International Journal of Inventive Engineering and Sciences (IJIES)
ISSN: 2319-9598 (Online), Volume-11 Issue-6, June 2024

A Review Study on Outbreak Prediction of Covid-19 By using Machine Learning

Kamal Narayan Kamlesh

Abstract: ***In December 2019, Wuhan City, China, discovered a new infectious disease, COVID-19. Over 70 million people have been infected, and one million people have died as a result of COVID-19. Defeating such a deadly, contagious disease requires accurate models that predict COVID-19 outbreaks. Using prediction models, governments can plan budgets and facilities for combating diseases, and take control measures to make informed decisions and take effective action. For example, they can determine the quantity of medicines and medical equipment to manufacture or import, as well as the number of medical personnel required to combat the disease. The COVID-19 outbreak has subsequently been predicted in several countries and continents using regression and classification models. A recent study that incorporated statistical and machine learning techniques was reviewed to predict future COVID-19 outbreaks. Ground truth datasets are used, their characteristics are investigated, models are developed, predictor variables are identified, statistical and machine learning methods are applied, performance metrics are calculated, and finally, comparisons are made. By applying machine learning methods, the survey results suggest that we can make predictions about whether a patient will contract COVID-19, how outbreak trends will evolve, and which age groups will be most affected.***

Keywords: ***Machine Learning, Classification, Regression, COVID-19, ARIMA***

I. INTRODUCTION

SARS-CoV-2 is also known as COVID-19. Coronavirus causes infectious diseases like this. Over 71 million people have been infected and died from COVID-19, the first major respiratory pandemic in the history of humanity [1]. Symptoms of the pandemic began to appear on March 12, 2020, when the World Health Organization declared it a pandemic. For governments to make better decisions, COVID-19 outbreaks need to be predicted as accurately as possible. To combat disease effectively, the government can (a) set up a budget, (b) make sure that facilities are adequate before they become overcrowded, (c) select the amount of medical equipment and medicines to manufacture, and finally (d) determine how many doctors are needed. Other diseases, including dengue fever, H1N1 flu, and swine fever, have been studied for developing prediction models in the past decades [2].

Manuscript received on 12 April 2024 | Revised Manuscript received on 13 June 2024 | Manuscript Accepted on 15 June 2024 | Manuscript published on 30 June 2024.
**Correspondence Author(s)*

Kamal Narayan Kamlesh*, Research Scholar, Department of CS & IT, Magadh University, Bodh Gaya (Bihar), India. E-mail: kamal.kamleshjha@gmail.com, ORCID ID: 0000-0001-8237-2905

COVID-19's unique characteristics and unprecedented spread necessitated several measures that had never been taken before. Curfews were imposed nationwide, businesses and public places were closed, masks were required outside of homes, and stadiums were converted into temporary hospitals.

To predict COVID-19 outbreaks, two types of prediction models are used: regression and classification. In regression models, continuous quantities are predicted. To assign them to appropriate classes, data are classified either by single-class labels (i.e., binary labels) or by multi-class labels. A time series prediction involves analysing how values have evolved to predict future changes. The combination of regular regressions, regular classifications, time series-based regressions, and time series-based classifications is used to obtain regular regression models and regular classification models.

In this study, recent statistical and machine learning research has been conducted to predict the COVID-19 outbreak. A description of the statistical methods and machine learning algorithms used in each study is provided, along with performance metrics. In addition to describing these characteristics, the ground-truth dataset, statistical models, predictor variables, and machine learning methods are also described. Finally, we provide results and conclusions. Throughout the paper, the studies are organised into the following categories: Indian cases, Eurasian cases, African cases, American cases, and intercontinental cases. At the end of the paper, a scope for future research is presented.

II. ANALYSES OF INDIAN CASES

A study was conducted to predict India's situation during the COVID-19 pandemic outbreak. The COVID-19 outbreak in India was analyzed using enhanced versions of the SIR epidemic model [3]. Rather than assuming values for parameters, sophisticated approaches based on curve fitting were employed to analyse the available data, yielding more realistic results. For optimal results, the Government of India implemented Zone bifurcation to identify susceptible patients. Compared with the number of people who should be tested for HIV, not enough people have been tested for HIV in India. According to the model, India has a low Return on Investment (RoI) value and a low fatality rate compared to other countries. In a single day, it is predicted that there could be around 800000 active cases and over 35000 deaths due to the outbreak. Another researcher believes that the situation in the country is acceptable compared to many other regions [4]. Lockdowns and distancing laws could have prevented this if enacted at the right time.

The model has predicted COVID-19 to cause large numbers of deaths in India, even at lower values of R0, due to the relatively high population. The models assume that the same conditions will persist, which can lead to a mismatch between the number of tests and the number of confirmed cases if there are too few tests conducted and a large number of asymptomatic patients.

Humans are infected with the COVID-19 virus, which damages the lungs and causes illness. The COVID-19 virus has killed a lot of people around the world. Using neural coronavirus analysis to study five different tasks, we propose Support Vector Regression [5]. Rather than using simple regression lines, this work employs support vectors to enhance classification accuracy. The first step in minimising virus spread is to determine the extent of the virus, allowing the government and citizens to better prepare for future outbreaks. Based on this information, the government and citizens will be able to make appropriate plans to minimize its spread. We are more likely to gain insight into the effectiveness of mitigation measures, the actions taken so far, and the number of cases prevented as a result.

Finally, by completing the third task, we will be able to estimate the number of patients who are likely to recover from this disease, as vaccines and cures have not yet been discovered. This will enable us to predict how long it will take for the pandemic to end by estimating the recovery time for all patients. In the fourth task, we will observe and analyze the speed of the virus's spread between regions, determining which areas need more attention [6]. To prevent the spread of infection, the environment of an infected patient should be improved. A warning can also be provided to indicate whether humidity and high temperatures should be avoided or not. Pearson's correlation method can be used to determine if disease spread is correlated with the weather. In comparison to other regression methods, the proposed Support Vector Regression method performed well in all tasks, as analysed using Coronavirus data.

Several nations have been devastated by COVID-19 outbreaks. COVID-19 has, however, shown signs of recovery in India, with an 88% recovery rate. Senapati used piecewise linear regression to predict the prevalence of positive cases and recovery cases in five Indian states. Instead of simple linear regression, we have used piecewise linear regression in the proposed scheme [7]. Consequently, both cases are accurately predicted by the proposed scheme. Therefore, our model can be applied to other COVID-19 parameters, regardless of their origin or intended use. We will develop advanced machine learning and deep learning models to enhance our response to COVID-19 and other pandemics [8]. Implementing piecewise linear regression involves finding where the data should be partitioned. Using the data from the previous 7 days, we have heuristically observed the slope of the point in this paper to predict the next day. By identifying an optimal partitioning point in the future, we aim to minimise errors.

