iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 10 **Issue:** V **Month of publication:** May 2022

DOI: https://doi.org/10.22214/ijraset.2022.42729

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 10 Issue V May 2022- Available at www.ijraset.com

COVID-19 Future Forecasting using Machine Learning

Nidhi Samani[1], Komal Yadav[2], Yash Bhamare[3], Meghana Patil[4], Dr. P P Halkarnikan[5]
[1, 2, 3, 4, 5]*Department of Computer Engineering, Dr. D Y Patil Institute of Engineering Management and Research, Akurdi*

***Abstract:** Machine learning (ML) based forecasting mechanisms have proved their significance to anticipate in perioperative outcomes to improve the decision making on the future course of actions. The ML models have long been used in many application domains which needed the identification and prioritization of adverse factors for a threat. Several prediction methods are being popularly used to handle forecasting problems. This study demonstrates the capability of ML models to forecast the number of upcoming patients affected by COVID-19 which is presently considered as a potential threat to mankind. In particular, four standard forecasting models, such as linear regression (LR), least absolute shrinkage and selection operator (LASSO), support vector machine (SVM), and exponential smoothing (ES) have been used in this study to forecast the threatening factors of COVID-19. Three types of predictions are made by each of the models, such as the number of newly infected cases, the number of deaths, and the number of recoveries in the next 10 days.*

I. INTRODUCTION

This problem of forecasting has been considered as a regression problem in this study, so the study is based on some state-of-art supervised ML regression models such as linear regression (LR), least absolute shrinkage and selection operator (LASSO), support vector machine (SVM), and exponential smoothing (ES). The learning models have been trained using the COVID-19 patient stats dataset taken from online websitehttps://www.worldometers.info/coronavirus/ The dataset has been pre-processed and divided into two subsets: training set (85records) and testing set (15) the current human crisis our attempt in this study is to develop a forecasting system for COVID-19. The forecasting is done for the three important variables of the disease for the coming 10 days the number of New confirmed cases, the number of death cases, the number of recoveries.

II. MOTIVATION

1) Government will easy to track covid Cases.
2) Contactless hygienic method.
3) Improved performance.

III. LITERATURE SURVEY

Suraj Bodapati, "COVID-19 Time Series Forecasting of Daily Cases, Deaths Caused and Recovered Cases using Long Short Term Memory Networks"[1] Novel Coronavirus (COVID-19) outbreak that emerged originally in Wuhan, the Hubei province of China has put the entire human race at risk. This virus was declared as Pandemic on 11th March 2020. Considering the massive growth rate in the number of cases and highly contagious nature of the virus, machine learning prediction models and algorithms are essential to predict the number of cases in the coming days. This could help in reducing the stress on health care systems and administrations by helping them plan better. In this paper the datasets used are obtained from the John Hopkins University's publicly available datasets to develop a state-of- the-art forecasting model of COVID-19 outbreak. We have incorporated data-driven estimations and time series analysis to predict the trends in coming days such as the number of cases confirmed positive, number of deaths caused by the virus and number of people recovered from the novel coronavirus. To achieve the estimations, we have used the Deep learning model long-short-term memory network (LSTM).

R.F. Sear1, N. Velasquez "Real-Quantifying COVID-19 content in the online health opinion war using machine learning"[2] A huge amount of potentially dangerous COVID-19 misinformation is appearing online. Here we use machine learning to quantify COVID-19 content among online opponents of establishment health guidance, in particular vaccinations ("anti-vax"). We find that the anti-vax community is developing a less focused debate around COVID-19 than its counterpart, the pro-vaccination ("pro-vax") community.

However, the anti-vax community exhibits a broader range of "flavors" of COVID-19 topics, and hence can appeal to a broader cross-section of individuals seeking COVID-19 guidance online, e.g. individuals wary of a mandatory fasttracked COVID-19 vaccine or those seeking alternative remedies. Hence the antivax community looks better positioned to attract fresh support going forward than the pro-vax community. This is concerning since a widespread lack of adoption of a COVID-19 vaccine will mean the world falls short of providing herd immunity, leaving countries open to future COVID-19 resurgences. We provide a mechanistic model that interprets these results and could help in assessing the likely efficacy of intervention strategies. Our approach is scalable and hence tackles the urgent problem facing social media platforms of having to analyze huge volumes of online health misinformation and disinformation.

