ANG WEE HWA 20/5/2021

Project Proposal (Predictive Model for Covid CTScan)

DS106

ANG WEE HWA 20/5/2021

Why?

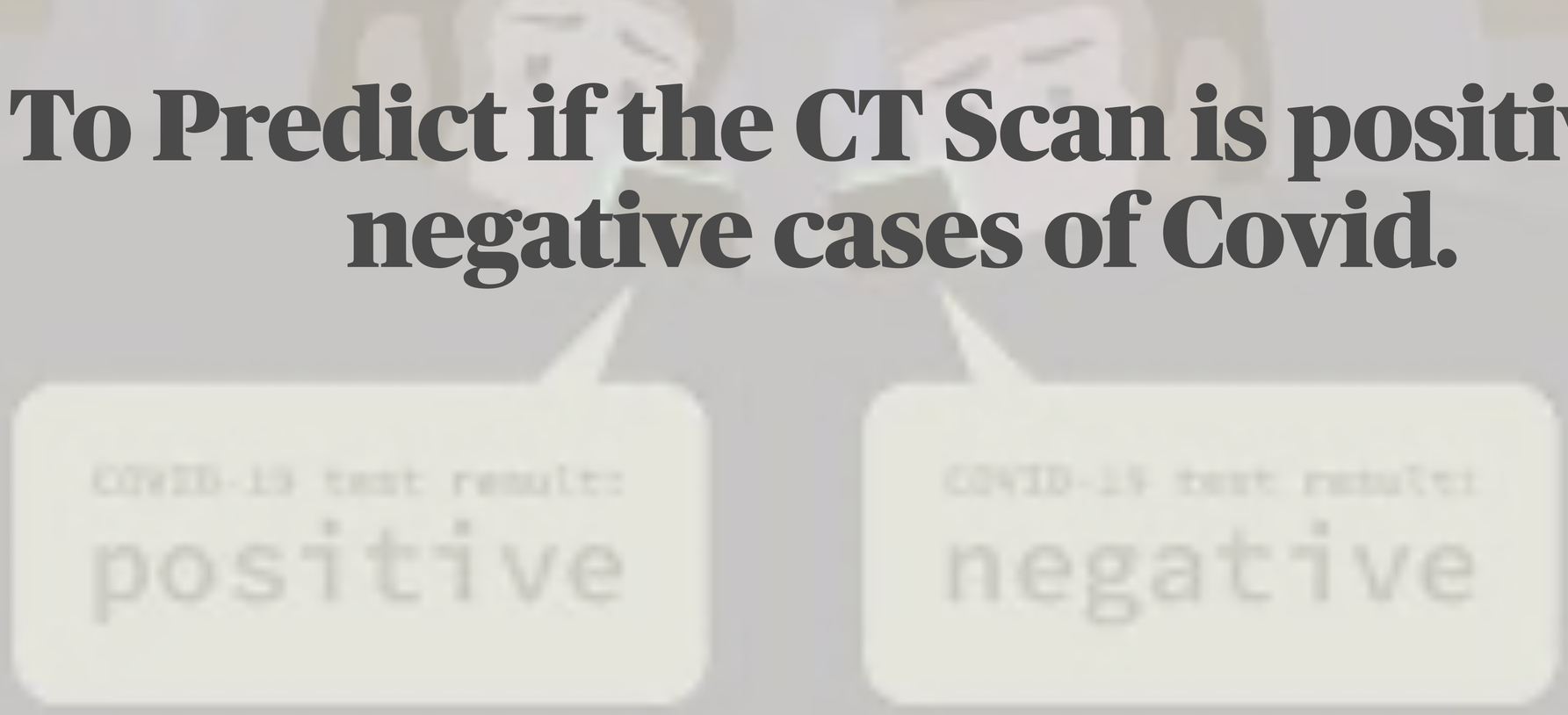
- Covid19 had been a issue with the world, as the virus had affected varies countries around the global. Businesses and individuals are losing customers as governments are issuing stricter measures to counter and prevent the spread of the virus.

What do we want to achieve?

