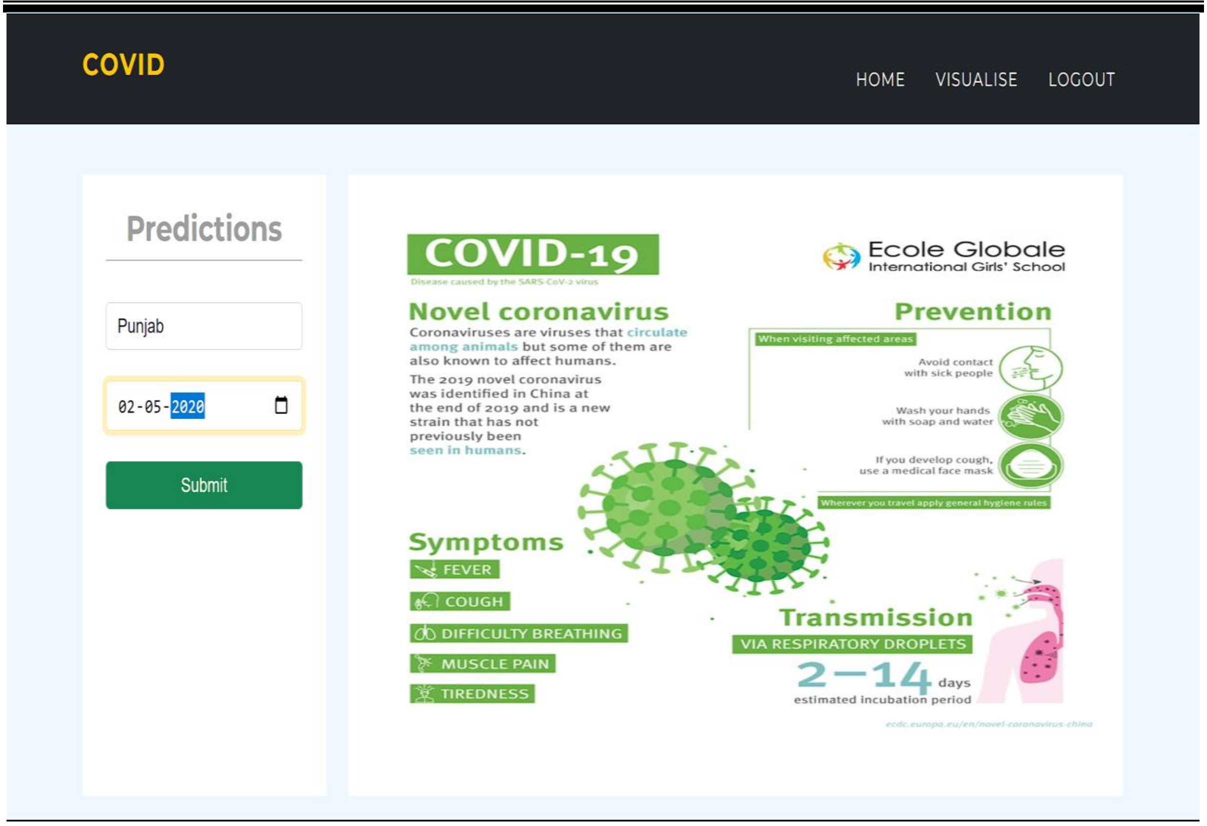
SCREENSHOTS

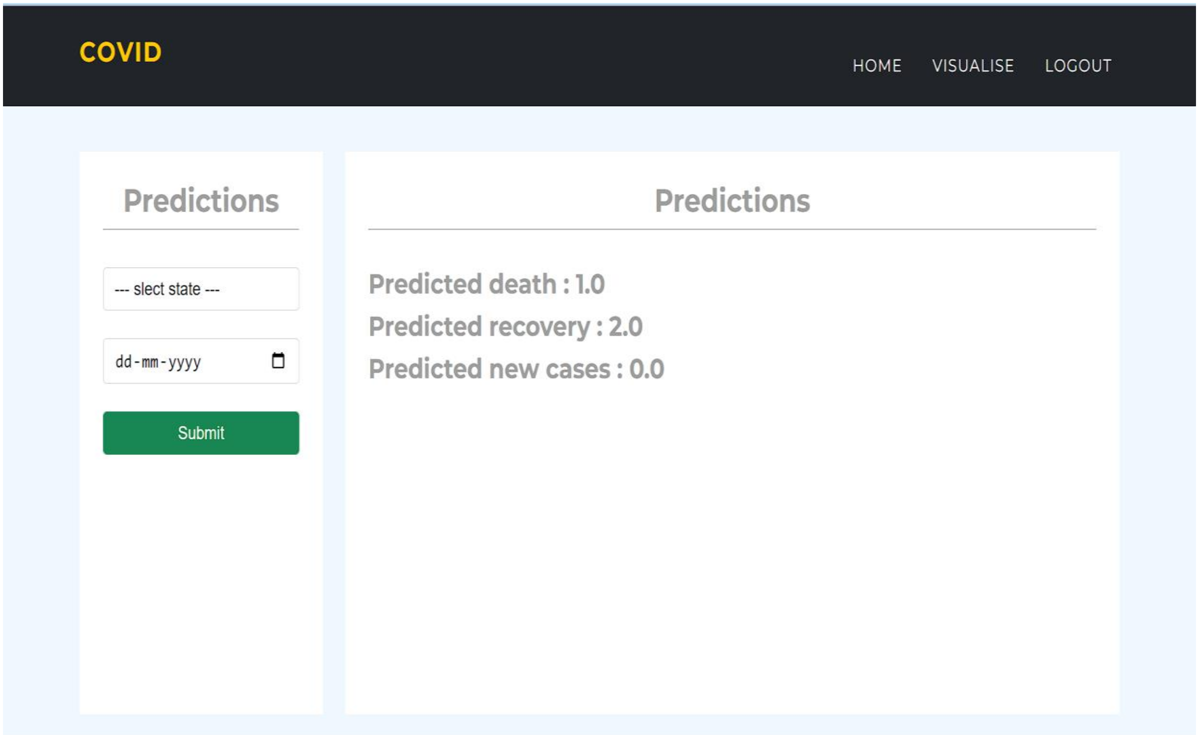


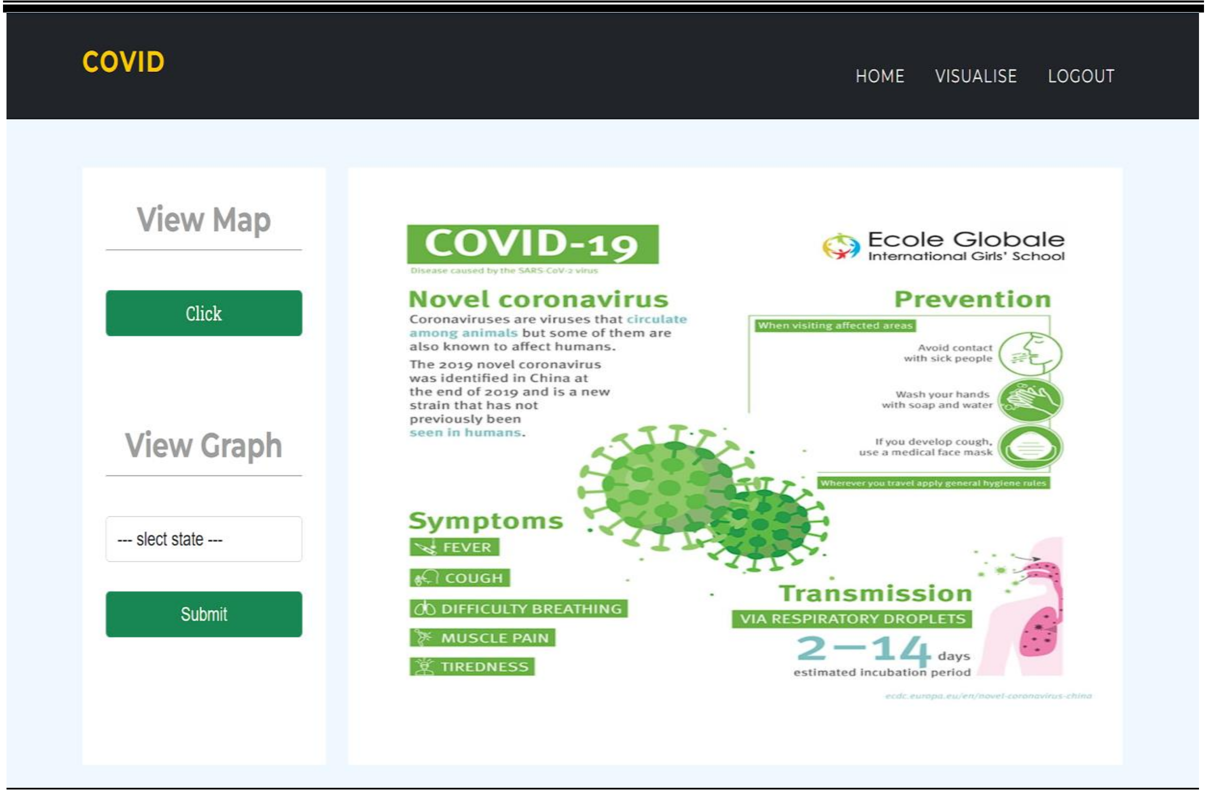


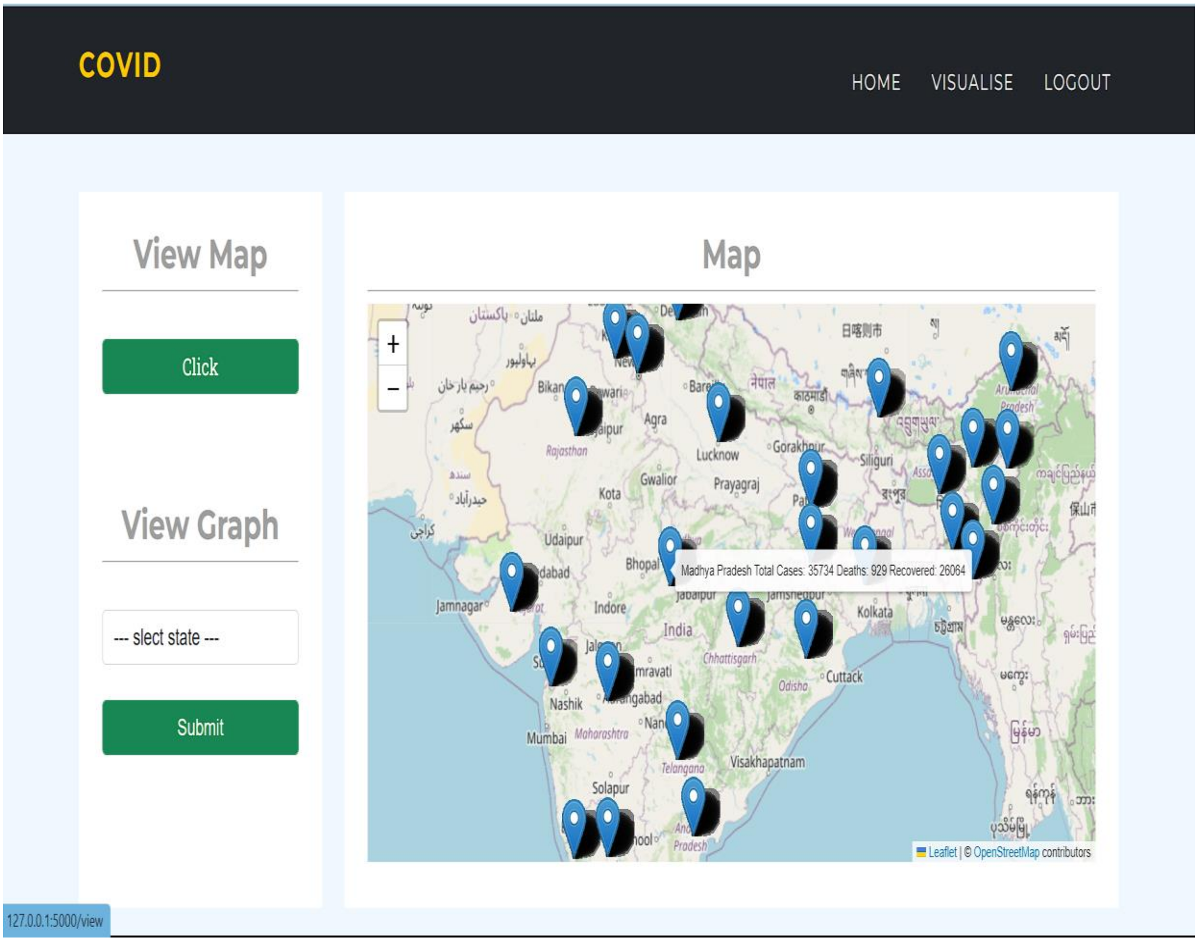

COVID
HOME
VISUALISE
LOGOUT
Predictions
--- slect state ---
dd-mm-yyyy
Submit
COVID-19
Ecole Globale
International Girls' School
Novel coronavirus
Coronaviruses are viruses that circulate among animals but some of them are also known to affect humans.
The 2019 novel coronavirus was identified in China at the end of 2019 and is a new strain that has not previously been seen in humans.
Prevention
When visiting affected areas
Avoid contact with sick people
Wash your hands with soap and water
If you develop cough, use a medical face mask
Wherever you travel apply general hygiene rules
Symptoms
FEVER
COUGH
DIFFICULTY BREATHING
MUSCLE PAIN
TIREDNESS
Transmission
VIA RESPIRATORY DROPLETS
2–14 days
estimated incubation period

