Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company’s public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.
Artificial Intelligence-Based Model for Predicting the Effect of Governments’ Measures on Community Mobility

Nada Osman\textsuperscript{b}, Marwan Torki\textsuperscript{b}, Mustafa ElNainay\textsuperscript{a,b,c}, Abdulrahman AlHaidari\textsuperscript{a}, Emad Nabil\textsuperscript{a,d}

\textsuperscript{a}Faculty of Computer Science and Information Systems, Islamic University of Madinah, Madinah, Saudi Arabia
\textsuperscript{b}Department of Computer and Systems Engineering, Alexandria University, Alexandria, Egypt
\textsuperscript{c}Faculty of Computer Science and Engineering, AlAlamein International University, Matrouh, Egypt
\textsuperscript{d}Faculty of Computers and Artificial Intelligence, Cairo University, Giza, Egypt

Received 2 December 2020; accepted 17 February 2021
Available online 3 March 2021

KEYWORDS
COVID-19; Community mobility; Time series prediction; Autoregression; Convolution neural network; Long short-term memory network

Abstract Mobility is considered one of the main reasons for the COVID-19 spread. Predicting the effect of control measures on mobility is essential to apply effective decisions. This work proposes an AI-based model for mobility changes prediction. The proposed CNN-LSTM with Autoregression has achieved the best results compared to other investigated models. Results show that the proposed model can predict the effect of precaution control measures on future community mobility with minimum loss. The mean absolute error over all countries in the study is 5.3. For Egypt and Saudi Arabia, the model achieved an MAE loss of 4.6 and 3.7 consecutively.

© 2021 THE AUTHORS. Published by Elsevier BV on behalf of Faculty of Engineering, Alexandria University. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

1. Introduction

Since the World Health Organization (WHO) announcement of COVID-19 outbreak as a pandemic, different countries responded differently, taking relaxed or aggressive control measures to reduce the disease spread. While some countries have decided to make general awareness of the disease and urged citizens to stay home and apply social distancing as much as they can, other countries have decided to go the other extreme by applying complete lockdown for extended periods in an attempt to contain the disease. Despite this, the number of infected cases kept growing in many countries, even those applying strict control measures to exceed three million cases in few months and hundreds of thousand death cases.

The Kingdom of Saudi Arabia, for example, has taken strong precautionary control measures to limit the spread of COVID-19, whose number of confirmed cases reached more than 350,000 cases until the first week of November 2020. Despite the relatively big number of cases, the number of deaths is very few as a result of the applied precautionary
and control measures; the percentage of death cases is about 1.6. The disease spread has a direct relation to the community mobility as the main source of infection is the human-to-human transfer. On the other hand, some control measures have major consequences on the economic sectors and the national income.

Many researchers have studied the effect of different government control measures for different countries on the community mobility patterns and the disease spread pattern using only mathematical models as a result of the scarcity of data available for one country. Therefore, we propose in this paper to collect data of many countries to obtain sufficient data to build an artificial intelligence-based model to predict the effectiveness of various control measures on community mobility patterns to provide insights to decision makers on the effectiveness of these control measures as compared to its potential economic consequences.

To achieve this goal, we utilize the mobility change data provided by Google [1] and correlate it with the control measures applied by the government to build an artificial intelligence model to predict future mobility changes for every control measure. In this work, artificial intelligence-based model has been proposed for the mobility changes prediction. First, an autoregressive model has been developed to be used as a baseline and then enhanced models using Convolutional Neural Network and Long Short Term Memory (CNN-LSTM) have been developed. The model has gone through iterations of tuning and then evaluated. The results show that the proposed parallel Convolutional Neural Network - Long Short Term Memory (CNN-LSTM) with Autoregression and extended features can predict the future mobility changes of potentially applied control measures with the average mean absolute error over all countries included in the study = 5.3. This model can be of great benefit during the current and future pandemics as decision makers may use it to predict the mobility changes of every potential control measure and choose the one that decrease the mobility with minimum economic consequences.

The paper is organized as follows. A survey of related work is discussed in the next section. The datasets preparation and preprocessing process is described in Section 3. The proposed prediction models and the used features are explained in Section 4. The conducted experiments and the evaluation results are presented in Section 5. Finally, conclusions and future work are presented in Section 6.

2. Related work

In this section, survey of the related work is presented. First, a survey on unsupervised clustering techniques proposed to cluster countries in response to COVID-19 is presented. Then, artificial intelligence-based time-series forecasting models are surveyed. Finally, the previous research work focuses on predicting the effectiveness of government control measures in response to pandemics is then presented.

As COVID-19 is spreading, countries have been impacted differently because of various factors, one of which is the nation’s healthcare capacity. Depending on the outbreak intensity and infection transmission rate, countries need to make decisions promptly to minimize the repercussions. Studies have explored the use of machine learning to draw similarities between countries, based on the available data [2]. Most earlier investigations have applied K-means and Principal Component Analysis (PCA) to cluster countries based on different combinations of data such as COVID-19, demographic, and socioeconomic. In [3], a simple clustering model, with data before the COVID-19 outbreak, such as a country’s air contamination and health insurance, was employed by Carollo-Larco et al. to cluster 155 countries. By using a one-way analysis of variance (ANOVA), authors clustered those countries using the K-means algorithm with three component analysis (PC3). Although the model could not capture the mortality rate, it was able to arrange those countries based on the number of confirmed COVID-19 cases.

