Pneumonia Detection by Chest X-Ray Images

Sibel Tanoglu

Problem Statement

- Pneumonia is an inflammatory condition of the lung affecting primarily the small air sacs known as alveoli caused by infection with viruses or bacteria.
- Diagnosed by symptoms, physical examination, chest X-ray, blood tests.
- Severity is variable especially in developing countries.
- Early stage diagnosis is very crucial.

Goal

A robust convolutional neural network that will be trained with thousands of labeled chest X-ray images to provide fast and accurate diagnosis for the benefit of both patients and health professionals.

Data

- Collected from **Mendeley** public datasets repository
- Labeled Optical Coherence Tomography (OCT) and Chest X-Ray Images for Classification
- Pediatric patients of one to five years old from Women and Children's Medical Center, Guangzhou.
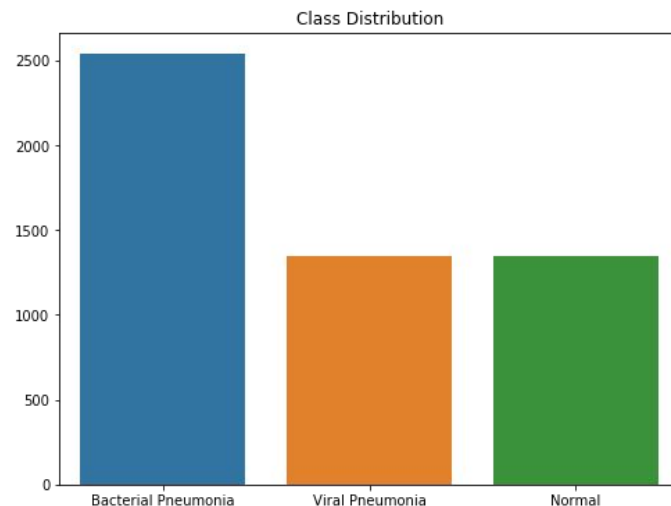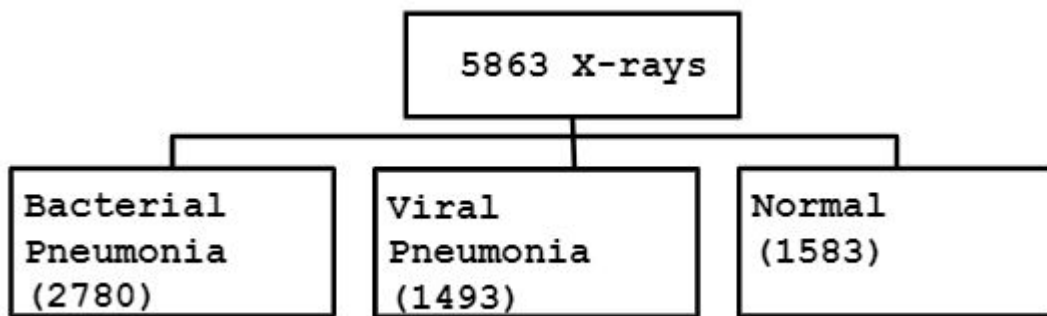
X-rays Visualization

- Images are almost indifferentiable for non-experts
- Pneumonia positive images seem to be slightly more faint due to soft tissue (water) around the lungs

