Predicting COVID-19 Disease Progression and Patient Outcomes Based on Temporal Deep Learning.

Course Name: Machine Learning

Course Code: CBS3006

Course Slot: L55+L56

ABSTRACT

Coronavirus 2019 (COVID-19) has caused worldwide health concern since December 2019. Since the onset of infection, patients will progress to a range of symptoms, such as fever, dyspnea or even death. Early detection of disease and predicting patient outcome helps guide treatment and resource allocation. However, there is no clear definition of the COVID-19 stage, and a few studies have talked about the continuation of COVID-19, which makes the need for this study apparent.

Ways

We developed an in-depth short-term learning method, based on a long-term conscious memory network (T-LSTM) and used an open online database, which included blood samples of 485 patients from Wuhan, China, for model training. Our approach can capture volatile relationships in a timely series of rare samples, which are ignored by existing tasks. Specifically, our approach predicted the effect of COVID-19 patients by considering both biomarkers and irregular time intervals. Then, we used patient presentations, extracted from T-LSTM units, to briefly type patient sections and explain the progression of COVID-19 disease.

Results

Using our method, the accuracy of the predictor outcome results was more than 90% for 12 days 98, 95 and 93% for 3, 6, and 9 days, respectively. Most importantly, we have identified 4 stages of COVID-19 progression with different patient conditions and mortality risks. We listed 40 disease-related biomarkers and assigned reference numbers for each category. Top 5 Lymph, LDH, hs-CRP, Indirect Bilirubin, Creatinine. In addition, we detected 3 complications - myocardial injury, liver damage, and kidney function damage. Predicting what 4 categories a patient is currently in can help doctors better diagnose and treat a patient.

Conclusions

To combat the COVID-19 epidemic, this paper aims to help physicians better diagnose and treat infected patients, provide qualified researchers with patterns of disease progression, and empower the effective use of medical devices. Our method predicted patient outcomes with high accuracy and identified the progression of the four-stage disease. We hope that the results and patterns obtained will help in the fight against this disease.

INTRODUCTION

The outbreak of Coronavirus 2019 (COVID-19) has caused widespread health concerns worldwide since December 2019; The disease was declared a pandemic by the World Health Organization (WHO) on March 11, 2020. More than seven million cases of COVID-19 have been reported worldwide, including more than 400,000 deaths (as of 15 June 2020). Although the disease is now under control in some countries, the WHO director warns that the epidemic is' still 'fast'. Due to its sudden appearance, many hospitals are still experiencing a shortage of medical supplies. For example, news in reported a shortage of medical services in New Delhi. Arizona met with a high hospital record as coronavirus cases increased. Proper allocation of resources according to patient status is required.

The solution to this problem involves determining the stages of disease progression by recording and predicting the outcome of COVID-19 patients. Thereafter, targeted treatment and distribution of medical resources can be performed on patients in different stages.

Longitudinal disease analysis is key to understanding disease progression, designing predictions and developing early diagnostic tools. The timing of the

disease may provide more information than the observation of stationary symptoms. Considering the complex patient regions, the number of interventions and the need for real-time, machine-readable learning methods inelectronic health records are desiderata to assist physicians.

Recently, an in-depth study, a recurrent neural network (RNN) could model effectively in a short-term sequence. Uses repetition to learn relationships between past, present and future. But primary RNN has long-term dependence problems . At that time, RNN only analyzes longitudinal data that are evenly distributed while blood samples from COVID-19 patient are randomly distributed and irregular time intervals between observations. Therefore, a method that can model this unusual time series for COVID-19 patients is required.

In this paper, we have also analyzed the blood samples of 485 patients from the stateof Wuhan, China. Medical records are compiled through standard case report forms, including disease information, demographic, clinical, laboratory and death information, from the online database under the MIT license.

We used an in-depth short-term study method of the Time-aware Short-Term Unit (T- LSTM) to demonstrate a unique time series for COVID-19 patients. T-LSTM can predict death with more than 98% accuracy within 3 days. During that time, we found four categories of COVID-19 patients. According to various categories, we have provided an analysis of the patient's condition and found related biomarkers and complications.

LITERATURE REVIEW

References	Method	Metrics	Limitations
1	We proposed a temporal deep learning method, based on a time-aware long short-term memory (T-LSTM) neural network and used an online open dataset, including blood samples of 485 patients from Wuhan, China, to train the model	Predicted the outcome of COVID-19 patients by considering both the biomarkers and the irregular time intervals.	More real clinical data are expected to be available for model validation and the model will be used to mine the inherent hidden features of other diseases.
2	The proposed predictive algorithm is a trained artificial intelligence-based network using 8,427 COVID-19 patient records from four healthcare systems.	The model provides a severity risk score along with likelihoods of various clinical outcomes, namely ventilator use and mortality.	V.S., D.C.,J.S.R.,E.R.,R.M.,S.V., D.C., and A.K. are employees of Siemens Healthineers, USA. R.S.,A.B., and P.S. were provided funding from Siemens Healthineers, USA.

3	Retrospectively analyzed the anamnestic data and laboratory parameters of 303 patients diagnosed with COVID-19 who were admitted to the Polyclinic Hospital of Bari during the first phase of the COVID-19 global pandemic.	After the pre-processing phase, we performed a survival analysis with Kaplan-Meier curves and Cox Regression, with the aim to discover the most unfavorable predictors. The target outcomes were mortality or admission to the intensive care unit (ICU).	The best model for predicting the risk of death was the decision tree, which resulted in ROC-AUC of 89.66%.
4	Research aimed to review and analyse articles about the occurrence of different types of infectious diseases, such as epidemics, pandemics, viruses or outbreaks, during the last 10 years	Articles on related topics were systematically searched in five major databases, namely, ScienceDirect, PubMed, Web of Science, IEEE Xplore and Scopus, from 1 January 2010 to 30 June 2020.	Articles generally do not provide the most correct assessments or assumptions.
5	Systematic review of artificial intelligence (AI) techniques used in the detection and classification of coronavirus disease 2019 (COVID-19) medical images in terms of evaluation and benchmarking.	Five reliable databases, namely, IEEE Xplore, Web of Science, PubMed, ScienceDirect and Scopus were used to obtain relevant studies of the given topic. Several filtering and scanning stages were performed according to the inclusion/exclusion criteria to screen the 36 studies obtained; however, only 11 studies met the criteria.	New and more efficient AI techniques were not recommended or suggested.
6	Study presents a new multi-biological laboratory examination framework for prioritising patients with COVID-19 on the basis of integrated MCDA methods.	The experiment was conducted on the basis of three phases. In the first phase, patient datasets containing eight biological laboratory examination criteria for six patients with COVID-19 were derived and discussed. The outcome of this phase was used to propose a decision matrix on the basis of the intersection between ‘biological	Prioritisation of patients with COVID-19 is a complex and multi-criteria decision-analysis (MCDA) problem due to (i) multiple biological laboratory examination criteria, (ii) criteria importance and (iii) trade-off amongst the criteria