Hu et al. [4] used a more sophisticated technique than Carollo-Larco et al. by utilizing a modified stacked autoencoder (MAE) to model COVID-19 contagion. Data from the World Health Organization for the period between 11 January to 27 February 2020, in addition to COVID-19 forecasting in China, was used to extract a model for finding corresponding provinces. As reported by [4], the model produced the total cases for three consecutive months, 20 January to 20 April 2020, using multiple-step forecasting. It was shown that out of the total of thirty-four cities, nine clusters were obtained, and the minimum error rate, the best, was a result of the 10-step prediction at 0.73%.

A similar approach was used in [5], where the authors proposed a trained Topological Autoencoder (TA), a streamlined soft-managed TA. The data was obtained by the Center for Systems Science and Engineering (CSSE) for COVID-19 patients from 240 countries. Then, this was utilized to create a 2-D clusters for those nations. In [6], Health data, death-causing diseases, of 146 countries and compiled it from different sources. Hierarchical clustering and gap statistical analyses were implemented to form the grouping, as the main goal is to determine the cause of death. Ultimately, they found the clusters to be statistically significant. In another work, Alonso et al. [7] applied Curvilinear Component Analysis (CCA) to cluster the European Union (EU) countries depending on their microeconomic data, then compared it with cluster obtained by self-organizing maps (SOM); however, CCA took advantage of dimensionality reduction where SOM did not. The author concluded that the clusters of both were identical, neglecting outliers. In our work, all countries that have their control measures indexed in the ACAPS (stands for Assessment Capacities Project) dataset [8] are used in an attempt to enlarge the dataset size used to develop the prediction model. Data preparation and preprocessing process has been applied as explained in Section 3 to unify the categories of control measures applied in different countries before using the data to develop our prediction models.

Number of time series prediction models have been proposed in the literature to develop prediction models for different domains. In [9], Barrow and Crone compared seven variants of the AdaBoost time series prediction models. They conducted experiments to explain the strength and weaknesses of each model. This comparison is a useful guide for using AdaBoost model variants. One of the options for time series forecasting is the support vector regression. In [10], Gupta et al. used support vector regression in financial time series forecasting. The authors developed a model called Twin support vector regression that can work with noisy data. The algorithm is compared to the standard support vector regression
using some performance metrics and proved its superiority. Forty-four open datasets for training and testing are used to develop the proposed model. In [11], Ebhuhonm et al. developed a model to predict the monthly malaria cases in KwaZulu-Natal, South Africa; they used the Seasonal Autoregressive Integrated Moving Average (SARIMA) time series prediction model by adopting the Box-Jenkins approach. The model trained on dataset in the time period (January 2005 - December 2013) and tested on a different dataset in the time period (January - December 2014).

Some AI-based models have been proposed to predict the future of mobility change based on the previous data. Google mobility change reports [1] published by Google contain a time series data about daily mobility pattern changes of many countries during the COVID-19 pandemic. In [12], the residential mobility change during COVID-19 quarantine in Philippine was predicted in the period from 12 April 2020 to 30 April 2020. Google Mobility reports from 15 February 2020 to 11 April 2020 are used and the Autoregressive Integrated Moving Average (ARIMA) time series prediction model has been utilized to make the prediction.

Several research papers have investigated the effectiveness of mobility-related features on the disease spread. All papers proposed deterministic or mathematical-based prediction modeling.

In [13], Nyabadza et al. have investigated the effectiveness of different levels of social distancing in South Africa on the transmission dynamics of COVID-19 using the SEIR mathematical model and Google mobility change reports. The study concluded that applying social distancing at the current level is not sufficient to flatten the growth rate of the cases, increasing the level by 2% will decrease the total number of cumulative cases by 18%. On the other hand, relaxing the social distancing level will significantly increase the number of new cases. A similar study on the effectiveness of social distancing has been presented in [14] for the US using optimal control theory. Prem et al. [15] have applied the SEIR mathematical model to investigate the effect of different social mixing patterns based on schools and events control measures on controlling the COVID-19 outbreak in Wuhan. They have used synthetic location-specific contact patterns to simulate the population mixing scenarios. Roques et al. [16] have estimated the effect of a one-month lockdown in France on the reduction of the contact rate and consequently, the reproduction number of the disease spread using SIRD mathematical model. The results show a reduction of a factor of 7 in the reproduction rate compared to before the lockdown period. The effect of travel restrictions has been studied in [17] using the SEIR mathematical model. Through correlating mobility to the COVID-19 spread model, the results show that unconstrained mobility of International travel would have significantly increased the COVID-19 spread.

In [18], Banholzer et al. have studied the impact of 7 different control measures on the reduction of reported new cases of COVID-19 for 20 countries. The study applied statistical methods and concluded that venue closures, gathering bans, and border closures are the most effective measures, while school closures and complete lockdowns are the least effective measures. The results though are not credible due to the insufficient data and the simplistic of the used analysis methods. A similar study for the impact of 6 different control measures has been conducted in the UK using SEIR mathematical model [19] and concluded that a short period of intense lockdown may keep the growth rate at a safe level.

