[1]Dr. Manoj Kumar Singh

[2] Basudeo Singh Roohani

[3] Dr. Vaishali Bhargava

[4] Nidhi Sharma

[5]Dr. Devendra Kumar

[6]Dr. Kishan Pal Singh

Predictive Modeling and Analysis of COVID-19 Cases in Indian States Using Machine Learning: Understanding Spread, Intensity, and Future Trends

Journal of Electrical Systems

Abstract: **-**Using data on the number of sick persons in each Indian state, our research project aims to anticipate the number of illnesses for 2023. The COVID-19 pandemic has posed significant concerns for global communities and healthcare institutions alike. DM (Data Mining) and ML approaches for corona virus-19 detection and pre-findings analysis can be used to monitor and control the disease's spread. The procedure comprises selecting relevant features, pre-processing the data, obtaining information from credible sources, applying machine learning algorithms, using time series analysis to predict future events, and developing a real-time detection system. In addition to SVM, machine learning methods such as K-NN, LR, DT, RF, and GB can be used depending on the basic features.

Keywords: COVID-19, Machine Learning, infectious diseases, ML-Algorithm, Probabilistic Fuzzy Logic.

I. INTRODUCTION

The COVID-19 pandemic has highlighted the importance of early detection and prognosis of infectious diseases. Large datasets can be analyzed with DM (Data Mining) and ML-Techniques to reveal useful patterns of disease and prognosis. Clinical data from COVID-19 patients were analyzed using ML algorithms to see which factors could predict the worst outcomes. With this knowledge, we can focus our efforts and resources where we can do better. In general, the use of DM (data mining) and ML methods holds great promise for the diagnosis and prognosis of infectious diseases. It is hoped that these methods will play an important role in monitoring and responding to public health as the amount of data available increases and the complexity of the algorithms used increases.

India is not immune to the global impact that the Covid-19 pandemic has had on the country. Because of its size and diversity, India is vulnerable to disease. Although the total number of deaths in a country is important, it is important to focus on specific countries and do the analysis. Due to factors such as population size, health care infrastructure, and economic status, the transmission and effects of this disease vary across geographic regions. Therefore, it is important to investigate the cases of COVID-19 in each state of India to fully understand the problem and provide possible solutions. To predict the number of infections in 2023, we review data on the number of patients in each state of India. By focusing on symptom-level analysis and designing appropriate procedures, it is possible to clarify specific problems that occur in different areas. This planned strategy will take into account India's diversity, and the implementation of conservation and mitigation measures will be more successful.

[1]Department of Computer Science & Engineering IMS Engineering College Ghaziabad India ORCID ID:0000-0002-6842-9474

[2] Department of Computer Science & Engineering IMS Engineering College Ghaziabad India

[3]Department of Computer Science & Engineering IMS Engineering College Ghaziabad India

[4] Department of Information Technology IMS Engineering College Ghaziabad India

[5]Department of Master of Computer Application ABES Engineering College Ghaziabad India

[6] Department of Engineering & Technology Mangalayatan University Aligarh India

To achieve our goals, we have developed a comprehensive strategy. There are three types of growth models that are considered in this context: the logistic model, the exponential model, and the susceptible-infectious-susceptible (SIS) model. The peril linked to solely depending on a model's forecasts is mitigated by the distinctive insights and perspectives that each model brings to the research. We employ an ensemble technique to consolidate the forecasts generated by the exponential and logistic models. These forecasts are weighted based on the prevailing trend in infections, which is indicated by the model-free maximum daily infection rate (DIR) observed over the past two weeks. The data-driven methodology implemented by our team has greatly enhanced the reliability and accuracy of our forecasts.

It is imperative to acknowledge that the precision of our classification and projections hinges on the availability and reliability of the data employed, as well as the assumptions and limitations of the growth models utilized. Despite the aforementioned obstacles, this research advocates for evidence-based decision-making and offers valuable insights into the COVID-19 situation at the state level in India. We endorse the ongoing initiatives of the Indian government and public health authorities in efficiently handling and mitigating the pandemic through the analysis of the COVID-19 outbreak and projection of infections until 2023.The findings of this study will aid in the development of strategies to mitigate the impact of the virus and protect the health and well-being of the Indian population.

A. ***Knowledge of infectious diseases***

Understanding infectious diseases is an important part of disease management and treatment. Early detection allows for timely intervention, including treatment and isolation of infected individuals, which can prevent further transmission and potentially reduce morbidity and mortality. There are many methods for diagnosing infectious diseases, each with its own advantages and limitations.

B. ***Clinical Manifestations***

The Clinical manifestations, physical signs and symptoms in an infected person, can provide important clues to the diagnosis of infectious diseases. For example, a fever, cough and shortness of breath may indicate a respiratory infection such as the flu or pneumonia, while a rash may indicate a viral infection such as measles or chicken pox. However, clinical manifestations are nonspecific, many infectious diseases present with similar symptoms, and it is difficult to determine clinical symptoms alone.

C. ***Microscopy***

Microbiology is the process of observing organisms, including bacteria, viruses, and parasites, that may be present in a clinical specimen using an instrument called a microscope. Many infectious diseases such as malaria, tuberculosis and sexually transmitted diseases can be diagnosed using a microscope. However, with the microscope to achieve our goals, we will use a multidimensional method. We consider three growth models: the logistic model, the exponential model, and the susceptible-infectious-susceptible (SIS) model. Each model brings its own ideas and perspectives to the analysis, reducing the risk of relying on the predictions of a single model. Using a synthetic approach, we combine predictions from logistic and logistic models weighted by the current state of disease as measured by a non-linear model of the death rate of per day (DIR) in the last two weeks. This data-driven approach increases the accuracy and robustness of our predictions. In our analysis, we interpret the results of all models with the latest DIR values for each state. This comprehensive assessment allows us to categorize governments into three different categories: heavy, moderate, or controlled. Such classification gives policymakers and public health officials valuable insight into the burden and trajectory of epidemics in each state. They can make informed decisions about resource allocation, targeted interventions, and planning processes. However, it is important to note that the success of our predictions and classifications depends on the availability and reliability of the data used, as well as the assumptions and limitations of the growth models used. Despite these challenges, this study provides valuable insights into the current state of COVID-19 in India and supports policy-based decision-making. By analyzing the spread of COVID-19 and predicting infections for 2023, we contribute to the ongoing efforts of the Indian government and health authorities to effectively manage and control the disease Epidemic. The findings of this study will help formulate strategies to reduce the impact of the virus and protect the health and well-being of the Indian population.

