# Predictions and visualization for confirmed, recovered and deaths COVID-19 cases in Iraq

**Wisam Dawood Abdullah[1], Abdulrahman Ahmed Jasim[2], Layth Rafea Hazim[1]**
[1]Computer Science Department, Computer Science and Mathematics College, Tikrit University, Tikrit, Iraq
[2]Department of Network Engineering, Collage of Engineering, Al-Iraqia University, Baghdad, Iraq

| Article Info | ABSTRACT |
|---|---|

The 2019-2020 coronavirus pandemic is an emerging infectious disease that has been referred to as the "COVID-19", which results from the coronavirus "SARS-CoV-2" that started in Wuhan, China, in Dec. 2019 and then spread worldwide. In this paper, an attempt for compiling and analyzing the information of the epidemiological outbreaks on "COVID-19" based upon datasets on "2019-nCoV" has been presented. An empirical data analysis with the visualizations was conducted for understanding the numbers of the variety of the cases that have been reported (i.e confirmed, deaths, and recoveries) in and outside of Iraq and carried out a dynamic map visualization of the " COVID-19" expansion in a global manner through the date wise and in Iraq. We an investigation has been carried out as well, which characterized the pandemic effects Iraq and the entire world, with the use of machine learning. A k-nearest neighbor (kNN) model and a linear regression (LR) model have been proposed. This paper included the precise analysis of the confirmed cases, as well as the recovered cases, deaths, predicting the pandemic viral attacks and how far it is expanding in Iraq and the world, the LR model got the highest results, reaching 100 percent.

*Corresponding Author:*

Wisam Dawood Abdullah
Computer Science Department, Computer Science and Mathematics College, Tikrit University
Street of Tikrit-Mosil, Al-Qadissiyah quarter, 009642, Tikrīt, Sallahaldin, Iraq
Email: wisamdawood@tu.edu.iq

## 1. INTRODUCTION

The coronavirus disease 2019 (COVID-19) can be defined as one of the emerging issues for the public health that affects the countries worldwide, and that includes Iraq, and results from the severe acute respiratory syndrome Coronavirus 2 (SARS-CoV2) [1], [2]. This infection spreads from one person to another, usually in the case of being nearby physically, but over long distances as well, typically indoors [3]. The most widespread symptoms can be fever, loss of smell, fatigue, muscle pains, loss of appetite, shortness of breath, dry cough, and coughing up sputum [4], [5].

COVID-19 has been found in Iraq, and the first COVID-19 infection has been confirmed in a student from Iran in Al-Najaf city on Feb. 24th of 2020. Also, a family that had returned recently from Iran was tested positive for coronavirus in Kirkuk city near Baghdad on Feb. 25th of 2020. The first case of death that had had been confirmed happened in a 70-year-old clergyman on Mar. 4th in Al-Sulaymaniyah.

This disease had spread from Mar. 27th on, in all of the 19 Governorates of Iraq. As infections were increasing steadily since the middle of Mar. 2020, the government of Iraq had taken some measures, which included a national partial or complete lockdown, travel restrictions and curfew. The universities, cinemas, and schools in Baghdad have been closed; in addition to other large religious and public gatherings have been banned in cities. The Ministry of Health of Iraq declared a total nationwide lockdown that started on Jul. 30th,

except for Kurdistan region. The high numbers of the reported cases had led towards studying the spreading of COVID-19 in Iraq.

Various researches were published for the evaluation of COVID-19 epidemiological characteristics and for mitigating its impact on the public health [6], [7]. Tosepu *et al*. [8] have researched the association between the weather and COVID-19 pandemic in Jakarta, Indonesia. Batista [9] used logistic growth model for the evaluation of final size and peak of COVID-19 epidemic in South Korea, China, and other countries. Boldog *et al*. [10], a computational method has been established for the assessment of the novel COVID-19 outbreak risks outside China. According to confirmed data, Almeshal *et al*. [11] have suggested a prediction of COVID-19 epidemic size in Kuwait; stochastic and deterministic methods of modeling have been utilized. A research that has been utilized for forecasting the peak of COVID-19 epidemic in Japan with the use of conventional model of the SEIR, which has been was carried out by Kuniya [12]. Röst *et al*. [13] have suggested a statistical and epidemiological study of early phase of the outbreak of COVID-19 in Hungary and advanced an age-structured compartmental model for the investigation of the alternative post-lockdown scenarios. Bantan *et al*. [14], have advanced an exponentiated M family of the continual distributions for providing new statistical models.

