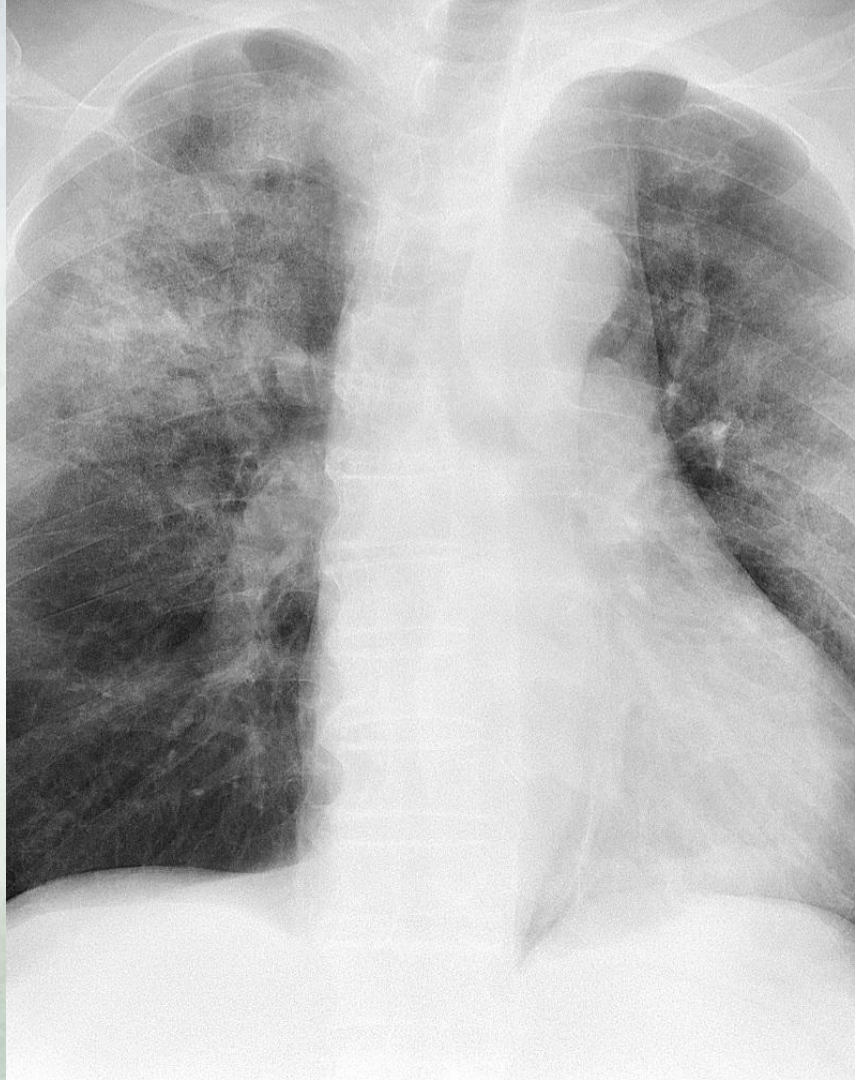
Detecting Pneumonia in Chest X-Rays

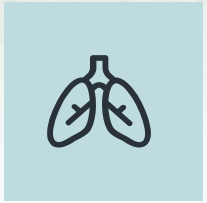
By: Cassie Nutter
March 2021

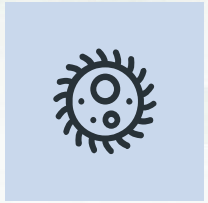


About Pneumonia

What is Pneumonia?

An infection that inflames the air sacs in one or both lungs and can cause lungs to fill with fluid





What causes it?

Many organisms, including bacteria, viruses and fungi

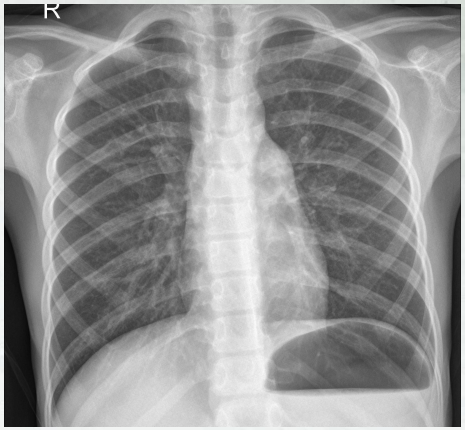
Is it contagious?

YES! Please cover your coughs and sneezes and wash your hands.