		laboratory examination criteria' and 'COVID-19 patients list'.	
7	This study presents new integrated decision-making framework that handles the prioritisation of patients with COVID-19 and can detect the health conditions of asymptomatic carriers.	Real and simulation datasets from various medical perspectives are integrated to produce a new dataset involving 56 patients with different health conditions and can be used to check asymptomatic cases that can be detected within the prioritisation configuration. The first phase aims to develop a new decision matrix depending on the intersection between 'multi-laboratory criteria' and 'COVID-19 patient list'. In the second phase, entropy is utilised to set the objective weight, and TOPSIS is adapted to prioritise patients in the third phase. Finally, objective validation is performed.	The patients are prioritised based on the selected criteria in descending order of health situation starting from the worst to the best. The proposed framework can discriminate among mild, serious and critical conditions and put patients in a queue while considering asymptomatic carriers.
8	The occurrence of a large number of such cases in India has allowed us to draw our own conclusions about issues related to handling the situation within the framework of international guidelines.	A review of weekly trends in the prevalence of the disease reveals that there have been peaks in the prevalence of the disease in August–September and November–December, 2009. An almost similar increase in the number of such cases has been noticed in 2010 as well.	In the absence of clear-cut guidelines with regard to maximal drug dose, maximal prolongation in duration of treatment and combination of multiple of anti-viral drugs etc, extensive variability in management patterns have been observed amongst physicians.
9	a systematic review meta-analysis research of both randomised controlled trials and non-randomised controlled trials. We will search MEDLINE, EMBASE, Allied & Complementary Medicine, COVID-19 Research and WHO database on COVID-19 for primary studies assessing the effects of SDMs	The PRISMA-P checklist will be used while preparing this protocol. Joanna Briggs Institute guidelines (JBI Critical Appraisal Checklists) are used to assess the methodological	Ethics approval and consent will not be required for this systematic review of the literature as it does not involve human participation.

		qualities and synthesised performing thematic analysis.	
10	Research of COVID-19 Outbreak in Wuhan, Chine in 2019-20.	Sharing experienceandlearning from all geographical regions and across disciplines will be key to sustaining and further developing the progress being made.	All authors have a specialist interest in emerging and re-emerging pathogens. FN, RK, OD, GI, TDMc, CD and AZ are members of the Pan-African Network on Emerging and Re-emerging Infections (PANDORA-ID-NET) funded by the European and Developing Countries Clinical Trials Partnership the EU Horizon 2020 Framework Programme for Research and Innovation. AZ is a National Institutes of Health Research senior investigator. All authors declare no conflicts of interest.

OBSERVATIONS

Based on the predictive results, we found:

1) In-depth learning modes (T-LSTM, RNN, PNN and BPNN) have higher accuracy predicting COVID-19 results than in-depth study modes (Cox, k-NN, SVM and DT) as they have completed the conversion of the most indirect element with neural junction structures.

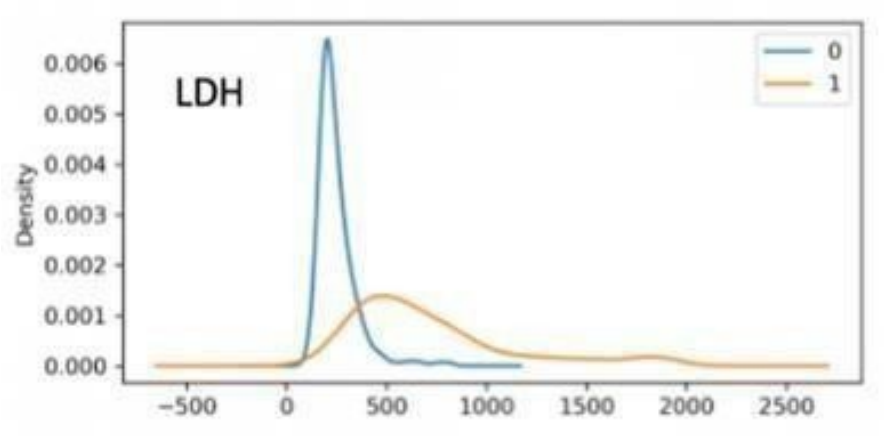
2) RNN-based models (T-LSTM and RNN) work better on time series data as they contain a regional connection to generate time delays and an output response link for loop.

3) The time-sensitive model (T-LSTM) has the best performance as it can model the time series at random intervals, which is a prominent feature of the COVID-19 blood sample database.

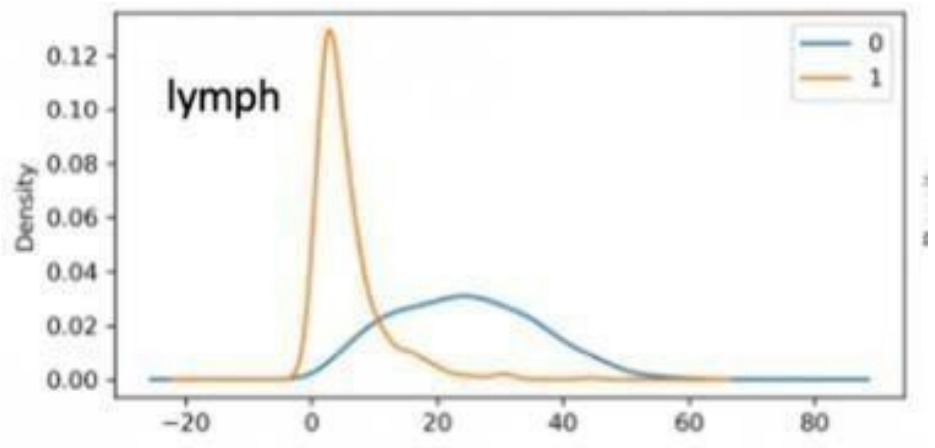
The dataset considered in this paper consists of the blood samples of COVID-19 patients of the Tongji Hospital of Tongji Medical College, Huazhong Universityof Science and Technology, Wuhan, China, collected between the 10th of January and 18th of February 2020. The dataset contains 80 characteristics from 375 patients with 6120 records as a training set. 11 patients with 757 records will be considered as a test case.
In the dataset, the average age of patients is 58.83, the survival rate is 53.6% and the ratio of male patients of female patients is 1:5:1.

Out of the features to be considered, the key features seem to be

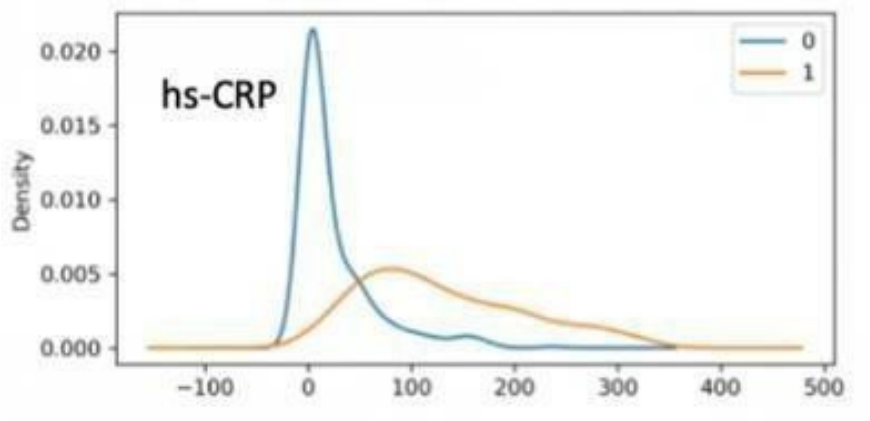
1. Lactic Dehydrogenase(LDH)

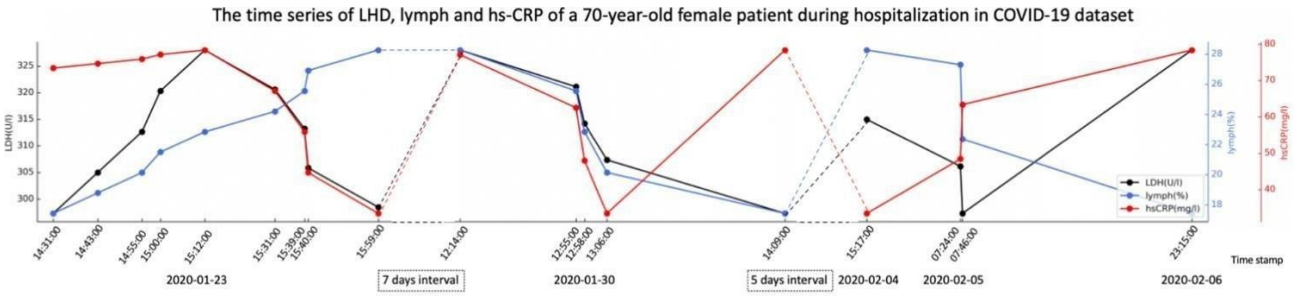


2. Lymphocytes (lymph)



3. High-sensitivity C-reactive protein (hs-CRP).





Dataset Description:

The dataset used in this experiment is a random dataset that displays the number of Active Cases, Confirmed Deaths and Recoveries throughout numerous countries in the world.

The data file 'covid_19_data.csv' was used to analyze and plot data. The dataset consists of the following columns in the headings:

1. ObservationDate
2. Province/State
3. Country/Region
4. Last Update
5. Confirmed Deaths
6. Recovered

(The dataset was collected from the databases of John Hopkins University)

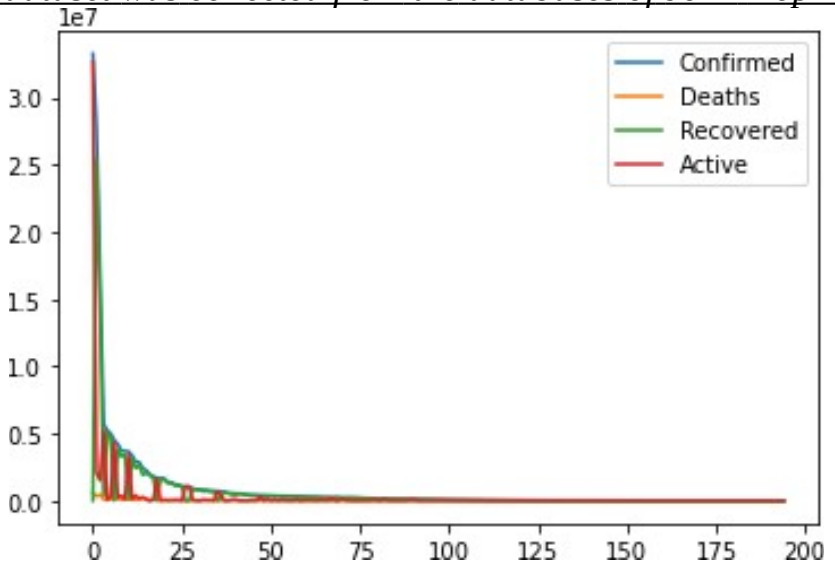


Fig: (Plotting the active, confirmed, deaths, recovered cases in the world and displaying the top 10 countries that are most affected.)

