

A decorative graphic on the left side of the slide consisting of two overlapping parallelograms. The front one is blue and the back one is a light green. They are positioned diagonally, with the blue one partially covering the green one.

X-Ray Image Classification

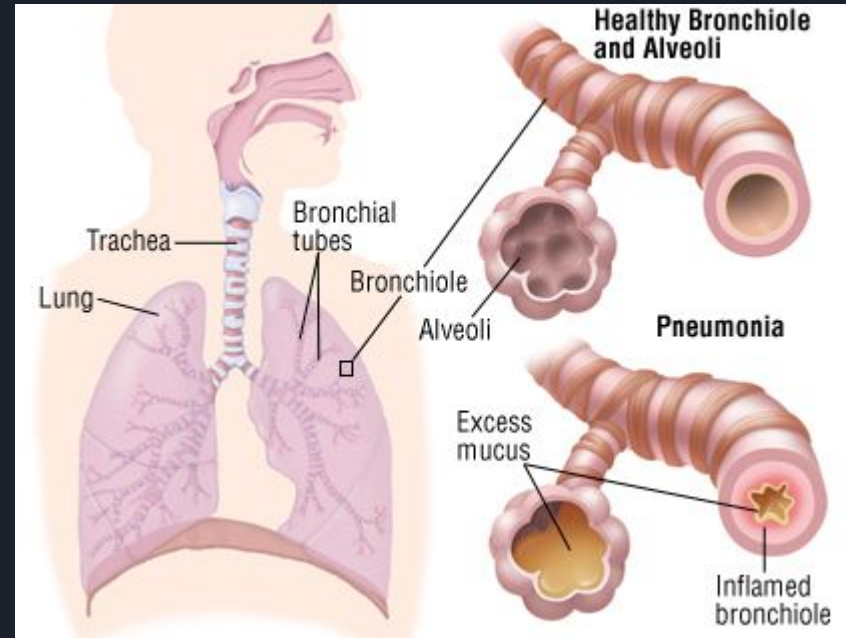
Eric Cusick and Trevor Mott

Business Case

A hospital reach out to a team of data scientists to create a learning model that would look at chest X-ray to predict whether or not the patient has pneumonia.

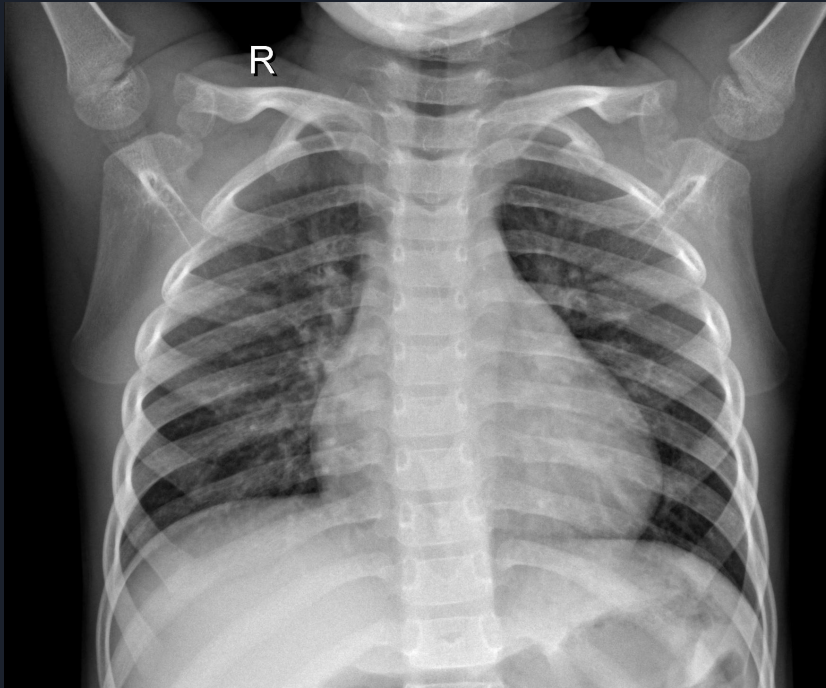
Given the dataset from [Kaggle](#), the team are expected to:

- Create an image classification model using Convolutional Neural Network (CNN)
- High accuracy

