Covid-19 outbreak prediction using Machine Learning.

Dr Anil Kumar Sagar *1, Disha Bansal*2, Vaishali Chhonkar*3 Khalilullah Ahmad Zai*4 Sayed Fayaz Sadat*5
*1Professor, Department of computer science& Engineering, Sharda University, Gr. Noida, U.P, India.
. *2 Student, Department of Computer Science & Engineering, Sharda University, Gr. Noida, U.P., India.
*3Student, Department of computer science& Engineering, Sharda University, Gr. Noida, U.P, India.
*4 Student, Department of Computer Science & Engineering, Sharda University, Gr. Noida, U.P., India.
*5Student, Department of computer science& Engineering, Sharda University, Gr. Noida, U.P, India.

Abstract

Covid19 is the worst highest infectious disease faced by the world. Covid19's first case was confirmed in Wuhan China in December 2019 and then open out all around the world. According to the history after every 100 years, the deadliest disease comes but covid is the highest death rate recorded disease till now there are diseases which came in1800'sthird chorale in 1900's sixth cholera, Hong kong flu, Asian flu in 2000's aids/HIV and now it is covid19. Covid19 has trigged panic across the world a large-scale quarantine lockdown has changed the life of the people. Due to covid, many people lost their lives and many people lost their employment covid leads to many factors including employment, political issues, cancellation of many exams etc which has destroyed the lives of people the people who earn daily for their living covid has killed them. Lockdown plays a very important role in getting rid of covid 19 but the lockdown has two effects positive and negative, negative is due to lockdown many people lost their jobs and faced economical and financial issues. Boards exams are cancelled first time due to which many students' dreams suffered. So here in this prediction, we are using machine learning for predicting the future forecast of this deathliest disease covid19.

Keywords

Covid19 Outbreak, covid disease, Machine learning, Covid prediction, Coronavirus, Covid outbreak

Introduction

Covid 19 has been the world's greatest severe health hazard since World War II. Keep the distribution cycle in mind while projecting a Covid 19 outbreak, as it can help you make critical steps to reduce the prevalence rate. We can effectively allocate financial resources to areas where public health officials are most required

(WHO). The proposed model is called "Covid 19 outbreak Prediction Using a Reading Machine." Two strategies are utilised to estimate line rotation and support vector mechanism accuracy in this supervised learning. In order to forecast outbreaks early, we evaluate the data and make estimates for the future. To anticipate future outcomes, we used data from Git hub and machine learning algorithms such as linear regression and support vector machine we analysed the data and make predictions for the future. here we have taken data from the git hub and used machine learning algorithms such as linear regression and vector support for future predictions that help us prepare for the next deadly disease that gives us an idea of the situation. Covid is the world's most serious global health concern. Countries make every effort to treat and test their citizens. Following the lockdown, several countries are experiencing economic hardship, financial unemployment, and political unrest. Many exams have been cancelled or postponed as a result of the lockdown, and social meeting places have been closed as a result of the lockdown.

Covid outbreak prediction analysis aids in predicting this pandemic and recognising the spread cycle, which may aid in requiring key measures to reduce the spread rate and ensuring that public health officials may allocate economic resources efficiently to areas of highest need. WHO Covid outbreak prediction using machine learning is the proposed model, and it employs algorithms such as linear Regression. support vector machine classifier SVM. This aim is to predict the out Break before it's observed that the most effective performance is of the XG boost classifier. however, accuracy may be improved using hyperparameter turning.

Literature Survey

Machine learning may be used to anticipate covid outbreaks in a variety of ways. Many forecasts have been made in the past to predict covid outbreaks using machine learning. [1] Machine-readable monitoring model electronic learning in machine learning was utilised in the last several predictions. SIER estimates employing intelligence, CT scan and chest X-ray studies, and AI image is also re-utilised in forecasts generated on the covid outbreak. Covid 19 is a global disease that is quickly expanding. Because there are no symptoms, it is extremely difficult to detect. estimating the number of patients that may be affected in the future By researching theory: This study used Covid-19 to forecast an infectious illness outbreak. [2] This project forecasts what will happen in the future. Machine learning supervised algorithm linear regression and support vector machine were used for 30 days (SVM). We used real-time data from GitHub to test both of these algorithms. By evaluating several graphs, it was shown that the linear regression technique provides the best prediction rate when compared to the support vector machine. Future forecasts for Covid-19 are based on machine-readable monitoring models. Machine learning (ML) contributes to the outcomes of working to improve decision-making by predicting the future course of action. ML models are used to predict the number of future patients who will be afflicted by COVID-19. [3] COVID-19 threats have been predicted using standard prediction models such as the total reduction and selection operator (LASSO), support vector support (SVM), linear regression (LR), and big display smoothing (ES). The fact that ES is doing so well in the

