International Journal of Research Publication and Reviews

Covid-19 Future Forecasting Using Supervised Machine learning Algorithms

Dr Dhananjay Maktedar, Sudha Rajkumar, Shreya N Chalwa,Shilpa Rani, Shivaganga Mantole.

A B S T R A C T

This study bespeak the ability of ML models to prognosticate the number of forthcoming patients affected by COVID-19 which is presently considered as a major harm to the humans. In consideration three major ML models have been used in this study those are mainly Linear Regression (LR), Support Vector Machine (SVM), K- nearest neighbors (KNN) for forecasting the threatening elements of COVID-19.In this study three various predictions are made such as the number of deaths, number of recovered patients, number of new cases.This study is the day-wise analysis.

1.Introduction

COVID-19 pandemic as well-known has caused health crisis globally, the essential and preventive ways are face mask wearing in the mostly crowded places and even while stepping out of house for any work, sanitizing frequently, social distancing as directed by world health organization(WHO). Due to pandemic Government were forced to impose lockdown to hinder the transmission of virus. The effective and economic approach of AI being used is to generate an environment that is safe in the society of human.

Forecasting is one of the most significant areas of ML, many ML algorithms that are standard have been used within this area for guiding the future course of actions which are needed in numerous applications such as disease prognosis.

2.Literature Survey

- This is the state-of-art of ML regression models which are supervised.
- Linear regression, Support vector machine, K nearest neighbors models.
- These learning models are trained by the covid-19 patient stats dataset.
- Preprocessed dataset is mainly split into two subsets 75% is recorded by training set and 25% is recorded by testing set.

3.System Specification

3.1 Software Requirements
1 .computer-processor, high speed preferred. 64-bit
2. RAM 4GB
3. hard disk-free space of5GB

3.2 Hardware Requirements
1. Python 3.7
2. OS-windows 8/10, 64-bitt
3. Python (IDLE or ANACONDA)

* *Corresponding author.* Tel:.+91 94487 54514
E-mail address: dhananjay.maktedar@gmail.com

4. System Design

4.1 Problem Definition

COVID-19 pandemic had adverse and negative impact on the economy of the country, it has also disrupted the mental and physical health of humans and even the well-being of an individual.

Due to cases being increased and the healthcare practitioners, Government under-going burden some measures are needed to lighten the burden, those measures can be predicting the number of infected COVID-19 cases which will be further helpful in the planning required hospital resources.

4.2 Proposed System

To handle forecasting problems several methods of prediction are being used. This study shows the capability of ML models for forecasting affected people by COVID-19 which is presently considered as a potential threat to mankind.

In general, standard forecasting models, such as logistic regression(LR), and support vector machine(SVM), have been used in this study to forecast the threatening factors of COVID-19.

4.3 Dataset

 The live data is taken from the below link:

 Novel corona virus (COVID-19): archive and API of India data

 https://thejeshgn.com/2020/03/16/novel-corona-virus-covid19-archive-api-india-data/

The data is converted into .csv from json by the following code:

```
Import urllib.request, json, csv
From pandas.io.json import json_normalize
Ur
I=https://data.thejeshgn.com/covid19/_design/india/_view/incidents?include_docs=true&descending=true
Responses=urllib.request.urlopen(url)
Data=json.loads(response.read())
```

5. Advantages

➢ It helps in predicting death rates, recovered cases, and new upcoming cases.

➢ Prediction will be helpful in providing some sort of information to the Government, hospitals, citizens of that particular areas.

➢ Government can impose some lockdowns which assures people to stay at home and discontinue the flow of transmission.

➢ The areas where the probability of risk of virus is high, hospitals can prepare themselves with effective medications , with better treating techniques and lot more.

➢ Responsible doctors can be shifted to the most affecting areas for the better care taking of patients.

➢ Responsible Citizens will be alert if the over- all prediction is brought to their notice, by following some rules such as wearing of masks, frequent sanitization and social distancing.

6. Applications

➢ These predicting models can be inserted into the systems of government and hospitals.

➢ Various workplaces or institutes can also have access to these implemented predicting models for their awareness and safety.

➢ Hospitals can keep a check on it based on day-to-day analysis.

7. Methodology

➢ Gathering COVID-19 dataset.

➢ Data preprocessing is to be done.

➢ Selection of layers, numbers and functions based on performance function.

➢ Data splitting (i) Training set (ii) Testing set.

➢ Training set includes applying of different ML learning models.

➢ Testing set includes evaluation of parameters.

➢ Selection of best architecture of network based on performance function.

➢ Forecast of covid-19 confirmed cases.

8. Implementation

- ➢ Implementation begins with the collecting dataset from different sources.
- ➢ The collected data set is then cleaned
- ➢ The cleaning of dataset is done by checking the existence of any null values.
- ➢ Next step is the visualization of the dataset.
- ➢ The visualization mainly includes graphs.
- ➢ After visualization data preprocessing is done.
- ➢ Applying ML models Logistic regression, Support vector machine and K-nearest neighbors.
- ➢ Finding which model gives the best accuracy.

9. Conclusion

The analyzation of dataset is done which contains the month wise actual previous data and prediction is done for upcoming days using machine learning algorithm.We conclude that support vector machine performs the best in forecasting with provided nature and size of the dataset.Logistic regression performs good at some level for predicting death rate and confirmed cases.Over all we conclude that model predictions are correct which are helpful in understanding the upcoming situation.It is also helpful in alerting the authorities to take the timely action and make decisions contain the crisis of covid-19.This study will be enhanced in future courses continuously.

REFERENCES

[1] 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.

[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] 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.

[4] P. Lapuerta, S. P. Azen, and L. Labree, ''Use of neural networks in predicting the risk of coronary artery disease,'' Comput. Biomed. Res., vol. 28, no. 1, pp. 38–52, Feb. 1995.

[5] K. M. Anderson, P. M. Odell, P. W. Wilson, and W. B. Kannel, ''Cardiovascular disease risk profiles,'' Amer. heart J., vol. 121, no. 1, pp. 293–298, 1991.

[6] H. Asri, H. Mousannif, H. A. Moatassime, and T. Noel, ''Using machine learning algorithms for breast cancer risk prediction and diagnosis,'' Procedia Comput. Sci., vol. 83, pp. 1064–1069, Jan. 2016.

[7] F. Petropoulos and S. Makridakis, ''Forecasting the novel coronavirus COVID-19,'' PLoS ONE, vol. 15, no. 3, Mar. 2020, Art. no. e0231236.

[8] 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, vol. 323, no. 16, p. 1545, Apr. 2020.

[9] Novel Corona Virus (COVID19) : Archive and API of India Data
https://thejeshgn.com/2020/03/16/novel-corona-virus-covid19-archive-api-india-data/[10] Johns Hopkins University Data Repository. Cssegisanddata.
Available: https://github.com/CSSEGISandData