Prediction and Forecasting of Coronavirus Cases using Artificial Intelligence Algorithm

Sunday Adeola AJAGBE[1*[0000-0002-7010-5540]], Joseph Bamidele AWOTUNDE[2[0000-0002-1020-4432]], Matthew Abiola OLADIPUPO[3[0000-0002-5984-6217]], Olawale, E. OYE[4]

[1]Computer Engineering Department, Ladoke Akintola University of Technology, Ogbomoso, Nigeria
[2]Computer Science Department, University of Ilorin, Ilorin, Nigeria
[3]Electronic and Electrical Engineering Department, Ladoke Akintola University of Technology, Ogbomoso, Nigeria
[4]Statistics Department, University of Ibadan, Ibadan, Nigeria

saajagbe@pgschool.lautech.edu.ng[1]
awotunde.jb@unilorin.edu.ng[2],
moladipupo@pgschool.lautech.edu.ng[3],
ooye157@stu.ui.edu.ng[4]

Abstract: Artificial intelligence (AI) and Internet of Things (IoT) are evolving technologies that are being adopted by different facets of human endeavor. Among other reasons, it is being used to increase productivity and enhance efficiency. Specifically, AI is used in the medical field for many professional activities like medical diagnosis, telemedicine and medical laboratory investigations. Coronavirus otherwise known as COVID-19 pandemic, being the first of such, putting nations of the world on lockdown almost at the same time, is a contagious respiratory and vascular disease caused by severe acute respiratory syndrome coronavirus. Consequently, academic, business and other socio-economic activities were paralyzed as physical distancing, new greeting style and the use of nose mask and face shield were not only introduced but established. However, various COVID-19 medical dataset that were collected by Johns Hopkins University, were yet to be fully investigated using cutting edge technologies that would bring out the hiding facts vis-à-vis the trend of the deadly pandemic. This study proposed a framework for comprehensive investigation of COVID-19 medical dataset, Exploratory Data Analysis was done to visualize how this infectious disease affected most nations of the world, Facebook Prophet model was built for prediction of active, confirmed, death and recovered cases. The model performance was evaluated using R-Score. The R-Score were achieved for active, confirmed, death and recovered cases was 0.9999, 0.9999, 0.9999 and 0.9998 respectively. This data-driven and empirical innovation study will assist engineers, scientists and medical experts especially, for deployment of vaccine, contact tracing and monitor the trend so as to curb the spread of the Corona virus pandemic that is ravaging the world.

Keywords: Artificial Intelligence, COVID-19, Healthcare System, Internet of Things, Machine Learning, Medical Dataset

1 Introduction

The COVID-19 can spread complication before and after the onset, and it is extremely infectious in. Monitoring and lockdown have to encompass anyone with symptoms and properly isolate persons who have been infected from those who are not, to allow good containment. Patients carrying the virus could either be minor symptomless (like fever, sore throat, and sneezing) or have serious clinical signs (such as respiratory failure pneumonia and eventually death) [1]. Transmittable SARS-CoV-2 condition is called "corona virus disease" (COVID-19) [2]. Gratitude to the recent developments in analytical techniques of Artificial Intelligence (AI), Information and Communication technologies (ICTs) and big data will aid manage the immense, unparalleled volume of data generated from patient monitoring, real-time tracking of disease outbreaks, now-casting/predicting patterns, daily briefings and public updates [3]. According to the recent report, ages between 30 and 79 years, approximately 86.60% are susceptible to COVID-19 pandemic of all patients infected so far, thus have a median age of 47 years [4].

The used of broad variety of advanced technology brought about the Internet of Things (IoT) that transform conventional objects into intelligent objects, from implanted devices and networking technologies to Internet protocols. Gul, et al., (2017) the potential impact of IoT in driving the economic development of IoT-based services are expected to offer many business opportunities. Furthermore, industries like transport, agriculture, urban infrastructure, defense, and retail, have an IoT market of approximately 15% overall. These estimates show the enormous and sharp progress of IoT products and produced data within a short period. In the years ahead, the estimated market share is expected to improve drastically. Kevin Ashton proposes the IoT for the first time in 1999, and many IoT methods have been created by various researchers, such as the Internet of Medical Things (IoMT), the Internet of People, the Internet of All, the Internet of Marks, the Internet of Info, the Internet of Services, etc [5, 6]. Therefore, the empirical investigation was done using artificial intelligence/machine learning (AI/ML) algorithm with the implementation of Plotly and Matplotlib visualization tool that may be used with Internet of Things (IoT) to support contact tracing in future were initiated, and results shown the locations that are affected. This data-driving innovation can assist engineers and scientists for clinicians, and healthcare experts around the world for deployment of vaccine, contact tracing and curb the spread of the corona virus that is ravaging the whole world.

Artificial intelligence (AI) will create smart framework to automatically track and forecast the spreading of this outbreak [7, 8, 9]. A genetic algorithm may also be built to remove the visual characteristics of this infection. It has the potential to provide patients with daily alerts and also to offer better options for COVID-19 epidemic follow-up. AI can easily determine this virus' level of transmission by recognizing the fragments and 'hot spots' and can effectively track the individuals' contacts and even monitor them. It can foresee this spread of the disease future path, and possibly reoccurrence. These technologies can also monitor and predict the existence of the epidemic from the data, social media and broadcasting channels present, about the threats of the outbreak and its probable spreading. It can also forecast the number of useful cases and deaths in any area. AI will help recognize the areas, citizens and communities most affected and take effective measures [10, 11].

Advent of Corona virus generated large amount of medica dataset that call for the adoption of machine learning (ML) techniques and applications as it enables computers to imitate and adapt human-like behaviour. ML techniques can interact with different applications to perform actions, and produce something the system can learn and use as experience for the next action(s). The overview of analytics data method which enables computers to learn and do what comes naturally to human's mind was carried out. It includes machine learning's nomenclature and applications' describing what, how and why. The technology roadmap of machine learning was discussed to understand and verified. The primary objective of the study was to give insight into why machine learning is the future [12, 13].

An empirical investigation was done using artificial intelligence/machine learning (AI/ML) with the implementation of Plotly and Matplotlib visualization tool that may be used with Internet of Things (IoT) to support contact tracing in future were initiated, and results shown the locations that are affected. This data-driving innovation can assist engineers and scientists for clinicians, and healthcare experts around the world for deployment of vaccine, contact tracing and curb the spread of the corona virus that is ravaging the whole world. The degrees of COVID-19 active cases vary in various counties of the world, Canada, Russia, United States, Brazil were among the countries of the world that have higher number of COVID-19 active cases while Madagascar, Angola and most of the African countries has low active cases. The extent of COVID-19 confirmed cases also varies in various counties of the world, Brazil, Canada, China, Russia, United States were among the countries with higher number of confirmed cases while Angola, Libya, Madagascar, Zambia and most of the African countries has low confirmed cases. The extent of COVID-19 death cases also varies in various counties of the world like other cases Brazil, Canada, China, Russia, United States were among the countries of the world that have higher death cases. Many African countries on the other hand have very low death cases.

This may due to the climatic and other environmental conditions in African. The extent of COVID-19 recovered cases also varies in various counties of the world as usual. Bolivia, Canada were among the countries with high recovery cases, very low recovery cases of COVID-19 in Algeria, Australia, Brazil, Kazakhstan and Russia. The higher level of recovery cases in Bolivia, Canada and few other countries may not be unconnected with the use and implemented contact tracing that might have reduces the spread. Most European countries implemented this at the early stage of the pandemic as a measure of curbing the spread. The forecasting results of active, confirmed, death and recovered cases for the next seven days were done using Artificial Intelligence/machine learning model built with Facebook prophet. The evaluation of the model built for prediction and forecasting shows high degree of accuracy, having used three evaluation metrics to measure its performance. The model performance for forecasting model was evaluated using three performance metrics, the R-score, mean square error and absolute mean error. These evaluation metrics help to measure performance of univariance time-series forecasting models in this AI/ML medical dataset study. The Active, Death, Confirmed and Recovered cases forecasting models were evaluated. The evaluation results of forecasting model shown that R-Score of 0.9999, 0.9999, 0.9999, and 0.9998, Mean square error of 32950433.47, 43988753.26, 724820.40 and 43611232.23 for Active, Death, Confirmed and Recovered cases respectively. Also, absolute mean error of 4229.98, 4776.94, 603.04 and 3814.23.