FURQAN RUSTAM1, AIJAZ AHMAD RESHI "COVID-19 Future Forecasting Using Supervised Machine Learning Models"[3] Machine learning (ML) based forecasting mechanisms have proved their significance to anticipate in perioperative outcomes to improve the decision making on the future course of actions. The ML models have long been used in many application domains which needed the identification and prioritization of adverse factors for a threat. Several prediction methods are being popularly used to handle forecasting problems. This study demonstrates the capability of ML models to forecast the number of upcoming patients affected by COVID-19 which is presently considered as a potential threat to mankind. In particular, four standard forecasting models, such as linear regression (LR), least absolute shrinkage and selection operator (LASSO), support vector machine (SVM), and exponential smoothing (ES) have been used in this study to forecast the threatening factors of COVID-19. Three types of predictions are made by each of the models, such as the number of newly infected cases, the number of deaths, and the number of recoveries in the next 10 days. The results produced by the study proves it a promising mechanism to use these methods for the current scenario of the COVID-19 pandemic. The results prove that the ES performs best among all the used models followed by LR and LASSO which performs well in forecasting the new confirmed cases, death rate as well as recovery rate, while SVM performs poorly in all the prediction scenarios given the available dataset.

QUOC-VIET PHAM1 , DINH C. NGUYEN "Artificial Intelligence (AI) and Big Data for Coronavirus (COVID-19) Pandemic: A Survey on the State-of-the-Arts"[4] The very first infected novel coronavirus case (COVID-19) was found in Hubei, China in Dec. 2019. The COVID-19 pandemic has spread over 214 countries and areas in the world, and has significantly affected every aspect of our daily lives. At the time of writing this article, the numbers of infected cases and deaths still increase significantly and have no sign of a well-controlled situation, e.g., as of 13 July 2020, from a total number of around 13.1 million positive cases, 571, 527 deaths were reported in the world. Motivated by recent advances and applications of artificial intelligence (AI) and big data in various areas, this paper aims at emphasizing their importance in responding to the COVID-19 outbreak and preventing the severe effects of the COVID-19 pandemic. We firstly present an overview of AI and big data, then identify the applications aimed at fighting against COVID-19, next highlight challenges and issues associated with state-of-the-art solutions, and finally come up with recommendations for the communications to effectively control the COVID-19 situation. It is expected that this paper provides researchers and communities with new insights into the ways AI and big data improve the COVID-19 situation, and drives further studies in stopping the COVID-19 outbreak.

Alaa A. R. Alsaeedy and Edwin K. P. "Detecting Regions At Risk for Spreading COVID-19 Using Existing Cellular Wireless Network Functionalities"[5] Goal: The purpose of this article is to introduce a new strategy to identify areas with high human density and mobility, which are at risk for spreading COVID-19. Crowded regions with actively moving people (called at-risk regions) are susceptible to spreading the disease, especially if they contain asymptomatic infected people together with healthy people. Methods: Our scheme identifies at-risk regions using existing cellular network functionalities—handover and cell (re)selection—used to maintain seamless coverage for mobile end-user equipment (UE). The frequency of handover and cell (re)selection events is highly reflective of the density of mobile people in the area because virtually everyone carries UEs. Results: These measurements, which are accumulated over very many UEs, allow us to identify the atrisk regions without compromising the privacy and anonymity of individuals. Conclusions: The inferred at-risk regions can then be subjected to further monitoring and risk mitigation.

