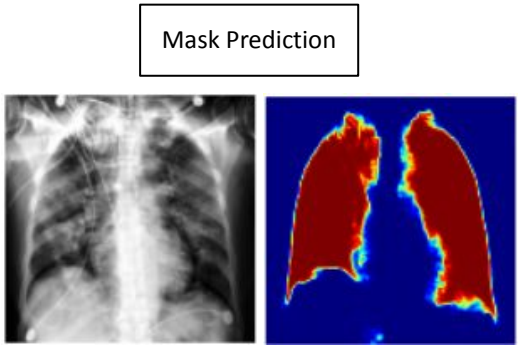
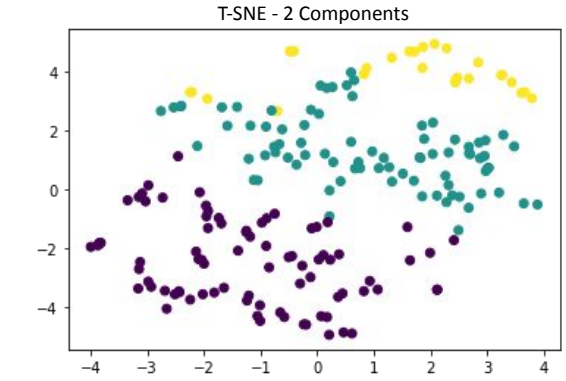
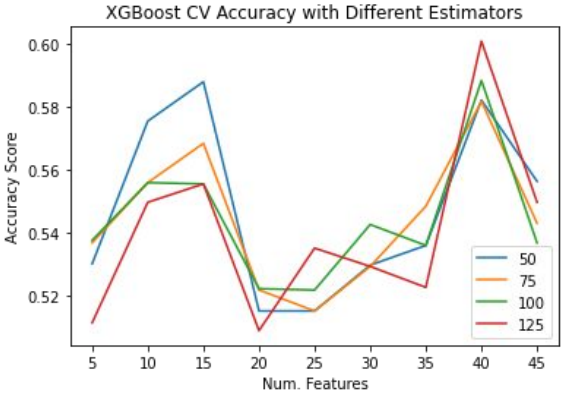
Covid-19 Lung Infection Severity Prediction

Luke Robbins, Yeyun Xu, Jeremiah Joseph, Aatifa Khan, Justin Pham

Goal: Create a prediction model that can predict a lung infection severity score based on features extracted from chest X-rays. Using deep learning techniques and several classification models, we attempt to classify each image according to a subset (0, 1, 2) of the Brixia infection severity score index.







Approach:

- Lung segmentation using U-Net neural net architecture.
- Mask creation for feature extractor.
- Map given Brixia score values to 0, 1, or 2 (severity).
- Use Python tool, PyRadiomics, to extract features by passing the X-Ray and mask. These features are necessary to explore supervised learning techniques.
- Perform T-SNE and PCA as dimensionality reduction techniques to visualize and filter 109 total features.
- Compare accuracies of models such as Lasso regression, XGBoost, and Random Forest by cycling through hyper-parameters, and choosing ANOVA selected features or Principal Components.
- The best performing model was an XGBoost classifier using 40 ANOVA selected features with a 60.1% CV accuracy, and 58.9% accuracy rate on new test data.