1.1 Organization of the Chapter

The organization of this chapter is as follows: Section 2 give the details literature review of contract tracing during COVID-19, the dataset created during this pandemic was also review. The AI and ML used during COVID-19 pandemic were also reviewed. Section 3 discusses the methodology used in this chapter, the data collection, data preprocessing, explanatory data analysis, models building, prediction and forecasting of the dataset, and the performance evaluation. Section 4 contained the results and discussion. Section 5 is the conclusion of this the chapter.

1.2 Concept of Artificial Intelligence

Artificial Intelligence (AI) is referred to the process of human intelligence simulation in machines that makes machine think, mimic human actions and execute simple and complex tasks [14]. Although, applications of AI are endless, but are specifically more applicable to any machine that has traits associated with human being, such as problem solving, reasoning, learning and robotics. AI is evolving and beneficial to different industries and its operations are guided by multi-disciplinary approach like computer science, linguistics, mathematics, psychology, and more [15]. AI is an advancement of deep learning (DL) and machine learning (ML) and are creating a paradigm shift in virtually every sector of the technological industry today [14, 16].

The branch of AI that has a versatile and central role to play in the investigation of a global pandemic like COVID-19 and medical challenges generally is Machine learning [17]. Data-driven innovation during pre-pandemic, pandemic and post pandemic period may be used by medical technologies and scientists to control the pandemic or cripple its spread. The far-reaching scientific discoveries and innovations in data science and AI/ML have capacity to investigate COVID-19 pandemic, put an end to its spread and predict or forecast likely new cases or any related outbreak using the available medical data [18].

2 Literature Review

In December 2019, series of atypical pneumonia cases occurred at Wuhan, Hubei, China, that were identified as Coronavirus Disease 2019 (COVID-19). It was first identified, as a cluster of unknown patients with beta coronavirus pneumonia linked to the seafood wholesale market. It is a contagious respiratory and vascular disease and very infectious caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) [19, 20, 21]. All evidence suggests that the Corona virus have already been actively disseminating some months earlier in Italy before it became a global problem [22]. As at March 5, 2020 in China, about 80,000 people have been infected, this includes about 60,000 cases at Hubei province. After several confirmation of many cases from different countries, the World Health Organization (WHO) declared this epidemic to be a Public Health Emergency of International Concern on 31st January, 2002 [23].

The new disease was not comprehended well for a while now. Although, the previous studies have shown that the main symptoms are chest tightness, cough, diarrhea, difficulties in breathing, fever, and fatigue, whole the upper respiratory track symptoms are obstruction of the nasal, loss of taste and smell, sore throat and running nose for treatment and diagnosis of COVID-19, the symptom being some of the first fourteen days of exposure to COVID-19, Polymerase Chain Reaction (PCR) can be used to diagnose infectious person, the test identified by the genetic fingerprint [24, 25, 26, 27]. There are numbers of ways by which COVID-19 spread, which primarily includes saliva, droplets of body fluid and excretions that can make up a small aerosols and droplets, which can also disperse through the breathes, coughs, sings, speaks or sneezes by an infected person(s). Corona virus disease originated from bat and it caused acute diarrhea syndrome in pig in 2018, and it was learnt to have transmitted to bat, birds, camels, cat, human, pigs and rats [28, 29].

The suspected mode of transmissions of the virus is through direct physical contact with an infected person or contaminated surface [22]. Infection happens mainly when people are close to each other, it can spread within two days before infected persons can show symptoms and from asymptomatic individuals. Preventive measures are covering sneezes and coughs, washing of hands, social distancing, quarantining, ventilation of indoor spaces sanitizers (alcoholic-based). Face shield used, face masks or coverings has been recommended and established in public places like never before to minimize transmission risks. Three vaccines for the treatments for COVID-19 have just been found in a very limited quantity, they are; Pfizer, Moderna and AstraZeneca. These vaccines have national regulatory authorities but are yet to receive WHO approval, and there are plans to supply them to various countries of the world [30, 31], though there are other managements for the symptoms; (experimental measures, isolation and supportive care before now), control of the spread (contact tracing) that have been used before now.

Some available contact tracing applications uses the Internet of Things (IoT) which has greater capacity to monitor infectious pandemic and it is a beneficial supplement to traditional surveillance system in data acquisition [32]. Corona virus has become a pandemic, spreading all over the world. Engineers, scientists and medical experts are working round the clock to develop the vaccine, kits for testing, enhance monitoring, prevent the spread. Internet-based applications such as CoronaApp, StopCovid and CovTracer have also been developed to monitor the Corona virus status of individuals with very little or no attention to possibilities of using IoT as many countries are yet to put in place IoT enables devices to operate, that can be used to monitor and later curb the spreading of COVID-19. Some of these devices are designed either to sense and record, monitor and respond. The review of COVID-19 and IoT related studies done in the research shows that much needs to be done to provide IoT based architecture and monitoring techniques to curtail the disperse of Corona virus. The use of mobile phone and smart devices for healthcare delivery that have been impactful significantly to the world has potential technologies to deliver in this regard too, thus the spread of COVID-19 can be curbed significantly and substantially to improve the healthcare sector through data-driven research that can be supported by IoT [28, 33]

There is a need to properly investigate any diagnosed and confirmed cases of Corona virus and all the previous contact traced so as to control the spread of the disease. According to the World Health Organization, the virus spread from infected persons to another primarily through droplet, discharge from nose through contact transmission or saliva [34]. To this end, contact tracing is healthcare sector essential tool to the control of the COVID-19 transmission [35]. The contact tracing process which is to identify and be wield people that are recently exposed to the infected Corona virus patient to limit further spread is done mostly through physical tracing and 14-days follow-up since the day of exposure to the infected person. If thoroughly employed, it can break the outsending chain of the current novel, Corona virus and suppress the outbreak by giving a higher chance of adequate control and helping to reduce the pandemic magnitude, but the efficiency of the physical tracing and the stress involved is of high magnitude [18]. Most of the developed and developing countries of the world started employing digital applications for contact tracing and detection of Corona virus, most of which are digital applications and IoT based. These applications include but not restricted to the internet-based application, utilizing different technologies like Global Positioning System (GPS), Bluetooth, contact details, social graph, mobile tracking data, network-based API, system physical address and card transaction data [18, 28].

The real-time digital contact tracing with the support internet of things achieves better performance and more accuracy than non-digital process. All these digital apps are developed to collect individual personal data, which can be analyzed by AI and ML tools to trace a person who is vulnerable to the COVID-19 due to their recent contact. Access to these personal data can be achieved with the use of IoT, just as businesses around the world are adopting IoT as a means of enhancing business activities and provide security by extension. Medical field and health institutions have been adopting IoT technology for healthcare delivery, a relevant example is the use of IoT for contact tracing, about thirty-six countries of the world launched contact tracing application for COVID-19 in five continents. Details is presented in Table 1 [14]. Lalmuanawma, Hussain, & Chhakchhuak (2020) presented 36 Countries of the world who had successfully employed digital contact tracing techniques to manage the disperse of Corona virus by using decentralized, centralized, or hybrid of both techniques to lessen the effort and augment the effectiveness of both the orthodox and traditional health care diagnosis procedures. Table 1 show the details of the continents, countries, contact tracing Apps as well as the dates that the applications were launch for the control of the spread of COVID-19.

Table 1: Countries with COVID-19 Contact Tracing Application

S/N	Continents	Country	Contact tracing App	Launch
1	Africa	Ghana	GH Covid-19 Tracker App	12th April, 2020
2	Asia	Bahrain	BeAware Bahrain	31st March, 2020
3		India	Aarogya Setu	April 2, 2020
4		China	Conjunction with Alipay	little information available
5		Cyprus	CovTracer	May, 2020
6		Iran	Mask.ir	May, 2020
7		Israel	HA Magen	March, 2020
8		Jordan	AMAN App - Jordan	May, 2020
9		Malaysia	MyTrace	3rd May, 2020
10		Qatar	Ehteraz	May, 2020
11		Singapore	Trace Together	20th March, 2020
12		Saudi Arabia	Corona Map	3rd April, 2020
13		South Korea	Non-App-Based	May, 2020
14		United Arab Emirate	TraceCovid	May, 2020
15	Australia/Oceania	Australia	COVIDSafe	14th April, 2020
16		New Zealand	NZ Covid Tracer	20th May, 2020

