International Journal of E-Health and Medical Communications
Volume 13 • Issue 2

Comparison of Active COVID-19 Cases per Population Using Time-Series Models

Sakinat Oluwabukonla Folorunso, Olabisi Onabanjo University, Ago Iwoye, Nigeria
Joseph Bamidele Awotunde, University of Iliorin, Iliorin, Nigeria
https://orcid.org/0000-0002-1020-4432
Oluwatobi Oluwaseyi Banjo, Olabisi Onabanjo University, Ago Iwoye, Nigeria
https://orcid.org/0000-0002-5440-7442
Ezekiel Adebayo Ogundepo, Data Science Nigeria, Nigeria
Nureni Olawale Adeboye, Federal Polytechnic, Ilaro, Nigeria

ABSTRACT

This research explored the precision of diverse time-series models for COVID-19 epidemic detection in all the 36 different states and the Federal Capital Territory (FCT) in Nigeria with the maximum count of daily cumulative of confirmed, recovered, and death cases as of 4 November 2020. A 14-step forecast system for active coronavirus cases was built, analyzed, and compared for six different deep learning-stimulated and statistical time-series models using two openly accessible datasets. The results obtained showed that based on RMSE metric, ARIMA model obtained the best values for four of the states (0.002537, 0.001969.12E-058, 5.36E-05 values for Lagos, FCT, Edo and Delta states, respectively). While no method is all-encompassing for predicting daily active coronavirus cases for different states in Nigeria, ARIMA model obtains the highest-ranking prediction performance and attained a good position results in other states.

KEYWORDS

ARIMA, COVID-19, Deep Learning, Epidemic, Exponential Smoothing, Forecasting

1. INTRODUCTION

The novel severe contagious respiratory syndrome coronavirus called (COVID-19) has cause greatest global challenge and public health, after the pandemic of influenza outbreak of 1918. The spread of COronaVirus Disease 2019 (COVID-19) has caused significant changes in the style of living of communities around the world. The onset of the epidemic was discovered in Wuhan, China in December, 2019 and was triggered due to Severe Acute Respiratory Syndrome Coronavirus 2 (SARSCoV-2) virus (Ogundokun, Lukman, Kibria, Awotunde, & Aladeitan, 2020). As stated by World Health Organization (WHO) there have been 30, 949,804 confirmed cases and 959,116

DOI: 10.4018/IJEHMC.20220701.oa6

deaths of COVID-19 Globally, as of 4:30pm CEST, 21 September 2020 and affecting more than 180 countries, and due to the exponential level of spread, WHO stated COVID-19 as an epidemic. These make governments of different countries to take immediate control actions like separating profoundly influenced regions, stopping the cross-border traffic between nations, shutting down schools, working environments, regular spots, limiting resident's movements by encouraging them to stay at as much as possible. The social life and economics had significantly affected due to these measures put in place. The societies face different challenges ranges from education, healthcare, manufacturing, supply chain management travel, tourism and service delivery under the prevailing conditions and in a post COVID-19 world. As an example, the overcrowded of hospitals and other healthcare facilities due to exponential increase of COVID-19 patients and the inability to provide medical assistance to normal patients due to limited movement are significant barriers to the fight against COVID-19 in the healthcare sector. Similarly, delays and the increased resource demand for manual contact tracing, and unavailability of effective and automated contact tracing software impede the actions for controlling the spread. Hence, it is the duty of different parties to act with their fullest potential to control the prevailing COVID-19 situation, such parties are healthcare workers, government authorities, students, researchers, engineers, technology managers and the general public. In order to not only safeguard, but also handle the post-COVID-19 environment, digitalization and the implementation of information and communication technologies will be imperative. Technologies like Big data, 5G communication, Internet of Things (IoT), cloud computing, Artificial Intelligence (AI), and blockchain play vital role to assist the environment adopting to different protection and improvement of people and economies. The technologist and engineers will have to tackle important challenges to implement these promising solutions and realize their benefits, and carry out elaborate findings regard to risk management, resources, cost, scope, and quality.

In order to predict the prevalence of coronavirus in Nigeria, some researchers recently used the Linear Regression (LR) and curve estimation model (Ogundokun, Lukman, Kibria, Awotunde, & Aladeitan, 2020; Lukman, et al., 2020). In India (Pandey, Chaudhary, Gupta, & Pal, 2020) used the SEIR and Regression model to predict COVID 19 and LR to estimate the count of deaths in India due to coronavirus. (Ghosal, Sengupta, Majumder, & Sinha, 2020) proposed AutoRegressive Integrated Moving Average (ARIMA) model is among the numerous statistical methods that are used to forecast disease events. In multiple disease outbreaks, many researchers have used the ARIMA model for disease prevalence / incidence prediction. These include (Guan, Huang, & Zhou, 2004; Earnest, Chen, Ng, & Sin, 2005; Gaudart, et al., 2009; Liu, Liu, Jiang, & Yang, 2011; Nsoesie, Beckman, Shashaani, Nagaraj, & Marathe, 2013; Zheng, Zhang, Zhang, Wang, & Zheng, 2015; Fang, et al., 2020; Polwiang, 2020; Cao, et al., 2020)

Hence, this study is intended to compared six (6) statistical and deep learning-stimulated time-series models to evaluate the proportion of active coronavirus cases with reverence to the entire populace for the thirty-seven (37) states with the most active cases with multi step forecasting of 14 days (2 weeks). Specifically, six time-series models namely ARIMA (Box & Jenkins, 1990), the Holt–Winters Additive model (HWAAS) (Chatfield, The Holt-Winters Forecasting Procedure, 1978), TBAT (De Livera, Hyndman, & Snyder, 2011), Facebook's Prophet (Taylor & Letham, 2018), DeepAR (Salinas, Flunkert, Gasthaus, & Januschowski, 2020) but implemented as GluonTS (Alexandrov, et al., 2019) and N-BEATS (Oreshkin, Carpov, Chapados, & Bengio, 2019) were applied and their performances were accessed based on Root Mean Square Error (RMSE) and Mean Absolute Error (MAE) to find the optimum model that can predict future infections with least error for active cases in Nigerian states. These models were evaluated on publicly available COVID-19 dataset from NCDC website from 2/23/2020 to 11/4/2020. As far as we could possibly know, this is one of the few studies to compare time series models on COVID-19 active cases on all states in Nigeria. Expectantly, our addition to disease forecasting would be aidful to researchers, health practioners and the government.

Section 2 presents related work time-series prediction while section presents the methodology adopted for this study, the dataset used and the description of ARIMA, HWAAS, TBAT, Prophet,