Is it deadly?

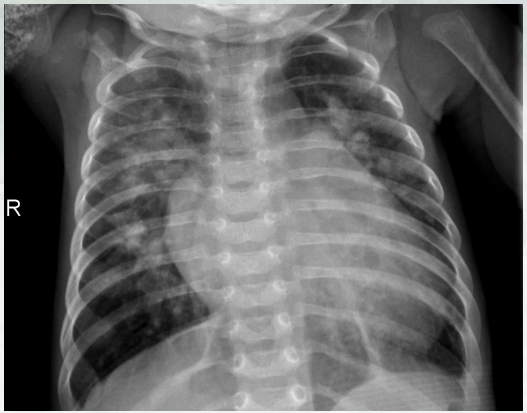
It can be. Especially for those with weakened immune systems

Pneumonia in Radiographs



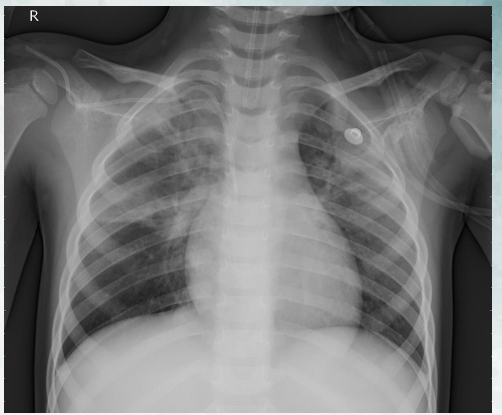
Normal

Clear lungs, no abnormal opacification



Bacterial

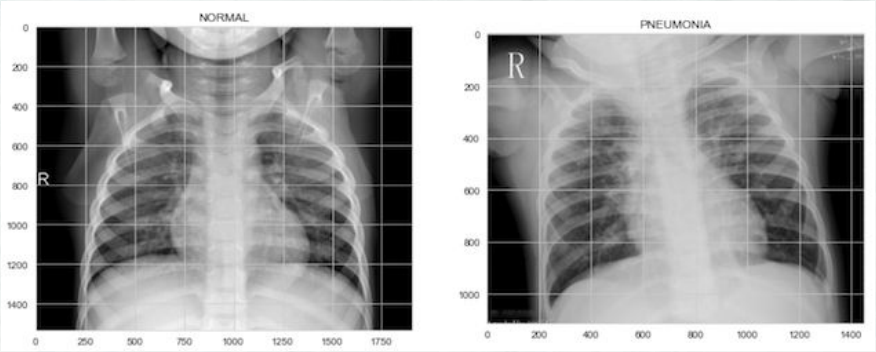
Focal lobar consolidation



Viral

Diffuse “interstitial” pattern in both lungs

Data



Data was obtained from Kaggle, though the much larger, original dataset is also available

Chest radiographs were performed on children of one to five years of age

5856 total images in the dataset

Radiographs labeled “normal” or “pneumonia

How is Data Used?