current field of prediction stems from study findings that show that, given the type and quantity of the database, it is capable of doing so. LR and LASSO also perform a good job of predicting death rates and confirming cases to some extent. The Covid-19 Case: Examining Rehabilitation Using Reading Theory This study looked at how to predict how long it would take to recover from an infectious disease outbreak. The error learning method is used here, with objection measures such as treatment, segregation, social isolation, and so on, specifically to harmonise control of virus spreads by reducing infection levels if this is effective, and the rate of infection, in the background to reach the top, down to follow what is known as the Universal Recovery Curve. [4] The threat of When the rate of infection is slow, infectious diseases are infrequent, a phenomenon known as 'loosening of the curve.' When the rate rises, it should fall as a result of effective countermeasures. [5] Data mining and data analysis are research fields in which data is mined and analysed. Data In the First Covid-19 Epidemic, Death, Prevention, and Drug Development Data mining of scientific literature records from the Core Science [6]Web Collection was utilised to obtain Covid-19 death facts, immunizations, and vaccines. The analysis compares records from throughout the Covid-19 study topics, with individual records being evaluated separately.

Methodology

We are using machine learning algorithms to predict the outbreak of covid19 using supervised classification algorithms linear regression as polynomial and support vector machine (SVM) as a classifier. we have created a virtual environment in the anaconda Jupiter notebook we are using python in the 3.8 version for coding [7]It is also used for classification and regression analysis. SVM uses labelled data. Regression analysis is a type of predictive modelling technique that looks at how a dependent and independent variable are related.
Regression analysis is used to determine the strength of predictors, forecast an effect, and forecast trends.
Types of Regression analysis:

- Linear regression.
- Logistic regression.

Consider a straight equation line that combines any two variables X and Y and may be written as:

LINEAR REGRESSION

The linear regression algorithm is a sort of regression analysis that involves establishing a link between the dependent and independent variables. It has a linear structure.

$$Y = aX + b$$

based on the input variables in a linear combination (x). The dependent variable is the one that wishes to be predicted.

The independent variable is the one we're utilising to predict another variable.

Linear regression is a technique for predicting a variable's value based on the value of another variable. This model is created with the use of a straight line [8]and a continuous variable. It is determined by the amount of money lost (R).

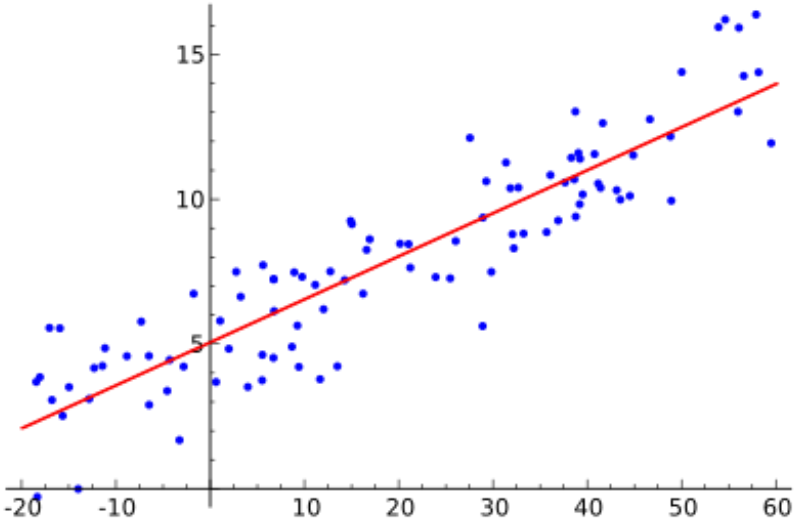

Figure :1

This example employs simple linear regression, where the space between the red line and each sample point is lowered by the square of the distance.