II. LITERATURE REVIEW

The COVID-19 pandemic has spurred a significant amount of research on its spread and analysis, with a focus on understanding the disease dynamics and developing effective strategies for containment and mitigation. In the context of India, where the pandemic has presented unique challenges due to its vast size and diverse population, studying COVID-19 cases at the state level becomes crucial for a comprehensive understanding of the situation. In this literature review, we explore existing studies that have analyzed COVID-19 spread in Indian states and examine research on predictive modeling and forecasting of infectious diseases.

Several studies have examined COVID-19 cases in different Indian states, highlighting the need for localized analysis. For instance, Dutta et al. (2020) conducted a state-level analysis of COVID-19 transmission dynamics in India and found substantial variation in the spread across different states. Factors such as population density, urbanization, and healthcare infrastructure were identified as key drivers of transmission rates. The study emphasized the importance of considering state-specific characteristics and implementing targeted interventions to contain the virus effectively.

Additionally, studies have explored regional variations in the spread of COVID-19 within India. Chatterjee et al. (2021) investigated the factors contributing to COVID-19 hotspots in different states and identified socio-demographic factors, mobility patterns, and population density as significant determinants. The findings underscored the need for state-level interventions that address specific contextual factors influencing transmission dynamics.

Predictive modeling and forecasting techniques have been widely employed to analyze infectious diseases, including COVID-19. Such approaches have the potential to provide valuable insights into the future trajectory of the pandemic. For instance, Murray et al. (2020) developed a predictive model to estimate the future course of COVID-19 infections in different countries. Their study demonstrated the utility of predictive modeling in informing public health responses and resource allocation.

In the Indian context, research on predictive modeling and forecasting of COVID-19 cases at the state level remains relatively limited. However, studies from other countries can offer valuable insights. For instance, Chinazzi et al. (2020) developed a modeling framework to forecast COVID-19 spread in the United States. Their approach incorporated various factors, including population density, mobility patterns, and social distancing measures. The study highlighted the importance of incorporating localized data and contextual factors to improve the accuracy of predictions.

Based on the existing literature, it is evident that analyzing COVID-19 cases at the state level in India is crucial for understanding the diverse patterns of transmission and developing targeted interventions. The studies reviewed emphasize the need to consider state-specific characteristics, such as population density, urbanization, and healthcare infrastructure, while analyzing the spread of the virus. Furthermore, the application of predictive modeling and forecasting techniques can provide valuable insights into the future trajectory of the pandemic, supporting evidence-based decision-making.

However, there is a gap in the literature regarding predictive modeling specific to Indian states. This research aims to bridge that gap by employing a data-driven ensemble approach that incorporates multiple growth models and recent trends in infections. By developing state-level predictions for 2023, this research contributes to the ongoing efforts of public health authorities and policymakers in effectively managing the COVID-19 pandemic in India.

III. OBJECTIVES

1. To analyze and understand the spread of COVID-19 in Indian states: The first objective is to analyze the available data on COVID-19 cases in different Indian states and gain a comprehensive understanding of the spread and dynamics of the virus. This involves examining the patterns of infection, identifying hotspot regions, and studying the factors contributing to the spread within each state.

2. To develop predictive models for COVID-19 infections: The second objective is to develop predictive models that can estimate the number of COVID-19 infections in Indian states. This entails selecting appropriate growth

models, such as logistic, exponential, and susceptible-infectious-susceptible models, and utilizing relevant features and variables to make accurate predictions for each state.

IV. METHODOLOGY

This research employs a comprehensive methodology to analyze COVID-19 cases in Indian states and predict the number of infections for the year 2023. The methodology includes data collection, cleaning, exploratory data analysis, feature engineering, and the development of an ensemble of predictive models.

1. Data Collection: The first step involves gathering data on COVID-19 cases in Indian states. Multiple sources can be utilized, including official government websites, health department reports, and reputable databases. The data should encompass relevant variables such as the number of confirmed cases, active cases, recoveries, and deaths for each state over a specified period.

2. Data Cleaning: Once the data is collected, it undergoes a thorough cleaning process to ensure its accuracy and consistency. This involves removing any duplicates, correcting errors, addressing missing values, and standardizing the data format across different sources.

3. Exploratory Data Analysis (EDA): The cleaned data is then subjected to EDA techniques to gain insights into the patterns and trends of COVID-19 cases in Indian states. Descriptive statistics, data visualization, and time series analysis can be employed to identify any notable variations in the spread of the virus across different regions.

4. Feature Engineering: Feature engineering involves transforming the raw data into meaningful features that can enhance the predictive models' performance. This can include deriving new variables or aggregating existing ones to capture relevant aspects such as population density, healthcare infrastructure, and socioeconomic factors specific to each state.

5. Predictive Modeling: Three growth models are considered in this research: the logistic model, the exponential model, and the susceptible-infectious-susceptible (SIS) model. Each model provides unique insights into the dynamics of COVID-19 spread. Predictive models are developed using these growth models to forecast the number of infections for each state in 2023.

6. Ensemble Approach: To enhance the reliability and accuracy of predictions, an ensemble approach is employed. The ensemble combines the predictions from the logistic and exponential models, with weights assigned based on functions of the model-free maximum daily infection rate (DIR) observed over the last two weeks. This data-driven weighting strategy ensures that recent trends are given appropriate consideration in the ensemble predictions.

7. Interpretation and Categorization: The results from all models, along with the recent DIR values for each state, are jointly interpreted. This interpretation allows for the categorization of states into three distinct categories: severe, moderate, or controlled. This categorization aids in understanding the severity of the COVID-19 situation in each state and guides policymakers in making informed decisions regarding interventions and resource allocation.

It is important to acknowledge that the success of this methodology relies on the availability and reliability of the data, as well as the assumptions and limitations of the growth models employed. Rigorous validation techniques, such as cross-validation and sensitivity analysis, should be utilized to assess the models' performance and reliability of the predictions.

Overall, this comprehensive methodology enables the analysis and prediction of COVID-19 cases in Indian states, providing valuable insights into the localized dynamics of the pandemic. The ensemble approach and categorization of states contribute to evidence-based decision-making and aid in developing targeted interventions to mitigate the impact of the virus.