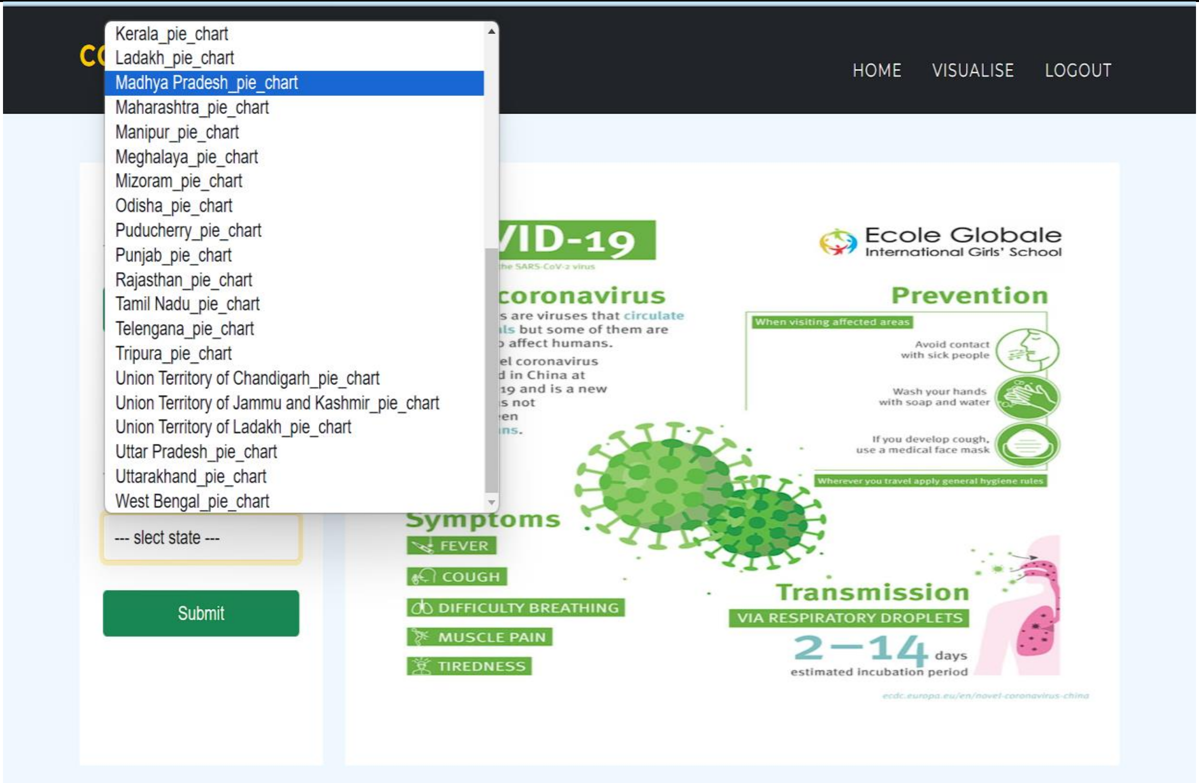

COVID
HOME
VISUALISE
LOGOUT
Predictions
Punjab
Haryana
Kerala
Meghalaya
Uttar Pradesh
West Bengal
Nagaland
Andhra Pradesh
Tripura
Chandigarh
Chhattisgarh
Odisha
Bihar
Rajasthan
Madhya Pradesh
Union Territory of Jammu and Kashmir
Manipur
Arunachal Pradesh
Gujarat
COVID-19
Ecole Globale
International Girls' School
Novel coronavirus
Prevention
When visiting affected areas
Avoid contact with sick people
Wash your hands with soap and water
If you develop cough, use a medical face mask
Wherever you travel apply general hygiene rules
FEVER
COUGH
DIFFICULTY BREATHING
MUSCLE PAIN
TIREDNESS
Transmission
VIA RESPIRATORY DROPLETS
2–14 days
estimated incubation period

