June 2020

Mod 4 Project - Pneumonia Detection

By Chamila Dharmawardhana

Vidya Menon

About Pneumonia

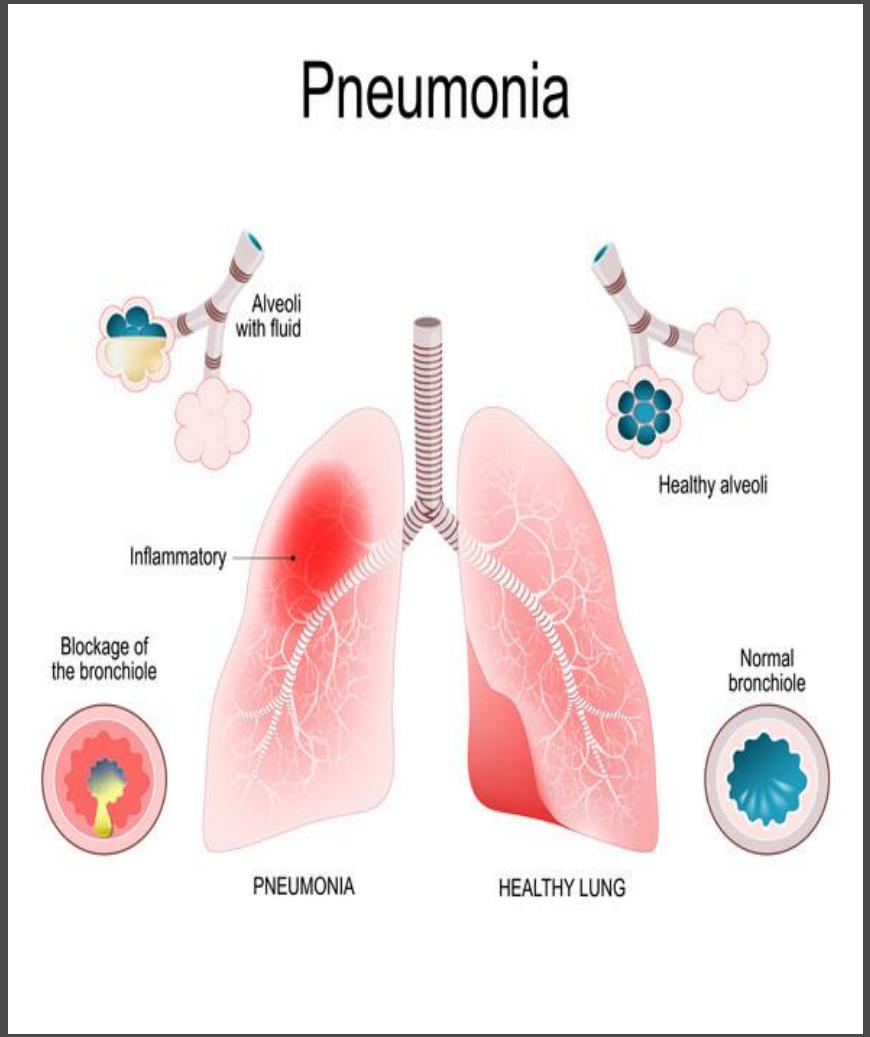
Key Facts:

- Pneumonia is an infection of the lungs that can cause mild to severe illness in people of all ages.
- Viruses, bacteria, and fungi can all cause pneumonia.
- Common signs of pneumonia can include cough, fever, and trouble breathing.

Who can get Pneumonia:

Certain people are more likely to become ill with Pneumonia: adults 65 or older, children younger than 5 years old; people who have ongoing medical conditions (like asthma, diabetes or heart disease); and people who smoke cigarettes.

In the United States, more than 250,000 people have to seek care in a hospital due to pneumonia each year. Unfortunately, about 50,000 people die from the disease each year in the United States. Most of the people affected by pneumonia in the United States are adults.



The Challenge

Build an algorithm to identify whether a patient is suffering from pneumonia or not, by looking at chest X-ray images.

The algorithm must be extremely accurate because lives of people is at stake.

Methodology

ROSEMED Methodology - This is the one of the most straightforward of the Data Science processes.

Research - Find out about various models and see which model works best for our data.