COVID-19 poses a threat to thousands of people worldwide. All governments in all countries should pay close attention to this disease to reduce its effects [9]. A study has been conducted to analyse and track the outbreak of this disease in the Indian region until May 10th, 2020, as well as to estimate the number of cases over the next three weeks. Government data collected between 29th January and 11th May was analyzed using a machine learning algorithm [10]. RMSLE was used to evaluate the models, and SEIR performed well with a value of 1.52, while regression achieved a value of 1.75. Regression and SEIR performed well with 2.01. Regression and SEIR performed well with 2.01. R0, which represents the rate of disease spread, was calculated to be 2.86. The expected number of cases is 176K–210K for the three-week test period [11]. Doctors and the government will benefit from this study in preparing their plans and strategies. A new coronavirus, COVID-19, was reported in Wuhan, China, on December 30, 2019.

This virus has an estimated fatality rate of 4.5%, but that increases to 8.1% in the 70-80 age group, and 14.9% in the 81+ age group. Diabetes, Parkinson's disease, and cardiovascular disease pose the most significant risk to elderly patients over 50 [12]. As a result of this disease, you may experience a fever, cough, shortness of breath, pneumonia, kidney failure, and even death. The virus spreads via respiratory droplets between people nearby, with an average of 1.5–3.5 infected individuals per patient; however, it is not considered an airborne illness. As there is no vaccine against COVID-19, flattening the epidemic curve and reducing the epidemic peak are critical to managing this pandemic [13]. Using data science and data mining, researchers can develop concrete plans and make accurate decisions by understanding viruses and their characteristics. To prevent future epidemics and create necessary infrastructure, facilities, and vaccines, more aggressive measures will be taken. In this study, a combination of regression and SEIR models is used to predict the number of cases and estimate the spread of the disease in India. Our empirical results were analysed using data from the government web portal (https://www.mohfw.gov.in/) and the Johns Hopkins University, USA. Data for the training period will be available on January 30, 2020. The test data will be available on May 11, 2020, and May 31, 2020. In the data, we include confirmed cases, deaths and recovered cases. Using data from Guhathakurata in his article, analyses and predicts COVID-19 cases in India [14]. There have been fewer than one COVID-19 case per million in India, which represents approximately 18.6% of the world's population.

Artificial intelligence has been utilised in the fight against the COVID-19 pandemic. The definition of Artificial Intelligence given by Tamhane et al. in Machine Learning, Natural Language Processing, and Computer Vision Applications [15]. Pattern recognition, prediction, and other tasks can be taught to computers using these models. This paper uses two machine learning techniques to predict COVID-19 outbreaks. Support vector regression and polynomial regression methods are used for data visualization and prediction. It would allow us to take precautions before COVID-19 spirals out of control, without losing control, by estimating future cases of the virus. In the fight against COVID-19, the use of artificial intelligence does require accurate data and human-computer interactions, but it has become a source of hope for the advancement of technology for the protection of mankind [16].

Retrieval Number: 100.1/ijies.E41240612523
DOI: 10.35940/ijies.E4124.11060624
Journal Website: www.ijies.org

Published By:
Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP)

Here, a prediction model is developed to predict the number of cases over the next 25 days. A polynomial regression model and a support vector regression model were used to obtain the results. The Python language simplifies the process of achieving desired results. As a result, it uncovers hidden insights that provide insights and predict future trends.

Table 1

Author's	Publishing Year	Model	Technique of the study	Predictor variables of the study	Target variables of the study
Kumar et. al.	2020	Nested Exponential Statistical Model	Machine Learning	Confirmed Cases	Total Cases
Gupta et. al.	2020	Regression Model	Machine Learning	Expected Cases	All Cases
Yadav et. al.	2020	Support Vector Regression Model	Machine Learning	Active Cases	Recoveries
Guhathakurata et. al.	2021	Multiple Criteria Decision-Making (MCDM) Technique	Machine Learning	Covid-19 Cases	Total Death
Senapati et. al.	2021	Linear Regression	Machine Learning	Confirmed Cases	Total Cases
Hussain et. al.	2021	Convolutional Neural Network	Deep Learning	X-ray and CT Scan	Non-COVID bacterial pneumonia

III. ANALYSIS OF EURASIAN CASES

COVID-19 spread in Indonesia can be predicted using different models [17]. Both ARIMA and PROPHET algorithms were used to create the model, which is an Autoregressive Integrated Moving Average (ARIMA) model. An ARIMA (p, d, q) formula is used to average a difference between two samples, where p is the lag time, d is the difference degree, and q is the size of the average window. Data for this study were obtained from the Kaggle website. They contain serial numbers, observation dates, provinces or states, countries or regions, last updates, confirmations, recoveries, and deaths. In the original dataset, 28.218 rows were included, but only 82 rows remained when only Indonesian data were included [18]. Since ARIMA uses only univariate data, the confirmed, recovery, and death data frames were separated. Afterwards, they checked if the data was stationary since ARIMA is better able to calculate results with stable data. According to their findings, there was no stationary trend in the data. As a next step, they transformed the data into a stationary form using a log-scale transformation and a time-shifting transformation. The prediction of future death counts was based on past confirmed, recovered, and recovered death counts [19].

Metrics such as coefficients of determination (R2), mean square errors (MSEs), mean absolute errors (MAEs), and mean forecast errors (MFEs) are used to measure program performance. Between April 22, 2020, and May 21, 2020, both models were evaluated over a period of 30 days. Both PROPHET and ARIMA are generally more accurate as more days are forecasted, but PROPHET is usually better than ARIMA. Using age, gender, nationality, and location as indicators of susceptibility, analyzed the demographics of individuals [20]. The Saudi Ministry of Health provided the dataset. Between January 2, 2020, and April 25, 2020, 229124 patient records in Saudi Arabia were included in the dataset. Cities, confirmatory results, dates of birth, genders, nationalities, subject identification numbers, result dates, screening results, and test dates are among the variables that predict the outcome. You can substitute the date of birth with any of the following age groups: 0-8, 9-20, 21-29, 30-39, 40-49, 50-59, 60-69, 70-79, or 80+. A total of 15 Saudi regions replaced the city. For the positive class, random oversampling was used because the negative test results were overwhelmingly high, resulting in a database with 51% positive results and 49% negative results. Using age, gender, nationality, and location variables, we predicted COVID-10 outbreaks [21]. Decision Trees (DTs), Random Forests (RFs), and Multi-Layer Perceptrons (MLPs) were the most popular models in this study. Five cross-validations were performed on the models during training and testing. In a survey of 2021, AUC (Area under the Curve) was used as a measurement of accuracy, precision, and recall [22]. A DT-based model has been determined to perform the best based on all performance metrics. In addition, only the mentioned predictor variables could be used to predict susceptibility without considering medical factors.

Among Italy, Spain, and France, COVID-19 is prevalent, according to Ceylan [23]. Ceylan collected data for Italy, Spain, and France from the WHO website between February 22, 2020, and April 15, 2020. A time series ARIMA regression model was created using 45 samples in this study. There are a variety of architectures used in the ARIMA model, where confirmed cases are used as predictors and total cases as dependent variables [24].

A forecasting model was developed by Fang et al. to prevent and control COVID-19 outbreaks. The coronavirus resource centre at Johns Hopkins University provided the dataset [25]. In the dataset, data are available from January 30, 2020, to June 30, 2020. This study utilised two categories of Russian cases: confirmed recovery cases and confirmed death cases.