17	Europe	Austria	Stopp Corona	March, 2020
18		Bulgaria	ViruSafe	May, 2020
19		Czech Republic	eRouska(sFacemask)	15th April, 2020
20		Estonia	Estinia's App	April, 2020
21		Finland	Ketju	May, 2020
22		France	StopCovid	May, 2020
23		Germany	CoronaApp	May, 2020
24		Hungary	StopCovid	13th April, 2020
25		Iceland	CoronaApp	April 2020
26		Ireland	HSE Covid-19 App	May, 2020
27		Italy	Immuni	May, 2020
28		Latvia	Apturi COVID	29th May, 2020
29		North Macedonia	StopKorona	13th April, 2020
30		Norway	Smittestopp	16th April, 2020
31		Poland	ProteGO	May, 2020
32		Switzerland	SwissCovid	20th May, 2020
33		Turkey	Hayat Eve Sigar	April, 2020
34		United Kingdom	NHS COVID-19 App	May, 2020
35	South America	Colombia	CoronaApp	12th April, 2020
36		Mexico	CovidTracer	May, 2020

2.1 Artificial Intelligence/ Machine Learning and Corona Virus Pandemic

Artificial intelligence (AI) empowered machine learning (ML) to make clever machines that imitate brilliant conduct and supports in dynamic with almost no human impedance. AI and machine learning are both extraordinary advances tools that makes significant additional intriguing using suitable dataset and data science field in particular [16, 36]. Beyond COVID-19, AI/ML could be of help in predicting future pandemics, by the use of statistical-based means and converging with artificial intelligence and available data. This could position AI as a key enabler of medical transformation, from a reactive to a proactive system in the nearest future [37, 38]. This current global pandemic (COVID-19) makes the clinicians, healthcare, scientists and other medical experts around the world to keep on searching for a new technology to support in tackling the COVID-19 medical pandemic. The evidence of Artificial Intelligence (AI) and Machine Learning (ML) tools and the previous diseases encourage researchers around the world to investigate and develop application to the novel COVID-19 outbreak [18]. Table 2 summarizes some of AI/ML prediction studies available.

Table 2: AI/ML Technology for Prediction and Forecasting of COVID-19

Author, year	**AI/ML Model**	**Data types**	**Predicted cases**	**Evaluation metrics**	**Results**
Ribeiro, da-Silva, Mariani, & Coelho, (2020)	SVR and stacking-ensemble	Clinical data	Cumulative confirmed cases	The improvement index, symmetric mean absolute percentage error criteria and absolute mean error.	0.87% - 3.51%
Gupta, Pandey, & Pal, (2020)	SEIR model and Regression model	Clinical data	Confirmed	Root mean squared log error	1.52 and 1.75
Khakharia, et al., (2020)	Auto-Regressive Moving Average (ARMA), XGBoost Regressor (XGB)	Clinical data	Confirmed	Not stated	99.93%
Abdulmajeed, Adeleke, & Popoola, (2020)	combines an Auto-Regressive Integrated Moving Average model (ARIMA), a Holt - Winters Exponential Smoothing model and Prophet - an additive regression model developed by Facebook,	Clinical data		Not stated	Not stated
Santos, (2020)	exponential smoothing model with multiplicative error and multiplicative trend components	Demographic and Clinical data	Death	Not stated	Above 80% for the 10 days

					prediction
Yan, et al., (2020)	XGBoost classifier	Clinical and demographic data	demographic, epidemiological, clinical, laboratory and mortality outcome information	The mortality of individual patients	Accuracy was 90%
Siess, (2020)	SIR model	Medical data	Mortality	Not stated	75%
Ardabili, et al., (2020)	SIR and SEIR model	Medical data	Infection	Not stated	SEIR model
Chimmula & Zhang, (2020)	Deep Learning network	Demographic and Clinical data	the trends and possible stopping time of the current COVID-19 outbreak in Canada and around the world	Not stated	Not stated
Chakraborty & Ghosh, (2020)	Regression tree model and Hybrid Wavelet-autoregressive integrated moving average model.	Demographic and Clinical data	New COVID-19 cases	Not stated	Not stated
Petropoulos & Makridakis, (2020)	exponential smoothing model with multiplicative error and multiplicative trend components	Demographic and Clinical data	Confirmed case	Information Criteria that measure the maximum likelihood of a model while penalizing for its complexity	Above 80% for the 10 days prediction

Anastasso poulou, Russo, Tsakris, & Siettos, (2020)	calibrating the parameters of the SIRD model	Clinical data	forecast the evolution of the outbreak	Not stated	Not Stated

So far, researchers use machine learning models to predict and forecast deaths, confirm, survival/ recovered of COVID -19 cases especially in a separate studies, there is no critical investigation of COVID-19 medical dataset across various countries of the world, that has forecasted four different cases (active, confirmed, death and recovered cases) of Corona virus at the same time with the same dataset using state-of-the-art techniques lying in AI/ML tools and visualizing the degrees of infection of the pandemic across various countries of the world. To address the concerns raised by the spread of COVID-19, this chapter proposed a comprehensive research framework responsible for the investigation of COVID-19 medical data across various countries of the world. Combatting Corona virus can be contributed to, immensely but its contributions have not been impactful due to the lack of robust COVID-19 medical data which is also hindered by the requirement of data privacy, healthcare sector and human-AI interaction. Massive diagnostic data gathering of infectious people and their contacts will be essential AI- based solutions to curtail economic loss and save lives. The proliferations of data on COVID-19 has cost World Health Organization (WHO) a huge investment to collect data for predicting, forecasting and monitoring of the Corona virus disease in other to ensure the curtailment of the disease [39, 40].

3 Methodology

This section provides method employed for the comprehensive investigation of the world COVID-19 medical data vis-à-vis artificial intelligence/ machine learning (AI/ML) application in this chapter. The world COVID-19 investigation was carried out in this chapter was an empirical investigation and experimental based. The outline of the methodology includes the experimental environment, data collection, data preprocessing, exploratory data analysis (EDA), building of model, prediction and forecasting, as well as evaluation of model.

3.1 Execution Environment

The execution environment was an hp envy machine with 16g ram, 1 terabyte ssd hd 8th gen. The machine also has python 3.8 installed, eight (8) important python libraries were imported for the experimentation. Matplotlib which is an ECoG

Visualization tool that was designed by John Hunter to overcome the problem single license of preoperatory software and can be used by multiple investigators was imported, as an important library for visualization. Pandas, Seaborn, Plotly and other important libraries were also imported for model building, plotting of the graphs, performance evaluation of the model and other experimentation purposes. Figure 1 reveals the details of the libraries imported for the prediction. This chapter also proposed framework for the prediction and forecasting investigation of COVID-19 that may be used for similar infectious diseases, the proposed framework is shown in figure 2. The prediction and forecasting infection across various countries of the world that was done through an empirical means in this chapter using suitable COVID-19 medical dataset which was sourced online and data wrangling which comprises of data preprocessing and cleaning were carried out. Prediction and forecasting as well as the visualization of the infectious area were anchored on the proposed framework depicted in figure 1. The model performance for prediction was evaluated using the R-score, mean absolute error and mean square error as the evaluation metrics. Four different cases were evaluated (Active, Death, Confirmed and Recovered cases).

```python
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
import plotly
import plotly.express as px
import plotly.graph_objects as go
from fbprophet import Prophet
from sklearn.metrics import mean_squared_error, r2_score, mean_absolute_error
```

Fig. 1. Important libraries for the study (Source: Authors)

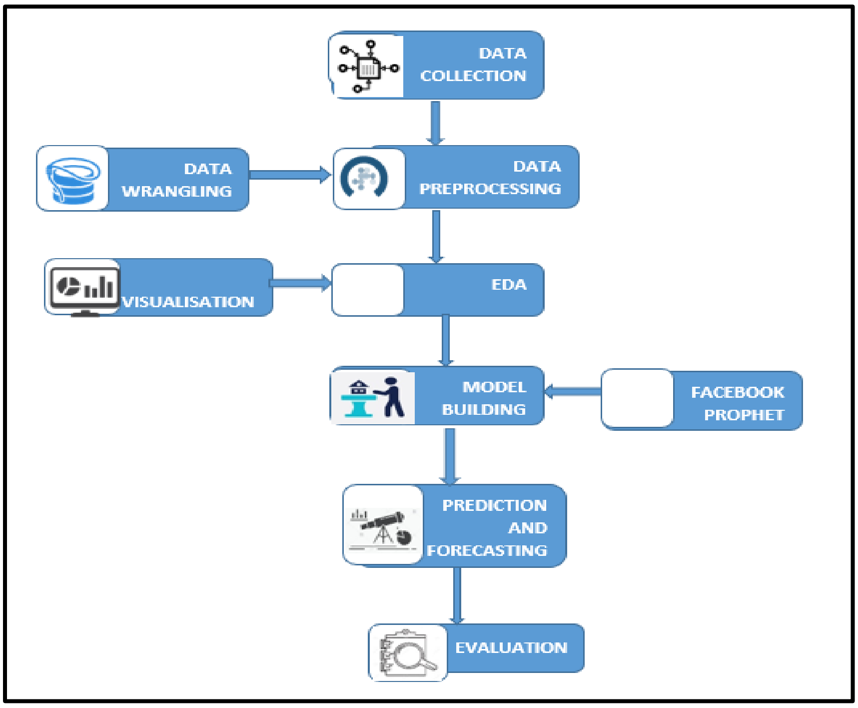

Fig. 2. Framework of the World COVID-19 AI based Investigation (Source: Authors)

3.2 Data Collection

COVID-19 world medical dataset from reliable source was retrieved from [41], the university has been providing quality data and other world class educational resources for research around the world, the data was collected between January and May 2020, through the website on the National Centre for Disease Control of Various Countries. The dataset comprises of the following features; province/states, country/region, Lat, Long, date, confirmed, death, recovered cases, this made it suitable for the studies compared to others available dataset. The population of the dataset stands at 172,481 people across various countries of the world within the period under review.