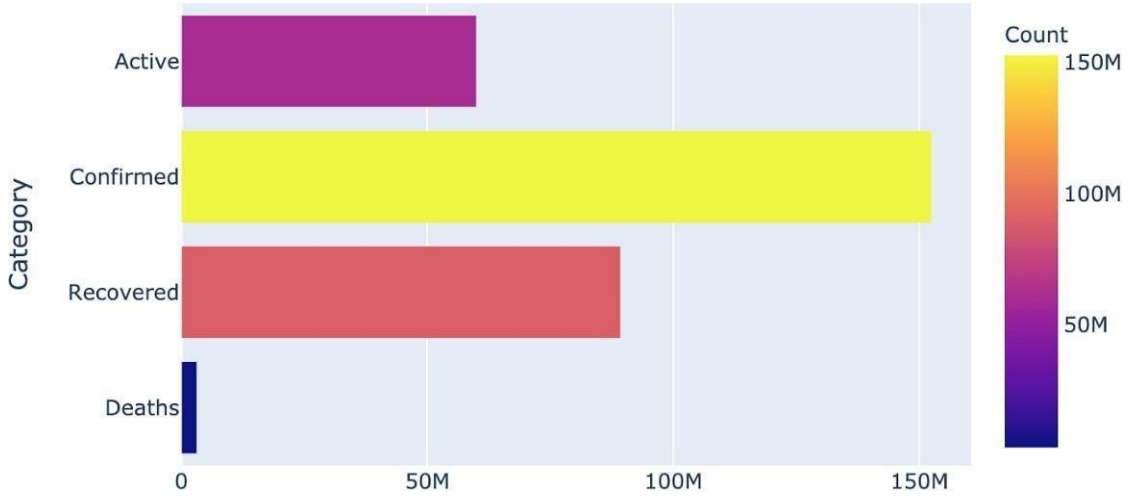


Fig: (Plotting the active, confirmed, deaths, recovered cases in the world and displaying the top 10 countries that are most affected.)

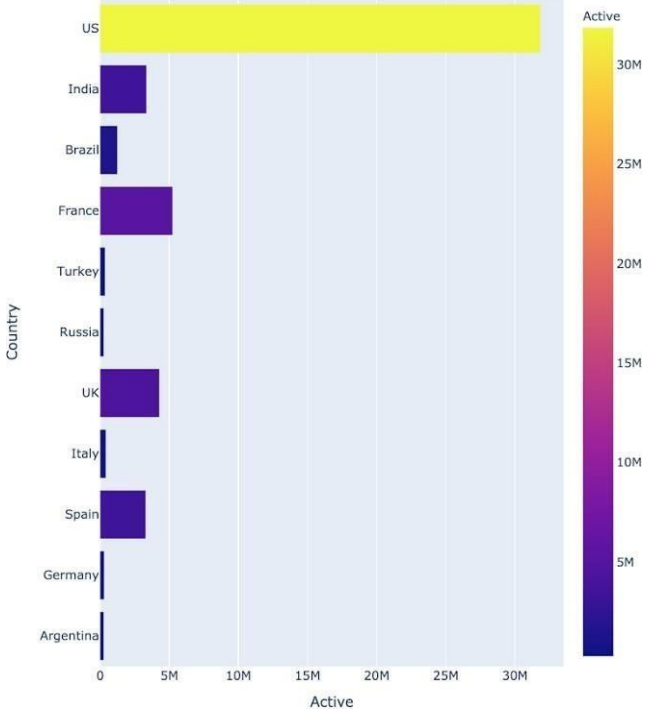


Fig: (Countries with Most Number of Confirmed Cases.)

IMPLEMENTATION OF ML TOOL

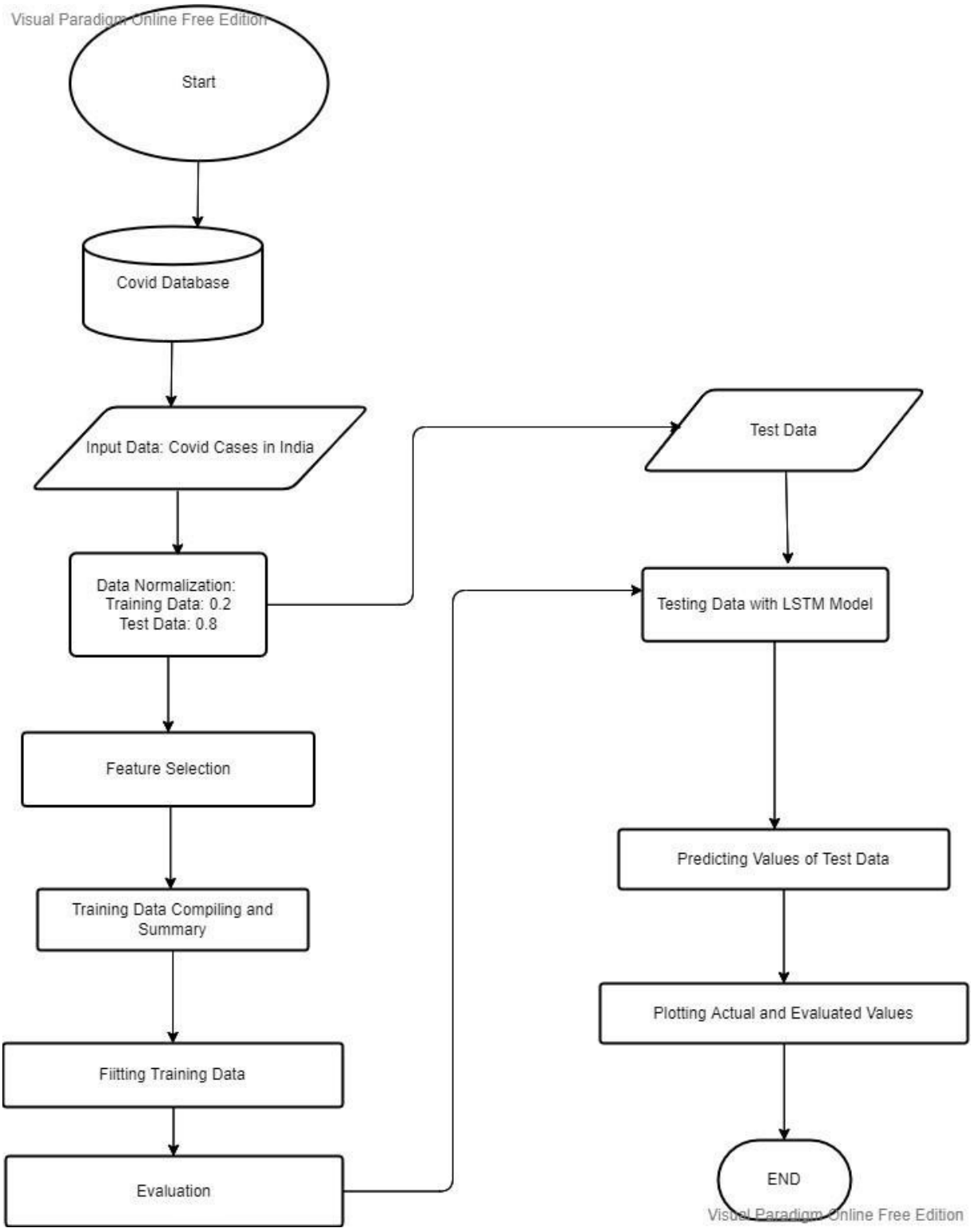
Long Short-Term Memory:
Long short-term memory (LSTM) is an artificial recurrent neural network (RNN) architecture used in the field of deep learning (DL). Unlike standard feedforward neural networks, LSTM has feedback connections. It can process not only single data points (such as images), but also entire sequences of data (such as speech or video). For example, LSTM is applicable to tasks such as unsegmented, connected handwriting recognition, speech recognition and anomaly detection in network traffic or IDSs

(intrusion detection systems).

A common LSTM unit is composed of a cell, an input gate, an output gate and a forget gate. The cell remembers values over arbitrary time intervals and the three gates regulate the flow of information into and out of the cell.

LSTM networks are well-suited to classifying, processing and making predictions based on time series data, since there can be lags of unknown duration between important events in a time series. LSTMs were developed to deal with the vanishing gradient problem that can be encountered when training traditional RNNs. Relative insensitivity to gap length is an advantage of LSTM over RNNs, hidden Markov models and other sequence learning methods in numerous applications.

Implementation of LTSM Model:



Implementation of LTSM Model in Python:

```
import pandas as pd

import numpy as np import seaborn as sns

import matplotlib.pyplot as plt
import plotly.express as px
import plotly.graph_objects as go
from pandas.plotting import register_matplotlib_converters
register_matplotlib_converters()
from plotly.offline import plot, iplot, init_notebook_mode
init_notebook_mode(connected=True)
pd.set_option('precision',0)
import warnings
warnings.filterwarnings('ignore')
table = pd.read_csv('covid_19_data.csv',parse_dates=['ObservationDate'])
pd.set_option('display.max_rows', table.shape[0])
pd.set_option('display.max_columns', None)
table.style.set_properties(subset=['ad_description'], **{'width-max': '100px'})
table.head(10).style.background_gradient(cmap='cool')

table['Country/Region'].value_counts().head(20)
table.isnull().sum()cases = ['Confirmed', 'Deaths',
'Recovered', 'Active']
table['Active'] = table['Confirmed'] - table['Deaths'] - table['Recovered']
table[['Province/State']] = table[['Province/State']].fillna('')
table[cases] = table[cases].fillna(0)
latest = table[table['ObservationDate'] == max(table['ObservationDate'])].reset_index()

latest_grouped = latest['Confirmed'] - latest['Deaths'] - latest['Recovered']
latest_grouped = latest.groupby('Country/Region')['Confirmed', 'Deaths', 'Recovered',
'Active'].sum().reset_index()




pred = latest_grouped.sort_values(by='Confirmed', ascending=False)
pred = pred.reset_index(drop=True)
cm = sns.light_palette("red", as_cmap=True)
pred.head(11).style.background_gradient(cmap=cm).background_gradient(cmap='Green
s',subset=["Recovered"])\
.background_gradient(cmap='Blues',subset=["Active"]).background_gradient(cmap='Ora
nges',subset=["Confirmed"])
```

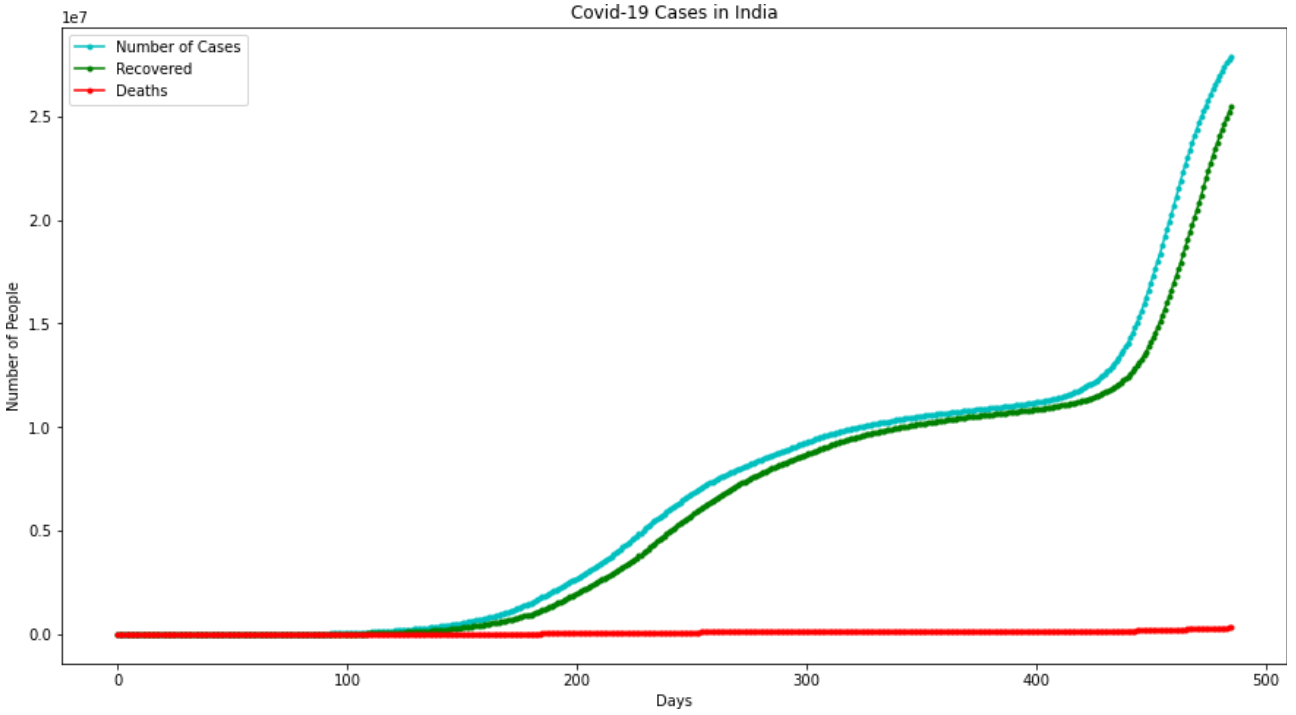
```
temp = table.groupby('ObservationDate')['Confirmed', 'Deaths', 'Recovered',
'Active'].sum().reset_index()
temp = temp.sort_values('ObservationDate', ascending=False)
cm = sns.light_palette("red", as_cmap=True)
temp.head(11).style.background_gradient(cmap=cm).background_gradient(cmap='Green
s',subset=["Recovered"]).background_gradient(cmap='Blues',subset=["Active"]).backgro
und_gradient(cmap='Oranges',subset=["Confirmed"])

total_count = pd.DataFrame({'Category':'Deaths', 'Count':temp.head(1)['Deaths']})
total_count = total_count.append({'Category':'Recovered','Count':int(temp.head(1)
['Recovered'])}, ignore_index=True)
total_count = total_count.append({'Category':"Confirmed",'Count':int(temp.head(1)
['Confirmed'])}, ignore_index=True)
total_count = total_count.append({'Category':"Active",'Count':int(temp.head(1)
['Active'])}, ignore_index=True)
fig = px.bar(total_count, x='Count', y='Category',
             hover_data=['Count'], color='Count',
             labels={}, orientation='h',height=400, width = 650)
fig.update_layout(title_text='Confirmed vs Recovered vs Death cases vs Active')
fig.show()


ind_confirmed = table[table['Country/Region'] ==
'India'].groupby(['ObservationDate'])['Confirmed'].sum().tolist()
ind_recovered = table[table['Country/Region'] ==
'India'].groupby(['ObservationDate'])['Recovered'].sum().tolist()
ind_deaths = table[table['Country/Region'] ==
'India'].groupby(['ObservationDate'])['Deaths'].sum().tolist()
plt.figure(figsize = (15,8))
plt.plot(ind_confirmed, color = 'c', marker = '.', label = 'Number of Cases')
plt.plot(ind_recovered, color = 'g', marker = '.', label = 'Recovered')
plt.plot(ind_deaths, color = 'r', marker = '.', label = 'Deaths')

plt.title('Covid-19 Cases in India')
plt.xlabel('Days')
plt.ylabel('Number of People')
plt.legend()
plt.show()
lstm_data =
table.groupby(['ObservationDate']).agg({'Confirmed':'sum','Recovered':'sum','Deaths':'su
m'})
```

```
training_set = lstm_data.iloc[:,0:1].values
from sklearn.preprocessing import MinMaxScaler
from sklearn.preprocessing import MinMaxScaler
sc = MinMaxScaler(feature_range=(0,1))
training_set_scaled = sc.fit_transform(training_set)

import numpy as np
X, y = [], []
time_steps = 45
for i in range(len(training_set) - time_steps):
    x = training_set_scaled[i:(i+time_steps), 0]
    X.append(x)
    y.append(training_set_scaled[i+time_steps, 0])
X = np.array(X)
y = np.array(y)

split = int(len(X) * 0.8)
X_train = X[:split]
X_test = X[split:]
y_train = y[:split]
y_test = y[split:]
X_train = np.reshape(X_train, (X_train.shape[0], 1, X_train.shape[1]))
X_test = np.reshape(X_test, (X_test.shape[0], 1, X_test.shape[1]))

import tensorflow as tf
from tensorflow.keras.layers import Dense, Dropout, Input, LSTM
from tensorflow.keras.models import Model
from tensorflow.keras.models import Sequential
from tensorflow.keras.optimizers import RMSprop
model = Sequential()
model.add(Input(shape=(1, time_steps)))
model.add(LSTM(48, return_sequences=True))
model.add(Dropout(0.4))
model.add(LSTM(48, return_sequences=True))
model.add(Dropout(0.2))

model.add(LSTM(48))
model.add(Dropout(0.2))
model.add(Dense(1, activation='relu'))
model.compile(loss = 'mean_squared_error',
              optimizer = 'adam',
              metrics = ['mean_squared_error'])
model.summary()
from keras.callbacks import ReduceLROnPlateau
batchsize = 100
epochs =  100
```

```
learning_rate_reduction = ReduceLROnPlateau(monitor='val_mean_squared_error',
                                            patience=3,
                                            verbose=1,
                                            factor=0.5,
                                            min_lr=1e-10)
history = model.fit(X_train,
                    y_train,
                    batch_size=batchsize,
                    epochs=epochs,
                    validation_split=0.2,
                    shuffle=False,
                    callbacks=[learning_rate_reduction])

y_pred =  model.predict(X_test)
y_pred = sc.inverse_transform(y_pred)
y_test = sc.inverse_transform(y_test.reshape(-1,1))
plt.plot(y_pred, color='red')
plt.plot(y_test, color='blue')
plt.title('Actual vs. Predicted Covid Cases (Test Data)')
plt.ylabel('Number of Cases')
plt.xlabel('Day')
plt.legend(['predicted', 'actual'])
```

RESULTS:

Fig: COVID Cases in India(Number of People Vs Days)



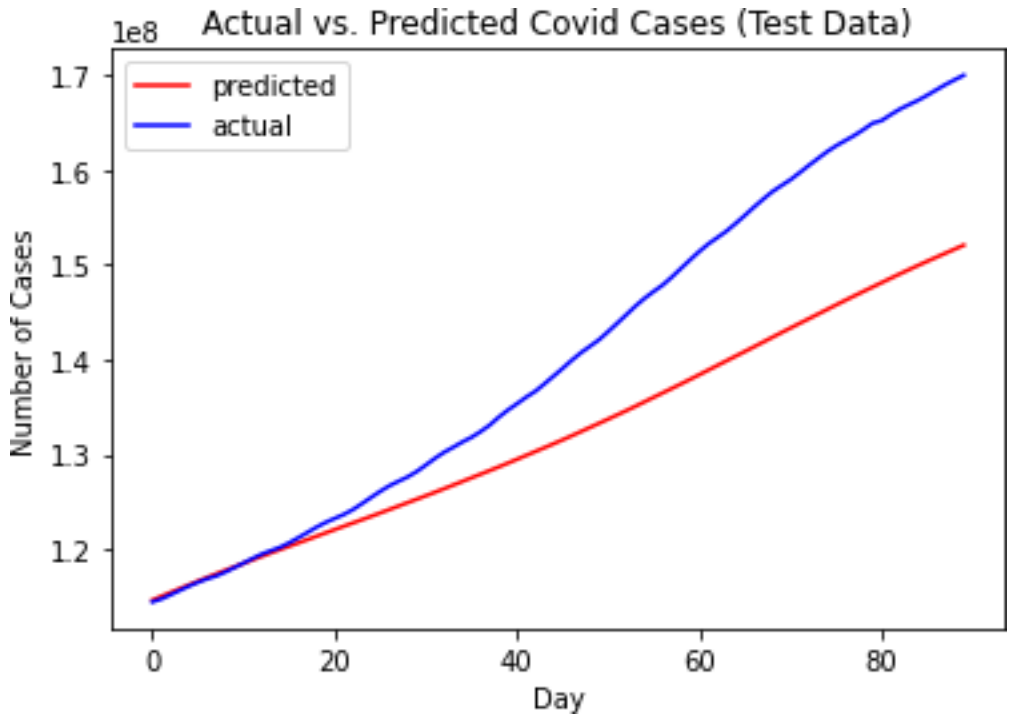


Fig: Prediction of Actual Cases Vs Predicted Cases in India (Number of Cases Vs Days)

By performing EDA we can get to know the dataset better, at the same time we can bring out meaningful information from the dataset and also could figure out if any flaw exists in the dataset or not.

For forecasting of the data, there are many algorithms available, even though there are many statistical models like Random Effect, Fixed Effect, etc, but all these models are linear models, therefore it can be difficult to adapt to multiple input forecasting problems.

The number of epochs used were 100 in this model. The LSTM model which is being used for forecasting has an exponential trend in the number of Covid-19 cases which is quite similar to the Real number of cases. This model can give better results if it is trained with more epochs.

CONCLUSION:

The COVID-19 pandemic and the ensuing lockdown has forced about several changes in the conventional means of achieving several human activities such as transportation, medical services, communication and so on, forcing most of the population to further rely on technology and machines. The contagious nature of the disease, its novelty and its effects have made research in the disease extremely difficult.

The recent advances in Machine Learning have allowed for extremely accurate predictions in the health and medical industries. Therefore, these concepts would be of great use to predict the progression of COVID-19(or of its variants) in the human body, thereby allowing for greater efficiency in the production, transportation and allocation of resources in times of decreased trade and scarcity. As proven by the results of the machine displayed in this paper, in-depth learning modes (T-LSTM, RNN, PNN and BPNN) have higher accuracy predicting COVID-19 results than in-depth study modes (Cox, k-NN, SVM and DT) as they have completed the conversion of the most indirect element with neural junction structures. The analysis of the patterns of LDH, Lymphocytes and hs-CRP in a patient have allowed for a fairly accurate prediction of the spread of COVID-19 in the human body These models would thereby help to prevent the spread of diseases in pandemics of the present and the future, by providing means of early diagnosis or by providing highly accurate predictions regarding the spread of the disease under consideration.

BIBLIOGRAPHY:

1. Predicting COVID-19 disease progression and patient outcomes based on temporal deep learning
   (Chenxi Sun, Shenda Hong, Moxian Song, Hongyan Li & Zhenjie Wang)
2. A deep learning approach for predicting severity of COVID-19 patients using a parsimonious set of laboratory markers
   (Vivek Singh, Rishikesan Kamaleswaran, Donald Chalfin, Parsa Hodjat, Dorin Comaniciu, Ali Kamen)
3. Predictive Machine Learning Models and Survival Analysis for COVID-19 Prognosis Based on Hematochemical Parameters

(Nicola Altini, Antonio Brunetti, Stefano Mazzoleni, Fabrizio Moncelli ,Ilenia Zagaria, Berardino Prencipe , Erika Lorusso , Enrico Buonamico ,Giovanna Elisiana Carpagnano , Davide Fiore Bavaro, Mariacristina Poliseno , Annalisa Saracino , Annalisa Schirinzi , Riccardo Laterza , Francesca Di Serio , Alessia D'Introno , Francesco Pesce ,and Vitoantonio Bevilacqua)

4. Alamoodi AH, Zaidan BB, Zaidan AA, et al. Sentiment analysis and its applications in fighting COVID-19 and infectious diseases: a systematic review. *Expert Syst Appl.* 2021;167:114155. [PMC free article] [PubMed] [Google Scholar]

5. Albahri OS, Zaidan AA, Albahri AS, et al. Systematic review of artificial intelligence techniques in the detection and classification of COVID-19 medical images in terms of evaluation and benchmarking: taxonomy analysis, challenges, future solutions and methodological aspects. *J Infect Public Health.* 2020;13:1381–96. [PMC free article] [PubMed] [Google Scholar]

6. Albahri AS, Al-Obaidi JR, Zaidan AA, et al. Multi-biological laboratory examination framework for the prioritization of patients with COVID-19 based on integrated AHP and group VIKOR methods. *Int J Inf Technol Decis Mak.* 2020;19:1247–69. [Google Scholar]

7. Albahri AS, Hamid RA, Albahri OS, et al. Detection-based prioritisation: framework of multi-laboratory characteristics for asymptomatic COVID-19 carriers based on integrated Entropy-TOPSIS methods. *Artif Intell Med.* 2021;111:101983. [PMC free article] [PubMed] [Google Scholar]

8. Suri JC, Sen MK. Pandemic influenza-Indian experience. *Lung India.* 2011;28:2–4. [PMC free article] [PubMed] [Google Scholar]

9. Regmi K, Lwin CM. Impact of social distancing measures for preventing coronavirus disease 2019 [COVID-19]: a systematic review and meta-analysis protocol [Preprint] doi: 10.1101/2020.06.13.20130294. Posted 2020 Jun 16. medRxiv 2020.06.13.20130294. [CrossRef] [Google Scholar]

10. Hui DS, I Azhar E, Madani TA, et al. The continuing 2019-nCoV epidemic threat of novel coronaviruses to global health: the latest 2019

novel coronavirus outbreak in Wuhan, China. *Int J Infect Dis.* 2020;91:264–6. [PMC free article] [PubMed] [Google Scholar]
11. Azhar E.I., Hui D.S.C., Memish Z.A., Drosten C., Zumla A. The Middle East Respiratory Syndrome (MERS) Infect Dis Clin North Am. 2019;33(4):891–905. - PMC - PubMed
12. Center for Health Protection, HKSAR . 2020. Press Release.https://www.info.gov.hk/gia/general/202001/11/P2020011100233.htm 11 Jan Accessed 12 Jan 2020. Available at:
13. Hui D.S.C., Zumla A. Severe acute respiratory syndrome: historical, epidemiologic, and clinical features. Infect Dis Clin North Am. 2019;33(4):869–889. - PMC - PubMed
14. Parr J. Pneumonia in China: lack of information raises concerns among Hong Kong health workers. BMJ. 2020;368:m56. doi: 10.1136/bmj.m56. Published 8 January 2020. - DOI - PubMed
15. WHO . Disease outbreak news. 2020. Emergencies preparedness, response. Pneumonia of unknown origin – China.https://www.who.int/csr/don/05-january-2020-pneumonia-of-unkown-cause-ch... 5 January, Accessed 12 Jan 2020. Available at:
16. WHO . Disease outbreak news. 2020. Emergencies preparedness, response. Pneumonia of unknown origin – China.https://www.who.int/csr/don/12-january-2020-novel-coronavirus-china/en/ 12 January, Accessed 12 Jan 2020. Available at:
17. World Health Organization . 2019. SARS (Severe Acute Respiratory Syndrome)https://www.who.int/ith/diseases/sars/en/ -accessed 12th January, 2020.
18. World Health Organization . 2019. MERS situation update.http://applications.emro.who.int/docs/EMROPub_2019_MERA_apr_EN_23513.pdf... (accessed August 20, 2019)
19. Zumla A., Dar O., Kock R. Taking forward a' One Health' approach for turning the tide against the Middle East respiratory syndrome coronavirus and other zoonotic pathogens with epidemic potential. Int J Infect Dis. 2016;47:5–9. - PMC - PubMed
20. ProMED "Novel coronavirus - China (01): (HU) WHO, phylogenetic tree". Archive Number: 20200112.6885385.
21. Zhao YF, Shou MH, Wang ZX. Prediction of the number of patients infected with COVID-19 based on rolling grey Verhulst models. *Int J Environ Res Public Health.* 2020;17:4582. [PMC free article] [PubMed] [Google Scholar]
22. Achterberg MA, Prasse B, Ma L, et al. Comparing the accuracy of several network-based COVID-19 prediction algorithms. *Int J Forecast.* 2020 Oct 9; doi: 10.1016/j.ijforecast.2020.10.001. [Epub]. [PMC free article] [PubMed] [CrossRef] [Google Scholar]
23. Fernandez A, Obiechina N, Koh J, et al. Survival prediction algorithms for COVID-19 patients admitted to a UK district general hospital. *Int J Clin Pract.* 2021;75:e13974. [PMC free article] [PubMed] [Google Scholar]

24. Li Y, Horowitz MA, Liu J, et al. Individual-level fatality prediction of COVID-19 patients using AI methods. *Front Public Health.* 2020;8:587937. [PMC free article] [PubMed] [Google Scholar]
25. Siwiak M, Szczesny P, Siwiak M. From the index case to global spread: the global mobility based modelling of the COVID-19 pandemic implies higher infection rate and lower detection ratio than current estimates. *PeerJ.* 2020;8:e9548. [PMC free article] [PubMed] [Google Scholar]
26. Bhandari S, Tak A, Gupta J, et al. Evolving trajectories of COVID-19 curves in India: prediction using autoregressive integrated moving average modeling [Preprint] Posted 2020 Jul 7. Research Square. c. [Google Scholar]
27. Muhammad LJ, Algehyne EA, Usman SS, et al. Supervised machine learning models for prediction of COVID-19 infection using epidemiology dataset. *SN Comput Sci.* 2021;2:11. [PMC free article] [PubMed] [Google Scholar]
28. Brownlee J. *Introduction to time series forecasting with Python: how to prepare data and develop models to predict the future.* San Juan (PR): Machine Learning Mastery; 2017. [Google Scholar]
29. Parham PE, Michael E. Outbreak properties of epidemic models: the roles of temporal forcing and stochasticity on pathogen invasion dynamics. *J Theor Biol.* 2011;271:1–9. [PubMed] [Google Scholar]
30. Ray S. Indore: Analytics Vidhya; 2015. Aug 14, 7 Regression Techniques you should know [Internet] [cited 2020 Aug 2]. Available from: https://www.analyticsvidhya.com/blog/2015/08/comprehensive-guide-regression/ [Google Scholar]

(Dataset Used in the Experiment: https://www.kaggle.com/datasets/sudalairajkumar/novel-corona-virus-2019- dataset)

(Code is Available at https://colab.research.google.com/drive/1gJKZG12yuIfdEGtCRLCwi9OCyR6rFCvk?usp=sharing)