A Classification of COVID-19 cases using Fine-Tune Different Model of Machine Learning

Km. Mehak Jain[1], Mr. Sumit Kumar Modi[2], Ms. Namita Sharma[3]

[1](MCA Department, Noida Institute of Engineering and Technology, India)
[2](MCA Department, Noida Institute of Engineering and Technology, India)
[3](MCA Department, Noida Institute of Engineering and Technology, India)

Abstract: *The Machine Learning Techniques are used for finding upcoming or futuristic Prediction. It's very powerful tools in finding of future prediction. From Last decade to recent days Data mining and Machine Learning play vital roles in many Industries. In last 2-3 Years overall World is suffering from a new Virus i.e., COVID-19. The first covid case was detected in Month of December. In very less time it spread out throughout the world. In 2-3 months, this illness is declared as Pandemic by WHO (World Health Organization). The behaviour of spreading of COVID-19 was very abnormal. it was growing Exponentially. Here in this Paper as a Researcher's we are trying to find out overall Analysis of COVID-19. Here our main motto to finding How many patients gets infected, how many of them get cure in sometimes & finally How many of them goes to death. Here our main concentration is to finding the features of the COVID. We will be applying Machine Learning Algorithms like LR, SVM, DT and many more. After applying the above methods, we will find the performance measures of each one and compare with their values. Finally, we will try to Implement some new Enhancement in Existing Algorithms.*

Keywords - Machine Learning, COVID-19, Training & Testing Data, Model or classifier, NB, SVM, DT, Performance Measures.

1. INTRODUCTION

This is World Known facts that Viruses having very huge family that may causes different types of illness. COVID-19 having also different variants which causes different types of illness. In this illness patient

may faces different symptom like common cold and throat infection [6]. If we talk about the variation that is MERS & MERS-Cov. SARS-Cov is also a different variation of COVID-19. if we talked about initiation of above two diseases one is originated from China in 2002 i.e., SARS-Cov- 2 and another one originated from Middle East with the name MERS in 2012 [7,8]. The very new and recent one was originated from city Wuhan which is situated in China. From here only this COVID-19 starting spreading into different Region of the World. If we talk about official claim, then this is claimed by China government and Reported to WHO to December 2019. In very less time this is spread to overall world and infection rate was very high. With infection death toll was also very high. As we all know this disease is spreading through touch so Here Artificial Intelligence Enable Devices i.e., Robotics play very important and Effective Role [13]. Electronics devices i.e. Robotics works upon Data gather and implement different Algorithms from Data mining and Machine Learning. In recent days many hospitals and organization start using the devices which play base upon ML Techniques. In Health Care we are using these techniques in very effective manner. Here Researcher's used a dataset with record values 44,25,485 cases. out of them 3,02,059 went dead. and 203 countries get infected in very less time. Due to huge number of increments in cases all hospitals get over burden and this industry is going to collapse in very less time in throughout the world. In even Developed countries they face the same problem. In developing countries like India and other faces more issues in this time. Due to collapse Machine like ventilators and different apparatus play very important roles in mechanical monitoring of patient.

Prediction means to analyse the future steps of any events depends upon the previous data and their behaviour. In this we have Different categories which is given below:

Short-term correlations and certain behavioral patterns that frequently result in a particular behavior can be explained by **Predictive models**.

According to the provided dataset, **Descriptive models** provide an explanation by dividing or segmenting the data.