3.3 Data Preprocessing

The data preprocessing and data wrangling of the COVID-19 medical data retrieved has some features like Date, which was converted from object format to date-time format to make the code to recognize it as date feature. This is necessary in the forecasting and exploratory data analysis during the investigation. The dataset was sorted to make the latest cases appeared first and later grouped by countries in preparing the dataset for data visualization. Some features were renamed to make them readable to the code and remove ambiguities, features like province/state was renamed to state, country/region was changed to country and ‘active’ was created as a new feature. Active case was created by removing recovered and death from the confirmed cases to get the number of active COVID -19 patient in each case.

3.4 Exploratory Data Analysis

The exploratory data analysis (EDA) was done to get more insight to the COVID-19 medical data investigation. The EDA of world COVID-19 medical data in this chapter was carried out on chloroplet map, which is one the AI data visualization tools employed for the investigation. The tool was used to visualized four types of COVID-19 cases that were investigated in over 200 countries of the world. Plotly visualization library of Python used as visualization tool for the geographical visualization on the chloroplet maps, the was made possible because the COVID-19 dataset used in this research contains longitude (long) and latitude (lat). Matplotlib for chats labelling and seaborn for chats plotting (line chat and bar chat) and easy interpretation (points plots).

3.5 Model Building

A frontline prediction and forecasting algorithm Facebook prophet was used to build models in this experiment, because of its univariate time series nature of the algorithm, this made it considered more suitable to be used for this study, it also has better accuracy and easy to build than most of the prediction and forecasting algorithm used in building machine learning model. Four distinct models were built, they are model to predict and forecast active, confirmed, death and recovered cases using Facebook prophet. The model scaling was done and the dataset was splitted into Test set and Training set, the Test set was 20% while the Training set was 80% so as to avoid overfitting and underfitting of the model.

3.6 Prediction and Forecasting

The prediction and forecasting of COVID-19 medical data in this investigation was to predict each case for seven days, ie. Active, Death, Confirmed and Recovered cases. After the model had been built, the prediction was done and the forecasting was done using point plotting and bar plotting.

4 Results and Discussion

The results of investigation of COVID-19 medical data across various countries of the world using AI/ML applications towards augmenting the investigation from multiple angles were presented. It also provides an enhanced COVID-19 medical dataset that can be rely on for a real-time world COVID-19 treatment around the world. The results of the experiment shown that the prevalence of the pandemic at US (North America) in the months of April, this could help in the understanding of the disease and consequently curb the menaces.

4.1 Results Exploratory Data Analysis (EDA)

The investigation results of exploratory data analysis of COVID-19 data various across countries of the world are shown on the chloroplet maps in figure 4-7 for the four cases; active, confirmed, death and recovered cases. In the figure 4-7, the higher the intensity of the color on the map, the higher the number of active, death, confirmed and recovered cases vice-vasal, and to know the number of confirmed cases in a country, the pointer/mouse can be used to point to that country and that will be provided as shown in figure 3. The implementation was done using Plotly for map visualization, Matplotlib for chats labelling and seaborn for chats plotting (line chat and bar chat) and easy interpretation (points plots).

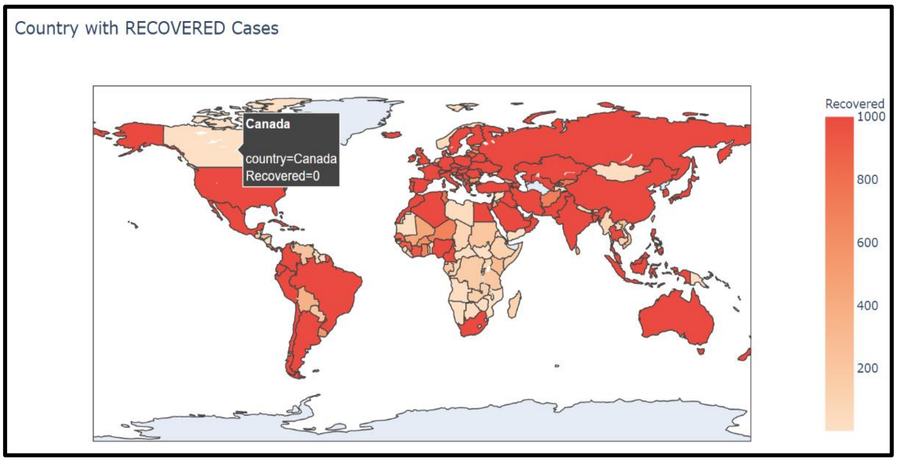

Fig. 3. world chloroplet map showing Canada and the number of the recovered cases per 12,481 and 37,742,154 inhabitants in a particular week (Source: Authors)

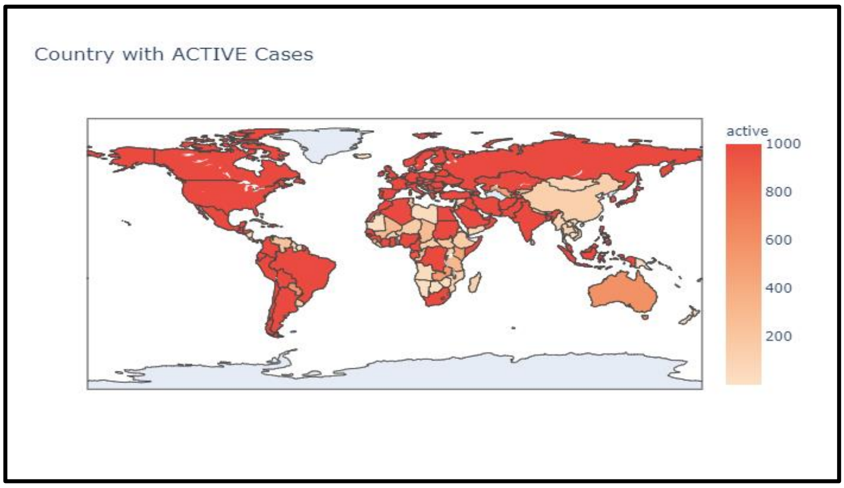

Fig. 4. World chloroplet map showing the degree of the active cases around the world (Source: Authors)

The degrees of COVID-19 active cases vary in various counties of the world as shown in figure 4. The is high level of active COVID-19 cases in Canada, Russia, United States, Brazil and so on. But Angola, Madagascar and most of the African countries has low active cases of COVID-19.

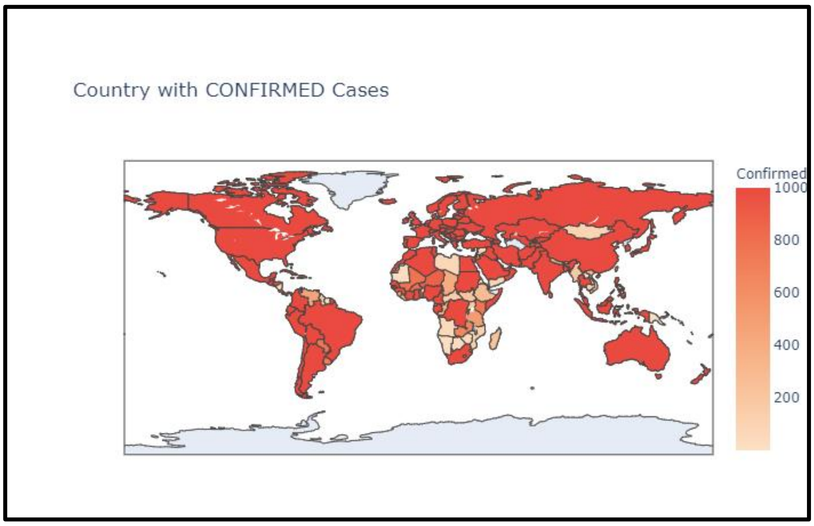

Fig. 5. World chloroplet map showing the degree of the confirmed cases around the world (Source: Authors)

The degrees of COVID-19 confirmed cases vary in various counties of the world as shown in figure 5. The is high level of confirmed COVID-19 cases in Brazil, Canada, China, Russia, United States and so on. But Angola, Libya, Madagascar, Zambia and most of the African countries has low confirmed cases of COVID-19.