In this study, we will visualize and analyse information of the epidemic outbreak about "COVID-19" based on data-sets that are available on "2019-nCoV". An exploratory analysis of the data will be conducted with visualizations for understanding the amount of the various reported cases (i.e confirmed, recovery and death) in and outside Iraq and perform a dynamic map visualization of the expansion of " COVID-19" globally by date in Iraq. An investigation will be conducted describing the effects of the epidemic on the world and Iraq utilizing the machine learning approach. Anticipating is one of the typical data science exercises assisting the management in planning jobs, setting goals, and detecting anomalies. We proposed a linear regression (LR) model and a closest k-nearest neighbor (kNN) model. The main benefits of the present work include the precise analyses of the confirmed cases in the world as well as Iraq, recovered cases, deaths, and a prediction of the pandemic viral attack and its extent of expansion globally and in Iraq.

## 2. METHODS

The method includes three steps; the first step is data collection. The data-set has been obtained from "Kaggle", the second one is to generate a dynamic map of expansion of "COVID-19" from the date of 22-01-2020 till 29-05-2021 globally, third step and finally aims at exploring "COVID-19" through data analysis, visualize and prediction to build a strong model which has the ability for predicting the way the virus could be spreading through several regions and countries could have the ability of helping the efforts of mitigation, Iraq is a model. Our study period was a year almost and the spreading of the virus locally and internationally has been visualized and analyzed throughout this study period with the confirmed cases, deaths and recovered cases. Finally, the expansion of the virus globally has been predicted using methods of data mining, such as k-nearest neighbors and linear regression by using python anaconda notebook.

### 2.1. Dataset introduction

We used the open dataset of "2019 Novel Corona Virus 2019 Dataset" that has been provided by "Johns Hopkins Univ.", they have made an excellent dashboard with the use of the data on the affected cases to date [15]. Besides that, they have provided as well, the chance for the researcher and the data analyst through the provision of data that has been available in the format of Google sheets. Which is why, dataset has daily level data on the number of influenced cases, recovery, and deaths from 22-01-2020 to 29-05-2021. It has a total of 306,430 tuples and 8 features as shown in Table 1.

Table 1. Description of dataset

| Features name | Description |
|---|---|
| Sno | Serial number. |
| Observation Date | In MM/DD/YYYY. |
| Region/Country | Observation Country. |
| State/ Province | State or Province of observation (May be left empty in the case where it is missing). |
| Last Update | Time in the UTC where row is updated for a certain country or province. (Not standardized, therefore, it should be clean prior to utilizing it). |
| Confirmed | Cumulative number of the confirmed cases to that date. |
| Recovered | Cumulative number of the recovered cases up to date. |
| Deaths | Cumulative number of the death cases up to date. |

## 2.2. Exploratory data analysis and visualize

We are going to visualize this data and make some statistics to find the relevant information. The data-sets has been analysed by the use of a variety of the exploratory approaches of data analysis and these data have been visualized for the purpose of providing an efficient consciousness in regards to COVID-19 outbreak worldwide. Visualization helps in understanding the spread of corona virus from one country to other and for how much [16]. In a response to that ongoing emergency for the public health, an on-line interactive dashboard has been developed, it has been hosted by Johns Hopkins University and CSSE (i.e. Centre for Systems Science and Engineering), for the visualization and tracking of COVID-19 reported cases in the real time [17]. A worldwide map has been reported cases and knowledge has been provided of the way through which "SARS"-"CoV2" spread between Jan. 22nd 2020 and May 29nd 2021 worldwide, as it has been listed in Table 2. Every one of the map segments represents an area, through the use of the visual data analytics it is helpful for individuals in understanding the COVID-19's epidemiological nature. From the representation of the map, it has been shown that the US had reported the maximum number of the confirmed cases with up to 33,251,939 (19.6%), India 27,894,800 (16.4%), Brazil 16,471,600 (9.69%) and the rest of the countries are like Russia, UK and France, as shown in Figure 1. We took Iraq as a model for our study to track the extent of the corona virus and the number of deaths as well, where the numbers of confirmed 1,193,608 spreads from the first case in 23 February 2020 to 29 May 2021. We also examine the time series data with the use of the visual exploratory data analyses for the purpose of providing an understandable and strong result of this extreme COVID-19 outbreak. It's obvious that the analysis of those data in real time is highly useful to capture an epidemic behaviour of that severe pandemic [18].

