Int. J. Nonlinear Anal. Appl. In Press, 1–10
ISSN: 2008-6822 (electronic)
http://dx.doi.org/10.22075/ijnaa.2024.21537.2268

A model for predicting the trend deaths of COVID-19

Mehran Saeidi Aghdam[a,*], Sherrie X.Y. Komiak[b], Alireza Bahiraie[c], Madjid Eshaghi Gordji[c], Ahmad Sadeghi[d]

[a]*Department of Entrepreneurship, Qazvin Branch, Islamic Azad University, Qazvin, Iran*

[b]*Faculty of Business Administration, Memorial University of Newfoundland, St. John's, NL, Canada*

[c]*Department of Mathematics, Semnan University, Semnan, Iran*

[d]*Department of Geography, Faculty of Earth Sciences, Shahid Beheshti University, Tehran, Iran*

(Communicated by Javad Damirchi)

Abstract

The novel coronavirus pneumonia (COVID-19) originated in Wuhan and rapidly disseminated across China and subsequently the globe. This study aims to predict the trend of COVID-19-related deaths by optimizing the parameters of deep learning algorithms, particularly focusing on integrating big data. The performance of long-short-term memory (LSTM) learning models was rigorously compared with the auto-regressive integrated moving average (ARIMA) model to forecast future trends in COVID-19 fatalities. Through extensive data analysis and model optimization, the results indicate that the experimental results highlight the performance differences between the ARIMA and LSTM models in predicting COVID-19 outcomes. Specifically, the ARIMA model demonstrates superior performance with an accuracy of 87 percent, compared to the LSTM model's 79 percent accuracy. However, this does not mean ARIMA is unequivocally better than LSTM across all metrics. The findings suggest that the implementation of these predictive models can significantly improve the timeliness of reporting in existing surveillance systems, thereby enhancing public health responses and reducing societal costs associated with the pandemic. The study highlights the potential of using advanced predictive modelling to support healthcare planning and intervention strategies during global health crises.

Keywords: Prediction, LSTM, ARIMA, COVID-19
2020 MSC: 60G25

1 Introduction

The novel Coronavirus disease (COVID-19) was first reported on 31 December 2019 in Wuhan, Hubei Province, China. It started spreading rapidly across the world. The outbreak of COVID-19 has experienced three stages since mid-December 2019: local outbreak, community transmission, and large-scale transmission. In December 2019, a novel coronavirus was found in a seafood wholesale market in Wuhan, China. WHO officially named this corona-virus as COVID-19. Since the first patient was hospitalized on December 12, 2019, China has reported many confirmed COVID-19 cases and many deaths as of August 2020. Wuhan's cumulative confirmed cases and deaths accounted for 61.1 and 76.5 of the whole China mainland, making it the priority center for epidemic prevention and control. Meanwhile, many

*Corresponding author
Email addresses: mehransaedi@gmail.com (Mehran Saeidi Aghdam), skomiak@mun.ca (Sherrie X.Y. Komiak), alireza.bahiraie@yahoo.com (Alireza Bahiraie), meshaghi@semnan.ac.ir (Madjid Eshaghi Gordji), sadeghi_ahmad@yahoo.com (Ahmad Sadeghi)

Received: October 2020 *Accepted:* August 2024

countries and regions outside China have reported many confirmed cases and deaths as of August 2020. The COVID-19 epidemic does great harm to people's daily lives and the country's economic development [7]. Future research on coronaviruses will continue to investigate many aspects of viral replication and parthenogenesis. understanding the propensity of these viruses to jump between species, establish infection in a new host, and identify significant reservoirs of coronaviruses will dramatically aid in our ability to predict when and where potential epidemics may occur. As bats seem to be a significant reservoir for these viruses, it will be interesting to determine how they seem to avoid clinically evident diseases and become persistently infected [13]. During the past years, artificial intelligence (AI), the capability of a machine to mimic human behaviour, has become a key player in high-techs like predicting disease. AI tools help scientists uncover the secret behind the big bio-logical data using optimized computational algorithms. AI methods such as deep neural networks improve decision-making in biological and chemical applications [9]. Artificial intelligence (AI) is defined as the technology that uses computer knowledge to represent intelligent behaviour with nominal human involvement, and deep learning is considered as a subset of AI techniques. Usually, this kind of intelligence is commonly acknowledged as having begun with the innovation of robotics [5]. The applications of AI in medicine are developing quickly. In 2016, AI projects coupled with medicine drew in more speculation from the global economy than other projects [2]. This behaviour of COVID-19 requires developing a robust mathematical basis for tracking its spread and automation of the tracking tools for online dynamic decision-making. Predicting disease trends for a complex human disease using data is an important, yet challenging, step in personalized medicine. Among many challenges, the so-called curse of dimensional problems results in unsatisfied performances of many state-of-the-art machine learning algorithms. A major recent advance in machine learning is the rapid development of deep learning algorithms that can efficiently extract meaningful features from high-dimensional and complex datasets through a stacked and hierarchical learning process. Deep learning has shown breakthrough performance in several areas including image recognition, natural language processing, and speech recognition. However, the performance of deep learning in predicting disease trends using datasets is still not well studied [16]. There is a need for innovative solutions to develop, manage and analyze big data on the growing network of infected subjects, patient details, and their community movements, and integrate with clinical trials and, pharmaceutical, genomics, and public health data.

