2022

Comparing Econometrics Approach Vs. Deep Learning Approach in Forecasting Covid-19 Infections and Deaths Horizon in Egypt

Shereen Nosier
Seragaldin Institute for Multidisciplinary Advanced Research (SIMAR), Bibliotheca Alexandrina, Alexandria, Egypt.
Department of Economics, Faculty of Economic Studies & Political Science, Alexandria University, Alexandria, Egypt.

Sara El-Shobaky
The International School of Information Science, Information Communication Technology (ICT) Sector, Bibliotheca Alexandrina, Alexandria, Egypt.

Reham Salah
Seragaldin Institute for Multidisciplinary Advanced Research (SIMAR), Bibliotheca Alexandrina, Alexandria, Egypt.

Abstract

Recently, researchers have applied the Artificial Intelligence, especially deep learning in forecasting, instead of the traditional econometrics' methods. They argue that deep learning approach can improve the forecasting results by reducing the forecasting errors and saving time and cost. However, there is no empirical evidence for that, since there is a lack of research comparing the forecasting performance of these approaches. Consequently, the aim of this paper is to examine the latter argument and to identify the best technique to be used in forecasting the Covid-19 pandemic and others. In fact, the Covid-19 pandemic is defining a global health crisis, which is the hugest the world has faced since World War II. In addition to being threatened by GDP decline and income losses; fears of fetal effects of this epidemic makes it critical to predict the potential spread and identify the best techniques to be applied for that purpose. To achieve the aim of this study, two different methods of forecasting, namely an econometric approach named Autoregressive-Distributed Lag (ARDL) and a deep learning model named Long Short-Term Memory (LSTM) are utilized to forecast the number of daily cases and deaths of Covid-19 in Egypt (March 2020 - March 2021). Consequently, the contribution of this paper is twofold; first, investigating the best way of forecasting the Covid-19 epidemic especially, and therefore real-life phenomena in general, second, assessing the impact of mobility on the incidence of the pandemic in Egypt. The results revealed that the LSTM method shows a slightly better forecasting performance even without using mobility data.

Keywords: Forecasting, Econometrics, Deep Learning, Community Mobility, ARDL, LSTM.

Funding sources: This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

الملخص

في الآونة الأخيرة استخدم الباحثون الذكاء الاصطناعي، ولاسيما التعلم العميق Long Short-Term Memory (LSTM)، في التنبؤ كبديل لأساليب الاقتصاد القياسي التقليدية. فقد ادعى البعض أن منهج التعلم العميق يمكن أن يحسن نتائج التنبؤ عن طريق تقليل أخطاء التنبؤ

وتوفير الوقت والتكلفة. ومع ذلك، فلا يوجد دليل تجريبي على ذلك الزعم، حيث يوجد ندرة في البحوث التي تقارن بين أداء نماذج التنبؤ باستخدام النماذج القياسية وباستخدام الذكاء الاصطناعي. ومن هنا فإن الهدف الرئيس لهذا البحث هو التحقق من ذلك بالقياس وتحديد أفضل طريقة تستخدم في التنبؤ بجائحة كورونا وغيرها من الظواهر الواقعية. ومن أجل تحقيق هدف الدراسة، تم استخدام مدخلين مختلفتين للتنبؤ: هما مدخل الاقتصاد القياسي، متمثلًا في منهج التكامل المشترك باستخدام نموذج الانحدار الذاتي للفجوات الزمنية الموزعة (ARDL) من ناحية، ومدخل الذكاء الإصطناعي متمثلًا في نموذج التعلم العميق (LSTM) من ناحية أخرى، للتنبؤ بعدد حالات الإصابات اليومية وحالات الوفيات الناتجة عن كوفيد -19 في مصر خلال الفترة (مارس 2020 - مارس 2021). تمثلت مساهمة هذا البحث في استكشاف أفضل طريقة للتنبؤ بحالات الإصابة وحالات الوفيات الناتجة عن كوفيد -19 بشكل خاص، ومن ثم إمكانية التطبيق على ظواهر الواقع العملي بشكل عام. بالإضافة إلى قياس تأثير حركة ونشاط الأفراد على انتشار الوباء في مصر. توصلت النتائج إلى أن التنبؤ باستخدام منهج التعلم العميق (LSTM) يعطي أداءً أفضل نسبيًا (أخطاء تنبؤ أقل) من التنبؤ باستخدام النموذج القياسي، إلا أن النماذج القياسية تتميز ببعض الخصائص التي تميزها عن نماذج الذكاء الاصطناعي، ومن ثم فإن تقنيات الاقتصاد القياسي لا يمكن الاستغناء عنها أواستبدالها بتقنيات الذكاء الاصطناعي في التنبؤ، لكن يجب أن يتم دمجهم والاستفادة بهم معًا.

1. Introduction

Since the mid of December 2019, the novel coronavirus (Severe Acute Respiratory Syndrome 2 “SARS-CoV-2”) from Wuhan, China, has continued to spread across the globe and turned into a pandemic currently known as Covid-19 (Yousaf, Zahir, Riaz, Hussain & Shah, 2020; Kumar et al., 2020). Because of the rapid increase of the number of infections worldwide, the contagious disease has turned the world upside down. In fact, during the first year of the pandemic, as of 15 August 2020, the novel virus has infected 21.63 million persons across the globe and killed more than 769 thousand people with 3.56% death to infection rate. On the other hand, 14.34 million have been recovered with 66.31% recovery rate as of the same date. A month later, the virus has infected 29.73 million individuals and the death toll reached 939,289 resulting in a lower mortality rate of 3.16% compared to 3.56%.

Additionally, on the same date the number of recovered people has reached 21.54 million, so that the recovery rate increased to 72.46%. Tracking the path of infections, it was depicted that the number of confirmed cases more than doubled only two months later by November 25 2020. For instance, the number of infections reached 60.78 million persons and the deaths crossed the million mark to 1.43 million deaths hence the death rate declined to 2.35%. Following, a 104% increase has occurred in the number of infections to hit 124.86 million after about four months- 25 March 2021. Whilst 2.7 million deaths were caused by the virus resulting in 2.2% death ratio. Moreover, the recovery rate reached a maximum of 80.8%. Up to the present time, by 31 January 2022, the number of infections reached 382.341 million with 5.706 million deaths. Further, recoveries of 303.208 million were recorded. These latest figures reveal that the recovery rate equals 79.30%, whilst the death ratio is 1.49%.

Based on the herein above numbers, it can be deduced that the virus may have become less lethal globally as the death rate kept on decreasing and the recovery rate has risen. To add more, due to the Covid-19 pandemic, everything has been impacted around the world. In order to prevent further transmission, strong preventive measures had been taken such as implementing lockdown and international travel bans. Despite being under lockdown for a long time of uncertainty, countries are still trying their best to stay opening their economies and relax movement restrictions to avoid poverty increase and job losses.

Indeed, the global economy is facing the worst economic catastrophe since the 1930s Great Depression (Breisinger et al., 2020). The case in Egypt is not different from the rest of the world. Despite being threatened by GDP decline and income losses; the Egyptian government has reacted early to restrain the pandemic outbreak. By mid-March 2019, many measures had been undertaken to contain the spread of the virus, including travel bans, a nighttime curfew, social gatherings ban, and the closure of some organizations, religious institutions and schools. More than three months after imposing them, Egypt began lifting many of the restrictions put in place to curb the spread of coronavirus. However, due to increased concerns about a possible second wave to hit the country, other measures of closure had been taken to control any potential infections.

Due to the fact that decisions made now, and in the months to come, will be some of the most important in addition to the fears of possible powerful waves, it is urgent that the government making those decisions has access to the best information available regarding the possible trajectory of the pandemic. Therefore, predictions on the potential spread of Covid-19 based on time series, econometric models as well as deep learning approaches can be important tools for designing and/or evaluating countermeasures. In the last few months, researchers have developed or utilized existing statistical methods and machine learning methodologies to predict the number of cases and deaths (Yousaf et al., 2020). Additionally, the released Google Community Mobility Reports (GCMR) offer a novel opportunity to investigate the relationship between mobility patterns and spread or transmission of Covid-19 (Paez, 2020).

2. Aim of The Study

Based on the previously mentioned reasons, the objective of this study is first using the most recent available data to find the best prediction models for cases and death in Egypt. Additionally, due to the fact that the pandemic is of dynamic nature, more and continuous epidemiological models are needed for forecasting; consequently, three models will be employed namely, Auto Regressive Integrated Moving Average (ARIMA), Autoregressive-Distributed Lag (ARDL) and Long Short-Term Memory (LSTM). Second, to analyze the impact of mobility on the incidence of the pandemic using GCMR to evaluate the results of easing lockdown restrictions. Third, to compare the forecasting capabilities of the three estimated models by means of different statistical assessment criteria.

3. Literature Review

In fact, early attempts have been undertaken in the beginning of the world-wide pandemic in order to model and forecast its possible short or long-run paths. That being the case, several statistical or mathematical models as well as Artificial Intelligence (AI) algorithms have been utilized by many scholars across the globe. In May 2020, Mohamadou, Halidou and Kapen conducted a comprehensive overview of the different employed techniques. Similarly, in October 2020, Tayarani-N carried out a literature review regarding the application of AI models in all horizons related to Covid-19.

On the other front, some work examined the performance of statistical models as compared to AI models. In their study, (Saba & Elsheikh, 2020) forecasted the possible number of cases in Egypt using statistical and AI methods namely; the ARIMA and the Nonlinear Autoregressive Artificial Neural Networks (NARANN). The study utilized the reported cases from 1 March to 9 May 2020 to forecast one month ahead until 8 June 2020. To assess and compare the performance of the two models, the authors utilized the following measures; Mean Absolute Error (MAE), Root Mean Square Error (RMSE), R2, Deviation Ratio (RD) and Coefficient of Residual Mass (CRM). The study concluded that the NARANN outperformed ARIMA as the latter resulted in an increasing error rate over time from 3.08% for the one-day ahead forecasting to 29.48% ten-days ahead. Conversely, the NARANN results showed an absolute percentage error fluctuating between 1.12% and 4.89%.

In a like manner, to estimate future numbers of infections, recovered cases and fatalities in Saudi Arabia, Elsheikh et al. (2021) employed the ARIMA, the NARANN and the LSTM models. Moreover, the trained models were utilized to forecast the total number of cases and deaths in Brazil, India, Saudi Arabia, South Africa, Spain and USA as they have dissimilar pandemic trends. In order to compare the prediction performance of the afore mentioned models, the authors utilized the RMSE, R2, MAE, Efficiency Coefficient (EC), Overall Index (OI), Coefficient of Variation (COV) and CRM. The findings of the paper pointed out that the LSTM proved to have better performance than the other two models in terms of the forecasting accuracy. Further, the coefficient of determinations (R2) of the LSTM models of different countries for the infections and deaths ranged between 0.976 to 0.998 and 0.944 to 0.998, respectively.

Similar effort was conducted by Marzouk et al. (2021) who employed LSTM as well as convolutional neural network and multilayer perceptron neural network techniques to forecast the spread of the virus in Egypt using data from mid-February 2020 to 30 June 2021. To evaluate the models' accuracy, R2 and RMSE were used. The LSTM proved to have the best performance in predicting the infections for one week and one month ahead. The authors claimed that the better prediction outcomes of the LSTM are due to its feedback connections that propagate the data in the backward pass. In addition to the fact that it has the ability to detect the nonlinear pattern in the data over time.

Likewise, Omran et al. (2021) conducted a comparative study of the LSTM and Gated Recurrent Unit (GRU) deep learning approaches in predicting the number of infections and deaths. These methods have been implemented on Egypt, Saudi Arabia, and Kuwait data from May 1 2020 to December 6 2020. The two algorithms were evaluated using Mean Absolute Percentage Error (MAPE), MAE and RMSE. The findings demonstrated that the first method has accomplished better performance in forecasting the infections in the three countries. In contrast, the latter outperformed in the deaths' predictions in Kuwait and Egypt.