Table 2. COVID-19 in the world from 22 Jan 2020 till 29 May 2021

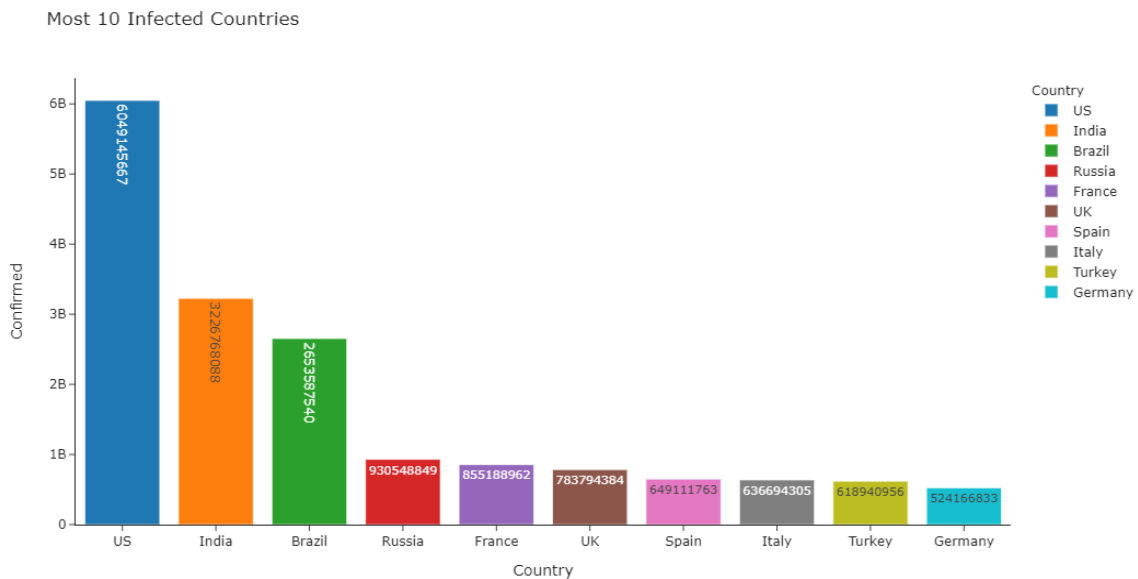| Last update | Confirmed | Active | Recoverd | Deaths |
|---|---|---|---|---|
| 2021-05-29T00:00:00 | 169951560 | 592777272 | 107140669 | 3533619 |



Figure 1. Most 10 infected countries

## 2.3. Building prediction model

We have to build a strong model, predicting the way through which this virus would have the ability of spreading over numerous regions and countries could have the ability of helping the efforts of mitigation. The aim of the present research is building a model predicting the virus progression during 2021 year. By using machine learning LR and kNN model in Python.

### 2.3.1. Linear regression (LR) model

LR can be defined as a regression modelling type, which is the most usable statistical approach for the predictive analyses in the machine learning. In modelling of regression, a target class will be predicated on independent features. Which is why, this approach may be utilized for finding out the relations between

the dependent and the independent variables as well as for the forecasting [19]. LR determines a linear relation between those independent and dependent variables. After that, the model of linear regression is utilized for the evaluation of relative effect of the active cases as a result of the daily-confirmed cases in worldwide, and particularly in Iraq. There have been 2 factors (x,y) involved in the analyses of the linear regression [20]. In (1), (2) exhibits the way by which y is related to x known as regression.

$$y = \beta_0 + \beta_1 x + \varepsilon \tag{1}$$

Or equivalently.

$$y = \beta_0 + \beta_1 x \tag{2}$$

Where, $\varepsilon$ represents the linear regression's error term. This term is used for accounting variability between x and y, $\beta_0$ denotes y- intercept, $\beta_1$ denotes the slope.

### 2.3.2. K-nearest neighbors model

kNN can be defined as supervised learning approach that has been mainly utilized in the detection systems, also was extensively used in COVID-19 problems [21]. In such approach, new instance query has been classified on the basis of popular kNN distance measures, such as Minkowski distance, Euclidean distance, and Manhattan distance; we used the last one in this work. The kNN have been considered as one of the and simplest and most efficient approaches for the classification of item. In the examples, kNN training are represented as points in feature space in numerous separate classes. For the purpose of predicting the new item Lx label, at first, it's projected in the feature space issue. After that, distances between the K-nearest examples and Lx are computed [22], [23]. After that, Lx has been classified with the use of its neighbours' majority vote. Despite the fact that it is simple and highly efficient, the conventional kNN may be trapped easily. kNN trapping could happen in 2 separate cases. The first in the case where there isn't any confidence in the decision of the classifier. Which could happen in the case where the chances that the item that has been tested is targeted numerous classes are nearly identical or quite close to one another. The second kNN trapping case happens when the classifier's decision is targeting the tested item to 2 classes or more as in (3), after its formulation that has been provided by [21].

$$D\left(x_i, x_j\right) = \sqrt{\sum_{k=1}^{n}(x_{ik} - x_{jk})^2} \; k = 1,2, \ldots \ldots, n \tag{3}$$

In this study, we will use four well-known measure to evaluate the methodology, which are r2_score (determination coefficient) function of regression score. The optimal score equals 1 and it may be negative (due to the fact that the model may be randomly worse). A constant model which always has the ability of predicting the expected y value, in spite of the features of the input, would result in an R^2 score of 0.0 [24].