While all previous work studying the effect of control measures have applied mathematical modeling techniques, Artificial intelligence-based modeling is proposed in our work to predict the community mobility change patterns in response to different potential control measures. The next section presents the datasets used in our study and the preparation and preprocessing steps applied before that data is used to develop the proposed model.

3. Data preparation and preprocessing

The work in this paper depends on two data sources. The first data source is COVID-19 community mobility reports provided by Google [1] that show the daily changes of the visit ratio (as an indicator for mobility) to main places categorized into six categories: Retail and recreation, Grocery and pharmacy, Parks, Transit stations, Workplaces, and Residential. The second data source is the dataset provided by independent information provide called ACAPS (stands for Assessment Capacities Project) [8] that puts together all the control measures applied by governments worldwide to reduce the spread of COVID-19. Measures are categorized into five main categories: Social distancing, Movement restrictions, Public health measures, Social and economic measures, and lockdowns, and each is broken into several types, giving a total of 34 subcategories. The following subsections explain the preprocessing steps applied to the data.

3.1. Control measures data filtering and categorization

Not all of the reported control measures in the ACAPS dataset affect community mobility. Control measures, such as the ones related to strengthening the public health systems or to the economy, do not have a considerable effect on community mobility. Therefore, a filtering procedure is carried out to exclude mobility-unrelated features, reducing the number of control measures to 13 types.

To further improve the control measures data, highly correlated measures are merged into a single category. For example, a single control measure is enough to represent all of the actions related to closing public services. Hence, the number of control measures used is reduced from 13 to 8 categories, as shown in Table 1.

The control measures in the ACAPS dataset are reported as the textual announcement made by the government (e.g. All flights to and from Saudi Arabia suspended for two weeks starting 15 March, Northern Italian regions (NUTS2) officially locked down, The Government of Egypt will begin enforcing a nighttime curfew from 7:00 p.m. to 6:00 a.m., etc). Therefore, the textual reported actions are manually engineered to be replaced by numerical levels in each defined category, based on the severity of the taken action, as summarized in the third column in Table 1.

3.2. Data sources merging and alignment

First, the data sources are filtered to include only the countries that have reports for both mobility changes and control measures, summing up to 119 countries. Afterward, the two data
sources are merged by the date of applying a control measure, where the base is the daily mobility change, and the control measures features are added on the day of its application and kept the same until a new measure is applied.

The start date of reporting mobility changes in all countries is the 15th of February, which is the start date of the created dataset as well. The end date of the dataset used in this work is the 27th of July, with a total of 164 days in each of the 119 countries.

3.3. Training and testing splitting

The data is split into training and testing by countries, where 85% of the countries (99 countries) are used for training, and 15% of them (20 countries) are used for testing. Each country is treated as a time series sample of 164 data points (the 164 reporting days of each country).

4. Prediction models

Community mobility forecasting can be formulated as a supervised machine learning prediction problem, given a time series of past days of size $W_i$ (input window size), an ML model could be trained to predict the future mobility changes in $W_o$ days (output window size). Four models are investigated in our work: an Autoregressive model is developed as a baseline model that uses the mobility changes data only, three enhanced models using Convolutional Neural Network and Long Short Term Memory (CNN-LSTM) have then been developed. The prediction of the later three models depend on both community mobility changes during the past days and the applied control measures. This section explains the architecture of the proposed models and the features used with each model.

4.1. Autoregressive model (baseline)

The simplest forecasting time series model is the autoregressive model. Therefore, it has been developed as the baseline for our community mobility prediction problem. The model uses only the time series of mobility changes in the $W_i$ past days to predict the changes in the $W_o$ future days. Applied control measures have not been used in this model. We started with a single linear layer model, where the prediction is a linear combination of the past time series, as shown in Fig. 1a. Additionally, an upgraded autoregressive model is used, where multiple RELU layers are used instead of the single linear layer as depicted in Fig. 1b.

4.2. CNN-LSTM model

The baseline model used only the mobility changes time series data to predict future changes. Since our target is to build a
model that predicts the mobility changes in response to different potential control measures, an enhanced CNN-LSTM model is developed. The model takes three types of inputs:

- Control measures features, represented as a window of size $W_i \times 8$, where 8 is the number of control measures features, as described in Table 1.
- Mobility changes features, represented as a window of size $W_i \times 6$, where 6 is the number of community mobility changes categories, as described in the previous section.
- Mobility changes separated time series, represented as six vectors of size $W_i$.

The control measures window and the mobility changes window are fed to parallel CNN models. Each CNN model consists of three different kernel sizes. The CNN networks capture the low-level relations inside the control measures window and the mobility changes window, where the output is a better feature embedding. The extracted features from the CNN models are passed to an LSTM network, capturing the input’s time-dependencies. Finally, the parallel models’ outputs are concatenated and passed to a fully connected network to produce the final prediction. Fig. 2 illustrates the architecture of the described model.
4.3. Parallel CNN-autoregressive model

A simple addition to the model is to benefit from the autoregressive model by using it parallel to the CNN-LSTM models, as shown in Fig. 3. The final prediction depends on the features extracted by the CNN-LSTM models and the autoregressive model’s output. This addition improved the prediction results, as will be discussed in the next section.