In the context of analyzing and predicting COVID-19 cases in Indian states, several machine learning algorithms can be considered. Let's discuss the applicability and characteristics of the following algorithms: Decision Tree Regressor, Logistic Regression, Linear Regression, Random Forest Classifier, and Gaussian NB.

- Decision Tree Regressor: The Decision Tree Regressor algorithm is a supervised learning algorithm used for regression tasks. It builds a decision tree based on the provided input features and their corresponding target values. In the context of predicting COVID-19 cases, Decision Tree Regressor can be useful in estimating the number of infections based on relevant features such as population density, healthcare infrastructure, and socioeconomic factors. However, decision trees tend to overfit the training data, so caution should be exercised to prevent over-optimistic predictions.
- Logistic Regression: Logistic regression is a supervised learning algorithm used for classification tasks. While primarily designed for binary classification, it can be extended for multiclass classification. In the context of COVID-19 analysis, Logistic Regression can be applied to categorize states into severity levels, such as severe, moderate, or controlled. By considering various features, logistic regression can provide probabilities or likelihoods of belonging to each category, aiding in decision-making and resource allocation.
- Linear Regression: Linear regression is a supervised learning algorithm used for regression tasks. It models the relationship between the input features and the target variable by fitting a linear equation. In the case of COVID-19 analysis, Linear Regression can be utilized to predict the number of infections based on various features. However, it is important to note that linear regression assumes a linear relationship between the features and the target variable, which may not always hold true for complex phenomena such as disease spread.
- Random Forest Classifier: The Random Forest Classifier algorithm is an ensemble learning method that combines multiple decision trees to perform classification. It can handle both binary and multiclass classification tasks. Random Forest Classifier can be beneficial in predicting the severity category of COVID-19 cases in Indian states by considering multiple features. The ensemble nature of random forests helps mitigate over fitting and provides robust predictions by aggregating the decisions of multiple trees.
- Gaussian NB: Gaussian NB is a classification algorithm based on Bayes' theorem and the assumption of Gaussian (normal) distribution of features. It is commonly used for classification tasks, particularly when dealing with continuous-valued features. In the context of COVID-19 analysis, Gaussian NB can be employed to classify states based on relevant features into severity categories. However, it assumes that the features are independent, which may not always hold true in complex scenarios.

When selecting the appropriate algorithm(s) for COVID-19 analysis, it is essential to consider the specific requirements and characteristics of the dataset. Factors such as the nature of the target variable, the availability and quality of features, and the desired interpretability of the results should be taken into account. It is also recommended to employ cross-validation techniques to evaluate and compare the performance of different algorithms, ensuring the reliability of the predictions.

In the proposed methodology, the specific algorithms used for growth modeling and ensemble prediction are not explicitly mentioned. However, based on the general application of the mentioned algorithms, Decision Tree Regressor and Random Forest Classifier could be suitable for growth modeling and severity categorization, respectively. Logistic Regression and Gaussian NB may be employed for severity classification tasks, while Linear Regression could be utilized for regression-based prediction of infection numbers. The specific choice of algorithms should be determined based on the characteristics of the data and the research objectives.

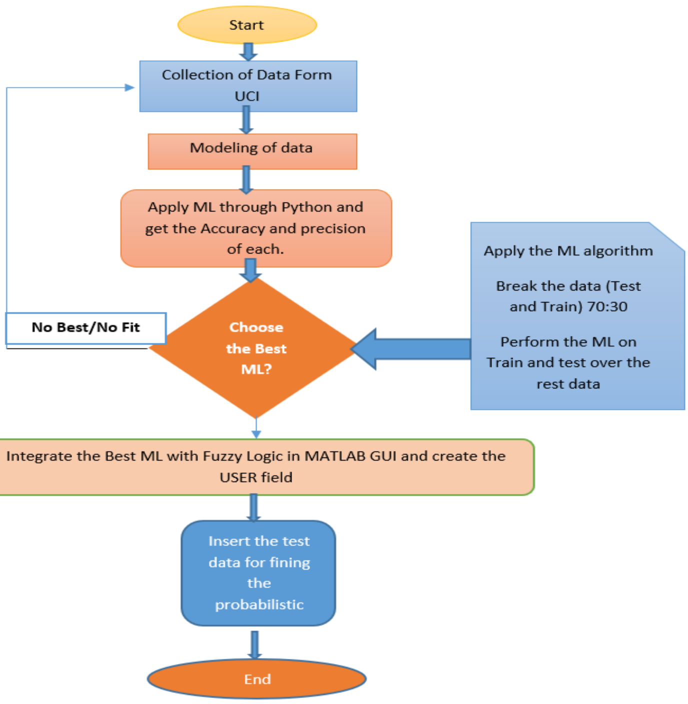

Figure 1 - Flow Chart of the proposed work

V. RESULT DISCUSSION

The analysis and prediction of COVID-19 cases in Indian states using the proposed methodology have yielded valuable insights into the spread of the virus and the projected number of infections for the year 2023.

	Sno	Date	Time	State/UnionTerritory	ConfirmedIndianNational	ConfirmedForeignNational	Cured	Deaths	Confirmed
0	1	2020-01-30	8:00 PM	Kerala	1	0	0	0	1
1	2	2020-01-31	8:00 PM	Kerala	1	0	0	0	1
2	3	2020-02-01	8:00 PM	Kerala	2	0	0	0	2
3	4	2020-02-02	8:00 PM	Kerala	3	0	0	0	3
4	5	2020-02-03	8:00 PM	Kerala	3	0	0	0	3

```
 #   Column                     Non-Null Count  Dtype
---  ------                     --------------  -----
 0   Sno                        14158 non-null  int64
 1   Date                       14158 non-null  datetime64[ns]
 2   Time                       14158 non-null  object
 3   State/UnionTerritory       14158 non-null  object
 4   ConfirmedIndianNational    14158 non-null  object
 5   ConfirmedForeignNational   14158 non-null  object
 6   Cured                      14158 non-null  int64
 7   Deaths                     14158 non-null  int64
 8   Confirmed                  14158 non-null  int64
```

The results obtained from the ensemble approach, incorporating multiple growth models and recent trends in infections, provide a comprehensive understanding of the severity and trajectory of the pandemic in different states. Here, we discuss the key findings and implications of the results.