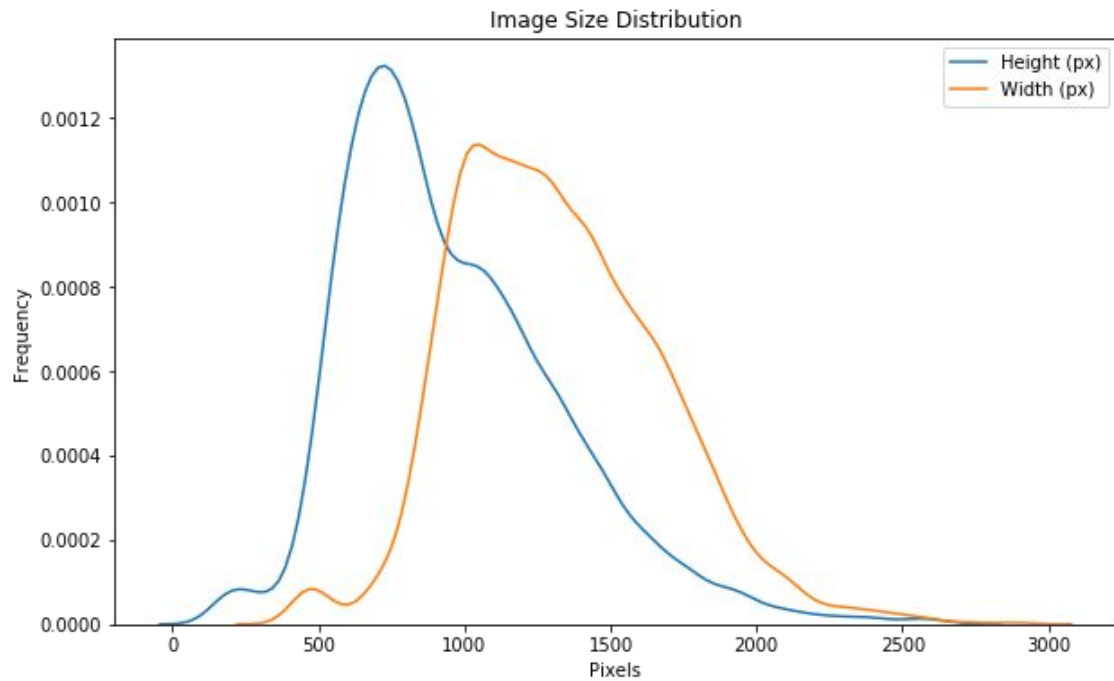Bacterial Pneumonia vs Viral Pneumonia vs Normal
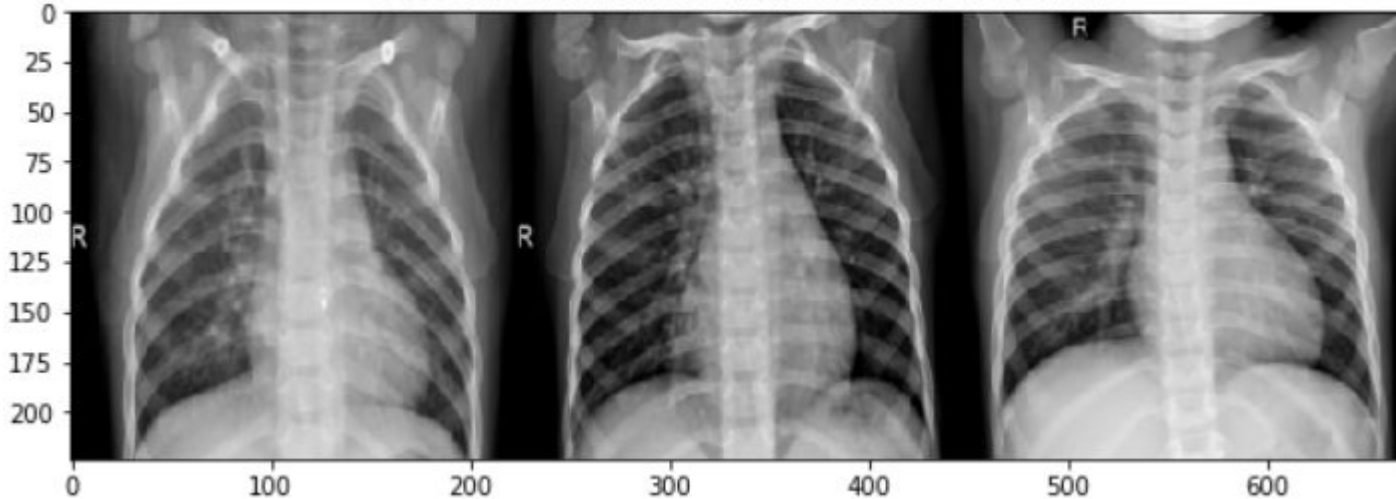
Black: Air