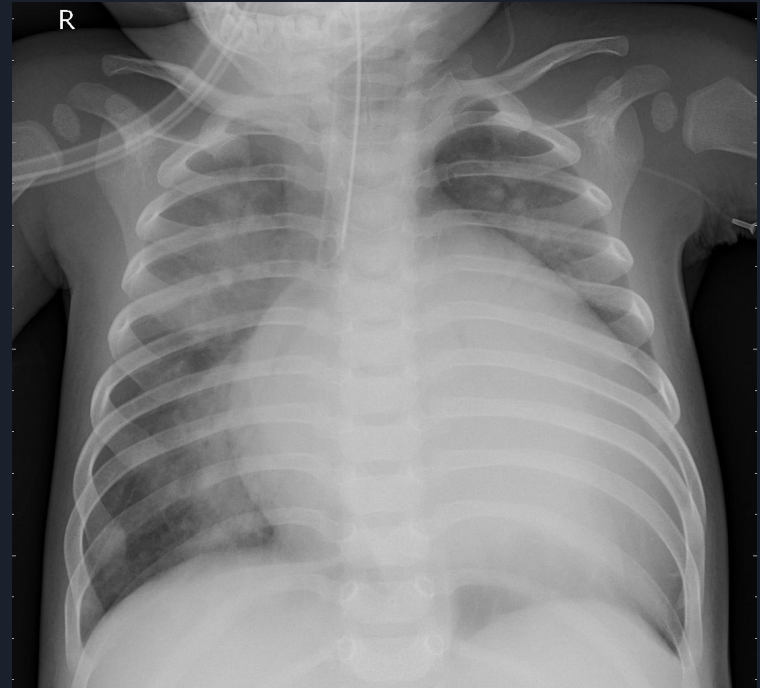


Exploratory Data Analysis

Normal



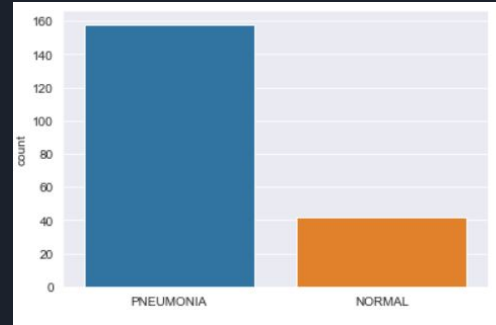
Pneumonia



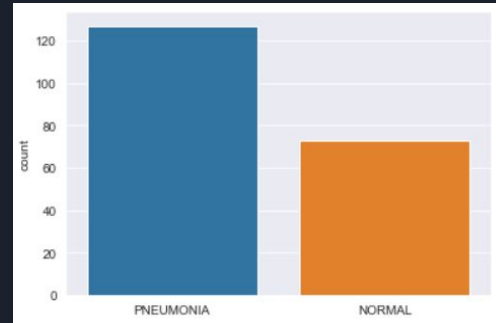
Exploratory Data Analysis

Upon further inspection of the datasets we found that there is a severe class imbalance within the train set and test set. Where the images of pneumonia is over saturated in comparison to the normal cases. Which can result in skew result from the image classification models.

Train Set

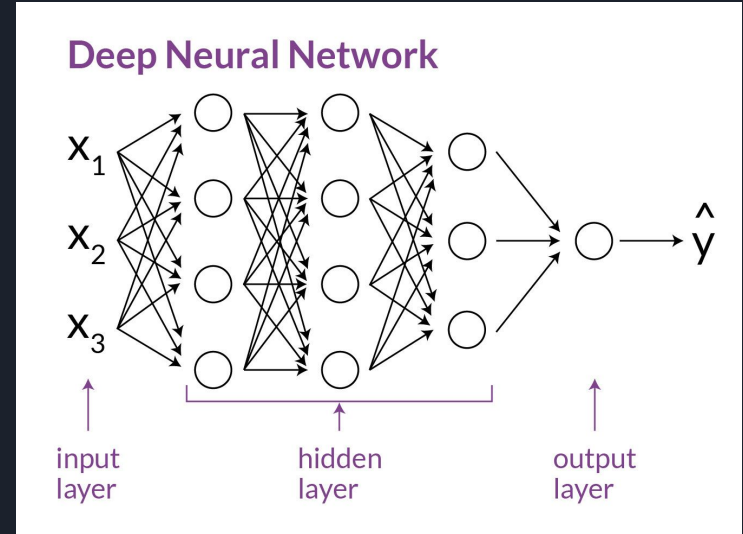


Test Set

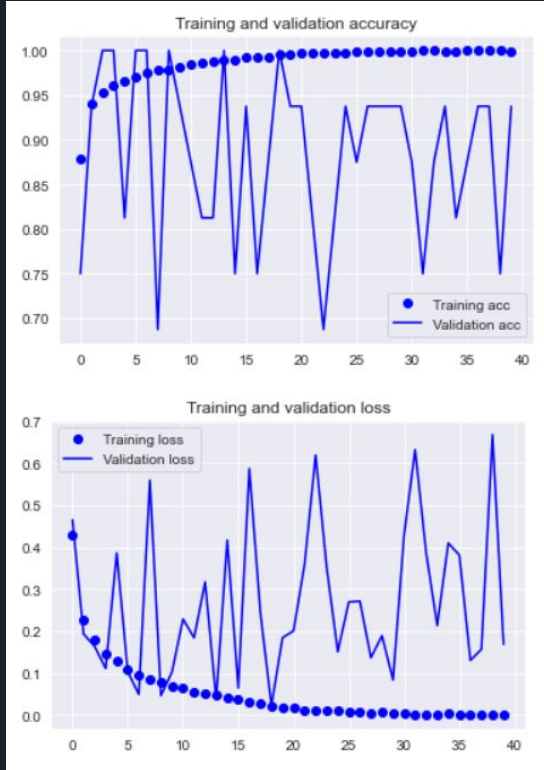


Modeling Process for CNN

- Convert all X-ray images to grayscale
- Resized all of the images to 128 X 128 pixels
- Created batch size for each datasets
- Gave more 'weight' to the Normal cases to balance the class imbalance
- Created multiple layers for the neural networks



Final CNN Model



Final model results:

- Training Accuracy: 99.56%
- Training Loss: 0.0191
- Validation Accuracy: 93.75%
- Validation Loss: 0.1840

Recommendations

Looking at the model, we would recommend:

We would recommend radiologist to use this as a supplementary tool to speed up the diagnostic process and as a secondary opinion.



Future Work

- Get more data for the validation set to gain a better/realistic model accuracy.
- Trying/explore different parameters and layers for the neural network modeling.
- Create more unique images by augmenting some of given data to give the machine more data to train from





Thank You

Thank you for giving us your time during the presentation today

Kaggle's Chest X-ray Dataset:

<https://www.kaggle.com/paultimothymooney/chest-xray-pneumonia>

Github Repo:

[https://github.com/Ericusick/Chest-X-Ray-Image Classification](https://github.com/Ericusick/Chest-X-Ray-Image-Classification)