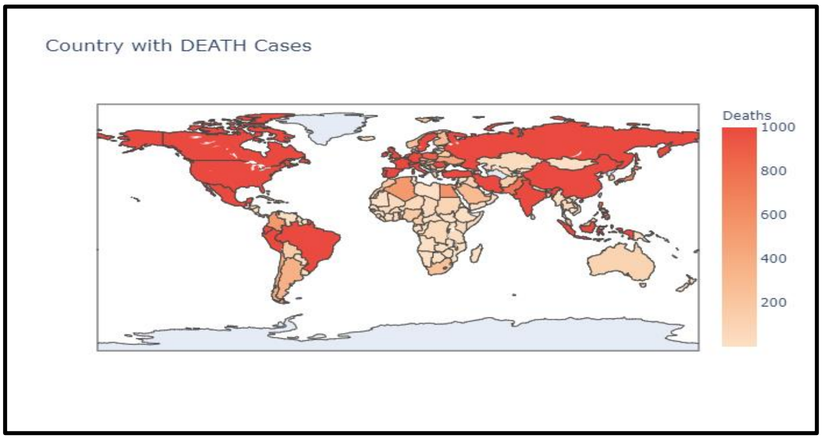

Fig. 6. World chloroplet map showing the degree of the death cases around the world (Source: Authors)

The degrees of COVID-19 death cases vary in various counties of the world as shown in figure 6. The is high level of death COVID-19 cases in Brazil, Canada, China, Russia, United States and so on. But African countries has very low death cases of COVID-19. This may due to the climatic and other environmental conditions in African

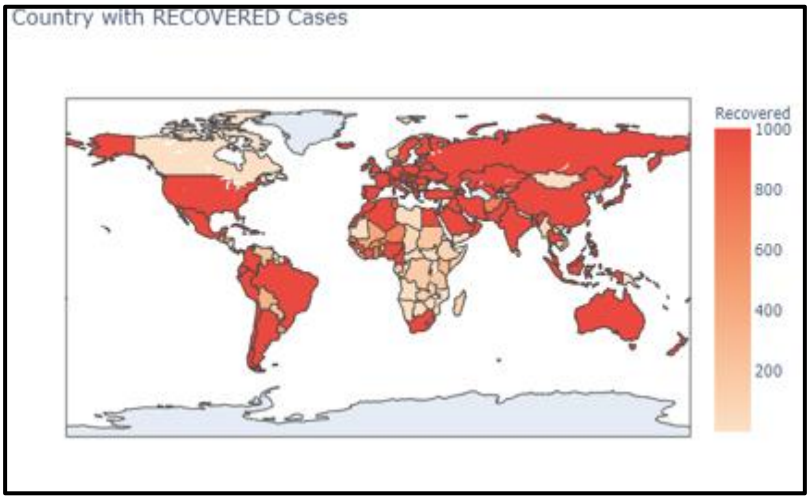

Fig. 7. World chloroplet map showing the degree of the recovered cases around the world (Source: Authors)

The degrees of COVID-19 recovered cases vary in various counties of the world as shown in figure 7. The is high level of recovered COVID-19 cases in Bolivia, Canada this may due to the level of the preparedness in these countries as these countries implemented contact tracing and other measures for the timely curtailment of COVID-19 of the disease. But there very low recovery cases of Corona virus in Algeria, Australia, Brazil, Kazakhstan and Russia.

4.1.1 Exploratory Data Analysis of World COVID-19 Results

The results of top twenty countries of the world that have highest number of COVID-19 for the four identified cases were reveal in the investigation of COVID-19 pandemic across the world with AI/ML tools. These were represented with bar chats in Figures 8-11.

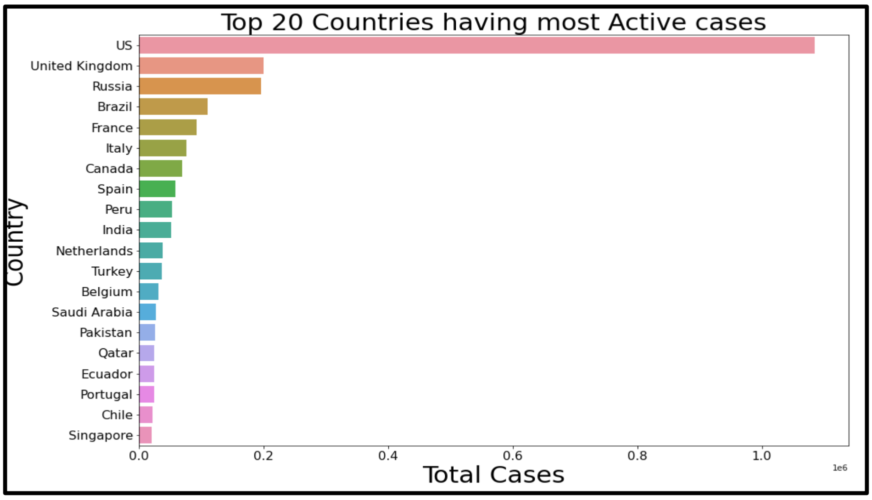

Fig. 8. Top twenty countries of the World with Active COVID-19 cases (Source: Authors)

The investigation results in figure 8 bar chart reveal that US has the highest number of Corona virus Active cases and Singapore was having the lowest number of active cases.

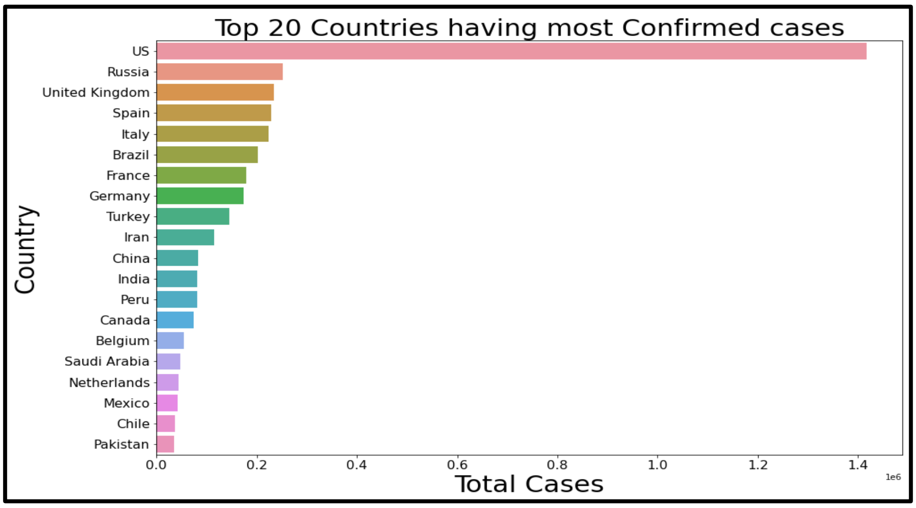

Fig. 9. Top twenty countries of the World with confirmed COVID-19 cases

The investigation results in figure 9 bar chart reveal that US has the highest number of Corona virus Confirmed cases and Pakistan was having the lowest number of active cases.

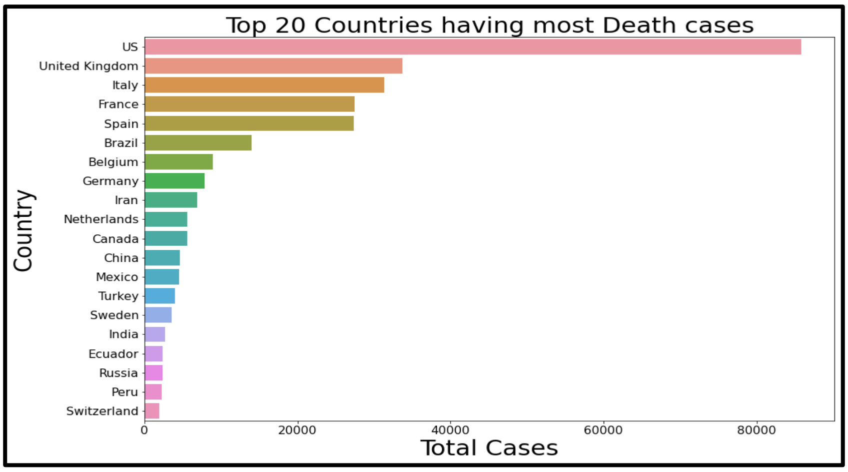

Fig. 10. Top twenty countries of the World with death COVID-19 cases

The investigation results in figure 10 bar chart reveal that US has the highest number of Corona virus Confirmed cases and Switzerland was having the lowest number of active cases.

