

Machine Learning-Based Data Science Model for Examination for Corona Virus Second Stage Spread Rate: A Contextual Investigation Utilizing West Bengal Dataset Till 15th May 2021

Ayan Kumar Ghosh¹, Subham Panda¹, Anup Das², Uttam Dey³, Subir Gupta^{1*}

¹Department of Masters of Computer Application, Dr. B. C. Roy Engineering College, Durgapur, West Bengal – 713206, India, subir2276@gmail.com^{*}

²CEO and founder of Anup Tech Tips

³Department of APC, Dr. B. C. Roy Engineering College, Durgapur, West Bengal – 713206, India

Abstract—Covid infection 2019 (COVID-19) has brought up pressing issues about regulation and relief, especially in nations where the infection has not yet settled human-to-human transmission. Machine learning-based data science models have been utilized to give an unequivocal system to comprehension Coronavirus second stage transmission elements in the human populace for more than three distinctive periods. These issues were tended to with regards to West Bengal, India, utilizing straightforward numerical models of irresistible illness transmission. In this article, comprehension of the spread and transmission of Coronavirus second stage have been examined. For this reason, the Machine learning (ML) way to deal with overseeing direct examination of this case. ML has different frameworks yet among those here we utilize Linear Regression for the current condition.

Keywords—Covid-19, Data Science, Linear model, Machine Learning

I. INTRODUCTION

The Coronavirus or the Roman Covid was verbalized as a pandemic by the World Health Organization in the significant segment of 2020[1][2][3]. In general, the sum of the nations of the world saw the essential mix of coronavirus that appeared in the spring, and therefore every little development increased its rate throughout the early summer[4][5]. In any case, the subsequent wave began stunningly in early 2021[6][7][8]. Activity, various assessments raising plan assessment and inspection for Indian region. The reviews in the West Bengal locality of India present models of extended length and temporarily without inhibitions[9][10].

According to various papers accessible in the writing, there are a couple of studies that attention

on pattern examination and estimating for the Indian locale[11]. Furthermore, the assessments in the Indian region from the past are more revolved around presenting time game plan examination subject to the overall data for the Indian region as opposed to covering various wellsprings of information isolated from basically pondering the number of corrupted patients, so they need to analyze the patient's experience and information is required for the experts to improve understanding about the situation[12][13][14]. Additionally, some additional mathematical models were made for separating the examples of the COVID-19 erupt in India[15][16]. In any case, it was furthermore attested that the Indian clinical benefits establishment isn't extraordinarily strong as indicated by the WHO rules and if there ought to be an event of the neighborhood, the Indian government may feel that it's difficult to manage the spread[17][18].

Machine learning (ML) is the mode for reviewing data that readies the machine to show models and settle on choices without human control[19]. Responded information is principal for man-made reasoning. As of now, ML anticipates the principal part in different spaces of assessment, like materials science, picture the board, environment appraisal, bio clinical game plan, and so on[20][21][22][23][24]. A linear equation is a famous system in ML just as data science displays for dissecting the information pattern[25][26]. It is normally used to distinguish the connection between things and determinations. Distinctive rebel models balance subjects with such a

connection among reliant and free gatherings they reflect.

II. METHODOLOGY

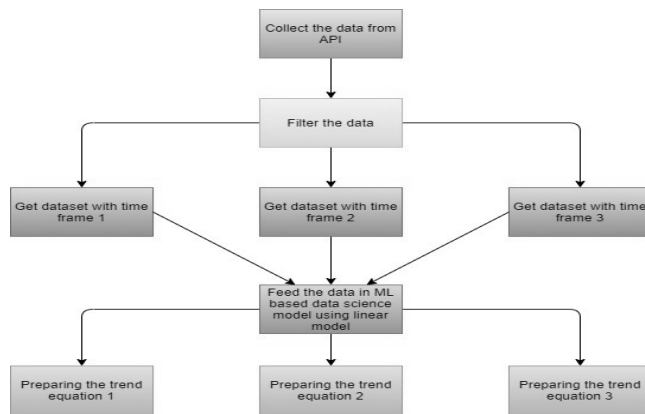


Fig. 1. Data flow diagram for preparing trained linear equation

In Figure 1, showing the strategies for setting up the dataset just as the linear equation model. From the outset, informational indexes are gathering from the public authority API. Structure the open-access dataset; separating strategies are utilized for discovering the information which is especially needed in this paper. Likewise, discover three diverse information outlines utilizing tried versus affirmed cases. For example, 01/10/2020 to 14/01/2021 is the 1st time frame, 15/01/2021 to 31/03/2021 is the 2nd time frame and 01/04/2021 to 15/05/2021 is the 3rd time frame. All the dataset goes through an ML-based information science model utilizing a direct condition. Direct conditions are completely founded on Equation no 1.

