

Corona Virus Outbreak: Analysis and Prediction

Dr. Sandeep Vanjale¹, Prof. Akash Surywanshi², Lakshay Goel³, Prashant Kumar⁴, Tavish Gautam⁵

Bharati Vidyapeeth deemed to be university college of engineering, Pune

Abstract: The Covid-19 pandemic which has affected more than 228 countries/territories infected more than million persons and consumed more than 548 million human lives. Since 2019 whole world had undergone several lockdowns. But that thing didn't work as expected and wrath of deadly virus still continued. The vaccination seems to be the robust solution for ending this situation. However vaccinating the whole world is still a challenging task. This paper basically analyzes how the Covid-19 virus spread all over the world and how it rapidly got increased in different countries/provinces and which country suffered the most with loss of its valuable asset as Human life. Based on the SVM and Regression algorithm, Death counts for all over the world is predicted for specific period and based on Machine learning techniques complete analysis is done. In the end comparison of Covid-19 is also done with similar epidemics on the basis of Confirmed Cases, Death Cases and Mortality Rate.

Keywords—Covid-19, Prediction, Dataset, Confirmed Cases, Death Cases, Recovered Cases, Epidemic, Python, Machine Learning

INTRODUCTION

Corona Virus disease (COVID-19) is a virus-borne infection caused by the SARS CoV-2 viral. Most patients infected with the virus experience a mild to severe respiratory infection and recover without the need for specific treatment. Some people, however, become extremely ill and require medical attention. People who are elderly or have underlying illnesses including cardiovascular disease, diabetes, chronic respiratory illness, or cancer are more susceptible to develop serious illnesses.

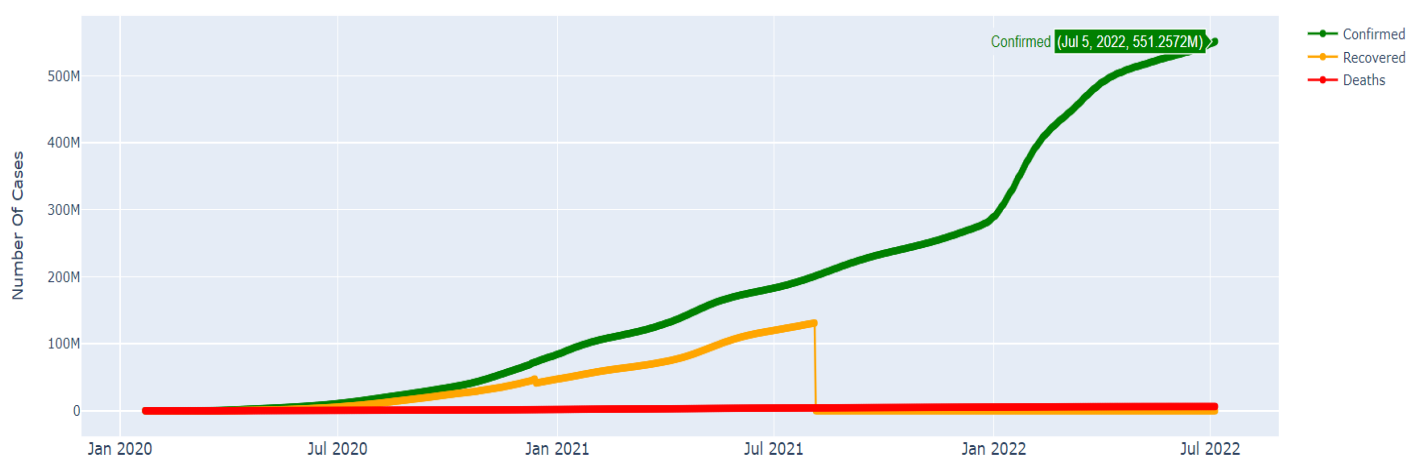
Healthcare systems and stakeholders should utilize the best predictive models to estimate case numbers and apply social distancing and blockade measures during a pandemic, according to the findings of a systematic literature review (SLR). It is necessary to take precautions. COVID-19 can infect anyone at any age and cause serious illness or death. The greatest method to avoid or delay infection is to learn as much as possible about the disease and how it spreads. Stay at least 1 meter away from people, wear a well fitted mask, and wash your hands or use an alcohol-based rub periodically to protect yourself and others from infection. When it's your turn, get vaccinated and follow local advice. When an infected person coughs, sneezes, speaks, sings, or breathes, the virus spreads in little liquid particles from their mouth or nose. Larger respiratory droplets to smaller aerosols are among the particles. Coughing on a bent elbow, for example, is a good way to develop breathing etiquette. If you're sick, stay at home and isolate yourself until you feel better.

CORONA SCENARIOS ALL OVER THE WORLD

According to WHO Weekly Epidemiological Update an overview of the global, regional and country-level COVID-19 estimates about total 551.2572 confirmed cases and total 6.342252 deaths till July 5 2022.

