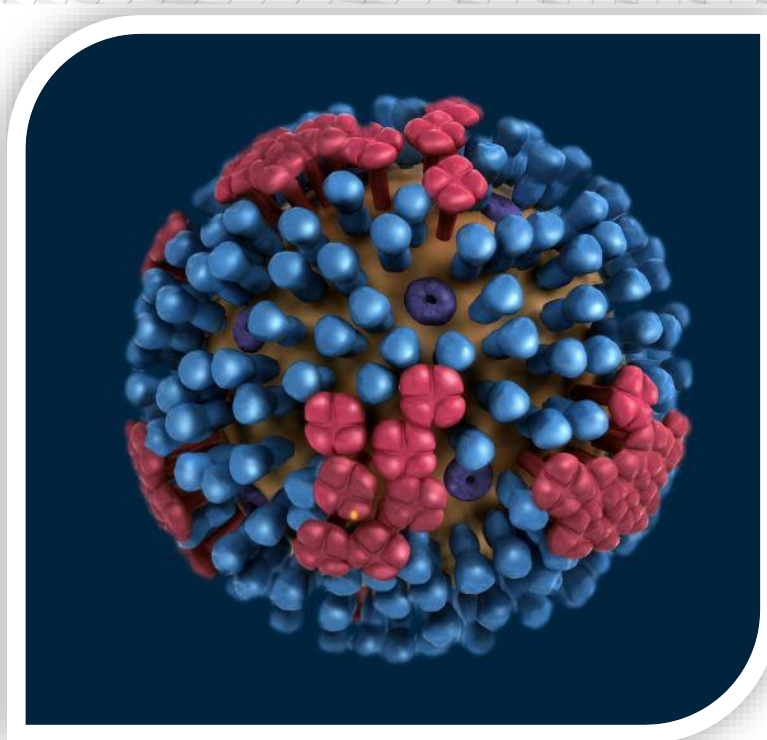


CT Scan Image Classification



Name: **Piyush Sanjay More**

Date: 20 Aug 2023

Course title:- **Full Stack Data
Science Program**

INSTRUCTIONS

In order to run this project, we need to follow some steps:

- Clone or download the project from this **Github** repository.
- Download the image data Zip file from here and save it to the root folder of the project.
- Download the image data **Zip** file from here and save it to the root folder of the project.



APPROACH TO THE PROBLEM

1. Data Preprocessing:

- Downloaded the data.zip file and extracted the COVID and Non-COVID directories.
- Loaded the images into list.
- Resized the image into fixed sized of 224 X 224 and then converting images to RGB mode.
- Normalized the pixel values of images.

2. Data Augmentation:

- Applied data augmentation techniques on images using Keras preprocessing 'ImageDataGenerator' function.
- Included rotation, shifting, shear, zoom, and flipping to increase data diversity.

3. Data Generator & Splitting the data:

- Combining COVID and Non-COVID data with corresponding labels.
- Splitting the data into Training and validation sets.

4. Model Building:

- We used ResNet50 Model architecture which offers better performance and capacity to capture complex features.
- Adding custom layers like Average pooling 2D layer, Dense Layers
- Freezing the layers of the ResNet50 base model to avoid overwriting learned features.

- Compiled the model using binary cross-entropy, Adam optimizer and performance metrics as accuracy.

5. Model Training:

- Defined early stopping and model checkpoint callbacks.
- Trained the model using fit method with data generators for training and validation sets.
- Saved the best model with 65.99% accuracy on validation set.

6. Model Evaluation and Prediction:

- Loaded the best saved model.
- Evaluated model using metrics like accuracy, precision, recall, F1-Score, and confusion matrix on validation set.
- Loaded a test image, preprocessed it and used model to predict the class whether COVID or Non-COVID of the test image.

7. Fine Tuning the best model:

- Firstly, we fine-tuned the model by unfreezing the last 20 layers of pretrained ResNet50 model and running for 60 epochs.
- Then again, we adjusted to 80 epochs and by unfreezing last 40 layers.
- Both times the model resulted with 77.86%, i.e almost 18% increase in accuracy of ResNet50 model.