- To Create a model to predict if the CTScan image is positive or negative case for Covid

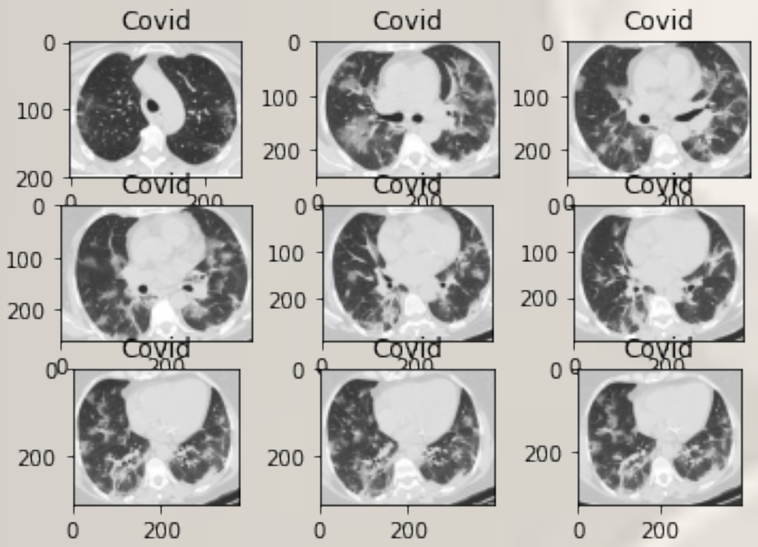
Problem Statement

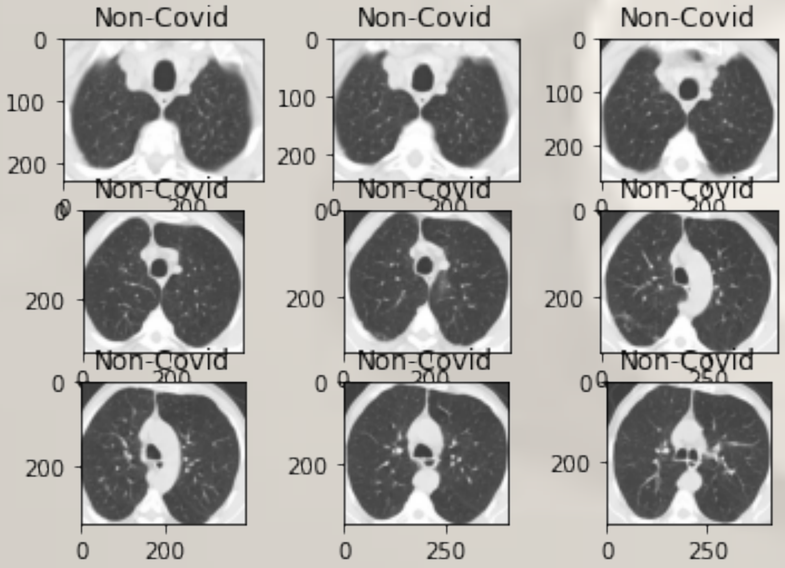
To Predict if the CT Scan is positive or negative cases of Covid.



ANG WEE HWA 20/5/2021

Dataset





FOR THIS PREDICTIVE MODEL WE DOWNLOADED THE DATA FROM KAGGLE
REF: HTTPS://WWW.KAGGLE.COM/C/COVIDCT

WE HAVE A TOTAL OF:
COVID CT SCAN IMAGES: 1252 (PNG)
NON-COVID CT SCAN IMAGES: 1230 (PNG)

WE WILL DIVIDE THEM INTO 3 FOLDERS MANUALLY:

1) TRAIN: (TO TRAIN THE MODEL)
COVID: 992 IMAGES
NON-COVID: 975 IMAGES

2) VAL: (TO VALIDATE THE MODEL AFTER EACH EPOCH)
COVID: 248 IMAGES
NON-COVID: 244 IMAGES

3) TEST: (TO MANUALLY TEST AND OBSERVE THE RESULT)
COVID : 12 IMAGES
NON-COVID : 10 IMAGES

Steps

WE WILL BREAK INTO 3 PARTS FOR THIS PROJECT NAMELY:

PART A : OBSERVE THE CT SCAN AND TO DIFFERENTIATE THE DIFFERENCE VISUALLY.