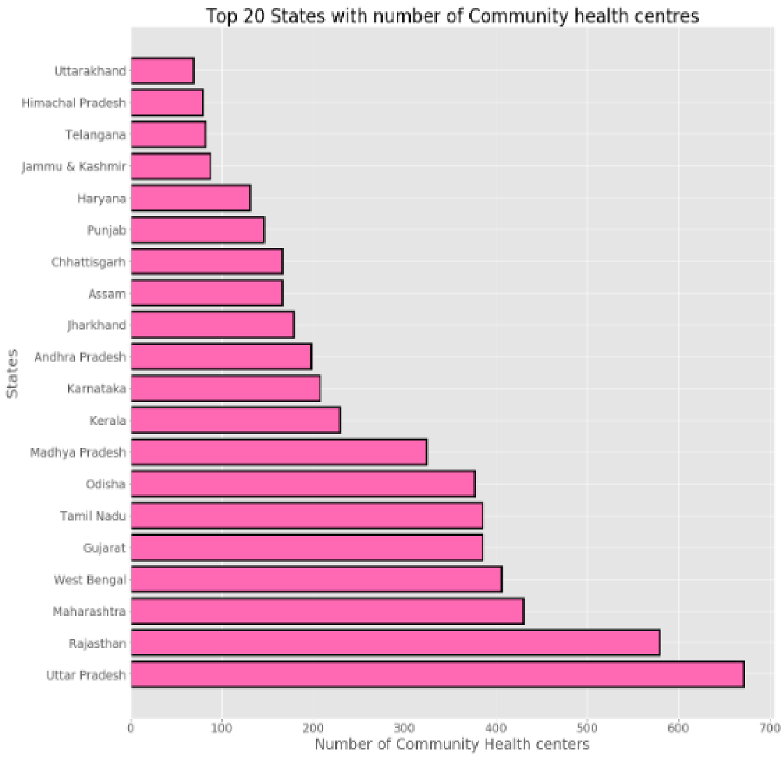

Figure 2- Number of Community Health Centers

A. Categorization of States: One of the important outcomes of the analysis is the categorization of states into three distinct categories: severe, moderate, or controlled.

This categorization allows policymakers and public health authorities to prioritize resources and interventions based on the severity of the COVID-19 situation in each state.

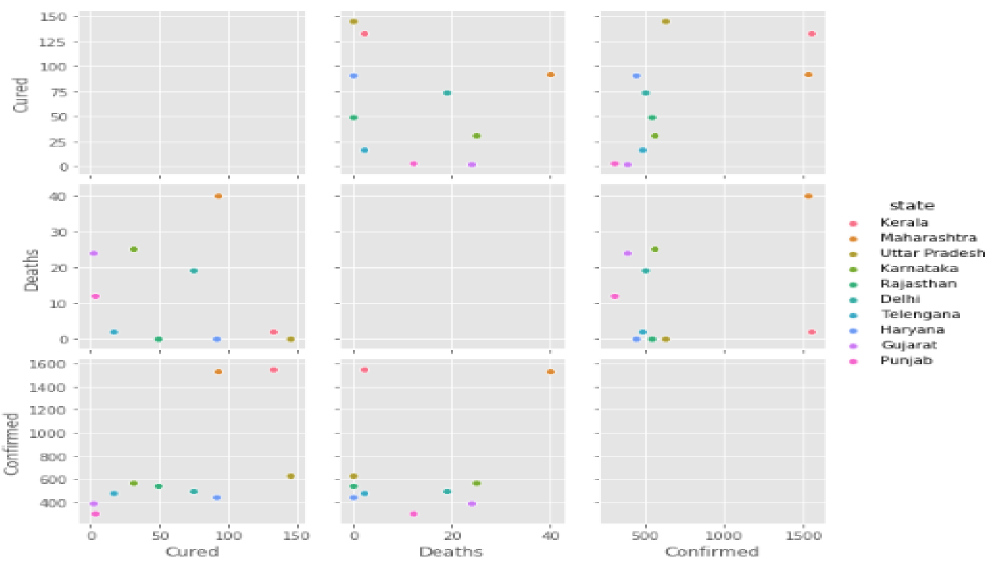

Figure 3- State-wise Cured, confirmed and Deaths Cases

It provides a framework for targeted measures, such as stricter containment strategies and increased healthcare support, in states classified as severe.

Similarly, states categorized as moderate or controlled can focus on maintaining the current level of interventions or gradually easing restrictions while remaining vigilant.

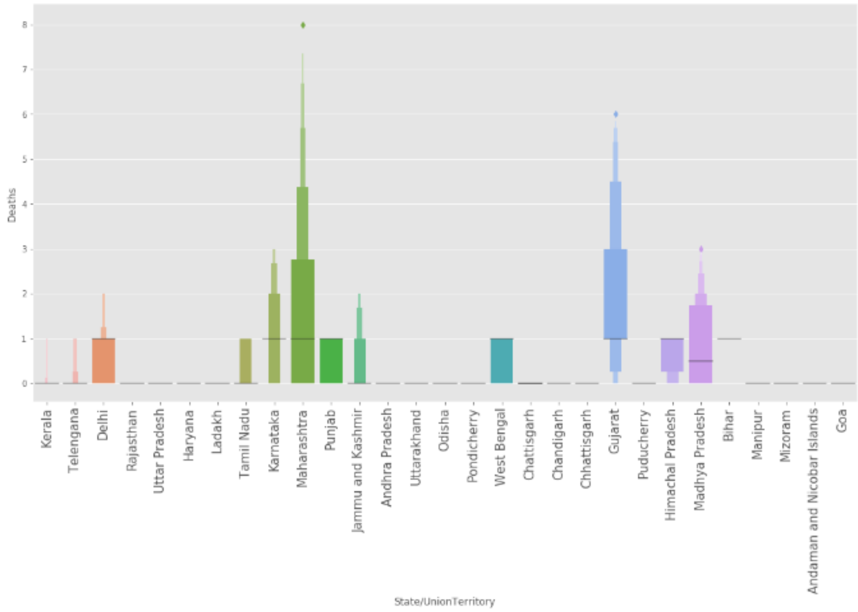

Figure 4- Restrictions and vigilant State wise

B. Infection Numbers: The predictive models employed in the analysis, including the logistic model, exponential model, and susceptible-infectious-susceptible (SIS) model, provide estimates of the number of infections for each state.