1.1 Techniques of Machine Learning for Sentiment Analysis

As everyone is aware, machine learning algorithms are playing a more and bigger role in many different industries. In this section, we used a dataset of movie reviews to work on sentiment analysis or polarity detection. We'll talk about a few of the many algorithms we have at our disposal to tackle this problem:
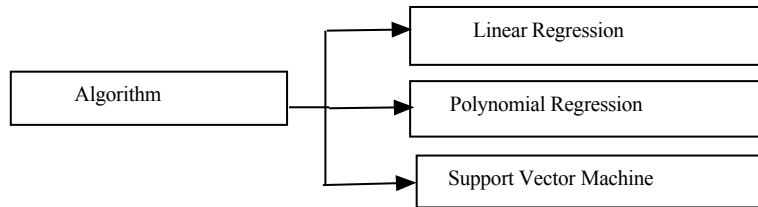
Figure 1: Algorithms used in sentiment analysis

2. RELATED WORK

Here explanation is given on results based upon reflective analysis of various patients from Wuhan itself. The name of Hospital is Jinyintan Hospital. Here the epidemiological data set is explained. Here patients having different features or symptom gathered and taken all lab results as a features or attributes of the given dataset [16]. By the above research they came to conclusion the pattern of clinical research which will support for further analyses. Here Researcher's go into deep dive for analysing the demographic impact by these diseases. We all knows its behaviour varies from one continent to another continent. Firstly, all data taken by Wuhan only. but analysing more we need data from different Region & Continent. Here Authors took different lab analysis like CT images. Here Machine learning play vital roles to detect any deep analysis of the given Image. By the research Author's want to contribute their research information to the society or doctors' community which will help to our people by developing medicine accordingly. In this Papers we explained that how anyone create a such framework which will detect the Corona Virus using different mechanical devices like smartphone enabled sensors. Here AI comes into existence to find the result from various sensor to predict the symptom of Corona will detect automatically without touch of any human being. so spreading ratio may be going deep down. and finally, we will get our result to further analysis.

3. GOAL

This paper uses machine learning approaches to predict whether an individual has COVID-19 or not. The clinical data for the patients forms the basis of the forecast. Finding out if a patient can receive a COVID-19 diagnosis is the goal.

4. OBJECTIVE

Finding the optimal machine learning method for patient clinical report prediction is the main objective of our work. creating a machine learning method that can correctly forecast patients' COVID-19.
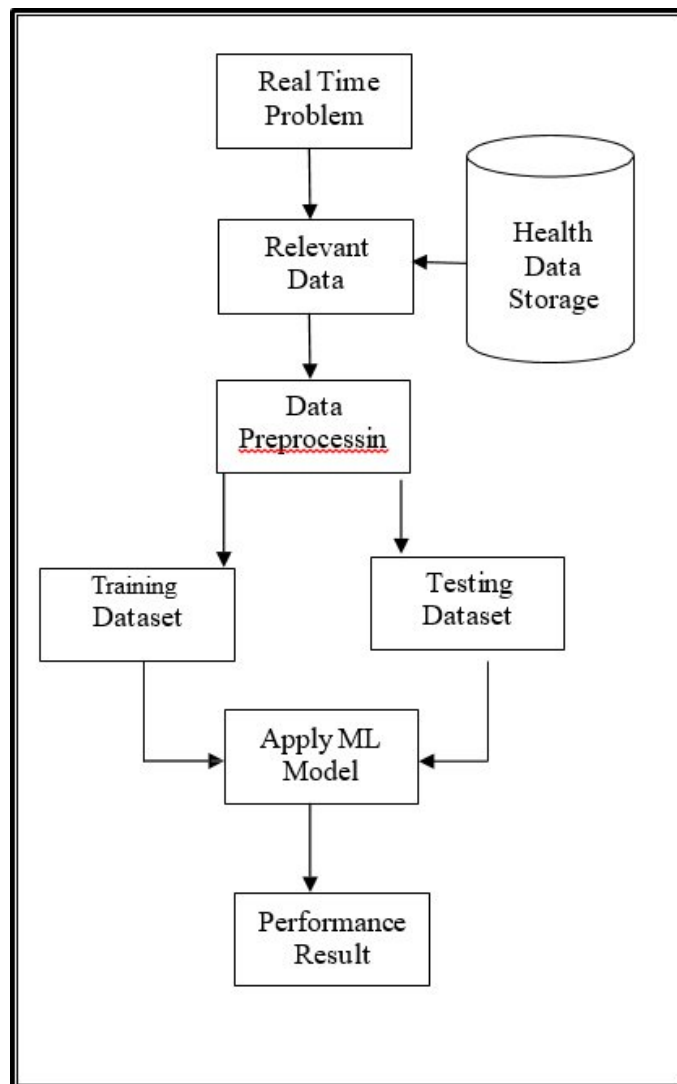