Recently, ArunKumar et al. (2022) performed a comparative study of statistical models represented by ARIMA as well as Seasonal ARIMA (SARIMA) and deep learning models –GRU and LSTM– in forecasting infections, deaths and recoveries in the USA, Brazil, India, Russia, South Africa, Mexico, Peru, Chile, United Kingdom (UK) and Iran. The study conducted prediction for 2 months ahead up to August the 22 2021. To select the best-fit models among the employed methods, MSE and RMSE were used. Speaking about the models predicting the total number of infections, the SARIMA outperformed the other models in UK, Russia, India, Chile and Peru. While in Brazil, ARIMA model proved to be the best. Further, the most accurate model in Mexico and Iran was the LSTM. In contrast, the GRU performed well in USA and South Africa. As regards the models of recovered patients, the ARIMA, SARIMA as well as the GRU models each excelled in three countries. Whereas the LSTM was the most accurate model in Iran only. Regarding the performance of the utilized models in predicting fatalities, the LSTM and the GRU models outpaced the other two with the first being the best in five of the ten countries and the other outperforms in the remaining five. That is, the ARIMA models did not prove to be suitable for forecasting deaths. To conclude, deep learning-based techniques outpaced statistical models in the majority of the countries with about 40 folds less RMSE values.

Resorting to a different technique, Negara et al. (2021) applied the ARDL model to predict the number of cases, recoveries and deaths inside a province in Indonesia until the end of 2020. As the results revealed, the prediction accuracy of the infections model is good according to its MAPE value of 11.25%. On the other side, the model of the recoveries provided very good performance.

Despite the importance of forecasting the possible trajectory of the pandemic, as far as we know, the studies conducting predictions of cases and deaths in Egypt are still parsimonious and not timely updated to reflect the most recent situation. In particular, applying the exponential growth rate model, El-Ghitany (2020) conducted a short-term prediction for the epidemic situation in Egypt based on data from February 14 until April 18, 2020. The research forecasted the daily cases from 19 April to 6 June. Additionally, in their study, Elmousalami and Hassanien (2020) provided daily forecasting models of Covid-19 cases based on different time series models such as Moving Average (MA), weighted moving average and single exponential smoothing.

To sum up, there was not any studies in the literature that has compared the econometric techniques to the machine learning algorithms. In this context, Nutarelli (2022) stated that although machine learning offers remarkable algorithms for prediction problems, econometrics employs robust statistical techniques for prediction, inference as well as causal modeling of economic relationships. According to Nutarelli (2022), data analytics in statistics and econometrics are mainly classified to prediction, summarization, estimation, and hypothesis testing. On the other side, machine learning can assist with the first one by detecting hidden patterns in the data. Moreover, it learns directly from data, without the restrictions and assumptions of model-based statistical methods.

4. Descriptive Analysis

This section provides real-time actual data about the confirmed cases and death of Covid-19 in Egypt since the outbreak of the virus in mid-February 2020 until 31 January 2022. In addition, a track of Egyptians mobility and infections is presented.

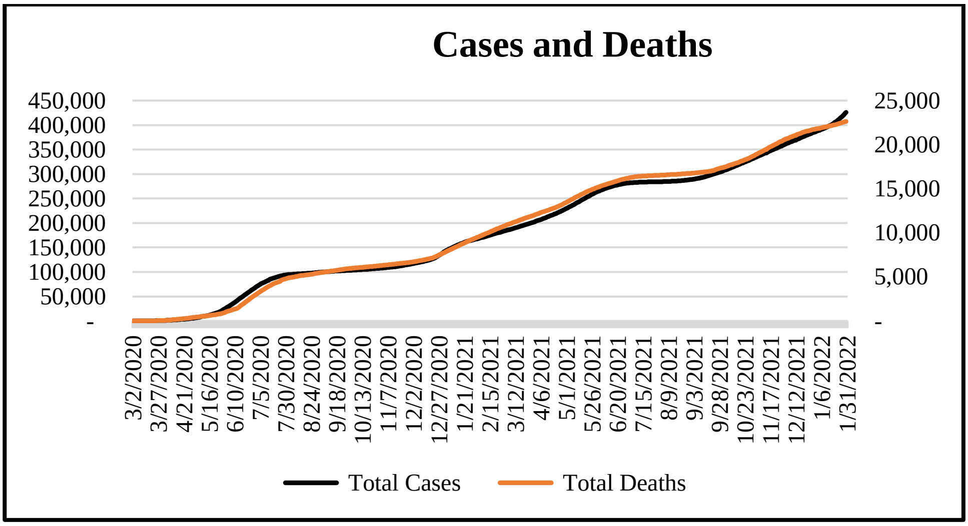
4.1 Cases and Deaths Trends

As depicted by Figure 1, on August 15, 2020, the total number of confirmed infections nationwide in Egypt reached 96,336 since the detection of the first one on 14 February. One month later, the number of infections has risen to 101,340. Further, the death toll amounted to 5,141 and 5,679 deaths with 5.34% and 5.60% death rates on 15 August and 15 September 2020 respectively. On the other hand, the recovered cases have increased to 58,835 and 85,745 on the same dates with 61.07% and 84.61% recovery ratio. About six months later, by mid-March 2021, the number of confirmed cases reached 196,709 whilst, 11,680 deaths were caused by the virus amounting to 5.94% death ratio. Further, the recovery rate reached 76.72%. According to these statistics, one can notice that while the death rate increased from 5.60% to 5.94%, the recovery rate decreased from 84.61% to 76.72%.

More than 10 months later, on 31 January 2022 the total number of cases stood at 425,911 with an increase of 119% than mid-March 2021. On the other side, the number of mortalities grown by 96% to reach 22,635. Moreover, a total of 356,274 persons have been recovered with a 136% increase in the recovery rate.

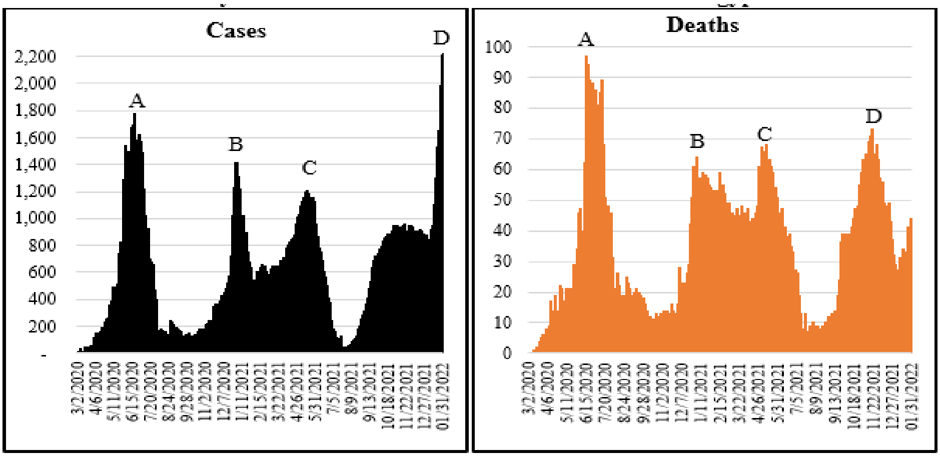
As regards the correspondence between the case in Egypt and worldwide, the most recent values of the pandemic reveal that the death rate in Egypt is higher than its global counterpart by 4.44%. Likewise, the proportion of recovered individuals is 1.82% lower in Egypt than its worldwide respective percentage.

Figure 1
Total Confirmed Cases and Death of Covid-19 in Egypt



Source: Author’s diagram using data of the Egyptian Ministry of Health and Population (MoHP) reports.

Figure 2
Daily Confirmed Cases and Death of Covid-19 in Egypt



Source: Author’s diagram using data of the Egyptian Ministry of Health and Population (MoHP) reports. Note: A, B, C and D point out to the peak points in each wave.

Figure 2 presents abundant information regarding the spread trajectory of the pandemic on a period of almost 2 years. It gives an insight about the waves of infections faced by Egypt as well as the time period between each wave. During 2020, the first year of the epidemic, the country experienced two waves. In fact, the highest number of infections witnessed on the first wave (A) was 1,774 on 20 June 2020. As for the second one, it started around mid-December 2020 while the peak of 1,418 infections has occurred on December 31 (B). Notably, the second wave started to decay by the twentieth of January 2021.

Again, a third wave begun during 2021, particularly on the tenth of April. The highest number of infections during that wave stood at 1,203 cases in mid-May (C). Following, another wave has hit the country starting 1st October and the number of infections kept on increasing to record an extreme of 2,223 cases on the 31st of January 2022 (D).

Important to realize that the maximum number of infections that Egypt has recorded is that of 2,223 cases over the two years of the pandemic. Further, the last wave is the longest with more than 4 months of growing infections.

Speaking of the number of deaths, it reached a maximum of 97 on June 16, 2020 (A). It is worth mentioning that, the number of deaths has been decreasing from 18 July 2020.Then it started rising since December 21 albeit with a smaller rate. Afterwards, it grown to 64 deaths by the 3rd of January 2021 (B). Another hike of 68 was reported on 15 May 2021 (C). Thereafter, an average of 19 mortalities was prevalent during the period (July 2021-September 2021). Subsequently, the death toll increased to 71 deaths by mid-November 2021 then diminished (D).

4.2 Mobility Trends

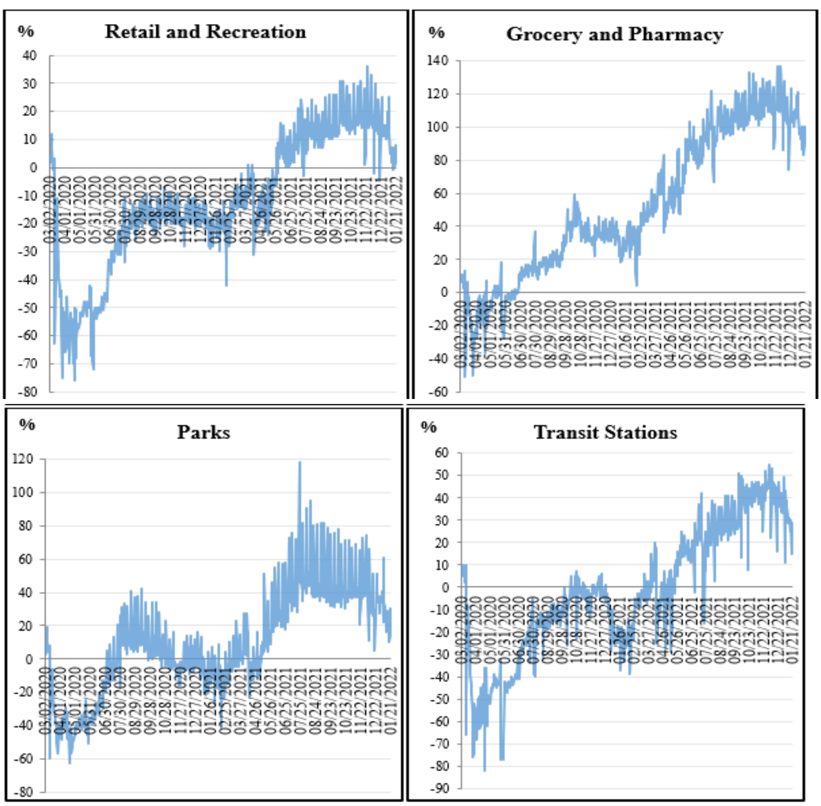
Obviously Egypt has experienced a long list of closures and policies aimed at encouraging citizens to stay at home. Consequently, the levels of mobility changed starting from 12 March 2020 when the government imparked on imposing some restrictions such as schools and universities closing.