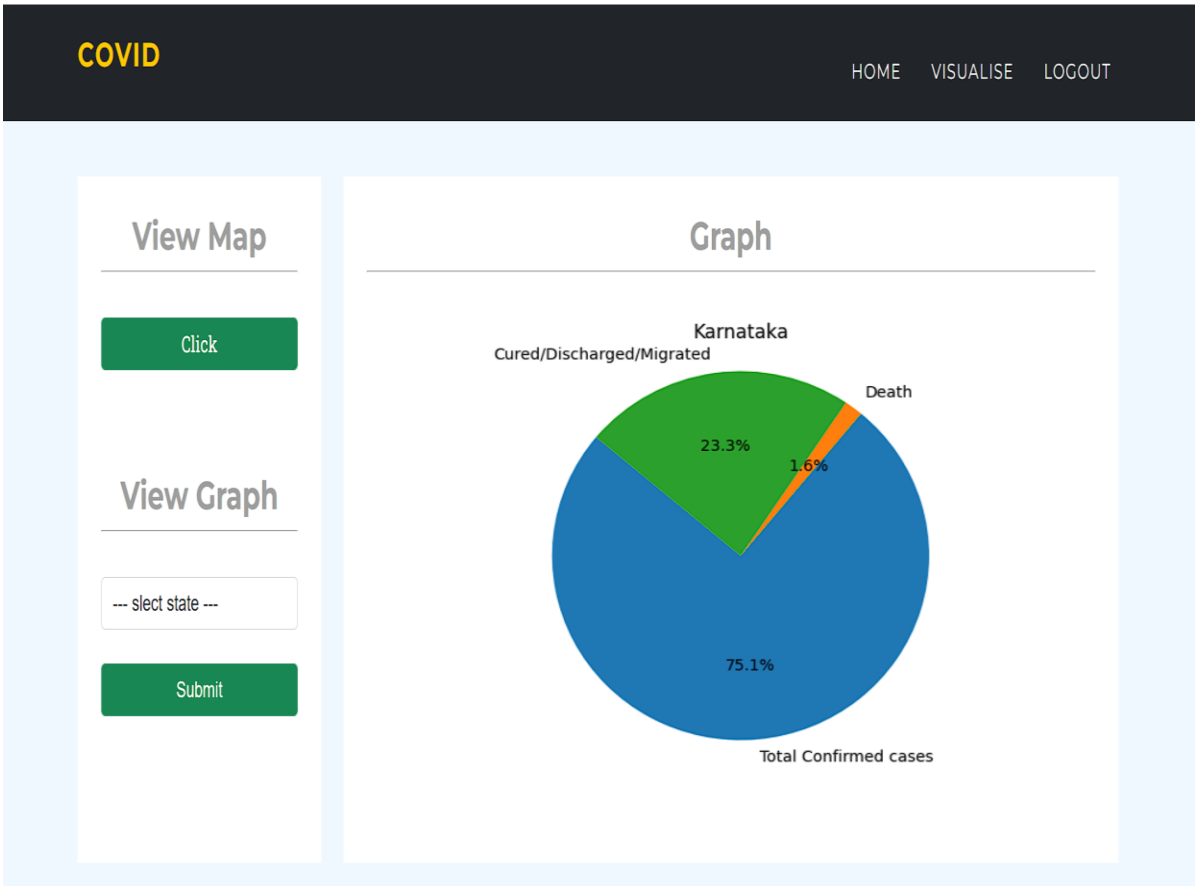

COVID
HOME
VISUALISE
LOGOUT
Predictions
Punjab
02-05-2020
Submit
COVID-19
Ecole Globale
International Girls' School
Novel coronavirus
Coronaviruses are viruses that circulate among animals but some of them are also known to affect humans.
The 2019 novel coronavirus was identified in China at the end of 2019 and is a new strain that has not previously been seen in humans.
Prevention
When visiting affected areas
Avoid contact with sick people
Wash your hands with soap and water
If you develop cough, use a medical face mask
Wherever you travel apply general hygiene rules
Symptoms
FEVER
COUGH
DIFFICULTY BREATHING
MUSCLE PAIN
TIREDNESS
Transmission
VIA RESPIRATORY DROPLETS
2–14 days
estimated incubation period


COVID
HOME
VISUALISE
LOGOUT
Predictions
--- slect state ---
dd-mm-yyyy
Submit
Predictions
Predicted death : 1.0
Predicted recovery : 2.0
Predicted new cases : 0.0


COVID
HOME
VISUALISE
LOGOUT
View Map
Click
View Graph
--- slect state ---
Submit
COVID-19
Ecole Globale
International Girls' School
Novel coronavirus
Coronaviruses are viruses that circulate among animals but some of them are also known to affect humans.
The 2019 novel coronavirus was identified in China at the end of 2019 and is a new strain that has not previously been seen in humans.
Prevention
When visiting affected areas
Avoid contact with sick people
Wash your hands with soap and water
If you develop cough, use a medical face mask
Wherever you travel apply general hygiene rules
Symptoms
FEVER
COUGH
DIFFICULTY BREATHING