Dark Grey: Fat

Light Grey: Soft tissue

Off white: Bone

Bright white: Metal

Image Size Distribution (px)

- Most images are at least 256 × 256 pixels and RGB color channels
- Smaller and grayscale images are removed from analysis
- Downsized all images to 256x256x3 tensor
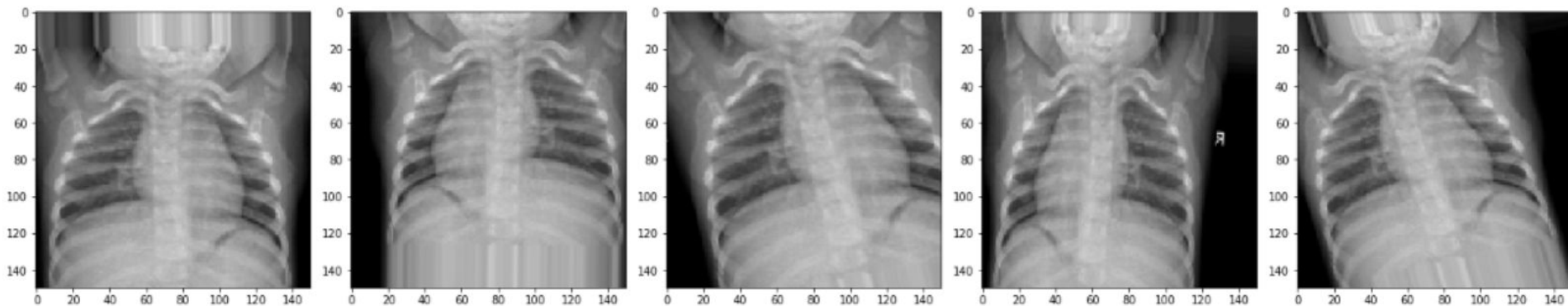- Multiplied channels by 1./255

Data Augmentation

Problem: A lot of variations for such a small dataset (angle, opaqueness, body position, zooming etc)

