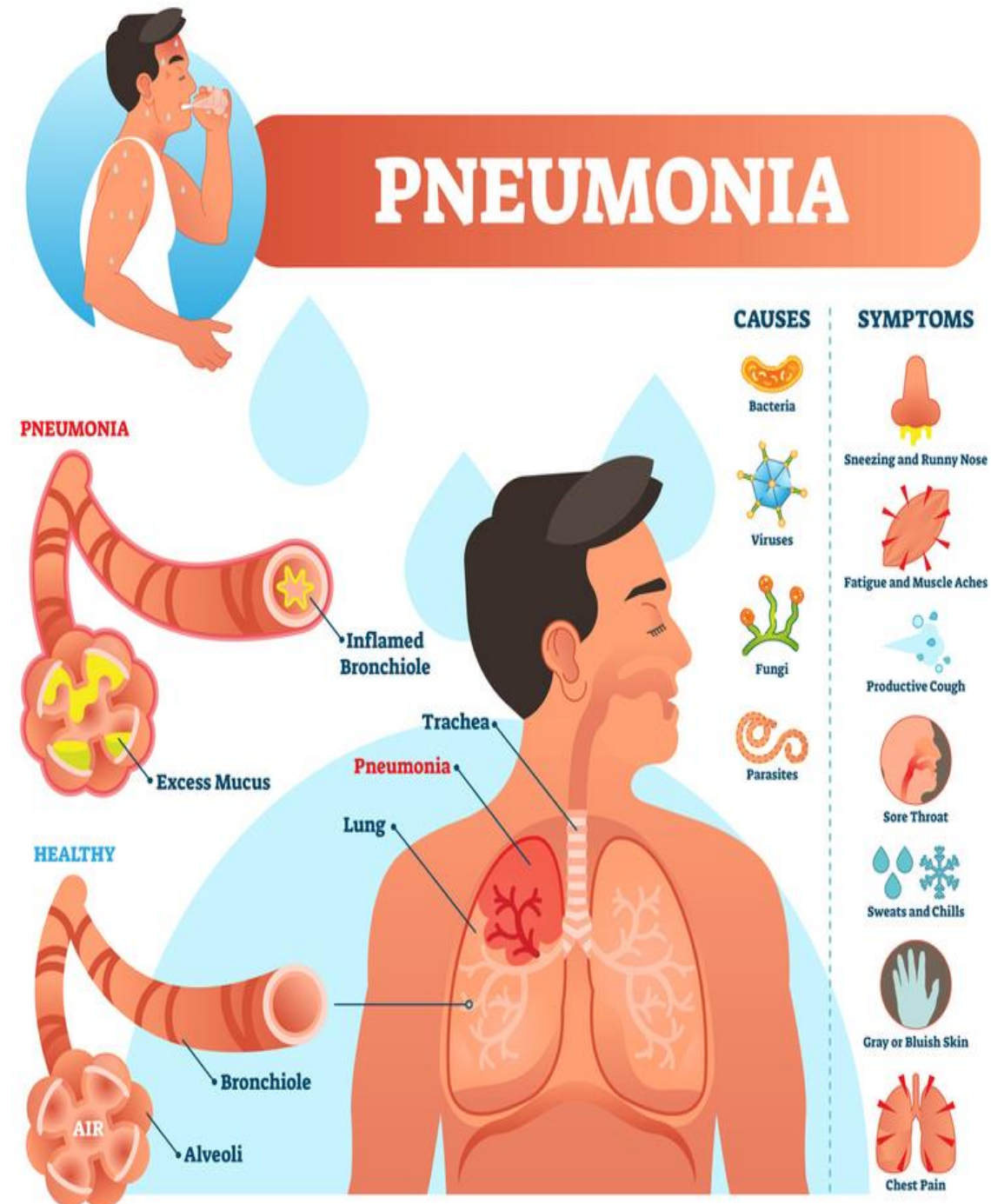


# Pneumonia Classification

By Alejandro Harrison

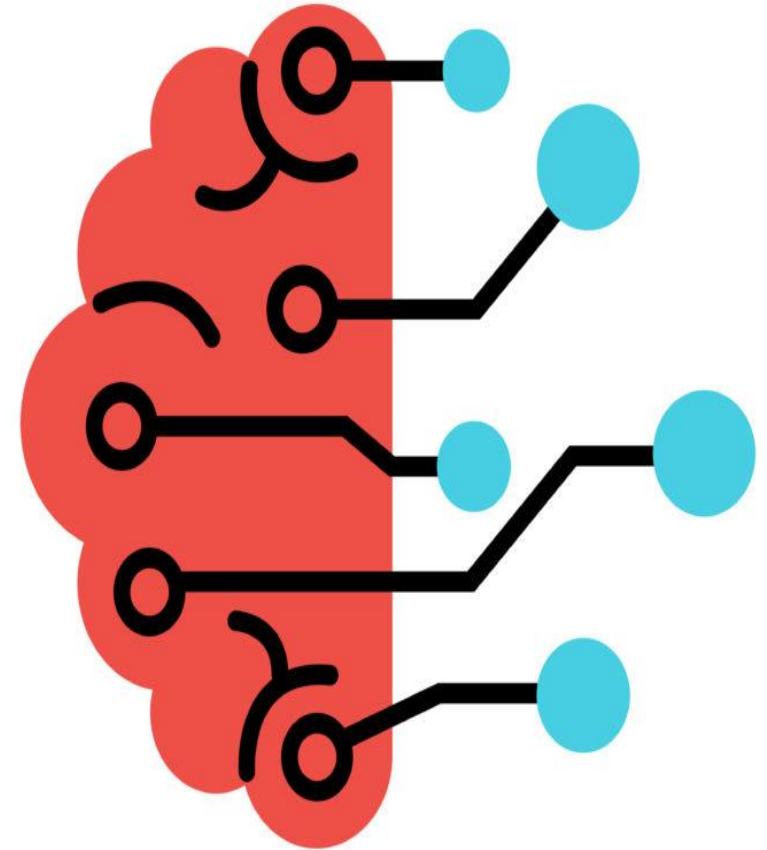
# Business Goal

- ▶ Pneumonia is common cause of hospitalization
- ▶ 1 million people in US affected a year
- ▶ Infection of lungs
- ▶ Requires chest x-ray, then interpretation by doctor
- ▶ Hospital wants to minimize time spent analyzing x-rays



# Method

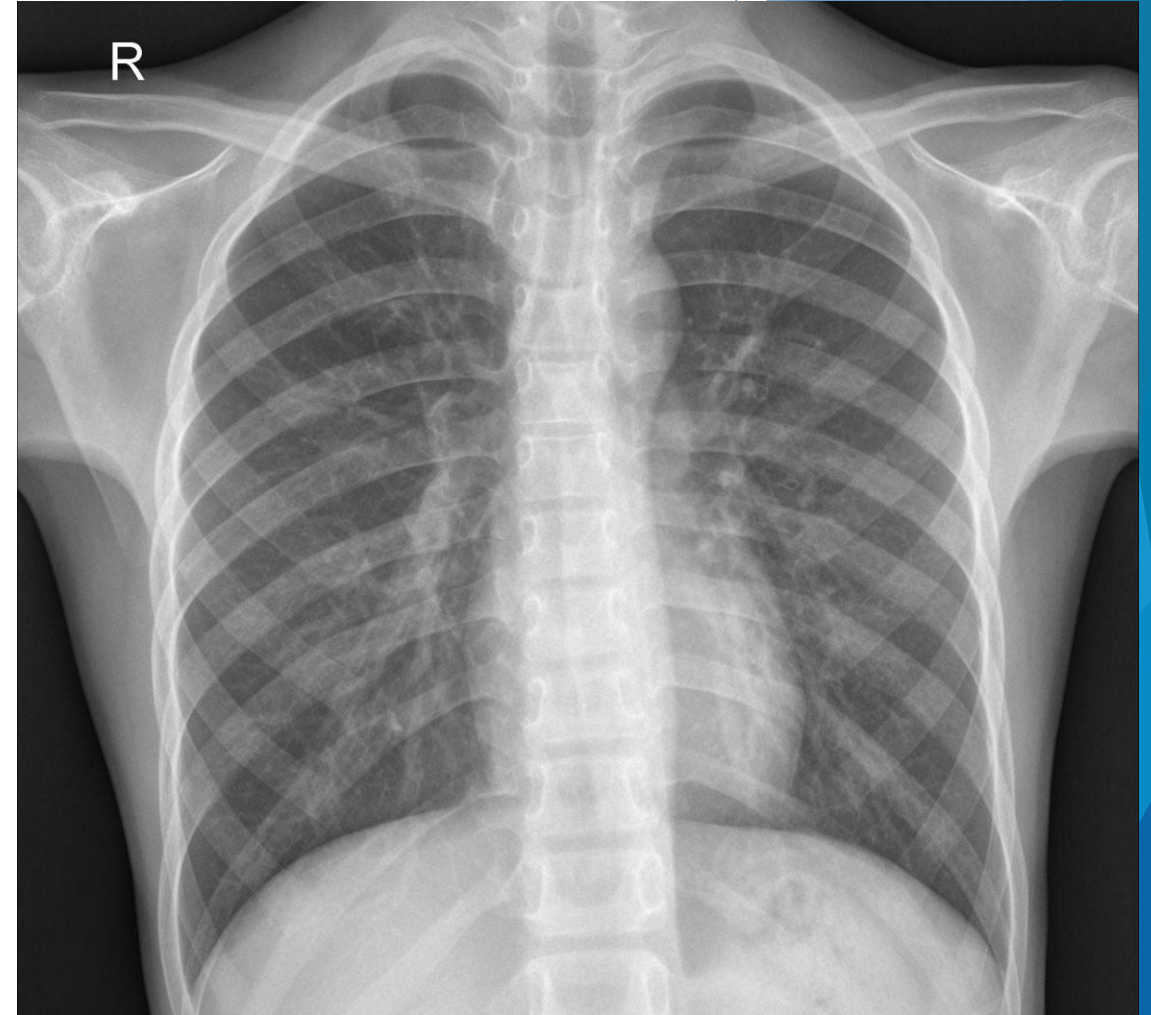
- ▶ Use of Image Classification through Neural Network models
- ▶ Deep Neural Network and Convolutional Neural Network(CNN)
- ▶ Pick the best one
- ▶ Accurately classify x-ray image



## NEURAL NETWORK

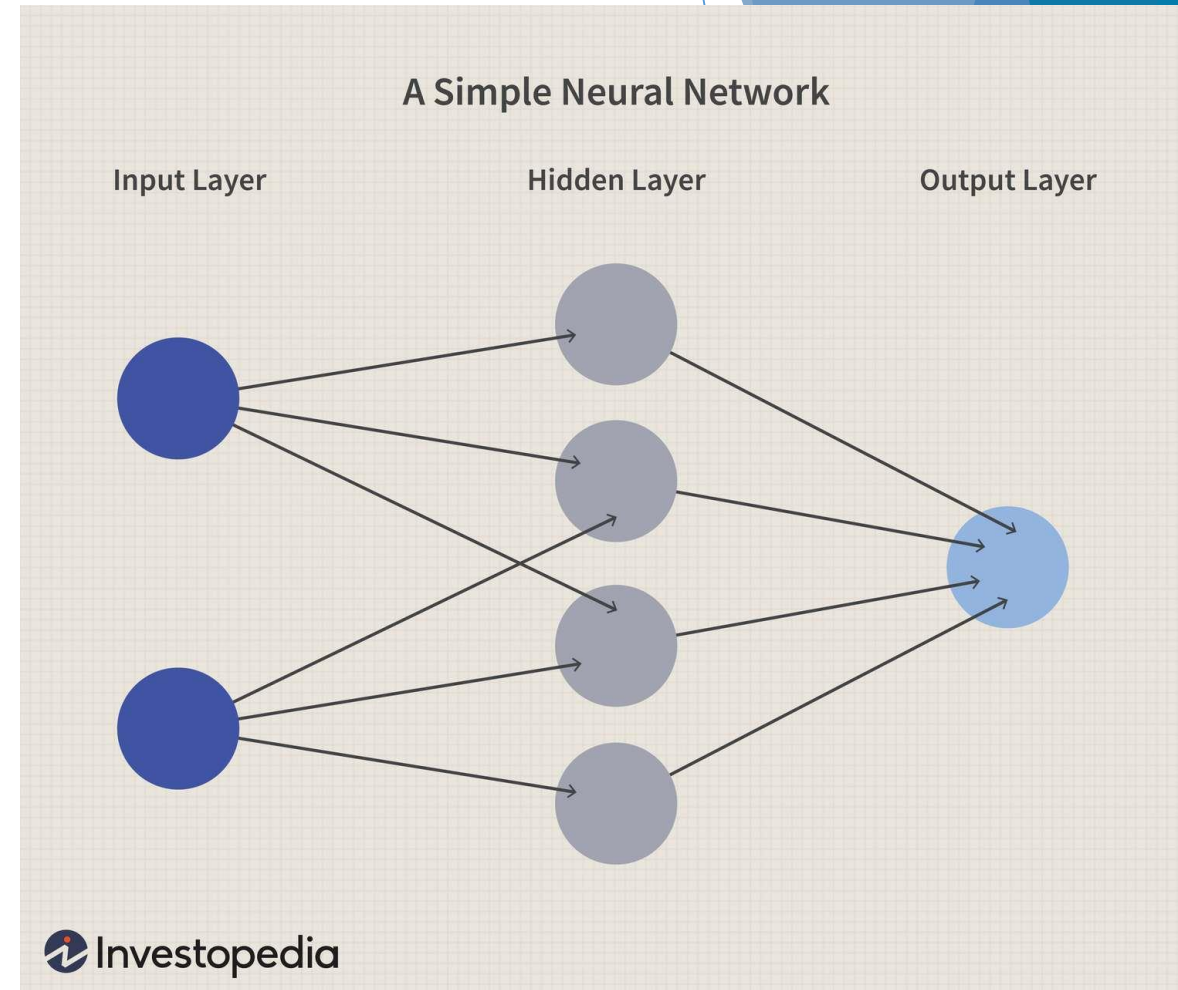
# Data

- ▶ Kaggle dataset of chest X-ray images
- ▶ 5,856 images
- ▶ 4,273 pneumonia images and 1,583 normal images
- ▶ What are DNN and CNN?



# A Simple Neural Network

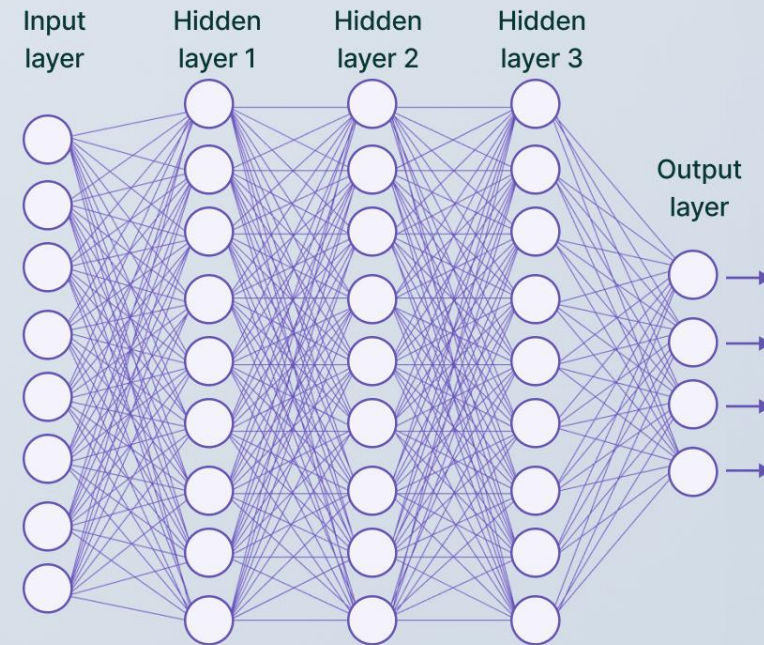
- ▶ Series of nodes (processing units) comprised into layers
- ▶ Input layer, 1 hidden layer, and output layer
- ▶ Model takes some input (Images in our case)
- ▶ Processes data (images) in hidden layer
- ▶ Output layer produces some output (image classification)





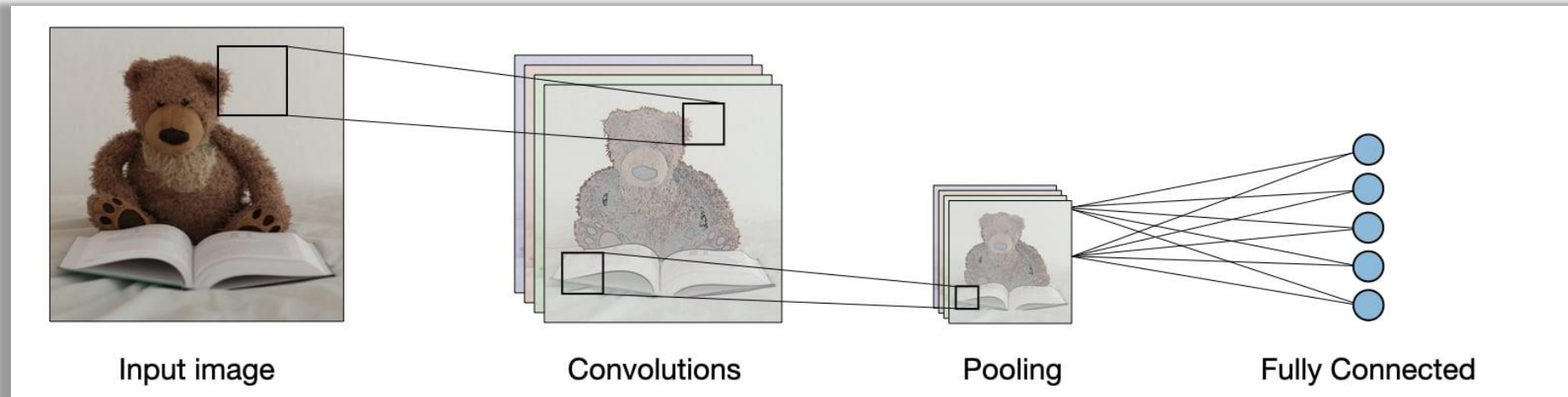
# Deep Neural Network

- ▶ Similar to simple Neural Network
- ▶ More hidden layers, therefore more processing
- ▶ Detection of more complex features of an image, and groups parts together to classify image



# Convolutional Neural Network

- ▶ Detects complex features better than a DNN
- ▶ Applies filters to image
- ▶ Filter moves across image from top to bottom
- ▶ Looks for edges or patterns
- ▶ Features mapped out and new smaller images created
- ▶ Images condensed and classification decision is made



# Best Model

- ▶ Compared DNN and CNN, tuned in different ways
- ▶ CNN had better accuracy
- ▶ Look at our results



# Training

- ▶ Different accuracy types
- ▶ Accuracy of model on classifying data it trained on
- ▶ Our training accuracy was 90%
- ▶ Comparing to DNN results, 51% improvement

# Validation

- ▶ Validation accuracy evaluates overall model performance
- ▶ Validation accuracy was at 81%
- ▶ 19% improvement compared to DNN results

# Testing

- ▶ Accuracy of model on classifying unseen data images
- ▶ Test accuracy was 91%
- ▶ Compared to DNN results, 57% improvement
- ▶ Overfitting decreased compared to DNN
- ▶ Can lead to model not performing well on new data

# Accuracy Recap

- ▶ Training accuracy was 90%
- ▶ Test accuracy was 91%
- ▶ Validation accuracy was 81%

# Recommendations

- ▶ CNN results showed improvements compared to DNN
- ▶ Larger difference between our training and test accuracy to reduce overfitting
- ▶ Higher testing accuracy
- ▶ See a higher validation accuracy (81%)
- ▶ Address by adding more images to our datasets
- ▶ Experiment with more tuning/ combination of tuning to models

# Next Steps

- ▶ Our model shows promising results
- ▶ Increase the accuracy of model even more
- ▶ Reduce time doctor spends on evaluating x-ray
- ▶ Start off with diagnosing pneumonia
- ▶ Can lead to diagnosing broken bones, bone cancer, even breast cancer



# Questions and Contact Info

- ▶ Questions?
- ▶ Any additional inquiries can be directed to linkedin:
  - ▶ <https://www.linkedin.com/in/alejandro-harrison-948034108/>