# An Analytical Survey on predicting Covid-19 using Deep Learning

1. Vishal Shinde , Computer Engineering , KJCOEMR, Pune. shindevt2009@gmail.com
2. Nagaraju Bogiri, Computer Engineering, KJCOEMR, Pune. bogirinagaraju.kjcoemr@kjei.edu.in

***Abstract —*** *Epidemics are extremely difficult and problematic circumstances because of the amount of stress that everyone experiences throughout these periods. The most concerning aspect for healthcare personnel and individuals is the uncertainty of the future. There was a paucity of tools that could be used to successfully anticipate the infection rate in the future during the current covid-19 infection outbreak. This has resulted in increasing ambiguity, which may be extremely difficult for medical professionals and governments in terms of planning. An effective predictive technique is required for this purpose in order to provide reliable estimates of future infection rates. This study paper arrived at an effective technique by analyzing previous studies on this issue, which have been described in this article. Our methodology will be expanded on in future versions of this research.*

***Keywords:*** *Pearson Correlation, Convolutional Neural Networks, Long Short Term Memory, Decision Tree*

## I INTRODUCTION

Some of the most important parts of a meaningful human existence are health and its upkeep. Having excellent health helps an individual to enjoy life to the fullest and reach the objectives that they have set for themselves. This enables for more effective and usable energy to be dispensed while doing helpful and productive tasks. When a person is not in complete good health, it can lead to a loss in performance and efficiency when compared to good health. As a result, staying healthy is one of the most important conditions for living a full life.

Numerous developments and studies have been conducted to enable this technique in the medical area. It has developed tremendously throughout the years, setting long-standing records in healthcare, illness prevention, and cure. The medical paradigm is one of the most important study topics that is continually improving the health and well-being of people all over the world. This study has been bolstered, resulting in the treatment of numerous severe ailments over the years.

There have long been pandemics that were effectively predicted to be extinction-level events for the human species. Virulent outbreaks, as as the current covid-19 epidemic, have demonstrated how a disease may be life-threatening on a worldwide scale. There are groups that monitor behavior and warn of the possibility of an epidemic in advance. This can assist to prepare the public for the breakout and avert a major disaster.

The current Covid-19 epidemic was not managed properly by the countries because they lacked the mechanisms to efficiently identify the virus's spread and discover strategies to limit it. These outbreaks have been extremely worrisome since they have resulted in a big number of deaths and a great deal of agony and suffering for the afflicted. There was a scarcity of integrated data that could be used to efficiently process the final result of real epidemiologists and other pandemic officials.

There have been studies that have been conducted for this goal in order to efficiently collect important data and develop models for future projections of the spread. They have been substantially developed in this study to achieve our approach for delivering future forecasts of virus transmission. However, the majority of these techniques have been determined to have some form of constraint that has impeded an efficient implementation. As a result, there is a need for an effective approach that can offer precise estimates for future infection rates, allowing governments to plan ahead of time.

Machine learning algorithms have been the finest choice for this purpose since they give very accurate forecasting capabilities. The CNN and Long Short Term Memory along with decision tree machine learning algorithms have been incorporated in this technique to produce an effective prediction of the correlation achieved by the Pearson correlation in our methodology. Our technique will be described in greater depth in future editions of this research artic le.

This literature survey paper segregates the section 2 for the evaluation of the past work in the configuration of a literature survey, and finally, section 3 provides the conclusion and the future work.

## II RELATED WORKS

According to Domenico Gaglione [1,] the recent worldwide pandemic outbreak caused by a new strain of Coronavirus has accelerated research into unique mathematical models and algorithms capable of properly estimating and forecasting the infection's epidemiological curve. In this paper, we introduced a Bayesian sequential estimating and forecasting method that, using daily data from authorities, can estimate the status of the epidemic and the parameters of the underlying model, as well as anticipate the evolution of the epidemiological curve. The authors created an efficient implementation of the previously proposed Bayesian framework, suited particularly to the stochastic SIR model of pandemic development. The proposed algorithm is validated using synthetic data simulating two epidemic scenarios, and on real data acquired during the recent COVID-19 outbreak.

