

## **PREDICT THE RISK LEVEL IN IRAQI GOVERNORATES ACCORDING TO THE SPREAD OF COVID-19 USING DATA MINING**

**(An article review)**

Ibtisam Abbas Othman ( [ibtisamabbas81@gmail.com](mailto:ibtisamabbas81@gmail.com))  
Ban Sharief Mustafa ( [Banmustafa66@uomosul.edu.com](mailto:Banmustafa66@uomosul.edu.com))

### **SUMMARY**

Recently, and due to its efficient and successful performance, Deep Learning (DL) algorithms have become widely used in many computing applications. In December 2019, the world witnessed a widespread spread of a new strain of the SARS virus, which was named COVID-19. Its rapid transmission from one person to another led to the spread of the virus on a large scale [1]. On 30 January, 2020, the World Health Organization declared that Covid-19 represents a health emergency of concern to the international community. Therefore, on basis of this rapid spread of the virus in many countries of the world, the Director-General of the World Health Organization announced that Coronavirus has become an epidemic that threatens the lives of millions of people in the world [2]. Non-clinical approaches such as data mining and machine and deep learning techniques can help reduce the spread of the virus. Data mining is used to find similar patterns and extract knowledge from large amounts of data. Machine and deep learning algorithms are data processing techniques that help in building analytical models [3]. In this research, the aim is to predict the number of infected persons or deaths in each Iraqi governorate and for a subsequent period determined by a range of days. According to the information available to us, Covid-19 dataset for Iraq's governorates has not been use till now in a research paper. The data available on the World Health Organization's HDX website has been used, processed and restructured to be suitable as inputs for training models. A deep learning model based on the LSTM network has been built to predict the path of disease spread, such as the number of new cases or the number of deaths in the coming days according to the Iraqi- governorates, and it has achieved very good performance rates. Predicting the current situation in Iraq is critical to containing the threat because it helps to take appropriate measures and decisions at the governorates level as well as at the personal level where travel to places of deployment

### **Deep Learning**

Machine and deep learning algorithms, can be considered as a branch of artificial intelligence. It is a field that is based on learning and updating depending on the analysis of mathematical and statistical algorithms. Deep learning algorithms differ from machine learning algorithms in that they require greater computing power and complexity than the latter. Despite this, progress in the field of big data has led to the emergence of larger and deeper networks, enabling algorithms to learn, monitor, and deal with complex data and cases more efficiently and faster than humans. In general, deep learning has been effective in applications of image classification, speech recognition, bioinformatics, etc. [8]. Deep learning is one of the most important tools for data scientists who specialize in collecting, analyzing and interpreting large amounts of data. This is due to its enormous efficiency and speed in the process of processing large data in a way that simulates and even surpasses the human mind. [9]. DL allows

computational models consisting of multiple processing layers to learn data representations with multiple levels of abstraction. These methods have greatly improved the latest technologies in speech recognition, visual object recognition, object detection and many other fields such as drug discovery and genetics. DL discovers the complex structure in large data sets using a reverse propagation algorithm to indicate how the machine works [10]. DL methods do not rely on human intervention; it consists of many layers of algorithms that provide a different interpretation of the data they feed on [9].

### **Long Short-Term Memory**

Currently, recurrent neural networks for long-term memory (LSTM) are one of the most important types of deep learning networks. It has been used in challenging areas such as language translation, image annotation and text production. LSTMs differ significantly from other intricate learning techniques, such as multi-layer cognition (MLPS) and convolutional neural networks (CNN), in that it is specifically designed for sequence prediction problems [11]. LSTM is a neural network. In each LSTM module, there are three gates: the forget gate, the input gate and the output gate. An input gate is interested to direct the memory of input data and prevent worthless information from incoming the storage unit. The forgetting gate is used to selectively discard the last coin information. The output gate is used to control the output of information each time [12].

## **METHODS**

### **1. Data Set**

The data set of (Covid-19) virus cases for the State of Iraq has been used for training the model.

### **2. Preprocessing step**

It has been assumed that the infection rate in any region follows the pattern of infections spread in the last days in that region.

### **3 . Constructing the prediction model**

The study methodology mainly consists of five main steps: data preparation, data pre-processing, model

constructing, model evaluation, and visualization.

### **References :**

[1] Leppard, Gary G. "Evaluation of electron microscope techniques for the description of aquatic colloids." In *Environmental particles*, pp. 231-289. CRC Press, 2019.

[2] Ahouz, Fatemeh, and Amin Golabpour. "Predicting the incidence of COVID-19 using data mining." *BMC public health* 21.1 (2021): 1-12.

[3] Widyanto, R. Arri, et al. "Data Mining Predicts the Need for Immunization Vaccines Using the Naive Bayes Method." *Journal of Applied Data Sciences* 2.3 (2021): 93-101.

[4] Ramchandani, Ankit, Chao Fan, and Ali Mostafavi. "Deepcovidnet: An interpretable deep learning model for predictive surveillance of covid-19 using heterogeneous features and their interactions." *IEEE Access* 8 (2020): 159915-159930.

[5] Roy, Santanu, Gouri Sankar Bhunia, and Pravat Kumar Shit. "Spatial prediction of COVID-19 epidemic using ARIMA techniques in India." *Modeling earth systems and environment* 7.2 (2021): 1385-1391 .

[6] Ramchandani, Ankit, Chao Fan, and Ali Mostafavi. "Deepcovidnet: An interpretable deep learning model for predictive surveillance of covid-19 using heterogeneous features and their interactions." *IEEE Access* 8 (2020): 159915-159930.

- [7] Ramadan, Rasha H., and Mona S. Ramadan. "Prediction of highly vulnerable areas to COVID-19 outbreaks using spatial model: A case study of Cairo Governorate, Egypt." *The Egyptian Journal of Remote Sensing and Space Science* (2021).
- [8] Alakus, Talha Burak, and Ibrahim Turkoglu. "Comparison of deep learning approaches to predict COVID-19 infection." *Chaos, Solitons & Fractals* 140 (2020): 110120
- [9] Alansari, Husain, Oksana Gerwe, and Anjum Razzaque. "Role of Artificial Intelligence During the Covid-19 Era." *The Big Data-Driven Digital Economy: Artificial and Computational Intelligence*. Springer, Cham, 2021. 157-173 .
- [10] LeCun, Yann, Yoshua Bengio, and Geoffrey Hinton. "Deep learning." *nature* 521.7553 (2015): 106-444 .
- [11] Brownlee, Jason. *Long short-term memory networks with python: develop sequence prediction models with deep learning*. Machine Learning Mastery, 2017.
- [12] Ghany, Kareem Kamal A., Hossam M. Zawbaa, and Heba M. Sabri. "COVID-19 prediction using LSTM algorithm: GCC case study." *Informatics in Medicine Unlocked* 23 (2021): 100566.
- [13] "Iraq: Coronavirus (COVID-19) Subnational Cases" variety, [data.humdata , variety, .org/dataset/iraq-coronavirus-covid-19-subnational-cases](https://data.humdata.org/dataset/iraq-coronavirus-covid-19-subnational-cases).