Predicting Pneumonia

Laura Brown

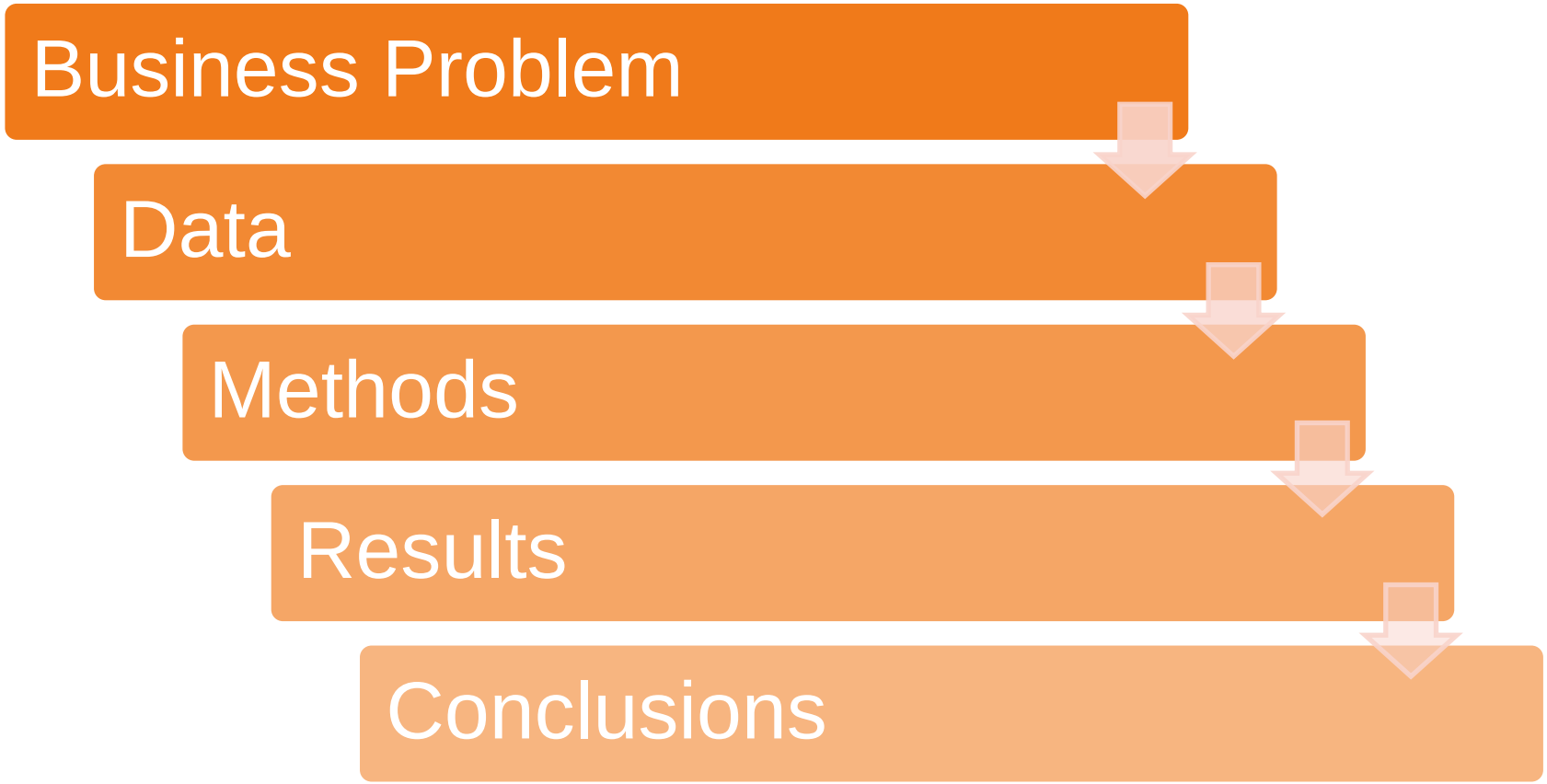
Summary

The goal for this analysis was to build a model that can classify whether a patient has pneumonia, given a chest X-Ray image.