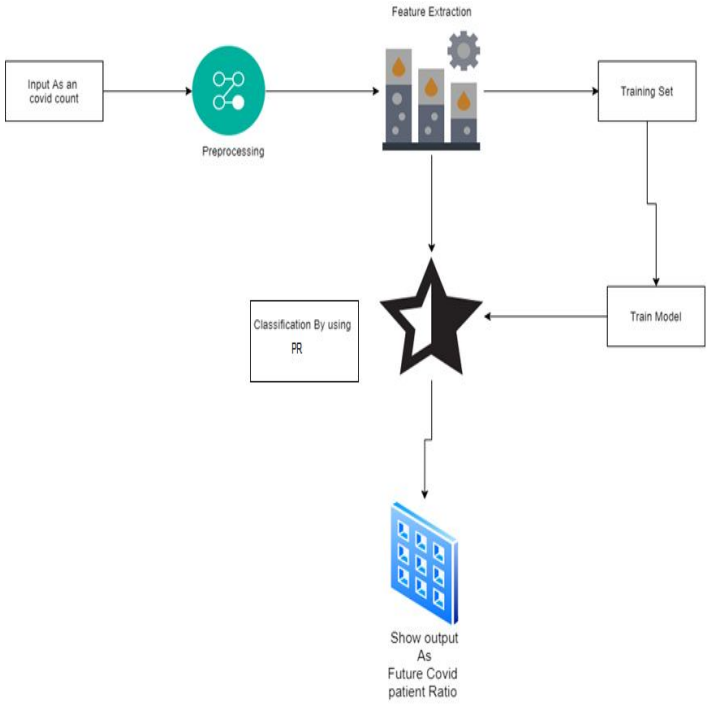
IV. PROBLEM STATEMENT

To predict the corona virus status by using machine learning algorithm

V. PROJECT OBJECTIVE

1) The forecasting is done for the three important variables of the disease for the coming 10 days:
2) The number of New confirmed cases.
3) The number of death cases
4) The number of recoveries.

VI. SYSTEM ARCHITECTURE



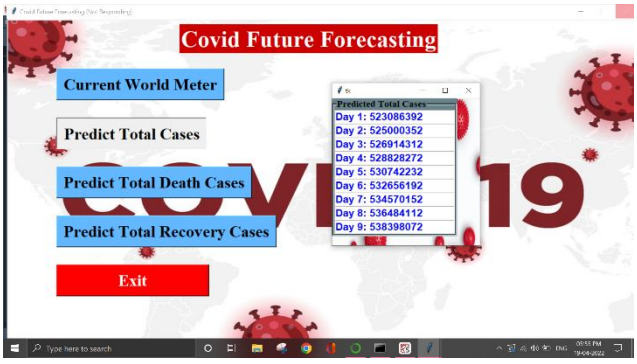
VII. METHOD AND ALGORITHM

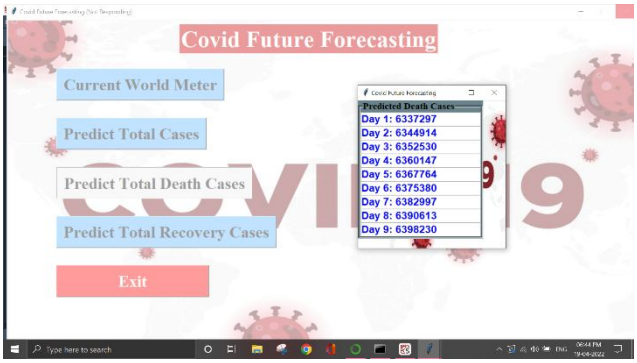
A. Polynomial Regression

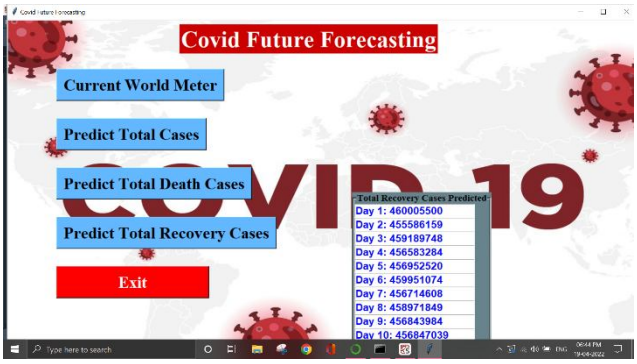
Polynomial Regression is a form of Linear regression known as a special case of Multiple linear regression which estimates the relationship as an nth degree polynomial. Polynomial Regression is sensitive to outliers so the presence of one or two outliers can also badly affect the performance. polynomial regression is a form of regression analysis in which the relationship between the independent variable x and the dependent variable y is modelled as an nth degree polynomial in x. Polynomial regression fits a nonlinear relationship between the value of x and the corresponding conditional mean of y, denoted E(y |x). Although polynomial regression fits a nonlinear model to the data, as a statistical estimation problem it is linear, in the sense that the regression function E(y | x) is linear in the unknown parameters that are estimated from the data. For this reason, polynomial regression is considered to be a special case of multiple linear regression.