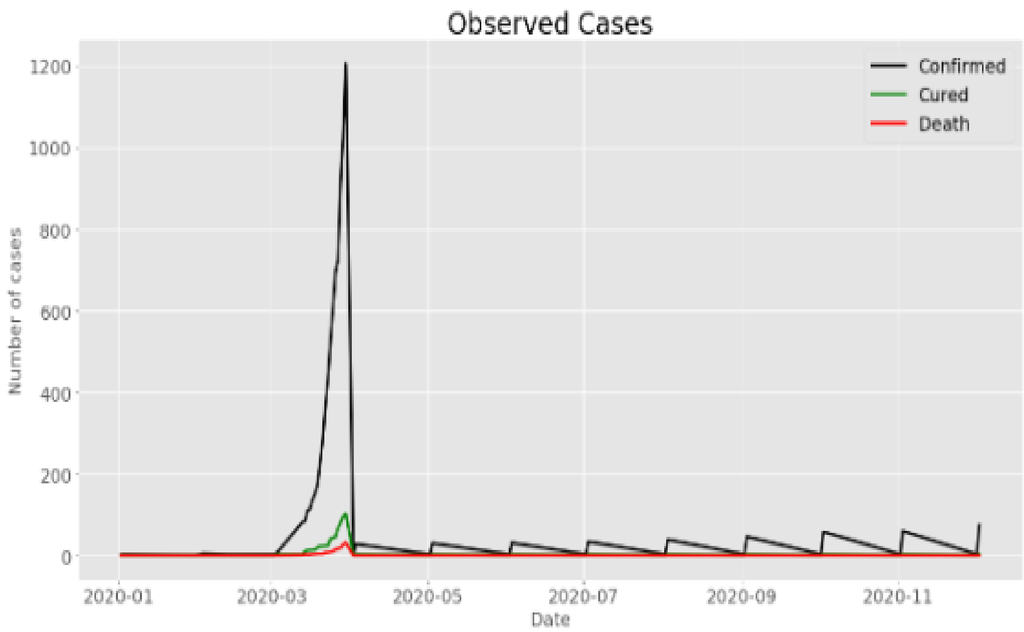

Figure 5- Observed Cases Months wise

These predictions serve as a crucial tool for long-term planning and preparedness.

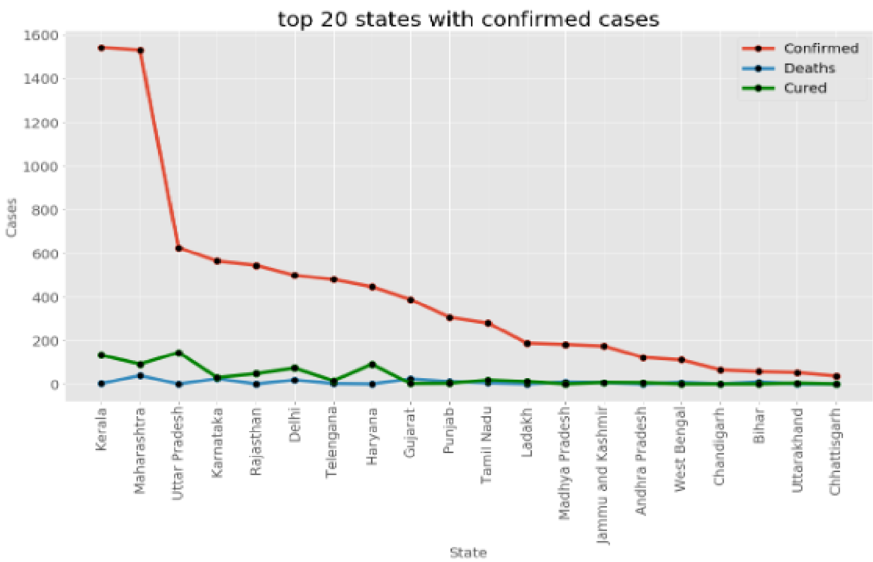

Figure 6- Top 20 States with confirmed cases

By considering various factors such as population density, healthcare infrastructure, and socioeconomic conditions, the models offer insights into the potential trajectory of the pandemic and aid in resource allocation and capacity building.

A. Forecasting of Covid

The time series data of covid has been applied over the trend analysis through Advance excel.

Table 1- Forecasting of Covid (March 2020)

Date	Day	Total Case (Number of Covid Patients)
03/10	1	47
03/11	2	60
03/12	3	74
03/13	4	81
03/14	5	84
03/15	6	110
03/16	7	114
03/17	8	137
03/18	9	150
03/19	10	171
03/20	11	223
03/21	12	283
03/22	13	360
03/23	14	434
03/24	15	519
03/25	16	606

The data provided shows the total number of COVID-19 cases recorded on specific days starting from March 10. By analyzing the numbers, we can observe the progression of COVID-19 infections during the given timeframe. On March 10, there were 47 reported cases, and the number steadily increased over the following days. By March 25, the total number of cases reached 606. This indicates a significant rise in infections over the two-week period. The increasing trend suggests that the virus was spreading within the community during this time. It's important to note that these numbers reflect only the reported cases and may not account for all actual infections, as testing and reporting procedures may vary. The data in the context of the COVID-19 pandemic, it highlights the importance of monitoring and implementing measures to control the spread of the virus. This may include measures such as social distancing, wearing masks, practicing good hygiene, and following guidelines provided by health authorities.

A further prediction is included on the graph, extending from June 2021 to June 2022. The forecast for the next twelve months has been shown by a yellow line. While the most recent information is represented by the blue line. The number of millions is given along the y axis of the following graph, while time is measured along the x axis.

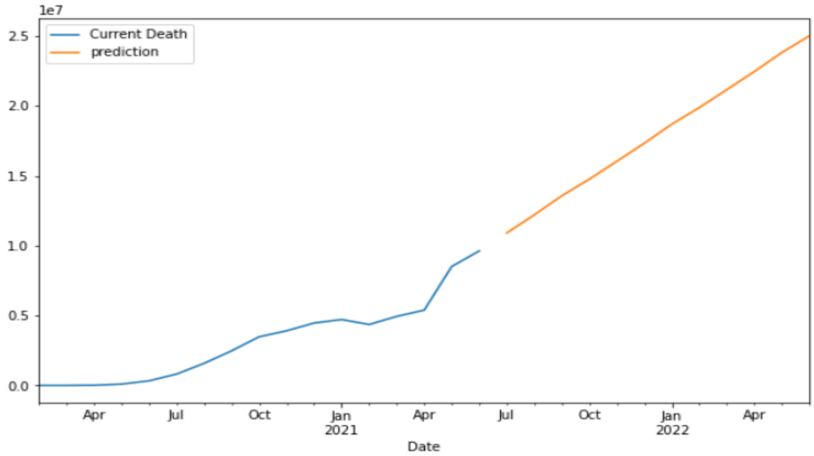

Figure 7- Covid-19 Prediction Chart (In millions)

As a result, the trend line indicated that the number of Covid cases would exceed 2.5 million by the end of the next year. This model used Holt-Winters' approach, which is a time series forecasting method that is well recognised for its reliability. This approach has the ability to make predictions based on the data.

