Project 4 – Pediatric Pneumonia Identification via Chest X-Ray Images

Group 1: Devin Fogerty, Seth Gard, Leah LoSchiavo

Our goal for this project is to train a machine learning model on a set of chest x-ray images containing healthy patients, patients infected with bacterial pneumonia, and patients infected with viral pneumonia, to be able to identify positive cases and eventually distinguish between the two types of infections.

Outline:

1. Importing Data
   * Import all training data and convert from raw images to arrays (dataset is pre-split between test and training data)
2. Cleaning Data
   * Ensure that all images are properly labeled and sorted
   * Low quality scans have already been removed by uploader
3. Model Training
   * Train convolutional network model using tensorflow to distinguish between healthy and pneumonia patients
   * Once accuracy of the model is high, enhance model to distinguish between bacterial and viral infections
4. Evaluate Model
   * Tweak model until accuracy is above the desired 75% mark
   * Ensure model isn’t overfitting to dataset
5. Analysis
   * Determine viability of model and discuss results
6. Create Slides
   * Create presentation for class

Dataset: https://www.kaggle.com/datasets/paultimothymooney/chest-xray-pneumonia/data