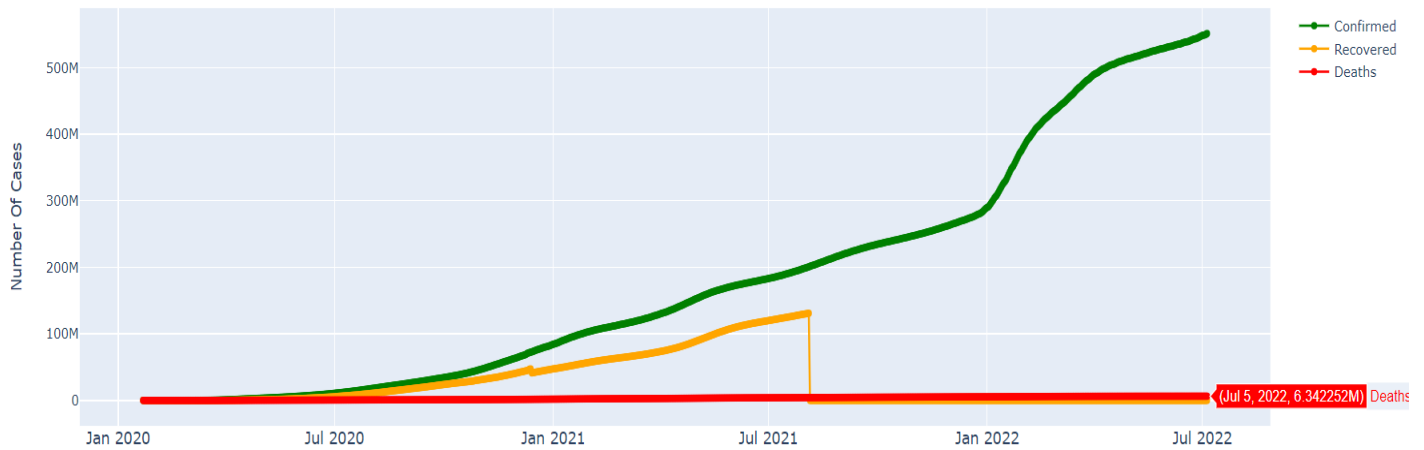
Confirmed Cases:-

Worldwide Covid-19 Cases



Death Cases:-

Worldwide Covid-19 Cases

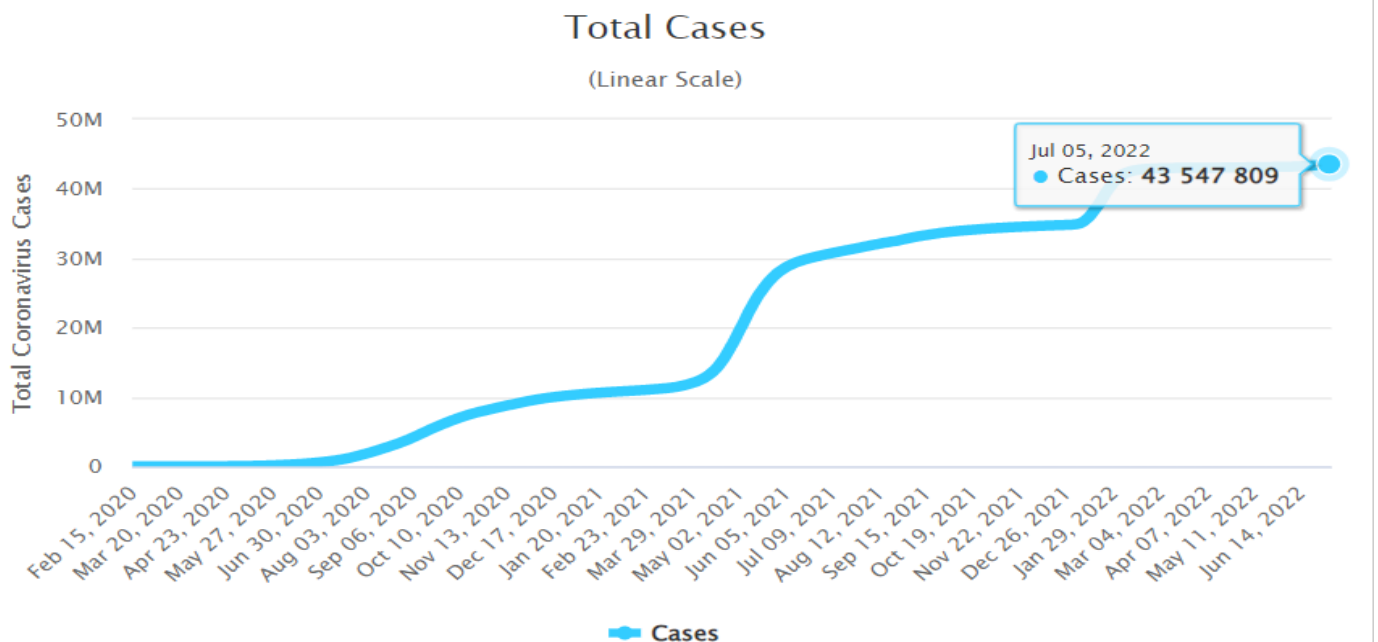


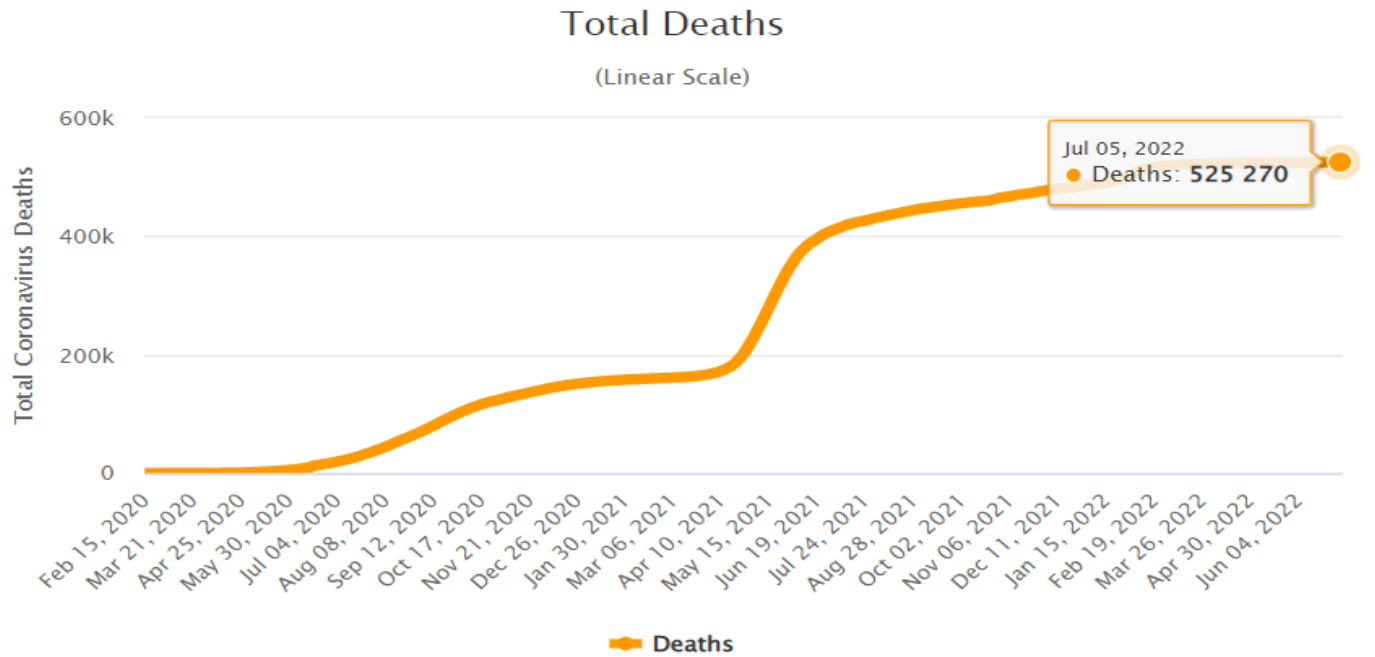
Entering the third year of the coronavirus pandemic, more than 551 million people have been infected and the virus has killed more than 6.3 million globally. Efforts many countries took to stamp out the pneumonia-like illness led to entire nations enforcing lockdowns, widespread halts of international travel, mass layoffs and battered financial markets. New variants of the virus have led to new waves of cases, though effective vaccines, new drugs and improved care are reducing the consequences for those who get sick.