LOGISTIC REGRESSION

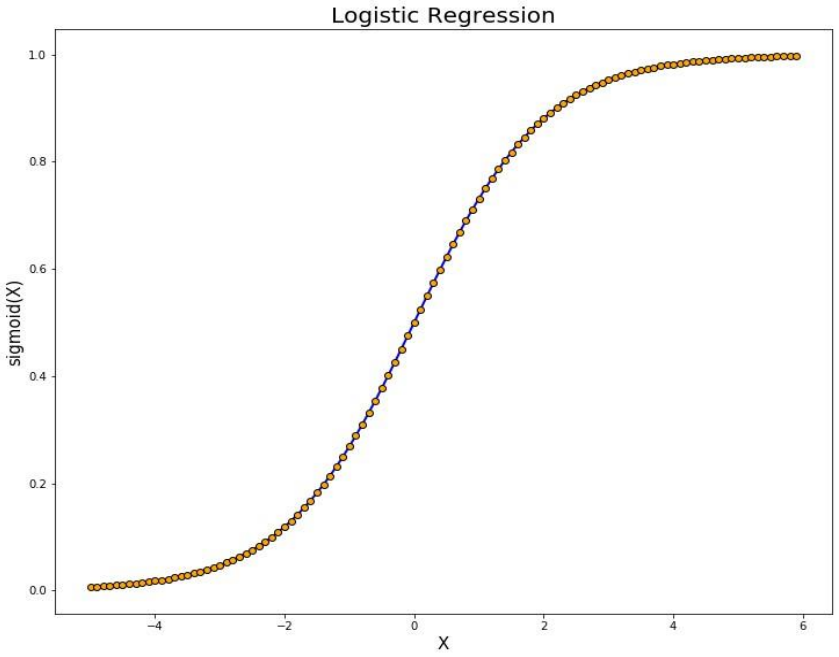

Figure :2

We don't draw a straight line through our data in linear regression. As a result, data was fitted to a sigmoid curve, an S-shaped bend line. Pseudo RR2 Equation

In statistics and machine learning, linear regression is one of the most well-known and well-understood techniques. A linear model is one in which the input variables (x) and the single output variable (y) are assumed to have a linear relationship (y). It is possible to calculate y in more detail.

When only one input variable (x) is used, Simple linear regression is the name of the method. When there are several input variables, the procedure is typically referred to as multiple linear regression in the statistics literature.

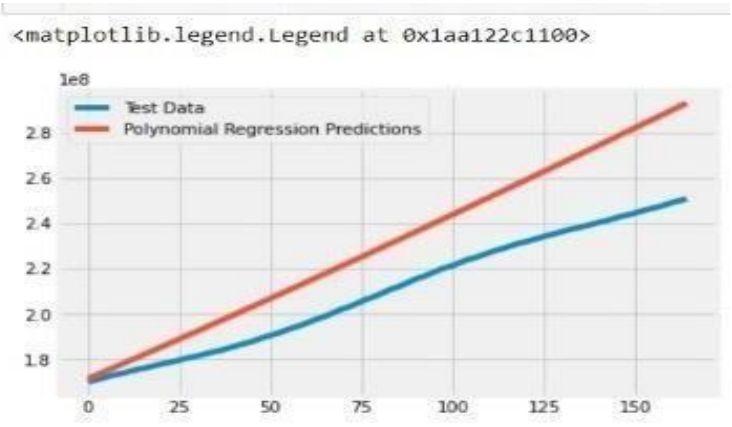

Figure:3

In this graph on the x-axis and y-axis, the increase of confirmed cases in the world is shown in comparison with the tested data and polynomial regression. [9]The best fit of polynomial regression prediction in shown.In this, the blue lines show the test data variation and red lines shows the polynomial regression variation.

SUPPORT VECTOR MACHINE

SVM comes in supervised classification.

A Support vector machine is a supervised machine learning model that employs a classification method to solve problems involving two groups. It's utilised for regression analysis and categorization. SVM models can categorise new text after being given sets of labelled training data for each category.