Obtain - Understanding stakeholder requirements, gathering information on the problem, and finally, sourcing data that we think will be necessary for solving this problem.

Scrub - Focus on preprocessing our data. Important steps such as identifying and removing null values, dealing with outliers, normalizing data, and feature engineering/feature selection are handled around this stage.

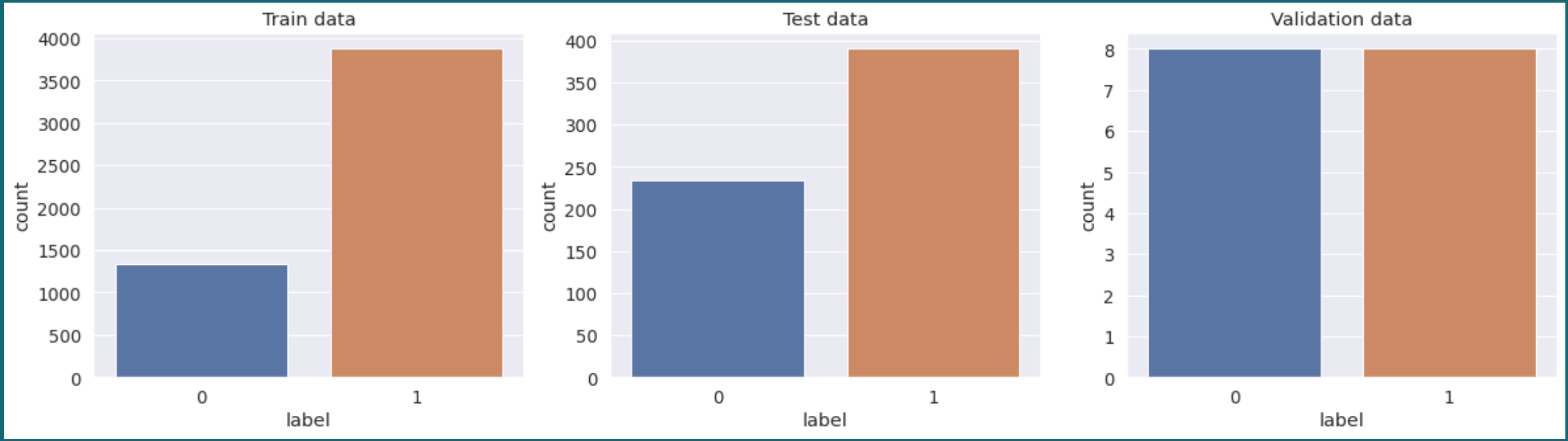
Explore - We create visualizations to really get a feel of the dataset. We focus on things such as understanding the distribution of different columns, checking for multicollinearity, and other tasks like that.

Model - It consists of building and tuning models using all the tools we have in our data science toolbox.

Evaluate - Interpret the results of models and communicate results to stakeholders.

Deploy - Deploying the Model.

Our Data



In the Train Set: Normal Images: 1341, Pneumonia Images: 3875
In the Test Set: Normal Images: 234, Pneumonia Images: 390
In the Val Set: Normal Images: 8, Pneumonia Images: 8

Solution

With Training on about 5216 Chest X-Rays: having Normal Images: 1341 and Pneumonia Images: 3875 and Testing on Normal Images: 234, Pneumonia Images: 390