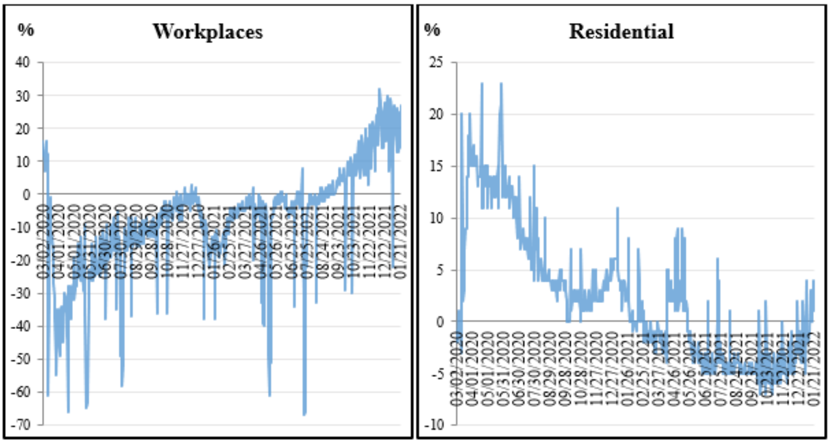
At that time, the reported cases were 7 solely. Following, a night-time curfew has been imposed since March 24, 2020. Additionally, the Ministry of Awqaf [Endowment] declared the closure of the mosques and the places of worship nationwide until further notice. Figure 3 depicts how mobility to diferent destinations has altered in terms of percentage change from baseline.

Clearly, the hike in the number of individuals going out mirrors a spike in the number of persons who have reportedly tested positive for the disease. Accordingly, the crowding that has occurred over the two weeks preceding the holy month of Ramadan has led the curve of the infection to grow. It is worth mentioning that, right before the holy month, Egyptians were supposed to celebrate the Easter holiday on 20 April. However, amid fears of coronavirus outbreak, complete closure of public gardens and beaches was directed. Figure 3 shows clearly the huge recession in mobility on that day. Importantly, this day witnessed the biggest drop in the number of people going out. Visits to retail and recreation venues, parks, transit stations and workplaces were down 76%, 63%, 82% and 66% respectively. Afterwards, on 29 May 2020, the number of cases surpassed the 1,000 marks; this was due to the increased mobility of the Egyptians before and during Eid Al-Fitr.

Figure 3

Mobility Trends' Percent Change from Baseline- the Median Value from the 5-Week Period Jan 3 – Feb 6, 2020- in Egypt





Source: Author's diagram using data of the Google Community Mobility Reports (GCMR).

As mentioned earlier, the highest number of infections taken place at the second half of June. Notably, ten days prior to this latter date, movement to retail and recreational outlets and parks was 2% and 6% higher on average than its counterpart was at the first weeks of June. At the same time, the number of people staying home has dropped. Foundationally, visits to grocery shops and pharmacies started to increase starting from the beginning of June marking a possible reason that individuals embarked on applying self-isolation at homes following the Egyptian Health Minister Treatment protocol.

Later and after about a three-month closure, by the end of June the government started reopening. After that, individuals started spending more time out and as a result time spent at home dropped by 5%. On the other hand, hanging to restaurants, cafes and other places increased by 20% on average during July than June. Likewise, the portion of individuals visiting parks has been raised by 25% on average for the same period. At the same time, grocery stores and pharmacies witnessed 16% increase in visits.

Additionally, use of transportation hubs jumped 18% on average. Despite resulting in upward movement as well as having Eid Al-Adha holiday, the number of infections kept on decreasing since Mid-July. This may be because an increasing number of Egyptians are taking social distancing measures seriously and committing to adhering to preventive measures to limit coronavirus outbreak as a result of increased level of awareness.

With the beginning of January 2021, the Egyptian government re-imposed some closure measures as a proactive step towards the possible second wave. Consequently, schools and universities have applied the blending learning iterating between online and face-to-face learning to reduce the rate of mobility in these places and after a while, they have been totally closed. Without delay, many work places considered lessening the number of employees in a rotational manner to avoid workspace and transportation traffic. As a result, one can observe from the hereinabove figure that the use of transportation facilities has dropped. In addition, going to workplaces has decreased while staying at homes again has grown.

With the coming of the holy month of Ramadan on 13 April 2021, the mobility of the Egyptians especially to grocery shops kept on rising during the month and so did the number of infections. It is worth mentioning that the Egyptians celebrated the Easter holiday on 3 May 2021 during the holy month. Figure 3 shows clearly the huge recession in mobility to workplaces on that day. Importantly, this day witnessed the biggest drop in the number of people going out. Afterwards, on Eid Al-Fitr, the peak of the 3rd wave occurred with maximum number of reported infections of 1,203 on 15 May 2021. Compared to the same days in 2020, the number of cases surpassed the 1,000 marks because of the soaring mobility of the Egyptians before and during Eid Al-Fitr.

On the contrary, the number of infections before, during and after Eid Al-Adha was very low. Notably, visits to parks have increased by 65%, 73% and 68% on the 2nd, the 3rd and the 4th days of Eid Al-Adha respectively. It is worth mentioning that the same days of 2020 have witnessed that exact declining trend of infections.

Investigating the influence of such decisions on the number of infections, it can be noted that these procedures have been successful in easing the effect of the second wave. In fact, during the first hit of the virus within which there was an increase in confirmed cases, the average number of daily infections between May 2020 and July 2020 was 1,122. Whilst, the corresponding number for the period December 2020 and February 2021 was 879. Later on, after smoothing the effect of the second wave, schools and universities have been re-opened at the mid of February. On the other side, the 3rd wave which, occurred during the period April 2021 to June 2021, was characterized by an average number of infections of 932. Moreover, the 4^{th} wave may have started by mid-September 2021 and is still continuing with an increasing rate. Despite the continued governmental regulations regarding the precautionary measures to contain the spread of the virus, it can be concluded that recently the Egyptians are not taking social distancing measures seriously and not committing to adhering to the preventive measures to limit coronavirus outbreak.

5. Methodology

Modeling and forecasting daily number of cases and deaths can help the health system provide services for the infected individuals and provide an insight for policy makers to design protective measures. Consequently, the forecasting models could be of help in predicting trajectories of the pandemic. In this study, firstly, the effect of mobility on the number of Covid-19 cases is estimated using the ARDL approach. In addition, the effect of the incidence of Covid-19 on the number of deaths is also discovered using the same approach. Secondly, ARIMA, ARDL and LSTM methods are employed to predict ex-post daily-confirmed cases and deaths of Covid-19. Therefore, a comparison between the forecasting performances of these utilized methods is conducted and the conclusion of which method obtains the most accurate forecasting and the least forecasting errors is drawn. Therefore, the best model is recommended to be used in other papers to predict the future values of Covid-19 cases and deaths.

5.1 Data

Speaking of the utilized data, a total number of 384 days (from March 2, 2020 to March 20, 2021) were used to develop the ARIMA, ARDL and LSTM models of new cases and deaths. The Egyptian Ministry of Health and Population (MoHP) reports the official data regarding Covid-19 on a daily basis for instance, daily-confirmed cases, recoveries and deaths, as well as the total figures. This period has been chosen because it is the period from the outbreak of the virus on March up to the spread of vaccines in Egypt.

To include the recent period in our sample, data about number of daily-vaccinated persons should be included as another important determinant in both models (cases and deaths); therefore, the time series should begin from March 2021 since the availability of this data. Longer time series is our choice, and another work may examine the effect of vaccination on the predicted value of cases and deaths.

In 2020, Google released GCMR for many countries including Egypt. These reports aim to give vision about the changes in response to the different policies aimed at alleviating Covid-19 consequences. The reports provide mobility alternations over time across different sets of places such as; retail and recreation covering visits to restaurants, cafes, shopping centers, museums, libraries and movie theaters, groceries and pharmacies, parks, transit stations including subway stops and bus and train stations, workplaces, and residential places with respect to a baseline (Google LLC). In other words, the dataset highlights the percent change in visits to different categories of places. It depicts how the frequency and/or length of visits to categorized places change compared to baseline days. The baseline day corresponds to the median value from the 5-week period Jan 3 – Feb 6, 2020 (Ibid). Using this data one can assess the effect of mobility patterns on the spread of Covid-19.

According to Lauer et al. (2020), the median incubation duration is 5.1 days (95% Confidence Interval (CI), 4.5 to 5.8 days), and 97.5% of individuals who develop symptoms will do so within 11.5 days (CI, 8.2 to 15.6 days) of infection. As a result, it is to be expected that any changes in mobility will have a lagged effect on detecting the new infections. Consequently, lagged values of the mobility elements and indicators are estimated by means of time series models.

5.2 Models' Construction

Number of new cases (NCases) and new deaths (NDeaths) on daily basis are used in the estimation as dependent variables in two different models, each model uses three different methods of estimation. Additionally, different independent variables - representing mobility- are included in daily cases model. Moreover, new cases variable is included to estimate and predict the deaths model.

5.2.1. ARIMA Model

Many researchers have utilized ARIMA model as benchmark to forecast the infection horizon of the virus in different countries (Dehesh et al., 2020; Yousaf et al., 2020; Kumar et al., 2020; Benvenuto et al., 2020; Saba & Elsheikh, 2020). Popularly known as the Box–Jenkins (BJ) methodology, the ARIMA methodology emphasizes on analyzing the probabilistic, or stochastic, properties of time series on their own under the philosophy *let the data speak for themselves* (Gujarati & Porter, 2009).

Unlike the regression models, in which Y_t is explained by k regressors $X_1, X_2, \dots, X_k$, the BJ-type time series models allow Y_t to be explained by past, or lagged, values of Y itself and stochastic error terms. The important point to note is that, the time series data used in ARIMA should have stationary and linear nature. Hence, a stationarity test is conducted: for instance, Autocorrelation Function (ACF), Correlogram, or the unit root test using the Augmented Dickey–Fuller (ADF) Test or the Phillips Perron Test. The ARIMA (*p,d,q*) model includes an autoregressive (AR) model and a moving average (MA) model where *p* denotes the number of autoregressive terms, *d* the number of times the series has to be differenced before it becomes stationary, and *q* the number of MA terms. The chief tools in identifying the parameters of the model are the ACF, the Partial Autocorrelation function (PACF) and the resulting correlograms. Having chosen a particular ARIMA model, and having estimated its parameters, the next step is checking whether the chosen model fits the data reasonably well. All the models that pass the residual tests should be compared using Akaike Information Criterion (AIC). The model which has the least AIC is selected as the best model. Afterwards, the forecasting step based on the previously best-fit model is carried out.

5.2.2. ARDL Model

Until the beginning of the 1990s, static regression dominated the literature. It assumed that all the variables in the model are stationary. It has been proved that the macroeconomic data are mostly non-stationary; therefore, the OLS estimator does not produce reliable estimates and the regression tends to be spurious. Researchers use differenced variables in the model to obtain stationary variables, but in this case important information related to the long-run analysis is lost.

The co-integration approach is very attractive since it retains the long-run relations and obtains highly consistent parameters in the long run (Stock, 1987). However, there are integration and co-integration restrictions that the models have to overcome in order to apply this approach. Firstly, the ADF test is used to examine the stationarity of the variables in the different models.

Pesaran, Shin and Smith (2001) proposed the ARDL approach of co-integration which can be applied and yields consistent estimates of the long-run parameters irrespective of whether the underlying variables are I(0), or I(1), or a combination of them. In addition, it permits different number of lags for each regressor to capture the data generation process in a general to specific framework (Feridun, 2009; Nosier, 2018).