CORONA SCENARIOS IN INDIA

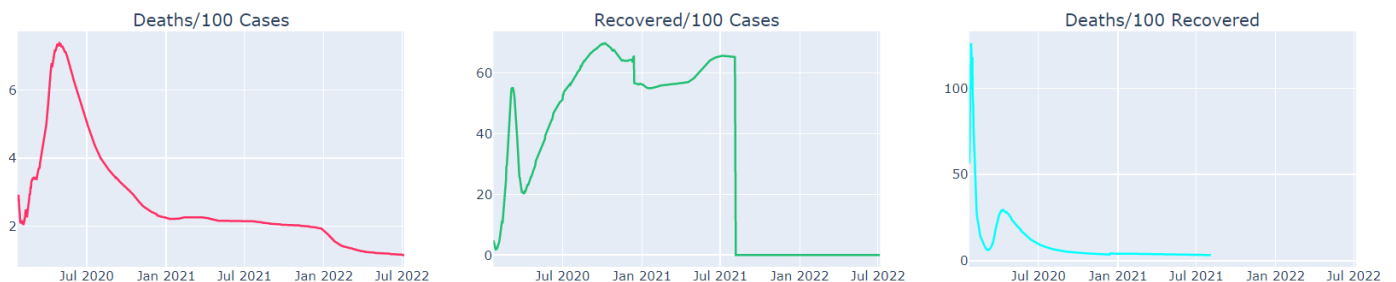
India has recorded 16,159 new coronavirus cases in the last 24 hours, taking the total number of COVID-19 cases to 4,35,47,809. India's COVID-19 death count rose to 5,25,270 with 28 new coronavirus deaths in the last 24 hours(July 5 2022). According to Union Health Ministry data, the active cases stand at 1,15,212, comprising 0.26 per cent of the total infections, while the recovery rate was recorded at 98.53 per cent, the ministry said.

Confirmed Cases:-

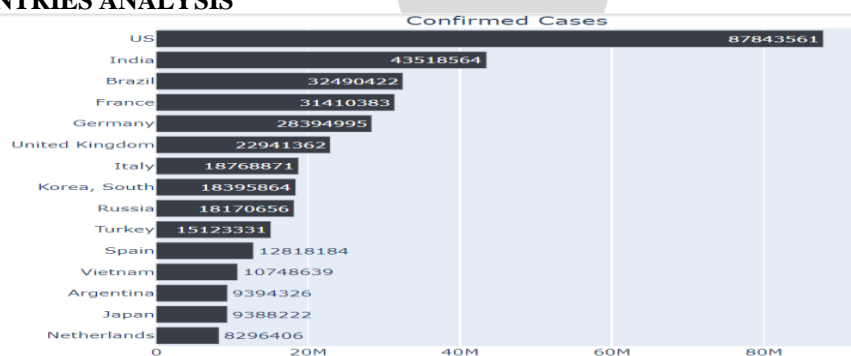


Death Cases:-**ANALYSIS OF CORONA SCENARIOS ON PER 100 FACTOR BASIS**

- Deaths/100 Cases:-** It is defined the total no. of Death cases estimated on per 100 positive cases observed.
- Recovered/100Cases:-** It is defined the total no. of Recovered cases estimated on per 100 positive cases observed.
- Deaths/100 Recovered:-** It is defined the total no. of Death cases estimated on per 100 Recovered cases observed.



As per the Deaths/100 cases graph it is concluded that before 5 Feb 2020 the graph increases linearly representing that Deaths keeps on increasing on per 100 cases. But after 29 April 2020 it linearly decreases not because of the reduction in death cases but because of drastically increase in positive cases. Similarly for Recovered/100 Cases graph it is observed that till Jan 2020 to August 2020 graph increases in unordered fashion but after August 2020 the graph becomes constant to 0 representing that as the positive cases increases the recovery factor decreases rapidly.