### 3. RESULT AND DISCUSSION

The present work analysed 3 separate data classes. Which include the confirmed, recovered and death cases in Iraq and the world between Jan. 22nd 2020 and May. 29nd 2021. Which provides as well a comparative analysis of all cases that have been reported in Iraq and the other countries.

### 3.1. Visualization result

The library Plotly Python has been utilized for the to development of a Choropleth Dynamic Map for the visualization of the way this pandemic has been spreading worldwide, country wise as well as continent wise. This dynamic map will show how it varies day by day with the help of legend. Based on the dataset, Figure 2 depicts the comparative analysis data of the confirmed cases worldwide, until 29 May 2021. The pie chart depicts those twelve countries have been named as the most infected countries around the world, proportion of confirmed cases by country. This pie plot is also useful to visualize the proportion of active cases, recovered cases and deaths by country. Figure 3 shows the percentage of recovered cases by country. Proportion of recovered cases by country, the tree map is similar to a pie plot; it also helps visualize the proportion of confirmed, active, recovered cases and deaths.

This representation demonstrates in Figure 2 that US endured the maximum amount of the infected patients, which was 33,251,939 and then comes India with 27,894,800 confirmed cases. After that, the country of the risk zone has been Brazil with 9.69% of the confirmed cases worldwide. Figure 3 illustrates the cases of recovery in the world according to the data-set; it has been seen that the cases of recovery in

India is maximal, where there has been 25,454,320 total recovery cases worldwide. This rate of recovery in India is rather satisfactory under the existing condition in the world. Not that the data that has been illustrated in Figure 3 gives no complete recovery rate scenario in various countries. This results from the indifferent countries, the number of new cases varies and several of those are under recovery.

By visualizing, the dataset on Figure 4 shows the general structure of "COVID-19" expansion explains the current scenario of "COVID-19" exposure till 29 May 2021. United States extinguishes all other nations with a rapid boom of 33,251,939 confirmed cases, also for almost a whole year, confirmed cases escalated in many countries come Russia, Brazil, Iran, Germany, India, Saudi Arabia, Egypt and Iraq, as an indicator in dark red. Light red colour the very less virus prone countries, such as Australia, Mali, Congo and Mauritania.

Figure 5 shows the Iraq country affected from the first recorded case the end of February 2020. By analysing the Figure 5, only single confirmed case has affected till 24 February 2020. While focusing on the status of the Iraq till 20 June 2020 in Figure 6. We can see that, it has been covered around 10 percent of the Iraq within the time period of 23 February 2020 to 20 June 2020 is occupying the peak position with 29,222 confirmed cases with the numbers of recovered and deaths 132,121,013 respectively.

The change of confirmed cases within five months is very big in Iraq and we can see the changes till 20 November 2020. By visualizing the data on Figure 7 we can see that it has spread globally within 10-11 months in the list with confirmed, recovered and deaths 531,769, 460,394 and 11,388 respectively. From these figures, we can see that how fast it is spreading all provinces in Iraq. Figure 7 provides a clear picture of the pandemic virus attack affected all provinces in Iraq till 20 November 2020. Finally, scenario of Covid-19 exposure until 29 May 2021, Iraq extinguishes with a rapid boom of 1,193,608 confirmed, 1,107,101 recovered and 16,334 death cases, as in Figure 8.

Figure 8 represents the confirmed, recovered, and death cases' graphical representation throughout the period of the study. Almost the 3 cases were going in a straight line from 23 February 2020 to 5 June 2020 and after that it suddenly increased exponentially. The proven cases reached large numbers within a year. From this, we can assume the expansion of this virus. The major number of deaths reported worldwide are people of age 55 above due to the lack of immunity.
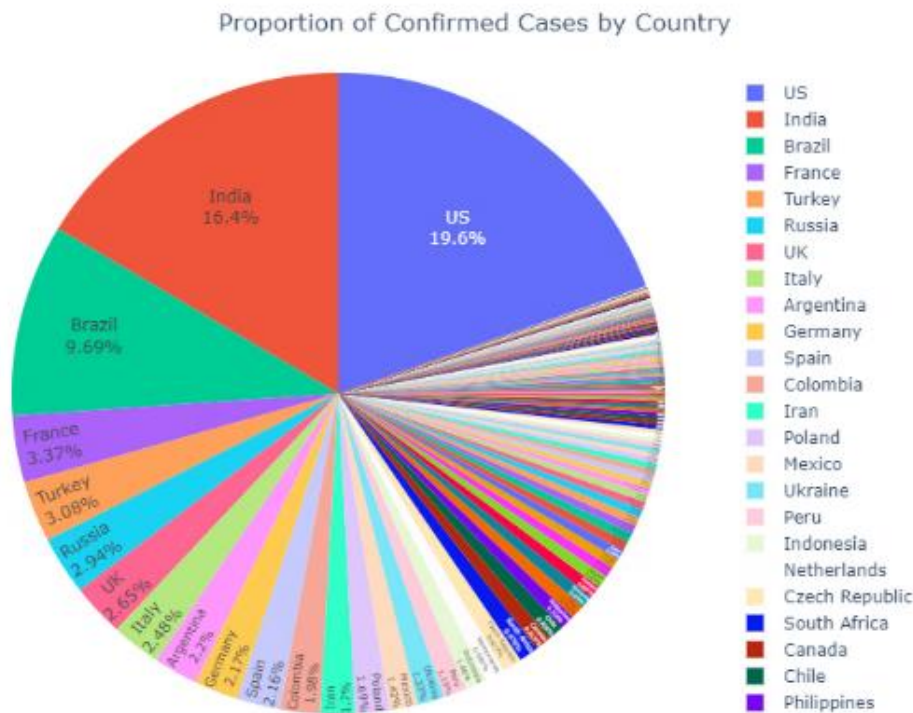