The ARDL approach involves several steps. First, the optimal number of lags for all level variables is selected, using the appropriate information criteria, mainly the AIC and Schwartz Information Criterion (SIC). The second step is the bounds test, which involves estimating the Conditional Unrestricted Error Correction Model (UECM) to test for the existence of a long-run steady state relationship between the dependent variable and all the explanatory variables. Then, we can proceed to estimate the UECM (p, q, m, n, s, v, u) as in Equation (1) and Equation (2).

$$\begin{aligned}\Delta LnNCases_t = {} & \sigma_0 + \sum_{i=1}^{p-1} \sigma_{1i}\Delta LnNCsaes_{t-i} + \sum_{i=0}^{q-1} \sigma_{2i}\Delta Grocery_{t-i} \\ & + \sum_{i=0}^{m-1} \sigma_{3i}\Delta Parks_{t-i} + \sum_{i=0}^{n-1} \sigma_{4i}\Delta Recreation_{t-i} \\ & + \sum_{i=0}^{s-1} \sigma_{5i}\Delta \mathrm{Resedential}_{t-i} + \sum_{i=0}^{v-1} \sigma_{6i}\Delta Transit_{t-i} \\ & + \sum_{i=0}^{u-1} \sigma_{7i}\Delta Workplaces_{t-i} + \rho LnNCases_{t-1} \\ & + \theta_1 Grocery_{t-1} + \theta_2 Parks_{t-1} + \theta_3 Recreation_{t-1} \\ & + \theta_4 \mathrm{Resedential}_{t-1} + \theta_5 Transit_{t-1} + \theta_6 Workplaces_{t-1} \\ & + \epsilon_t \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots (1)\end{aligned}$$

$$\begin{aligned}\Delta LnNDeaths_t = {} & \sigma_0 + \sum_{i=1}^{p-1} \sigma_{1i}\Delta LnNDeaths_{t-i} + \sum_{i=0}^{q-1} \sigma_{2i}\Delta lnNCases_{t-i} \\ & + \rho LnNDeaths_{t-1} + \theta_1 LnNCases_{t-1} + \epsilon_t \dots\dots\dots\dots\dots (2)\end{aligned}$$

where p, q, m, n, s, v, u are the optimal lags of level of the regressors: $LnNCases_{t-i}$, $LnNDeaths_{t-i}$ $Grocery_t$, $Parks_t$, $Recreation_t$, $Resedential_t$, $Transitt_t$and $Workplaces_t$ respectively. Moreover, Δ is the first difference operator and σ_0 is a drift component.

Further, the left-hand side is the natural log of new cases ($LnNCases_t$) in Equation (1) and the natural log of new deaths ($LnNDeaths_t$) in Equation (2). The right-hand side of the equation represents the explanatory variables in one lag in level and in differences with the optimal lags for each variable; the independent variables represent six different measures of mobility alternations over time across different sets of places in Egypt in Equation (1) and the natural log of new cases in Equation (2). The parameters σ_{si} correspond to the short-run relations, whereas ρ and θ_s correspond to the long-run relations; ϵ_t is the random error.

The Wald or F-statistic is used to test the joint significance of lagged levels of the variables in the UECM, and determine the existence of the long-run equilibrium under the null hypothesis of no co-integration (H_0: $\rho = \theta_1 = \theta_2 = \theta_3 = \theta_4 = \theta_5 = \theta_6 = 0$) against the alternative that a long-run relation exists (H_1: $\rho \neq \theta_1 \neq \theta_2 \neq \theta_3 \neq \theta_4 \neq \theta_5 \neq \theta_6 \neq 0$) in Equation (1). However, as discussed by Pesaran et al. (2001), both statistics have no standard distribution, irrespective of whether the regressors are purely I(0), purely I(1) or mutually co-integrated.

Therefore, Pesaran et al. (2001) computed two types of asymptotic critical values for a given significance level in the case of including and excluding trend. The first type assumes that all the variables are I(1), and the other assumes that all the variables are I(0). If the computed Wald or F-statistics exceed the upper critical value, the null hypothesis is rejected and the underlying variables are co-integrated. If the Wald or F-statistics are below the lower critical value, the null cannot be rejected and the variables are not co-integrated.

Finally, if the Wald or F-statistic values lies between the two bounds, the test is inconclusive. Once a long-run relationship has been established by the bounds test, the long-run relations can be estimated as $\beta_j = \frac{-\theta_j}{\rho}$. Since j represents the independent variables in Equation (1), β_{js} are the long-run coefficients of the different independent variables in the equation. Finally, the models have to undergo several statistical checking such as autocorrelation, heteroscedasticity and stability test in order to ascertain their statistical reliability to be used.

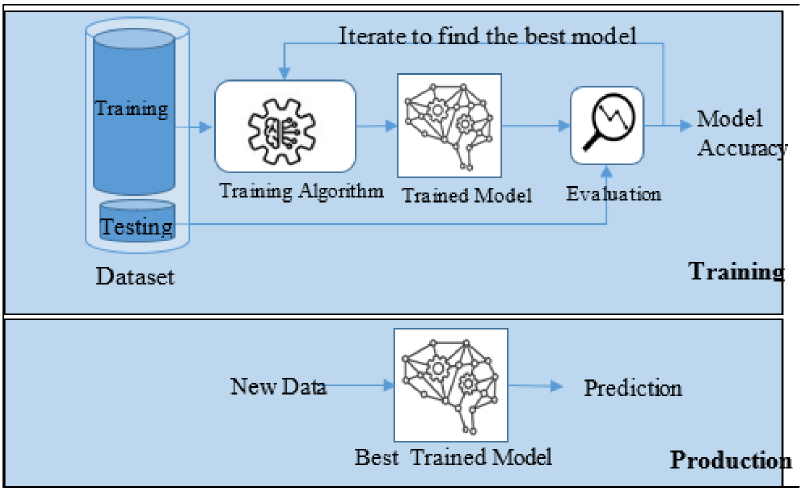
5.2.3. Deep Learning Model

In fact, time series and econometric models such as ARIMA and ARDL among others depend heavily on assumptions. Consequently, they are problematic for forecasting real-time transmission rates of pandemics like Covid-19. Moreover, while they can provide satisfactory predictions to some extent, the complications present in the data are typically hard to detect with classic methods. Hence, the necessity arises for the usage of the deep learning-based network approaches, which are more proficient of catching these intricacies. Deep learning is a type of machine learning which is a subfield of AI. Their main goal was to make computers learn from a provided dataset.

Error! Reference source not found. 4 gives a high-level overview of the learning process. Initially a dataset is split into two different parts. Most of the data is used for training the model and the rest is used to test and evaluate its behavior. Initially the training data is fed to a learning algorithm. That algorithm applies a nonlinear transformation to its input and uses what it learns to build a model as output. The algorithm iterates until the output reaches an acceptable level of accuracy. The most accurate model is then used in the production phase. It can predict the output of any new provided data.

Figure 4

Learning Process

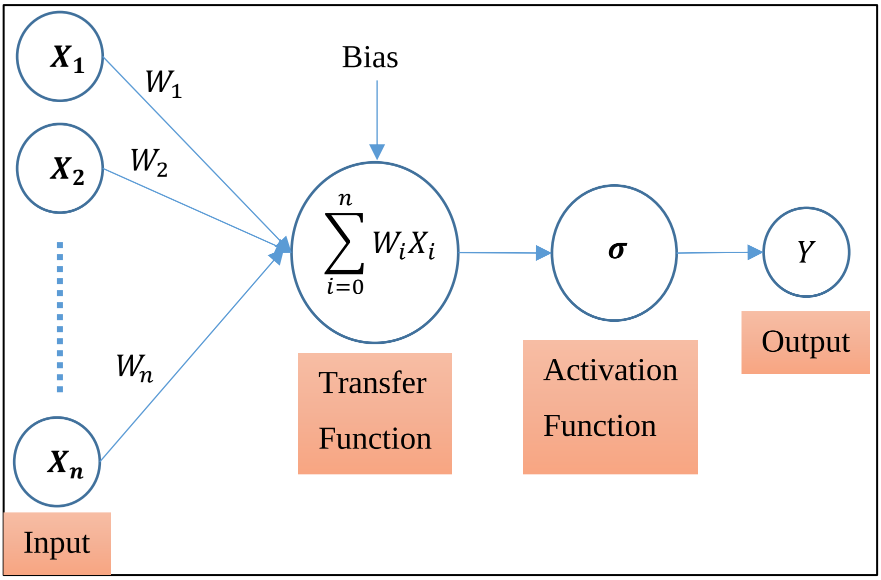


Source: Author’s diagram using the learning process idea.

An Artificial Neural Network (ANN) consists of artificial neurons stacked next to each other in three different layers; an input, output and a hidden layer. The network starts by assigning random values to the weights of the connections between nodes. The node multiplies the inputs with the random weights, calculates them and adds a bias. Finally, a nonlinear activation function is applied to determine the output. If the generated output is different from the expected output, the error is calculated and the weights are readjusted in order to minimize the error between the network output and the target output. This process is iterated until an expected output is obtained. The main network computation graph is depicted in Figure 5.

Figure 5

Neural Network Structure



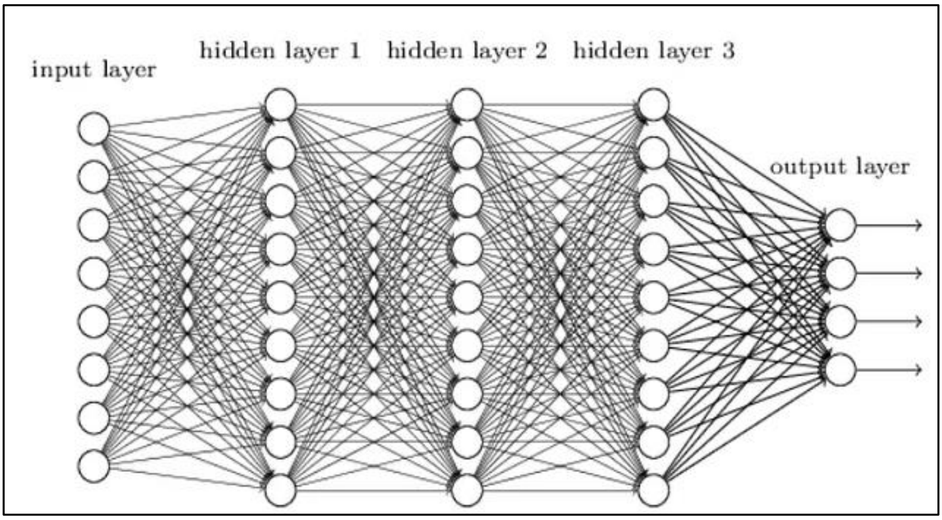
Source: Author's diagram using the Neural Network Structure idea.

Deep neural network is a neural network with more than three nodes' layers. The first one is the input layer and the last one is the output layer. The middle layers are the hidden layers. Those layers are connected together and used to make the model learns to give the final result. An example of a deep neural network is shown in Figure 6.

Deep learning models make use of several algorithms that each is well suited to perform a specific task. In fact, the Recurrent Neural Network (RNN) is the most widely used type of ANN. They are networks with loops, in order to store the output data of the processing nodes and learns to improve its functionality. RNN have shown a great potential in time series prediction by storing information in its internal state. However, it exhibited problems that lead to large training time or not working. Therefore, to get rid of this shortcoming Hochreiter and Schmidhuber designed the LSTM in 1997.

