Forecasting and Prediction of COVID-19 Severity Using Machine Learning and Time-Series Models

Sabin TT
Assistant Professor
ISE Dept
SJCIT
Chickballapur, India
sabintt@sjcit.ac.in

Dr. Sunitha BS
Professor
CSE Dept
PESITM
Shivamogga, India
sunitha74123@gmail.com

Abstract-It's great to see the use of data science and deep learning algorithms being applied to important real-world problems like identifying COVID-19 in medical images and forecasting the number of deaths due to COVID-19. The use of image segmentation with VGG16 classifier is a powerful technique to detect defects in scanned images, and the ability to classify CT images into COVID and non-COVID classes is a valuable tool for medical professionals in diagnosing the disease. The forecasting system described, which uses linear regression and Fbprophet models to predict the number of deaths due to COVID-19, is also a valuable contribution to the field. By providing accurate predictions of the impact of COVID-19 on lung health, governments and medical professionals can better prepare for and respond to the ongoing pandemic. Overall, the work described demonstrates the potential of data science to make a positive impact on society by providing insights and solutions to important real- world problems.

Key Words-Fbprophet, COVID-19, CT-Scan.

I. INTRODUCTION

The detailed explanation of deep learning and the VGG architecture, as well as the use of linear regression and Fbprophet for forecasting is as followes.

It's great to see how these tools can be used to solve real-world problems, and the advantages of using Fbprophet for time series forecasting, such as handling outliers, shifts in trend, and missing data.

In addition, the use of graphs and visualizations is important for making the results of these models more easily interpretable and understandable for end-users.

This can help to avoid confusion or misinterpretation, and ensure that the insights gained from these models are actionable and useful.

Overall, the explanation highlights the power and potential of deep learning and other machine learning algorithms for solving complex problems in a variety of domains, and the importance of using appropriate models and visualizations to ensure that the results are accurate, reliable, and understandable.

II. RELATED WORK

Asif, S., Wenhui, Y., Jin, H., & Jinhai, S. [4], The chest X-ray image classification system was created with the purpose of categorizing images into one of three classes: normal, pneumonia, and COVID-19. This system assists in the identification of COVID-19 patients by analyzing their chest X-ray images. Since a basic model was developed, the accuracy was low with anomalies. Yasin, R., Gouda,W [5], In the study, Chest X-Rays (CXRs) were utilized to forecast the specific lung disease type. The application of AI techniques enabled the automated detection of COVID-19 patients through their X-Ray images. Chest X-Ray images can be easily analyzed using CXRs, but the diagnostic process demands robust analysis parameters. Kwee, T.C., & Kwee, R. M. [6], The proposed image processing algorithm utilized by the system scrutinizes the lung's view and eliminates CT images with poor visibility. Chest CTs are conducted to detect COVID-19 pneumonia in patients and report. The discussion did not include any techniques for categorizing the different kinds of lung abnormalities. Guefrechi, S., Jabra, M. B., Ammar, A., Koubaa, A., & Hamam, H. [7], o identify COVID-19 through chest X-ray images, a deep learning system was employed for feature extraction. Three robust networks, namely ResNet50, InceptionV3, and VGG16, were fine-tuned using an augmented dataset. This dataset was created by gathering normal and COVID-19 chest X-ray images from various public databases. Despite utilizing sophisticated classification algorithms, X-Rays provide an incomplete depiction of the lungs' intricate structure, leading to the potential misclassification of classes.

III. CLASSIFICATION OF CT-SCAN IMAGES INTO COVID OR NON-COVID

The paper is using deep learning techniques to classify Covid and Non-Covid CT scanned images. VGG16 is a powerful CNN architecture that has been proven to perform well on a variety of computer vision tasks, including image classification. Pre- processing of the image dataset is

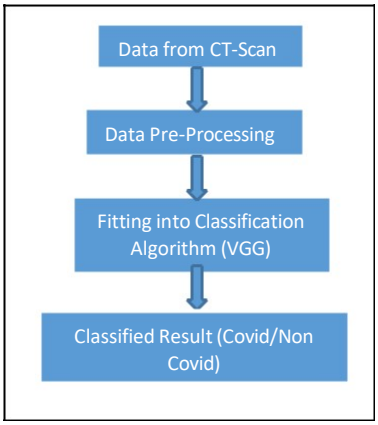

Fig 1 Classification of COVID or Non-COVID

It is also important to note that the performance of the deep learning model heavily depends on the quality and quantity of the dataset used to train it. It is recommended to use a large and diverse dataset for training, validation, and testing of the deep learning model to ensure its robustness and generalization to unseen data.