$$\hat{y} = bx + a \quad (1)$$

A. Abbreviations and Acronyms

Define abbreviations and acronyms the first time they are used in the text, even after they have already been defined in the abstract. Abbreviations that incorporate periods should not have spaces: write “C.N.R.S.,” not “C. N. R. S.” Do not use abbreviations in the title unless they are unavoidable.

B. Other Recommendations

Use one space after periods and colons. Hyphenate complex modifiers: “zero-field-cooled magnetization.” Avoid dangling participles, such as, “Using (1), the potential was calculated.” Write instead, “The potential was calculated by using (1),” or “Using (1), we calculated the potential.”

Use a zero before decimal points: “0.25,” not “.25.” Use “cm³,” not “cc.” Indicate sample dimensions as “0.1 cm × 0.2 cm,” not “0.1 × 0.2 cm².” The abbreviation for “seconds” is “s,” not “sec.” Use “Wb/m²” or “webers per square meter,” not “webers/m².” When expressing a range of values, write “7 to 9” or “7-9,” not “7~9.”

A parenthetical statement at the end of a sentence is punctuated outside of the closing parenthesis (like this). (A parenthetical sentence is punctuated within the parentheses.) In American English, periods and commas are within quotation marks, like “this period.” Other punctuation is “outside”! Avoid contractions; for example, write “do not” instead of “don’t.” The serial comma is preferred: “A, B, and C” instead of “A, B and C.”

If you wish, you may write in the first person singular or plural and use the active voice (“I observed that ...” or “We observed that ...” instead of “It was observed that ...”). Remember to check spelling. If your native language is not English, please get a native English-speaking colleague to carefully proofread your paper.

III. RESULT

For examination, the pattern of expanding three linear equations is arranged to utilize three distinctive periods. Equations 2, 3, and 4 are showing the direct condition of 01/10/2020 to 14/01/2021, 15/01/2021 to 31/03/2021, and 01/04/2021 to 15/05/2021 periods separately. Besides, the looking at line diagram shows up in Figure 2, Figure 3, and Figure 4 independently. From which we finish up the spread of the novel infection in West Bengal.

$$\hat{y} = 3110.19291X - 2294107682.58438 \quad (2)$$

$$\hat{y} = 240.87654X - 177154634.26617 \quad (3)$$

$$\hat{y} = 12276.7527X - 9058283946.89087 \quad (4)$$

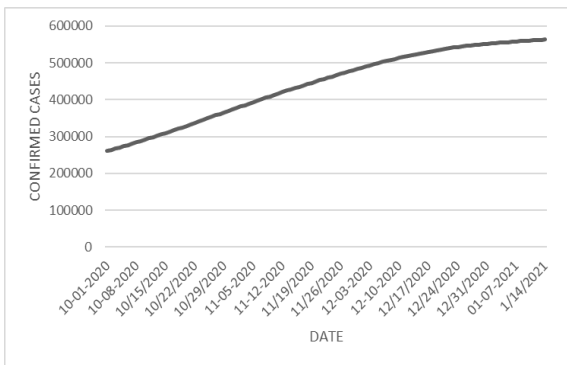


Fig. 2. Linear line of West Bengal for covid-19 spread from 01/10/2020 to 14/01/2021

Change of confirmed case concerning test case, where both are variable i.e. dy/dx of equations 2, 3, and 4 are separately 3110, 240, and 12276. Which guaranteed that the spread rate moderate in the 1st time frame, while its particularly steady in the 2nd time frame, and the primary upheaval or high change in the 3rd time frame. During the third schedule opening, the infection spread 51 multiple times quicker than the second time allotment and during the third time allotment, the infection spread four multiple times quicker than the first time allotment.