Solution: Image Data generator

rotation_range=20, width_shift_range=0.1, height_shift_range=0.1, shear_range=0.1, zoom_range=0.2, horizontal_flip=True, fill_mode='nearest'
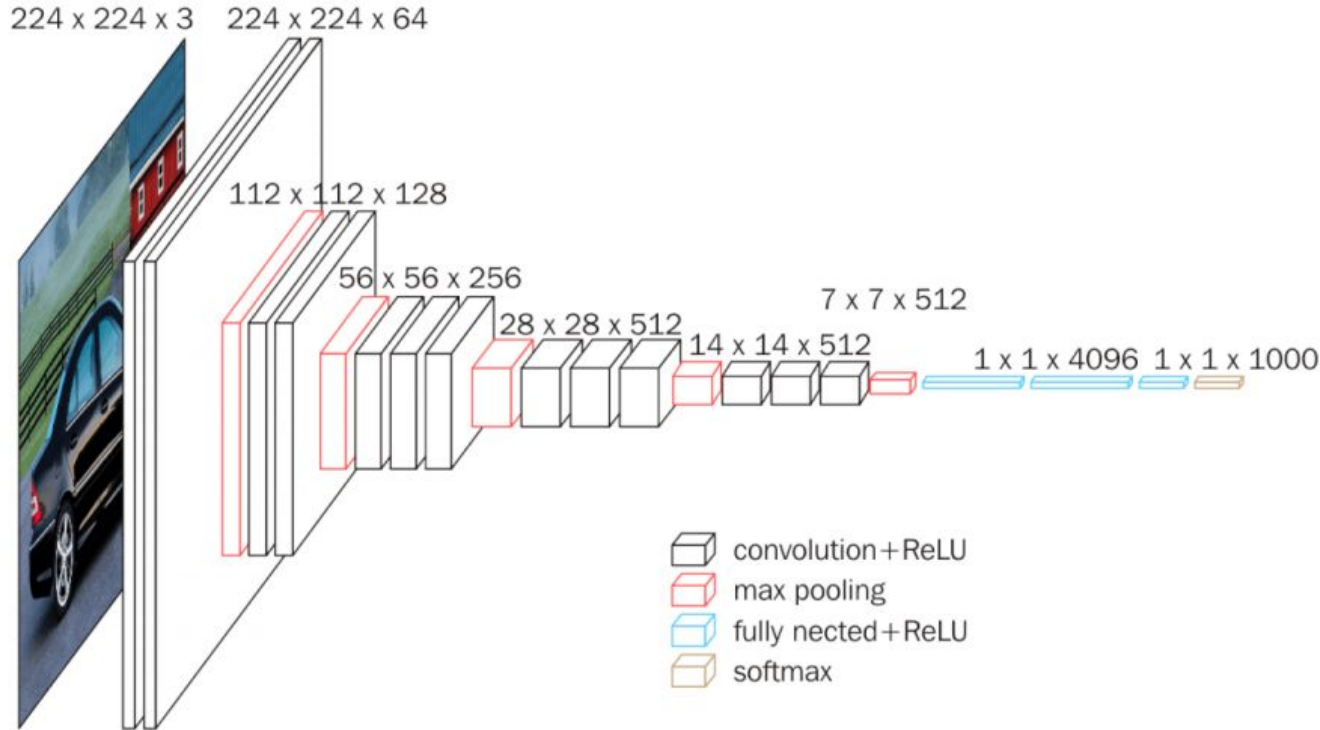
Different versions of same image:
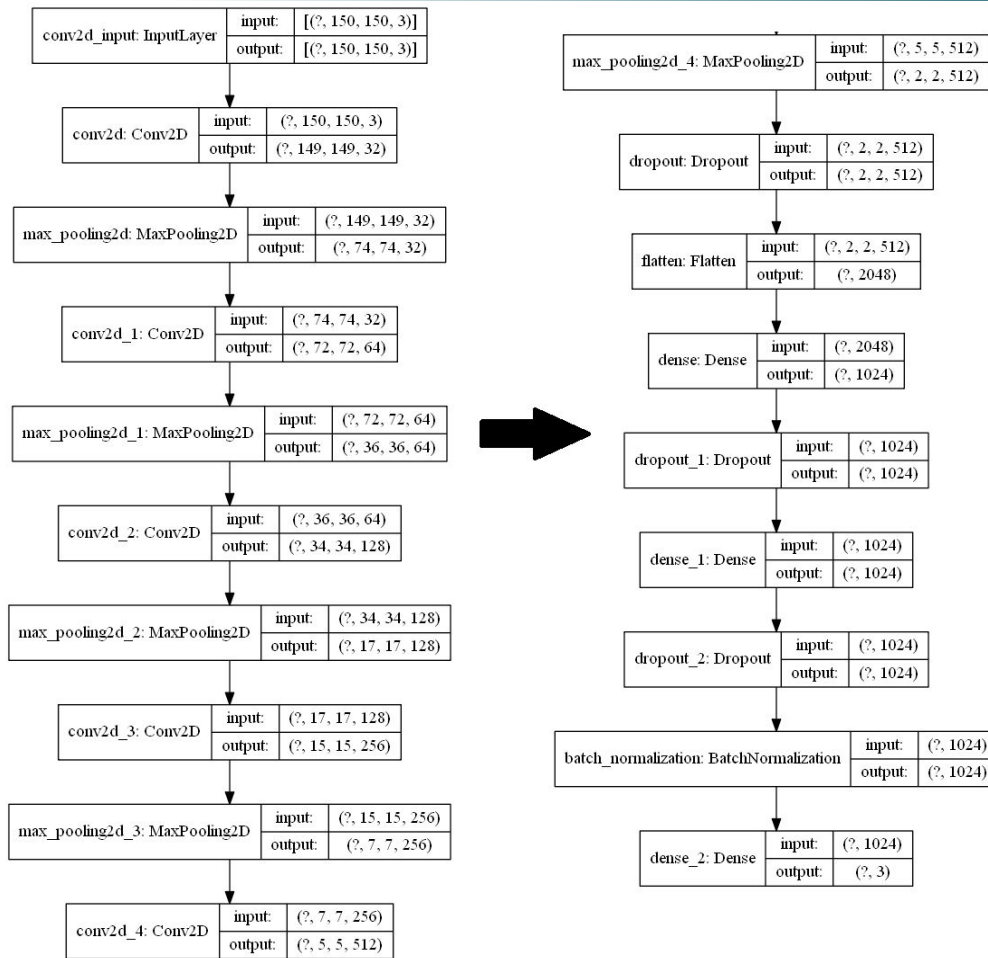
Modeling