4.4. Extra features

To further enhance the model, features engineering is applied to add extra features that improve the prediction performance. Two types of features are added:

- The first is related to the control measures features, which is the time duration of applying a control measure for all measures (extra eight features per day).
- The second is related to the mobility changes features, where two averages over the mobility changes time series are applied: a fixed window average of five days and an average computed from the last date of update of any control measure.

The extra features add extra channels to the CNN model; the control measures features are two channels window (the measure and its duration), and the mobility changes features are three channels window (mobility change, fixed average, last control measure updated average). Fig. 4 illustrates the final community mobility forecasting model’s overall structure with all components added.

4.5. Training and parameters settings

The dataset has 99 training samples, with each sample consists of 164 data points. In each training epoch, the samples are chosen at random, and the model begins to slide a window of size \( W_i + W_o \) over the training data points. If the window starts at \( X_i \), the model uses the points \((X_i, X_{i+1}, X_{i+2}, \ldots, X_{i+W_i})\) as the input, and the points \((X_{i+W_i+1}, \ldots, X_{i+W_i+W_o})\) as the output. The epoch ends when the model passes over all of the training samples, and the number of epochs is set by early stopping when the model starts to overfit. To choose the best values for \( W_i \) and \( W_o \), we have conducted control experiments using different values, and the ones producing the lowest loss have been chosen as the final windows sizes.

We used Adam optimizer and conducted tuning experiments for the best learning rate and step decay values. For regularization, all layers use dropout with the value of 0.2.

5. Experiments and evaluation

The accuracy of the proposed models’ predictions is measured by the average difference between the predicted mobility changes and the actual changes, using two evaluation measures: mean square error (MSE), and mean absolute error (MAE). First comparison of the four developed models have been conducted to choose the best model. Next, experiments to tune the model parameters have been conducted focused on the training and the validation sets (the 99 countries). The proposed model is evaluated using the countries in the test set (the 20 countries). Each community mobility category of the six categories presented in Section 3 is evaluated independently. The summation of the six categories MSE presents the overall MSE.

5.1. Models evaluation experiments

Different models presented in the previous section are evaluated and compared to each other, starting with the baseline autoregressive model to the final parallel CNN-LSTM autoregressive prediction model with the extended features. All of the evaluation experiments in this subsection use a fixed learning rate of 0.001, an input window size of 10 days \( (W_i = 10) \), and an output window size of 5 days \( (W_o = 5) \).

5.1.1. Autoregressive models evaluation

Experiments have been conducted to evaluate the two variations of the baseline Autoregressive model: single-linear-layer and multi-RELU-layers autoregressive models. AS explained before in Section 4, the Autoregressive models are evaluated using the mobility change data only. Table 2 illustrates the comparison between the two variations. The results show that
Artificial intelligence-based model for predicting the effect of governments’ measures

Fig. 4 The final overall structure of the community mobility prediction model.

### Table 2 MSE of the autoregressive models.

| Mobility Change Category      | One linear layer | Multi-RELU layers |
|------------------------------|------------------|-------------------|
|                              | Training | Validation | Training | Validation |
| Retail and Recreation         | 1638.02   | 1167.81     | 263.93   | 147.41     |
| Groceries and Pharmacies      | 1028.99   | 839.27      | 235.85   | 174.75     |
| Parks                        | 801.73    | 554.95      | 817.26   | 335.68     |
| Transportation                | 11574.27  | 8910.62     | 294.47   | 236.27     |
| Workplaces                    | 418.56    | 333.43      | 371.2    | 173.15     |
| Residential                   | 187.73    | 124.04      | 55.51    | 24.88      |
| Overall                      | 15649.29  | 11930.12    | 2038.23  | 1092.15    |

### Table 3 MSE of the CNN-LSTM model compared to the MSE of the baseline autoregressive model.

| Mobility Change Category      | Autoregressive Model | CNN-LSTM Model |
|------------------------------|----------------------|----------------|
|                              | Training | Validation | Training | Validation |
| Retail and Recreation         | 263.93   | 147.41     | 220.98   | 117.52     |
| Groceries and Pharmacies      | 235.85   | 174.75     | 243.95   | 172.73     |
| Parks                        | 817.26   | 335.68     | 650.75   | 304.03     |
| Transportation                | 294.47   | 236.27     | 208.46   | 118.87     |
| Workplaces                    | 371.2    | 173.15     | 252.42   | 163.21     |
| Residential                   | 55.51    | 24.88      | 66.71    | 25.3       |
| Overall                      | 2038.23  | 1092.15    | **1643.26** | **901.66** |
adding multiple RELU layers to the model produces a noticeable reduction in the overall MSE loss of more than 85%.

5.1.2. CNN-LSTM model evaluation
Second, we compare the performance of the CNN-LSTM model to the baseline results (autoregressive model). As shown in Table 3, the CNN-LSTM model’s performance is much better than the baseline model’s performance, proving that the community mobility is affected by the control measures applied. Extending the model input features to include the applied control measures improved the prediction. Moreover, the CNN-LSTM structure is more complex, and hence it is reasonable to produce better results with reduction in the overall MSE of 10–20% compared to the multi-RELU-layers autoregressive model.