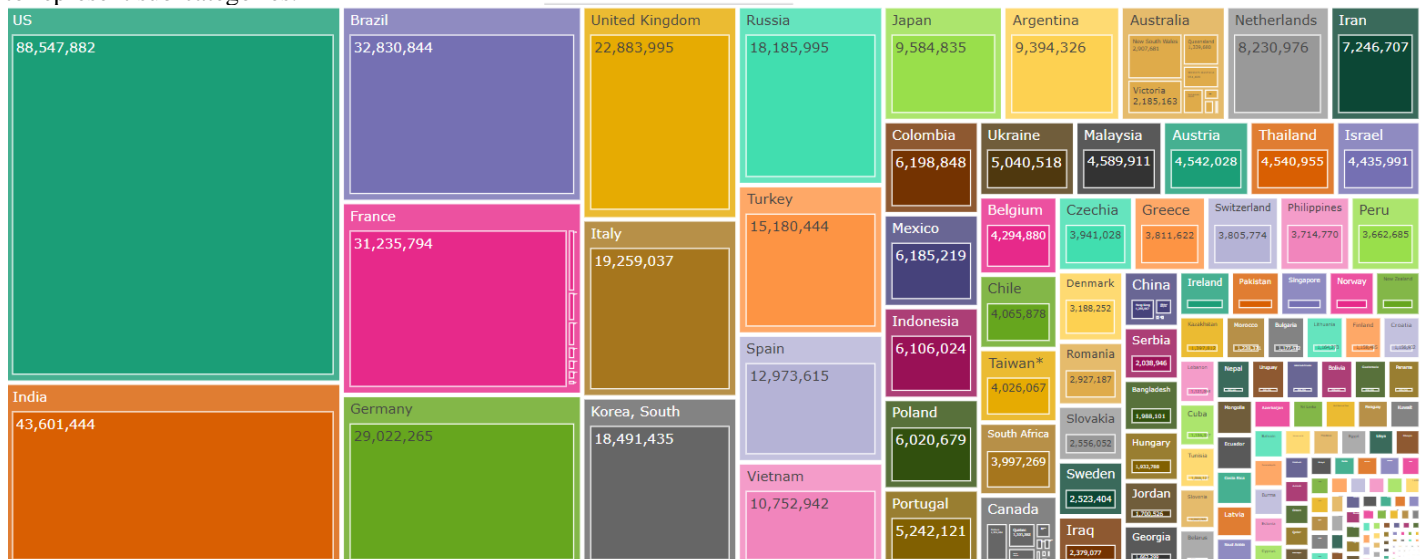
TOP 15 COUNTRIES ANALYSIS



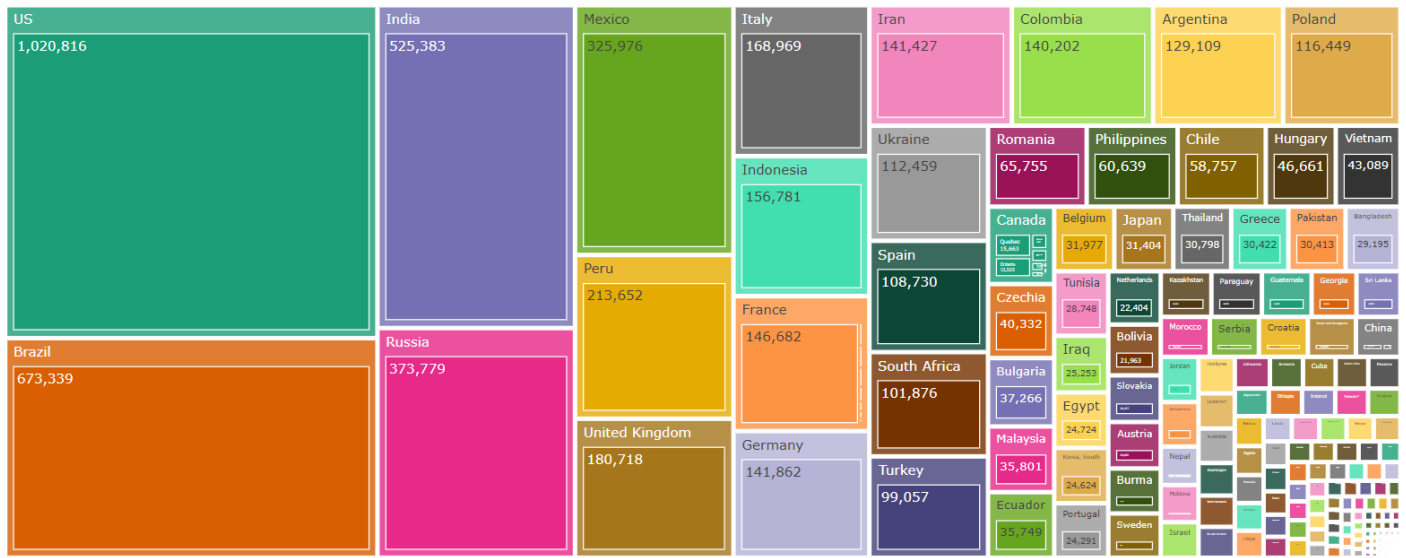
The identified Top 15 countries with overall cases based on confirmed, deaths, recoveries, deaths/100 cases, new cases, cases/100 million, 1 week change and 1 week % increase from Covid-19 were compared. The United States, India, Brazil, France, and Germany saw a relatively fast spread of the disease. There was a fast recovery ratio in United States, India, Brazil, France, and Germany and slow recovery ratio in the Spain, Vietnam, Japan, Argentina and Netherlands. There was a high death rate ratio in The United States, Brazil, India, Russia and Mexico. The rapidly increase in New Cases is observed in Italy, Taiwan, Mexico and Australia. However if we observe the Cases/Million people then Denmark, Portugal, Israel, Slovenia and Austria topped the list. If we observe the 1 week change in corona cases it is found that The United States, France, Germany, Italy and Brazil holds the first top five position all over the world.

TREE-MAP ANALYSIS OF CONFIRMED AND DEATH CASES

Tree Map analysis provides you the effective way of depicting the information. Basically, it is used to show the composition of a whole when several components are present. In this type of arrangement a rectangular area is divided into sub smaller rectangles to represent sub-categories.



Confirmed Cases



Death Cases

Above Tree-Map analysis represents the confirmed and death cases all over the in different rectangular components. It is observed that United States is having most of the confirmed and death cases all around the Globe. But if we compared the aspect of Confirmed and Death cases between India and Brazil it is observed that India is having more confirmed cases than Brazil and Brazil is having more Death cases than India.

