



Detection of Novel Corona Virus Using Machine Learning and Image Recognition

Dhruv Garg | Saurabh Gautam

¹Department of Information Technology, Delhi, India

To Cite this Article

Dhruv Garg and Saurabh Gautam, "Detection of Novel Corona Virus Using Machine Learning and Image Recognition", *International Journal for Modern Trends in Science and Technology*, 6(12): 394-397, 2020.

Article Info

Received on 16-November-2020, Revised on 09-December-2020, Accepted on 12-December-2020, Published on 18-December-2020.

ABSTRACT

In the recent past whole of the world has come to a standstill due to a novel airborne virus. The airborne nature of this disease has made it highly contagious which has led to a great number of people being infected very fast. This requires a new method of testing that is faster and more precise. Machine Learning has allowed us to develop sophisticated self-learning models that can learn from data being fed and decide on entirely new options. In the past we have used different Machine Learning algorithm to make models on different biomedical dataset to detect various kind of acute or chronic diseases. Here we have developed a model that successfully detects severe cases of Novel corona virus affected person with great precision.

I. INTRODUCTION

This project uses the libraries of Machine learning in python to develop a model that can parse through hundreds of x-rays in fraction of seconds and provide result on how many of the people were infected with virus. The model was trained using biomedical dataset of the x-rays of infected patients. In the current scenario the x-ray of a patient has to analyzed by hand by a doctor which is slow and susceptible to various human error.

When looking at image interpretation by human experts, the quality is quite limited due to its variety, extensive variations exist across different interpreters, differentiability, complexity of the images, and fatigue. Checking each and every x-ray can consume a lot of time which is not the best practice when dealing with a patient on ventilator. As the mortality rate was just 2-3 percent during the treatment of

COVID-19, it was noticed that doctors were refraining from treating corona inflicted cases until extremely necessary. But when the numbers of cases reach in million, even a small percentage can mean a whole lot of patients.

Due to the complexity and overlapping of regions of a human organ in x-ray, it sometimes becomes difficult for doctors to judge the correct data within a stipulated time frame, that's why we required a system that is not only reliable but also swift in deciding the classification of patient based of available data. Here we explicitly state and discuss the findings and the result of the project. In this a few other paper on the use of Machine Learning and bio medical dataset to detect various disease have also been investigated and discussed.

II. RELATED WORKS

Over past a few time we have seen quite a few research and research papers on use of various models of machine learning to detect several acute or chronic diseases. The prime compliancy or different models was to provide least numbers of false negative while maintaining certain level of precision too.

Parkinson's disease

In the following paper the author defines the use of machine learning to detect Parkinson disease, in this disease there is a difference between core motor function of human body, so a special software is used and the patient is ask to draw certain images, then several images by healthy patients are also fed into the system to create a biomedical dataset then these drawings are analyzed after clearing noises by various machine learnings models like Optimum-Path Forest, Support Vector Machine remove, and Naive Bayes use the dataset to search and learn of the characteristics for the process of classifying individuals distributed into two classes: sick and healthy.

Image Based Cardiac Diagnosis

In the following paper the author defines the use of machine learning to detect cardiovascular disease. Cardiovascular imaging performs a pivotal role in diagnostic decision making in the case of cdc. Present image analysis techniques are mostly dependent on the qualitative visualized assessment of the images and the crude quantitative measures of the cardiac structure and function. If we want to optimize the diagnostic capacity of the cardiac imaging, we need to increase advancement of image analysis techniques that allows us deeper quantification of imaging phenotypes. We can achieve the same with the help of Machine Learning.

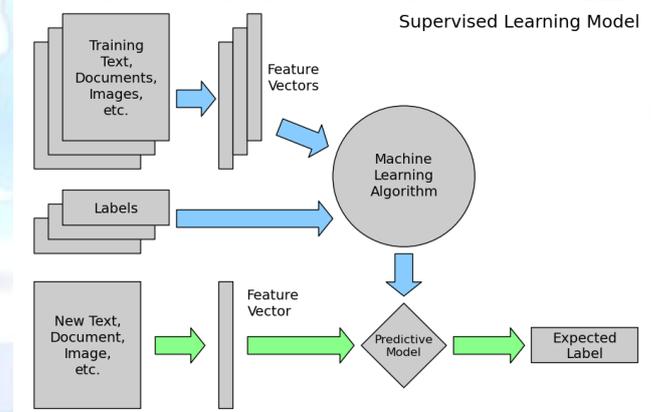
Alzheimer's disease

In the following paper the author defines the use of machine learning to detect Alzheimer in human beings. A person suffering from Alzheimer may start to loose memory and later loose complete memory, Loss of memory is also related with different brain damage or problems like brain tumor. So it becomes important to correctly differentiate between the two and all. Here author proposed a novel AD detection method by displacement field (DF) estimation between a normal brain and an AD brain. The DF was treated as the AD-related features,

reduced by principal component analysis (PCA), and finally fed into three classifiers: support vector machine (SVM), generalized eigenvalue proximal SVM (GEP SVM), and twin SVM (TSVM). The 10-fold cross validation repeated 50 times.