Hamdi Friji [2] developed a modelling approach composed of eight states to categorize the COVID-19 distributed in multiple nations. To quantitatively represent the interactions between the distinct states, an Ordinary Differential Equation system is developed. To find the precise fitting solutions and estimate the parameters of the Ordinary Differential Equation model, a curve fitting employing real-world observed data sets is devised. The Levenberg-Marquardt technique was used by the researchers to solve and estimate the Ordinary Differential Equation system parameters. The authors have conducted tests that indicate that the algorithms used surpass the Broyden-Fletcher-Goldfarb-Shanno reference model. The fitting approach is depicted beside the model, which seeks to validate its predicting findings by concentrating on the most recent data patterns.

Yi-Cheng Chen [3] provides the detailed mathematical and numerical analysis for the Covid -19 infection rates. In compared to traditional Susceptible-Infected-Recovered models, the time dependent Susceptible-Infected-Recovered model is helpful and adaptable. These SIR models are also contrasted to direct estimating approaches, resulting in the SIR being much more robust. To measure the influence of undetected infections in COVID-19, the authors updated the SIR model by taking into account two groups of infected people: detectable infected people and untraceable infected people.

S. Dash [4] developed the best Forecasting methods for six high-incidence countries around the world and six seriously afflicted states in India to estimate 90 days in advance of future values to explain pandemic transmission. Except in a few cases, the theory predicts quantities that are reasonably close to the actual values. The ARIMA model used in the Great Britain is not the best fit. The expected RMSE, AIC, and BIC numbers are suitable for all six nations. Similarly, ARIMA models developed for Indian states have large predicted RMSE, AIC, and BIC values. The effectiveness of the ARIMA Model is shown by statistical performance measures. The model's conclusions might be utilized to develop potential ways to improve the medical system and management.

A. Ramchandani [5] provide a novel deep learning framework for analyzing multiple county-level parameters and anticipating the spread of infected cases in the future. The proposed approach extracts annotations both multivariate time series including multivariate spatial time series data in a novel way by utilizing both of the spatial and temporal representation of the model. This embedding extraction strategy might be used in other deep learning projects to evaluate similar sorts of data. Furthermore, in contrast to previous models, the proposed method incorporates a large number of input characteristics and learns correlations between them. During the COVID-19 pandemic, the model's accuracy in anticipating the increase in the number of cases reported in U.S. counties was demonstrated.

O. Tutsoy [6] develops a model of Suspicious-Infected-Death (SpID) with completely unknown dynamics. Because suspicious, contaminated, and death fatalities are all strongly dependent on one another, the accepted SpID model is closely interrelated. Each SpID sub-model adds 2nd order internal workings to describe the peaks and variations in COVID-19 casualties. The precise bases correlating to the model's parameter space are created, and the design variables are learnt by training the SpID model's unknown parameters that used a batch type Least Squares (LS) estimator. The model with the established characteristics was thoroughly studied employing mathematical methods, and prospective COVID-19 casualties for Turkey were predicted using the built model.

The findings of X. Chen's work improve past asymptomatic infector estimates driven by concerns specific to the COVID-19 pandemic [7]. To begin, this paradigm considers a more realistic components causing where the asymptomatic infector community is obtained from the sick instead of the affected population. Second, the rehabilitation and mortality groups are evaluated individually, with the deceased serving as an important to systematically of the impact of different response approaches. Third, assess the effectiveness of various preventative and response approaches under a range of implementation time constraints, since this may give important information for selecting the proper epidemic preparedness and safety strateg ies in reality.