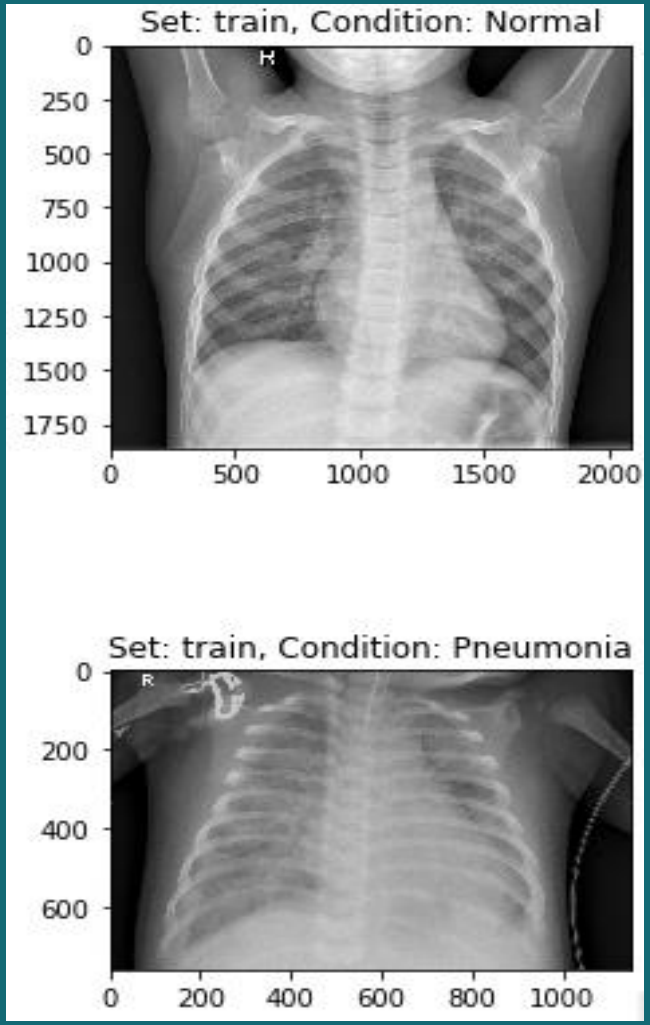
- **Our Model Summary :**

Model	Loss	Accuracy	Validation Loss	Validation Accuracy
Base Model	0.27	0.94	0.89	0.73
Improved Model	0.21	0.92	0.29	0.89

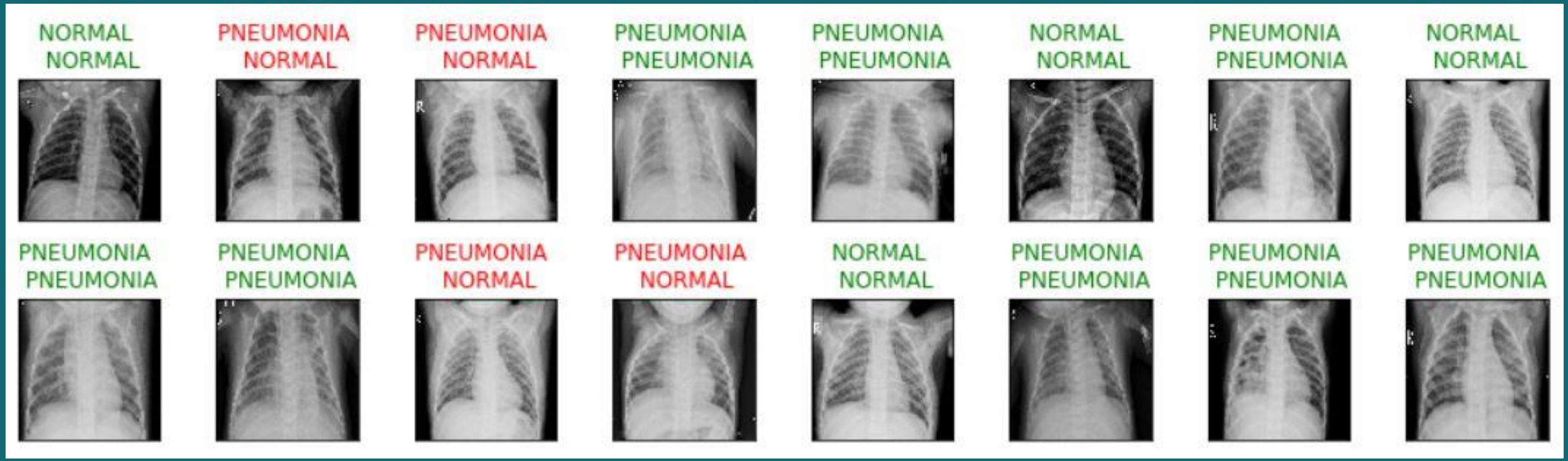
Precision: 91%

Recall: 90%

For Our model Recall is more important.