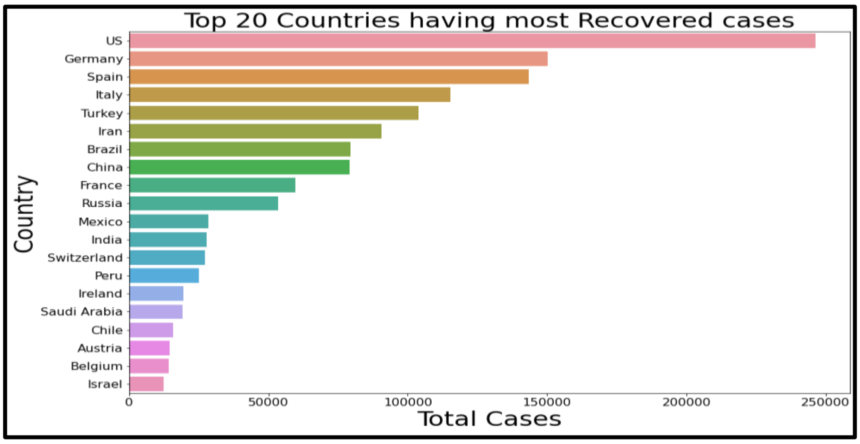

Fig. 11. Top twenty countries of the World with recovered COVID-19 cases

The investigation results in figure 11 bar chart reveal that US has the highest number of Corona virus recovery cases and Israel was having the lowest number of active cases.

4.2 Forecasting investigation results

The forecasting results of the four identified cases using AI/ML tools were shown in Figure 12-15 respectively. The points on the figure shows the trend of the forecast for the whole world. Point plotting and Bar plotting were used to represent the seven days forecast. The Total case in 100,000 were plotted against Dates of each cases.

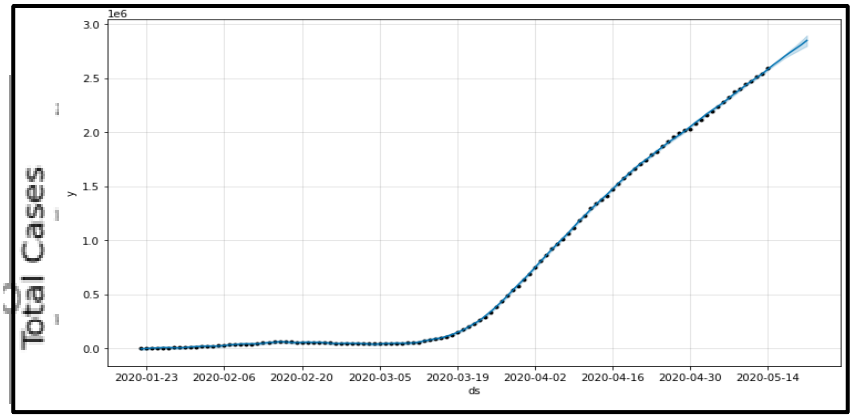

Fig. 12. World COVID-19 Active Cases forecast (Source: Authors)

The active cases forecast for the next seven days was on the increase as shown in figure 12, the areas that point plotting did not appear towards the tail end of the plot. The total cases in million was plotted against the dates (which was on interval of 15 days)

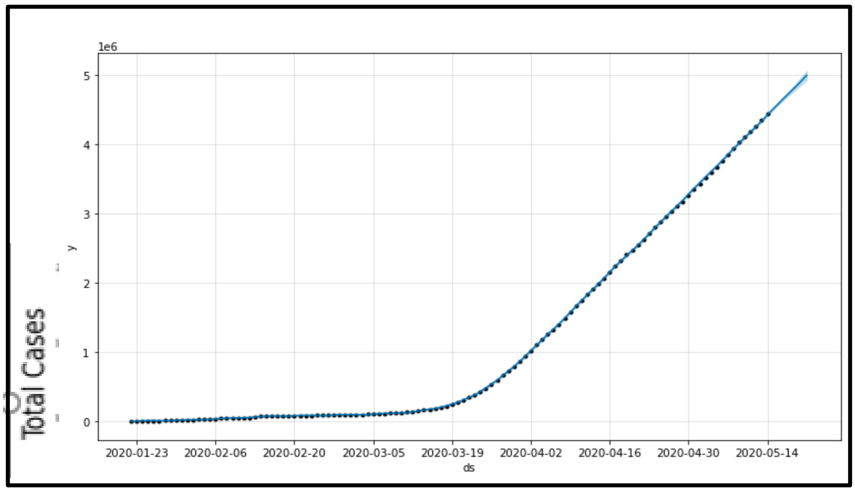

Fig. 13. World COVID-19 Confirm Cases forecast (Source: Authors)

The confirm cases forecast for the next seven days was on the increase as shown in figure 13, the areas that point plotting did not appear towards the tail end of the plot. The total cases in million was plotted against the dates (which was on interval of 15 days)

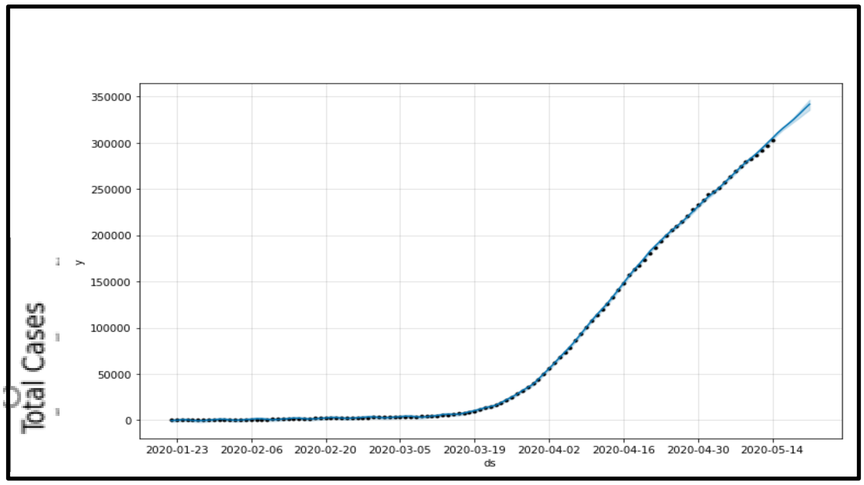

Fig. 14. World COVID-19 Death Cases forecast (Source: Authors)

The death cases forecast for the next seven days was on the increase as shown in figure 14, the areas that point plotting did not appear towards the tail end of the plot. The total cases in million was plotted against the dates (which was on interval of 15 days)

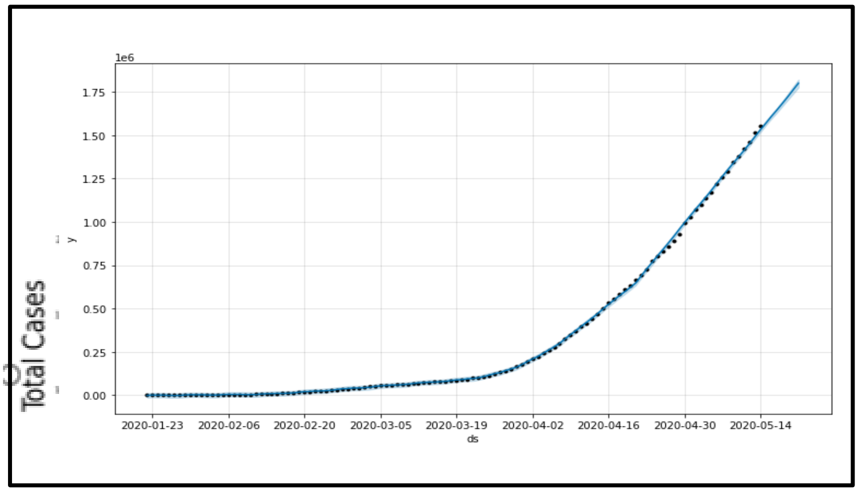

Fig. 15. World COVID-19 Recovered Cases forecast (Source: Authors)

The recovery cases forecast for the next seven days was on the increase as shown in figure 15, the areas that point plotting did not appear towards the tail end of the plot. The total cases in million was plotted against the dates (which was on interval of 15 days)

4.3 Performance Evaluations

The model performance for prediction was evaluated using the R-score, mean square error and absolute mean error as the metrics. These evaluation metrics help to measure performance of univariance time-series forecasting models and they were used in this investigation on the dataset. Four different cases forecasted were evaluated (Active, Death, Confirmed and Recovered cases). The evaluation results of forecasting model shown that R-Score of 0.9999, 0.9999, 0.9999, 0.9998 for Active, Death, Confirmed and Recovered cases respectively. Mean square error of 32950433.47, 43988753.26, 724820.40 and 43611232.23 for Active, Death, Confirmed and Recovered cases respectively. Also, absolute mean error of 4229.98, 4776.94, 603.04 and 3814.23. The summary of the evaluation result is presented in Table 3.

