Project Proposal

Background:

The covid-19 virus has universally impacted every person throughout the world. With 2 million confirmed cases globally and 144,000 deaths, covid-19 is no doubt the biggest global health crisis this decade. Healthcare workers are especially in danger by being on the frontline and interacting with patients with the lack of personal protective equipment. This problem is especially distressing due to how infectious the virus is. As a result, there is a drastic need for a non-social diagnostic method that would help health care workers in patients.

Goal:

The goal of our project is to generate a model that can differentiate between healthy and covid-19 infected lung X-rays with the highest accuracy. With Covid-19 having a striking impact on the lungs, we hope to create models that can learn these features and differentiate between healthy and sick lungs.

Proposed approach:

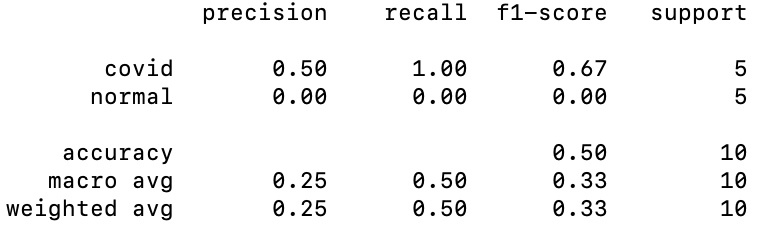
In this project, a 9-layer convolutional neural network (CNN) with dropout, batch normalization and data augmentation will be constructed. Transfer learning will be applied by importing trained VGG and Densenet to obtain a better result.

Dataset:

The dataset used in our project includes Covid-19 image data collection and RSNA Pneumonia Detection Challenge Dataset. 201 chest X-ray images of suspected Covid-19 patients, 155 chest X-ray images of sick patients confirmed with Covid-19 and over 10000 healthy chest X-ray images are used.

Initial Results of Baseline:

For our baseline model we decided to utilize the ResNet architecture. After training the model on our Covid-19 dataset, our model was able to predict images with Covid-19 with f1 score of 67%. In addition the precision of the model on Covid-19 data was 50%. On the other hand, our model had entirely predicted normal images to have Covid-19. The accuracy of our model on test data is 50%. There is much that can be improved upon as this was our baseline run.



Timeline of key milestones

|  |  |  |
| --- | --- | --- |
| Milestone | Date | Status |
| Setup environment to develop models | 4/12 | Completed |
| Familiarize self with data | 4/14 | Completed |
| Create and Train baseline model to obtain a baseline statistic(ResNet) | 4/16 | Completed |
| Train models for all those suggested in the proposal | 4/26 | In-Progress |
| Optimize data to improve output | 5/1 | To Do |
| Optimize models to improve output | 5/7 | To Do |