EXPERIMENTAL DATA ANALYSIS

We need a model that predicts how the virus could spread across different countries and regions. The goal of this task is to build a model that predicts the spread of the virus in next 10 days. For this we are considering the data 10 days after 23 March 2020. For this there are certain prerequisite like collecting the databases of Confirmed, Death and Recovered cases from 22/01/2020 to 15/03/2020.

Database imported:- 1. time_series_covid-19_confirmed, 2. time_series_covid-19_deaths, and 3. time_series_covid-19_recovered.

For analysis part Support Vector Machines & Linear Regression models are prepared and based on the actual data we compared the predicted value and estimate the efficiency of our built model.

Support Vector Machines:- Support vector machines are a set of supervised learning methods used for classification, regression, and outlier's detection. All of these are common tasks in machine learning. You can use them to detect cancerous cells based on millions of images or you can use them to predict future driving routes with a well-fitted regression model. There are specific types of SVMs you can use for particular machine learning problems, like support vector regression (SVR) which is an extension of support vector classification (SVC). The main thing to keep in mind here is that these are just math equations tuned to give you the most accurate answer possible as quickly as possible SVM's are different from other classification algorithms because of the way they choose the decision boundary that maximizes the distance from the nearest data points of all the classes. The decision boundary created by SVMs is called the maximum margin classifier or the maximum margin hyper plane.

SVM uses different parameters to build a model like kernel, c, gamma, epsilon, shrinking, svm_grid.

- Kernel specifies the kernel type to be used in algorithm. It must be linear, poly, sigmoid, rbf, precomputed or callable. If nothing is given rbf is used.
- c is a regularization parameter.
- Gamma is kernel coefficient of poly, rbf and sigmoid.
- Epsilon specifies the epsilon tube within which no penalties are associated in training loss function.
- Shrinking takes boolean values as true or false.
- svm_grid has all the values passed as parameters to it.

```
# Building the SVM model

kernel = ['poly', 'sigmoid', 'rbf']
c = [0.01, 0.1, 1, 10]
gamma = [0.01, 0.1, 1]
epsilon = [0.01, 0.1, 1]
shrinking = [True, False]
svm_grid = {'kernel': kernel, 'C': c, 'gamma': gamma, 'epsilon': epsilon, 'shrinking': shrinking}

svm = SVR()
svm_search = RandomizedSearchCV(svm, svm_grid, scoring='neg_mean_squared_error', cv=3, return_train_score=True, n_jobs=-1, n_iter=40, verbose=1)
svm_search.fit(X_train_confirmed, y_train_confirmed)
```

Using the function `svm_search.best_params_` as `{'C': 10, 'epsilon': 0.1, 'gamma': 0.01, 'kernel': 'poly', 'shrinking': True}` we are providing the best parameter to our model to predict the data.

With the help of `randomizdsearchCV` function we are building a model by passing necessary parameters and fitting the SVM search fit function by passing `X_train confirmed` and `Y_train confirmed`.

Using the `predict(future_forecast)` function allow the function to predict the model.

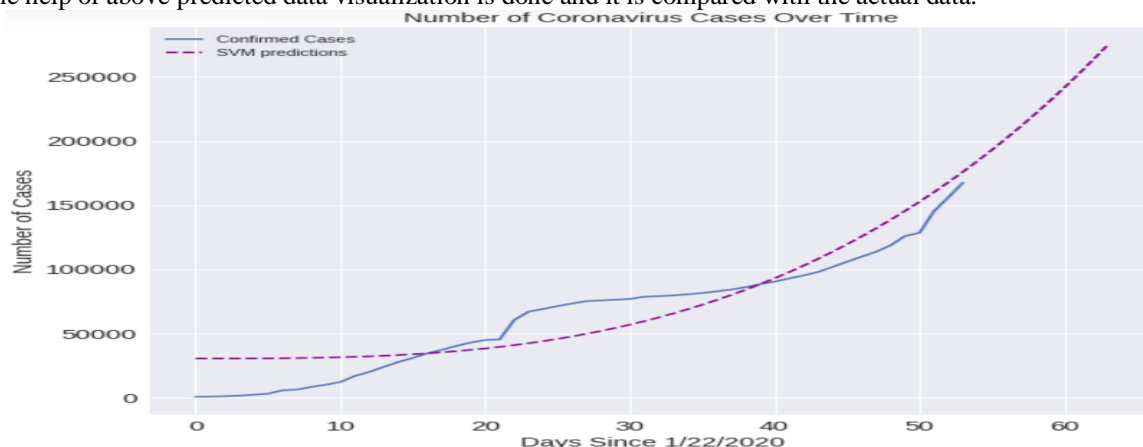
```
svm_confirmed = svm_search.best_estimator_
svm_pred = svm_confirmed.predict(future_forecast)
```

```
svm_confirmed
```

```
svm_pred
```

```
array([[ 30388.34440129,  30389.32185622,  30396.1640407 ,  30414.73568429,
         30450.90151654,  30510.52626702,  30599.47466527,  30723.61144085,
         30888.80132331,  31100.90904222,  31365.79932711,  31689.33690756,
         32077.38651311,  32535.81287332,  33070.48071775,  33687.25477594,
         34391.99977746,  35190.58045185,  36088.86152868,  37092.7077375 ,
         38207.98380786,  39440.55446933,  40796.28445144,  42281.03848376,
         43900.68129585,  45661.07761726,  47568.09217753,  49627.58970624,
         51845.43493293,  54227.49258716,  56779.62739848,  59507.70409644,
         62417.58741062,  65515.14207054,  68806.23280579,  72296.7243459 ,
         75992.48142043,  79899.36875894,  84023.25109098,  88369.99314611,
         92945.45965388,  97755.51534385, 102806.02494557, 108102.8531886 ,
        113651.86480248, 119458.92451679, 125529.89706107, 131870.64716488,
        138487.03955776, 145384.93896929, 152570.210129 , 160048.71776646,
        167826.32661123, 175908.90139284, 184302.30684088, 193012.40768487,
        202045.06865439, 211406.15447898, 221101.52988821, 231137.05961162,
        241518.60837877, 252252.04091922, 263343.22196252, 274798.01623822])
```

With the help of above predicted data visualization is done and it is compared with the actual data.