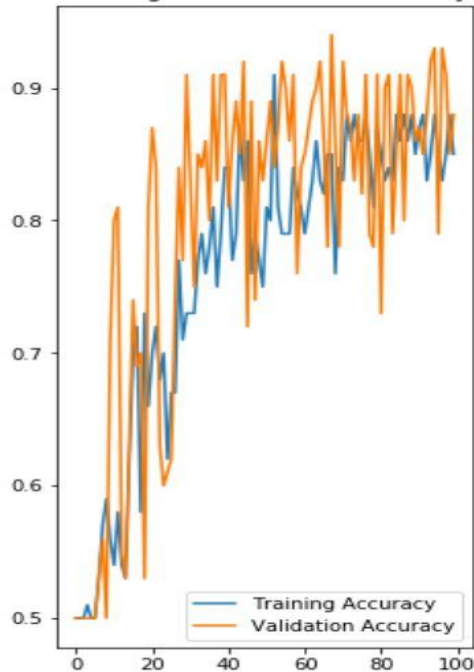# VGG16 (Transfer Learning)

- ImageNet weights used only in first 2 layers (first few layers capture general details)
- Binary classification (P/N)
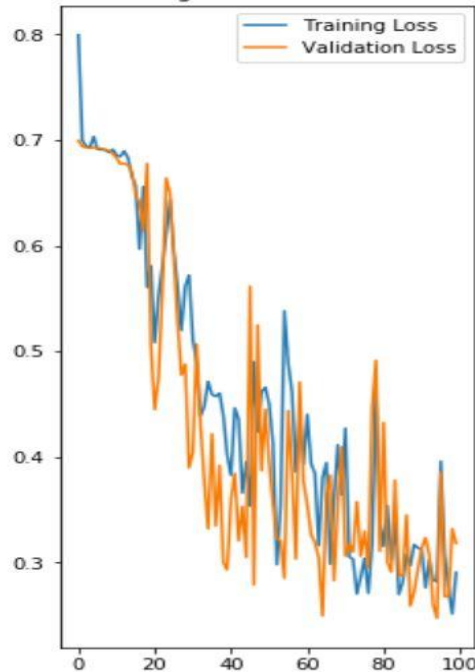- Acc = 90.4%
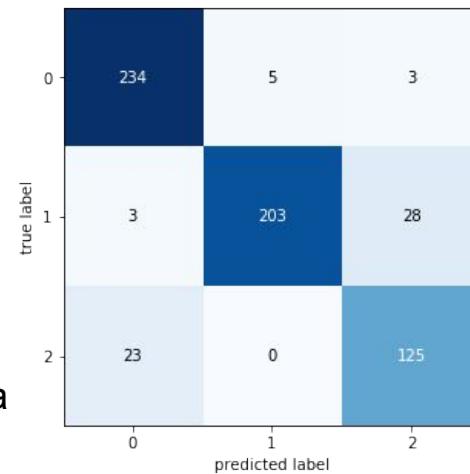- Val Acc = 88.5%

Final CNN

Model Evaluation

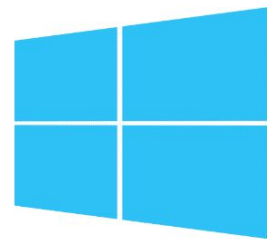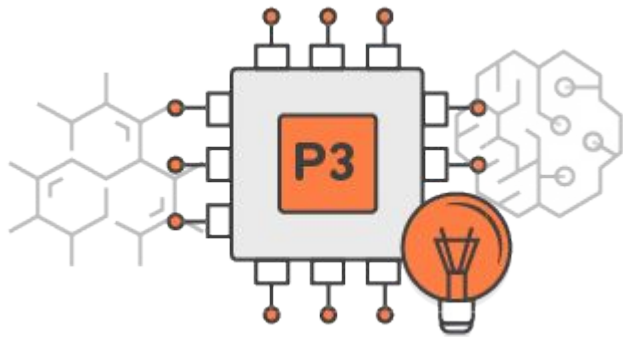
Accuracy = 92.4 %
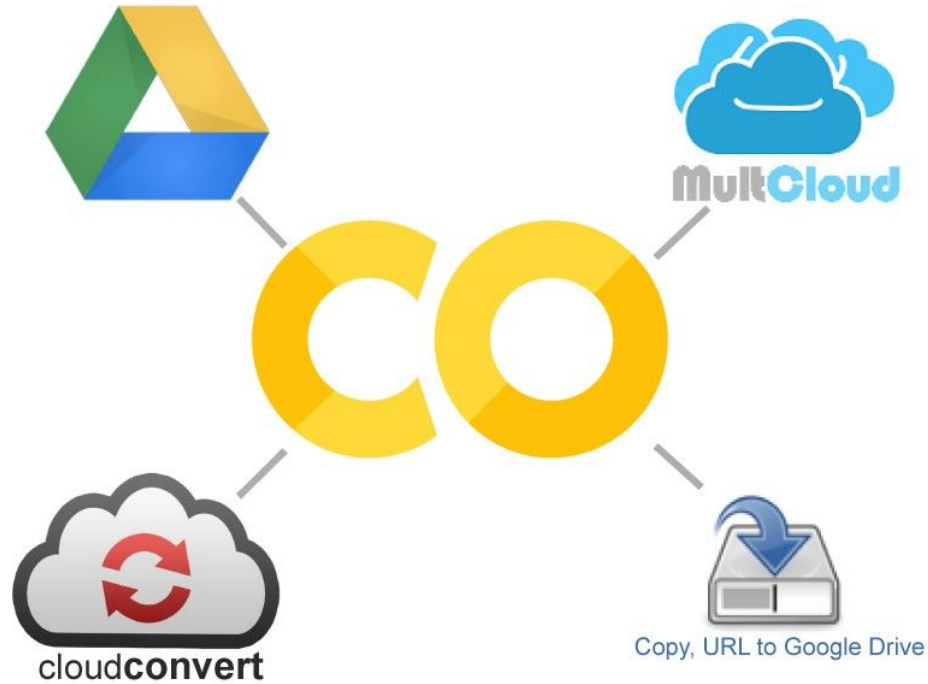Val Accuracy = 90.1 %