Validation Dataset



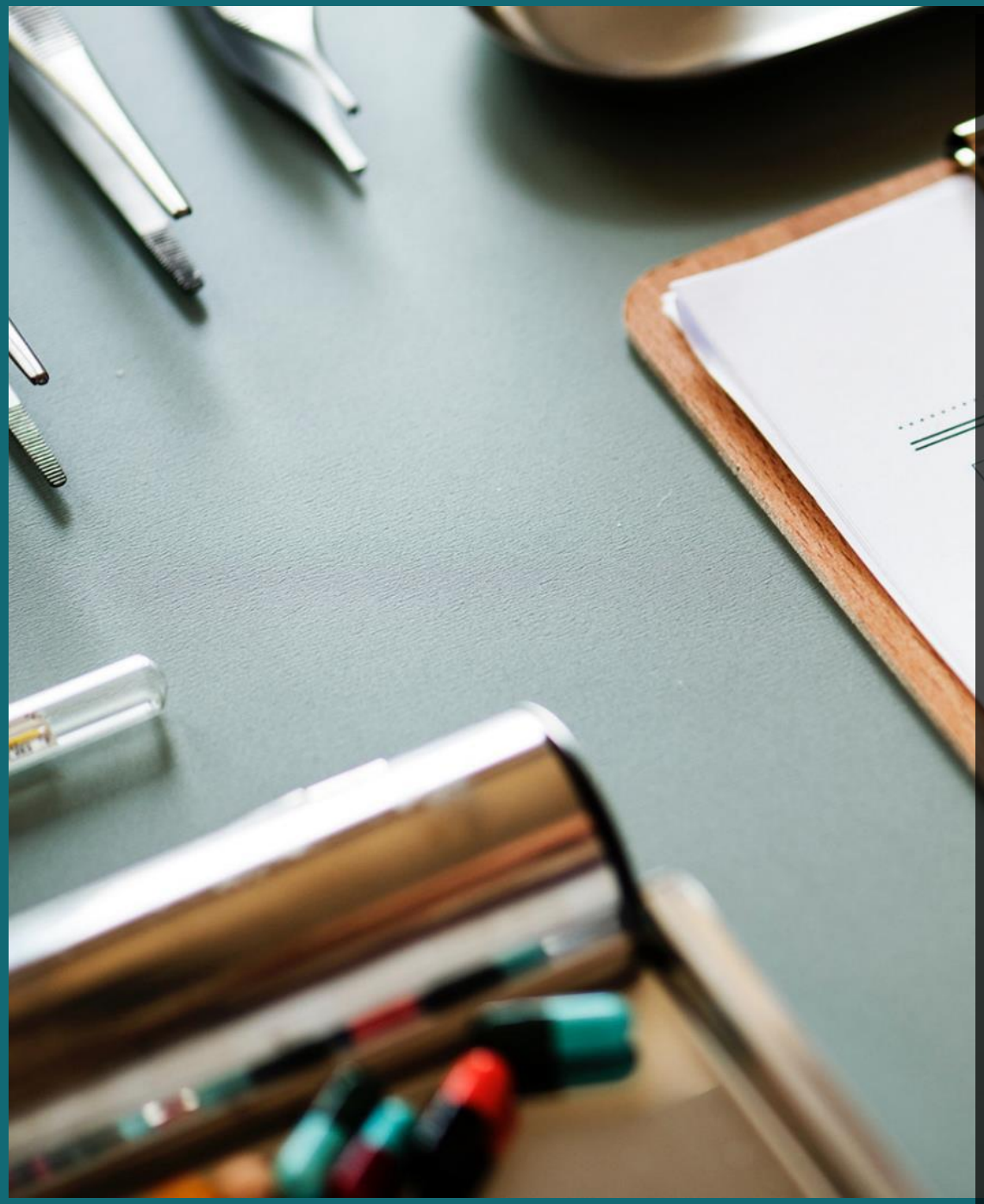
Our Model for the Validation Set has about :

- Accuracy: 75%
- Precision: 67%
- Recall: 100%

Conclusion

This study presents a deep Convolutional Neural Network (CNN) based approach for the automatic detection of Pneumonia. We constructed a CNN model from scratch to extract features from a given Chest X-Ray image and classify it to determine if a person is infected with Pneumonia.

We have demonstrated how to distinguish between Normal and Pneumonia Chest X-Rays with our model having an Accuracy of 89% and a Recall of 90%.

