Multi-Label Thoracic Disease Classification from Chest X-Rays

1. Problem Statement

Develop a deep learning-based diagnostic model to identify multiple thoracic diseases from chest X-rays. Each image can exhibit several pathologies, making it a multi-label classification task. The objective is to assist radiologists in minimizing diagnostic error, workload, and reporting time.

2. Context

Thoracic diseases such as pneumonia, pleural effusion, cardiomegaly, and pneumothorax are leading causes of global mortality. Manual interpretation of X-rays is slow and prone to human variability. AI-driven models can enhance diagnostic precision and accessibility, especially in underserved regions.

3. Criteria for Success

- Achieve ≥85% average AUC across all disease categories.
- Provide explainability via Grad-CAM visualizations.
- Ensure reproducibility through notebooks and scripts.
- Deliver clear executive summary for clinical use.

4. Scope of Solution Space

Implement CNN and transformer models (DenseNet121, EfficientNet, ViT) using transfer learning. Techniques such as class weighting, focal loss, and augmentation will mitigate imbalance. Evaluate using AUC, F1-score, precision, and recall.

5. Constraints

- High computational requirements for GPU-based training.
- Imbalanced data across disease classes.
- Strict ethical and interpretability requirements.
- Usage limited to open-source, de-identified medical datasets.

6. Stakeholders

Primary: Radiologists, hospitals, and healthcare networks.
Secondary: AI researchers, imaging solution providers, regulatory bodies.

7. Data Sources

The NIH ChestX-ray14 dataset (112,000+ frontal-view scans labeled across 14 diseases) will be used. Supplementary datasets such as CheXpert and MIMIC-CXR will improve robustness. Data will undergo preprocessing, normalization, and augmentation.

8. Proposed Approach

- Perform EDA on disease frequency, correlation, and sample visualization.
- Preprocess and encode labels for multi-class learning.
- Fine-tune pre-trained CNN or transformer model.
- Evaluate performance using AUC, F1, precision, recall, Grad-CAM.
- Deploy prototype using Streamlit for clinician interaction.

9. Deliverables

- EDA, modeling, and evaluation Jupyter notebooks.
- Model performance report (PDF).
- Stakeholder presentation slides.
- Version-controlled ML pipeline repository.

10. Expected Impact

This AI-powered tool can streamline diagnostic workflows, improve accuracy, and reduce radiological delays. By supporting rather than replacing clinicians, it promotes trustworthy human-AI collaboration in medical imaging.

Performance Summary

Key Metric	Target Value
Average AUC	≥ 0.85
Precision-Recall (per class)	> 0.80
Explainability (Grad-CAM)	Heatmaps per disease label
Deployment	Streamlit-based clinical dashboard