Normalization is an important step in pre-processing image data for machine learning models. It involves scaling the pixel values of an image so that they fall within a specific range, typically between 0 and 1 or - 1 and 1. This helps to ensure that the model can learn from the data effectively and can also improve the convergence of the training process. In the case of the COVID-19 classification model mentioned, encoding the RGB values as 8-bit integers is a form of normalization as it limits the range of possible pixel values. However, it is also important to consider other forms of normalization such as subtracting the mean pixel value or dividing by the standard deviation.

VGG16 is a powerful CNN architecture that is widely used in computer vision and image classification tasks due to its deep structure and the use of small convolution filters. The 8-bit integer encoding and resizing of image normalization are common techniques used to prepare image data for deep

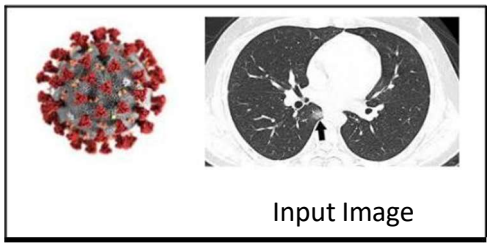

Fig 2. Classification result for COVID-19

Important to ensure that the images are properly prepared for input into the algorithm. Resizing and learning algorithms. Also help to improve the efficiency of the algorithm and ensure that the input images have consistent data distribution. By passing the input images through the various layers of the VGG16 model, the algorithm is able to extract features and make predictions about whether the image belongs to the COVID or Non COVID class. The accuracy of the model is an important metric that reflects how well it is able to classify the input images.

Snapshot of results obtained:

Table 1: Classification Report

VGG Classification Report for CT Scan				
	Precision	Recall	F1-Score	Support
0	0.93	0.93	0.93	79
1	0.95	0.96	0.95	80
Accuracy	-	-	0.94	150
Macro Average	0.94	0.94	0.94	150
Weighted Average	0.94	0.94	0.94	150
VGG Prediction Result				
99.99% COVID				

IV. FATALITY FORECASTING

While working on time series forecasting using the Fbprophet algorithm, the dataset one working with contains five attributes: date, Country/region, confirmed, recovered, and deaths. For prediction task, consider only the attributes of date and deaths.

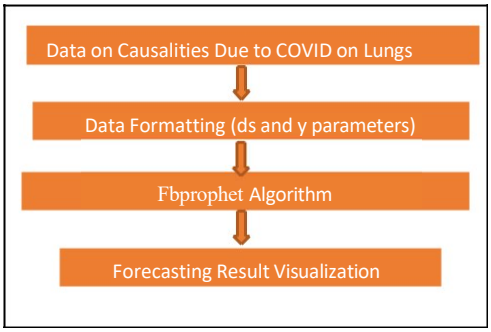

Fig 3. Forecasting of Fatility

To use the Fbprophet algorithm for forecasting, need to format date parameter as 'ds' and target variable

(deaths) as 'y'. Then create a data frame with future dates for prediction task, using the periods parameter to specify how many days into the future need to forecast. In this case, forecasting 365 days into the future.

Once predictions available, one can visualize them using a scatter plot. This will help to see how well model is performing and identify any potential trends or patterns in the data.

Overall, the work has a solid methodology for time series forecasting task. Just make sure to validate results and adjust model as needed based on the accuracy of predictions.

The working on forecasting COVID-19 fatality using linear regression. Dataset contains the date (formatted as DD-MM-YYYY) and death values as features.

In linear regression, there is a dependent variable (in this case, the fatality count) thatone want to predict based on an independent variable (the date in this case). Here using the scikit-learn library to fit data into a linear regression model, using the LinearRegression() function.

Once one have fit data into the model, then can visualize results using matplotlib. This will help to

The figure below shows forecasting using Linear Regression:

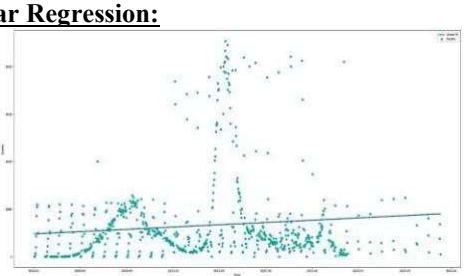

Fig 4 Forecasting using Linear Regression

To use Fbprophet, one need to format date parameter as 'ds' and target variable (death count) as 'y'. One also need to set the daily seasonality to true, as have daily death data. Additionally, one can set the period for forecasting as 200 days, which is the number of days need to forecast into the future.

Once have fit data into the Fbprophet model, one can use the predict() function to generate the forecasted values for the next 200 days. This will give a complete table of various useful parameters such as yhat_upper, yhat_lower, and trend.