2 Literature

2.1 COVID-19

In 2019, the Centers for Disease Control and Prevention (CDC) started monitoring the outbreak of a new coronavirus, SARS-CoV-2, which causes respiratory illness now known as COVID-19. Authorities first identified the virus in Wuhan, China. More than 74,000 people have contracted the virus in China. Health authorities have identified many other people with COVID-19 around the world, including many in the United States. On January 31, 2020, the virus passed from one person to another in the U.S. The World Health Organization (WHO) have declared a public health emergency relating to COVID-19. Since then, this strain has been diagnosed in several U.S. residents. The CDC has advised that it is likely to spread to more people. COVID-19 has started causing disruption in at least 100 other countries. The first people with COVID-19 had links to an animal and seafood market. This fact suggests that animals initially transmitted the virus to humans. However, people with a more recent diagnosis had no connections with or exposure to the market, confirming that humans can pass the virus to each other [13].

2.2 Long Short-Term Memory (LSTM)

Long short-term memory (LSTM) is an artificial recurrent neural network (RNN) architecture used in the field of deep learning. Unlike standard feed-forward neural networks, LSTM has feedback connections. It can not only process single data points (e.g. images), but also entire sequences of data (such as speech or video inputs). LSTM models can store information over a period of time. LSTM is a type of model or structure for sequential data that has emerged from the development of RNNs and improved by Gers, Schmidhuber, and Cummins [5]. Long-term memory refers to learned weights and short-term memory refers to internal states of cells. LSTM was created for the vanishing gradient problem in RNNs whose main change is the replacement of the RNN mid-layer with a block that is called an LSTM block [6]. The main feature of LSTM is the possibility of long-term affiliation learning, which was impossible with RNNs. To forecast the next time step, it is required to update the weight values in the network, which requires maintenance of the initial time step data. An RNN could just learn a limited number of short-term affiliations; however, long-term time series, such as 1000 time steps, cannot be learned by RNNs; in contrast, LSTMs could properly learn these long-term affiliations [12]. The LSTM structure includes a set of recurrent sub-networks, called memory blocks. Each block includes one or more auto-regressive memory cells and three multiple units of 'input, output, and forgetting' that present the analogues of continuous writing, reading, and regulation of the cells'

functions. Moreover, there are various types of LSTM blocks, including stacked LSTMs, encoder-decoder LSTMs, bidirectional LSTMs, CNN LSTMs, and generative LSTMs [11]. Peephole connections allow the gates to access the constant error carousel (CEC), whose activation is the cell $h_{(t-1)}$ is not used, $C_{(t-1)}$ is used instead in most places.

$$f_t = \sigma_g(W_f x_t + U_f c_{t-1} + b_f)$$

$$i_t = \sigma_g(W_i x_t + U_i c_{t-1} + b_i)$$

$$o_t = \sigma_g(W_o x_t + U_o c_{t-1} + b_o)$$

$$c_t = f_t o c_{t-1} + i_t o \sigma_c(W_c x_t + b_c)$$

$$h_t = o_g o \sigma_h(c_t)$$

A peephole LSTM unit with input (i.e. i), output (i.e. o), and forget (i.e. f) gates. Each of these gates can be thought of as a "standard" neuron in a feed-forward (or multi-layer) neural network: that is, they compute an activation (using an activation function) of a weighted sum. i_t, o_t and f_t represent the activations of respectively the input, output and forget gates, at time step t The 3 exit arrows from the memory cell c to the 3 gates i.o and f represent the peephole connections. These peephole connections actually denote the contributions of the activation of the memory cell c at time step $t-1$, i.e. the contribution of $c_(t-1)$ and not c_t, as the picture may suggest). In other words, the gates i.o and f calculate their activation at time step t (i.e., respectively, $i_t.o_t$ and f_t also considering the activation of the memory cell c at time step $t-1$, i.e. $C_(t-1)$. The single left-to-right arrow exiting the memory cell is not a peephole connection and denotes c_t. The little circles containing an a$\times$ symbol represent an element-wise multiplication between its inputs. The big circles containing an S-like curve represent the application of a differentiable function (like the Sigmund function) to a weighted sum. There are many other kinds of LSTMs as well [10].