Figure 6

A Deep Neural Network

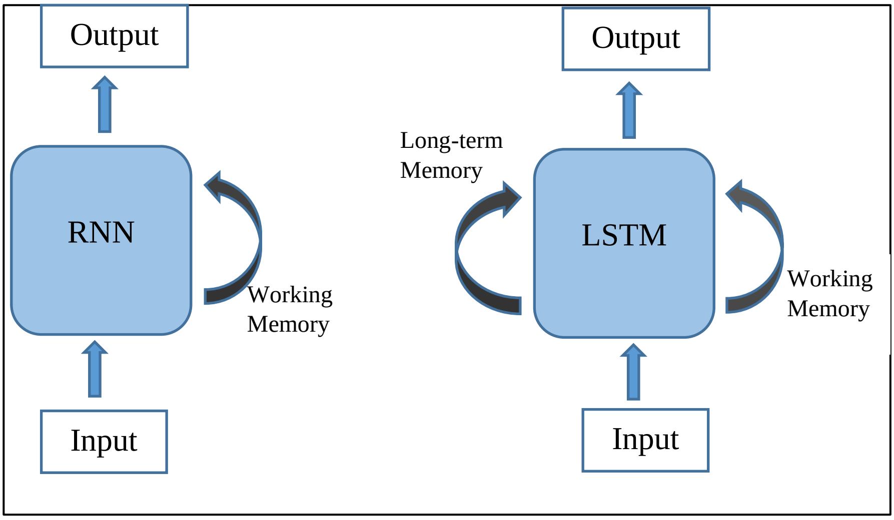


Source: https://www.kdnuggets.com/2020/02/deep-neural-networks.html

LSTM is a kind of RNN that is designed to capture sequential dependencies between the data. It has achieved a great success in many fields including time series forecasting (Elsworth & G¨uttel, 2020; Masum, Liu & Chiverton, 2018; Namini, Tavakoli & Namin, 2018). It has a more complex structure to help remembering the values from earlier stages for future use. Each node in the LSTM network takes three different inputs, the current input data, the short-term memory from the previous cell similar to the RNN, and lastly the long-term memory. A comparison between the RNN and LSTM is described in

Figure 7.

Figure 7
RNN vs. LSTM



Source: Yasrab, Pound, French, and Pridmore (2020).

LSTMs are a complex area of deep learning. A common LSTM unit is composed of four gates, the input gate, the forget gate, the control gate and the output gate as described in Figure8 (https://colah.github.io/posts/2015-08-Understanding-LSTMs/).

Initially, the LSTM decides what information from the input of previous memory has to be neglected. It uses a sigmoid layer which is the 'forget gate layer'. This layer decides on the received output from the last LSTM module h_{t-1} at a time step $t-1$ in addition to the present input signal X_t at time t. This gate outputs f_t -a number between 0 and 1- for each number in the cell state C_{t-1} based on the information carried by the input. The '1' means keep this, while '0' means forget it. The function of the forget gate is defined by the following equation:

$$f_t = \sigma\left(W_f \times [h_{t-1}, X_t] + b_f\right) \quad (3)$$

where, σ is the sigmoid layer, b_f is the applied bias of the gate and W_f is the weight matrix of the gate.

Next, in order to decide what to store in the cell state, a sigmoid layer which is the "input layer gate" decides which values to update. In addition, a tanh layer creates the new candidate values $\tilde{C}_t$ that has to be added to the state. Their functions are defined by the following:

$$i_t = \sigma\,(W_i \times [h_{t-1}, X_t] + b_i) \quad (4)$$

$$\tilde{C}_t = \tanh(W_c \times [h_{t-1}, X_t] + b_c) \quad (5)$$

Then, in order to update the new cell state C_t, the old state C_{t-1} is multiplied by the f_t to forget what was decided earlier. Then $(i_t \times \tilde{C}_t)$ are added to the results which are the new candidate values scaled by how much decided to update each state value.

$$C_t = f_t \times C_{t-1} + i_t \times \tilde{C}_t \quad (6)$$

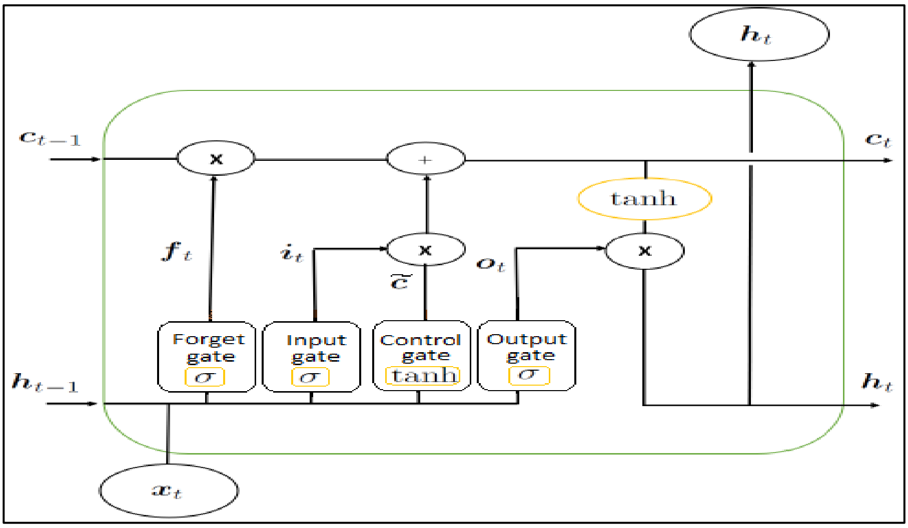
Finally, a sigmoid layer is used to decide which part of the cell to output. And, a tanh layer is used to scale the values into ranges between -1 and 1. Then, the result is multiplied by the output of the sigmoid gate to only output what was decided to.

$$O_t = \sigma\,(W_o \times [h_{t-1}, X_t] + b_o) \quad (7)$$

$$h_t = O_t \times \tanh(C_t) \quad (8)$$

It has to be noted that the ***'tanh'*** is used to scale the values into range from -1 to 1, ***'σ'*** is the activation function which is taken as sigmoid, ***'W'*** are the corresponding weight matrices.

Figure 8
LSTM unit



Source: https://colah.github.io/posts/2015-08-Understanding-LSTMs/

5.3 Ex-Post Forecasting Procedures

An ex-post forecast (after the event forecast) is a forecast that is run in past periods for which actual values of the new cases and new deaths, in addition to the explanatory variables are available. The comparison of ex-post forecasts between different methods of estimation allows researchers to decide which method generates the best forecast, and therefore it can be used to produce ex-ante (before the event) forecasts (Nosier, 2012). Ex-post forecasting in the present study aims to evaluate the forecasting performance of the LSTM, ARDL co-integration estimates and ARIMA (*p d q*) estimates for comparison.

The three methods are re-estimated for the new cases and new deaths of Covid-19 models using daily data for the period from 2nd of March 2020 to 31st of December 2020. The estimated parameters are used to generate ex-post forecasting over 79 days for the period (1st of January 2021 – 20th of March 2021). Dynamic forecasting calculates dynamic, multi-step forecasts starting from the first period in the forecast sample. In dynamic forecasting, previously forecasted values for the lagged dependent variables are used in generating forecasts of the current value. This choice will be available for our models, since the estimated equation contains dynamic components, such as lagged dependent variables and ARIMA terms.

The appropriate method of forecasting is chosen as the method which has the least forecasting errors according to the forecasting error criteria, such as Root Mean Square Errors (RMSE), Mean Absolute Errors (MAE), Mean Absolute Percentage Errors (MAPE) and Symmetric Mean Absolute Percentage Errors (SMAPE).

Forecasting errors are defined as the differences between the actual (A_t) and forecasted (F_t) value over the period of the forecasting horizon, which are defined as:

$$e_t = A_t - F_t \quad (9)$$

where $LnNCases_t$ is the actual value of the natural log of new cases in time t, $LnNCasesF_t$ is the forecasted value of the natural log of new cases in time t, e_t is the forecasting errors in time t. In theory, for a well specified model, it is expected that the forecasting error has a mean of zero over a certain forecasting horizon. However, very small forecasting errors can be obtained even if the models are not well specified as a result of the existence of positive and negative forecasting error values, which cancel each other out. To solve this problem, measures of forecasting errors accuracy have been improved and the errors of Equation (9) transformed either to squared values (e_t^2) as in the RMSE in Equation (10), or to absolute values $|e_t|$ as in the MAE in Equation (11) (Song, Witt & Li, 2009).

$$RMSE = \sqrt{\sum \frac{e_t^2}{n}} \quad (10)$$

$$MAE = \frac{\sum |e_t|}{n} \quad (11)$$

where n is the number of forecasts. The main difference between the two forecasting error measures is that the MAE gives equal weight to all errors, whereas the RMSE gives more weight to larger errors. Therefore, the RMSE is more sensitive to one extremely bad forecast (Li, Song & Witt, 2005). The MAPE is another error forecast measure in which the errors of the forecast are divided by the actual values of the dependent value, as in Equation (12), to produce unit independent measures (percentage errors), so one can compare the errors of fitted models that differ in level.

$$MAPE = \left(\frac{1}{n} \sum \frac{|e_t|}{A_t}\right) \times 100\% \quad (12)$$

Finally, the SMAPE is an accuracy measure based on relative or percentage errors. Similar to the MAPE, relative error is the absolute error divided by the actual values, therefore it is unit independent (percentage-based measure), and hence, it can be used to compare forecast performances between datasets. In contrast to the MAPE, SMAPE has both a lower bound and upper bound as illustrated in Equation (13).

$$\text{SMAPE} = \left(\frac{1}{n}\sum\frac{|e_t|}{|A_t|+|F_t|}\right) \times 100\% \qquad (13)$$

In the present study, the three measures are used to compare forecasting accuracy between different methods for the same model, whereas the MAPE is used to compare forecasting accuracy even across the two models. Moreover, Lewis (1982) suggested the following interpretation of the resulting MAPE values: Less than 10% is highly accurate forecasts, 10% - 20% is a good forecasting, 20% - 25% is a reasonable forecasting, 50% or more is an inaccurate forecasting (Lewis 1982, quoted in Tideswell, Mules & Faulkner, 2001).

6. Results

In this section, the estimation results of different methods are presented. Diagnostic tests are performed as well. If the used methods overcome the diagnostic and stability tests, we proceed by using the fitted models in generating ex-post forecasting of daily cases and deaths in Egypt over a period of 79 days from January 1 to March 20 as stated before. Comparisons between the ex-post forecasts in Cases and deaths models using the three different methods are also presented and discussed. Finally, the results of the APE are presented to compare the performance of the three methods of forecasting at different time horizons. EViews version 10 is used to estimate the econometrics methods- ARIMA and ARDL, while Python version 3.7 and the Keras version 2.3 are used to estimate the LSTM models.

6.1 Results of ARIMA

ARIMA (3, 0, 3) is selected - according to the AIC- as the most appropriate specification to forecast the number of new cases of Covid-19 in Egypt, as illustrated in Figure 9 (A). The model as a whole is significant at the 1% level of significance and it has high explanatory power (R^2= 0.954). The residuals of the model have no first order autocorrelation according to DW statistic (1.79), but the residuals are heteroscedastic (F-statistic of Autoregressive Conditional Test of Heteroscedasticity (ARCH) = 90.721, P-value = 0.000).

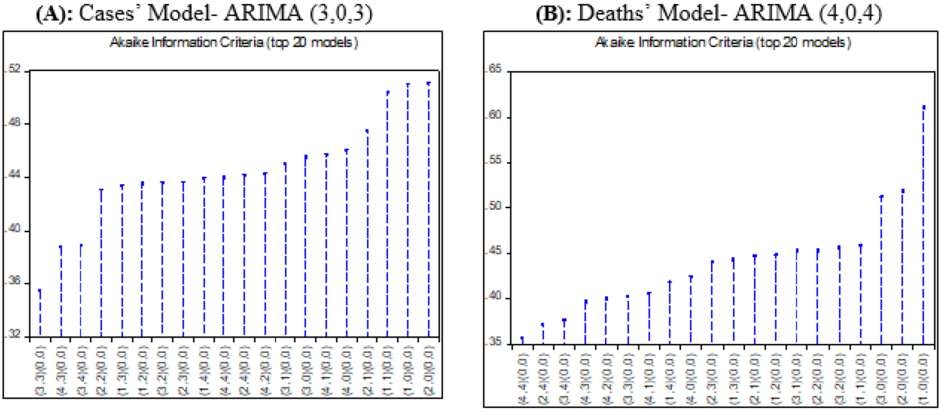
In addition, ARIMA (4, 0, 4) is used to forecast the number of new deaths, as it is proved to have the best specification using the AIC, as shown by Figure 9 (B). The model is significant at the 1% level of significance, and it has strong explanatory power (R^2= 0.920). The residuals have no first order autocorrelation; DW statistic = 2.01. However, using ARCH test, the residuals are heteroscedastic (F-statistic = 16.572, P-value=0.004). Finally, the model is stable (F-statistic of RESET test=1.118, P-value=0.291). Therefore, the ARIMA models have no econometric problems, except for heteroscedasticity. Solving for the problem of heteroscedasticity in the two models; GLS estimator is utilized to obtain homoscedastic residuals.