PART B : WE WILL TEST OUR 3 MODELS NAMELY:

1) VGG16

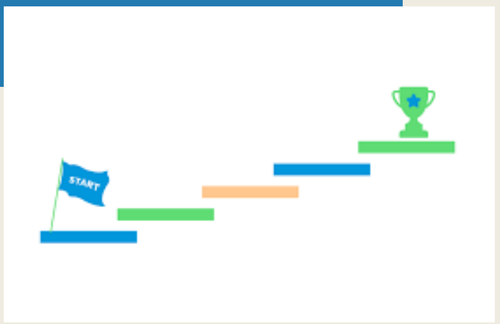
2) INCEPTIONV3

3) RESNET50

AND SELECT THE BEST ONE FOR HYPER PARAMETER TUNING.

PART C : WE WILL USE THE BEST MODEL TO FINE TUNE THE TRAINING RATE AND USE THE BEST MODEL.

TO AVOID BIAS, WE WILL BE USING 50 EPOCH TO TRAIN ALL THE MODELS AND CONFUSION METRIC FOR PREDICTION OF THE TEST DATASET. ACCURACY AND VALIDITY ACCURACY ARE THE MEASUREMENTS FOR THE PERFORMANCE OF THE MODEL.



Difficulties

- From the dataset we might be able to spot the difference however with my limited medical knowledge it might not be the only factor or feature
- There is only CT Scan of positive and negative omitting the other types of lung diseases which might cause the image to change hence unable to get a complete model

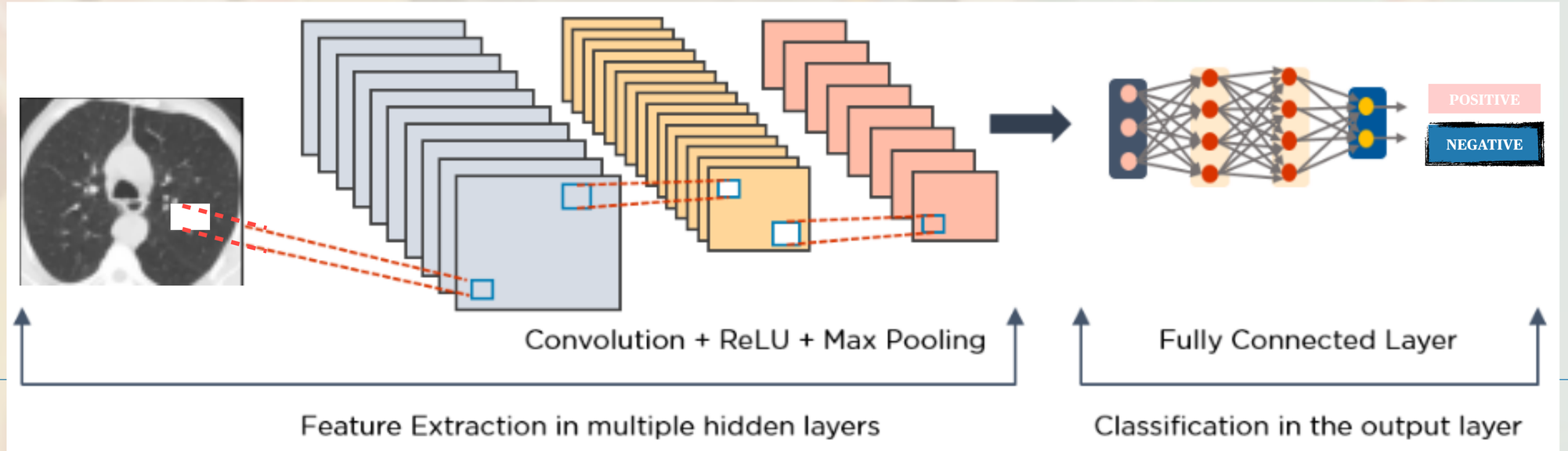
Machine Learning Model

For this problem statement I will be using deep learning method Convolutional neural network(CNN)

Will prepare and fit the respective model

- VGG 16
- ResNet50
- Inception V3

After comparing the performance of the 3 models will select the one with the best prediction to do hyper parameter tuning



“

The doctor of the future will give no medicine but will interest his patients in the care of the human frame, in diet and in the cause and prevention of disease.

- Thomas Edison