01

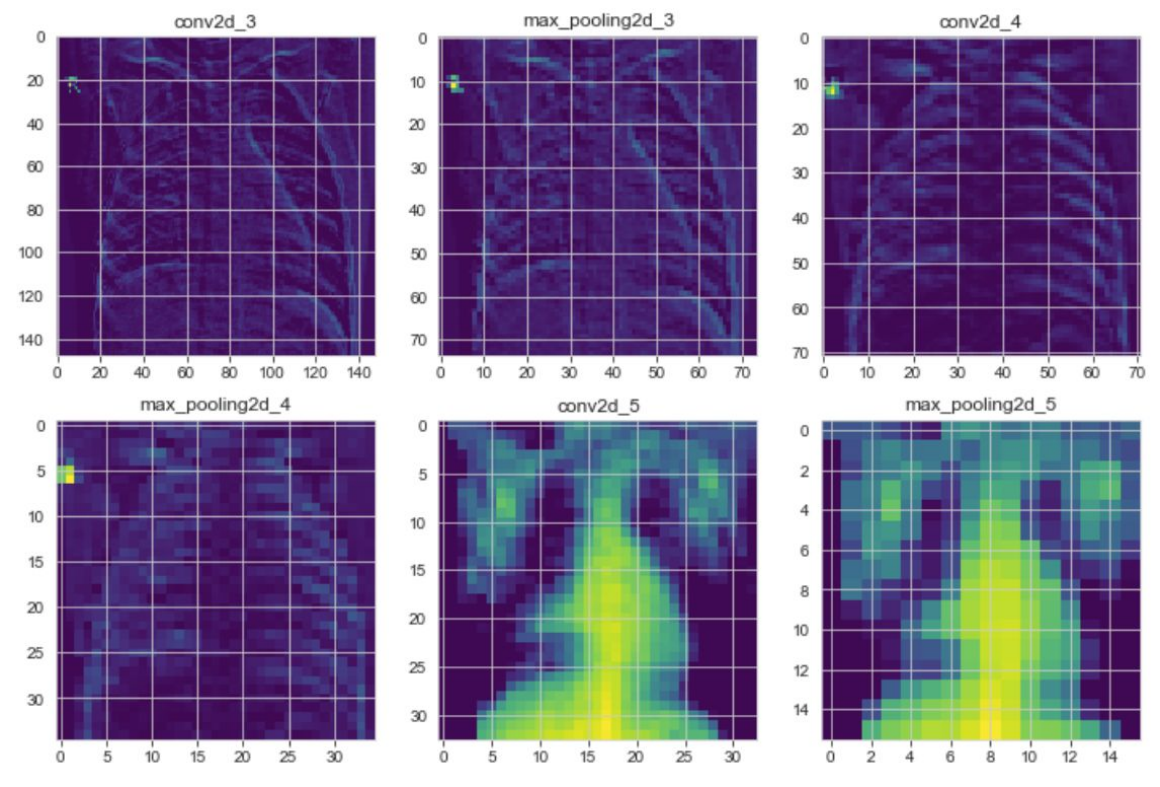
Patterns

Models will go through multiple layers looking for patterns and edges

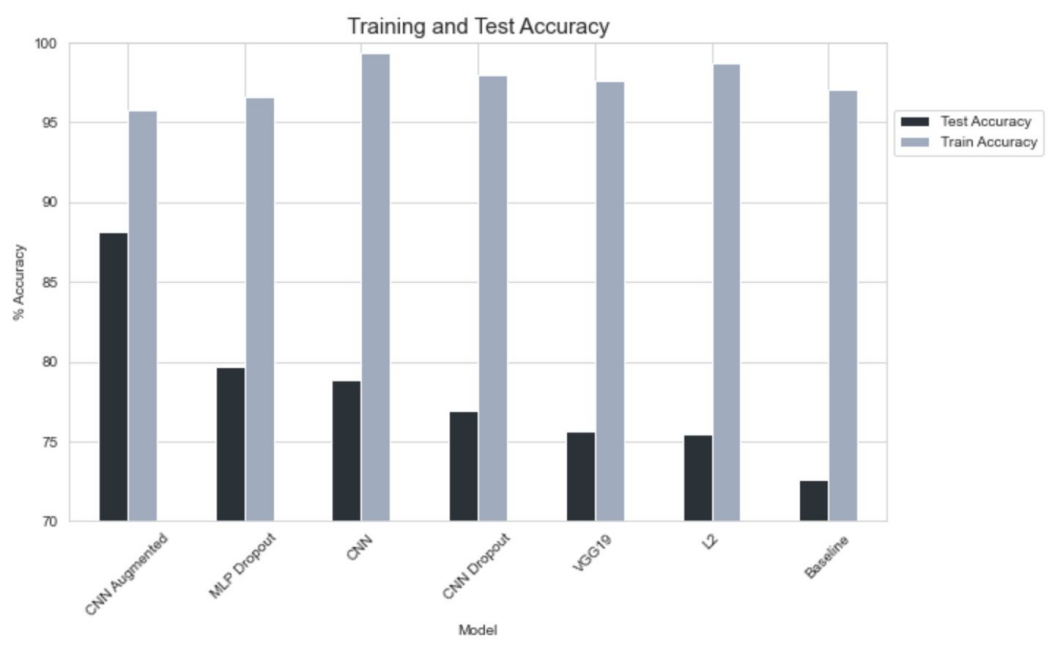
02

Learning

Using the patterns, the model learns what normal and pneumonia X-rays look like



Results



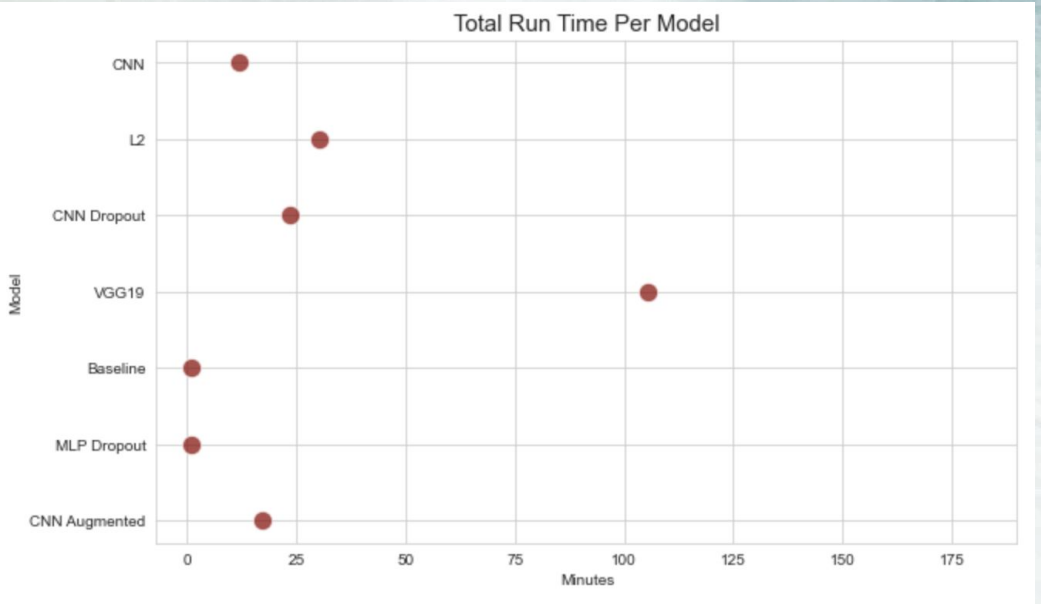
When models use data to learn then checked their results, all models scored above 95%

However, when tested against unseen data, only one model was able to obtain above 80% accuracy

Computational Cost

It is important to note that some models that are able to correctly classify radiographs may have downfalls, like the time it takes to train them

The VGG19 model was one of the best at decreasing false negatives, but took over 1½ hours to learn

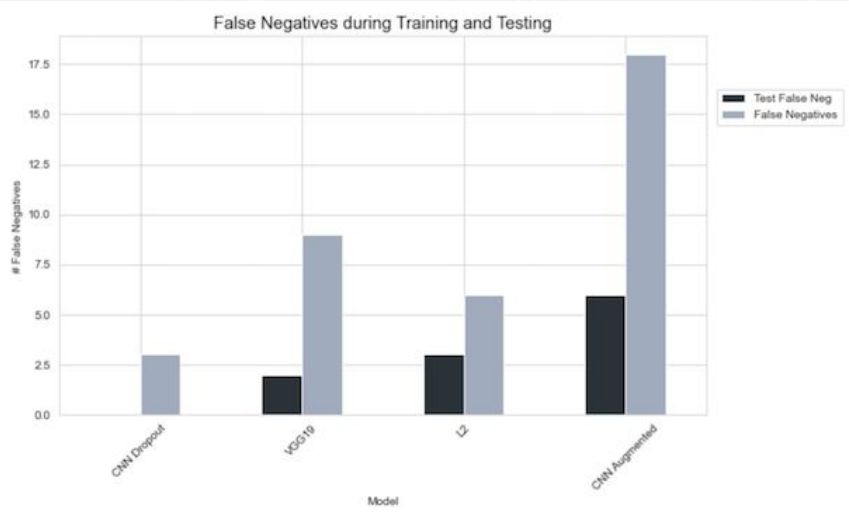


Accuracy is not everything

Accurate models are great, but we really need to make sure the model decreases false negatives

In terms of our problem, a false negative would be the model saying the X-ray is normal when it is truly has pneumonia

This could be dangerous and life-threatening to miss pneumonia in a patient



Putting it all Together

Which model
out-performed the others?

CNN with Augmentation

88%	Highest accuracy
1.6%	Low false negative rate
29%	Lowest false positive rate
17	Reasonable computational cost, measured in minutes

Future Work

Help with COVID-19?

Research how this model or dataset could help diagnose COVID-19-related issues

Transfer Learning

See if other pre-trained networks improve performance

Increase Data

Add more images from original dataset

Confirm Diagnosis

Use other methods to diagnose pneumonia in conjunction with physician diagnosis

Thanks!

Questions?

Cassie Nutter
cassie.m.nutter@gmail.com