5.1.3. Parallel CNN-autoregressive model evaluation
The integration between the CNN-LSTM model and the autoregressive model as parallel models, depicted in Fig. 2, improved the prediction performance, as shown in Table 4. The autoregressive model captures the direct relations between the mobility changes over time, allowing for better predictions. The parallel model results in reduction in the overall MSE of 9–10%.

5.1.4. Parallel CNN-autoregressive model with extra features evaluation
Finally, we evaluate the proposed model with the extra features, depicted in Fig. 3. The extra features are evaluated in two stages. First, we evaluate the addition of only the duration of the applied measures. Then, the mobility changes’ averages are added, and the overall model is evaluated. As shown in Table 5, adding the duration window reduced MSE loss, which indicates that the community mobility is affected by how long a control measure is applied. Furthermore, the mobility changes’ averages improved the results, which is reasonable, as the averages are taken over different window sizes (fixed = 5 days, and a changing window depending on the control measures’ updates). The different window sizes introduced new information over the mobility changes from older days, which resulted in better predictions in terms of the overall MSE.

5.2. Tuning experiments
After finalizing the model structure, the hyper-parameters’ tuning remains. The main parameters to tune are activation function, learning rate and step decay, number of layers, and input and output window sizes.

5.2.1. Activation function
Table 6 compares the linear and the RELU activation functions. As shown in the table, the linear function gives better performance, meaning that the mobility changes have an approximately linear relationship with the features space.

5.2.2. Learning rate and step decay
The second experiment is to find the best learning rate and step decay values. A higher or a lower learning rate than 0 produces a very high loss. Table 7 reports the results of different step decay values with different number of epochs. Based on the results, the proposed model uses a learning rate of 0.01 and a step decay of 0.7 every 1000 epochs.

| Table 4 | MSE of the CNN-LSTM model alone VS the MSE of the parallel CNN-Autoregressive model. |
|-----------------|-----------------|-----------------|-----------------|-----------------|
| Mobility Change Category | CNN-LSTM Model | CNN-Autoregressive Model |
| | Training | Validation | Training | Validation |
| Retail and Recreation | 220.98 | 117.52 | 205.88 | 95.24 |
| Groceries and Pharmacies | 243.95 | 172.73 | 201.67 | 168.89 |
| Parks | 650.75 | 304.03 | 604.7 | 284.55 |
| Transportation | 208.46 | 118.87 | 180.85 | 89.84 |
| Workplaces | 252.42 | 163.21 | 254.02 | 151.85 |
| Residential | 66.71 | 25.3 | 49.65 | 21.01 |
| Overall | 1643.26 | 901.66 | 1496.76 | 811.37 |

| Table 5 | Comparing the MSE of the model with and without adding the extra features. |
|-----------------|-----------------|-----------------|-----------------|-----------------|
| Mobility Change Category | No Extra Features | Duration Features | All Extra Features |
| | Tr. | Val. | Tr. | Val. | Tr. | Val. | Tr. | Val. |
| Retail-Recreation | 205.88 | 95.24 | 186.54 | 82.82 | 127.46 | 78.53 |
| Groceries-Pharmacies | 201.67 | 168.89 | 196.86 | 167.25 | 157.76 | 152.84 |
| Parks | 604.7 | 284.55 | 608.14 | 288.76 | 440.14 | 261.72 |
| Transportation | 180.85 | 89.84 | 169.3 | 76.96 | 113.85 | 88.62 |
| Workplaces | 254.02 | 151.85 | 242.18 | 142.25 | 172.26 | 129.22 |
| Residential | 49.65 | 21.01 | 47.01 | 20.83 | 27.00 | 129.22 |
| Overall | 1496.76 | 811.37 | 1450.02 | 778.87 | 1038.48 | 726.97 |
5.2.3. Number of layers
As shown in Table 6, the linear activation function performed better than the RELU function. Therefore, there is no need to keep multiple linear layers. Instead, they can be replaced with a single linear layer. Table 8 shows that a single linear layer produced lower loss. This can be attributed to two reasons: the linear relation between the mobility changes and the input features and that removing the multiple layers reduced the used drop out of the layers and hence, reduced the regularization; enabling the model for better learning.

5.2.4. Input window size
An essential parameter in the model is the input window’s size, how many past days gives the best prediction results. Therefore, a simple control experiment is conducted to choose between three different values: 7 days, 10 days, and 15 days. For this experiment, the number of epochs is fixed to 500 epochs with no decay and a fixed learning rate of 0.01. As shown in Table 9, the best results are achieved with the 10 days window size.