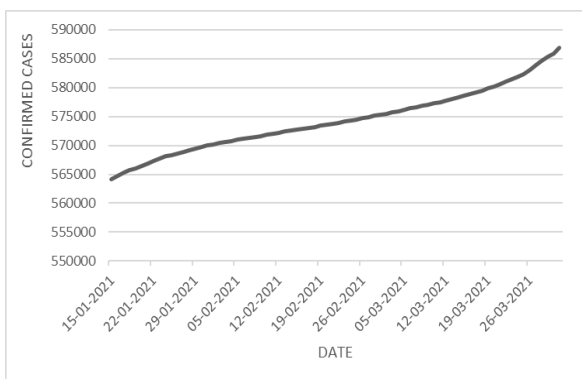


Fig. 3. Linear line of West Bengal for covid-19 spread from 15/01/2021 to 31/03/2021.

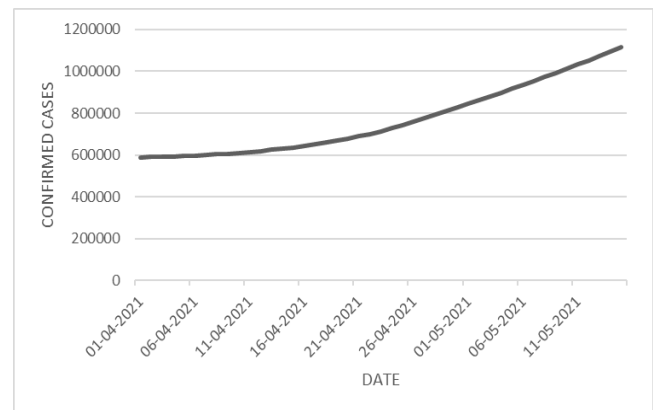


Fig. 4. Linear line of West Bengal for covid-19 spread from 01/04/2021 to 15/05/2021

IV. CONCLUSION

This investigation presented the most recent things of the COVID-19 scene from 01/10/2020 to 15/05/2021 as imagined in API data with a background concerned with an examination of West Bengal. As indicated by results, it can doubtlessly see 01/04/2021 to 15/05/2021 i.e. the second surge of Covid-19 is an upsetting condition in West Bengal. There are satisfactory purposes behind called full lockdown by the West Bengal government. This paper is at this point persistent assessment as much more assessments concerning this disease should be possible. To go about as a representation for hitter results and finding the security course of action certainly a better curve (non-linear curve) ought to be thought about similarly as construct a differential equation is required for finding a trustworthiness plan.

REFERENCES

- [1] "Coronavirus." https://www.who.int/health-topics/coronavirus#tab=tab_1 (accessed Apr. 29, 2021).
- [2] "Coronavirus disease (COVID-19)." <https://www.who.int/emergencies/diseases/novel-coronavirus-2019> (accessed Apr. 28, 2021).
- [3] "Coronavirus disease (COVID-19): How is it transmitted?" <https://www.who.int/news-room/q-a-detail/coronavirus-disease-covid-19-how-is-it-transmitted> (accessed Mar. 29, 2021).
- [4] S. J. Fong, N. Dey, and J. Chaki, "An Introduction to COVID-19," in *SpringerBriefs in Applied Sciences and Technology*, Springer, 2021, pp. 1–22.