Confusion Matrix
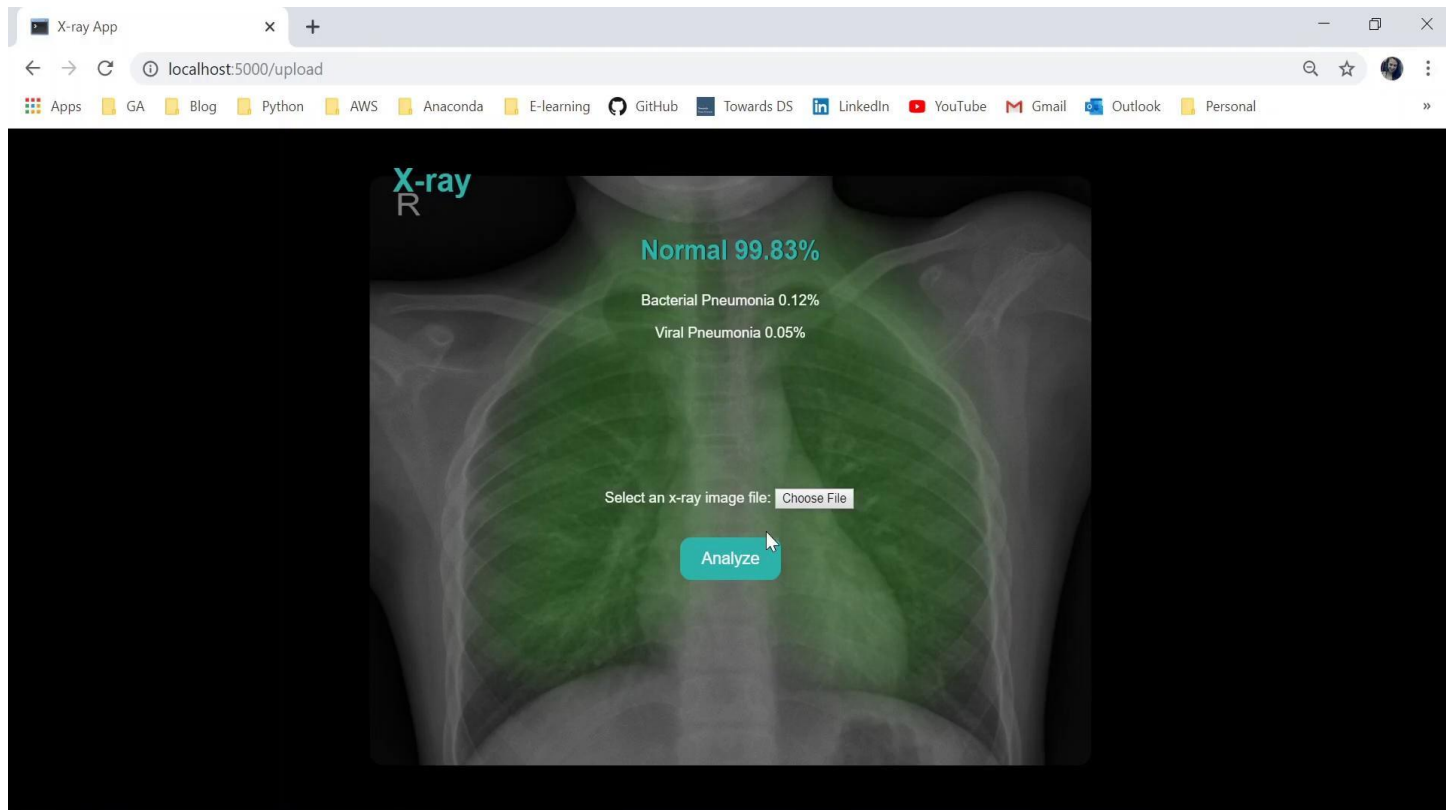
0 = Bacteria
1 = Normal
2 = Virus

Workspace 1

- **AWS EC2 P3 Instance pre-configured with ML packages**
- **Windows Server 2016**
- **NVIDIA Tesla V100GPU / CUDA Runtime**
- **64 GB RAM**
- **100 GB HDD**

Workspace 2

X-ray App Demo

Next Steps

- Collecting more data.

- Adding more complexity in model and using Keras gridsearch functionality.

- Creating additional model with PyTorch library and compare the results.

- Researching on how to detect most informative pixels on X-ray (highest weights) and emphasize (highlight) them on image.

- Partnering with radiologists to classify same X-rays and compare model's performance with professional interpretation.

Thank you...

Deep learning is shallow next to the depths of your greatness.

Repo: https://github.com/sibeltan/Capstone/