III. ANALYSIS

During the detection of Covid-19, blood test proved to be extremely slow and were only able to detect presence of virus in the body (or if virus was present at some time which may or may not have been cured by using Anti-gen test), also the test were not very accurate and provided with a lot of false positive. Here we tried to optimize the ever-expanding field of using Machine Learning to generate better results in the analysis of medical image-based records. Here we try to distinct cases of Covid-19 which might require better care at the doctor's end. We try to arrive at a model which can provide a better percentage of correct result for this medical problem At this time X-ray based Covid-19 detection model will not only be able to detect virus but also state that if the patient requires special care and ventilator. Here we tried various machine learning model to suite our requirement.



Our main objective while selecting the model was to select a model which would give us least amount of false negative. We used supervised learning method as our end result required us to classify our patient as infected or healthy only. We tried different model including Logistic Regression, Decision Tree, Support Vector Machine, Naïve Bayes and K nearest neighbor.

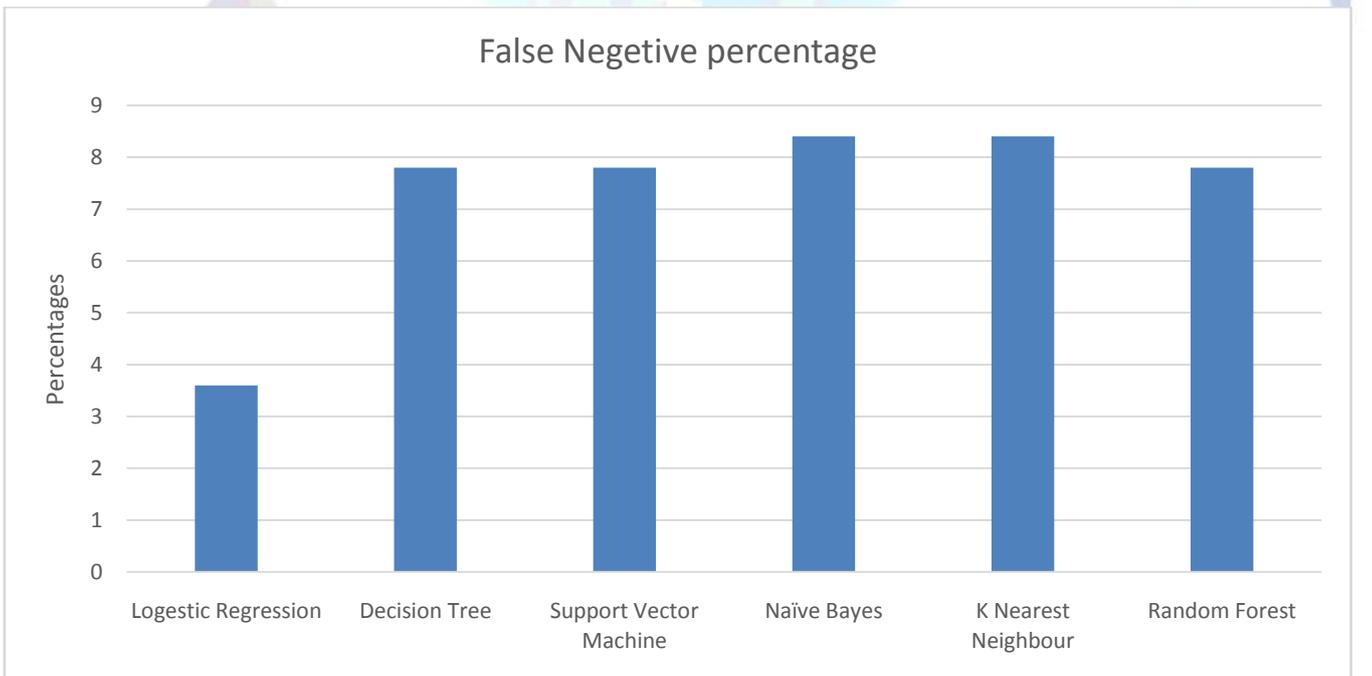
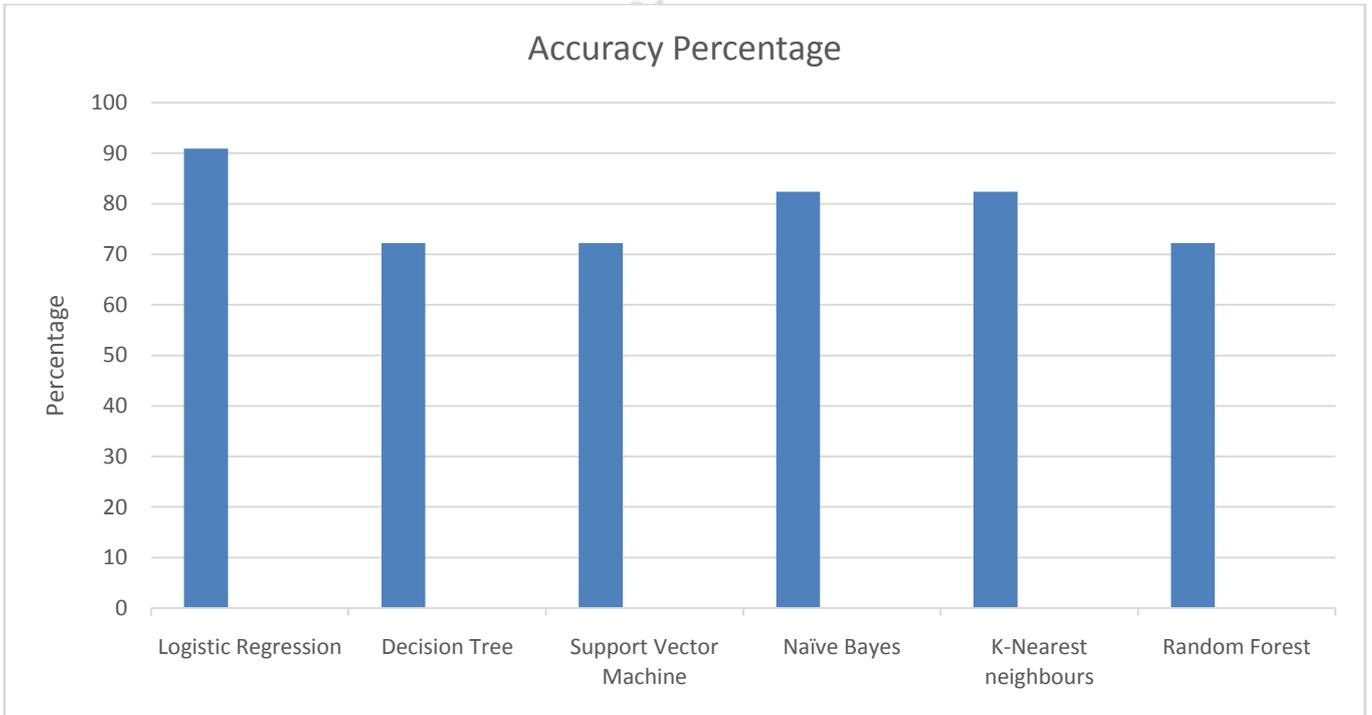
IV. RESULT