Figure 2. COVID-19 confirmed cases around the world until 29 May 2021
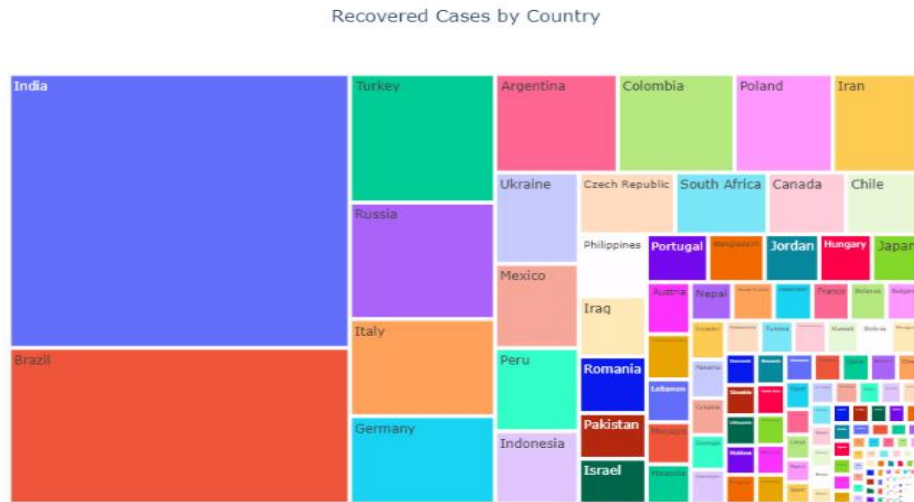
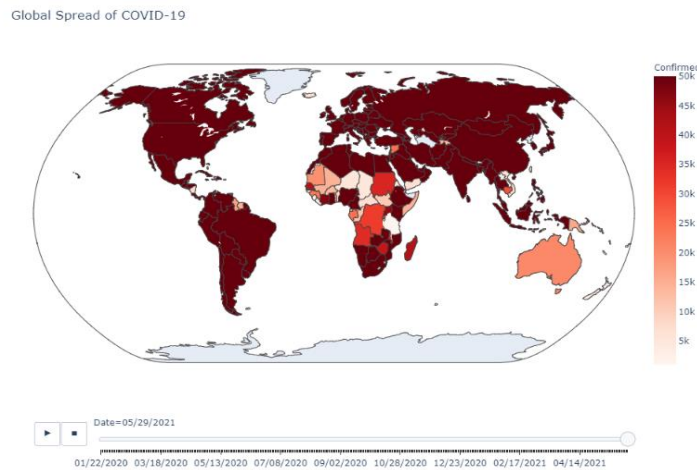Figure 3. COVID-19 recovered cases worldwide until 29 May 2021



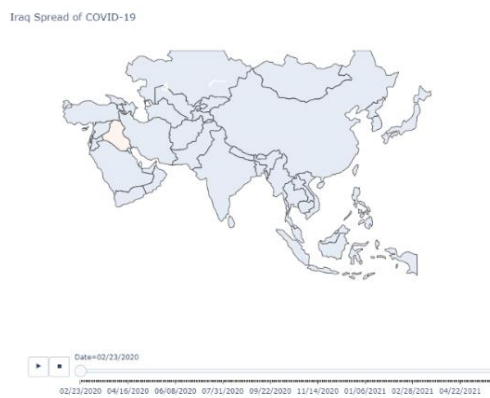Figure 4. Global spread of COVID-19 until 29 May 2021



Figure 5. Iraq spread of COVID-19 until 23 February 2020
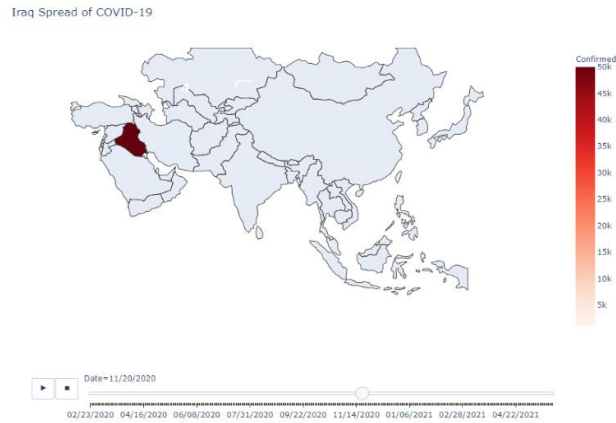


Figure 6. Iraq spread of COVID-19 until 20 June 2020

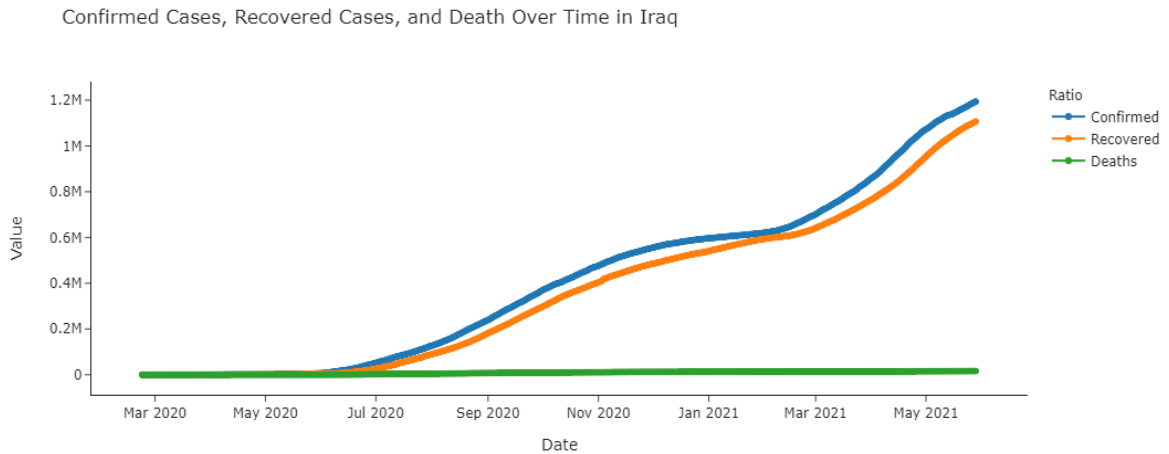Figure 7. Iraq spread of COVID-19 until 20 November 2020



Figure 8. Iraq spread of the confirmed, recovered and death cases of COVID-19

## 3.2. Prediction model result

The present research attempted at the development of a system for future predictions of numbers of the COVID-19 cases with the use of the approaches of machine learning. The data-set that has been utilized in this research includes the information on daily reports of numbers of the newly confirmed cases, the numbers of recovered, and the number of the death cases from "COVID-19". The present research attempts at forecasting the methods of machine learning which may be affected according to the new infected and death cases, which include the numbers of the expected recovery cases for the near future. There have been 2 machine learning models kNN and LR utilized for the prediction of the numbers of the newly infected cases, the numbers of the death cases, as well as the numbers of the recovered.

Table 3 lists the prediction results of LR and kNN models that have been utilized in the present research, it was applied to the whole world dataset first, and then it was applied to Iraq dataset only. LR lead the table according to the performance, kNN also performed very good, in terms of all metrics of evaluation. Therefore, when building a forecasting system, we recommend using linear regression (LR) technique, as it has achieved very high results in terms of performance [25].

Table 3. Models' performance on the future prediction two model

| Country | Iraq | World |
|---------|------|-------|
| Models | R2_Score | R2_Score |
| LR | 100.0% | 100.0% |
| kNN | 99.88% | 99.97% |

## 4. CONCLUSION

In conclusion, this research a model was introduced for visualize and analyse the information of the epidemiological outbreaks on "COVID-19" according to the data-sets on "2019-nCoV", Model building process has three steps, first one is the data collection, second step was to generate a dynamic map of expansion of " COVID-19" from the date of 22-01-2020 till 29-05-2021 globally, third step has a prediction of COVID-19pandemic infection with the use of the algorithms of machine learning reveals the comparative discussion on the confirmed cases, recovered cases, and the deaths status across various countries including Iraq as a model, by using LR and kNN machine learning techniques. Such visualize and analyse the activities may be helpful for the generation and dissemination of the detailed knowledge about to scientific community, particularly in advanced outbreak stages, in the case where there's a data-set to an available year, which allows for the independent evaluation of the key parameters influencing the interventions. We concluded, in the case of confirmed COVID-19 status, the United States, India, Russia, Brazil, UK, France, Italy, Turkey, Spain and Germany showing the highest number of cases with more than 50% of the world's countries and it is in the peak position while checking the mortality status too. African countries are in the least affected countries of confirmed cases. Early indications that the response is being strengthened in Iraq have been investigated as well based on a smaller number of the cases in the beginning and for 4 months' case of the time of detection and rapid management that has been identified by Iraqi government, the number of confirmed cases has been then 29,222 cases at 20 June 2020. The active measurements planned and executed against the epidemic attack by Iraqi government were appreciable and should be followed by other nations. Unfortunately, the epidemic then spread and increased dramatically in Iraq after June 2020, this great climb, which reached 1,193,608 cases at 29 May 2021, is due to the lack of commitment of people to safety conditions, including masks, proper social distancing and personal hygiene. The effective management using symptomatic therapy and quarantine system can control the spread of the disease state up to a limit. People must adhere to these procedures for their safety.

