Problem Statement Worksheet (Hypothesis Formation)

How can HUP's Lung Center implement a system to correctly classify chest x-rays as either pneumonia detected or normal with an 80% F1 score by the end of 2021?

Context

The Paul F. Harron Jr. Lung Center is a department at the Hospitals of the University of Pennsylvania (HUP). The Lung Center offers patient-focused care combining research and medical advancements to provide cutting-edge treatments. The center asked that a machine learning mechanism be developed to screen patients for pneumonia. Pneumonia can be caused by either bacteria or virus. Bacterial pneumonia typically results in focal lobar consolidation. Viral pneumonia typically is more diffused and results in a pattern present in both lungs. Due to the complex nature of diagnosing images, the center just wants the program to detect abnormalities caused by pneumonia for radiologists to examine further.

Criteria for success

- Classify chest x-ray images as pneumonia detected or normal with least an 80% F1 score
- Be available for clinical use by the end of 2021

Scope of solution space

- Correctly classify chest x-rays into pneumonia detected or normal

Constraints within solution space

- Only investigating the presence of pneumonia or normal, no other chest diseases/ abnormalities
- Data not labelled in terms of disease progression or severity

Stakeholders to provide key insight

- **Jason D Christie, MD, MS.**, Chief, Division of Pulmonary, Allergy and Critical Care
- **John I. Gallin, M.D.**, Clinical Center Chief Scientific Officer
- **Jon W. McKeeby, DCc**, Biomedical Translational Research Information Systems
- **Elizabeth Jones, MD**, Director of Radiology and Imaging Services Department

Key data sources

- http://www.cell.com/cell/fulltext/S0092-8674(18)30154-5