Table 3: Performance of prediction model

Metrics Cases	R-Score	Mean square error	Absolute mean error
Active	0.9999	32950433.47	4229.98
Confirmed	0.9999	43988753.26	4776.94
Death	0.9999	724820.40	603.04
Recovered	0.9998	43611232.23	3814.23

5 Conclusion

This chapter has provided a comprehensive investigation of COVID-19 dataset, using AI/ML experimental-based method through visualization, AI/ML also many algorithms for prediction and forecasting. These algorithms are used for building different machine learning models used in prediction and forecasting, some of these algorithms has been used for prediction of Corona virus cases in a clinical or medical dataset, with the aims analyzing the dataset for planning and control the spread of the pandemic. Corona virus disease that is ravaging the whole world was studied with a view of predicting and forecasting four (4) possible cases of Corona virus: the active, death, confirmed and recovered cases using AI/ML tools, and visualize the extent active, death, confirmed and recovered cases of the infection in various countries of the world. Matplotlib, Pandas, Seaborn, Plotly and other important python libraries were used for model building, plotting of the graphs, performance evaluation of the model and other experimentation purposes in the prediction, forecasting and visualization of the world Corona virus medical dataset. A study framework was proposed and used in this chapter. The active cases, confirmed cases, and recovered cases of the world COVID-19 medical dataset were visualized using chloroplet map and showing the extent of each these cases on the world map. This chapter has used Facebook prophet algorithm to build a machine learning algorithm for prediction and forecasting of active, death, confirmed and recovered/survival cases of COVID-19 cases. This study provides a strong insight into COVID-19 medical data making adequate plans toward combatting the pandemic. This study also reveal that contact tracing can be aided with IoT for controlling the spread and data access or data capturing of COVID-19 or another related pandemic. The possibility predicting four (4) different cases of COVID-19 was revealed. The Point plotting and Bar plotting were used to represent the seven days forecast on the chat. The prediction investigation results reveal that US has the highest numbers of Corona virus active cases and Singapore was predicted to have the lowest number of active cases in seven predictions. The prediction of active cases for top 20 countries of the world, the prediction investigation results reveal that US has the highest numbers of Corona virus confirmed cases and Pakistan was predicted to have the lowest number of confirmed cases in seven predictions. The prediction of confirmed cases for top 20 countries of the world. The prediction investigation results reveal that US has the highest numbers of COVID-19 death cases and Switzerland was predicted to have the lowest number of death cases in seven predictions. The prediction of death cases for top 20 countries of the world. The prediction investigation results reveal that US has the highest numbers of COVID-19 recovered cases and Israel was predicted to have the lowest number of recovered cases in seven predictions, the prediction of recovered cases for top 20 countries of the world. There is still room for deployment of real-time prediction and forecasting solutions to assist the medical personnel in the discharge of their duties, this is hope to be achieve through integration of IoT to machine learning for contact tracing and other data science techniques that may lead to variation in the active, confirmed, death and recovered cases which can help health care personnel in their plan towards the control of the pandemic.

References

[1] A. Perrella, N. Carannante, M. Berretta and M. Rinaldi, "Novel Coronavirus 2019 (Sars-CoV2): a global emergency that needs new approaches?," *European Review for Medical and Pharmacological Sciences,* vol. 24, no. 4, pp. 2162-2164, 2020.

[2] C. Wirawan and Shierly, "Neurological manifestations in COVID-19 infection," *International Journal of Research in Medical Sciences,* vol. 8, no. 11, pp. 4168-4173, 2020.

[3] S. Kannan, P. Shaik Syed Ali, A. Sheeza and K. Hemalatha, "COVID-19 (Novel Coronavirus 2019) – recent trends," *European Review for Madical and Pharmacological Sciences,* vol. 24, no. 4, pp. 2006-2011, 2020.

[4] C. Arunkumar and S. Ramakrishnan, " Prediction of cancer using customised fuzzy rough machine learning approaches.," *Healthcare Technology Letters,* vol. 6, no. 1, pp. 13-18, 2018.

[5] J. Gubbi, R. Buyya, S. Marusic and M. Palaniswami, "Internet of Things (IoT): A vision, architectural elements,and future directions.," *Future Generation Computer System,* vol. 29, pp. 1645-1660, 2013.

[6] J. A. Stankovic, "Research directions for the internet of things.," *IEEE Internet Things Journal,* vol. 1, pp. 3-9, 2014.

[7] F. S. Lu, S. Hou, K. Baltrusaitis, M. Shah, J. Leskovec, R. Sosic, J. Hawkins, J. Brownstein, G. Conidi, J. Gunn, J. Gray, A. Zink and M. Santilana, "Accurate Influenza Monitorin and Forecasting Using Novel Internet Data Steams: A Case Study in the Boston Metropolis," *JMIR Public Health and Survelance,* vol. 4, no. 1, pp. 5-18, 2018.

[8] F. E. Ayo, R. O. Ogundokun, J. B. Awotunde, M. O. Adebiyi and A. E. Adeniyi, "Severe Acne Skin Disease: A Fuzzy-Based Method for Diagnosis," *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics),* pp. 320-334, 2020.

[9] J. B. Awotunde, O. E. Matiluko and O. W. Fatai, "Medical diagnosis system using fuzzy logic.," *African Journal of Computing & ICT,* vol. 7, no. 2, pp. 99-106, 2014.

[10] T. O. Oladele, R. O. Ogundokun, J. B. Awotunde, M. O. Adebiyi and J. K. Adeniyi, "Diagmal: A Malaria Coactive Neuro-Fuzzy Expert System.," *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics),* pp. 428-441, 2020.

[11] J. B. Awotunde, F. E. Ayo, R. O. Ogundokun, S. O. Folorunso and A. O. Adekunle, "A decision support system for multi-target disease diagnosis: A bioinformatics approach.," *Heliyon,,* vol. 6, no. 3, 2020.

[12] J. Alzubi, A. Nayyar and S. Kumar, "Machine Learning from Theory to Algorithms: An Overview," in *Second National Conference on Computational Intelligence (NCCI 2018) . In Journal of physics: conference series* , 2018.

[13] M. Zivkovic, N. Bacanin, K. Venkatachalam, A. Nayyar, A. Djordjevic, I. Strumberger and F. Al-Turjman, "COVID-19 cases prediction by using hybrid machine learning and beetle antennae search approach.," *Sustainable Cities and Society, ,* Vols. 66, , no. 102669, 2020.

[14] J. Frankenfield, "Investopedia," 13 March 2020. [Online].

[15] J. B. Awotunde, F. E. Ayo, R. G. Jimoh, R. O. Ogundokun, O. E. Matiluko, I. D. Oladipo and M. Abdulraheem, "Prediction and classification of diabetes mellitus using genomic data.," *Intelligent IoT Systems in Personalized Health Care,* pp. 235-292, 2020.

[16] S. A. Ajagbe, I. R. Idowu, J. B. Oladosu and A. O. Adesina, "Accuracy of Machine Learning Models for Mortality Rate Prediction in a Crime Dataset," *International Journal of Information Processing and Communication (IJIPC),* vol. 10, no. 1 & 2, pp. 150-160, 2020.

[17] A. O. Babarunde, I. D. Oladipo, T. O. Aro and J. B. Awoniyi, "Enhanced Neuro-Fuzzy Inferential System for Diagnosis of Diabetes

Mellitus (DM)," *International Journal of Information Processing and Communication (IJIPC),* vol. 8, pp. 17-25, 2020.

[18] S. Lalmuanawma, J. Hussain and L. Chhakchhuak, "Applications of machine learning and artificial intelligence for Covid-19 (SARS-CoV-2) pandemic: A review," *Chaos, Solitons and Fractals,* vol. 139, pp. 1-7, 2020.