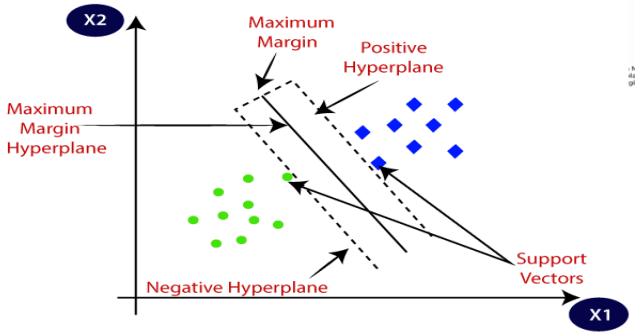

Figure:4

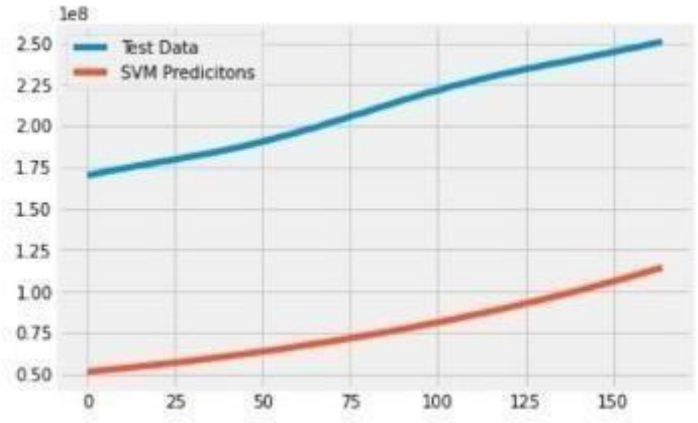

Figure:5

In this graph on x-axis and y-axis the increase of confirmed cases in world is shown in compare with the tested data and Support Vector Machine (SVM) . [10]The best fit of SVM prediction is shown. In this the blue lines shows the test data variation and red lines shows the SVM variation .

Result Analysis

This project “covid19 outbreak prediction using machine learning "is a prediction project. After the prediction and the final testing has been done using ML supervised algorithm which is linear regression and SVM. And after doing the testing with live data which we have taken from GitHub and than compared both the Algorithm with live data we found that linear regression algorithm is best for prediction covid19 outbreak in this research.

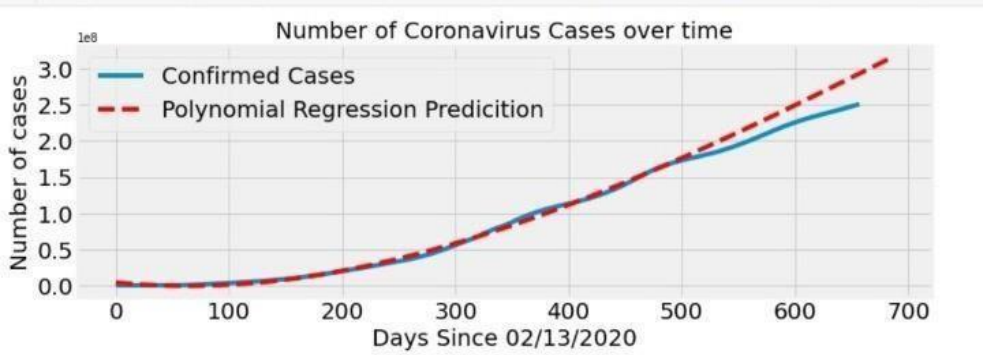

Figure:1 Plot in compare to confirmed cases polynomial regression.

Description- In this graph on x axis and y axis the increase of confirmed cases in world is shown in compare with the tested data and polynomial regression The best fit of polynomial regression prediction in shown. In this the blue lines shows the test data variation and red lines shows the polynomial regression variation.

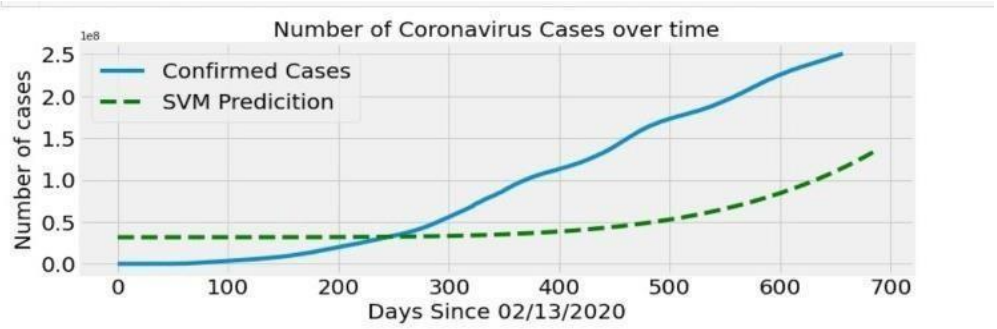

Figure:2 Plot in compare to Confirmed cases and SVM.

Description- In this graph on x axis and y axis the increase of confirmed cases in world is shown in compare with the tested data and Support Vector Machine (SVM). The best fit of SVM prediction is shown. In this the blue lines shows the test data variation and red lines shows the SVM variation .