- [5] "Coronavirus." https://www.who.int/health-topics/coronavirus#tab=tab_1 (accessed Mar. 29, 2021).
- [6] "Second wave of COVID-19 may peak by April 15-20: What we know so far." <https://www.businesstoday.in/current/economy-politics/second-wave-of-covid-19may-peak-by-april-15-20-what-we-know-so-far/story/436166.html> (accessed Apr. 29, 2021).
- [7] "India to see worst of COVID-19 second wave in April-May, what data experts predict | The News Minute." <https://www.thenewsminute.com/article/india-see-worst-covid-19-second-wave-april-may-what-data-experts-predict-147391> (accessed Apr. 23, 2021).
- [8] "Coronavirus Second Wave: India Likely to Witness 2,320 Daily Deaths by First Week of June, Says Report." <https://www.india.com/news/india/coronavirus-second-wave-india-likely-to-witness-2320-daily-deaths-by-first-week-of-june-says-report-4589701/> (accessed Apr. 29, 2021).
- [9] "West Bengal issues revised schedule for COVID-19 lockdown in August. New dates are here." <https://www.livemint.com/news/india/west-bengal-issues-revised-schedule-of-covid-19-lockdown-new-dates-are-here-11597233643769.html> (accessed Apr. 03, 2021).
- [10] "West Bengal Lockdown On 7 Different Days In August, Opposition Says 'It's All Whimsy.'" <https://www.ndtv.com/india-news/west-bengal-lockdown-on-7-different-days-in-august-opposition-says-its-all-whimsy-2270597> (accessed Apr. 03, 2021).
- [11] "Sci-Hub | India under COVID-19 lockdown | 10.1016/S0140-6736(20)30938-7." [https://www.sci-hub.ren/10.1016/S0140-6736\(20\)30938-7](https://www.sci-hub.ren/10.1016/S0140-6736(20)30938-7) (accessed Apr. 23, 2021).
- [12] K. H. Byeon, D. W. Kim, J. Kim, B. Y. Choi, B. Choi, and K. D. Cho, "Factors Affecting the Survival of Early COVID-19 Patients in South Korea: An Observational Study based on the Korean National Health Insurance Big Data," *Int. J. Infect. Dis.*, vol. 105, pp. 588–594, 2021, doi: 10.1016/j.ijid.2021.02.101.
- [13] S. Mandal, R. Sarkar, and S. Sinha, "Mathematical models of malaria - A review," *Malar. J.*, vol. 10, pp. 1–19, 2011, doi: 10.1186/1475-2875-10-202.
- [14] M. Res, C. Ruts, C. R. Hospital, M. Sciences, I. E. Committee, and S. Crh-smims, "Prevalence of," no. May, pp. 517–520, 2018, doi: 10.4103/ijmr.IJMR.
- [15] J. Pre-proof, "Coronavirus lockdown and virus suppression: an international analysis," *Technol. Forecast. Soc. Chang.*, p. 120861, 2021, doi: 10.1016/j.techfore.2021.120861.
- [16] Y. Xu, Y. Zhou, P. Sekula, and L. Ding, "Machine learning in construction: From shallow to deep learning," *Dev. Built Environ.*, vol. 6, no. April 2020, p. 100045, 2021, doi: 10.1016/j.dibe.2021.100045.
- [17] "Coronavirus: Here Are The Steps Taken Across India To Control The Spread of COVID-19 | News." <https://swachhindia.ndtv.com/coronavirus-here-are-the-steps-taken-by-india-to-control-the-spread-of-covid-19-42304/> (accessed Apr. 29, 2021).
- [18] A. P. Singh, "Model for prediction of death rate due to COVID-19 transmission and required precautions," *Mater. Today Proc.*, vol. 37, no. Part 2, pp. 2318–2320, 2020, doi: 10.1016/j.matpr.2020.07.731.
- [19] N. J. Nilsson, "INTRODUCTION TO MACHINE LEARNING AN EARLY DRAFT OF A PROPOSED TEXTBOOK," 1998.
- [20] S. Gupta et al., "Modelling the steel microstructure knowledge for in-silico recognition of phases using machine learning," *Mater. Chem. Phys.*, vol. 252, p. 123286, 2020, doi: 10.1016/j.matchemphys.2020.123286.
- [21] S. Gupta, J. Sarkar, M. Kundu, N. R. Bandyopadhyay, and S. Ganguly, "Automatic recognition of SEM microstructure and phases of steel using LBP and random decision forest operator," *Meas. J. Int. Meas. Confed.*, vol. 151, p. 107224, 2020, doi: 10.1016/j.measurement.2019.107224.
- [22] X. Li and T. Liu, "Sports image recognition based on FPGA and Machine Learning," *Microprocess. Microsyst.*, no. November, p. 103490, 2020, doi: 10.1016/j.micpro.2020.103490.
- [23] Z. Khandezamin, M. Naderan, and M. J. Rashti, "Detection and classification of breast cancer using logistic regression feature selection and GMDH classifier," *J. Biomed. Inform.* vol. 111, p. 103591, 2020, doi: 10.1016/j.jbi.2020.103591.
- [24] S. Panda, A. K. Ghosh, A. Das, U. Dey, and S. Gupta, "Machine Learning-based Linear regression way to deal with making data science model for checking the sufficiency of night curfew in Maharashtra, India," vol. 1, no. 2, pp. 168–173, 2021.
- [25] J. Burrell, "How the machine 'thinks': Understanding opacity in machine learning algorithms," *Big Data Soc.*, vol. 3, no. 1, p. 205395171562251, Jan. 2016, doi: 10.1177/2053951715622512.
- [26] D. A. Otchere, T. O. Arbi Ganat, R. Gholami, and S. Ridha, "Application of supervised

machine learning paradigms in the prediction of petroleum reservoir properties: Comparative analysis of ANN and SVM models,” J. Pet. Sci. Eng., p. 108182, 2021, doi: 10.1016/j.petrol.2020.108182.