ARIMA forecasting was used to develop time series regression models. The MAPE metric is preferable for verifying the fit of the ARIMA model [26]. With a MAPE value of 0.65, the forecast model for confirmed cases is highly accurate and robust. It has been reported that death forecast models have an MAPE value of 3.90, while recovery forecast models have an MAPE value of 2.40 [27]. These models have also been reported as being reasonably robust.

Table 2

Author's	Publishing Year	Model	Technique of the Study	Predictor Variables of the Study	Target Variables of the Study
Satrio et al.	2021	Time Series Regression	Prophet Forecasting Model	Death Confirmed, Recovered	Death Confirmed, Body Recovered
Althnian et al.	2020	Regular Classification	Decision Tree, MultiLayer Perceptron	Location, Age, Gender, and Nationality	Susceptibility
Ceylan	2020	Time Series Regression	Autoregressive Integrated Moving Average	Cases Confirmed	The Total Number of Cases
Fang et al.	2020	Time Series Regression	Autoregressive Integrated Moving Average	Death Confirmed, Recovered	Death Confirmed, Recovered
Fayyoumi et al.	2020	Regular Classificati	MultiLayer Perceptron	Test Result Positive For PCR, Age, Gender, Smoking, Positive X-Ray, Fever, Breathing	COVID-19 Positive or Negative
Gupta et al.	2020	Time Series Regression	Prophet Forecasting Model	Cases Confirmed, Death Cases, And Cases Recovered	Cases Confirmed, Death Cases, and Cases Recovered
Önder	2020	Time Series Regression	Sub-Exponential Growth	Cases Confirmed	Reproductive Number
Pinter et al.	2020	Time Series Regression	Multi-Layered Perceptron-Imperialist Competitive Algorithm	Cases and Deaths per Day	Mortality Rate and Number of Cases
Pourghasemi et.al.	2020	Time Series Classification	Support Vector Machines, Autoregressive Integrated Moving Average	Coldest Month's Maximum Temperature, Warmest Month's Maximum Temperature, Wettest Month's Precipitation, Distance from Banks, Distance from Bakeries, Distance from Attraction Sites	Areas with High or Low Risk
Tamhane et al.	2020	Time Series Regression	Polynomial Regression, Support Vector Machines	Cases Reported	Amount of Cases

For predicting the future occurrence of COVID-19, Fayyoumi developed a regular classification model based on machine learning and statistics [28]. This study utilised a dataset collected through an online survey conducted among various eligible Jordanians, from which 105 participants were selected for participation. Machine learning models were built with SVM and MLP, while statistical learning models were built with Logistic Regression [29]. The dataset also contains predictor variables such as the presence of a PCR test, age, gender, smoking status, x-rays showing a positive result, fever, bronchitis, diarrhoea, feeling unwell, nasal congestion, dry cough, loss of smell, aching and pains, and sore throat. Tenfold cross-validation was used to test all models. According to Kumar, this study aims to identify patients who are infected with COVID-19 [30]. The geometric mean (G_Mean), precision, and accuracy were assessed for the models. Due to the performance metrics used, MLP was selected as the best model.

In the Fars Province of Iran, Pourghasemi investigated risk factors for COVID-19 using

Retrieval Number: 100.1/ijies.E41240612523
DOI: 10.35940/ijies.E4124.11060624
Journal Website: www.ijies.org

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)

an SVM model [31].

The growth trends of COVID-19 in Fars Province were also compared with those in other regions. Data from the time series was classified with the help of these models. A combination of ARIMA and SVM was used. Analyzing COVID-19 spread patterns using ARIMA models and SVMs helped create a risk map of the outbreak. PWM (Precipitation of Wettest Month), PDM (Precipitation of Driest Month) and MTCM (Minimum Temperature of Coldest Month) are predictor variables. Road distance, mosque distance, hospital distance, petrol station distance, bus station distance, bank distance, bakery distance, attraction distance, ATM distance, footprints of people, and density of towns. SVM coefficients were analyzed using accuracy metrics, AUC, and t-statistics, whereas ARIMA coefficients were analyzed using standard error, t-statistics, and probability metrics. MTCM and village density moderately influence COVID-19 risk mapping, but ATMs, distances from attractions, fuel stations, mosques, and roads are least affected by climate factors. There is a strong correlation between MTCM and distance from bus stations, bakeries, and hospitals [32]. The infection isn't expected to demonstrate a rapid explosion, but the general trend may continue for some time, according to ARIMA.

To predict the number of cases for the next 25 days, Yeung used a time series regression model [33]. Using dates between January 1, 2020, and May 25, 2020, John Hopkins University provided the dataset. A polynomial regression (PR) model was constructed using SVM and Polynomial Regression (SVM), while the number of cases was used as a predictor [34]. To analyze the data, we divide it into two sets: test data and train data. MAE measures a performance metric, and MSE measures a performance metric. A critical tool for dealing with the crisis may be machine learning, as demonstrated in this study.

IV. ANALYSIS OF AFRICAN CASES

A prediction of COVID-19's future trends in Ethiopia was made by Gebretensaeby using ARIMA [35]. A study of confirmed, recovered, and death cases in Ethiopia was conducted using data from the Ethiopian Public Health Institute's website. The data included in the dataset span from March 13, 2020, to August 31, 2020. A time-series regression model was constructed using the ARIMA model. Based on Box and Jenkins' description, the authors estimated model parameters using the least squares method after determining the appropriate ARIMA order. We used Partial Autocorrelation (PACF) and Autocorrection Function 8ACF to confirm the stationary nature of the data in confirmed cases [36]. ACF and PACF graphs were used to select ARIMAs (5, 1, 0) and (3, 1, 2) for predicting confirmed and recovered cases. The Bayesian information criteria (BIC) and RMSE are not the only performance metrics we examined. It is expected that Ethiopia will experience an increase in confirmed and recovered cases within the next 60 days, as indicated by the ARIMA (5, 0, 0) and ARIMA (3, 1, 2) models.

A deep learning algorithm was used by Marzouk to predict COVID-19 outbreaks [37]. It is an essential source of confirmed, recovered, and fatal cases from Egypt, according to the Egyptian Ministry of Health and Population. Over 504 days have passed since February 13, 2020, when the dataset was collected, and it covers the period from February 13, 2020, to June 30, 2021. There was a 90 percent training portion of the data, and a 10 percent testing portion [38]. A time-series regression analysis was conducted in this study. As part of our model-building process, we utilised Multi-Level Perceptrons, Convolutional Neural Networks, and Long Short-Term Memory (LSTM). The past values of confirmed, recovered, and death counts were used to predict future counts. Performance metrics have been RMSE and R2. Based on their ability to capture nonlinear patterns in input data over time, LSTM models outperformed CNN and MLP models in predicting a week and a month.