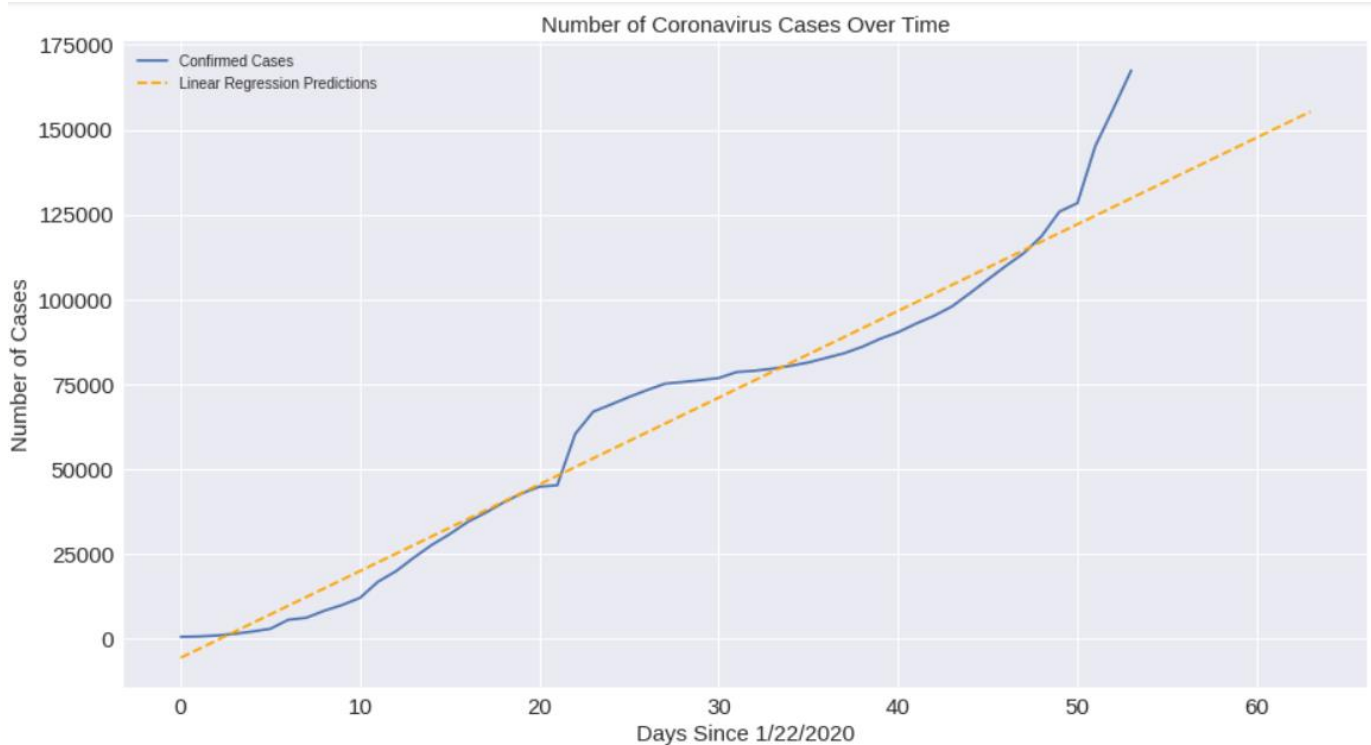


Linear Regression Model:- Linear regression is one of the easiest and most popular Machine Learning algorithms. It is a statistical method that is used for predictive analysis. Linear regression makes predictions for continuous/real or numeric variables such as sales, salary, age, product price, etc. Linear regression algorithm shows a linear relationship between a dependent (y) and one or more independent (x) variables, hence called as linear regression. Since linear regression shows the linear relationship, which means it finds how the value of the dependent variable is changing according to the value of the independent variable.

Using the Linear model with the help of predict function as future_forecast we will allow the model to predict the data.

Linear regression future predictions:

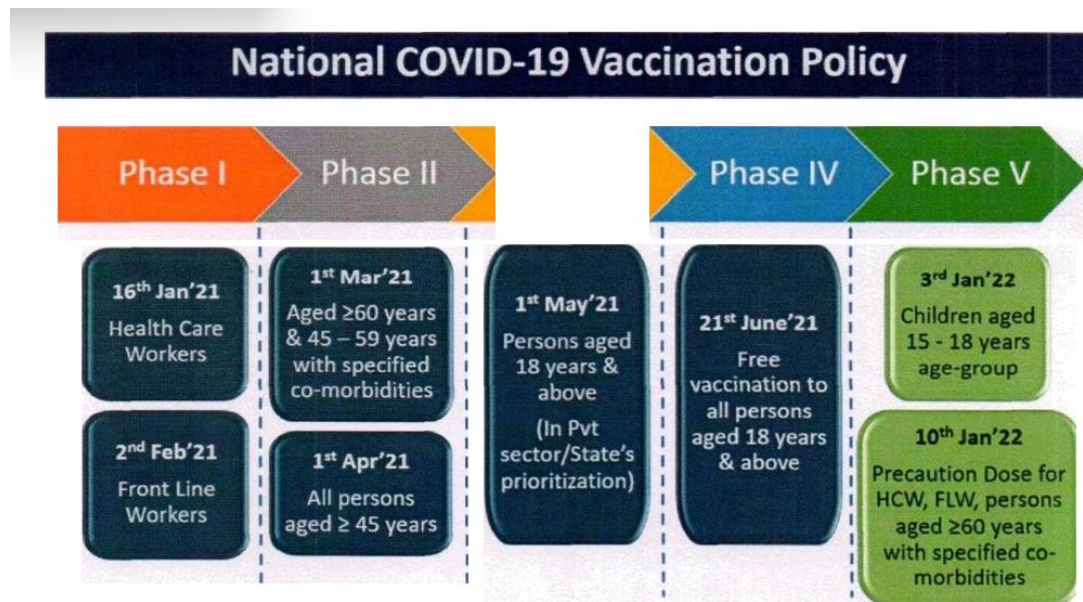
```
[[132336.25252525]
 [134890.72222222]
 [137445.19191919]
 [139999.66161616]
 [142554.13131313]
 [145108.6010101 ]
 [147663.07070707]
 [150217.54040404]
 [152772.01010101]
 [155326.47979798]]
```



