



e-ISSN: 2319-8753 | p-ISSN: 2347-6710

IJIRSET

International Journal of Innovative Research in
SCIENCE | ENGINEERING | TECHNOLOGY



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

Volume 10, Issue 7, July 2021

ISSN INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA

Impact Factor: 7.569



9940 572 462



6381 907 438



ijirset@gmail.com



www.ijirset.com

Prediction Base Analysis of Covid-19 by Machine Learning

Gokul Patil¹, Manish Vaidya², Ankita Rane³, Vandana Nathjogi⁴, Bhumi Patil⁵, Dr.Satpalsing Rajput⁶

U.G. Student, Department of Computer Engineering, SSBT's COET, Jalgaon, Maharashtra, India¹

U.G. Student, Department of Computer Engineering, SSBT's COET, Jalgaon, Maharashtra, India²

U.G. Student, Department of Computer Engineering, SSBT's COET, Jalgaon, Maharashtra, India³

U.G. Student, Department of Computer Engineering, SSBT's COET, Jalgaon, Maharashtra, India⁴

U.G. Student, Department of Computer Engineering, SSBT's COET, Jalgaon, Maharashtra, India⁵

PhD Teacher, Department of Computer Engineering, SSBT's COET, Jalgaon, Maharashtra, India⁶

ABSTRACT: Several outbreak prediction models for COVID-19 are being used by officials around the world to make informed-decisions and enforce relevant control measures. Due to a high level of uncertainty and lack of essential data, standard models have shown low accuracy for longterm prediction. Peoples are facing various problem regarding accuracy of covid cases so, by using this portal, people can easily find out updated, live and predicted cases and it also predict the state wise situation after next day which are available on one portal therefore this portal is designed. By using this portal we can easily maintain the data without any loss.user can choose the state according to their own requirement. Users access the information by giving any input to the portal.

KEYWORDS: Prediction of Cases, Upcoming, Updated, live and predicted cases, Covid cases.

I. INTRODUCTION

The World Health Organisation (WHO) has declared the Coronavirus disease 2019 (COVID19). A Pandemic is defined as accuring over a wide geographic area. There are solving many very complex and sophisticated real-world problems. The study points out the key factors that are affect the spread of the virus, such as the Positive Cases, Negative Cases, and Recovered Cases. Relevant models and data analysis can provide some basis and guidance for the related state about epidemic prevention and control.

Paper is organized as follows. Section II describes prediction base analysis of covid-19 by Machine Learning, The flow diagram represents the step of the algorithm. After prediction of upcoming cases, how many cases are active, recovered and death using Machine Learning, SectionIII present experimental result. Section IV presents conclusion.

II. METHODOLOGY

The study is about novel coronavirus also known asCOVID-19 predictions. The COVID-19 has proved a present potential threat to human life. It causes tens of thousands ofconfirmed, active, recovered and death cases is increasing day by day throughout the globe. To contribute to this pandemic situation control, this study attempts to perform future forecasting on the confirmed, active, recovered and death cases in the upcoming days.

The forecasting has been done by using four ML approachesthat are appropriate to this context. The dataset used in the study contains daily summary tables, includingthe number of confirmed, active, recovered and deaths cases in the past number of days from which the pandemic started.Initially, the dataset has been preprocessed for this studyTo find the countries wise, state wise and district wise statistics of the daily number of confirmed cases, active, recovered and deaths cases and symptoms wise statistics. The statistics has been extracted from the reported data as shown in Table 1, 2, 3 and the samples of the resulted dataset are shown in Tables 1, 2, 3 respectively.

	Day	Month	Year	Detected Country	Current Status	Total Cases
0	1	1	2021	Afghanistan	Active	7608
1	1	1	2021	Afghanistan	Confirm	51526
2	1	1	2021	Afghanistan	Death	2191
3	1	1	2021	Afghanistan	Recover	41727
4	1	1	2021	Albania	Active	23501

TABLE 1. Sample of data of countries.**TABLE 2.** Sample of data of state.

	Day	Month	Year	Detected State	Current Status	Num Cases
0	1	1	2021	Andaman and Nicobar Islands	Active	1
1	1	1	2021	Andaman and Nicobar Islands	Recovered	4
2	1	1	2021	Andhra Pradesh	Active	326
3	1	1	2021	Andhra Pradesh	Recovered	350
4	1	1	2021	Arunachal Pradesh	Recovered	2

TABLE 3. Sample of data of district.

	Day	Month	Year	Detected District	Current Status	Num Cases
0	1	1	2021	Adilabad	Confirmed	6
1	1	1	2021	Agar Malwa	Confirmed	3
2	1	1	2021	Agar Malwa	Recovered	8
3	1	1	2021	Agra	Confirmed	28
4	1	1	2021	Agra	Recovered	40

After the initial data preprocessing step, the dataset has been divided into two subsets: a training set to train the models and testing set. The learning models such as LR, RFR, and MNLr have been used in this study. These models have been trained on the days, month and year, and confirmed cases, active, recovered, and death patterns. The learning models have then been evaluated based on important metrics such as R^2 -score, R^2 adjusted score, MSE, RMSE, and MAE and reported in the results. The proposed approach used in the study has been shown as a block diagram Figure.

After the initial data preprocessing step, the dataset has been divided into two subsets: a training set to train the models and testing set. The learning models such as LRC, RFC, and DTC have been used in this study. These models have been trained on the symptoms patterns. The learning models have then been evaluated based on important metrics such as f1-score, precision, recall, Accuracy and reported in the results.

III. EXPERIMENTAL RESULTS

This study attempts to develop a system for the future forecasting of the number of cases affected by COVID-19 using machine learning methods. The dataset used for the study contains information about the daily reports of the number of confirmed cases, the number of active, the number of recovered, and the number of deaths due to COVID-19 countries wide, state wide and district wide. As the confirmed cases active, recovered and deaths cases are increasing day by day which is an alarming situation for the country. The whether a people is affected with the corona virus in pandemic. This study is an attempt to forecast the number of people that can be affected in terms of confirmed cases, active cases, recovered cases, and deaths cases for the upcoming days. Three machine learning models LR, RFR and MNLr have been used to predict the number of confirmed cases, active cases, recovered cases and deaths cases.

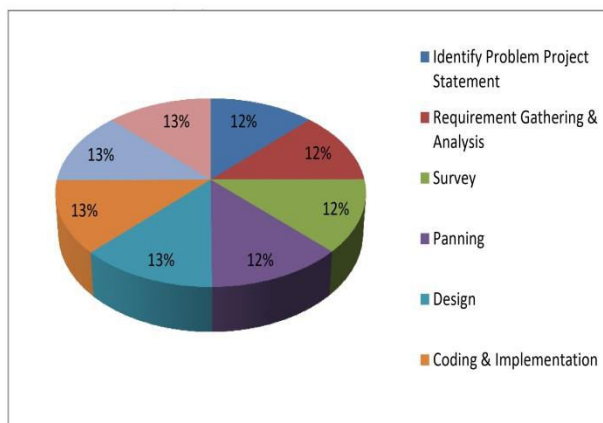


Fig. Pie Chart of Experimental Results

IV. CONCLUSION

In this project web application is created. This web application provides the facility to user for accurate prediction regarding how many COVID-19 cases are present currently in our country, state and district and whether a user is infected with corona virus. It also provides the information regarding confirmed, recovered, active and death corona cases. This web Application is visible to all. User can access information by choosing the country, state and district wise criteria. The proposed system can be expanded with availability over worldwide. This study will be enhanced continuously in the future course, next we plan to explore the prediction methodology using the updated dataset and use the most accurate and appropriate ML methods for forecasting. Real time live forecasting will be one of the primary focuses in our future work.

REFERENCES

- [1] F. E. Harrell Jr, K. L. Lee, D. B. Matchar, and T. A. Reichert, "Regression models for prognostic prediction: Advantages, problems, and suggested solutions" Cancer Treat. Rep., vol. 69, no. 10, pp. 1071, 1077, 1985.
- [2] G. Bontempi, S. B. Taieb, and Y.-A. Le Borgne, "Machine learning strategies for time series forecasting", in Proc. Eur. Bus. Intell. Summer School. Berlin, Germany: Springer, 2012, pp. 62-77
- [3] S. Makridakis, E. Spiliotis, and V. Assimakopoulos, "Statistical and machine learning forecasting methods: Concerns and ways forward", PLoS ONE, vol. 13, no. 3, Mar. 2018, Art. no. e0194889.
- [4] G. Grasselli, A. Pesenti, and M. Cecconi, "Critical care utilization for the COVID19 outbreak in lombardy, Italy: Early experience and forecast during an emergency response", JAMA, vol. 323, no. 16, p. 1545, Apr. 2020.
- [5] L. van der Hoek, K. Pyrc, M. F. Jebbink, W. Vermeulen-Oost, R. J. Berkhout, K. C. Wolthers, P. M. Wertheim-van Dillen, J. Kaandorp, J. Spaargaren, and B. Berkhout, "Identification of a new human Coronavirus "Nature Med., vol. 10, no. 4, pp. 368-373, 2004.
- [6] F. Petropoulos and S. Makridakis, "Forecasting the novel coronavirus COVID-19", PLoS ONE, vol. 15, no. 3, Mar. 2020, Art. no. e0231236.
- [7] Y. Grushka-Cockayne and V. R. R. Jose, "Combining prediction intervals in the m4 competition", Int. J. Forecasting, vol. 36, no. 1, pp. 178-185, Jan. 2020.
- [8] S. Baran and D. Nemoda, "Censored and shifted gamma distribution based EMOS model for probabilistic quantitative precipitation forecasting", Environmetrics, vol. 27, no. 5, pp. 280-292, Aug. 2016
- [9] E. Cadenas, O. A. Jaramillo, and W. Rivera, "Analysis and forecasting of wind velocity in chetumal, quintana roo, using the single exponential smoothing method", Renew. Energy, vol. 35, no. 5, pp. 925-930, May 2010.



**Impact Factor:
7.569**



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

 **9940 572 462**  **6381 907 438**  **ijirset@gmail.com**



www.ijirset.com

Scan to save the contact details