In a study published by Anki, Ethiopian cases were used to predict COVID-19 spread [39]. A dataset for this analysis can be found in Johns Hopkins University's official GitHub repository. This dataset only contains data about Ethiopia. Based on data collected on January 25, 2020, the dataset includes the following information. A training and testing set was created based on the dataset. Training data comprised approximately two-thirds to three-fourths of the dataset, while the testing data made up the remainder. Time series regression models were developed using SVM and PR. A confirmed case, a recovered case, and a death case are the predictors and output variables. Performance evaluation was based on MAE and MSE metrics. To make predictions for COVID-19, a deep learning model utilising LSTMs, GRUs, and Bi-LSTMs is employed. Journal of Chaos, Solitons, and Fractals, 141, 120232. Therefore, both confirmations, recovered cases, and death cases perform better with SVM than PR.

In Algeria, new COVID-19 cases were predicted by Ballı [40]. Using the Algeria health ministry's public health database, we obtained the dataset used in this study. An Extreme Learning Machine (ELM) method is used to build a time series regression model. New cases for COVID-19 were calculated using the cumulative number of confirmed cases, as well as the index day. The model generates a new COVID-19 case. Several performance metrics were preferred, including MSE, RMSE, MAE, NSE, OI, and R2, as well as Nash–Sutcliffe coefficients of efficiency (NSE). It has been concluded that new COVID-19 cases can be predicted using the ELM architecture.

Authorities can make better decisions with the help of the short-term prediction model developed by Saba [41]. Based on data collected between March 10, 2020, and May 15, 2020, a dataset was obtained from the Egyptian Ministry of Health and Population. To develop time series regression models, the NARANN and ARIMA algorithms were employed. A predictor variable was based on reported cases, and a target variable was based on new cases [42]. Model performance was evaluated using MAE, RMSE, R2, and coefficient of residual mass (CRM) analyses. Using NARANN as a predictor of COVID-19 case numbers outperforms ARIMA.

Table 3

Author's	Publishing Year	Model	Technique of the study	Predictor variables of the study	Target variables of the study
Gebretensae et al.	2021	Time Series Regression	Autoregressive Integrated Moving Average	Recovered, Confirmed,	All Confirmed, Recovered
Marzouk et al.	2021	Time Series Regression	Long Short-Term Memory	Confirmed, Recovered, Death	Confirmed, Recovered, Death
Ahmed	2020	Time Series Regression	Support Vector Machines	Death Confirmed, Recovered	All Confirmed, Recovered, and Dead
Djeddou et al.	2020	Time Series Regression	ELM	New COVID-19 Cases	Total Cases
Saba et al.	2020	Time Series Regression	NARANN	Cases Reported	New Cases
Takele	2020	Time Series Regression	Autoregressive Integrated Moving Average	Confirmed Cases	Total Confirmed Cases

V. ANALYSIS OF AMERICAN CASES

According to Luo, COVID-19 is predicted to be on the rise in the US shortly [43]. We obtained this time-series dataset from the WHO website. As we initiated no isolation or treatment measures between January and March, new confirmed cases were considered 19 or not between April 2, 2020, and September 29, 2020. The performance metrics used were precision, accuracy, recall, and AUC. Based on the results, MLP outperformed other methods.

COVID-19 had been predicted to spread across countries in the Americas by Jojoa [44]. An open data repository of the European Union was used to download the dataset. To build the models, we employed regression-based prediction models utilising MLPs and SVMs. As predictors and output variables, COVID-19 confirmed cases were used. In this study, Pearson's correlation coefficients, mean percent errors, and mean percent errors (MPEs) were used [45]. Using an optimization algorithm to determine the hyperparameters improves MLP performance; however, if not, SVM must be used. In Chile, Mexico, and the United States, MLP performed better than SVM. When the same performance metrics are applied to Brazil, Colombia, and Peru, SVM outperforms MLP.

Moreau predicted the COVID-19 pandemic in a study he conducted [46]. Our World in Data Project provided the dataset used in this study. Daily modelling objects are included in the dataset. Ninety per cent of the data were training data, while ten per cent was test data. With time-series regression, LSTM and XGBoost models were built. Based on past data, we forecast the number of confirmed cases over the next 25 days. Several performance metrics have been used, including MAE, MSE, RMSE, and MAPE. When the MAPE value is lower, LSTM performs better than XGBoost, indicating that LSTM outperforms XGBoost.

Using machine learning, Gomes attempted to detect COVID-19 early using early symptoms [47]. A total of 55.676 patients are included in the dataset, which was created in Brazilian. Regular classification techniques were employed in conjunction with RF, SVM, MLP, KNN, DT, Gradient Boosting Machine (GBM), and XGBoost. According to Mohammedqasim, sore throats, dyspnea, fever, cough, headache, taste disorders, olfactory disorders, and the presence of sore throats were all used as predictor variables. COVID-19 infection can be predicted through the use of this study. In addition to precision, accuracy, recall, and AUC, performance metrics were used. A comparison of MLP with other methods revealed that it performed the best [48].

Various countries in America are affected by COVID-19, according to Jojoa [49]. A dataset from the European Union's open data repository was used for the analysis. Models based on time series regression and SVM were constructed for the prediction process. A COVID-19 case study serves as a predictor and output variable in this study. The performance metrics used in this study were Pearson's Correlation Coefficients, Mean Percentage Errors (MPE), and Mean Absolute Errors (MAE). A hyperparameter optimization algorithm improves SVM performance, whereas it deteriorates MLP performance [50]. MLP performed better than SVM in Chile, Mexico, and the United States. An SVM performance metric outperforms an MLP performance metric when applied to Brazil, Colombia, and Peru.

Moreau's study predicted the COVID-19 pandemic. The dataset for this study was provided by the 'Our World in Data Project'. The dataset includes all confirmed cases and deaths each day [52]. Beginning with February 26, 2020, the first confirmed case in Brazil, this dataset includes data going back in time. Weibull Distribution models are derived using time series regression. We used the number of confirmed cases and deaths as predictors, with no new cases or deaths being recorded daily.

The diagnostic-death lag correlates with the mortality rate, so we use this parameter.

You can calculate this parameter by measuring the distance between new cases and deaths on a given day. A

Retrieval Number: 100.1/ijies.E41240612523
DOI: 10.35940/ijies.E4124.11060624
Journal Website: www.ijies.org

Published By:
Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP)

total of four scenarios were tested in this study. Using the daily death rate as a basis, these scenarios were developed [51].

A At the maximum point of the curve, the first, second, third, and fourth scenarios would have daily death rates of 1250, 1550, 1760, and 250, respectively. Performance was evaluated using R2 [53]. According to this author's study, the first scenario produced the most optimistic results, while the fourth scenario produced the most pessimistic outcomes. A similar pattern of results was found in the second and third scenarios, which were based on actual daily data. The similar precision of these scenarios makes it impossible to select one over the other as the most likely. According to the diagnostic-death lag value, there are between 1750 and 2000 deaths per day in a prospective scenario.

Forecasting models were updated to include exogenous climatic variables by Silva [54]. Datasets for confirmed cases include cumulative cases for five states in the USA and Brazil. Data is available in the datasets up to May 29, 2020. An API was used to collect COVID-19 information about all Brazilian states.