### Table 6  MSE of using linear activation function VS using RELU activation function.
| Mobility Change Category | Linear | Validation | RELU | Validation |
|--------------------------|--------|------------|------|------------|
| Retail and Recreation    | 115.9271 | 76.4833 | 127.4648 | 78.5267 |
| Groceries and Pharmacies | 152.6817 | 152.2067 | 157.7581 | 152.8403 |
| Parks                    | 389.4060 | 261.9567 | 440.1389 | 261.7222 |
| Transportation           | 106.8011 | 66.4929 | 113.8462 | 88.6238 |
| Workplaces               | 171.1694 | 123.7523 | 172.2645 | 129.2225 |
| Residential              | 24.9797  | 14.1786  | 27.0044  | 129.2225 |
| **Overall**              | 960.9650 | 695.0705 | 1038.4769 | 726.9678 |

### Table 7  MSE of different step decay values.
| Mobility Change Category | No decay | 0.75/100 epochs | 0.7/1000 epochs |
|--------------------------|----------|---------------|-----------------|
|                          | Tr.      | Val.          | Tr.             | Val.           |
| Retail-Recreation        | 115.93   | 76.48         | 96.56           | 74.14          |
| Groceries-Pharmacies     | 152.68   | 152.21        | 130.36          | 122.51         |
| Parks                    | 389.41   | 261.96        | 325.66          | 184.31         |
| Transportation           | 106.8    | 66.5          | 85.92           | 70.63          |
| Workplaces               | 171.17   | 123.75        | 154.95          | 133.42         |
| Residential              | 24.98    | 14.18         | 21.42           | 15.45          |
| **Overall**              | 960.97   | 695.07        | 815.07          | 600.45         |

### Table 8  MSE of using single linear layer VS using multiple linear layers.
| Mobility Change Category | Single Layer | Multiple Layers |
|--------------------------|--------------|-----------------|
|                          | Training | Validation | Training | Validation |
| Retail and Recreation    | 57.1809  | 58.7202       | 81.8457  | 70.1072  |
| Groceries and Pharmacies | 104.8443 | 123.2545      | 6        | 123.9067 |
| Parks                    | 242.2163 | 146.2602      | 300.9999 | 159.1872 |
| Transportation           | 44.1125  | 44.5865       | 69.5861  | 54.5180  |
| Workplaces               | 118.8109 | 103.0173      | 140.3398 | 108.8    |
| Residential              | 13.1419  | 12.3903       | 18.9168  | 16.4213  |
| **Overall**              | 580.3127 | 488.2290      | 734.0540 | 532.9451 |

### Table 9  MSE of using different input window sizes (Validation results).
| Mobility Change Category | $W_i = 7$ | $W_i = 10$ | $W_i = 15$ |
|--------------------------|----------|-----------|-------------|
| Retail and Recreation    | 67.9160  | 59.4479   | 61.2684     |
| Groceries and Pharmacies | 122.2598 | 117.4052  | 121.0365    |
| Parks                    | 199.2722 | 170.6434  | 172.1133    |
| Transportation           | 54.7785  | 48.2316   | 51.4652     |
| Workplaces               | 127.1406 | 115.5476  | 120.8443    |
| Residential              | 14.6893  | 13.5251   | 14.2987     |
| **Overall**              | 586.0565 | 524.8007  | 541.0262    |
5.2.5. Output window size

Another control parameter is the output window size. We conducted two experiments: one-day look ahead and 5-days look ahead. Table 10 shows that the model makes better prediction for a single day ahead.

5.3. Test set evaluation

There are 20 countries in the test set, where Table 12 reports the overall testing MSE and MAE of all countries for the four models discussed in Section 4. As shown in the table, the proposed model (Illustrated in Fig. 4 and its configuration summarized in Table 11) achieves the lowest error and the best prediction performance.

To demonstrate the statistical significance of the results, Table 13 reports the mean values of both the actual and the predicted mobility changes, the P-value of the two-tail Welch T-test [20] between the actual and the predicted distributions, and the R-squared value [21]. As shown in the table, the mean of the actual values is almost equal to the mean of the predicted value for all mobility categories. For the T-test P-value, the lowest P-value is 0.2245 (< 0.05), which means that the null hypothesis of having almost the same distribution for actual and predicted values is true. Additionally, the R-squared value is high for all categories (min = 0.72), proving that the actual and the predicted mobility changes are fit to the 1:1 line, as shown in Fig. 5, where the figure illustrates the predicted values on the X-axis and the actual values on the Y-axis. These outcomes prove that there is no significant difference between the predicted and the actual mobility changes.

Furthermore, Table 14 reports the two countries having the best and the worst performance: Cameroon and Slovenia, where the model achieved the best prediction results in Cameroon and the worst results in Slovenia. Many factors affect the prediction results, such as the starting date of applying control measures, the period between the applied measures, and the country’s different demographics. For example, in Slovenia, only five control measures were applied after more than a

---

### Table 10 MSE of using different output window sizes (Validation results).

| Mobility Change Category | \(W_o = 1\) | \(W_o = 5\) |
|--------------------------|-------------|-------------|
| Retail and Recreation    | 35.0544     | 59.4479     |
| Groceries and Pharmacies | 69.1847     | 117.4052    |
| Parks                    | 108.6678    | 170.6434    |
| Transportation           | 26.6296     | 48.2316     |
| Workplaces               | 45.8255     | 115.5476    |
| Residential              | 5.6830      | 13.5251     |
| Overall                  | 291.0450    | 524.8007    |