MUSCLE PAIN
TIREDNESS
Transmission
VIA RESPIRATORY DROPLETS
2—14 days
estimated incubation period


COVID
HOME
VISUALISE
LOGOUT
View Map
Click
View Graph
--- slect state ---
Submit
Map
+
−
Madhya Pradesh Total Cases: 35734 Deaths: 929 Recovered: 26064
Rajasthan
Agra
Lucknow
Prayagraj
Gwalior
Kota
Udaipur
Bhopal
Indore
Jamnagar
India
Chhattisgarh
Kolkata
Cuttack
Odisha
Nashik
Mumbai
Maharashtra
Telangana
Visakhapatnam
Solapur
Siliguri
Leaflet | © OpenStreetMap contributors
127.0.0.1:5000/view


HOME
VISUALISE
LOGOUT
Kerala_pie_chart
Ladakh_pie_chart
Madhya Pradesh_pie_chart
Maharashtra_pie_chart
Manipur_pie_chart
Meghalaya_pie_chart
Mizoram_pie_chart
Odisha_pie_chart
Puducherry_pie_chart
Punjab_pie_chart
Rajasthan_pie_chart
Tamil Nadu_pie_chart
Telengana_pie_chart
Tripura_pie_chart
Union Territory of Chandigarh_pie_chart
Union Territory of Jammu and Kashmir_pie_chart
Union Territory of Ladakh_pie_chart
Uttar Pradesh_pie_chart
Uttarakhand_pie_chart
West Bengal_pie_chart
--- slect state ---
Submit
Ecole Globale
International Girls' School
Prevention
When visiting affected areas
Avoid contact with sick people
Wash your hands with soap and water
If you develop cough, use a medical face mask
Wherever you travel apply general hygiene rules
Symptoms
FEVER
COUGH
DIFFICULTY BREATHING
MUSCLE PAIN
TIREDNESS
Transmission
VIA RESPIRATORY DROPLETS
2–14 days
estimated incubation period


COVID
HOME
VISUALISE
LOGOUT
View Map
Click
View Graph
--- slect state ---
Submit
Graph
Karnataka
Cured/Discharged/Migrated
Death
23.3%
1.6%
75.1%
Total Confirmed cases