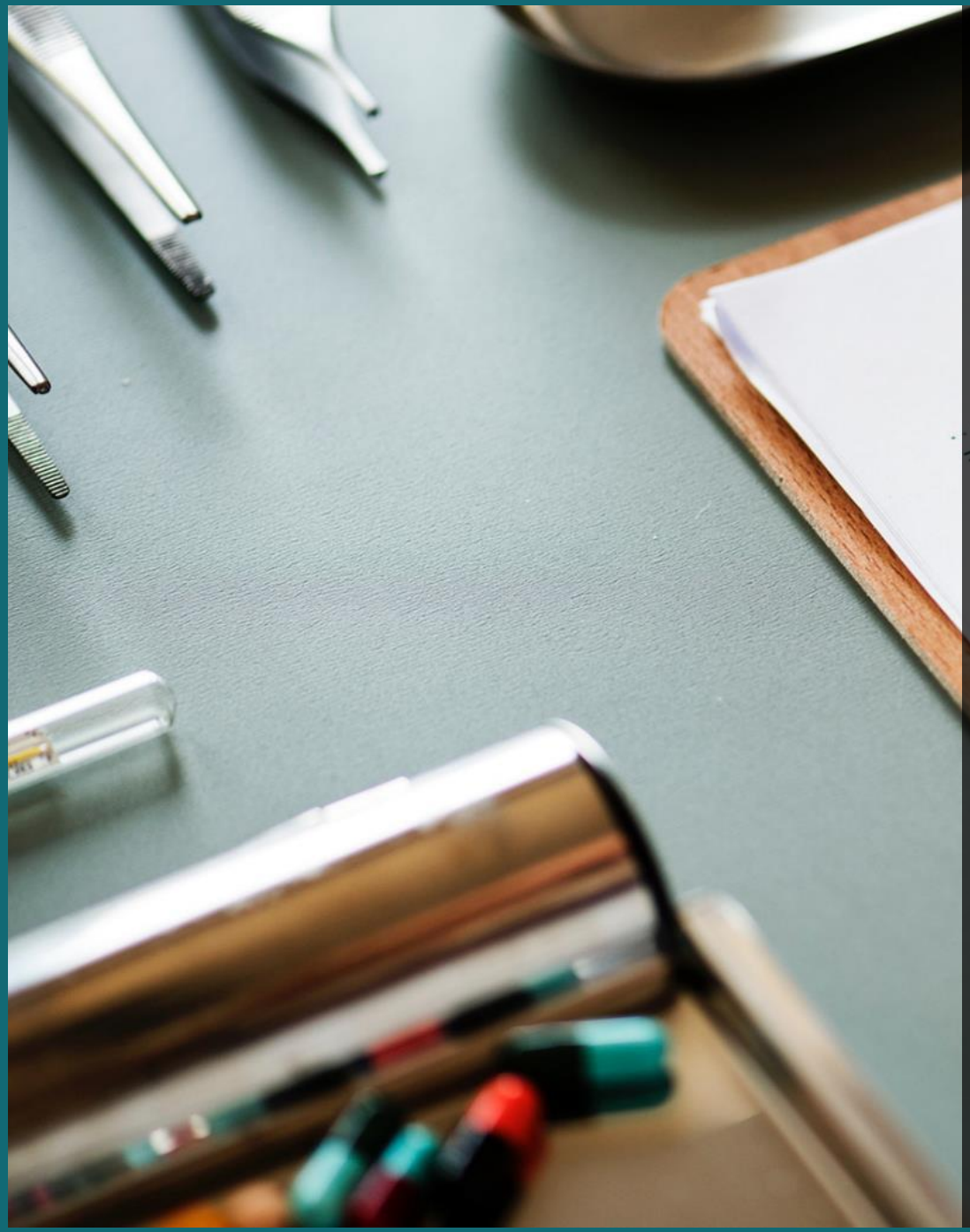


Recommendations:

Incorporate our model to see how it works in hospitals so that it can assist health professionals diagnose patients with Pneumonia.

Of course these reports must be validated. This is not the ultimate test. This needs to be certified by health professionals.

