



# PNEUMONIA CLASSIFIER

Protecting Lung Health through CNN

# The Dataset

- Training Data contains 1341 images of normal lungs and 3875 images of Pneumonia infected lungs of the bacterial and viral strains.
- The data is from pediatric patients in China



# The Problem



Coronavirus, air pollution and many other aspects of modern life have increased the risk to our personal health and put enormous pressure on medical systems to automate processes such as triage.



Classifying a scan is a critical part of the triage process and a mistake can cost a life. 99% recall is required.



The goal is to build an image classifier that maximizes a recall score to support. We need to reduce load on limited staff in times of crisis.

# The Solution

01

Test various models against the dataset for result optimization

02

Assign a probability that the image is pneumonia to address the risk of false positives

03

Understand how the tool fits into the daily medical imaging process in hospitals

Used a model defined in the sequential class.




Layer types:

Desne: Fully connected  
layer

Dropout: Setting fraction  
of inputs to zero to reduce  
overfitting

Merge inputs into a single  
model



Results: 97% accuracy, 86% recall

The Results  
of the Best  
Model

# Integrating in the Hospital Toolkit



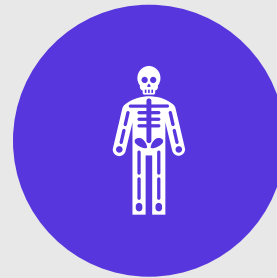
The recall remains 16% inaccurate in identifying pneumonia



The current tech would be best suited as a tool for stack-ranking the probability of severe pneumonia and feeding those images to the medical technician in order of 'probability' rather than a binary classifier.



This can help medical professionals view the scans in order of severity.



Further work with medical professionals to enrich the data could improve the classifier.