APPENDIX-C

ENCLOSURES

CONFERENCE PAPER

ISSN: 2582-3930

ACCEPTANCE CERTIFICATE

Impact Factor: 8.176
DOI Prefix: 10.55041

INTERNATIONAL JOURNAL OF SCIENTIFIC RESEARCH IN ENGINEERING & MANAGEMENT (IJSREM)

An Open Access Scholarly Journal || Index in major Databases & Metadata

We are pleased to inform you that your Research Paper titled

IMPLEMENTATION OF A REAL-TIME, DATA-DRIVEN ONLINE EPIDEMIC CALCULATOR FOR TRACKING THE SPREAD OF COVID-19

has been ACCEPTED for publication in

INTERNATIONAL JOURNAL OF SCIENTIFIC RESEARCH IN ENGINEERING & MANAGEMENT (IJSREM)

VOLUME 08 ISSUE 01 JANUARY - 2024

We are delighted to see your commitment & hardwork to share your research is being recognized. We look foreward to helping you with all of your publication needs. Thank you for choosing IJSREM!!

Crossref CiteFactor

Editor-in-chief
IJSREM

www.ijsrem.com ijsremjournal@gmail.com

Similarity Index / Plagiarism Check report

G 08

ORIGINALITY REPORT

13%	10%	8%	6%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	link.springer.com Internet Source	1%
2	robot.unipv.it Internet Source	1%
3	cs230.stanford.edu Internet Source	1%
4	Manikandan Rajagopal, BaigMuntajeeb Ali, S.Sharon Priya, W.Aisha Banu, Madhavi G. M, Punamkumar. "Machine Learning Methods for Online Education Case", 2023 Eighth International Conference on Science Technology Engineering and Mathematics (ICONSTEM), 2023 Publication	<1%
5	mdpi-res.com Internet Source	<1%
6	Kamran Etemad, Rama Chellappa. "Discriminant analysis for recognition of human face images", Journal of the Optical Society of America A, 1997 Publication	<1%

7	www.researchgate.net Internet Source	<1%
8	fastercapital.com Internet Source	<1%
9	www2.mdpi.com Internet Source	<1%
10	www.mdpi.com Internet Source	<1%
11	A.K. Jain, A. Ross, S. Prabhakar. "An Introduction to Biometric Recognition", IEEE Transactions on Circuits and Systems for Video Technology, 2004 Publication	<1%
12	futur.upc.edu Internet Source	<1%
13	Nenghuan Zhang, Yongbin Wang, Peng Yu. "A Review of Human Action Recognition in Video", 2018 IEEE/ACIS 17th International Conference on Computer and Information Science (ICIS), 2018 Publication	<1%
14	Submitted to Simpson University Student Paper	<1%
15	journals.sagepub.com Internet Source	<1%

16	Submitted to Southern New Hampshire University - Continuing Education Student Paper	<1 %
17	Rein-Lien Hsu, M. Abdel-Mottaleb, A.K. Jain. "Face detection in color images", IEEE Transactions on Pattern Analysis and Machine Intelligence, 2002 Publication	<1 %
18	dokumen.pub Internet Source	<1 %
19	Submitted to Queen Mary and Westfield College Student Paper	<1 %
20	cbqt.vista.gov.vn Internet Source	<1 %
21	www.scpe.org Internet Source	<1 %
22	www.kluniversity.in Internet Source	<1 %
23	Ying Li, Xiang-lin Qi, Yun-jiu Wang. "Eye detection by using fuzzy template matching and feature-parameter-based judgement", Pattern Recognition Letters, 2001 Publication	<1 %
24	"Intelligent Data Engineering and Automated Learning - IDEAL 2018", Springer Science and	<1 %

SUSTAINABLE DEVELOPMENT GOALS

The project work carried out here is mapped to SDG-3 Good Health and Well-Being.

Implementation of a real-time, data-driven online epidemic calculator for tracking the spread of COVID-19, directly contributes to Goal 3: Good Health and Well-Being. As Goal 3 focuses on ensuring healthy lives and promoting well-being for all at all ages. By delivering accurate predictions and timely precautions, especially in areas with limited healthcare access. Tracking the spread of COVID-19 in real-time allows for a better understanding of the disease's impact, facilitating timely interventions and healthcare planning.