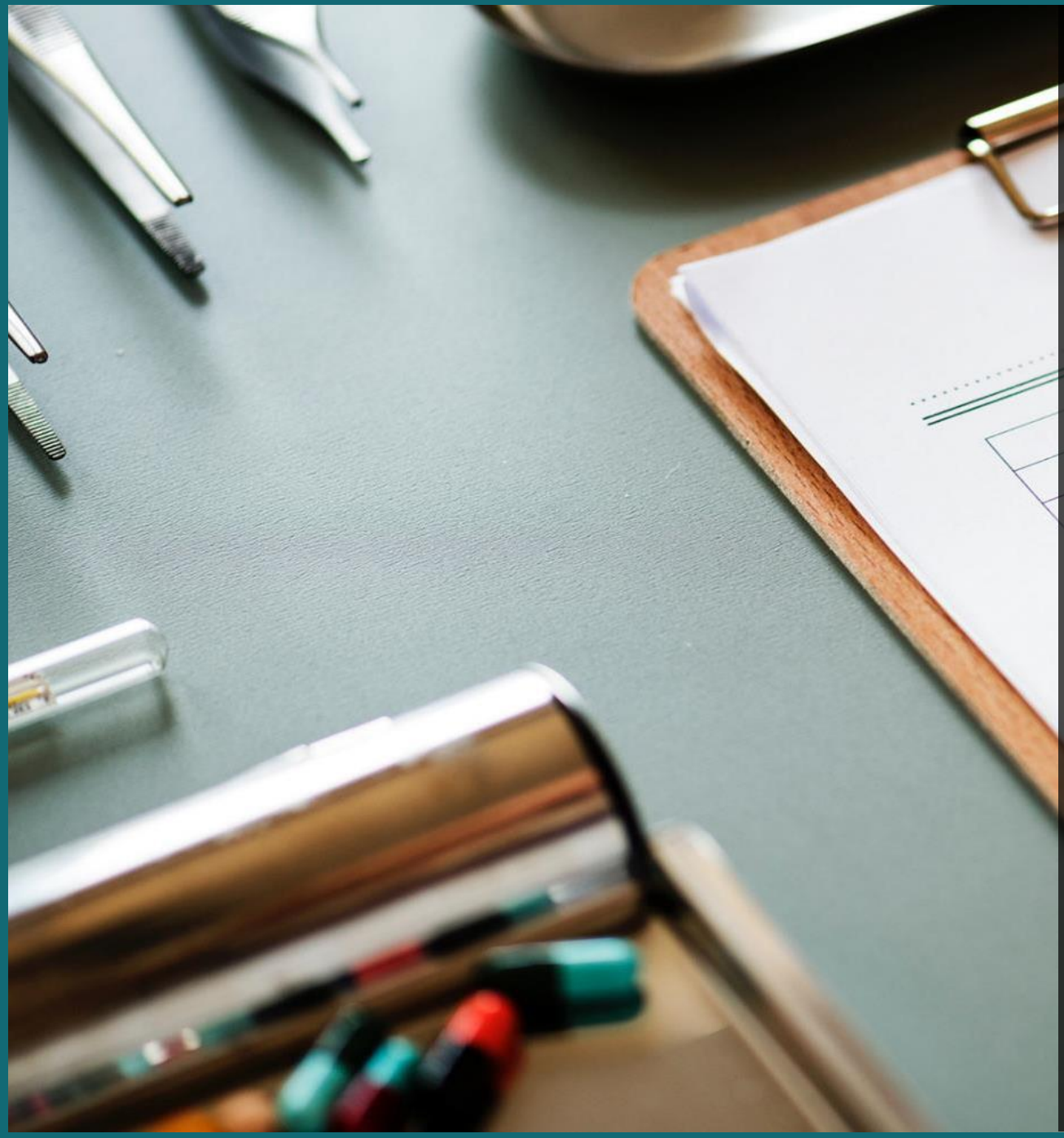
This model should be run under the supervision of a radiologist to enhance accuracy/recall to improve treatment outcomes which will increase hospitals ratings and fundings.



Future Work

We need more data to run our validation set on so that we can be sure of the way the model is predicting.

Also, work towards getting our Transfer Learning model to work better.



Thank You

- Amber Yandow
- Ben Jacobson
- Luay Matalka
- Jesse Neumann

Any Questions ☺ ???