Finally, the specified ARIMA models are estimated for the period of 305 days from 2nd of march to 31 of December. Then, the estimated models are used to create the ex-post forecasting of the values of Cases and Deaths for the period of 79 days from January 1, 2021, to March 20, 2021.

Figure 9

ARIMA Optimal Lags Selection Using AIC

(A): Cases' Model- ARIMA (3,0,3) **(B):** Deaths' Model- ARIMA (4,0,4)



Source: Authors' estimation using EViews 10.

6.2 Results of ARDL

It has to be noted that the ARDL specification model, which includes all the mobility alternations in addition to weather temperature is examined first. However, it has been found empirically that including the insignificant mobility indicators and temperature in the model reduces the accuracy of the forecasting results. The more parsimonious model is proved to be better for forecasting purposes.

6.2.1. Unit Root Results

The results of the ADF unit root tests are reported in Table 1. At the 5% level of significance, the results indicate that LnNCases, lnNDeath, and Parks are non-stationary, but stationary in first differences; I(1) variables according to the appropriate trend-specification. Other variables in the two models are found to be stationary variables; I(0). However, we proceed by estimating the co-integration relation of the two models, since we use the ARDL technique which permits a co-integration relationship to exist irrespective of whether variables have the same integrating order or not.

Table 1

Unit Root Tests According to the Appropriate Deterministic Trend

Variables	Trend Specification	ADF	Variables	Trend Specification	ADF
$LnNCases_t$	Constant	-2.164 (0.220)	**$\Delta LnNCases_t$**	None	-4.135 (0.000)
$LnNDeath_t$	Constant	-2.333 (0.162)	**$\Delta LnNDeath_t$**	None	-7.240 (0.000)
$Workplaces_t$	Constant + trend	-5.420 (0.000)			
$Parks_t$	Constant	-1.697 (0.432)	**$\Delta Parks_t$**	None	-4.227 (0.000)
$Residential_t$	Constant + trend	-4.932 (0.000)			

Source: Authors' estimation using EViews 10.

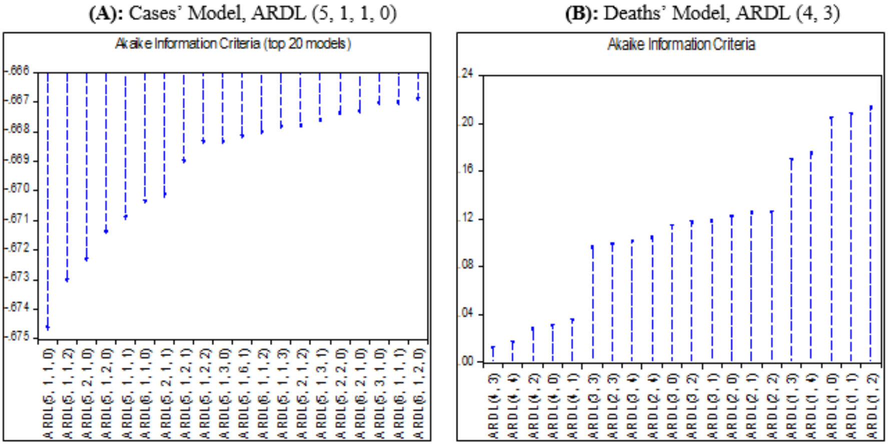
6.2.2. Bounds Test Results

The first step in ARDL approach of co-integration is to determine the optimal lags of the variables included in the models. Six lags are set as a maximum lag length, and different ARDL specifications are chosen separately for each model and each variable according to the AIC as illustrated in Figure 10. Therefore, ARDL (5, 1, 1, 0) has been selected as the best-fit model to forecast daily new cases, while, ARDL (4, 3) has been preferred for predicting daily new deaths.

Figure 10

ARDL Optimal Lags Selection Using AIC

(A): Cases' Model, ARDL (5, 1, 1, 0) **(B):** Deaths' Model, ARDL (4, 3)



Source: Authors' estimation using EViews 10.

The second step of ARDL approach is to test the existence of the long-run Equilibrium relationship –co-integration relationship– between LnNCases and its determinants in the first model, and between LnNDeaths and its determinants in the second model. The results of bounds tests (computed F-statistics) for the two models are presented in Table 2.

The null hypothesis of no co-integration is rejected in the two models, indicating that there is a co-integration relation at the 1% level of significance for the two models over the period of the study. Therefore, we can proceed by estimating the long-run relationships between these variables.

Table 2

The Results of F-Statistics for Co-Integration Relationship

	LnNCases' Model			**LnNDeaths' Model**		
Critical value	**Significance level**		**F-Statistic**	**Significance level**		**F-Statistic**
	1%	5%		1%	5%	
Minimum value I (0)	4.29	3.23	15.629	6.84	4.94	17.525
Maximum value I (1)	5.61	4.35		7.84	5.73	

Source: Authors' estimation using EViews 10. Note: The critical value of F is calculated at 3 independent variables in the Cases' model, and one independent variable in the Deaths' model. Constant without time trend is included in both models.

6.2.3. Estimating the Long-Run Equilibrium Relationships

Having determined the best ARDL specification for the two models and the long-run co-integration relation is detected, long-run steady-state parameters were estimated and examined statistically in the two models over the period of 305 days from 2nd of march to 31 of December as illustrated in Table 3.

Table 3
Long-Run Results of ARDL Co-Integration Approach

Independent Variables	LnNCases' Model ARDL (5,1,0,1)		Independent Variable	LnNDeaths' Model	
	Coefficient	T-statistics		Coefficient	T-statistics
Workplaces	0.104 (0.004)	2.899	**LnNCases**	0.772 (0.000)	11.463
Parks	-0.077 (0.000)	-3.554			
Residential	0.089 (0.268)	1.109			
Goodness of fit & diagnostic tests					
R^2	0.969		**R^2**	0.911	
Adj. R^2	0.968		**Adj. R^2**	0.909	
Durbin-Watson stat	2.005		**Durbin-Watson stat**	2.124	
B-G LM (F-statistics)	8.819 (0.000)		**B-G LM (F-statistics)**	6.419 (0.002)	
ARCH hetero (F-statistics)	52.192 (0.000)		**ARCH hetero (F-statistics)**	19.542 (0.000)	
RESET (F-statistics)	3.343 (0.068)		**RESET (F-statistics)**	1.288 (0.198)	

Source: Authors' estimation using EViews 10. Note: B-G LM test is the Breusch-Godfrey Lagrange Multiplier statistics for residual autocorrelation. P-value between Brackets.

The models perform reasonably. The adjusted R^2 is very strong in both models, indicating strong explanatory power and the diagnostic tests indicate that the two models are stable. However, autocorrelation and heteroscedasticity are detected in both models. Both problems are corrected in the two models, and the reported results are the results after corrections. Therefore, forecasting using these models can be performed accurately.

For the cases' model, all the long-run coefficients are significant at the 1% level of significance - but the residential variable is not significant - indicating that mobility of population is affecting the incidence of new cases of Covid-19 significantly over the period of the study. As expected, a 1% increase in visits to workplaces compared with the base line is associated with higher incidence of Covid-19 by 10% at the 1% level of significance. However, the long-run coefficients of parks have an unexpected negative sign. In the deaths' model, an increase in new cases by 1% tends to increase the new deaths by 0.77% on average in the long run.

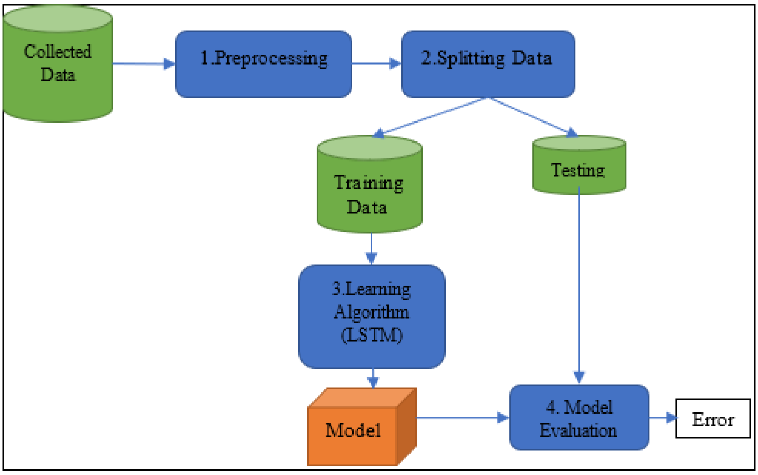
Therefore, the spread of Covid-19 is caused by increasing mobility, especially to workplaces. Moreover, the spread of Covid-19 is more likely to increase the possibility of deaths.

Finally, the estimated ARDL models are used to create the ex-post forecasting of the values of Cases and Deaths for the period of 79 days from January 1 to March 20.

6.3 Results of LSTM

The proposed system includes four different steps. The first one is the data pre-processing step. It prepares the data in a format applicable to feed the LSTM Network for learning. The second step is splitting the processed data into two different datasets one for training the model and the other one for evaluating it. The third step is the deep learning process itself using the LSTM network. The last step is to evaluate the generated models in order to use the best one in our experiments. All those steps are applied in each of the New Cases and New Death datasets separately. The workflow of the proposed system is depicted in Figure 11.

Figure 11
Proposed System Workflow



Source: Authors' diagram of the system workflow.

6.3.1. Data pre-processing:

The original dataset contains the daily reported number of new cases and new death since March 2, 2020, to March 20, 2021. It has to be noted that the data set also included the daily mobility and weather temperature for the same period. However, it has been found empirically that including the mobility and temperature in the data features didn't highly improve the accuracy of the results.

In our experiments, the input features used to train the model are a number of recent past values that are used to predict the value of the next time period. For example, given a look back window=3 -referred as lag=3- the values at time (t-1, t-2 and t-3) are used as input features to predict the output value at time (t). In each experiment, the dataset is adjusted with lags ranging from 1 to 15 to choose the one that generates a model with the lowest error.
Furthermore, all the time series data were re-scaled to values between -1 and 1, this is because the LSTM model requires data to be within the scale of the activation function of the network.

6.3.2. Data Splitting

After preparing the dataset with a specific lag parameter for the cases and deaths models, the ordered dataset is split into a training dataset and a testing dataset with 79.5% and 20.5% of the observations respectively.

6.3.3. Model Training

The LSTM network has been implemented in Python using the open source Keras deep learning library (Keras, n.d.) and the Scikit-learn library (Scikit-learn: Machine Learning in Python, n.d.). The implemented network has one input layer, followed by a hidden layer with 50 LSTM neurons and one output layer that predicts a single value. The network is trained for 200 epochs. The activation function of the gates is hyperbolic tangent. The model is trained using the Adam optimizer. The Mean Squared Error (MSE) is the loss function for the model. All those parameters are chosen empirically to choose the most accurate model with the minimal RMSE. For new cases forecasting, the LSTM model with lag=13 reported the lowest RMSE. For the new death forecasting, using the number of new cases as a feature with lag =15 reported the lowest RMSE.

All computations are performed on the BA High Performance Computing cluster (The Bibliotheca Alexandrina Supercomputing Facility, n.d.). The used Python version is 3.7 and the Keras version is 2.3. Due to the stochastic nature of the algorithms that depends on the random initialized weights, the results may vary. Hence, we considered comparing the average of eight consecutive runs in order to choose the most accurate network parameters.

6.3.4. Model Evaluation

Once the model is fit using the training dataset, their corresponding testing dataset is used to evaluate the model. In such process, the model is used to predict the current number of cases and deaths in the current time step. Then, this current predicted value is used to update the input features of the next step. Finally, all the predicted values are compared with the original values to calculate the RMSE error to evaluate the model.