COVID VACCINE STATS

As at 22 May 2022, almost one billion people in lower-income countries remain unvaccinated. Only 57 countries have vaccinated 70% of their population – almost all of them high-income countries. It is necessary that all countries to reach 70% vaccination coverage as soon as possible, including 100% of those aged over 60; 100% of health workers; and 100% of those with underlying conditions. However as of July 2 2022 a total of 12,037,259,035 vaccine doses have been administered yet.

India's National COVID Vaccination Program is built on scientific and epidemiological evidence, WHO guidelines and global best practices. The program started with the vaccination of health care workers (HCWs) and frontline workers (FLWs). The program ambit expanded in a phased manner, with the inclusion of all individuals aged 18 years and above from 1 May 2021. The program coverage was further expanded to children 15 to 18 years of age from 3 January 2022 and administration of precaution dose to HCWs, FLWs and people aged 60+ years with co morbidities based on medical advice from 10 January 2022.



COMPARISON OF COVID-19 WITH SIMILAR EPIDEMICS

SARS-: Severe acute respiratory syndrome (SARS) is a viral respiratory disease caused by a SARS-associated coronavirus. It was first identified at the end of February 2003 during an outbreak that emerged in China and spread to 4 other countries. SARS is an airborne virus and can spread through small droplets of saliva in a similar way to the cold and influenza. It was the first severe and readily transmissible new disease to emerge in the 21st century and showed a clear capacity to spread along the routes of international air travel. SARS can also be spread indirectly via surfaces that have been touched by someone who is infected with the virus. Most patients identified with SARS were previously healthy adults aged 25–70 years. A few suspected cases of SARS have been reported among children under 15 years. The case fatality among persons with illness meeting the current WHO case definition for probable and suspected cases of SARS is around 3%.

EBOLA-: Ebola virus disease (EVD), formerly known as Ebola haemorrhagic fever, is a severe, often fatal illness affecting humans and other primates. The virus is transmitted to people from wild animals (such as fruit bats, porcupines and non-human primates) and then spreads in the human population through direct contact with the blood, secretions, organs or other bodily fluids of infected people, and with surfaces and materials (e.g. bedding, clothing) contaminated with these fluids. The average EVD case fatality rate is around 50%. Case fatality rates have varied from 25% to 90% in past outbreaks. The first EVD outbreaks occurred in remote villages in Central Africa, near tropical rainforests. The 2014–2016 outbreak in West Africa was the largest and most complex Ebola outbreak since the virus was first discovered in 1976. There were more cases and deaths in this outbreak than all others combined. It also spread between countries, starting in Guinea then moving across land borders to Sierra Leone and Liberia. It is thought that fruit bats of the Pteropodidae family are natural Ebola virus hosts.

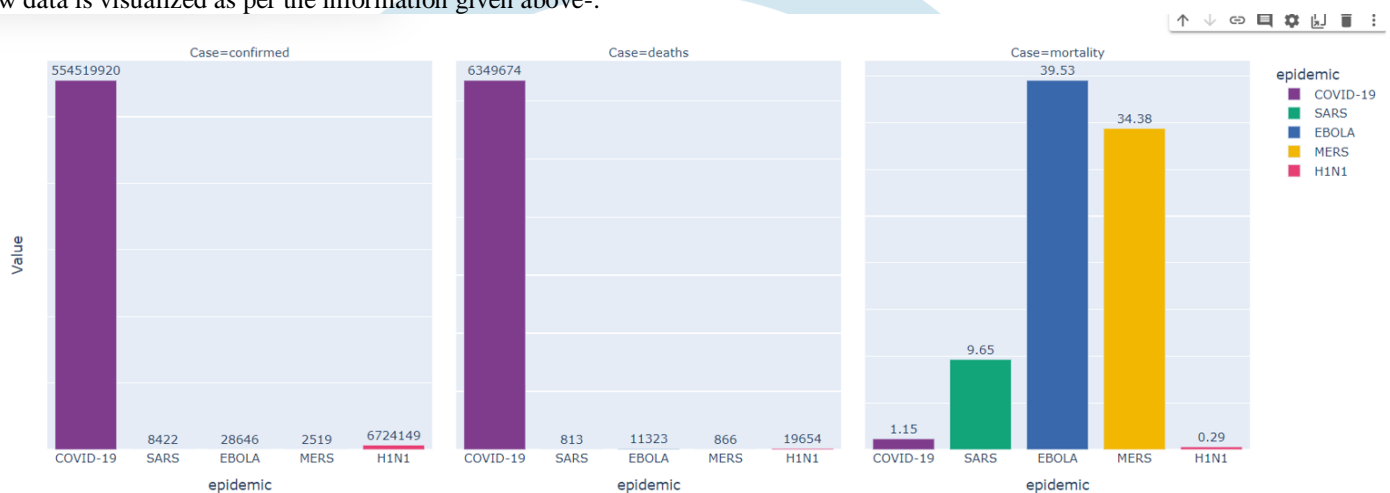
MERS-: Middle East respiratory syndrome coronavirus (MERS-CoV) is a virus transferred to humans from infected dromedary camels. It is a zoonotic virus, meaning it is transmitted between animals and people, and it is contractible through direct or indirect contact with infected animals. MERS-CoV has been identified in dromedaries in several countries in the Middle East, Africa and South Asia. In total, 27 countries have reported cases since 2012, leading to 858 known deaths due to the infection and related complications. The origins of the virus are not fully understood but according to the analysis of different virus genomes it is believed that it may have originated in bats and later transmitted to camels at some point in the distant past. Human-to-human transmission is possible, but only a few such transmissions have been found among family members living in the same household. In health care settings, however, human-to-human transmission appears to be more frequent.

