Chest X-Ray Image Classification: Project Proposal

Zach Wimpee

1 Motivation, Problem, and Client

The COVID-19 pandemic has presented seemingly all areas of society with a host of new and unprecedented challenges, many of which are faced in the healthcare field.

For example, while many people living in developed countries are fortunate enough to have relatively easy access to COVID-19 testing, those living in developing or impoverished areas may not be afforded this same luxury. The virus, however, does not discriminate, and therefore additional tools are needed to aid in diagnosing infected patients when available tests are limited or perhaps even nonexistent.

2 Proposed Solution and Approach

Consider the following hypothetical scenario: Dr. Smith runs a clinic for the impoverished in a developing country, and has a new patient exhibiting a variety of symptoms including chest pain, coughing, and fever. Dr. Smith suspects that the patient has pneumonia, but is unsure if it is a viral or a bacterial infection. He would like to rule out the possibility of a COVID-19 infection, but the clinic is still waiting to be resupplied with a shipment of tests. The patient's condition is declining, and the fastest diagnostic tool available is a chest X-ray.

In this hypothetical scenario, Dr. Smith and his patient would benefit from a tool that could determine from an X-ray whether or not the pneumonia is being caused by a COVID-19 infection. The project being proposed here is an exploration into the potential use of neural network machine learning models in developing such a tool.

2.1 Data

A brief search identified a promising dataset to use for this project's development. It is a curated set of chest X-rays for which each image is labeled one of either COVID-19, normal, viral-pneumonia, or bacterial-pneumonia.[1] The dataset has already been checked for duplicates samples and defective images, and therefore minimal cleaning and preprocessing will be required.

2.2 Deliverables

The final goals for this project are twofold: The first is to design and train a neural network model to classify chest X-rays, and achieve an acceptable level of performance upon test set

---

[1] https://data.mendeley.com/datasets/9xkhgts2s6/2

evaluation. After acquiring a decent model, the final goal will be to implement visualizations that give insight into what the model deems important in its decisions so as to provide model interpretability.