6.4 Ex-Post Forecasting Results

6.4.1. Forecasting Errors According to Different Criteria

The performance of the chosen models of daily cases and deaths respectively is compared using different measures of forecasting error accuracy. Table 4 shows the forecast error of these models in terms of RMSE, MAE, MAPE and SMAPE over the period of 79 days, which represent almost 20% of all the available observations.

Table 4

Forecasting Error Criteria of the Models (1 January - 20 March)

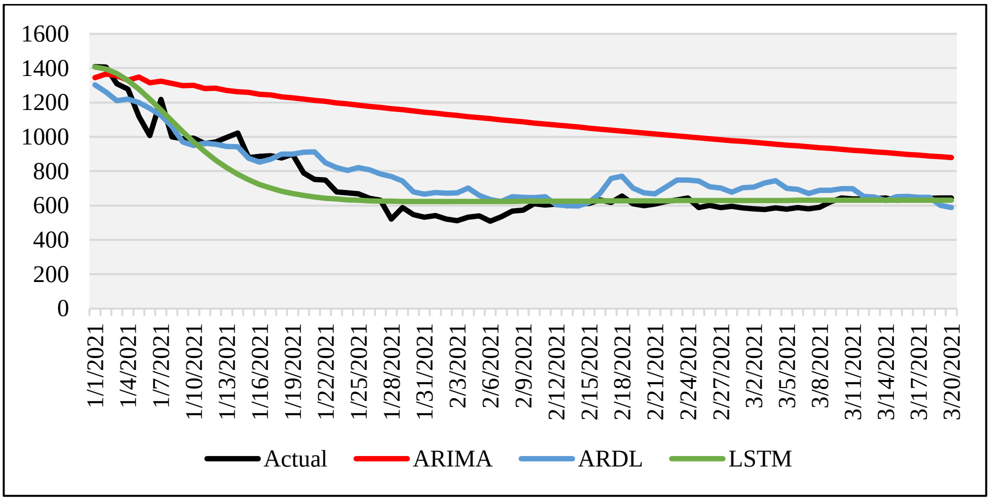
	Cases' Model				Deaths' Model			
Forecasting Model	RMSE	MAE	MAPE	SMAPE	RMSE	MAE	MAPE	SMAPE
ARIMA	0.492	0.456	7.078	6.795	0.355	0.304	7.796	7.500
ARDL	0.148	0.121	1.876	1.851	0.237	0.198	5.073	5.261
LSTM	0.109	0.081	1.238	1.238	0.115	0.097	2.479	2.455

Source: Authors' estimation using EViews 10.

It is noticed that based on the different error measures, the LSTM model is the best to forecast daily cases followed by the ARDL. Likewise, the LSTM model is the most accurate one for daily deaths followed by the ARDL model. The graphical investigation confirms the latter results. Figure 12 and Figure 13 compare the results of the preceding methods for the ex-post period (1 January - 20 March), and illustrate that LSTM and ARDL models provide the closest results to the actual data for daily cases and deaths. Moreover, across the two models together, the MAPE criterion compare all the methods in the two models and indicates that the LSTM method of new cases model is the best one across all models (it obtains the least errors), followed by the LSTM method for the new deaths model. Moreover, according to the Lewis (1982) criteria, all the methods of the two models are accurate since they have forecasting errors less than 10%.

Figure 12

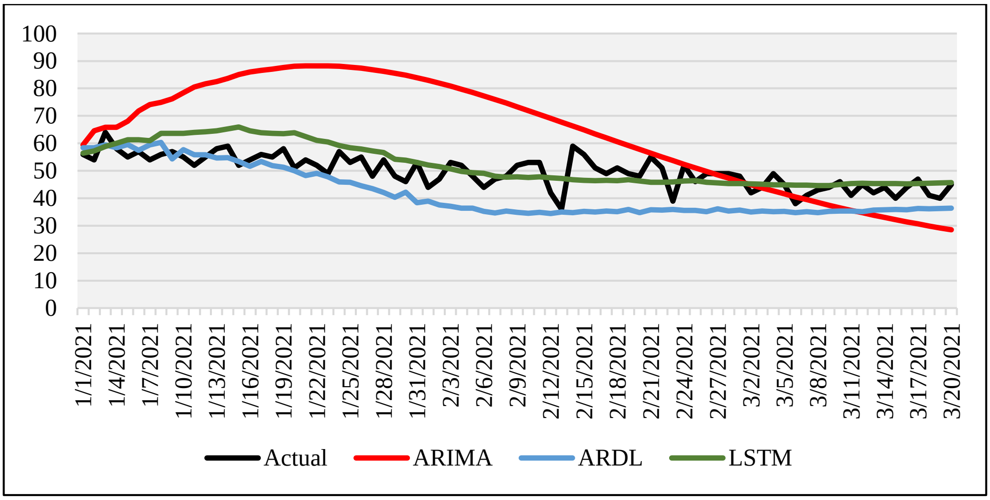
Ex-Post Forecasting of Daily Cases Using Different Methods



Authors' diagram using the estimated ex-post results.

Figure 13

Ex-Post Forecasting of Daily Deaths Using Different Methods



Authors' diagram using the estimated ex-post results.

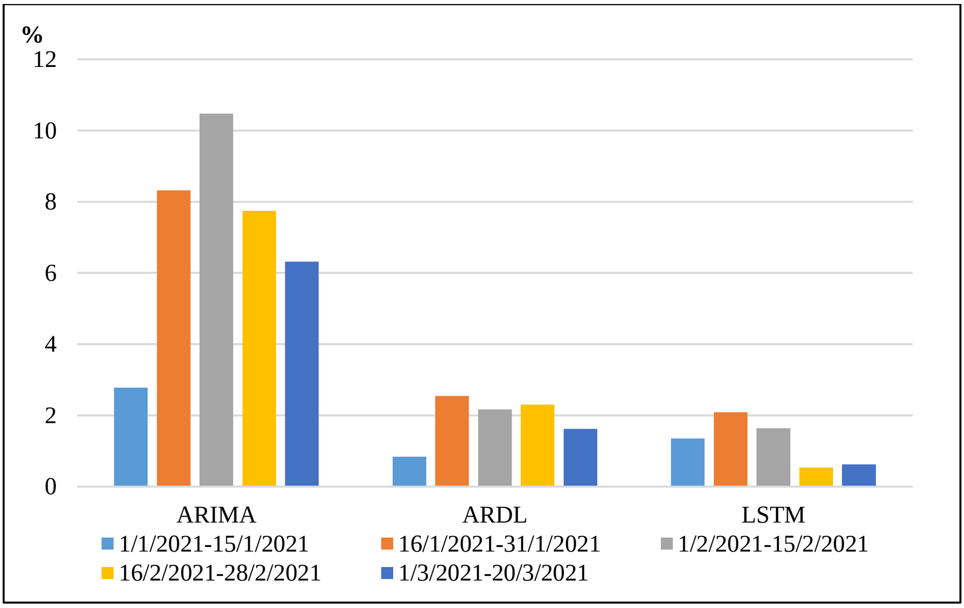
6.4.2. Forecasting errors at different time horizons "Absolute Percentage Error"

To compare the forecasting performance of the estimated methods over time, the Absolute Percentage Error (APE) between the predicted and actual values of cases and deaths respectively for the three methods in the two models has calculated on a daily basis along the ex-post forecasting period from 1 January - 20 March 2021. Figures 14 and 15 show the average of the APE every two weeks of the 79 ex-post period for cases and deaths models. The error of the results obtained by the LSTM model is the least among other models.

Regarding the cases' model, the APE of the ARIMA model has increased with time from 2.78% in the first period to reach a maximum of 10.48% at the middle (third period), then kept decreasing to reach 6.32% at the last period of the ex-post forecasting. Speaking about the APE of the ARDL model, it started at 0.84% at the first period, reaching its highest value of 2.55% in the second period, and finally it reached 1.62% on the last period. By same token, the LSTM-APE started at 1.35%, reaching a maximum of 2.09% in the second period then embarked on decreasing to reach 0.29%. Talking about the time horizons, the ARDL has the lowest APE in the first period, but the LSTM is the best in all other periods and its APE decreases continually over time. Consequently, it can be concluded that the LSTM can be used to forecast the number of infections of Covid-19 especially in the long term.

Figure 14

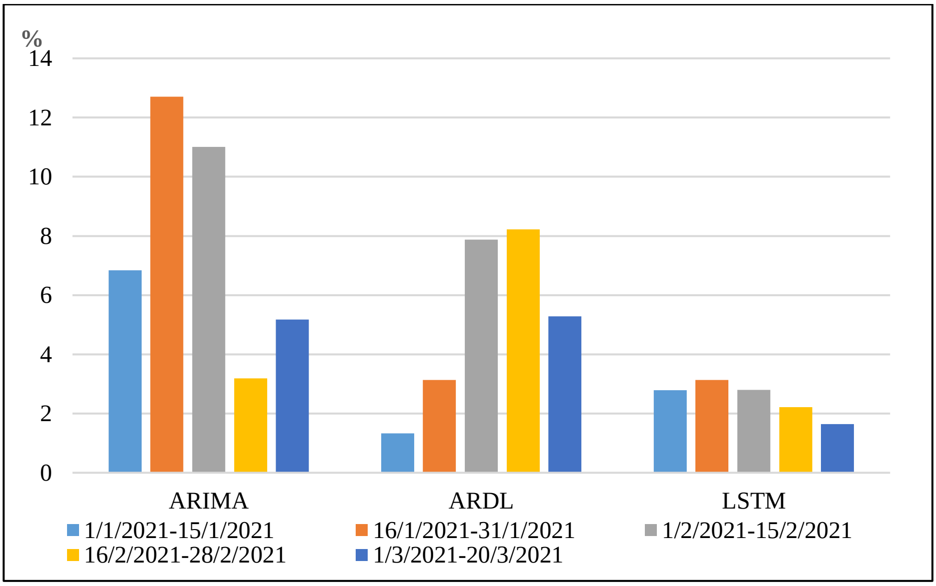
Absolute Percentage Error for ARIMA, ARDL and LSTM- Cases' Model



Authors' diagram using the estimated ex-post results.

Figure 15

Absolute Percentage Error for ARIMA, ARDL and LSTM- Deaths' Model



Authors' diagram using the estimated ex-post results.

Speaking of the deaths' model, the APE of the ARIMA model has increased with time from 6.84% in the beginning to reach a maximum of 12.71% after one month then fluctuated to reach 5.17% at the last period of the ex-post forecasting. Speaking about the APE of the ARDL model, it started at 1.33%, reaching its highest value of 8.23% after 2 months, and finally it approached 5.28% on the last period. Similarly, the LSTM-APE started at 2.77%, reaching a maximum of 3.14% in the second period then it oscillated to reach 1.65%. Comparing the different methods, the LSTM is the best in the long term again, except for the first and second periods; the ARDL has the lowest forecasting errors. Hence, it can be deduced that the LSTM can be utilized to forecast the death toll of Covid-19 precisely in the long run, whereas the ARDL is the best in the short run.

7. Discussion

Time series and econometric models can be useful in predicting confirmed infections, recoveries and deaths as well as future pandemic spread rates if the spread of the virus does not change very oddly. However, various machine-learning models have been utilized for the same purpose. Undoubtedly, it is globally well known that this virus is new and has the potential to be highly transmissible in addition to growing fears of a powerful other waves.

Regarding the situation in Egypt, the infection and death rates kept on rising. For instance, the infection rate increased form 0.1% on September 2020 to 0.19% after six months –on March 2021. Resulting in about 94% increase during the same period. For the same time horizon, the death rate had been raised from 5.6% to 5.9%. On the other side, the recovery rate has dropped by 9.3%. More than 10 months later, on 31 January 2022 the total number of cases stood at 425,911 with an increase of 119% than March 2021. On the other side, the number of mortalities grown by 96% to reach 22,635.
Speaking about the case in Egypt and globally, the current values of the pandemic disclose that the death rate in Egypt is higher than its worldwide counterpart by 4.44%. Likewise, the proportion of recovered individuals is 1.82% lower in Egypt than the worldwide percentage.

As regards the spread of the pandemic on a 2-years course, Egypt has faced many waves of infections. The country experienced two waves during the first year of the epidemic. Following, a third wave begun during 2021, particularly on the tenth of April. Later on, another wave has hit the country starting 1st October and the number of infections kept on increasing to record an extreme of 2,223 cases on the 31st of January 2022.

In fact, Egypt has reacted early to tackle the spread of the virus. Three-months after directing the closure decisions, at the end of June 2020 the government started reopening to avoid GDP decline, income losses and poverty increase. Despite resulting in an upward movement, the number of infections kept on decreasing since mid-July 2020. This may be because a growing number of Egyptians is taking social distancing procedures seriously and committing to protective measures to limit coronavirus outburst as a result of amplified level of awareness. At the beginning of January 2021, the Egyptian government re-imposed some closing measures as a proactive step towards the likely second wave. Consequently, the use of transportation facilities has dropped. Additionally, going to workplaces has declined while staying at homes again has increased. Such decisions have been effective in easing the effect of the second wave. Later on, after smoothing the outcome of the second wave, schools and universities have been re-opened by mid-February. On the other side, the 3rd wave which occurred during the period April 2021 to June 2021 was characterized by an average number of infections of 932. Moreover, the 4th wave may have started by mid-September 2021 and is still ongoing with a growing rate. Thus, it can be deduced that recently the Egyptians are not taking social distancing measures seriously and not committing to following the preventive measures to limit coronavirus outbreak despite the assured governmental regulations regarding various preventive restrictions.

Chiefly, the results of the utilized ARDL model conclude that mobility of population is affecting the incidence of new cases of Covid-19 significantly. Especially, an increase in mobility related to workplaces by 1% is associated with higher incidence of Covid-19 by 10%. Therefore, the spread of Covid-19 is caused by increasing mobility to workplaces. Moreover, the spread of Covid-19 by 1% more tends to increase the possibility of death by 0.77% on average in the long run.

Ex-post forecasting is performed for a period of 79 days (1 Jan 2021-20 March 2021) using time series, econometric and deep learning techniques to choose the most accurate forecast for each model to be used in the ex-ante forecasting of the new cases and new deaths in other papers. It was found that the best fit model with the least values of the error measures to model daily cases and deaths is the deep learning LSTM one followed by ARDL then ARIMA. Moreover, ARDL performs better than others in the short run, whereas the LSTM is the best in the long-run horizon.

In spite of showing better prediction performance and the fact that machine learning techniques can be applied in order to handle many problems that are targeted by traditional econometric and statistical models, there are some differences between them in terms of their objectives, approaches and settings among others. To explain, according to Athey and Imbens (2019), the so called classic statistical or econometric models provide estimates with formal properties of a type that many of the machine learning algorithms do not usually deliver. Such properties include the asymptotic –large sample– properties of estimators and tests including normality, efficiency as well as consistency. On the contrary, the main attention of machine learning methods is the properties of the algorithms. Further, econometric models establish probabilistic models intended to depict various phenomena, while machine learning uses algorithms that depend on learning from their mistakes.

Other key points regarding the differences between machine learning models and statistical or econometric models are their estimation steps and outcomes. To clarify, having machine learning models trained, it is difficult to understand their inner working steps. That being the case, they are often thought of as being "black boxes", making them hard to comprehend and to clarify their behavior. Speaking of the outcomes, econometric models result in estimates of population parameters that can be easily interpreted and instantly give insights about the effect of each regressor on the target variable. Thus, econometric and statistical techniques result in an estimated function that once reported can be utilized by any stakeholder, while machine learning models provide a written code that needs more effort to be operationalized-employed for prediction purposes. Finally, it can be recommended that statistical and econometric techniques are irreplaceable by machine learning techniques. Yet, they should be integrated.

References

1. ArunKumar, K. E., Kalaga, D. V., Kumar, C. M. S., Kawaji, M., & Brenza, T. M. (2022). Comparative analysis of Gated Recurrent Units (GRU), Long Short-Term Memory (LSTM) cells, Autoregressive Integrated Moving Average (ARIMA), Seasonal Autoregressive Integrated Moving Average (SARIMA) for forecasting COVID-19 trends. *Alexandria Engineering Journal*.
2. Athey, S., & Imbens, G. W. (2019). Machine learning methods that economists should know about. *Annual Review of Economics*, *11*, 685-725.
3. Benvenuto, D., Giovanetti, M., Vassallo, L., Angeletti, S., & Ciccozzi, M. (2020). Application of the ARIMA model on the COVID-2019 epidemic dataset. Data in brief, 105340.
4. Breisinger, C., Raouf, M., Wiebelt, M., Kamaly, A., & Karara, M. (2020). Impact of COVID-19 on the Egyptian economy: Economic sectors, jobs, and households (Vol. 6). Intl Food Policy Res Inst.
5. Dehesh, T., Mardani-Fard, H. A., & Dehesh, P. (2020). Forecasting of covid-19 confirmed cases in different countries with arima models. *medRxiv*.
6. El-Ghitany, E. (2020). A Short-Term Forecast Scenario for COVID-19 Epidemic and Allocated Hospital Readiness in Egypt.
7. Elmousalami, H. H., & Hassanien, A. E. (2020). Day Level Forecasting for Coronavirus Disease (COVID-19) Spread: Analysis, Modeling and Recommendations. *arXiv*. 2020 mar.
8. 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 COVID-19 outbreak in Saudi Arabia. *Process Safety and Environmental Protection*, *149*, 223-233.
9. Elsworth, S., & G¨uttel, S. (2020). Time Series Forecasting Using LSTM Networks: A Symbolic Approach. *arXiv*.
10. Feridun, M. (2009), ‘Unit Roots, Structural Breaks and Cointegration Analysis: A Review of the Available Processes and Procedures and an Application’, Presented at The Macroeconomics and Financial Economics Workshop: Recent Developments in Theory and Empirical Modelling workshop, held from October 8 to October 9 at Eastern Mediterranean University.
11. Google LLC "Google COVID-19 Community Mobility Reports". https://www.google.com/covid19/mobility/ Accessed: <16 August 2020>
12. Gujarati, D. N., & Porter, D. C. (2009). *Basic econometrics* 5th Ed." McGraw-Hill International Edition, Boston.
13. Han, E., Tan, M. M. J., Turk, E., Sridhar, D., Leung, G. M., Shibuya, K., ... & Legido-Quigley, H. (2020). Lessons learnt from easing COVID-19 restrictions: an analysis of countries and regions in Asia Pacific and Europe. *The Lancet*, *396*(10261), 1525-1534.
14. Hochreiter, S., & Schmidhuber, J. (1997). Long Short-Term Memory. Neural Computation, 9, 1735--1780.
15. https://www.worldometers.info/coronavirus/ Accessed: <16 August 2020>
16. Keras. (n.d.). Retrieved 3 2021, from https://keras.io/
17. Kumar, P., Kalita, H., Patairiya, S., Sharma, Y. D., Nanda, C., Rani, M. & Bhagavathula, A. S. (2020). Forecasting the dynamics of COVID-19 Pandemic in Top 15 countries in April 2020: ARIMA Model with Machine Learning Approach. *medRxiv*.
18. Lauer, S. A., Grantz, K. H., Bi, Q., Jones, F. K., Zheng, Q., Meredith, H. R. & Lessler, J. (2020). The incubation period of coronavirus disease 2019 (COVID-19) from publicly reported confirmed cases: estimation and application. *Annals of internal medicine*, *172*(9), 577-582.
19. Li, G., Song, H., and Witt, S.F. (2005), ‘Recent Developments in Econometric Modelling and Forecasting’, *Journal of Travel Research*, Vol.44: 82-99.

20. Marzouk, M., Elshaboury, N., Abdel-Latif, A., & Azab, S. (2021). Deep learning model for forecasting COVID-19 outbreak in Egypt. *Process Safety and Environmental Protection*, *153*, 363-375.
21. Masum, S., Liu, Y., & Chiverton, J. (2018). Multi-step Time Series Forecasting of Electric Load Using Machine Learning Models. International Conference on Artificial Intelligence and Soft Computing (ICAISC 2018), (pp. 148-159).
22. Mohamadou, Y., Halidou, A., & Kapen, P. T. (2020). A review of mathematical modeling, artificial intelligence and datasets used in the study, prediction and management of COVID-19. *Applied Intelligence*, *50*(11), 3913-3925.
23. Namini, S., Tavakoli, N., & Namin, A. (2018). A Comparison of ARIMA and LSTM in Forecasting Time Series. 17th IEEE International Conference on Machine Learning and Applications.
24. Negara, H. R. P., Kusuma, J. W., Apriansyah, D., & Tamur, M. (2021, May). Computing the auto regressive distributed lag (ARDL) method in forecasting COVID-19 data: A case study of NTB Province until the end of 2020. In *Journal of Physics: Conference Series* (Vol. 1882, No. 1, p. 012037). IOP Publishing.
25. Nosier, S. A. (2012), Estimating the International Tourism Demand for Egypt 'An Econometric Approach', PhD thesis, the University of Hull: UK.
26. Nosier, S. A. (2018), The Effect of Electoral Democracy On Economic Growth in Egypt (1970 – 2015), *The Afro Asian Journal of Social Sciences,* AAJOSS, VOLUME IX, NO I, Quarter I.
27. Nutarelli, F. (2022). At the intersection between machine learning and econometrics: theory and applications.
28. Omran, N. F., Abd-el Ghany, S. F., Saleh, H., Ali, A. A., Gumaei, A., & Al-Rakhami, M. (2021). Applying Deep Learning Methods on Time-Series Data for Forecasting COVID-19 in Egypt, Kuwait, and Saudi Arabia. Complexity, 2021.
29. Paez, Antonio. 2020. "Using Google Community Mobility Reports to Investigate the Incidence of COVID-19 in the United States." Transport Findings, May.
30. Pesaran, M.H., Shin, Y., and Smith, R.J. (2001), 'Bounds Testing Approaches to the Analysis of Level Relationships', *Journal of Applied Econometrics*, Vol.16: 289-326.
31. Saba, A. I., & Elsheikh, A. H. (2020). Forecasting the prevalence of COVID-19 outbreak in Egypt using nonlinear autoregressive artificial neural networks. *Process Safety and Environmental Protection*.
32. scikit-learn: Machine Learning in Python. (n.d.). Retrieved 3 2021, from https://scikit-learn.org
33. Song, H., Witt, S.F., and Li G. (2009), *The Advanced Econometrics of Tourism Demand*, London: Taylor & Francis.
34. Stock, J.H. (1987), 'Testing for Common Trends', *Journal of the American Statistical Association*, Vol.83: 1097-1107.
35. The Bibliotheca Alexandrina Supercomputing Facitliy. (n.d.). Retrieved 3 2021, from https://hpc.bibalex.org/
36. Tideswell, C., Mules, T., and Faulkner, B. (2001), 'An Integrative Approach to Tourism Forecasting: A Glance in the Rearview Mirror', *Journal of Travel Research*, Vol. 40: 162-171.
37. Tayarani-N, M. H. (2020). Applications of artificial intelligence in battling against covid-19: a literature review. *Chaos, Solitons & Fractals*, 110338.
38. Yasrab, R., Pound, M. P., French, A. P., & Pridmore, T. P. (2020). PhenomNet: bridging phenotype-genotype gap: a CNN-LSTM based automatic plant root anatomization system. *bioRxiv*.
39. Yousaf, M., Zahir, S., Riaz, M., Hussain, S. M., & Shah, K. (2020). Statistical analysis of forecasting COVID-19 for upcoming month in Pakistan. *Chaos, Solitons & Fractals*, 109926.