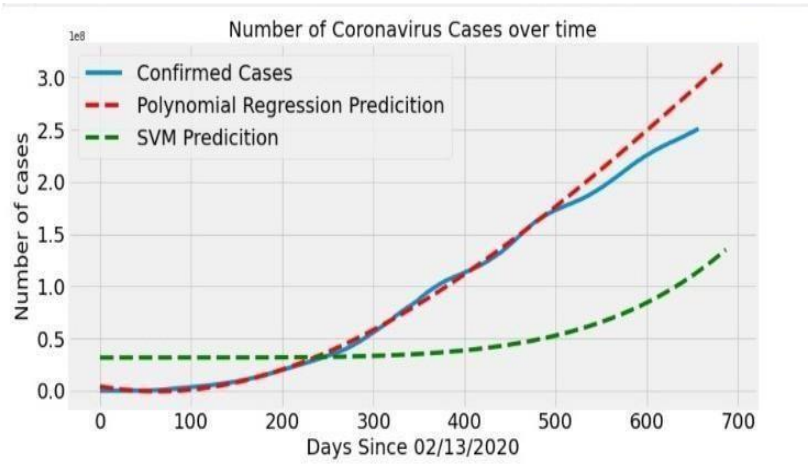

Figure: 3 Prediction graph in comparison of confirmed cases and polynomial regression prediction.

Description-In this graph on x axis the scale of 100 difference till 700 is taken with heading Days since 02/13/2020 and on y axis scale of 0.5 till 3.0 difference variation is taken this graph size is 12 on x by 6 on y, with heading number of cases is shown. This graph represents the prediction rate of coronavirus confirmed cases in the world over time in comparison with the SVM and Polynomial Regression Prediction. This outcome has been analyzed with the data we have taken for the prediction. In this graph the blue line shows the best fit line of confirmed cases with the help of data taken and the green lines shows about the prediction rate of SVM and the red line shows the Polynomial Regression Prediction. With the help of this Graph we can easily analyzed the best prediction rate of confirmed cases among these two algorithms (SVM, Polynomial Regression Prediction) in our project with the help of this graph we can say that Polynomial Regression Prediction gives the best accuracy in prediction, Because it touches the blue line of confirmed cases more accurately.

Conclusion

The Coronavirus Disease (COVID-19) Outbreak in India and outside India was studied in this report. COVID19's transmission in India and worldwide is influenced by a variety of variables which can function as a superspreader of the virus. Many unknown criteria might lead to huge uncertainty in the projection; the prediction is meant to aid the government in making future decisions and dealing with the continuing coronavirus spread in India. The conclusion of this study will aid in the planning of healthcare resources and the development of effective prediction of virus and knowledge about it. The environment is under the control of the COVID-19. These prediction aims are to use ML models to examine the epidemic with data from GitHub. In conclusion, the Polynomial Regression (PR) method is best for prediction in this research.

REFERENCES

[1] International Journal of Computer Information Systems and Industrial Management Applications. ISSN 21507988 Volume 12 (2020).

[2] MIR Labs, www.mirlabs.net/ijcisim/index.html Machine Learning and Deep Learning Covid-19 Project-Final Report Project: Spread Visualization and Prediction of the Novel Coronavirus Disease COVID-19(2019).

[3] Sahas Ramanan A., & Kumar, N. (2020). The network structure of COVID19 spread and the lacuna in India's testing strategy.https://arxiv.org/ftp/arxiv/papers/2003/2003.09715 on 3rd April 2020.

[4] Machine learning-based prediction of COVID-19 diagnosis based on symptomsYazeed Zoabi(2020)

[5] Rajan Gupta, Saibal K. Pal and Gaurav Pandey. A Comprehensive Analysis of COVID19 Outbreak situation INDIA Access from WHO Coronavirus disease (COVID-19) Pandemic (2020).

[6] Janmajay Nayak, Vighnaraaj Naik, Paidi Dinesh,Kanithi Vakula, B. Kameswara Rao, Weiping Ding, Danilo Pelusi. Intelligent system for COVID-19 prognosis: a state-of-the-art survey(2020)

[7] Singh, R., & Adhikari, R. (2020). Age-structured impact of social distancing on the COVID-19 epidemic in India. Ar Xiv preprint arXiv:2003.12055 on 4th April 2020.

[8] Shreshth Tuli, Shikhar Tuli, Rakesh Tuli, and Sukhpal Singh Gillard. Predicting the growth and trend of the COVID-19 pandemic using ML and cloud computing (2020)

[9] short-Term Prediction of COVID19 Cases Using Machine Learning Models Md. Shahriar Satu Koushik Chandra Howlader, Mufti, Mahmud, M.Sha mim Kaiser, Sheikh Mohammad Shafiullah, Julian M.W.(2020)

[10] Quinn, Salem A. Alyami and Mohammad Ali Moni , Citation: Satu, M.S.; how lade K.C.; Mahmud, M.; Kaiser, M.S.; Shaiful Islam, S.M.; Quinn, J.M.W.; alyami, S.A.; Moni, M.A(2019)