## REFERENCES

[1] Z. Wu and J. M. McGoogan, "Characteristics of and important lessons from the Coronavirus Disease 2019 (COVID-19) outbreak in China: Summary of a report of 72314 cases from the Chinese center for disease control and prevention," *JAMA - Journal of the American Medical Association*, vol. 323, no. 13, pp. 1239–1242, 2020, doi: 10.1001/jama.2020.2648.

[2] A. E. Gorbalenya *et al.*, "Severe acute respiratory syndrome-related coronavirus: The species and its viruses – a statement of the Coronavirus Study Group," *bioRxiv*, 2020, doi: 10.1101/2020.02.07.937862.

[3] "How Coronavirus Spreads | CDC," *https://www.cdc.gov*, 2021. https://www.cdc.gov/coronavirus/2019-ncov/prevent-getting-sick/how-covid-spreads.html (accessed Feb. 15, 2021).

[4] "Coronavirus." https://www.who.int/health-topics/coronavirus#tab=tab_1 (accessed Feb. 15, 2021).

[5] F. A. Binti Hamzah *et al.*, "CoronaTracker: World-wide Covid-19 outbreak data analysis and prediction," *Bulletin of the World Health Organization*, 2020, doi: 10.2471/BLT.20.255695.

[6] M. K. Yusof, W. M. A. F. W. Hamzah, and N. S. M. Rusli, "Efficiency of hybrid algorithm for COVID-19 online screening test based on its symptoms," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 25, no. 1, pp. 440–449, 2022, doi: 10.11591/ijeecs.v25.i1.pp440-449.

[7] A. Rajeshkumar and S. Mathi, "Smart solution for reducing COVID-19 risk using internet of things," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 25, no. 1, pp. 474–480, 2022, doi: 10.11591/ijeecs.v25.i1.pp474-480.

[8] R. Tosepu *et al.*, "Correlation between weather and Covid-19 pandemic in Jakarta, Indonesia," *Science of the Total Environment*, vol. 725, p. 138436, Jul. 2020, doi: 10.1016/j.scitotenv.2020.138436.

[9] M. Batista, "Estimation of the final size of the second phase of the coronavirus COVID 19 epidemic by the logistic model," *medRxiv*, pp. 1–13, 2020, doi: 10.1101/2020.03.11.20024901.

[10] P. Boldog, T. Tekeli, Z. Vizi, A. Dénes, F. A. Bartha, and G. Röst, "Risk assessment of novel Coronavirus COVID-19 outbreaks outside China," *Journal of Clinical Medicine*, vol. 9, no. 2, p. 571, 2020, doi: 10.3390/jcm9020571.

[11] A. M. Almeshal, A. I. Almazrouee, M. R. Alenizi, and S. N. Alhajeri, "Forecasting the spread of COVID-19 in kuwait using compartmental and logistic regression models," *Applied Sciences*, vol. 10, no. 10, 2020, doi: 10.3390/APP10103402.

[12] T. Kuniya, "Prediction of the epidemic peak of Coronavirus disease in Japan, 2020," *Journal of Clinical Medicine*, vol. 9, no. 3, p. 789, 2020, doi: 10.3390/jcm9030789.

[13] G. Röst *et al.*, "Early phase of the COVID-19 outbreak in hungary and post-lockdown scenarios," *Viruses*, vol. 12, no. 7, pp. 1–30, 2020, doi: 10.3390/v12070708.

[14] R. A. R. Bantan, C. Chesneau, F. Jamal, and M. Elgarhy, "On the analysis of new COVID-19 cases in Pakistan using an exponentiated version of them family of distributions," *Mathematics*, vol. 8, no. 6, 2020, doi: 10.3390/math8060953.

[15] "Novel Corona Virus 2019 Dataset | Kaggle." https://www.kaggle.com/sudalairajkumar/novel-corona-virus-2019-dataset (accessed Feb. 15, 2021).