### Table 11 Proposed Model Configuration.

| Parameter                        | Value                        |
|----------------------------------|------------------------------|
| Optimizer                        | Adam                         |
| Learning Rate                    | 0.01                         |
| Step Decay                       | 0.7/1000 epochs              |
| Number of epochs                 | 2000 (early stopping)        |
| Activation Function              | Linear                       |
| Number of feed-forward dense layers | 1                           |
| Control-Measures-CNN kernel sizes | (3,8)-(5,8)-(7,8)            |
| Mobility-CNN kernel sizes        | (3,6)-(5,6)-(7,6)            |
| Number of CNN Kernels            | 8                            |
| Number of LSTM Hidden Layers     | 32                           |
| Dropout                          | 0.2                          |
| Input Window Size                | 10 days                      |
| Output Window Size               | 1 day                        |

### Table 12 Test set overall evaluation.

| Mobility Change Category         | Autoregressive | CNN-LSTM | CNN-Autoregressive | Proposed Model |
|----------------------------------|----------------|----------|--------------------|----------------|
| Retail and Recreation            | 10.9988        | 9.4554   | 7.101              | 4.6167         |
| Groceries and Pharmacies         | 8.2783         | 8.4564   | 7.4054             | 6.5008         |
| Parks                            | 12.7616        | 14.3092  | 12.6567            | 9.7092         |
| Transportation                   | 13.7464        | 10.8689  | 9.2425             | 4.2555         |
| Workplaces                       | 6.3286         | 8.2676   | 6.4171             | 4.9156         |
| Residential                      | 4.0438         | 3.9981   | 2.5506             | 1.807          |
| Average Absolute Error           | 9.3596         | 9.226    | 7.5622             | 5.3008         |
| Mean Square Error                | 216.3692       | 197.2609 | 155.0921           | 94.3249        |

### Table 13 Statistical Analysis Values.

| Mobility Change Category         | Mean(Actual) | Mean(Pred) | P-Value | R-Squared |
|----------------------------------|--------------|------------|---------|-----------|
| Retail and Recreation            | −32.0745     | −31.4136   | 0.3212  | 0.9243    |
| Groceries and Pharmacies         | −14.6092     | −13.9929   | 0.2245  | 0.7168    |
| Parks                            | −5.9467      | −6.1648    | 0.8759  | 0.8945    |
| Transportation                   | −39.1667     | −38.5931   | 0.369   | 0.9409    |
| Workplaces                       | −26.3794     | −25.8806   | 0.3635  | 0.8714    |
| Residential                      | 11.3111      | 11.3211    | 0.9661  | 0.8966    |
month from the starting date of reporting the mobility changes, and that could be the reason for its poor prediction results.

Moreover, Table 15 reports two Middle Eastern countries’ prediction results: Egypt and Saudi Arabia. The model produces a higher loss for Egypt than the KSA, where the same explanation in the case of Slovenia applies here. The large population size, in addition to the unique nature of public places, workplaces, and transportation in Egypt, has a dominant effect on the community mobility and introduces a difficulty in forc-

| Table 14 | Evaluation of best and worst prediction results. |
|----------------|---------------------------------------------|
| Mobility Change Category | Cameroon(best) | Slovenia(worst) |
| Retail and Recreation | 2.226252 | 6.109125 |
| Groceries and Pharmacies | 2.591267 | 12.88361 |
| Parks | 3.385782 | 19.96535 |
| Transportation | 2.307079 | 6.53274 |
| Workplaces | 2.977999 | 5.190136 |
| Residential | 1.034193 | 2.144939 |
| Average Absolute Error | 2.420429 | 8.804316 |
| Mean Square Error | 13.1848 | 229.5093 |

| Table 15 | Evaluation of Egypt and Saudi Arabia. |
|----------------|---------------------------------------------|
| Mobility Change Category | Egypt | Saudi Arabia |
| Retail and Recreation | 3.88095 | 4.075632 |
| Groceries and Pharmacies | 5.479142 | 5.0427 |
| Parks | 5.774594 | 4.856126 |
| Transportation | 5.1092 | 2.625397 |
| Workplaces | 5.775543 | 4.087503 |
| Residential | 1.62474 | 1.616538 |
| Average Absolute Error | 4.607363 | 3.717316 |
| Mean Square Error | 66.54512 | 36.631 |

Fig. 5 Predicted VS Actual Plots of the six mobility categories.
ing some of the control measures, which might be the reason for the more considerable loss than the KSA. However, the loss is still relatively small for both Egypt and Saudi Arabia.

Additionally, Figs. 6 and 7 illustrate the actual and the predicted community mobility changes for the six mobility categories in the two countries (Egypt and KSA). As shown in the figures, the predictions follow the same pattern of the actual changes, illustrating the small loss, and proving the effectiveness of the proposed model in predicting community mobility changes, given the past mobility information and applied government measures.

6. Conclusion and future work

This paper presents artificial intelligence-based prediction model for the mobility patterns changes during COVID-19 pandemic. Autoregressive prediction model based on mobility data only has been developed as baseline. Then, three enhanced models using Convolutional Neural Network and Long Short Term Memory (CNN-LSTM) have been proposed using the mobility change data and applied control measures data. The parallel CNN-LSTM with autoregressive model and extended features has achieved the best results among the four models. Results show that the proposed model can predict the effect of governments taken control measures on community mobility pattern changes in the six categories found in the used data source with minimum loss, based on mean square error and mean absolute error evaluation metrics.