Model produced results that ranged from 80-88% accurate.

The Recall score for this analysis fluctuated from 85-90%

Outline



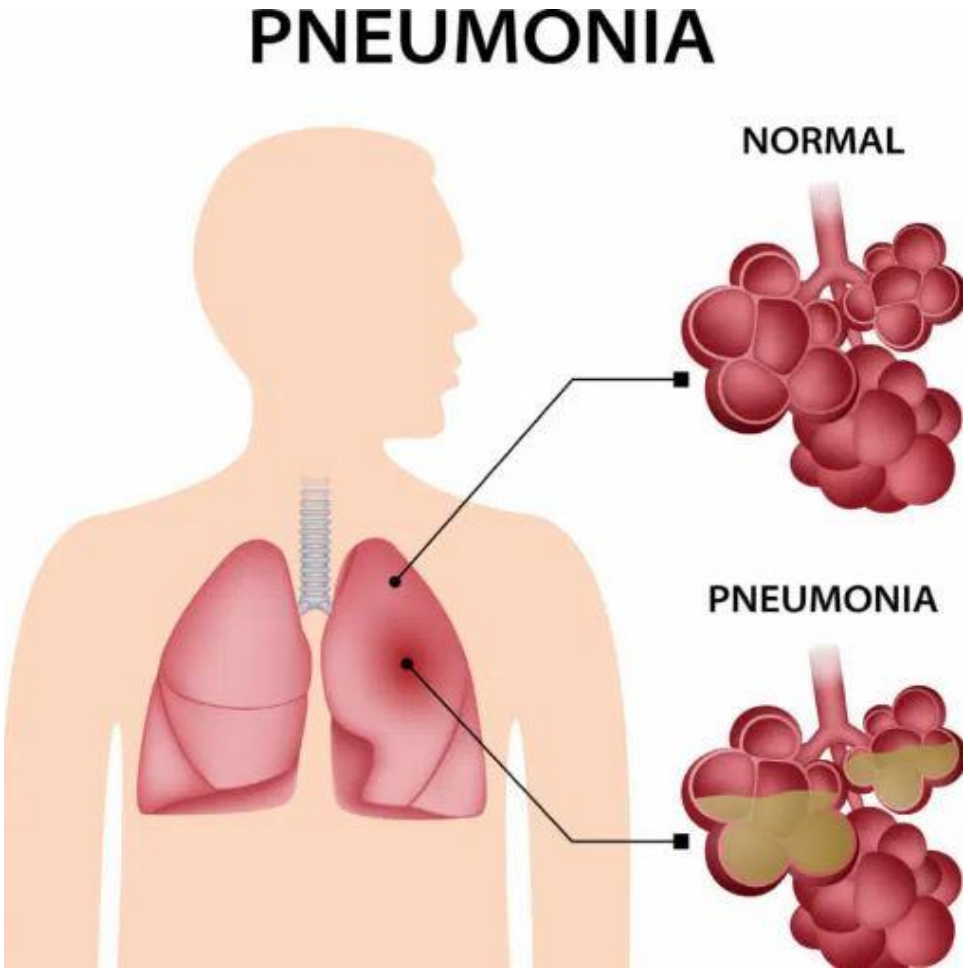
Business Problem

Given X-Ray images, this project aims to determine whether machine learning can differentiate the differences between normal patients and patients whose lungs have been infected with pneumonia with high accuracy.

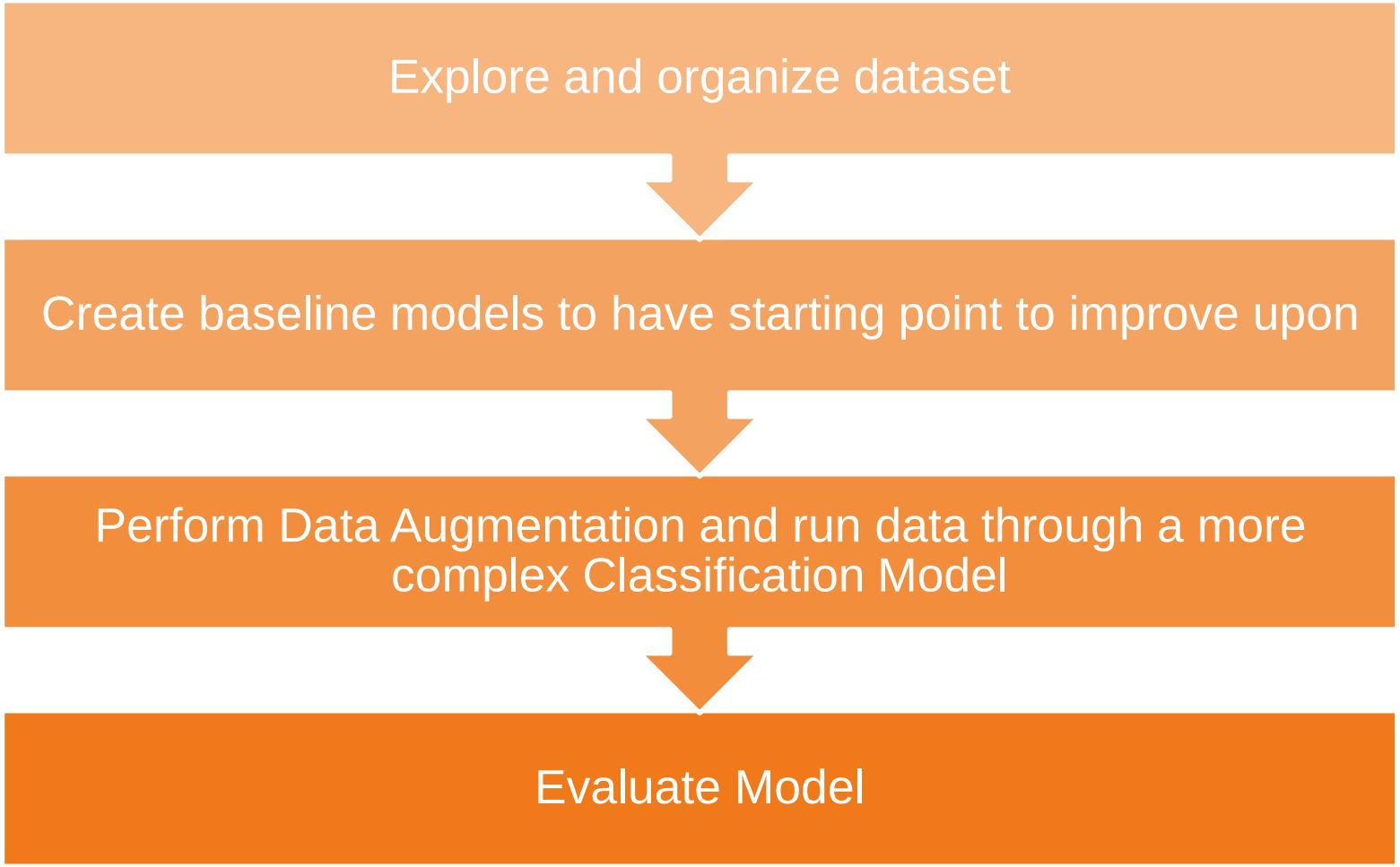
Objective:
Create a model that can accurately predict whether a patient's X-Ray image is positive for pneumonia.

Data

- Source: Mendeley Data
- Approximately 6,000 X-Ray images
- Pediatric patients 1 to 5 years old from Guangzhou Women and Children's Medical Center
- 2 Categories of images
  - Normal
  - Pneumonia

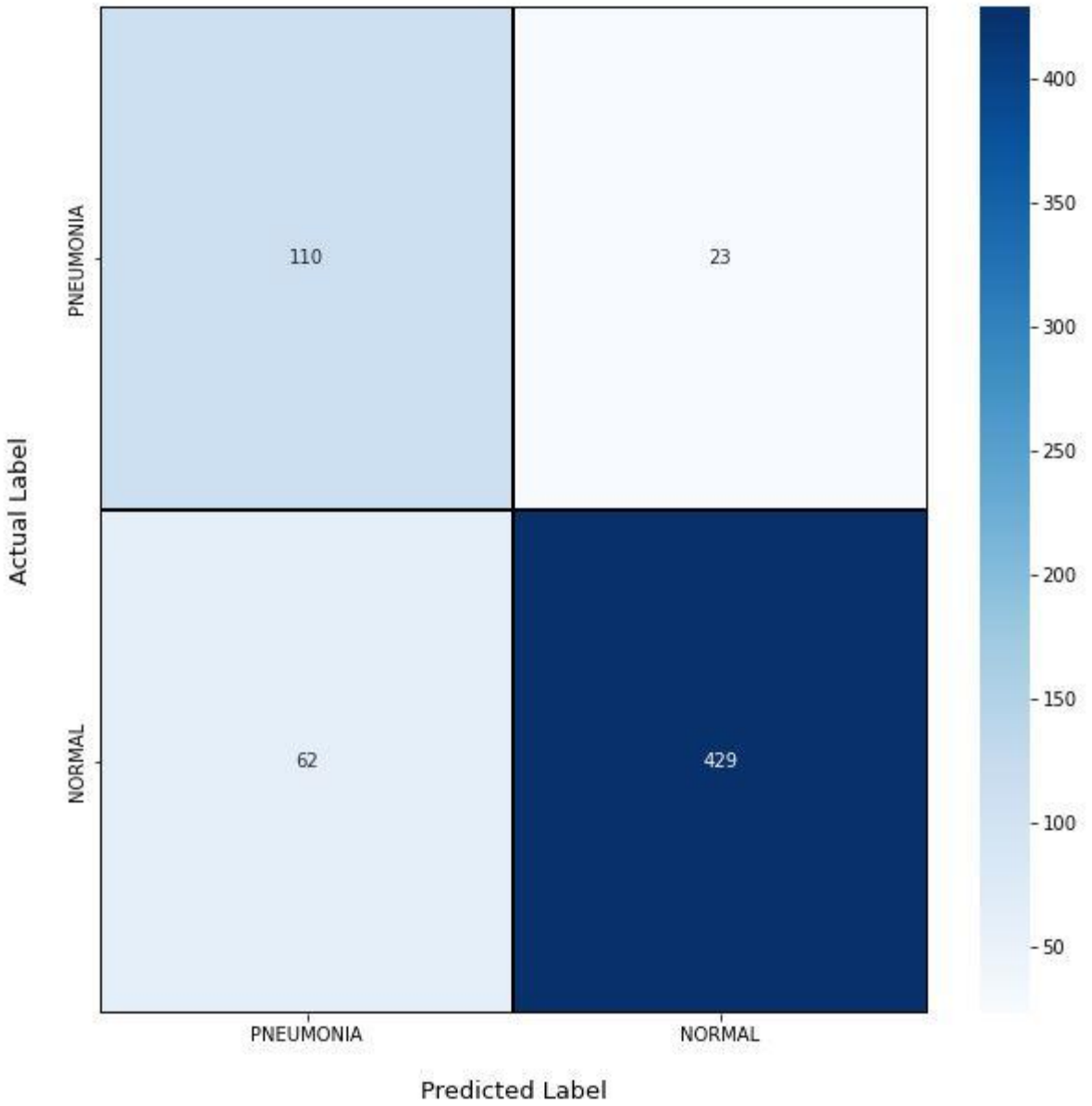


Methods


Explore and organize dataset
Create baseline models to have starting point to improve upon
Perform Data Augmentation and run data through a more complex Classification Model
Evaluate Model

Results – Final Model

- Accuracy ranged from 80-88%
- Recall ranged from 85-90%



Conclusions

- Due to the randomness of neural networks and CNNs, the accuracy rate varied from 80 % to 88%
- A high recall indicates a low false negative rate, meaning the model is good at catching positive cases. In this case, the recall score ranged from 85% to 90%

Next Steps:

- Creating a model that uses different data types alongside the X-Ray images, such as CT scans, patient metadata and audio data from lungs to more accurately predict positive cases.

Thank You!

Email: lpb3393@gmail.com
GitHub: @lpb3393
LinkedIn: linkedin.com/in/lpb3393