The National Centres for Environmental Information (NCEI) of the National Oceanic and Atmospheric Administration, along with Johns Hopkins University's climate data set, were used to analyse the climate in the United States. The Brazilian National Institute of Meteorology provided meteorological data for this study. Precipitation and temperature minimums and maximums were exogenously input into each model. In this study, a variety of time series regression models have been developed, such as Bayesian regression neural networks (BRNNs), Cubist regression neural networks (CUBISTs), Kakura neural networks (KNNs), Quantile Random Forests (QRFs), SVRs, and variational mode decompositions (VMDs). Climate variables such as rainfall, maximum temperature, and minimum temperature were predictive of COVID-19 cases [55]. The output variable is the cumulative confirmations.

Table 4

Author's	**Publishing Year**	**Model**	**Technique of the study**	**Predictor variables of the study**	**Target variables of the study**
Luo et al.	2021	Time Series Regression	Long Short-Term Memory	Confirmed Cases	Cases Confirmed
Santana et al.	2021	Regular Classification	MultiLayer Perceptron	Health Professional, Gender, Sore Throat, Dyspnea, Fever, Cough, Headache, Taste Disorder, Olfactory Disorder	Positive or Negative COVID-19
Jojoa et al.	2020	Time Series Regression	MultiLayer Perceptron, Support Vector Machines	Confirmed Cases	Confirmed Cases
Moreau	2020	Time Series Regression	Weibull distribution	Confirmed New Cases, Confirmed New Deaths	Daily New Cases, Daily New Deaths
Silva et al.	2020	Time Series Regression	BRNN, CUBIST, KNN, QRF, SVR, VMD-based Models	COVID-19 Cases, Precipitation, Maximum Temperature, Minimum Temperature	Cumulative Confirmed Cases
Souza et al.	2020	Regular classification	LogReg, LDA, XGBoost, Support Vector Machines	Symptoms Include Fever, Respiratory Distress, Coughing, Runny Nose, Sore Throat, Diarrhea, Headache, Pulmonary Disease, Cardiovascular Disease, Kidney Disease, Diabetes, and Smoking.	Is the Patient Recovered or Dead?
WollensteinBetech et al.	2020	Regular Classification	Support Vector Machines, LogReg, RF, XGBoost	Age, Pregnancy, Chronic Renal Failure, Diabetes, Immunosuppression, Chronic Obstructive Pulmonary Disease, Obesity, other, Hypertension, Tobacco Use, etc.	Hospitalization, Mortality, ICU Need, Ventilator Need

Retrieval Number: 100.1/ijies.E41240612523
DOI: 10.35940/ijies.E4124.11060624
Journal Website: www.ijies.org

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)

VI. ANALYSIS OF INTERCONTINENTAL CASES

In their work, Ayyubi investigate gated recurrent units and long short-term memories (LSTMs). COVID-19 outbreak forecast models proposed [56]. From the WHO website, several columns were included in the dataset, including dates reported, country codes, countries, regions, cumulative cases, and cumulative deaths. For the development of the models, data from Australia and Iran were combined. A period of data for the Australians covers January 25, 2020, to August 18, 2020, and a period of data for the Iranians covers January 2, 2020, to October 5, 2020 [57]. Approximately 71% of the data came from training, while the remaining 29% came from testing. Our models were built by combining bidirectional extensions and time-series regression with GRU, LSTM, and Conv-LSTM. Based on new cases, cumulative cases, new deaths, and cumulative deaths, we predict future trends in these metrics. The performance metrics used were MSLE, MAPE, RMSLE, and Explained Variance (EV). The performance of bidirectional models is usually better than that of non-bidirectional models most of the time. Moreover, the best-performed prediction method depends on the scenario, one, three, and seven days ahead, so no technique may always produce the best predictions. An outbreak of COVID-19 is predicted in any country using a model developed by Bala [58]. Data on COVID-19 deaths and confirmed cases are available from All Our World in Data (OWID). Published by Oxford University under the supervision of the European CDC, this dataset is updated daily. We developed time series regression models based on the date of each case and the number of cases. There was approximately 69:31 training time and 31 percent testing time for this database. XGBoost Regression was built using linear regression, SVM, and random forest. A performance evaluation metric was based on MAE and RMSE. According to the test results, XGBoost performed best when compared to RF, SVM, and Linear Regression (LinReg).

To help humanity survive the COVID-19 pandemic, Hassan proposed a prediction model for outbreaks of COVID-19 [59]. Johns Hopkins University provided the dataset on its GitHub repository. Between the beginning of the pandemic and October 15, 2020, daily data were collected on confirmed, recovered, and euthanasia cases worldwide. Our training dataset consisted of an initial training set of 225 days and an additional testing set of 39 days. In addition to MLP and SVM, Bayesian Networks (BN), Principal Regression (PR), and Linear Regression (LinReg) were used to build time-series regression models [60]. To predict future confirmation, recovery, and death counts, past values were used as inputs. R2 value, EV, RMSE, MAE, and MSE were used as performance metrics. MLP generally predicted confirmed cases more accurately than recovered cases and death cases, according to the results.

Using Apple Mobility Trends Reports, we compiled a dataset of mobility data. During the period of 13.01.2020 - 10.05.2020, data was collected to train the model. Data were collected for testing models between May 13, 2020, and May 20, 2020. The United Kingdom, the United States, Brazil, France, Germany, Italy, Spain, and the United Kingdom were among the countries analyzed. Data from a population database and a mobility index were used as inputs. A time series regression model, as well as Gaussian Process Regression, SVM, and decision tree models, were also developed. MAPE was used to assess the model's performance. The mobility index cannot predict daily cases across all countries using machine learning methods.

According to the above discussions, a cloud-based prediction model offers more realistic predictions in real-time. A daily WHO situation report is used to generate the dataset derived from Hannah Ritchie's "Our World in Data." This study was conducted on an Azure B1 single-core virtual machine with 1 GB of RAM and SSD storage, running Windows Server 2016 64-bit. Using the FogBus framework, the HealthFog framework performed multiple analysis tasks to predict various metrics. To develop time series regression models, Robust Weibull and Gaussian Fits were applied. New cases, new cases per year, and mortality rates are all determined by input and output variables. Metrics such as MSE, R2, and MAPE are used to evaluate performance. Weibull Robust models are more effective as performance evaluation metrics than Gaussian Fit models. A Robust Weibull model is found to work better than a Gaussian Fit model.

VII. CONCLUSION AND FUTURE SCOPE

This study summarises recent research on predicting the COVID-19 outbreak using machine learning and statistical methods. Numerous studies have been conducted recently to predict whether patients will contract COVID-19, which outbreak trend is most likely to occur, and which age groups are most susceptible to the disease. The four main types of regression models are linear regression models, semilinear regression models, regression models based on time series data, and classification models based on time series data. COVID-19 outbreaks can be predicted by analysing variables such as confirmed cases, recovered cases, and deaths, in addition to pre-existing conditions like asthma, obesity, hypertension, tobacco smoking, chronic renal insufficiency, diabetes, pregnancy, and demographics, including gender, nationality, and location. Additionally, it is essential to consider an individual's current health conditions, such as fever, breathing difficulties, diarrhoea, vomiting, loss of smell, nasal congestion, dry cough, headache, and sore throat, to predict COVID-19 infection accurately. Many performance metrics can be applied to a model; however, the most commonly used performance metrics are regression, MAE, MSE, MAPE, and RMSE. In contrast, the most widely applied performance metrics in classification are accuracy, precision, recall, and F1 scores. The MLP method yields relatively good results in most models, but cannot be ranked among these methods. A case count, a recovered case, and a death are the most common variables.