To visualize the results, one can use the fbprophet.plot() function, which generates an easy-to-

The Fbprophet is based on decomposable time-series. The equation is as follows:

$$y(t)= g(t) + s(t) + h(t) + \varepsilon t$$

The three main components are:

g(t): represents the trend function; piecewise linear or logical growth to fit non-periodic changes in the value of the time series.

s(t): represents the seasonality that is periodic changes as a week and/or year seasonality. h(t): represents the holiday effects.

εt: represents any unusual change which is not accommodated by the model.

Better understand the relationship between the date and the fatality count, and to evaluate the accuracy of the model.

It's important to note that linear regression assumes a linear relationship between the independent and dependent variables. In real-world scenarios, this may not always be the case, and more complex models may be necessary to accurately forecast future values. Additionally, it's important to carefully consider any factors that may impact the accuracy of model, such as changes in testing or reporting protocols over time.

understand graph that displays the predicted values on top of the observed data points. Then can adjust the period values and zoom in or out as needed to better understand the results.

Overall, Fbprophet is a powerful algorithm for time series forecasting that can provide improved accuracy over linear regression and other simpler algorithms.

The figure below shows forecasting using Fbprophet:

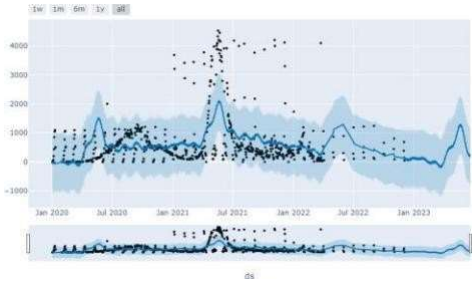

Fig 5: Visualization of Fbprophet

It's important to carefully consider data and any potential confounding factors when using Fbprophet or any other forecasting algorithm to ensure that the predictions are as accurate as possible.

V. CONCLUSION

It's great to see that deep learning models are being developed to aid in the diagnosis of COVID-19. By using CT scan images of the lungs, the model can automatically classify whether a patient has COVID-19 or not, saving time and resources for healthcare workers and providing an accurate diagnosis. This can help reduce the burden on healthcare systems and allow for more informed clinical decisions. In addition to the classification model, also proposed a forecasting model for COVID-19 to predict the confirmed and death cases of the disease. While the data acquired may be limited, the model has still been able to provide realistic predictions of the trajectory of the pandemic. With more data being generated over time, the accuracy of the model can be improved, which can be useful in forecasting peak values of the pandemic and developing strategies to flatten the curve and control the spread of the virus. With more data being generated over time, the accuracy of the model can be improved, which can be useful in forecasting peak values of the pandemic and developing strategies to flatten the curve and control the spread of the virus. Overall, the use of advanced technologies such as deep learning and forecasting algorithms can greatly assist in the fight against COVID-19. By providing accurate diagnosis and forecasting, healthcare workers can make better- informed decisions and take appropriate actions to prevent the spread of the virus and save lives. It's important to continue to explore new technologies and methods to combat this pandemic and future outbreaks.

REFERENCES

[1]. Maulud, D., & Abdulazeez, A. M. (2020). A review on linear regression comprehensive in machine learning. *Journal of Applied Science and Technology Trends*, *1*(4), 140-147.

[2]. Taylor SJ, Letham B.2017.Forecasting at scale.*PeerJ Preprints* 5:e3190v2 https://doi.org/10.7287/peerj.preprints.3190v2

[3]. Krishnaswamy Rangarajan, A., Purushothaman, R. Disease Classification in Eggplant Using Pre-trained VGG16 and MSVM. *Sci Rep* 10**,** 2322 (2020). https://doi.org/10.1038/s41598-020-59108-x

[4]. Asif, S., Wenhui, Y., Jin, H., & Jinhai, S. (2020) "Classification of COVID-19 from ChestX-ray images using Deep Convolutional Neural Network",2020IEEE6thInternational Conference on Computer and Communications (ICCC). Doi:10.1109/iccc51575.2020.934487.

[5]. Yasin, R., Gouda,W. "Chest X-ray findings monitoring COVID-19 disease course and severity", Egypt J Radiol Nucl Med51, 193(2020). https://doi.org/10.1186/s43055-020-00296-x.

[6]. Kwee, T. C., & Kwee, R. M. (2020). "Chest CT in COVID-19: What the Radiologist Needs to Know", RadioGraphics, 40(7), 1848-1865. doi:10.1148/rg.2020200159.

[7]. Guefrechi, S., Jabra, M. B., Ammar, A., Koubaa, A., & Hamam, H. (2021). "Deep learning based detection of COVID-19 from chest X-ray images". *Multimedia tools and applications*, *80*(21-23), 31803–31820. https://doi.org/10.1007/s11042-021-11192-5