Figure 2: Program Diagram

5. ALGORITHM

Step 01: Save the data from the Kaggle Repository

Step 02: Add Prior Libraries

Step03: Add Required Dataset

\# Let's read the data set

data = pd.read_csv('covid-19.csv')

data. shape

Step04: Apply Feature Extraction

   a)   Bivariate Analysis

   b)   Multivariate Analysis

Step 05: Visualize Data to improve understanding

Step06: Implementing Machine Learning Algorithms

Step07: Implement different Model

Step08: Repeat Step 07 several times using various algorithms

Step09: Lastly, compare the results to performance metrics like as the RMSE Score & R2 Score.
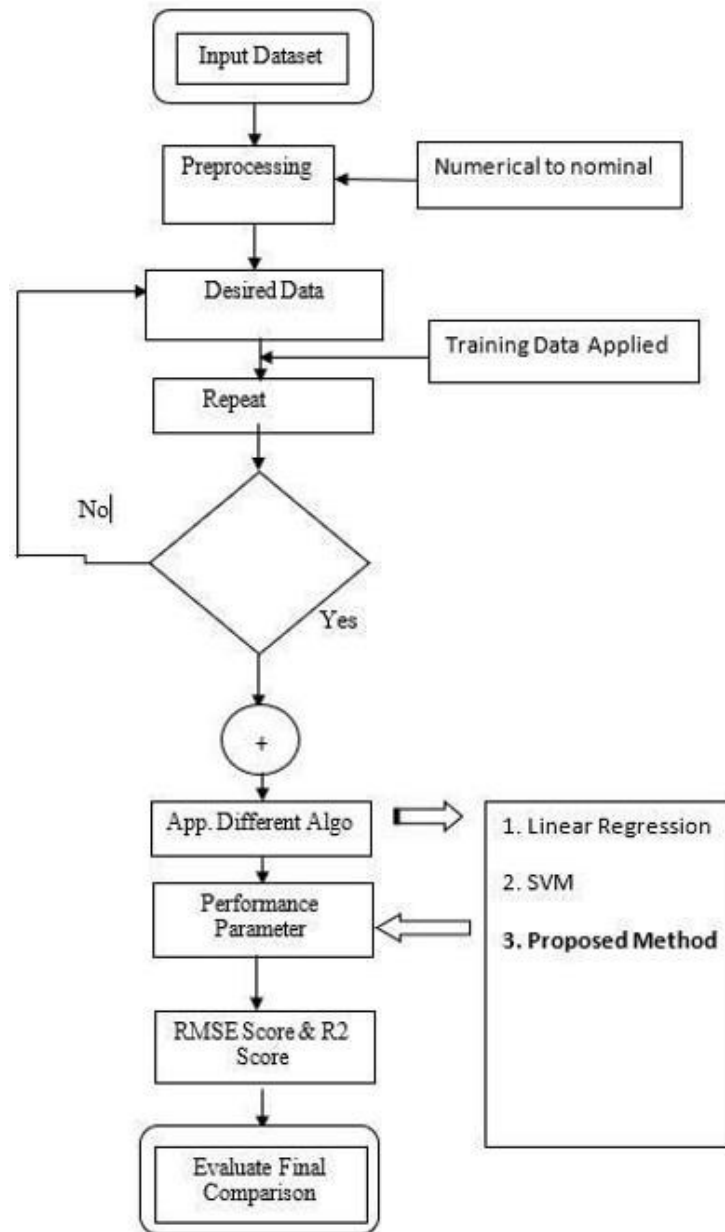
6. FLOWCHART

Figure 3: Flow Diagram

The first step in Fig. 3 is to get data from the local market or any other source; however, we are pulling data from Kaggle in order to improve analysis. That is a highly trustworthy global data source. Getting many libraries to process our data is the next stage. at order to be ready for the next step, we need to process our data at the third stage. Our processing mechanisms are diverse. We apply both univariate and multivariate data processing, as well as numerical to nominal data conversion. We must repeat it for every distinct data split in order to move on to the fourth stage. Next, we'll apply a number of machine learning algorithms in step 5 before applying our own suggested techniques—Gradient Boosting with Weighted Average values—in the last phase. The average values of the previously defined mechanism will be added here. Finding performance metrics like the RMSE Score and R2 Score, which offer distinct viewpoints on the new and current approaches, will be the last stage, or ninth phase. Finally, we shall compare the results that were provided. We can state that the outcomes of our suggested procedures are better.

7. PROCEDURE FOR IMPLEMENTATION

Filters are used by the model to expedite evaluation and shorten total time. The final results of classifiers' evaluation are greatly influenced by the preprocessing steps and filter applications (ML-based models). Feature extraction, nominal to binary conversion, and cleaning are a few of these filters.

```python
[1]: #Lets import the required Libraries

# for mathematical operations
import numpy as np
# for dataframe manipulations
import pandas as pd

# for data visualizations
import seaborn as sns
import matplotlib.pyplot as plt
import plotly.express as px

# setting parameters for visualization
plt.rcParams['figure.figsize']=(16,5)
plt.style.use('fivethirtyeight')
```

Figure 4: Calling Libraries

In the Fig. 4 we called Libraries which will help you to call all functionality which required.

	SNo	ObservationDate	Province/State	Country/Region	Last Update	Confirmed	Deaths	Recovered
0	1	01/22/2020	Anhui	Mainland China	1/22/2020 17:00	1.0	0.0	0.0
1	2	01/22/2020	Beijing	Mainland China	1/22/2020 17:00	14.0	0.0	0.0
2	3	01/22/2020	Chongqing	Mainland China	1/22/2020 17:00	6.0	0.0	0.0
3	4	01/22/2020	Fujian	Mainland China	1/22/2020 17:00	1.0	0.0	0.0
4	5	01/22/2020	Gansu	Mainland China	1/22/2020 17:00	0.0	0.0	0.0

Figure 5: Major Columns

In the Fig. 5 we try to show number of major columns available in our DataSet. We have 8 columns in our data set.

```
Basic Information
Totol number of countries with Disease Spread:  226
Total number of Confirmed Cases around the World:  96158735.0
Total number of Recovered Cases around the World:  53035240.0
Total number of Deaths Cases around the World:  2056996.0
Total number of Active Cases around the World:  41066499.0
Total number of Closed Cases around the World:  55092236.0
Approximate number of Confirmed Cases per Day around the World:  264172.0
Approximate number of Recovered Cases per Day around the World:  145701.0
Approximate number of Death Cases per Day around the World:  5651.0
Approximate number of Confirmed Cases per hour around the World:  11007.0
Approximate number of Recovered Cases per hour around the World:  6071.0
Approximate number of Death Cases per hour around the World:  235.0
Number of Confirmed Cases in last 24 hours:  599088.0
Number of Recovered Cases in last 24 hours:  400355.0
Number of Death Cases in last 24 hours:  16889.0
```