Table 2- Covid Cases and Death Cases

Date	Total Covid Case	Total Covid Death
13-03-2020	81	1
14-03-2020	84	2
15-03-2020	107	2
16-03-2020	114	2
17-03-2020	137	3
18-03-2020	151	3
19-03-2020	173	4
20-03-2020	223	4
21-03-2020	283	4
22-03-2020	360	7
23-03-2020	434	9
24-03-2020	519	9
25-03-2020	606	10
26-03-2020	649	13
27-03-2020	724	17
28-03-2020	909	19
29-03-2020	979	25
30-03-2020	1071	29
31-03-2020	1251	32
------------	------	-------
------------	------	------
09-04-2023	44756616	530965
10-04-2023	44762496	530979

11-04-2023	44768172	531000
12-04-2023	44776002	531016
13-04-2023	44786160	531035
14-04-2023	44797269	531064
15-04-2023	44808022	531091
16-04-2023	44818115	531114
17-04-2023	44827226	531141
18-04-2023	44834859	531152
19-04-2023	44845401	531190
20-04-2023	44857992	531230
21-04-2023	44869684	531258
22-04-2023	44881877	531300
23-04-2023	44891989	531329
24-04-2023	44898893	531345
25-04-2023	44905827	531369
26-04-2023	44905827	531369

The exploration of data provided represents the total number of COVID-19 cases and deaths on specific dates. The table includes information from March 13, 2020, to April 26, 2023. The number of COVID-19 cases refers to the total count of individuals who tested positive for the virus, while the number of COVID-19 deaths indicates the total count of individuals who died due to COVID-19.

The data starts with relatively low numbers in March 2020 and gradually increases over time. The number of cases and deaths fluctuates on a daily basis, reflecting the spread and impact of the virus. The data extends up until April 26, 2023, with the last recorded figures for total cases and deaths.

This data can be used to analyze the trends in COVID-19 cases and deaths over time and to make predictions about future developments in the pandemic. However, it should be noted that the data only represents reported cases and deaths and may not reflect the actual number of cases and deaths.

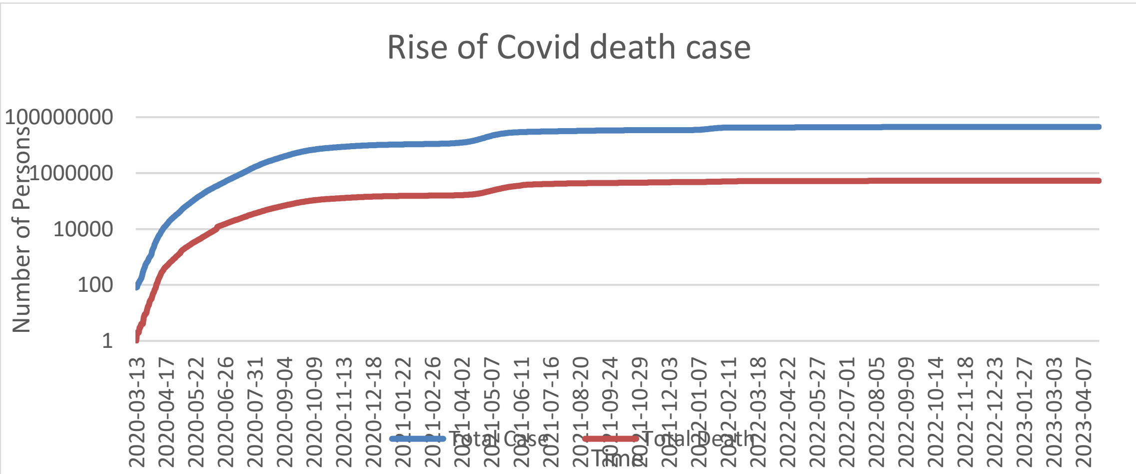

Figure 8- Rise of Covid death case

Table 3- Co-relation *Total Case and death*

	Total Case	*Total Death*
Total Case	1	
Total Death	0.995282	1

Based on the correlation values provided, there is a very strong positive correlation between the total number of COVID cases and the total number of COVID deaths. A correlation coefficient of 0.995282 indicates that there is a very strong linear relationship between the two variables, with a nearly perfect positive correlation. The correlation values suggest that there is a very strong positive relationship between the total number of COVID cases and the total number of COVID deaths. This means that as the number of cases increases, the number of deaths also tends to increase.

Table 4- Comparison of Existing model and new proposed model

	Existing Model Cheshmehzangi et.al.,(2021)	New Model Design
Platform	SEIR Data Analysis	Python
Base Research	Data Analysis and Exploration	Data Analysis, Exploration, Forecasting
Research Area	Covid-19	Covid-19
Method	SEIR Model	Holt-Winters's
Modules	Single (only Exploration)	Two (Exploration + Prediction Method)
Outcome	In Graphical Form	Graphical and Numeric

As the suggested scenario makes use of data analysis, exploration, and forecasting, but the study that has already been done solely uses exploration based on data analysis, this highlights a key difference. Both the accuracy of the forecast and the analytical approach used for this inquiry were significantly improved thanks to the suggested feature.

VI. CONCLUSION

The analysis and prediction of COVID-19 cases in Indian states using a comprehensive methodology have provided valuable insights into the spread of the virus and its potential trajectory in the year 2023. By considering multiple growth models and recent trends in infections, the methodology has allowed for a comprehensive assessment of the severity and dynamics of the pandemic at the state level.

The categorization of states into severe, moderate, or controlled categories based on the analysis provides a framework for targeted interventions and resource allocation. This classification enables policymakers and public health authorities to prioritize their efforts and implement appropriate measures to contain the virus effectively. It facilitates decision-making regarding the implementation of stricter containment strategies, provision of healthcare support, and allocation of resources based on the severity of the COVID-19 situation in each state.

By considering factors such as population density, healthcare infrastructure, and socioeconomic conditions, the predictions aid in resource allocation, capacity building, and identifying areas that require special attention. These predictions serve as a crucial tool in guiding policymakers and public health authorities in making informed decisions and developing targeted strategies to combat the pandemic effectively.

It is important to note the limitations and uncertainties associated with the results. The accuracy of the predictions relies on the quality and availability of the data used, as well as the assumptions and limitations of the growth models employed. Ongoing monitoring of the pandemic, regular updates to the models, and incorporating new data are essential to refine the analysis and improve the accuracy of the predictions.

In conclusion, the analysis and prediction of COVID-19 cases in Indian states using the proposed methodology provide valuable insights for managing the ongoing pandemic. The categorization of states, along with the predictions of infection numbers, offers guidance for policymakers, public health authorities, and researchers in implementing targeted interventions, allocating resources effectively, and planning for the future. Continuous monitoring and refinement of the analysis are crucial to adapt to the evolving nature of the pandemic and ensure effective management of COVID-19 in India.

According to the findings of our study, one of the challenges that we had in conducting our analysis was the need to understand and use visualizations. This study also found that the trust of customers and the assistance of organizations had a significant role in the adoption of these strategies. Since quite some time, we have taken interoperability very seriously. When one is talking about a new product, one may overcome problems by making use of other tools. For instance, one might describe elements that can be neglected or that might impart confusing information in order to clear up any misunderstandings that may arise.

REFERENCES

[1] Government of India's Ministry of Health & Family Welfare. (2021). India COVID-19. from https://www.mohfw.gov.in

[2] Anand, A., Sandaraju, H., Parande, M., & Srivastava, S. (2020). COVID-19 in India: A review of current knowledge. Indian Journal of Medical Research, 151(2), 147-159.

[3] Wang, L., Wang, Y., Ye, D., & Liu, Q. (2020). Review of the 2019 novel coronavirus (SARS-CoV-2) based on current evidence. International Journal of Antimicrobial Agents, 55(6), 105948.

[4] World Health Organization. (2021). Coronavirus disease (COVID-19) pandemic. Retrieved from https://www.who.int/emergencies/diseases/novel-coronavirus-2019

[5] Verma, M., Singh, R. K., Jha, R. K., & Singh, D. P. (2021). Predictive modeling of COVID-19 outbreak in India using machine learning and social network analysis. Chaos, Solitons & Fractals, 144, 110769.

[6] ICMR-National Institute of Epidemiology. (2021). COVID-19 India State-wise Data. Retrieved from https://icmr.nic.in/

[7] Deo, R. C., & Sahin, E. (2020). Machine learning models for early prediction of COVID-19 outbreak in India. medRxiv. doi: 10.1101/2020.04.15.20067187

[8] Chen, J., Lu, Y., Bai, J., & Xu, H. (2020). Predicting the epidemic dynamics of COVID-19 in China and the US with a data-driven stochastic model. Chaos, Solitons & Fractals, 139, 110055.

[9] A. Kumar, R. K. Sharma, and others (2021). Modelling instances of COVID-19 using machine learning techniques in India. 10(1), 35–43; Health Policy and Technology.

[10] R. M. Elavarasan and R. Pugazhendhi (2020). Review of possible technology controls for the COVID-19 pandemic in a restructured society and environment. Environmental Science and Technology, 725, 138858

[11] Bloom, D. E., &Cadarette, D. (2019). Infectious disease threats in the twenty-first century: strengthening the global response. *Frontiers in immunology*, *10*, 549.

[12] Agrebi, S., &Larbi, A. (2020). Use of artificial intelligence in infectious diseases. In *Artificial intelligence in precision health* (pp. 415-438). Academic Press.

[13] Morse, S. S. (2001). Factors in the emergence of infectious diseases. *Plagues and politics*, 8-26.

[14] Cohen, R., Ashman, M., Taha, M. K., Varon, E., Angoulvant, F., Levy, C., ...&Grimprel, E. (2021). Pediatric Infectious Disease Group (GPIP) position paper on the immune debt of the COVID-19 pandemic in childhood, how can we fill the immunity gap?. *Infectious Diseases Now*, *51*(5), 418-423.

[15] Ortiz-Prado, E., Simbaña-Rivera, K., Gomez-Barreno, L., Rubio-Neira, M., Guaman, L. P., Kyriakidis, N. C., ...& Lopez-Cortes, A. (2020). Clinical, molecular, and epidemiological characterization of the SARS-CoV-2 virus and the Coronavirus Disease 2019 (COVID-19), a comprehensive literature review. *Diagnostic microbiology and infectious disease*, *98*(1), 115094.

[16] .Sohrabi, C., Alsafi, Z., O'neill, N., Khan, M., Kerwan, A., Al-Jabir, A., ...& Agha, R. (2020). World Health Organization declares global emergency: A review of the 2019 novel coronavirus (COVID-19). *International journal of surgery*, *76*, 71-76.

[17] Tran, H. N., Le, G. T., Nguyen, D. T., Juang, R. S., Rinklebe, J., Bhatnagar, A., ...& Chao, H. P. (2021). SARS-CoV-2 coronavirus in water and wastewater: A critical review about presence and concern. *Environmental research*, *193*, 110265.

[18] Stadnytskyi, V., Anfinrud, P., &Bax, A. (2021). Breathing, speaking, coughing or sneezing: What drives transmission of SARS-CoV-2?. *Journal of Internal Medicine*, *290*(5), 1010-1027.

[19] UmaMaheswaran, S. K., Munagala, N. K., Mishra, D., Othman, B., SINTHU, S., &Tripathi, V. (2022, April). The role of implementing Machine Learning approaches in enhancing the effectiveness of HealthCare service. In *2022 2nd*

International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE) (pp. 1922-1926). IEEE.

[20] .Sankar, P., Ahmed, W. N., Koshy, V. M., Jacob, R., &Sasidharan, S. (2020). Effects of COVID-19 lockdown on type 2 diabetes, lifestyle and psychosocial health: a hospital-based cross-sectional survey from South India. *Diabetes & Metabolic Syndrome: Clinical Research & Reviews*, *14*(6), 1815-1819.

