Notebook 1: Stratified Sampling and Exploratory Data Analysis (nih-phase1-stratified-sampling-and-eda.ipynb)

- **Objective:** To create a manageable, representative subset of the full NIH Chest X-ray dataset and perform initial exploratory data analysis.
- **Stratified Sampling:** A stratified sampling technique was employed to downsize the original dataset of 112,120 images from over 30,000 unique patients. I sampled approximately 40% of the patients, resulting in a downsized dataset of 57,302 images from 13,224 patients. This method ensured that the distribution of pathologies in the smaller dataset mirrored that of the original, which is crucial for unbiased model training.
- **Data Preparation:** All 57,302 images in the downsized dataset were resized to a standard 256x256 resolution and converted to grayscale to ensure consistency and computational efficiency.
- **Exploratory Data Analysis (EDA):**
    - Class Distribution: An analysis of the pathology distribution revealed a significant class imbalance, with "No Finding" being the most common category.
    - Co-occurrence of Pathologies: A heatmap was generated to visualize the correlation between different pathologies. It was observed that certain conditions, such as Infiltration and Atelectasis, frequently co-occur.
    - Data Cleaning: During EDA, an issue with inconsistent labeling for "Pleural_Thickening" was identified and dropped from the classification process due to inconsistency caused by the class.

Notebook 2: Model Training (nih-phase2-model-training.ipynb)

- **Objective:** To train a baseline deep learning model on the downsized dataset for multi-label pathology classification.
- **Model Architecture Evaluation:** Several architectures were evaluated before finalizing the choice. An EfficientNet model was initially tested but was discarded due to unsatisfactory accuracy in the preliminary training phase. Similarly, while MobileNetV3 Mini offered a low training time, its accuracy was also found to be insufficient. On the other end of the spectrum, MobileNetV3 Large proved to be computationally prohibitive, with an estimated training time exceeding 12 hours for just 5 epochs, making it unsuitable for our project timeline and resources.
- **Final Model Selection:** A MobileNetV2 architecture, pre-trained on ImageNet, was ultimately selected as the base model. This comparative analysis confirmed that MobileNetV2 offered the best balance of performance and efficiency for our needs.
- **Training Process:**
    - The model's top layers were replaced with a new classification head suitable for our 13-label classification task.
    - The model was initially trained for 5 epochs with the base layers frozen, allowing the new classifier to adapt to the X-ray data.
    - The training process was monitored using AUC-ROC as the primary metric, achieving a validation AUC-ROC of approximately **0.73.**

Notebook 3: Fine-Tuning (nih-phase-3-fine-tuning.ipynb)

- **Objective:** To improve the model's performance by unfreezing the base layers and fine-tuning it with a lower learning rate.
- **Fine-Tuning Process:**
    - The pre-trained MobileNetV2 base was unfrozen to allow for end-to-end training.
    - The model was re-compiled with a very low learning rate (1e-5) to prevent catastrophic forgetting of the learned features.
    - The model was then fine-tuned for an additional 5 epochs.
- **Results:** Fine-tuning resulted in a notable improvement in performance, with the validation AUC-ROC increasing to approximately **0.77**. This indicates that allowing the model to adapt its core feature extractors to the specifics of chest X-ray images was a successful strategy.

Notebook 4: Further Fine-Tuning (nih-phase-4-further-fine-tuning.ipynb)

- **Objective:** To attempt to extract additional performance gains by fine-tuning the model again with an even lower learning rate.
- **Process:** The model from Phase 3 was loaded and re-compiled with an extremely low learning rate of 1e-6. It was then trained for an additional 5 epochs.
- **Results:** This final fine-tuning phase yielded a marginal but positive improvement, pushing the final validation AUC-ROC to approximately **0.801**.

Current Work and Next Steps:

- **Training with Full Dataset** (MobileNetV2-NIH-Full-Dataset.ipynb): The primary focus is now on training the fine-tuned model architecture on the entire NIH Chest X-ray-14 dataset. I have completed the initial model training with an AUC-ROC value of **0.66**. I am trying to improve this value with subsequent training.
- **Front-End Development:** In parallel, the development of the front-end user interface has begun and I am confident I will be able to complete it within this week.