Figure 6: Columns Description

In the Fig. 6 We make an effort to display disease spread on a global, national, and other scales.
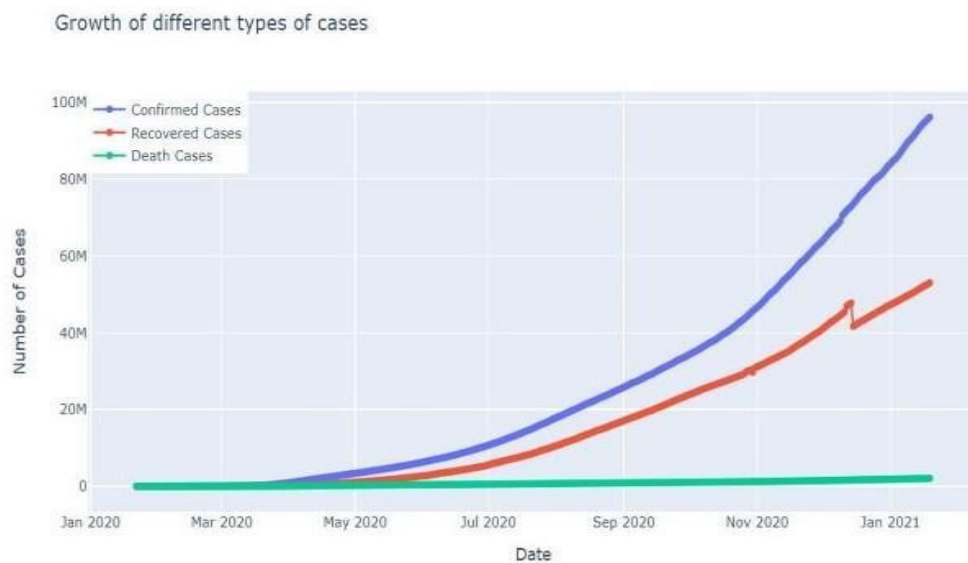
Figure 7: Growth of different types of cases

Table 1: Comparison of Algorithms

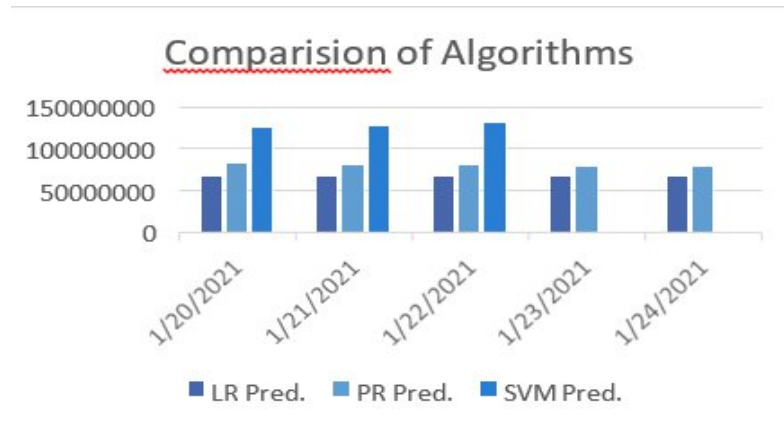| Dates	LR Pred.	PR Pred.	SVM Pred
2021-01-20	64534315.76411	80161880.46839	124201136.41425
2021-01-21	64755138.31053	79272179.85058	126099178.5141
2021-01-22	64975960.85694	78281099.64779	128023400.138555

Figure 8: Comparison Analysis

We may deduce from Fig. 8 that different outcomes are obtained using SVM Prediction values, Polynomial Regression, and Liner Regression. We will now go over a few tuning techniques to enhance the outcomes.

8. CONCLUSION

After testing several machine learning methods to obtain the best performance outcomes, we have come to this conclusion. We found that the state, country, confirmed cases, and death cases are the main characteristics of the dataset. The most crucial elements in determining the case distribution of a patient. In the end, we used a variety of machine learning techniques, including polynomial and linear regression, and our suggested approach yielded superior outcomes.

9. FUTURE SCOPE

Subsequent research endeavors will concentrate on utilizing alternative methodologies to optimize these technologies' performance to the maximum degree feasible. Learning can be employed as an alternative to machine learning technologies. This is because the most efficient and successful methods are already being employed. Recently, deep learning has also been developed, and its use to categorization is growing. Consequently, deep learning can also be used in future projects.

