# **Project Proposal: Chest X-Ray Classification Model with Web App Integration**

## **Goal**

My goal for this project is to create a deep learning model for classifying chest X-ray images to achieve a target accuracy of 99%. The Dataset I will be using is the “VinBigData Chest X-ray Abnormalities Detection dataset from Kaggle” from Kaggle. The final product will ideally include a web application to showcase the model’s functionality. However, if I fail to achieve my target accuracy within the next month, I will transition to a Google Slide as a fallback.

## **Project Objectives**

1. **Model Optimization**: Enhance my already worked on model to achieve 99% accuracy by fine-tuning hyperparameters and experimenting with the models architecture.
2. **Dataset Details**: The VinBigData Chest X-ray Abnormalities Detection dataset contains a diverse set of chest X-ray images with labeled abnormalities ranging from 0 - 14(Limited number of readings).
3. **Web Application/Presentation**: Build a web app to allow users to upload chest X-ray images and or view classification results, displaying the model’s accuracy in a “fashionable” way. If deadlines aren’t met, a simple Google Slide will be used to present the results.

## **Methodology**

### **Dataset**

* **Source**: [Link](https://www.kaggle.com/competitions/vinbigdata-chest-xray-abnormalities-detection).
* **Description**: The dataset includes thousands of chest X-ray images with annotations for various abnormalities, suitable for training a robust classification model.
* **Preprocessing**: Apply image normalization, resizing, and data augmentation to improve model generalization.

### **Model Development**

* **Current Status**: The model has been built already, but still needs tweaking to improve the accuracy from 90% to 99%. Web app not developed at all.
* **Optimization Strategies**:
  + Experiment with architectures like ResNet, DenseNet, or EfficientNet.
  + Adjust learning rates and batch sizes
  + Implement techniques like transfer learning, and batch normalization to reduce overfitting.
* **Evaluation Metrics**: Accuracy, precision, recall, F1-score, and confusion matrix to visualize the models performance.

### **Web Application**

* **Framework**: Use Flask for the backend since it’s python based, and React for the frontend for more interactivity.
* **Functionality**: Allow users to upload X-ray images, process them through the model, and display classification results with scores.
* **Deployment**: Host the app on a platform like Heroku (better for python apps) for accessibility during the presentation.
* **Fallback Plan**: If the web app is not completed by the deadline, a Google Slides presentation will be prepared, including:
  + Project overview and objectives.
  + Dataset and model details.
  + Performance results.
  + Visualizations.
  + Future improvements and applications.

## **Timeline**

* **Week 1-3**: Finalize model architecture and achieve target accuracy.
* **Week 4-5**: Develop the web application.
* **Week 6**: Final touches on the web app or prepare Google Slides presentation.
* **End of Next Month**: Submit project and present deliverables.

## **Resources**

* **Software**: Python, TensorFlow/Keras, Flask, React, Kaggle API for dataset access.
* **Dataset**: VinBigData Chest X-ray Abnormalities Detection (Kaggle).
* **References**: Research papers on chest X-ray classification, online tutorials for web app development.

## **Potential Challenges**

* **Model Accuracy**: Achieving 99% accuracy may take longer than expected/want.
* **Web App Development**: Limited time to learn and implement a web application to the extent that I would like.
* Adding an LLM using Gemini Pro api to provide insight on the readings of the predictions my model made.
* In the future, I plan to add a section of the metrics and brief descriptions of the code with snippets.

## **Conclusion**

This project is a reflection of my work ethic and my growing knowledge in the field of Machine Learning. It is designed not only to showcase my technical skills but also to contribute to real-world healthcare applications. The model I’m developing has the potential to assist in medical diagnostics by accelerating the diagnostic process, ultimately helping radiologists and patients receive quicker treatments. This goal is also personal to me, as my sister is a doctor, and I hope to create tools that will be able to support her and her colleagues in the medical field. With this proposal, I hope to have gained your approval to proceed with the project and I’m open to any feedback or suggestions you may have.

**Student Name**: Andrew Lin  
**Date**: May 8, 2025  
**Course**: AI and Machine Learning Course  
**Instructor**: Michael Pascuzzi