[16] L. Rafea, A. Ahmed, and W. D. Abdullah, "Classification of a COVID-19 dataset by using labels created from clustering algorithms," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 21, no. 1, pp. 164–173, 2021, doi: 10.11591/ijeecs.v21.i1.pp164-173.

[17]  E. Dong, H. Du, and L. Gardner, "An interactive web-based dashboard to track COVID-19 in real time," *The Lancet Infectious Diseases*, vol. 20, no. 5, pp. 533–534, 2020, doi: 10.1016/S1473-3099(20)30120-1.
[18]  S. K. Dey, M. M. Rahman, U. R. Siddiqi, and A. Howlader, "Analyzing the epidemiological outbreak of COVID-19: A visual exploratory data analysis approach," *Journal of Medical Virology*, no. February, pp. 1–7, 2020, doi: 10.1002/jmv.25743.
[19]  F. Rustam *et al.*, "COVID-19 future forecasting using supervised machine learning models," *IEEE Access*, vol. 8, pp. 101489–101499, 2020, doi: 10.1109/ACCESS.2020.2997311.
[20]  S. Rath, A. Tripathy, and A. R. Tripathy, "Prediction of new active cases of coronavirus disease (COVID-19) pandemic using multiple linear regression model," *Diabetes and Metabolic Syndrome: Clinical Research and Reviews*, vol. 14, no. 5, pp. 1467–1474, 2020, doi: 10.1016/j.dsx.2020.07.045.
[21]  O. Ata and L. Hazim, "Comparative analysis of different distributions dataset by using data mining techniques on credit card fraud detection," *Tehnicki Vjesnik*, vol. 27, no. 2, pp. 618–626, 2020, doi: 10.17559/TV-20180427091048.
[22]  O. F. C. Atalaya *et al.*, "K-NN supervised learning algorithm in the predictive analysis of the quality of the university administrative service in the virtual environment," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 25, no. 1, pp. 521–528, 2022, doi: 10.11591/ijeecs.v25.i1.pp521-528.
[23]  W. M. Shaban, A. H. Rabie, A. I. Saleh, and M. A. Abo-Elsoud, "A new COVID-19 patients detection strategy (CPDS) based on hybrid feature selection and enhanced kNN classifier," *Knowledge-Based Systems*, vol. 205, p. 106270, 2020, doi: 10.1016/j.knosys.2020.106270.
[24]  "sklearn.metrics.r2_score — scikit-learn 0.24.1 documentation." https://scikit-learn.org/stable/modules/generated/sklearn.metrics.r2_score.html (accessed Feb. 15, 2021).
[25]  N. Sureja, B. Chawda, and A. Vasant, "A novel salp swarm clustering algorithm for prediction of the heart diseases," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 25, no. 1, pp. 265–272, 2022, doi: 10.11591/ijeecs.v25.i1.pp265-272.

# BIOGRAPHIES OF AUTHORS

**Wisam Dawood Abdullah** 🆔 🇬 SC 🅿 Associate Professor, he received his B.Sc. degree in computer science from Tikrit University, Iraq and his M.S. degree in Information Technology (with concentration in Telecommunications and Networks) from the University Utara Malaysia (UUM). He received an expert certification from Cisco Networking Academy CCNP, CCNA, CCNA Security, IoT, Entrepreneurship, Grid, Voice, Wireless Cloud, Linux, CCNA Cybersecurity and IT, also he is a NetAcad administrator in Cisco Networking Academy, Iraq, currently he is associate professor in the Tikrit University, and recently he selected as AWS Community Builder in Amazon. Research interest: Protocol Engineering, Network Analysis, Cybersecurity, Cloud Computing, Network Traffic Analysis, Data Mining, Future Internet, Internet of Things, AI, ML. He can be contacted at email: wisamdawood@tu.edu.iq.
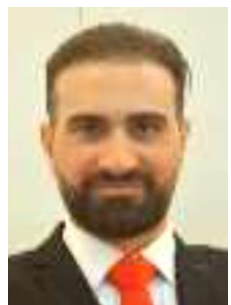
**Abdulrahman Ahmed Jasim** 🆔 🇬 SC 🅿 is a lectururer at Department of Network Engineering\Collage of Engineening, Al-Iraqia University. He received his BSc degree in Computer Engineering from Dijlah University in 2012, Msc degree in Elecrtrical and computer Engineering from Altinbas University, Turkey in 2018. He can be contacted at email: Abdulrahman.alsalmany@aliraqia.edu.iq.

**Layth Rafea Hazim** 🆔 🇬 SC 🅿 is a lecturer at the Cisco Networking Academy, Tikrit University, Iraq. He received his BSc degree in Computer Science from Tikrit University in 2007, Msc degree from the Altinbas University, Turkey in 2018. He worked as a head of Elctronic Computer Cemter ECC at Tikrit University during the period 2020 until now. He can be contacted at email: Layth.r.hazim@tu.edu.iq.