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.

Based on the study:

- ➔ The Convolutional Neural Network (CNN) model serves as the best algorithm for image classification, yielding 84.5% accuracy on tests.

Outline

- Business Problem
- Data
- Methods
- Results
- Conclusions

Business Problem

- ➔ Build a classifier that can classify X-ray images.

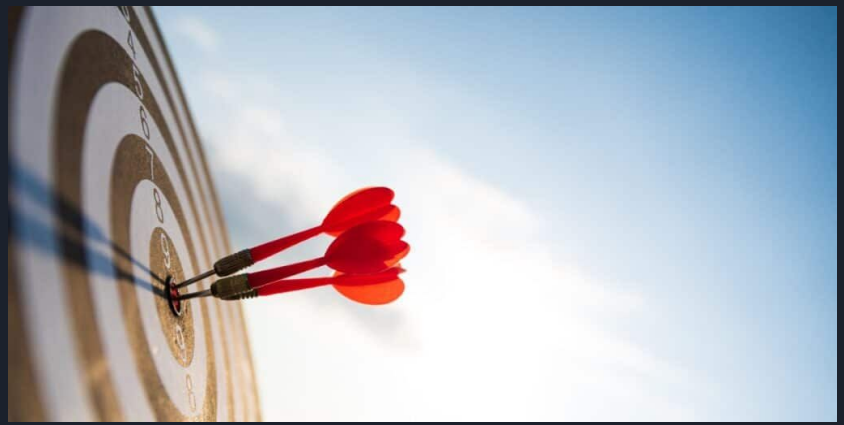
- ➔ Significant to physicians and imaging labs!

Data

- The initial dataset was gathered from Kaggle, and contains 3 directories, each with two sub-directories.
  - 5,216 training images
  - 58 validation images
  - 624 testing images

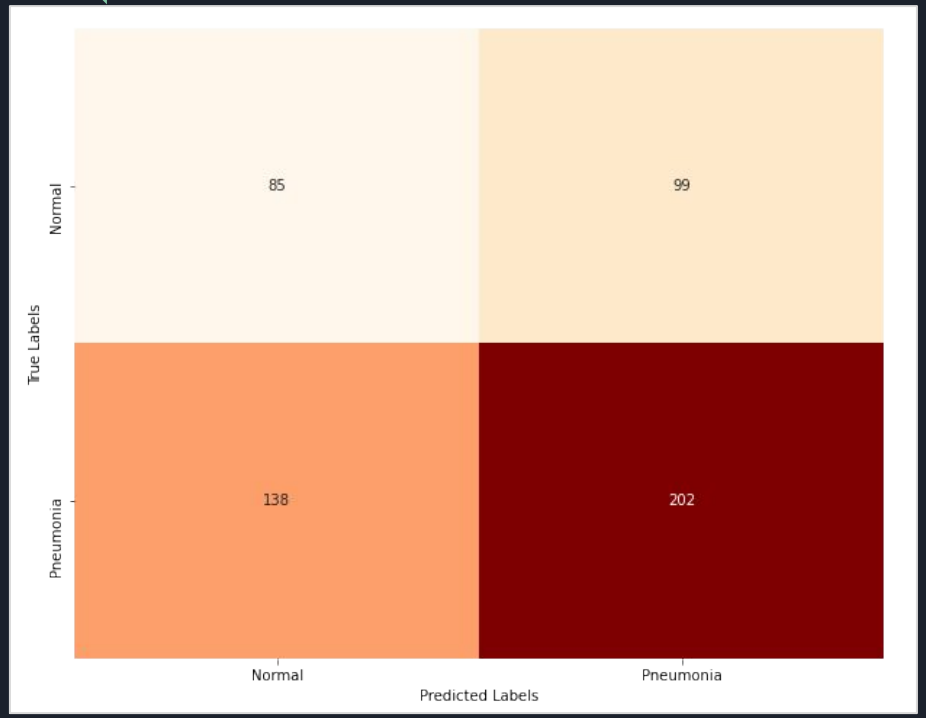
- The final dataset is scaled, reshaped, and augmented.
  - Validation data size was increased.

Modeling



- Models were evaluated based on accuracy, and their loss (margin of error) was monitored.
- I selected the CNN algorithm based on its high accuracy score.

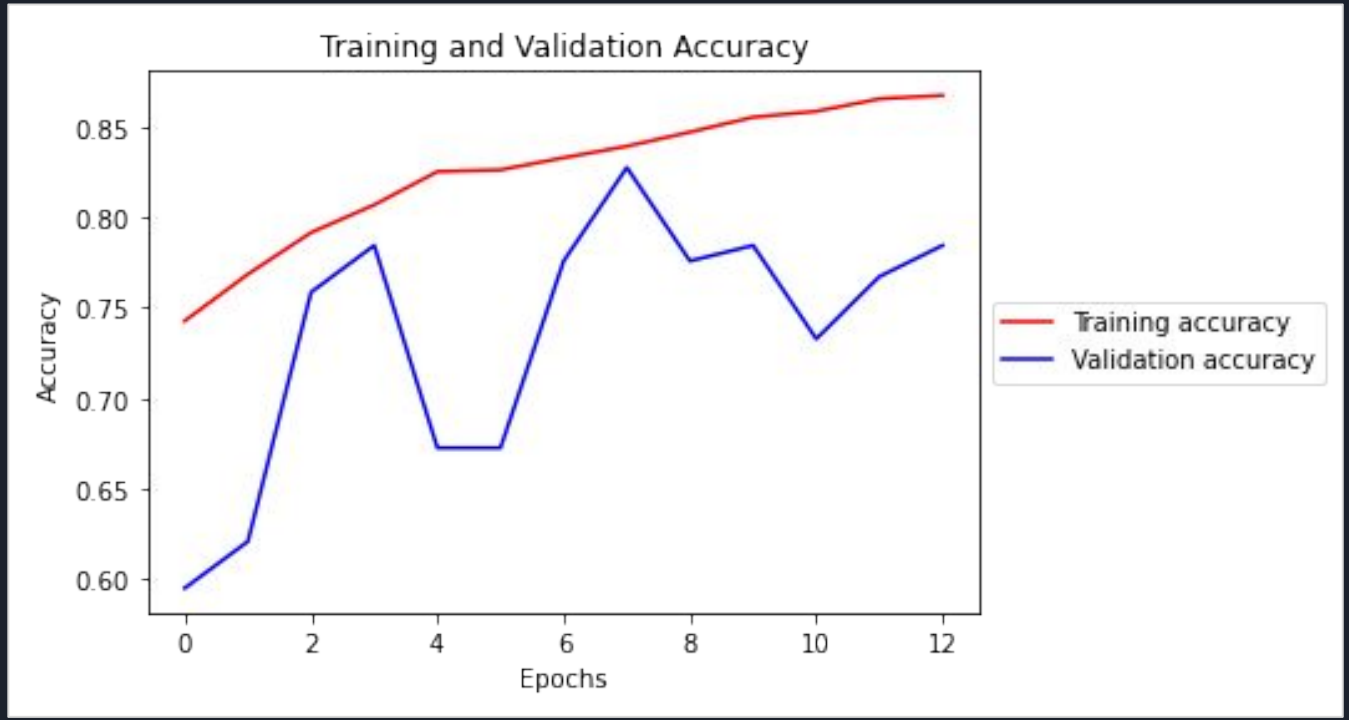
Modeling Results



- ➔ The model is 84.5% accurate on testing images.
- ➔ There were 138 false negative cases.

Modeling Results (cont.)

- The accuracy for both datasets increase as we run the model through more epochs.



Conclusions

➔ The CNN model is best for classifying chest X-ray images, and is 84.5% accurate on unseen data.

➔ Further work could include:

- Adding more images/data to train the model on.
- Investigate other pre-trained networks for modeling.
- Experiment with various hyperparameters for model tuning.

Thank You!

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