REFERENCES

[1] Bhardwaj, S., Bhardwaj, H., Bhardwaj, J. and Gupta, P., November. Global Prediction of COVID-19 Cases and Deaths using Machine Learning. In *2021 Sixth International Conference on Image Information* Processing (ICIIP) (Vol. 6, pp. 422-426). IEEE.

[2] Sen, A., Kala, U. and Manchanda, A., 2021, February. Analysis and prognosis of COVID-19 pandemic in India-A machine learning approach. In 2021 International Conference on Advances in Electrical, Computing, Communication and Sustainable Technologies (ICAECT) (pp. 1-6). IEEE.

[3] Gambhir, E., Jain, R., Gupta, A. and Tomer, U., 2020, September. Regression analysis of COVID-19 using machine learning algorithms. In *2020 International conference on smart electronics and communication (ICOSEC)* (pp. 65-71). IEEE.

[4] Mary, L.W. and Raj, S.A.A., 2021, October. Machine Learning Algorithms for Predicting SARS-CoV-2 (COVID- 19)–A Comparative Analysis. In *2021 2nd International Conference on Smart Electronics and Communication (ICOSEC)* (pp. 1607-1611). IEEE.

[5] Lalmuanawma, S., Hussain, J. and Chhakchhuak, L., 2020. Applications of machine learning and artificial intelligence for Covid-19 (SARS-CoV-2) pandemic: A review. *Chaos, Solitons & Fractals*, *139*, p.110059.

[6] Kumari, P. and Toshniwal, D., 2020, November. Real-time estimation of COVID-19 cases using machine learning and mathematical models-The case of India. In *2020 IEEE 15th International Conference on Industrial and Information Systems (ICIIS)* (pp. 369-374). IEEE.

[7] Patel, D., Kher, V., Desai, B., Lei, X., Cen, S., Nanda, N., Gholamrezanezhad, A., Duddalwar, V., Varghese, B. and Oberai, A.A., 2021. Machine learning based predictors for COVID-19 disease severity. *Scientific reports*, *11*(1) pp1-7.

[8] Subudhi, S., Verma, A., Patel, A.B., Hardin, C.C., Khandekar, M.J., Lee, H., McEvoy, D., Stylianopoulos, T., Munn, L.L., Dutta, S. and Jain, R.K., 2021. Comparing machine learning algorithms for predicting ICU admission and mortality in COVID-19. *NPJ digital medicine*, *4*(1), pp.1-7.

[9] Fernandes, F.T., de Oliveira, T.A., Teixeira, C.E., Batista, A.F.D.M., Dalla Costa, G. and Chiavegatto Filho, A.D.P., 2021. A multipurpose machine learning approach to predict COVID- 19 negative prognosis in São Paulo, Brazil. *Scientific reports*, *11*(1), pp.1-7.

[10] Coronavirus Disease (COVID-19) - events as they happen. Library Catalog:www.who.int.

[11] Countries where Coronavirus has spread - Worldometer. Library Catalog:www.worldometers.info.

[12] COVID-19 situation reports. Library Catalog: www.who.int.

[13] Diagnosis of covid-19 and its clinical spectrum dataset.url=https://kaggle.com/einsteindata4u/covid19.

[14] WHO Director-General's remarks at the media briefing on 2019-nCoV on 11 February 2020. Library Catalog: www.who.int.

[15] WHO EMRO | Questions and answers | COVID-19 | Health topics.

[16] Support Vector Machine Machine learning algorithm with example and code,January 2019. Library Catalog: www.codershood.info Section: Machine learning.

[17] Ali Al-Hazmi. Challenges presented by MERS corona virus, and SARS corona virus to global health. Saudi journal of biological sciences, 23(4):507–511, 2016.
Publisher: Elsevier.

[18] Sina F Ardabili, Amir Mosavi, Pedram Ghamisi, Filip Ferdinand, Annamaria R Varkonyi-Koczy, Uwe Reuter, Timon Rabczuk, and Peter M Atkinson. Covid-19 outbreak prediction with machine learning.

Available at SSRN 3580188, 2020.