MAIS 202 - PROJECT DELIVERABLE 2

1. Problem statement

I will be implementing a model that classify X-Ray scans from patients with pneumonia. I will use the dataset from Kaggle. <https://www.kaggle.com/paultimothymooney/chest-xray-pneumonia>

2. Data Preprocessing

Firstly, I read the images from 3 folders (Train, Val, Test). Each of them contains chest images of people who have pneumonia and people who don’t. Since the input dimensions of the images are too big, I applied image augmentation and changed the size of the images. Finally, I plotted the images corresponding to their labels of NORMAL and PNEUMONIA.

A screenshot of a social media post

Description automatically generated

3. Machine learning model & Preliminary results

I trained the dataset by creating a CNN model in KERAS.

A screenshot of a social media post

Description automatically generated

A screenshot of a social media post

Description automatically generated

If the accuracy of the validation set is similar to the accuracy of the training set, my model is good. However, if it is much lower than the training set, it means my model is overfitting. So I plotted the graph of the model performance to improve the accuracy.

A screenshot of a map

Description automatically generated

A picture containing screenshot, map

Description automatically generated

4. Next steps

- code a webapp

- try to implement a better model