2.3 Auto regressive Integrated Moving Average (ARIMA)

ARIMA combines the Autoregressive (AR) process and Moving Average (MA) processes and builds a composite model of the time series. As the acronym indicates, ARIMA ($p.d.q$) captures the key elements of the model:1) Auto regression. A regression model that uses the dependencies between observation and several lagged observations (p). 2) I: Integrated. To make the time series stationary by measuring the differences of observations at different times (d).3) MA: Moving Average. An approach that takes into account the dependency between observations and the residual error terms when a moving average model is used to the lagged observations (q) [6]. A simple form of an AR model of order$p$, i.e., AR(p), can be written as a linear process given by:

$$x_t = c + \sum_{i=1}^{p} \emptyset_i x_{t-i} + \epsilon_t$$

where x_t is the stationary variable, c is constant, the terms in $\emptyset_i$ are auto correlation coefficients at lags 1, 2, ..., p and ϵ_t, the residuals, are the Gaussian white noise series with mean zero and variance σ_ϵ^2. An MA model of order q, i.e., $MA(q)$, can be written in the form:

$$x_t = \mu + \sum_{i=1}^{q} \theta_i \epsilon_{t-1}$$

where μ is the expectation of x_t (usually assumed equal to zero), the θ_i terms are the weights applied to the current and prior values of a stochastic term in the time series, and $\theta_0 = i$. This research assume that ϵ_t is a Gaussian white noise series with mean zero and variance σ_ϵ^2. It can combine these two models by adding them together and form an ARIMA model of order ($p.q$) [14]:

$$x_t = c + \sum_{i=1}^{p} \emptyset_i x_{t-i} + \epsilon_t + \sum_{i=1}^{q} \theta_i \epsilon_{t-1}$$

3 Methodology

In research, the methodology consists of five steps. Stage 1-Raw Data: In this stage, the historical data is collected from: https://ourworldindata.org/coronavirus-source-data. The data are from 2020/12/31 till 2020/8/3.and this historical data is used for the prediction of future deaths. Stage 2- Data Reprocessing: The prepossessing stage involves a) Data discretization: Part of data reduction but with particular importance, especially for numerical data

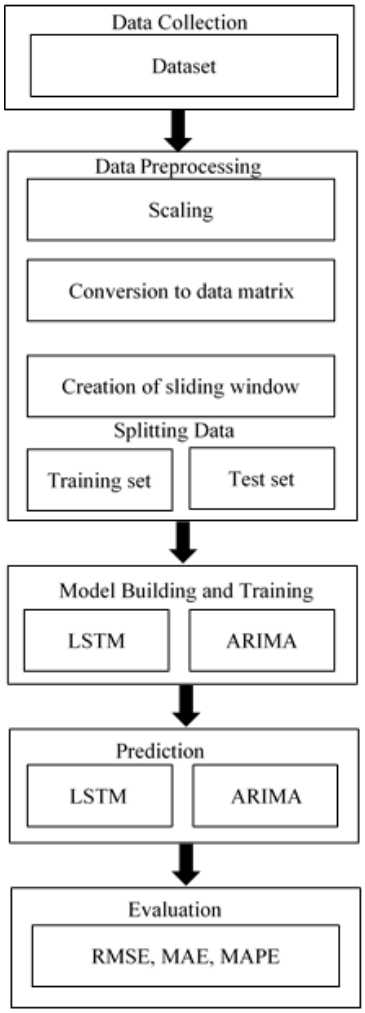

Figure 1: The general process of research model Experimental procedure

b) Data transformation: Normalization. c) Data cleaning: Fill in missing values. d) Data integration: Integration of data files. After the dataset is transformed into a clean dataset, the dataset is divided into training and testing sets so as to evaluate. Creating a data structure with 60 time steps and 1 output. After the dataset is transformed into a clean dataset, the dataset is divided into training and testing sets so as to evaluate. Stage 3- Feature Extraction: In this layer, only the features which are to be fed to the neural network are chosen. In this research will choose the feature from Date and deaths number. Stage 4-Training: In this stage, the data is fed to the neural network and trained for prediction assigning random biases and weights with ARIMA and LSTM models. Stage 5- Output Generation: In this layer, the output value generated by the output layer of the ARIMA and LSTM is compared with the target value evaluation by RMSE, MAE and MAPE. apply the model and develop the predicted approach. The flow chart of the steps involved in the proposed method for each model is shown in Figure (1).