A COVID-19 mutation may impact the severity of future outbreaks, which is a potential area for future research. Additionally, the survey results indicate that most forecasting models are developed within a very short timeframe.

Retrieval Number: 100.1/ijies.E41240612523
DOI: 10.35940/ijies.E4124.11060624
Journal Website: www.ijies.org

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)

To build more accurate models, longer timeframes, such as six or nine months, can be considered in new studies. Additionally, future studies may consider developing new vaccines. Lastly, it would be beneficial to conduct further studies based on feature selection to identify relevant indicators of COVID-19.

DECLARATION STATEMENT

Funding	No, I did not receive.
Conflicts of Interest	No conflicts of interest to the best of our knowledge.
Ethical Approval and Consent to Participate	No, the article does not require ethical approval or consent to participate, as it presents evidence that is not subject to interpretation.
Availability of Data and Materials	Not relevant.
Authors Contributions	All authors have equal participation in this article.

REFERENCES

1. Kumari, R., Kumar, S., Poonia, R. C., Singh, V., Raja, L., Bhatnagar, V., & Agarwal, P. (2021). Analysis and predictions of spread, recovery, and death caused by COVID-19 in India. *Big Data Mining and Analytics*, *4*(2), 65-75. https://doi.org/10.26599/BDMA.2020.9020013
2. Ardabili, S. F., Mosavi, A., Ghamisi, P., Ferdinand, F., Varkonyi-Koczy, A. R., Reuter, U., ... & Atkinson, P. M. (2020). COVID-19 outbreak prediction with machine learning. *Algorithms*, *13*(10), 249. https://doi.org/10.3390/a13100249
3. Kumar, A., Rani, P., Kumar, R., Sharma, V., & Purohit, S. R. (2020). Data-driven modelling and prediction of COVID-19 infections in India, and correlation analysis of virus transmission with socio-economic factors. *Diabetes & Metabolic Syndrome: Clinical Research & Reviews*, *14*(5), 1231-1240. https://doi.org/10.1016/j.dsx.2020.07.008
4. Singh, S., Raj, P., Kumar, R., & Chaujar, R. (2020, July). Prediction and forecast for the COVID-19 Outbreak in India based on Enhanced Epidemiological Models. In *2020, the Second International Conference on Inventive Research in Computing Applications (ICIRCA)* (pp. 93-97). IEEE. https://doi.org/10.1109/ICIRCA48905.2020.9183126
5. Yadav, M., Perumal, M., & Srinivas, M. (2020). Analysis on novel coronavirus (COVID-19) using machine learning methods. *Chaos, Solitons & Fractals*, *139*, 110050. https://doi.org/10.1016/j.chaos.2020.110050
6. Temsah, M. H., Al-Sohime, F., Alamro, N., Al-Eyadhy, A., Al-Hasan, K., Jamal, A., ... & Somily, A. M. (2020). The psychological impact of the COVID-19 pandemic on health care workers in a MERS-CoV endemic country. *Journal of Infection and Public Health*, *13*(6), 877-882. https://doi.org/10.1016/j.jiph.2020.05.021
7. Senapati, A., Nag, A., Mondal, A., & Maji, S. (2021). A novel framework for COVID-19 case prediction through piecewise regression in India. *International Journal of Information Technology*, *13*(1), 41-48. https://doi.org/10.1007/s41870-020-00552-3
8. Hussain, E., Hasan, M., Rahman, M. A., Lee, I., Tamanna, T., & Parvez, M. Z. (2021). CoroDet: A deep learning based classification for COVID-19 detection using chest X-ray images. *Chaos, Solitons & Fractals*, *142*, 110495. https://doi.org/10.1016/j.chaos.2020.110495
9. Priyadarshini, I., Mohanty, P., Kumar, R., Son, L. H., Chau, H. T. M., Nhu, V. H., ... & Tien Bui, D. (2020, May). Analysis of the outbreak and global impacts of COVID-19. In *Healthcare* (Vol. 8, No. 2, p. 148). MDPI. https://doi.org/10.3390/healthcare8020148
10. Gupta, R., Pandey, G., Chaudhary, P., & Pal, S. K. (2020). Machine learning models for the government to predict the COVID-19 outbreak. *Digital Government: Research and Practice*, *1*(4), 1-6. https://doi.org/10.1145/3411761
11. Zhu, N., Zhang, D., Wang, W., Li, X., Yang, B., Song, J., ... & Tan, W. (2020). A novel coronavirus from patients with pneumonia in China, 2019. *New England journal of medicine*. https://doi.org/10.1056/NEJMoa2001017
12. Chan, A. K. Y., Tamrakar, M., Jiang, C. M., Lo, E. C. M., Leung, K. C. M., & Chu, C. H. (2021). Common medical and dental problems of older adults: a narrative review. *Geriatrics*, *6*(3), 76. https://doi.org/10.3390/geriatrics6030076
13. Anderson, R. M., Heesterbeek, H., Klinkenberg, D., & Hollingsworth, T. D. (2020). How will country-based mitigation measures influence the course of the COVID-19 epidemic?. *The Lancet*, *395*(10228), 931-934. https://doi.org/10.1016/S0140-6736(20)30567-5
14. Guhathakurata, S., Saha, S., Kundu, S., Chakraborty, A., & Banerjee, J. S. (2021). South Asian Countries are less fatal concerning COVID-19: a fact-finding procedure integrating machine learning & multiple criteria decision-making (MCDM) technique. *Journal of The Institution of Engineers (India): Series B*, 1-15. https://doi.org/10.1007/s40031-021-00547-z
15. Tamhane, R., & Mulge, S. (2020). Prediction of the COVID-19 outbreak using machine learning. *International Research Journal of Engineering and Technology (IRJET)*, *7*(5).
16. Khemasuwan, D., Sorensen, J. S., & Colt, H. G. (2020). Artificial intelligence in pulmonary medicine: computer vision, predictive model and COVID-19. *European respiratory review*, *29*(157). https://doi.org/10.1183/16000617.0181-2020
17. Satrio, C. B. A., Darmawan, W., Nadia, B. U., & Hanafiah, N. (2021). Time Series Analysis and Forecasting of Coronavirus Disease in Indonesia Using the ARIMA Model and the PROPHET Model. *Procedia Computer Science*, *179*, 524-532. https://doi.org/10.1016/j.procs.2021.01.036
18. van Hell, N. (2021). *Text Data Measured with Error: Empirical Strategies with an Application in S&P 500 Implied Volatility Forecasting* (Doctoral dissertation, Queen's University).
19. Çetin, U. A., & Fatih, A. B. U. T. (2022). A survey of recent studies on COVID-19 outbreak prediction using statistical and machine learning methods. *Niğde Ömer Halisdemir Üniversitesi Mühendislik Bilimleri Dergisi*, 1-1. https://doi.org/10.28948/ngumuh.1025095
20. Althnian, A., Abou Elwafa, A., Aloboud, N., Alrasheed, H., & Kurdi, H. (2020). Prediction of COVID-19 individual susceptibility using demographic data: A case study on Saudi Arabia. *Procedia Computer Science*, *177*, 379-386. https://doi.org/10.1016/j.procs.2020.10.051
21. Iwendi, C., Huescas, C. G. Y., Chakraborty, C., & Mohan, S. (2022). COVID-19 health analysis and prediction using machine learning algorithms for Mexico and Brazil patients. *Journal of Experimental & Theoretical Artificial Intelligence*, 1-21. https://doi.org/10.1080/0952813X.2022.2058097
22. Song, J., Gao, Y., Yin, P., Li, Y., Li, Y., Zhang, J., ... & Pi, H. (2021). The random forest model has the best accuracy among the four pressure ulcer prediction models using machine learning algorithms. *Risk Management and Healthcare Policy*, 1175-1187. https://doi.org/10.2147/RMHP.S297838
23. Ceylan, Z. (2020). Estimation of COVID-19 prevalence in Italy, Spain, and France. *Science of The Total Environment*, *729*, 138817. https://doi.org/10.1016/j.scitotenv.2020.138817
24. Papastefanopoulos, V., Linardatos, P., & Kotsiantis, S. (2020). COVID-19: a comparison of time series methods to forecast the percentage of active cases per population. *Applied sciences*, *10*(11), 3880. https://doi.org/10.3390/app10113880
25. Mohamadou, Y., Halidou, A., & Kapen, P. T. (2020). A review of mathematical modelling, artificial intelligence and datasets used in the study, prediction and management of COVID-19. *Applied Intelligence*, *50*(11), 3913-3925. https://doi.org/10.1007/s10489-020-01770-9
26. ArunKumar, K. E., Kalaga, D. V., Kumar, C. M. S., Chilkoor, G., Kawaji, M., & Brenza, T. M. (2021). Forecasting the dynamics of cumulative COVID-19 cases (confirmed, recovered and deaths) for top-16 countries using statistical machine learning models: Auto-Regressive Integrated Moving Average (ARIMA) and Seasonal Auto-Regressive Integrated Moving Average (SARIMA). *Applied Soft Computing*, *103*, 107161. https://doi.org/10.1016/j.asoc.2021.107161
27. Fang, L., Wang, D., & Pan, G. (2020). Analysis and Estimation of COVID-19 Spreading in Russia Based on an ARIMA Model. *SN Comprehensive Clinical Medicine*, *2*, 2521-2527. https://doi.org/10.1007/s42399-020-00555-y
28. Fayyoumi, E., Idwan, S., & AboShindi, H. (2020). Machine learning and statistical modelling for the prediction of novel COVID-19 patient cases: A case study from Jordan. *International Journal of Advanced Computer Science and Applications*, *11*(5). https://doi.org/10.14569/IJACSA.2020.0110518

Retrieval Number: 100.1/ijies.E41240612523
DOI: 10.35940/ijies.E4124.11060624
Journal Website: www.ijies.org

Published By:
Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP)

29. Erdem, E., & Bozkurt, F. (2021). A comparison of various supervised machine learning techniques for prostate cancer prediction. *Avrupa Bilim ve Teknoloji Dergisi*, (21), 610-620.
30. Kumar, A., Gupta, P. K., & Srivastava, A. (2020). A review of modern technologies for tackling the COVID-19 pandemic. *Diabetes & Metabolic Syndrome: Clinical Research & Reviews*, *14*(4), 569-573. https://doi.org/10.1016/j.dsx.2020.05.008
31. Pourghasemi, H. R., Pouyan, S., Farajzadeh, Z., Sadhasivam, N., Heidari, B., Babaei, S., & Tiefenbacher, J. P. (2020). Assessment of the outbreak risk, mapping and infection behaviour of COVID-19: Application of the autoregressive integrated-moving average (ARIMA) and polynomial models. *Plos one*, *15*(7), e0236238. https://doi.org/10.1371/journal.pone.0236238
32. Chimmula, V. R. (2021). *A Data-Driven Comparison Study of Statistical and Deep Learning Based Time Series Forecasting Methods for Infectious Disease Modelling and Financial Data*. The University of Regina (Canada).
33. Yeung, A. Y., Roewer-Despres, F., Rosella, L., & Rudzicz, F. (2021). Machine learning–based prediction of growth in confirmed COVID-19 infection cases in 114 countries using metrics of nonpharmaceutical interventions and cultural dimensions: model development and validation. *Journal of Medical Internet Research*, *23*(4), e26628. https://doi.org/10.2196/26628
34. Hassan, A. H. M., Qasem, A. A. M., Abdalla, W. F. M., & Elhassan, O. H. (2021). Visualization & prediction of COVID-19 future outbreak by using machine learning. *Int. J. Inf. Technol. Comput. Sci*, *13*(3), 16-32. https://doi.org/10.5815/ijitcs.2021.03.02
35. Gebretensae, Y. A., & Asmelash, D. (2021). Trend analysis and forecasting the spread of the COVID-19 pandemic in Ethiopia using the Box–Jenkins modelling procedure. *International journal of general medicine*, 1485-1498. https://doi.org/10.2147/IJGM.S306250
36. Al-Turaiki, I., Almutlaq, F., Alrasheed, H., & Alballa, N. (2021). Empirical evaluation of alternative time-series models for COVID-19 forecasting in Saudi Arabia. *International Journal of Environmental Research and Public Health*, *18*(16), 8660. https://doi.org/10.3390/ijerph18168660
37. Marzouk, M., Elshaboury, N., Abdel-Latif, A., & Azab, S. (2021). Deep Learning Model for Forecasting the COVID-19 Outbreak in Egypt. *Process Safety and Environmental Protection*, *153*, 363-375. https://doi.org/10.1016/j.psep.2021.07.034
38. Sánchez-Luna, M., Fernández Colomer, B., de Alba Romero, C., Alarcón Allen, A., Baña Souto, A., Camba Longueira, F., ... & SENEO COVID-19 Registry Study Group. (2021). Neonates born to mothers with COVID-19: data from the Spanish Society of Neonatology Registry. *Pediatrics*, *147*(2). https://doi.org/10.1542/peds.2020-015065
39. Anki, P., Bustamam, A., & Buyung, R. A. (2021). Looking for the link between the causes of the COVID-19 disease using the multi-model application. *Commun. Math. Biol. Neurosci.*, *2021*, Article-ID.
40. Ballı, S. (2021). Data analysis of the COVID-19 pandemic and short-term cumulative case forecasting using machine learning time series methods. *Chaos, Solitons & Fractals*, *142*, 110512. https://doi.org/10.1016/j.chaos.2020.110512
41. Elsheikh, A. H., Saba, A. I., Abd Elaziz, M., Lu, S., Shanmugan, S., Muthuramalingam, T., ... & Shehabeldeen, T. A. (2021). Deep Learning-Based Forecasting Model for the COVID-19 Outbreak in Saudi Arabia. *Process Safety and Environmental Protection*, *149*, 223-233. https://doi.org/10.1016/j.psep.2020.10.048
42. Santra, A., & Dutta, A. (2022). A Comprehensive Review of Machine Learning Techniques for Predicting the Outbreak of COVID-19 Cases. *International Journal of Intelligent Systems & Applications*, *14*(3). https://doi.org/10.5815/ijisa.2022.03.04
43. Luo, J., Zhang, Z., Fu, Y., & Rao, F. (2021). Time series prediction of COVID-19 transmission in America using LSTM and XGBoost algorithms. *Results in Physics*, *27*, 104462. https://doi.org/10.1016/j.rinp.2021.104462
44. Jojoa, M., Lazaro, E., Garcia-Zapirain, B., Gonzalez, M. J., & Urizar, E. (2021). The impact of COVID-19 on university staff and students from Iberoamerica: Online learning and teaching experience. *International Journal of Environmental Research and Public Health*, *18*(11), 5820. https://doi.org/10.3390/ijerph18115820
45. Jojoa, M., & Garcia-Zapirain, B. (2020). Forecasting COVID-19 confirmed cases using machine learning: The case of America. https://doi.org/10.20944/preprints202009.0228.v1
