

Journal of Data Science and Information Systems



Edited by
Ulrich Böck
Dimitris Gunz
Victor Hugo C. de Albuquerque
Antonio Ribeiro

Data Science for **COVID-19**

Volume One: Computational Perspectives



Actions for selected chapters

Select all / Deselect all

Download PDFs

Export citations

Book chapter Full text access
26 - Received signal strength indication—based COVID-19 mobile application to comply with social distancing using bluetooth signals from smartphones

Alvaro Asplicueta Narvaez and Jorge Guerra Guerra
Pages 483-501

[Download PDF](#) [View abstract](#)

Book chapter Full text access
27 - COVID-19 pandemic in India: Forecasting using machine learning techniques

H.S. Hota, Richa Handa and A.K. Shrivastava
Pages 503-525

[Download PDF](#) [View abstract](#)

Book chapter Full text access
28 - Mathematical recipe for curbing coronavirus (COVID-19) transmission dynamics

Emmanuel Bala Gudu, Daniel Dauda Wisdom, ... Ebenezer Ajayi Akinyemi
Pages 527-545

[Download PDF](#) [View abstract](#)

FEEDBACK

COVID-19 pandemic in India: forecasting using machine learning techniques

H.S. Hota¹, Richa Handa², A.K. Shrivastava³

¹ATAL BIHARI VAJPAYEE UNIVERSITY, BILASPUR, CHHATTISGARH, INDIA; ²DR. C.V. RAMAN UNIVERSITY, BILASPUR, CHHATTISGARH, INDIA; ³GURU GHASIDAS UNIVERSITY, BILASPUR, CHHATTISGARH, INDIA

1. Introduction

The novel coronavirus disease 2019 (COVID-19) has been declared a world pandemic threat by the World Health Organization and is spreading rapidly across the world. It is an infectious disease caused by a virus called a novel coronavirus, which was also the reason for severe acute respiratory syndrome (SARS) in 2003. The SARS epidemic [1,2] was predicted by many authors using various machine learning (ML) techniques. Because of COVID-19, the number of confirmed cases and deaths are rapidly increasing in all most all countries; India has not escaped this pandemic. The first case of COVID-19 emerged in India on Jan. 30, 2020; 2 more cases were found on Feb. 3, 2020 and the number was stable until Mar. 1, 2020. The disease has since spread to most states and cities of India [3]. Few states or cities remain untouched by the COVID-19 outbreak in India. There is no vaccine to treat COVID-19 to prevent infection from one infected person to another. India has learned a lesson from other countries, and therefore the Indian government has made appropriate decisions and implemented various strategies to prevent the pandemic from spreading across the country well in advance. Many approaches were taken to stop spreading cases of COVID-19 from affected to unaffected parts of the country, including a citywide lockdown, closing all transports such as airports, railways, and local transportations, and closing markets, malls, cinemas, productions, and so on. Moreover, the isolation or quarantine of suspected patients is being done. The entire machinery of the government is fully involved to stop spreading it in the community; despite this, cases of COVID-19 positive are increasing every day. By the time of writing, there were more than 26,283 confirmed cases in India, 825 of whom died and about 5938 of whom had recovered, and the numbers keep rising [4].

Various statistical and mathematical analysis and studies are ongoing to forecast the future trend of COVID-19 in India, and models are being developed to predict the future situation, known as N-days in forecasting. Because of the nonlinear behavior of COVID-19 data, various ML techniques could be useful to develop a robust forecasting model. Research shows that the application of computational intelligence methods is the basis for constructing a predictive model. In this, the neural network is useful for predicting time series data because it has the ability to learn from data and capture the various dynamics of time series data [5]. Evolutionary computations, fuzzy logic, and other models are also crucial owing to their principal differences from existing mathematical approaches. Hybrid models of various intelligent techniques are also widely used in forecasting, because accuracy and efficiency are the most important criteria of focus by researchers [6].

Most work has been done on the basis of either trend already experienced by other countries such as China or statistical theory and analysis. Roosa et al. [7] worked on the real-time forecasting of the COVID-19 pandemic in China and developed models for 5-, 10-, and 15-days ahead forecasting based on the cumulative number of confirmed cases in Hubei and other provinces of China using three different techniques: the Richards model, the subepidemic model, and generalized logistic growth model (GLM). They concluded that each model predicts that the pandemic has reached saturation in Hubei and other provinces of China. Benvenuto et al. [8] performed an autoregressive integrated moving average (ARIMA) on Johns Hopkins epidemiological data to predict the epidemiological trend of the prevalence and incidence of COVID-19. Abdulmajeed et al. [9] proposed an online forecasting mechanism that streams data from the Nigeria Center for Disease Control, which provides updated COVID-19 forecasts every 24 h. The authors combine ARIMA, Prophet (an additive regression model developed by Facebook), and a Holt–Winters exponential smoothing model combined with generalized autoregressive conditional heteroscedasticity. In other work, Tuli et al. [10] applied an improved mathematical model to analyze and predict the growth of the epidemic of COVID-19 and deployed the model on a cloud computing platform for more accurate and real-time prediction of the growth behavior of the epidemic. Ardabili et al. [11] presented a comparative analysis of ML and soft computing models to predict the COVID-19 outbreak and found ML to be an effective tool to model the outbreak. Zhou et al. [12] explained the challenges to geographic information systems (GIS) with big data on COVID-19. Other authors [13] analyzed and forecast COVID-19 spread in China, Italy, and France, and concluded that the infection rate needed to be cut down drastically and quickly to observe an appreciable decrease in the pandemic peak and mortality rate.

Tobías et al. [14] analyzed the trends of incident cases, deaths, and intensive care unit admissions in Italy and Spain before and after their respective national lockdowns using an interrupted time-series design. Data were analyzed with quasi-Poisson regression using an interaction model to estimate the change in trends. Chintalapudi et al. [15] highlighted the importance of lockdown and isolation by forecasting registered and recovered cases of COVID-19 after 60 days' lockdown in Italy adopting a seasonal ARIMA forecasting package with the R statistical model. Koczkodaj et al. [16] predicted the