4 Data and Models

Traditionally most machine learning (ML) models use as input features some observations (samples/examples) but there is no time dimension in the data. Time-series forecasting models are the models that are capable of predicting future values based on previously observed values. Time-series forecasting is widely used for non-stationary data. Non-stationary data are called the data whose statistical properties e.g. the mean and standard deviation are not constant over time but instead, these metrics vary over time. These non-stationary input data (used as input to these models) are usually called time series. Some examples of time series include the temperature values over time, stock price over time, price of a house over time, etc. So, the input is a signal (time series) that is defined by observations taken sequentially in time. A time series is a sequence of observations taken sequentially in time. dataset: this research

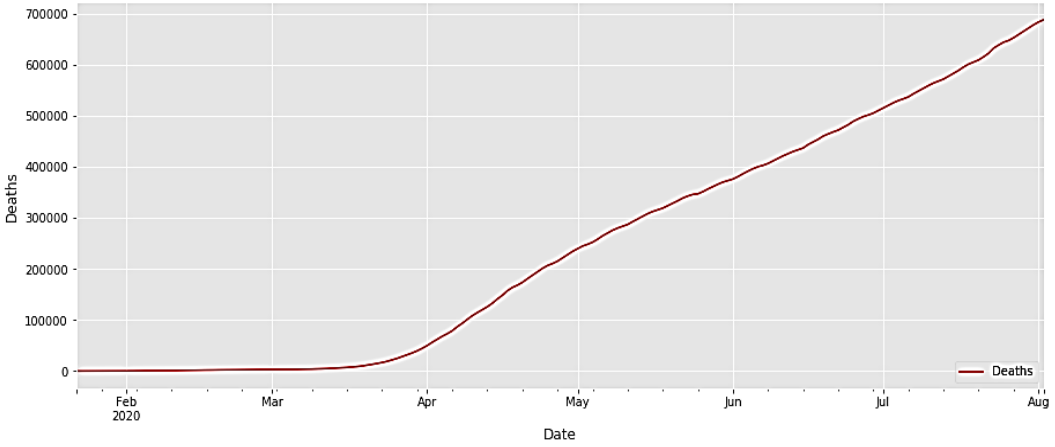

Figure 2: Time series plot of COVID-19 cases

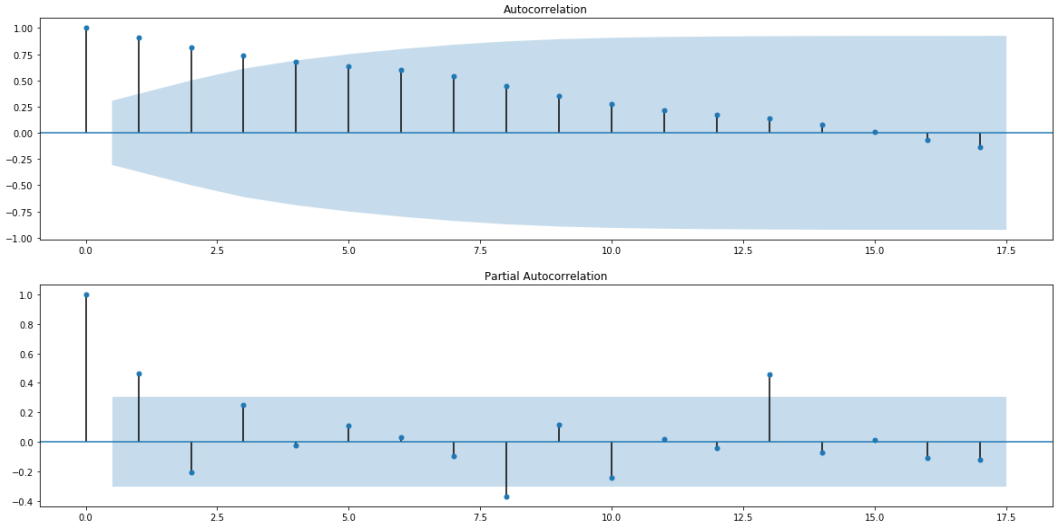

Figure 3: Auto correlation and Partial auto correlation factor

acquired the data from https://ourworldindata.org/coronavirus-source-data. this research has collected the historical data of COVID-19. Data ranges from 2020/12/31 to 2020/8/3.

Sequence data: there are 194 sequences from 2020/12/31 to 2020/8/3. From this dataset, it used 135 samples for training purposes and 59 samples for validation purposes. Training Detail: For training, the model in research used LSTM and ARIMA algorithm. For the experiment, it used various sets of parameters with a different number of epochs to measure the RMSE of the Training and Testing dataset. Comparison algorithms: in this stage compare LSTM and ARIMA algorithms. ARIMA: Data regarding the number of cases reported in COVID-19 till 02nd Aug 2020. This data was plotted on a graph to see the trend, as shown in Figure (2). A partial autocorrelation is a combination of the relationship between an observation in a time series with observations being excluded at the initiation phase with the relationships of intervening observations. The Auto Correlation Function given in Figure (3) shows that the series has positive autocorrelations to a large number of lags, i.e., 10, so a higher order of differentiation is required.

Figure (5) shows that the autocorrelation of lag-1 is small and patternless, so the series does not need a higher order of differentiation. If the autocorrelation of lag-1 is zero or more negative, then the series may be over-differentiated. The partial autocorrelation function of the differences series shows a sharp cut-off due to the positive lag-1 autocorrelation, and the series appears to be slightly under-differentiated, so one or more AR terms should be added to the model. The lag beyond which the partial autocorrelation function cut off is the number of AR terms indicated. The Auto Correlation function and partial autocorrelation function cut-off showed an irregular increasing pattern in the number of cases of COVID-19. Henceforth, ARIMA models (p, d, q), apt for such a scenario was applied. In terms of choosing a Box-Jenkins model, the smaller the goodness-of-fit measures the better. The best suitable Box-Jenkins model was

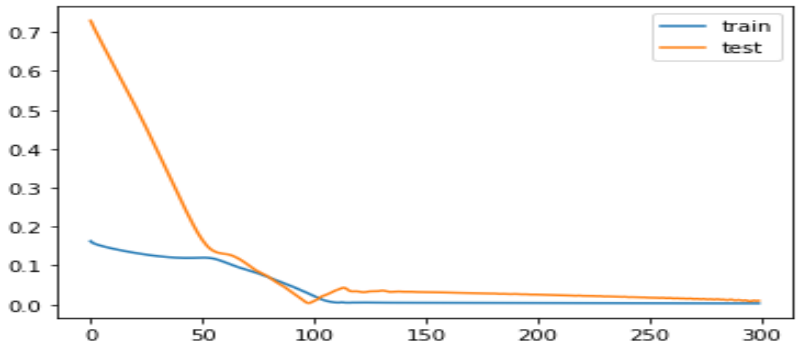

Figure 4: Training and Validation Loss of LSTM Model

selected based on minimal Bayesian Information Criteria (BIC) value. In this study, the least BIC value is 3202.294 as given below in Table 1, and the corresponding model is ARIMA $(1, 1, 0)$ with the overall highest R^2 values of 0.95.

Table 1: Model selection

ARIMA(p.d.q)	BIC	R-Squared
1,0,0	3624.157	0.18
1,1,0	3202.294	0.95
1,1,1	493.036	0.62
0,1,1	378.224	0.24
0,1,0	366.371	0.28

It is evident from the figure that all the lags are well within the 95 confidence level. This implies that residuals are random, i.e., white noise, indicating that the model is a good fit. It is also observed that all autocorrelation coefficients are not statistically significant, implying that residuals are not autocorrelated with each other. A model with the lowest value of normalized BIC is found to be ARIMA $(1, 1, 0)$, which can be considered as the best-fit model and can be further used to generate the forecasts. Based on The Auto Correlation function and partial autocorrelation function cut-off, the daily prediction of COVID-19 cases is calculated, as shown in Figure (4). LSTM: LSTM is widely used for sequence prediction problems and has proven to be extremely effective. The reason they work so well is because LSTM can store past important information and forget the information that is not. LSTM has three gates:

(1) The input gate: The input gate adds information to the cell state.

(2) The forget gate: It removes the information that is no longer required by the model.

(3) The output gate: The output Gate at LSTM selects the information to be shown as output.

The LSTM layer is added with the following arguments: 50 units are the dimensionality of the output space, return sequences=True is necessary for stacking LSTM layers so the consequent LSTM layer has a three-dimensional sequence input, and input shape is the shape of the training dataset. Specifying 0.2 in the Dropout layer means that 20 percent of the layers will be dropped. Following the LSTM and Dropout layers, it adds the Dense layer that specifies an output of one unit. To compile the model, it uses the Adam optimizer and sets the loss as the mean squared error. After that, it fits the model to run for 300 epochs (the epochs are the number of times the learning algorithm will work through the entire training set) with a batch size of 32. results represent the plots of epochs vs. loss for the LSTM network. The validation loss is found to be less than the training loss at some initial epochs because it is calculated at the end of each epoch and before the start of a new epoch when the model is not optimised. The comparison results of the predicted and observed flow data for a one-day forecasting model are shown in Figure (4).

Next stage, visualize the result of our predicted value and the actual value. While the exact value points from predicted price weren't always close to the actual value, this model did still indicate overall trends such as going up or down. the result showed the LSTMs can be some-what effective in times series forecasting.

Evaluation measures: The accuracy, precision, recall, F-value, AUC, APRS, and MCC are calculated from the

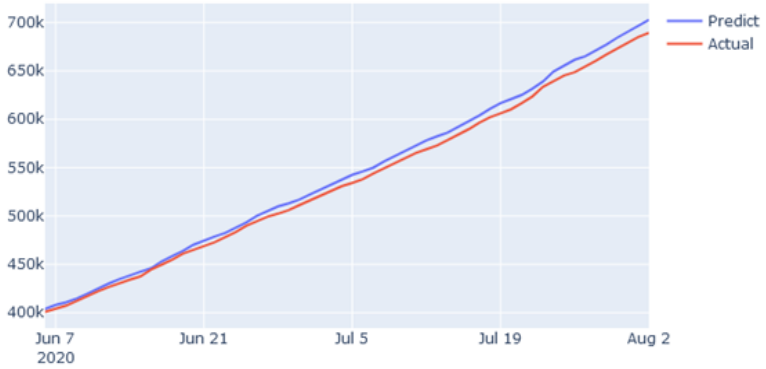

Figure 5: Comparison between the predict value and actual value

confusion matrix.

Table 2: Comparison between the predict value and actual value

Algorithm	Accuracy	Precision	Recall	F-score	AUC	APRS	MCC
LSTM	0.79	0.72	0.871	0.834	0.972	0.934	0.776
ARIMA	0.87	0.74	0.962	0.842	0.941	0.759	0.758

The experimental results show that the ARIMA model has the best performance in outcome predictions, and the accuracy reaches 87 percent. while LSTM's accuracy is 79 percent. There are huge differences between the advantages and disadvantages of algorithms. From the comparison of advantages, the ARIMA model is not always superior to the LSTM algorithm. Among evaluation indicators, the ARIMA model has advantages in accuracy, recall, and F-score but LSTM shows advantages in AUC, APRS and MCC. The result indicates that ARIMA has a stronger power to predict positive cases, and LSTM has shown significant advantages in precision. Mean Absolute Percentage Error (MAPE), Mean Absolute Error (MAE), Root Mean Squared Error (RMSE) and Mean Squared Error (MSE) are used to evaluate the performance of these prediction models. Formulas of these evaluation measures are shown in Eqs.

$$\text{MAPE} = \frac{1}{N}\sum_{t=1}^{n}\left|\frac{A_t - F_t}{A_t}\right| \times 100$$

$$\text{MAE} = \frac{1}{N}\sum_{t=1}^{n}\left|\frac{A_t - F_t}{A_t}\right|$$

$$\text{RMSE} = \sqrt{\frac{1}{N}\sum_{t=1}^{n}\left(\frac{A_t - F_t}{A_t}\right)^2}$$

The value of MSE (Mean Square Error) is used as the loss function for models used in the experiment. The metrics used for the evaluation of the models are Mean Absolute Error (MAE) and Mean Absolute Percentage Error ($MAPE$). $RMSE$ values obtained for the Trend of COVID-19 after evaluating each model are shown in Table 2.

Table 3: Performance statistics of ARIMA and LSTM models

Model	RMSE	MAPE	MAE
LSTM	10.2453	0.0114	7.3652
ARIMA	10.0834	0.0057	6.9647

The average performance criteria for each model were calculated and are presented in Table 1. The performance criteria $RMSE$, $MAPE$, and MAE obtained by the ARIMA model were calculated as 10.2543, 0.0114, and 7.3652. Performance indexes calculated by LSTM are 10.0834, 0.0057, and 7. 6.9547 for $RMSE$, $MAPE$, and MAE respectively. Theoretically, a forecasting model is regarded as good when $RMSE$, and MAE are small. It can be seen from Table 3 that the ARIMA model has smaller errors than those of the LSTM model. The performance criteria indicate that the assessed result is highly correlated and precise. Expected utility theory: The theory of Prospect was introduced by Kahneman, and Tversky [8] with the following key components: And (2) nonlinear transformation of the probability scale, which gives more weight to small probabilities and weighs less to high and medium probabilities. These components lead to risk aversion in the case of profit and risk-taking in the case of loss. Tversky and Kahneman [15] state that individuals behave differently depending on profit and loss, so the behaviour of investors may change in different market conditions, which also means that the beta is variable. If investors are risky at the time of the loss, the beta is expected to be higher in a downtrend. In addition, if investors are risk averse at the time of profit, the beta is expected to be lower in bullish conditions. If the beta exhibits temporal or pattern-dependent behaviour, it is possible that the relationship between risk and return is not as positive as is suggested in classical financial knowledge and is negative for certain periods. Hypothesis 1: The behaviour of COVID-19 negatively affects risk. Hypothesis 2: The different behaviour of COVID-19 affects the risk beta. In this research, first, the following quantile regression model is used to test the first and second hypotheses:

$$r_{it} = a_i + b_i r_{Mt} + V_{it}$$

$$Q_r(\tau | r_{it}) = a_i(\tau) + b_i(\tau) r_{Mt}$$

Thus, $Q_r(\tau|r_{it})$ represents the conditional quantile of the company's return r_{it} at quantile τ, which is assumed to be linearly dependent on the market return r_{Mt}. The model is estimated using the quarterly regression method, and as a result, the effect of r_{Mt} on different quantiles of r_{it} can be evaluated. In other words, the effect of market conditions specific to each company on the return can be examined. Also, to test the third hypothesis of the research, the Merton ICAPM (1980) model is used to extract risk information.

$$E_{t-1}[r_t] = \gamma E_{t-1}[\sigma_t^2]$$

Equation (2) establishes a dynamic relationship in which the investor demands higher risk when the market is riskier. To test this relationship, the following linear regression model can be fitted:

$$r_t = \beta_0 + \beta_1 \sigma_{e,t}^2 + \epsilon_t$$

where the dependent variable is the index surplus return on the expected volatility date. The parameters β_0 and β_1 are constant, and ϵ_t is a random error term. When estimating an equation, an unexpected term is usually added to reflect new information. This news can include economic news, news about monetary policy changes, and other economic shocks. Instead of adding a new variable, it follows the practice of French et al. (1987) and assumes that news and its impact on individuals' decisions are reflected in unexpected volatility [4]. Thus, to test the third hypothesis, the following equation is estimated:

$$r_t = \beta_0 + \beta_1 \sigma_{e,t}^2 + \beta_2 \sigma_{u,t}^2 + \epsilon_t$$

where the dependent variable is the index surplus yield expected in the history of volatility, which is estimated by fitting the following GARCH/mean model:

$$r_t = \mu_t + \beta_1 h_t^{1/2} + \epsilon_t, \quad \epsilon_t \sim N(0, h_t)$$

$$h_t = \omega_0 + \omega_1 \epsilon_{t-1}^2 + \omega_2 h_{t-1}.$$

In this research, to test the significance of the coefficients estimated through the linear regression model and the quarterly regression model, the T statistics and the related critical values are used. Additionally, the F statistic and its critical value are used to test the significance of the coefficient of determination of linear regression models, and the M-statistic and its critical value are used to test the significance of the standard coefficient Quasi-LR and determination of quadratic regression models.

It can be seen that by moving from 0.25 quart to 0.75 quart, the mean coefficient of determination first decreases relatively and then increases significantly. In other words, systematic risk is higher in the returns of the distribution of returns and in the positive (ascending) range of the distribution of returns, it is significantly higher. The results of quarterly regression analysis showed that in different quarters, risk (both expected and unexpected) has a significant

Table 4: Coefficients for regression model determination

model	Quartile 0.25	Quartile 0.50	Quartile 0.75	Linear regression model
Mean coefficient of determination	0.054	0.050	0.073	0.128
The ratio of non-systematic risk to total risk	0.945	0.947	0.927	0.874

and different effect on return. In other words, the relationship between risk and return is different between different quarters of returns. The results show that in negative risk conditions, the average beta increases. As a result, the mortality rate from COVID-19 increases. Therefore, the people should pay more attention and the government should support the people more.

5 Results and Discussion

The world was surprised in early 2020 when a novel coronavirus (COVID-19) quickly spread from China to other parts of the world [3]. Unlike previous coronaviruses that were largely contained to specific geographic regions such as SARS in North Asia and MERS in the Middle East, this new form of coronavirus rapidly spread to other parts of the world [16]. This disrupted global interaction with countries closing borders and regions being shut down [1]. Results showed that LSTM works better if a huge amount of data and enough training data are available, while ARIMA is better for smaller datasets. ARIMA requires a series of parameters (p.d.q) which must be calculated based on data, while LSTM does not require setting such parameters. However, some hyperparameters need to be tuned for LSTM. One major difference between the two is that ARIMA could only perform well on stationary time series (where there is no seasonality, trend, etc.) Validation of ARIMA and LSTM models was performed with the testing data and the results of actual and forecast values of both studied models are shown in the table. 3. Results indicated the forecasting values of the Deaths index obtained from the testing dataset for ARIMA and LSTM models are in excellent correlation with actual experimental values. From the results, ARIMA yields better results in forecasting the short term, whereas LSTM yields better results for prediction of the Trend of COVID-19 data. The presence of a large number of network parameters in the LSTM network makes it computationally expensive. Although the LSTM network is very good and consistent for problems based on sequence and time, the ARIMA networks provide a faster and cheaper alternative to the LSTM network for the Trend of the COVID-19 prediction problem. Traditional time series forecasting methods (ARIMA) focus on univariate data with linear relationships and fixed and manually diagnosed temporal dependence. It is observed that using layers of different models to obtain an ARIMA deep neural network gives precise results in terms of RMSE, MAE, and MAPE. Neural networks (LSTMs and other deep learning methods) with huge datasets offer ways to divide it into several smaller batches and train the network in multiple stages. The batch size/each chunk size refers to the total number of training data used. The term iteration is used to represent several batches needed to complete training a model using the entire dataset. LSTM is undoubtedly more complicated to train and in most cases does not exceed the performance of a simple ARIMA model. Classical methods like ARIMA outperform machine learning and deep learning methods for one-step forecasting on univariate datasets. Classical methods like ARIMA outperform machine learning and deep learning methods for multi-step forecasting on univariate datasets. Classical methods like ARIMA focus on fixed temporal dependence: the relationship between observations at different times, which necessitates analysis and specification of the number of lag observations provided as input. As LSTMs are equipped to learn long-term correlations in a sequence, they can model complex multivariate sequences without the need to specify any time window. The study also explores the impact of risk on COVID-19 mortality rates. It is observed that under negative risk conditions, the average beta (a measure of systematic risk) increases, correlating with higher COVID-19 mortality rates. This finding suggests that as the risk environment deteriorates, mortality rates tend to rise. Therefore, it underscores the need for heightened public awareness and stronger governmental support during periods of increased risk to mitigate the impact on public health.