H1N1-: Before the H1N1 pandemic in 2009, the influenza A(H1N1) virus had never been identified as a cause of infections in people. Genetic analyses of this virus have shown that it originated from animal influenza viruses and is unrelated to the human seasonal H1N1 viruses that have been in general circulation among people since 1977. After early reports of influenza outbreaks in North America in April 2009, the new influenza virus spread rapidly around the world. By the time WHO declared a pandemic in June 2009, a total of 74 countries and territories had reported laboratory confirmed infections. Unlike typical seasonal flu patterns, the new virus caused high levels of summer infections in the northern hemisphere, and then even higher levels of activity during cooler months. The new virus also led to patterns of death and illness not normally seen in influenza infections. The H1N1 (2009) virus continues to circulate as a seasonal virus and is included in the vaccines against seasonal influenza.

With the help of provided information from various datasets comparison is made between similar epidemics.

	epidemic	start_year	end_year	confirmed	deaths	mortality
0	COVID-19	2019	2022	554519920	6349674	1.15
1	SARS	2002	2004	8422	813	9.65
2	EBOLA	2013	2016	28646	11323	39.53
3	MERS	2012	2020	2519	866	34.38
4	H1N1	2009	2010	6724149	19654	0.29

Now data is visualized as per the information given above:-



- As per the data it is concluded that confirmed cases for COVID-19 is much greater than any other epidemics depicting that Covid-19 virus is more lethal than any other virus. More than 551.2572 Million cases are observed. However, confirmed cases for MERS (Middle East Respiratory Syndrome Coronavirus) are least as compared to others.
- If we compare death cases of COVID-19 with other epidemics it is observed that most of the death cases are reported due to COVID-19 depicting coronavirus is a world-widespread pandemic. Death cases for COVID-19 is most while for SARS (Severe Acute Respiratory Syndrome) is least which shows that its virus is less lethal than any other.
- Mortality rate for Ebola virus is higher than any epidemic which shows that out of observed confirmed cases max. people faced deaths.

CONCLUSION AND FUTURE SCOPE

In this research paper, we have built a model using Support Vector Machines (SVM) and Linear Regression Model. With help of past inputs, training of the model is done to best predict the result. The parameters provided help the model to build the computational algorithm which in turn helps to predict the best possible outcome.

ACKNOWLEDGEMENTS

We would like to extend our gratitude and sincere thanks to all those who helped in timely completion of this research paper. It has been a great honor and privilege to work along with our fellow project mates and a special thanks to our project guide Mr. Akash Suryawanshi, and head of the department Dr. Sandeep Vanjale without whose guidance, stimulating suggestions and constant encouragement, this project would not have turned out to be this successful.

We would also like to acknowledge the crucial role of the World Health Organization (WHO) and John Hopkins University who provided us with all the required data that helped in the stimulation of our project. At last, we would express our special thanks to our parents for cooperating and helping us with our project.

REFERENCES

- [1] Ching-Sung(2019) , Journal of the Chinese Medical Association, The Outbreak of Covid-19: Overview
- [2] Jeffrey Chu(March 25 2021), Journal Pone, A statistical analysis of the novel coronavirus(COVID-19) in Italy and Spain
- [3] Lucia Russo, Athanasios Tsakris,Journal Pone, Data based analysis, modeling and forecasting of the COVID-19 outbreak.
- [4] Amir Darki(2021) , Research article, Predictive Role of Blood Cellular Indices and Their Relationship with Endogenous as Determinants of Inflammatory Biomarkers in Pulmonary Embolism
- [5] Alan Meot, Patrick Bonin(June 23 2022), Find In PubMed, Covid-19 and Memory: A Novel Contamination Effect in Memory.
- [6] Wendy D. Manning(June 20 2022), Research Article, Covid-19 stress and sexual Identities
- [7] Lynda Lattke(May 1 2022), Scandinavian Journal of Public Health, Health Promoting Schools and Covid-19: preparing for the future
- [8] Eliuti Gahetau(June15 2022), Asia Pacific Journal of Public Health, A short Covid-19 Outbreak in an isolated pacific Island
- [9] Nadia Hassan Ali Awad (June 20 2022), Nursing ethics, Crisis, ethical leadership and moral courage
- [10] Dongbu Wo, wei Wang(May 26 2022), Predicting progression to severe Covid-19 using the Paint Score
- [11] Robert Sykes, Catherine Bagot(May 23 2022), Nature medicine, A multisystem, cardio-renal investigation of post- COVID-19 illness.
- [12] Ziyad-Al-Aly, Yan Xie(May 25 2022), Nature Medicine, Long Covid after breakthrough SARS Cov-2 infection
- [13] Ph.D Ori Magen(My 29 2022), The New England Journal of Medicine, Vaccine Effectiveness against Omicron in children from 5 – 11 years.
- [14] World Health Organization. COVID-19 disease in children and adolescents: scientific brief. September 29, 2021
- [15] Siegel DA, Reses HE, Cool AJ, et al. Trends in COVID-19 cases, emergency department visits, and hospital admissions among children and adolescents aged 0-17 years — United States, August 2020–August 2021. MMWR Morb Mortal Wkly Rep 2021;70:1249-1254.
- [16] UK Health Security Agency. COVID-19 vaccine surveillance report week 5. February 3, 2022
- [17] Andrews N, Stowe J, Kirsebom F, et al. Effectiveness of COVID-19 vaccines against the Omicron (B.1.1.529) variant of concern. December 14, 2021
- [18] Reis BY, Barda N, Leshchinsky M, et al. Effectiveness of BNT162b2 vaccine against Delta variant in adolescents. N Engl J Med 2021;385:2101-2103.