VIII. RESULTS

The below pictures shows the output of this system with the upcoming confirmed cases, death cases and recovery cases predicted using machine learning based polynomial regression algorithm.





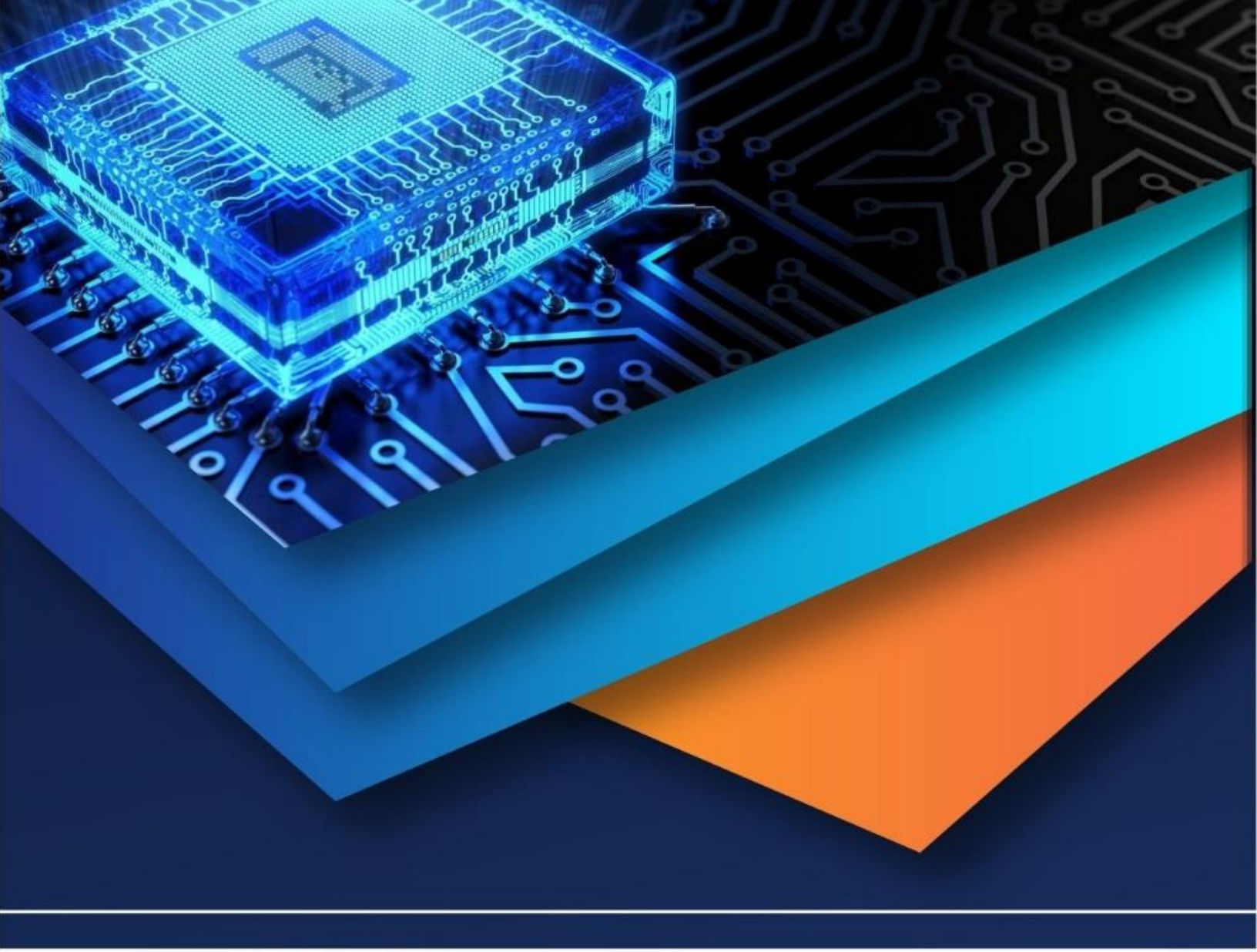


IX. CONCLUSION

In this study, an ML-based Polynomial Regression algorithm prediction system has been implemented for predicting the risk of COVID19 outbreak globally. This system analysis the dataset taken from the online website https://www.worldometers.info/coronavirus/ and predicts the new confirmed cases, new death cases and new recovery cases for the upcoming 10 days using machine learning algorithm i.e Polynomial Regression. This algorithm is preferred over other algorithms as it gives an accuracy of 94%. The results of the study prove that PR performs best in the current forecasting domain given the nature and size of the dataset. LR and LASSO also perform well for forecasting to some extent to predict death rate and confirm cases.

REFERENCES

[1] H. Asri, H. Mousannif, H. Al Moatassime, and T. Noel, "Using machine learning algorithms for breast cancer risk prediction and diagnosis," Procedia Computer Science, vol. 83, pp. 1064–1069, 2016.
[2] F. Petropoulos and S. Makridakis, "Forecasting the novel coronavirus covid-19," Plos one, vol. 15, no. 3, p. e0231236, 2020.
[3] G. Grasselli, A. Pesenti, and M. Cecconi, "Critical care utilization for the covid-19 outbreak in lombardy, italy: early experience and forecast during an emergency response," Jama, 2020.
[4] WHO. Naming the coronavirus disease (covid-19) and the virus that causes it. https://www.who.int/emergencies/diseases/novelcoronavirus-2019/technical-guidance/naming-the-coronavirus-disease(covid-2019)-and-the-virus-that-causes-it
[5] C. P. E. R. E. Novel et al., "The epidemiological characteristics of an outbreak of 2019 novel coronavirus diseases (covid-19) in china," Zhonghua liu xing bing xueza zhi= Zhonghua liuxingbingxue zazhi, vol. 41, no. 2, p. 145, 2020.