46. Moreau, V. H. (2020). Forecast predictions for the COVID-19 pandemic in Brazil by statistical modelling using the Weibull distribution for daily new cases and deaths. *Brazilian Journal of Microbiology*, *51*(3), 1109-1115. https://doi.org/10.1007/s42770-020-00331-z
47. Gomes, J. C., de Santana, M. A., Masood, A. I., de Lima, C. L., & Dos Santos, W. P. (2023). COVID-19's influence on cardiac function: a machine learning perspective on ECG analysis. *Medical & Biological Engineering & Computing*, 1-25. https://doi.org/10.1007/s11517-023-02773-7
48. Mohammedqasim, H., & Ata, O. (2022). Real-time data of COVID-19 detection with IoT sensor tracking using an artificial neural network. *Computers and Electrical Engineering*, *100*, 107971. https://doi.org/10.1016/j.compeleceng.2022.107971
49. Jojoa, M., Lazaro, E., Garcia-Zapirain, B., Gonzalez, M. J., & Urizar, E. (2021). The impact of COVID-19 on university staff and students from Iberoamerica: Online learning and teaching experience. *International Journal of Environmental Research and Public Health*, *18*(11), 5820. https://doi.org/10.3390/ijerph18115820
50. Yang, L., & Shami, A. (2020). On hyperparameter optimization of machine learning algorithms: Theory and practice. *Neurocomputing*, *415*, 295-316. https://doi.org/10.1016/j.neucom.2020.07.061
51. Brat, G. A., Weber, G. M., Gehlenborg, N., Avillach, P., Palmer, N. P., Chiovato, L., ... & Kohane, I. S. (2020). International electronic health record-derived COVID-19 clinical course profiles: the 4CE consortium. *NPJ digital medicine*, *3*(1), 109. https://doi.org/10.1038/s41746-020-00308-0
52. Moreau, V. H. (2020). Forecast predictions for the COVID-19 pandemic in Brazil by statistical modelling using the Weibull distribution for daily new cases and deaths. *Brazilian Journal of Microbiology*, *51*(3), 1109-1115. https://doi.org/10.1007/s42770-020-00331-z
53. Kafieh, R., Arian, R., Saeedizadeh, N., Amini, Z., Serej, N. D., Minaee, S., ... & Haghjooy Javanmard, S. (2021). COVID-19 in Iran: forecasting the pandemic using deep learning. *Computational and mathematical methods in medicine*, *2021*. https://doi.org/10.1155/2021/6927985
54. Da Silva, R. G., Ribeiro, M. H. D. M., Mariani, V. C., & dos Santos Coelho, L. (2020). Forecasting Brazilian and American COVID-19 cases based on artificial intelligence coupled with climatic exogenous variables. *Chaos, Solitons & Fractals*, *139*, 110027. https://doi.org/10.1016/j.chaos.2020.110027
55. Gomez-Cravioto, D. A., Diaz-Ramos, R. E., Cantu-Ortiz, F. J., & Ceballos, H. G. (2021). Data analysis and forecasting of the COVID-19 spread: A comparison of recurrent neural networks and time series models. *Cognitive Computation*, 1-12. https://doi.org/10.1007/s12559-021-09885-y
56. Ayoobi, N., Sharifrazi, D., Alizadehsani, R., Shoeibi, A., Gorriz, J. M., Moosaei, H., ... & Mosavi, A. (2021). Time series forecasting of new cases and new death rates for COVID-19 using deep learning methods. *Results in physics*, *27*, 104495. https://doi.org/10.1016/j.rinp.2021.104495
57. Cui, F., Salih, S. Q., Choubin, B., Bhagat, S. K., Samui, P., & Yaseen, Z. M. (2020). A newly explored machine learning model for river flow time series forecasting at Mary River, Australia. *Environmental Monitoring and Assessment*, *192*, 1-15. https://doi.org/10.1007/s10661-020-08724-1
58. Bala, S. (2021). COVID-19 Outbreak Prediction Analysis using Machine Learning. *International Journal for Research in Applied Science and Engineering Technology*, *9*(1). https://doi.org/10.22214/ijraset.2021.32690
59. Hassan, A. H. M., Qasem, A. A. M., Abdalla, W. F. M., & Elhassan, O. H. (2021). Visualization & prediction of COVID-19 future outbreak by using machine learning. *Int. J. Inf. Technol. Comput. Sci*, *13*(3), 16-32. https://doi.org/10.5815/ijitcs.2021.03.02
60. Cruz, R. C., Reis Costa, P., Vinga, S., Krippahl, L., & Lopes, M. B. (2021). A Review of recent machine learning advances for forecasting harmful algal blooms and shellfish contamination. *Journal of Marine Science and Engineering*, *9*(3), 283. https://doi.org/10.3390/jmse9030283

International Journal of Inventive Engineering and Sciences (IJIES)
ISSN: 2319-9598 (Online), Volume-11 Issue-6, June 2024

AUTHOR PROFILE

Kamal Narayan Kamlesh is pursuing a PhD at the Department of CS & IT, Magadh University, Bodh Gaya. He holds a B.Tech in IT and an M.Tech in Software Engineering from Jawaharlal Nehru Technological University, Hyderabad. He has over 7 years of experience in the teaching field. His areas of interest include Machine Learning, Computer Networks, Database, Artificial Intelligence, Data Science, and Deep Learning. He has published many papers in international journals indexed in various reputable databases.

Retrieval Number: 100.1/ijies.E41240612523
DOI: 10.35940/ijies.E4124.11060624
Journal Website: www.ijies.org

Published By:
Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP)