**Adaovi - Final Project**

# **Problem Statement**

The primary objective is to develop an advanced deep-learning model that significantly improves the accuracy and efficiency of lung cancer classification based on medical imaging data. By addressing the outlined challenges, the model aims to enable early and reliable detection of lung cancer, ultimately contributing to better patient outcomes and reducing the burden on healthcare systems associated with late-stage diagnoses.

# **Dataset**

The dataset is organized into 3 folders (train, test, val) and contains subfolders for each image category (Pneumonia/Normal). There are 5,863 X-Ray images (JPEG) and 2 categories (Pneumonia/Normal).

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. To account for any grading errors, the evaluation set was also checked by a third expert.

# **Acceptance Criteria**

Predict - Normal / Pneumonia

***Model Accuracy/Precision/Recall: more than 80***