

# Pneumonia Detection

## Description Of Data Set

Sangram Patil  
Vikas Ramaswamy  
Adarash Pal

3. April 2023

- 1 What is Dataset ?
- 2 Description Of Dataset
- 3 Specification

# What is Dataset ?

# Dataset I

- A dataset is a collection of data that is organized and presented in a structured.
- It is typically used for analysis, research, and machine learning purposes. A dataset can be as simple as a spreadsheet of numbers or as complex as a database of multimedia files.
- A dataset is simply a set of information that can be used to make predictions about future events or outcomes based on historical data.
- Data can be collected from a variety of sources, such as surveys, sensors, social media, or web scraping, and can be stored in various formats, including text, images, audio, and video.
- The data set used for our project is Image data which is taken from Kaggle dataset

# Description Of Dataset I

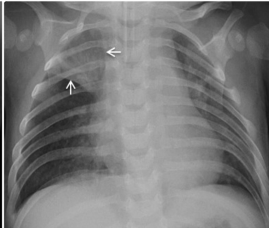
- For Pneumonia Detection Chest X ray images are used . Chest X-ray images (anterior-posterior) were selected from retrospective cohorts of pediatric patients of one to five years old from Guangzhou Women and Children's Medical Center, Guangzhou. All chest X-ray imaging was performed as part of patients' routine clinical care.
- For the analysis of chest x-ray images, all chest radiographs were initially screened for quality control by removing all low quality or unreadable scans. The diagnoses for the images were then graded by two expert physicians before being cleared for training the AI system. In order to account for any grading errors, the evaluation set was also checked by a third expert. **[kermay2018identifying]**

# Example I

Normal



Bacterial Pneumonia



Viral Pneumonia



Illustrative Examples of Chest X-Rays in Patients with Pneumonia

# Example I

- The normal chest X-ray (left panel) depicts clear lungs without any areas of abnormal opacification in the image. Bacterial pneumonia (middle) typically exhibits a focal lobar consolidation, in this case in the right upper lobe (white arrows), whereas viral pneumonia (right) manifests with a more diffuse “interstitial” pattern in both lungs. **[kermany2018identifying]**

# Specification I

- The dataset is organized into 3 folders : train, test, validation
- Subfolders for each image category : Pneumonia and Normal
- There are 5,863 X-Ray images in JPEG format and 2 categories : Pneumonia and Normal



Thank you  
for your attention