Identifying Pneumonia with Deep Learning

Author: Andre Layton

Summary

Objectives:

- ➔ To build a classification algorithm that correctly classifies chest X-ray images between two groups (“NORMAL” and “PNEUMONIA”).
- ➔ To determine what model accurately identifies pneumonia-ridden patients.

Results:

- ➔ The Convolutional Neural Network (CNN) model serves as the best algorithm for image classification, yielding 40.3% accuracy.

Outline

- Business Problem
- Data
- Methods
- Results
- Conclusions

Business Problem

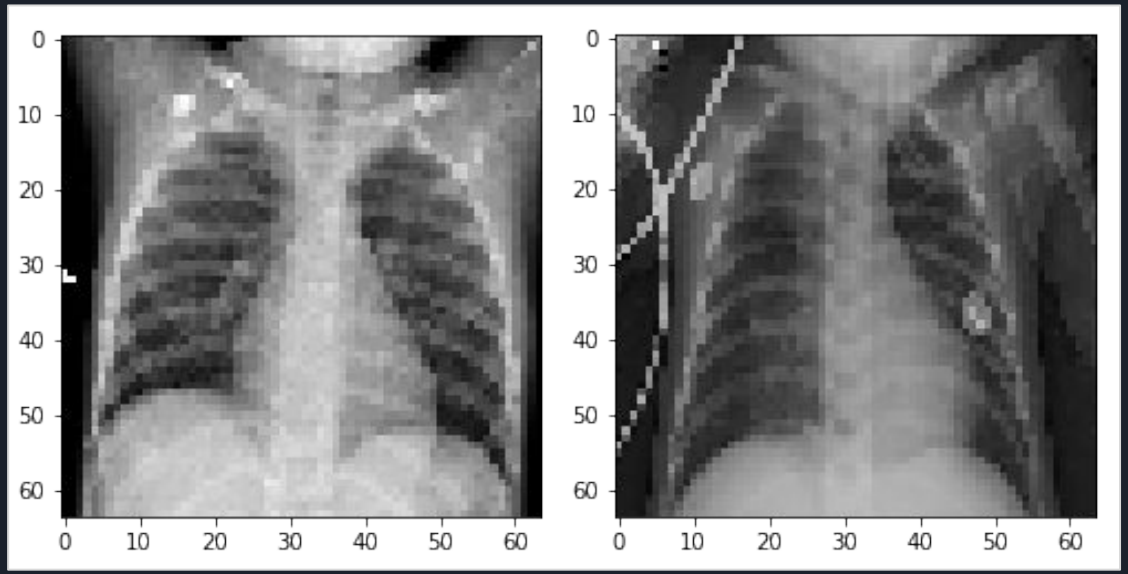
- A hospital has approved an initiative to build an in-house radiology unit that incorporates A.I.

- Use deep learning to create algorithms that will analyze and classify X-ray images.
  - Accuracy and false negatives are the important metrics to monitor!

Data

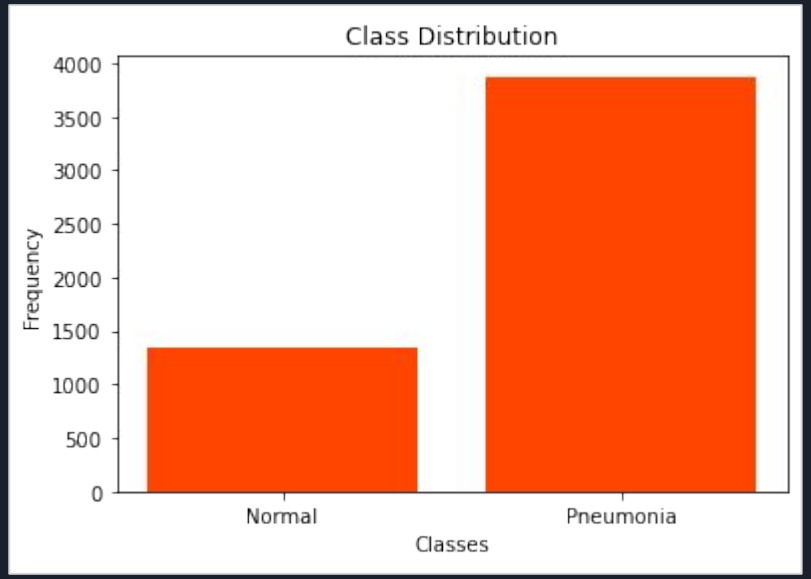
- The initial dataset was gathered from Kaggle, and contains:
  - 5,216 training images
  - 58 validation images
  - 624 testing images

- I added more images to the validation dataset by moving 100 images from the testing data.