[6] R. Schraer and E. Lawrie, "Coronavirus: scientists brand 5G claims 'complete rubbish,'" BBCNews,2020. https://www.bbc.com/news/52168096.
[7] World Health Organization, "Coronavirus disease (COVID-19) advice for the public: myth busters," World Health Organization, 2020. [Online]. Available: https://www.who.int/emergencies/diseases/novelcoronavirus-2019/advice-for-public/myth-busters.
[8] K. Benner and M. Shear, "After threats, Anthony Fauci to receive enhanced personal security," The New York Times, 2020. [Online]. Available: https://www.nytimes.com/2020/04/01/us/politics/coro navirus-fauci-security.html. [Accessed: 02-Apr-2020].
[9] S. Almasy, H. Yan, and M. Holcombe, "Coronavirus pandemic hitting some African-American communities extremely hard," CNN Health, 2020. [Online]. Available: https://www.cnn.com/2020/04/06/health/uscoronavirus-updates-monday/index.html.
[10] A. Maqbool, "Coronavirus: why has the virus hit African Americans so hard?," BBC News, 2020. [Online]. Available:
[11] J. Guy, "East Asian student assaulted in 'racist' coronavirus attack in London," CNN.com, 2020. [Online]. Available: https://www.cnn.com/2020/03/03/uk/coronavirusassault- student-london-scli-intl-gbr/index.html. [Accessed: 13-Apr-2020].
[12] H. Yan, N. Chen, and D. Naresh, "What's spreading faster than coronavirus in the US? Racist assaults and ignorant attacks against Asians," CNN.com, 2020. [Online]. https://www.cnn.com/2020/02/20/us/coronavirusracist-attacks-against-asian-americans/index.html.
[13] M. Rajagopalan, "Korean interpreter says men yelling 'Chinese' tried to punch her off her bike," BuzzFeed News, 2020. [Online]. Available: https://www.buzzfeednews.com/article/meghara/coro navirus-racism-europe-covid-19. [Accessed: 13-Apr2020].



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY