Nicholas Kellogg

DS4002 - Hook Document

4/28/2025

Task: Train a central neural network model that can quickly differentiate between COVID-19 infected lung x-ray images and Normal or Pneumonia infected lung x-ray images.

The Scenario: It's 2021 and the global COVID-19 pandemic is raging around the world. You have been brought in as a data science consultant to train a model that can identify different infections within lung x-ray images. Your mission is urgent and vital: build a machine learning model that can automatically distinguish between COVID-19 infected lungs, Pneumonia-infected lungs, and healthy lungs based solely on x-ray images. Your work could help overwhelmed hospitals triage patients faster, prioritize care more effectively, and ultimately save lives. Your dataset will be based on previously sourced images from lung x-rays, with around 6,000 images already found.

You'll need to dive into real-world medical imaging data, understand how infections present themselves visually, and train a convolutional neural network (CNN) to detect these subtle differences. Through this, your model's diagnosis will be referred to doctors who can then administer treatment in record time.

Your Materials: https://github.com/kellogg9/DS4002-CS3