The work presented in this paper can be extended in various directions. COVID-19 data can be included as a third data source to enhance the mobility prediction. Additional features can also be included to enhance the accuracy including information about national holidays and weekend days. Furthermore, the model can be extended to predict future infections data of COVID-19 based on the applied control measures and community mobility.

Fig. 6 Actual and predicted community mobility changes for the six mobility categories in Egypt.
Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

This work is supported by the Deanship of Scientific Research, Islamic University of Madinah, Program of Funding Research and Studies on Corona Virus/COVID 19, Project number 41/13.

References

[1] Google, Covid-19 community mobility reports. https://www.google.com/covid19/mobility/.

[2] M. van der Schaar, A. Alaa, How artificial intelligence and machine learning can help healthcare systems respond to covid-19 (2020). https://www.vanderschaar-lab.com/NewWebsite/covid-19/post1/paper.pdf.

[3] R.M. Carrillo-Larco, M. Castillo-Cara, Using country-level variables to classify countries according to the number of confirmed covid-19 cases: An unsupervised machine learning approach, Wellcome Open Res. 5 (56) (2020) 56.

[4] Z. Hu, Q. Ge, L. Jin, M. Xiong, Artificial intelligence forecasting of covid-19 in china, arXiv preprint arXiv: 2002.07112.

[5] P. Hartono, Mixing autoencoder with classifier: conceptual data visualization, IEEE Access 8 (2020) 105301–105310.

[6] S.Y. Işikhan, D. Gülec, The clustering of world countries regarding causes of death and health risk factors, Iran. J. Public Health 47 (10) (2018) 1520.

[7] D.B. Alonso, A. Androniceanu, I. Georgescu, Sensitivity and vulnerability of european countries in time of crisis based on a new approach to data clustering and curvilinear analysis, Administratie si Management Public 2016 (27) (2016) 46.

Fig. 7 Actual and predicted community mobility changes for the six mobility categories in Saudi Arabia.
[8] ACAPS. Covid-19 government measures dataset. https://www.acaps.org/covid19-government-measuresdataset.

[9] D.K. Barrow, S.F. Crone, A comparison of adaboost algorithms for time series forecast combination, Int. J. Forecast. 32 (4) (2016) 1103–1119.

[10] D. Gupta, M. Pratama, Z. Ma, J. Li, M. Prasad, Financial time series forecasting using twin support vector regression, PloS One 14 (3) (2019) e0211402.

[11] O. Ebuahuma, M. Gebreslasie, L. Magubane, A seasonal autoregressive integrated moving average (sarima) forecasting model to predict monthly malaria cases in Kwazulu-natal, South Africa, South African Medical Journal 108 (7) (2018).

[12] A. Dela Cruz, Forecasting philippine residential mobility percentage change during covid-19, Available at SSRN: https://ssrn.com/abstract = 3579517.

[13] F. Nyabadza, F. Chirove, W.C. Chukwu, M.V. Visaya, Modelling the potential impact of social distancing on the covid-19 epidemic in south africa, medRxiv.

[14] A. Perkins, G. Espana, Optimal control of the covid-19 pandemic with non-pharmaceutical interventions, medRxiv.

[15] K. Prem, Y. Liu, T.W. Russell, A.J. Kucharski, R.M. Eggo, N. Davies, S. Flasche, S. Clifford, C.A. Pearson, J.D. Munday, et al, The effect of control strategies to reduce social mixing on outcomes of the covid-19 epidemic in wuhan, china: a modelling study, Lancet Public Health 5 (5) (2020) e261–e270.

[16] L. Roques, E.K. Klein, J. Papaix, A. Sar, S. Soubeyrand, Effect of a one-month lockdown on the epidemic dynamics of covid-19 in france, medRxiv.

[17] K. Linka, M. Peirlinck, F. Sahli Costabal, E. Kuhl, Outbreak dynamics of covid-19 in europe and the effect of travel restrictions, Comput. Methods Biomech. Biomed. Eng. (2020) 1–8.

[18] N. Banholzer, E. van Weenen, B. Kratzwald, A. Seeliger, D. Tschernutter, P. Bottrighi, A. Cenedese, J. Puig Salles, W. Vach, S. Feuerriegel, Impact of non-pharmaceutical interventions on documented cases of covid-19 [preprint]. medrxiv. 2020 apr 28, DOI 10 (2020.04) (Apr. 2020) 16-20062141.

[19] N.G. Davies, A.J. Kucharski, R.M. Eggo, A. Gimma, W.J. Edmunds, T. Jombart, K. O’Reilly, A. Endo, J. Hellewell, E.S. Nightingale, et al, Effects of non-pharmaceutical interventions on covid-19 cases, deaths, and demand for hospital services in the UK: a modelling study, Lancet Public Health 5 (7) (2020) e375–e385.

[20] D.W. Zimmerman, B.D. Zumbo, Rank transformations and the power of the student t test and welch t test for non-normal populations with unequal variances, Can. J. Exp. Psychol./Revue canadienne de psychologie expérimentale 47 (3) (1993) 523.

[21] A.C. Cameron, F.A. Windmeijer, An r-squared measure of goodness of fit for some common nonlinear regression models, J. Economet. 77 (2) (1997) 329–342.