In the context of the current epidemic, E. Karaçuha [8] wished to provide practical guidance on how to estimate case numbers in order to establish a better plan for allocating health resources to patients. This simulation might be used not ju st for the current COVID-19 epidemic, but also for future local or worldwide outbreaks. This study models and forecasts the number of confirmed cases, fatalities, and recoveries from the COVID-19 outbreak for eight countries: China, France, Germany, Italy, Spain, Turkey, the United Kingdom, and the United States. First, the authors used their previously published Deep Assessment Methodology, which is based on Fractional Calculus, to model the COVID-19 data from the first confirmed case date through April 19$^{th}$. Following that, the DAM and Long-Short Term Memory (LSTM) were used to evaluate DAM effectiveness in a one-step prediction. The final section of the study concentrated on short-term pandemic forecasting, with the Time-Dependent Dynamic model and a Gaussian simulation projecting the next 30 days based on the elevation of the continuous number of reported cases received from DAM.

Michael Small [9] provided a framework with a small – potentially fundamental – number of factors that captures the observed patterns of pandemic sickness spread. Whenever employed correctly from the international coronavirus outbreak, the model shows good qualitative agreement overall observable values across population centers. Despite this, same simulations with different starting conditions yield significantly different outcomes. The volatility in the reported model forecasts is significantly larger than the variance seen across other epidemiological parameter settings. As a result, after adequately characterizing contact patterns and creating a transmission mitigation strategy, determining optimal transmission rates becomes a secondary concern. Their findings imply that simulations of models asserting predictive capability inside that prediction boundary may be susceptible to over-interpretation.

B. Wang [10] examined ways to harness social ties amongst mobile devices in SIoT to reduce infection rates by recognizing possible COVID19 infections early. The authors then turned the optimization problem into a MWVC problem for dynamic network design and devised a RAI approach to solve it. Using two actual datasets, they demonstrate that their strategy reduces the epidemic infection rate dramatically across both huge and limited circumstances when contrasted to the benchmarks . Finally, by relying on early identification of COVID-19 cases, the proposed strategy is ideal for disease and prevention.

F. Rustam [11] created a machine learning-based forecasting approach for assessing the risk of a global COVID19 outbreak. The system analyzes a dataset containing day-by-day real historical data and makes predictions for the upcoming days using machine learning techniques. The study's findings suggest that ES performed the best in the current forecasting domain given the kind and size of the dataset. To some extent, LR and LASSO are indeed good at forecasting death rates and confirming instances. According to the conclusions of these two models, fatality rates will climb in the next days, whereas health outco mes will slow. SVM produces poor results in all circumstances because to the peaks and troughs in the dataset values. Creating an exact higher dimensional space between the dataset's provided values proved difficult.

Learning theory was established by R. B. Duffey [12] to characterize the reduction of pandemic infections such as CoVid-19. Both the disease incidence, as a meter of wrong outcome, and time, as a measurement of experience/knowledge gained or risk exposure that supports learning, are critical concerns. According to reviews of existing data, the Co Vid-19 subsequent increase data nearly exactly follow the Universal Training Time detailing the long decline of several other situat ions people learn to deploy appropriate measures. The learning curve for various modern technological systems handled by peo ple is about the same (with universal parameter k3) as for any learning opportunity, avoiding results, accidents, and incidents.

N. Zheng's [13] hybrid AI architecture for forecasting COVID-19 is based on the ISI framework and incorporates an NLP component that combines key information obtained via central and local government activities, as well as extensive public engagement in the prediction computation process. The model's estimations are indeed very close to the real epidemic cases, demonstrating that the proposed hybrid model can more accurately analyze the virus's transmission law and development trend than past models, and that language data processing of related news can dramatically improve the model's accuracy.

## III CONCLUSION AND FUTURE SCOPE

This research paper specifies the process for generating accurate and exact infection rates for the covid-19 pandemic. The lack of an appropriate method for highly precise covid-19 infection rates and an adequate pandemic forecasting model resulted in large-scale loss of life and suffering for a lot of people worldwide. This was an issue that was found in both emerging and industrialized countries. This is because the majority of the researchers in the linked works that we have investigated on this issue have suffered from some of the other shortcomings and limitations. This has resulted in erroneous forecasts for mistakes, lowering the overall predictive model's efficiency. A variety of similar studies have been extensively reviewed for this aim in order to achieve our technique for giving an accurate covid-19 infection rate forecasting model. For generating accurate and exact covid-19 forecasts, our methodology has concentrated on the application of long short term memory and decision tree as machine learning methodologies. The approach will be detailed in greater depth in later revisions of this study paper.