[19] H. Qiu, L. Yuan, Q. Wu, Y. Zhou, R. Zheng, X. Huang and Q. Yang, "Using the internet search data to investigate symptom characteristics of COVID-19: A big data study," *KeAi Chinese Roots Global Impact,* vol. xxx, no. xxx, pp. 1-9, 2020.

[20] Wuhan Municipal Health Commission., "Situation Report on Unexplained Viral Pneumonia EB/OL; 2020-01-11 [2020-2-24]," 2020.

[21] N. Zhu, D. Zhang, W. Wang, X. Li, B. Yang, J. Song, X. Zhao, B. Huang, W. Shi, R. Lu, P. Niu, F. Zhan, X. Ma, D. Wang, W. Xu, G. Wu, G. F. Gao and W. Tan, "A Novel Coronavirus from Patients with Pneumonia in China, 2019," Th e new england journal o f medicine, Massachusetts, 2020.

[22] Wikipedia, "Coronavirus disease 2019," November 2020. [Online]. Available: https://en.wikipedia.org/wiki/Coronavirus_disease_2019.

[23] Chinese Center for Disease Control and Prevention, "Pandemic Mapping of Coronavirus Disease 2019. EB/OL; 2020-03-5 [2020-3-5]," 2020. [Online]. Available: http://2019ncov.chinacdc.cn/2019-nCoV/.

[24] National Health Commission of the People’s Republic of China., "Diagnosis and Treatment of Pneumonia with New Coronavirus Infection (Trial Version 7). EB/OL; 2020-03-03 [2020-04-05]," 2020. [Online].

[25] C. Huang, Y. Wang, X. Li, L. Ren, J. Zhao, Y. Hu, L. Zhang, G. Fan, J. Xu, X. Gu, Z. Cheng, T. Yu, J. Xia, Y. Wei, W. Wu, X. Xie, W. Yin, H. Li, M. Liu, Y. Xiao, H. Gao, L. Guo, J. Xie and G. J. R. G. Z. J. Q. W. J. C. B. Wang, "Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China.," *Lancet.,* vol. 395 , no. 10223, pp. 497-506, 2020.

[26] N. S. Chen, M. Zhou, X. Dong, Q. Jieming, F. Gong, Y. Han, Y. Qiu, J. L. Y. Wang, Y. Wei, J. Xia, T. Yu, X. Zhang and L. Zhang, "Epidemiological and clinical characteristics of 99 cases of 2019 novel coronavirus pneumonia in Wuhan, China: a descriptive study.," *Lancet.,* vol. 395, pp. 507-513, 2020.

[27] World Health Organisation, "Q & A on Coronaviruses," 2020.

[28] K. Kumar, N. Kumar and R. Shah, "Role of IoT to avoid spreading of COVID-19," *International Journal of Intelligent Networks,* vol. 1, p. 32–35, 2020.

[29] N. Doremalen, T. Lambe, A. Spencer, S. Belij-Rammerstorfer, J. N. Purushotham, J. R. Port, V. Avanzato, T. Bushmaker, A. Flaxman, M. Ulaszewska, F. Feldmann, E. R. Allen, H. Sharpe, J. Schulz, M. Holbrook, A. Okumura and K. Meade-White, "ChAdOx1 nCoV-19 vaccination prevents SARS-CoV-2 pneumonia in rhesus macaques, bioRxiv," *The US Government Work,* pp. 1-23, 2020.

[30] World Health Organization , "Coronavirus Disease (COVID-19): Vaccines," 12 December 2020. [Online].

[31] W. G. Santos, "Natural history of COVID-19 and current knowledge on treatment therapeutic options," *Biomedicine & Pharmacotherapy ,* vol. 129, 2020.

[32] G. J. Milinovich, G. M. Williams, A. C. Clements and W. Hu, "Internet-based surveillance systems for monitoring emerging infectious diseases," *Lancet Infect Disease,* vol. 14, no. 2, pp. 160-168, 2014.

[33] M. Neyja, S. Mumtaz, K. M. S. Huq, S. A. Busari, J. Rodriguez and Z. Zhou, "An IoT-based e-health monitoring system using ECG signal.," in *GLOBECOM 2017-2017 IEEE Global Communications Conference* , Singapore , 2017.

[34] World Health Organisation, "Health Topic, Coronavirus disease overview," May 2020. [Online]. Available: www.who.int/health-topics/coronavirus#tab=tab _ 1.

[35] World Health Organization, "Contact tracing in the context of COVID-19," May 2020. [Online]. Available: www.who.int/publications-detail/contact-tracing-in-the- context- of- covid- 19.

[36] J. Pati, "Gene Expression Analysis for Early Lung Cancer Prediction Using Machine Learning Techniques: An Eco-Genomics Approach.," *IEEE Access,,* vol. 7, pp. 4232-4238, 2018.

[37] K. Chebib, "IoT applications in the fight against COVID-19," 25 September 2020. [Online]. Available: www.gsma.com/mobilefordevelopment/blog/iot-applications-in-the-fight-against-covid-19/.

[38] C. Li, X. Hu and L. Zhang, "The IoT-based heart disease monitoring system for pervasive healthcare service.," *Procedia computer science,,* vol. 112, pp. 2328-2334., 2017.

[39] W. Naudé, "Artificial intelligence vs COVID-19: limitations, constraints and pitfalls," *Artificial Intelligence & SOCIETY,* pp. 1-5, 2020.

[40] S. A. Ajagbe and O. A. Adesina, "Design and Development of an Access Control Based Electronic Medical Record (EMR)," *Centrepoint Journal (Science Edition),* vol. 26, no. 1, pp. 98-119, 2020.

[41] Johns hopkins university, "center for systems science and engineering website.," November 2020. [Online]. Available: https://coronavirus.jhu.edu/data/new-cases.

[42] M. H. D. M. Ribeiro, R. G. da-Silva, V. C. Mariani and L. D. S. Coelho, "Short-term forecasting COVID-19 cumulative confirmed cases: perspectives for Brazil," *Chaos, Solitons and Fractals,* pp. 1-18, 2020.

[43] R. Gupta, G. Pandey and S. K. Pal, "Machine Learning Models for Government to Predict COVID-19 Outbreak," *Digital Government: Research and Practice,* vol. 1, no. 4, pp. 1-6, 2020.

[44] A. Khakharia, V. Shah, S. Jain, J. Shah, A. Tiwari, P. Daphal, M. Warang and N. Mehendale, "Outbreak Prediction of COVID-19 for Dense and Populated Countries Using Machine Learning," *Annals of Data Science,* 2020.

[45] K. Abdulmajeed, M. Adeleke and L. Popoola, "ONLINE FORECASTING OF COVID-19 CASES IN NIGERIA USING LIMITED DATA.," *Data in Brief,* pp. 1-8, 2020.

[46] L. Yan, H.-T. Zhang, J. Goncalves, Y. Xiao, M. Wang, Y. Guo, C. Sun, X. Tang, L. Jing, M. Zhang, X. Huang, Y. Xiao, H. Cao, Y. Chen, T. Ren, F. Wang, Y. Xiao, S. Huang and H. N. J. B. C. C. Z. Y. L. A. M. L. J. J. C. Tan X, "An interpretable mortality prediction model for COVID-19 patients.," *Nation Mach Intell,* 2020.

[47] V. Siess, "On the calibration of a confinement SIR model with respect to COVID-19 mortality curve in Italy," pp. 1-8, 2020.

[48] S. F. Ardabili, A. Mosavi, P. Ghamisi, F. Ferdinand, A. R. Varkonyi-Koczy, U. Reuter, T. Rabczuk and P. M. Atkinson, "COVID-19 Outbreak Prediction with Machine Learning," pp. 1-39, 2020.

[49] V. K. R. Chimmula and L. Zhang, "Time series forecasting of COVID-19 transmission in Canada using LSTM Networks.," *Chaos, Solitons & Fractals,* 2020.

[50] T. Chakraborty and I. Ghosh, "Real-time forecasts and risk assessment of novel coronavirus (COVID-19) cases: a data-driven analysis," *Chaos, Solitons and Fractals,* 2020.

[51] C. Anastassopoulou, L. Russo, A. Tsakris and C. Siettos, "Data-based analysis, modelling and forecasting of the COVID-19 outbreak," pp. 1-21, 2020.

[52] S. Gul, M. Asif, S. Ahmad, M. Yasir, M. Majid and M. S. A. Malik, "A Survey on Role of Internet of Things in Education," *Unternational Journal of Computer Science and Network Security (IJCSNS(,* vol. 17, no. 5, pp. 157-165, 2017.

[53] F. Petropoulos and S. Makridakis, "Forecasting the novel coronavirus COVID-19," pp. 1-8, 2020.