After applying various Machine Learning Models on our Biomedical Dataset we found out that Logistic Regression Model was able to give us the best result. Here our primary goal was to arrive

or rather select a model which would be able to provide us with least amount of false negative, which means that the model should be able not classify infected patients as not infected. On using various models we found that Logistic regression was able to provide us with a specificity of 90.9% with only 3% of the results being classified as false negative.

The various models result accuracy with confusion matrix for the given dataset were as follows.

Dataset 500 Images



	True Positive	False Positive	True Negative	False Negative
Logistic regression	85	9	65	6
Decision Tree	66	28	83	15
Support Vector Machine	66	28	83	15
Naïve Bayes	74	17	82	16
K nearest Neighbor	74	17	82	16
Random Forest	66	28	83	15

V. CONCLUSION AND FUTURE SCOPE

It is clear that the use of Machine Learning very useful when one is tasked with producing accurate decisions based on complicated data sets. But they come with some significant challenges and limitations that you either have to accept or try to overcome. When deploying any machine learning model, one is instantly faced with challenges related to data access, privacy issues, data protection, and more. And when the data is someone's personal data like here the question of security becomes really important. Though here we were able to achieve about 91% of accuracy there is still scope for improvement especially with new upcoming technologies like Neural Network and Deep Learning.

So it becomes important at our part as developer to continue to create new systems as well as optimize the older ones.

REFERENCES

- [1] By yudong Chang, shuhiua wang in 2015 "Detection of Alzheimer's disease by machine learning". PMID: 26401461PMCID: PMC4579022DOI: 10.7717/peerj.1251
- [2] By Lucas s Bernado Angeles Quezada Robert Munoz Fernanda Martinis in 2017 "Handwritten pattern recognition for early Parkinson's disease diagnosis".
- [3] By Carlos Martin, Victor M campello, Cristian Izquierdo. Image-Based Cardiac Diagnosis With Machine Learning in 2017, Research Gate.
- [4] Hu, Gang & Antikainen, Riitta & Jousilahti, Pekka & Kivipelto, M. & Tuomilehto, J.. (2009). PARKINSON DISEASE Reply. Neurology