## REFERENCES

[1] D. Gaglione et al., "Adaptive Bayesian Learning and Forecasting of Epidemic Evolution—Data Analysis of the COVID-19 Outbreak," in IEEE Access, vol. 8, pp. 175244-175264, 2020, DOI: 10.1109/ACCESS.2020.3019922.

[2] H. Friji, R. Hamadi, H. Ghazzai, H. Besbes, and Y. Massoud, "A Generalized Mechanistic Model for Assessing and Forecasting the Spread of the COVID-19 Pandemic," in IEEE Access, vol. 9, pp. 13266-13285, 2021, DOI: 10.1109/ACCESS.2021.3051929.

[3] Y. -C. Chen, P. -E. Lu, C. -S. Chang and T. -H. Liu, "A Time-Dependent SIR Model for COVID-19 With Undetectable Infected Persons," in IEEE Transactions on Network Science and Engineering, vol. 7, no. 4, pp. 3279-3294, 1 Oct.-Dec. 2020, DOI: 10.1109/TNSE.2020.3024723.

[4] S. Dash, C. Chakraborty, S. K. Giri, S. K. Pani, and J. Frnda, "BIFM: Big-Data Driven Intelligent Forecasting Model for COVID-19," in IEEE Access, vol. 9, pp. 97505-97517, 2021, DOI: 10.1109/ACCESS.2021.3094658.

[5] A. Ramchandani, C. Fan, and A. Mostafavi, "DeepCOVIDNet: An Interpretable Deep Learning Model for Predictive Surveillance of COVID-19 Using Heterogeneous Features and Their Interactions," in IEEE Access, vol. 8, pp. 159915-159930, 2020, DOI: 10.1109/ACCESS.2020.3019989.

[6] O. Tutsoy, Ş. Çolak, A. Polat and K. Balikci, "A Novel Parametric Model for the Prediction and Analysis of the COVID-19 Casualties," in IEEE Access, vol. 8, pp. 193898-193906, 2020, DOI: 10.1109/ACCESS.2020.3033146.

[7] X. Chen, "Infectious Disease Modeling and Epidemic Response Measures Analysis Considering Asymptomatic Infection," in IEEE Access, vol. 8, pp. 149652-149660, 2020, DOI: 10.1109/ACCESS.2020.3016681.

[8] E. Karaçuha et al., "Modeling and Prediction of the Covid-19 Cases With Deep Assessment Methodology and Fractional Calculus," in IEEE Access, vol. 8, pp. 164012-164034, 2020, DOI: 10.1109/ACCESS.2020.3021952.

[9] Small, Michael & Cavanagh, David. (2020). Modeling Strong Control Measures for Epidemic Propagation With Networks—A COVID-19 Case Study. IEEE Access. PP. 1-1. 10.1109/ACCESS.2020.3001298.

[10] B. Wang, Y. Sun, T. Q. Duong, L. D. Nguyen, and L. Hanzo, "Risk-Aware Identification of Highly Suspected COVID-19 Cases in Social IoT: A Joint Graph Theory and Reinforcement Learning Approach," in IEEE Access, vol. 8, pp. 115655-115661, 2020, DOI: 10.1109/ACCESS.2020.3003750.

[11] F. Rustam et al., "COVID-19 Future Forecasting Using Supervised Machine Learning Models," in IEEE Access, vol. 8, pp. 101489-101499, 2020, DOI: 10.1109/ACCESS.2020.2997311.

[12] R. B. Duffey and E. Zio, "Analyzing Recovery from Pandemics by Learning Theory: The Case of CoVid-19," in IEEE Access, vol. 8, pp. 110789-110795, 2020, DOI: 10.1109/ACCESS.2020.3001344.

[13] N. Zheng et al., "Predicting COVID-19 in China Using Hybrid AI Model," in IEEE Transactions on Cybernetics, vol. 50, no. 7, pp. 2891-2904, July 2020, DOI: 10.1109/TCYB.2020.2990162.

*****