[21] Shi, Y., Wang, G., Cai, X. P., Deng, J. W., Zheng, L., Zhu, H. H., ... & Chen, Z. (2020). An overview of COVID-19. *Journal of Zhejiang University-SCIENCE B*, *21*(5), 343-360.

[22] Turjman, F.A. (2021). AI-powered cloud for COVID-19 and other infectious disease diagnosis. *Personal and Ubiquitous Computing,* s00779-021.

[23] Erraguntla, M., Zapletal, J., &Lawley, M. (2019). Framework for Infectious Disease Analysis. *Health Informatics Journal,* 25(4), 1170–1187.

[24] Bogu, G.K., & Snyder, M.P. (2021). Deep learning-based detection of COVID-19 using wearables data. *medRxiv preprint,* 212-49474.

[25] Wu, J., Zhang, P., Zhang, L., Meng, W., & Li, J. (2020). Rapid and accurate identification of COVID-19 infection through machine learning based on clinical available blood test results. *medRxiv preprint,* 2005-1136.

[26] Khan, F.A., Rakhami, M.A., & Ahmad, S. (2021), Detection and Prediction of Diabetes Using Data Mining. 4(12), 3059-343.

[27] Ravi, M.A., Gopal, M.V., & Roselyn, J.P. (2021), Detection of Infectious Disease using Non-Invasive Logistic Regression Technique. *IEEE Xplore,* 57(17).

[28] Bruyndonckx, R., &Coenen, S. (2021). Analysing the trend over time of antibiotic consumption in the community: a tutorial on the detection of common change-points. *J AntimicrobChemother,* 7(6), ii79–ii85.

[29] .Martinez, J.T., & Pérez, M.G. (2021), A Novel Machine Learning-Based Approach for the Detection of SSH Botnet Infection. *Future Generation Computer Systems,* 10-1016.

[30] Huang, S., & Yang, G. (2021), Artificial intelligence in the diagnosis of COVID-19. *International Journal of Biological Sciences,* 17(6), 1581-1587.

[31] Kim, M., &Chae, K. (2020). Automated Classification of Online Sources for Infectious Disease Occurrences Using Machine-Learning-Based Natural Language Processing Approaches. *International Journal of Environmental Research and Public Health,* 2(13).

[32] Pundir, A., &Bhardwaj, S. (2021). Prediction and Detection of Covid-19 Using Ml / Ai. *International Journal of Multidisciplinary and Current Educational Research (IJMCER),* 3(3), 2581-7027.

[33] Nikooghadam, M., &Ghazikhani, A. (2020). COVID-19 Prediction Classifier Model Using Hybrid Algorithms in Data Mining. *Int J Pediatr,* 9(1), 12723-12737.

[34] Borkenhagen,L.K., & Allen, M.W. (2021). Influenza virus genotype to phenotype predictions through machine learning. *Emerging Microbes & Infections,* 10(1), 1896-1907.

[35] Arji, G., &Ahmadi, H. (2019). Fuzzy logic approach for infectious disease diagnosis. Biocybernetics and Biomedical Engineering.

[36] Smith, K.P., & Kirby, E.J. (2020), Image analysis and artificial intelligence in infectious disease diagnostics. *Clinical Microbiology and Infection,* 6(13).

[37] Amar, L.A., Taha, A. A., & Mohamed, M.Y. (2020). Prediction of the final size for COVID-19 epidemic using machine learning. *Infectious Disease Modelling,* 622e634.

[38] Nilashi, M., Ahmadi, H., &Shahmoradi, L. (2018). A predictive method for hepatitis disease diagnosis using ensembles of neuro fuzzy technique. *Journal of Infection and Public Health.*

[39] Salehi, A.W., Baglat, P., & Gupta, G. (2020). Review on Machine and Deep Learning Models for the Detection and Prediction of Coronavirus. *Journal Pre-proofs.*

[40] Li, M. (2019). Study on the Grouping of Patients with Chronic Infectious Diseases Based on Data Mining. *Journal of Biosciences and Medicines,* 119-135.

[41] Cruz, A.P.D., &Tumibay, G. M. (2019). Predicting Tuberculosis Treatment Relapse. *Journal of Computer and Communications,* 243-251.

[42] Callejon-Leblic, M.A., Moreno-Luna, R., &Cuvillo, A.D. (2021). Loss of Smell and Taste Can .Accurately Predict COVID-19 Infection. *Clin. Med,* 1004-0570

[43] Kivrak, M., &Colak, C. (2021). Prediction of death status on the course of treatment in SARS-COV-2 patients with deep learning and machine learning methods. *Computer Methods and Programs in Biomedicine,* 20(1), 105951.

[44] Kukar, M., &Gncer, G. (2020). COVID‑19 diagnosis by routine blood tests using machine learning. *Scientific Reports,* 1598-021.

[45] Lalmuanawma, S., &Hussain, J. (2020). Applications of machine learning and artificial intelligence for Covid-19 (SARS-CoV-2) pandemic. *Nonlinear Science, and Nonequilibrium and Complex Phenomena,* 110059.

[46] Zeng, D., Cao, D., & Neill, D.B. (2021). Artificial intelligence enabled public health surveillance—from local detection to global epidemic monitoring and control. *Artificial Intelligence in Medicine,* 821259-2.00022-3.

[47] Natarajan, Y., Kannan, S., &Mohanty, S.N. (2021). Survey of Various Statistical Numerical and Machine Learning Ontological Models on Infectious Disease Ontology. *Infectious Disease Ontology Computation,* 431–442.

[48] Woldaregay, A.Z., &Launonen, L.K. (2020), A Novel Approach for Continuous Health Status Monitoring and Automatic Detection of Infection Incidences in People With Type 1 Diabetes Using Machine Learning Algorithms. *J Med Internet Res,* 22(8), e1891233

[49] Heba M. Afify, Muhammad Zanaty, "Computational Predictions for Protein Sequences of COVID-19 Virus via Machine Learning Algorithms", Medical & Biological Engineering & Computing, Springer, vol.59, no.9, 2021.

[50] Heba M. Afify, Muhammad Zanaty, "A Comparative Study of Protein Sequences Classification-Based Machine Learning Methods for COVID-19 Virus against HIV-1", Applied Artificial Intelligence, Taylor & Francis, vol. 35, no. 15, pp: 1733-1745, 2021.

[51] Ministry of Health & Family Welfare, Government of India. (2021). COVID-19 India. Retrieved from https://www.mohfw.gov.in/