NORMAL PNEUMONIA

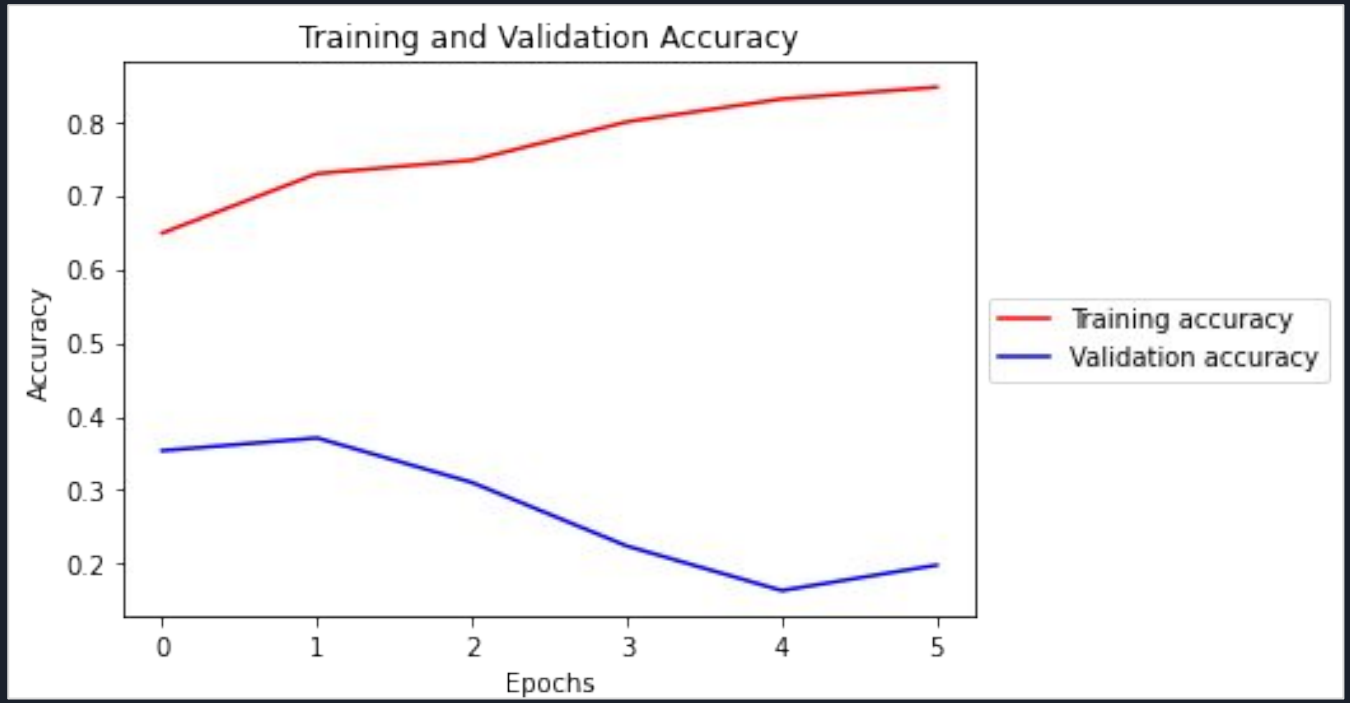
Data (cont.)



➔ The two groups (“NORMAL” and “PNEUMONIA”) are initially distributed as shown to the left.

➔ Data was balanced, reshaped, and augmented to prepare for the modeling process.

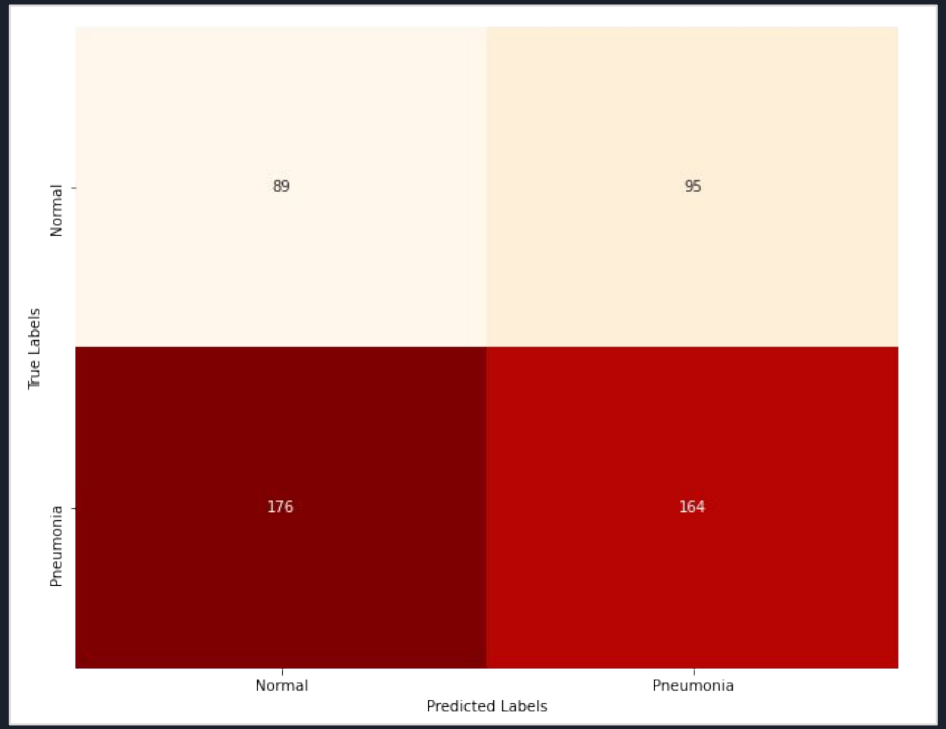
Modeling



- The CNN model was determined to be the best classifier.
- The model is 40.3% accurate when analyzing unseen images.

Modeling Results (cont.)

- ➔ Out of 524 predictions, the model predicted 176 false negatives (predicting “NORMAL” when the patient actually has pneumonia!).



Conclusions

➔ The CNN model is the best determined model for classifying chest X-ray images, and is 40.3% accurate on unseen data.

➔ Further work could include:

- Adding more images and balanced data to train the model on.
- Analyzing the dataset with a stronger computer.
- Experimenting with various parameters for model tuning.

Thank You!

Email: alaygt6@gmail.com
GitHub: @therookiescientist-andre
LinkedIn: linkedin.com/in/ak-layton/