This study used perspective theory to explain the results of the analysis of different return patterns, specifically the separation of positive and negative return patterns, and it was observed that the results are consistent with the inverse effect and the first and second hypotheses of the research were not rejected. Early data from the COVID-19 pandemic, such as those reported from Wuhan in December 2019 and early 2020, often suffered from incompleteness and inaccuracy. Limited testing capabilities led to underreporting of cases and deaths. Many early cases may not have been diagnosed or recorded, and the true extent of the outbreak was likely underestimated. In the initial stages, the scientific community had limited knowledge about the virus, its transmission dynamics, and its clinical

manifestations. This lack of understanding could have led to inconsistent data collection methods and reporting criteria, further compromising the quality of early data. Additionally, the early phase of the outbreak saw significant delays in case reporting due to overwhelmed healthcare systems and logistical challenges. This delayed reporting could skew the perceived trend of infections and deaths, making early predictions less reliable. As the understanding of COVID-19 evolved, the case definitions were updated, impacting the reported number of cases. Early on, cases were identified based on specific symptoms and exposure history, but as testing expanded and the definition of COVID-19 broadened, the reported cases increased, complicating trend analysis. Moreover, early data heavily focused on Wuhan and Hubei Province, where the outbreak was first detected. This regional bias means that early predictions may not be generalized to other regions with different demographic, healthcare, and socioeconomic contexts.

References

[1] I. Alon, M. Farrell, and S. Li, Regime type and COVID-19 response, FIIB Bus. Rev. **9** (2020), no. 3, 152–160.

[2] V.H. Buch, I. Ahmed, and M. Maruthappu, *Artificial intelligence in medicine: Current trends and future possibilities*, Br. J. Gen. Pract **68** (2018), 143–144.

[3] R. Cortez and W. Johnston, *The coronavirus crisis in B2B settings: Crisis uniqueness and managerial implications based on social exchange theory*, Ind. Market. Manag. **88** (2020), 125–135.

[4] K.R. French, G. William Schwert, and R.F. Stambaugh, *Expected stock returns and volatility*, J. Financ. Econ. **19** (1987), no. 1, 3–29.

[5] F. Gers, N. Schraudolph, and J. Schmidhuber, *Learning precise timing with lstm recurrent networks*, J. Machine Learn. Res. **3** (2000), 115–143.

[6] R.J. Hyndman and G. Athanasopoulos, *8.9 Seasonal ARIMA models*, oTexts **19** (2015).

[7] L. Jia, K. Li, Y. Jiang, X. Guo, and T. Zhao, *Prediction and analysis of corona virus disease 2019*, PloS one **15** (2020), no. 10, e0239960.

[8] D. Kahneman and A. Tversky, *Prospect theory: An analysis of decision under risk*, Econometrica **47** (1979), 263–291.

[9] M. Koohi-Moghadam, H. Wang, Y. Wang, X. Yang, H. Li, J. wen Wang, and H. Sun, *Predicting disease-associated mutation of metal-binding sites in proteins using a deep learning approach*, Nature Machine Intell. **1** (2019), no. 12, 561–567.

[10] X. Li and X. Wu, *Constructing long short-term memory based deep recurrent neural networks for large vocabulary speech recognition*, IEEE Int. Conf. Acoustics Speech Signal Process., IEEE, 2015, pp. 4520–4524.

[11] H. Sak, A. Senior, and F. Beaufays, *Long short-term memory recurrent neural network architectures for large scale acoustic modeling*, arXiv preprint arXiv:1402.1128 (2014).

[12] J. Schmidhuber, F. Gers, and D. Eck, *Learning nonregular languages: A comparison of simple recurrent networks and lstm*, Neural Comput. **14** (2002), no. 9, 2039–2041.

[13] T. Singhal, *A review of coronavirus disease-2019 (COVID-19)*, Ind. J. Pediat. **87** (2020), no. 43, 281–286.

[14] S. Swain, S. Nandi, and P. Patel *Development of an ARIMA model for monthly rainfall forecasting over Khordha district, Odisha, India*, Recent Findings in Intelligent Computing Techniques, Springer, Singapore, 2018, pp. 325–331.

[15] A. Tversky and D. Kahneman, *Rational choice and the framing of decisions*, J. Bus. **59** (1986), 251–278.

[16] Q. Wu, A. Boueiz, A. Bozkurt, A. Masoomi, A. Wang, D.L. DeMeo, S.T. Weiss, and W. Qiu, *Deep learning methods for predicting disease status using genomic data*, J. Biomet. Biostatist. **9** (2018), no. 5.