March 19, 2025

C964: Computer Science Capstone

Brad Pfahl

Task 2 parts A, B, C and D

Part A: Letter of Transmittal

Dr. John Dorian
Sacred Heart Hospital
San DiFrangeles, California

March 19, 2025

Dear Dr. Dorian,

This letter proposes the implementation of an artificial intelligence (AI)-driven system designed to enhance our radiology department's efficiency and diagnostic accuracy. This initiative aims to address the increasing workload faced by our radiologists and improve patient outcomes through timely and precise diagnoses. The problem that this proposal seeks to address is the ever-increasing workload put on radiologists in the clinical setting. Our health network is facing a shortage of specialists, including radiologists. We face the department being short-staffed and overworked, leading to reduced turnaround times on imaging interpretations and decreased patient care. To solve this, we need to reduce their workload so they can efficiently perform their tasks.

We propose developing and integrating an AI-based image recognition system that automatically analyzes chest X-rays to identify abnormalities. In the clinical setting, these chest X-rays make up many of the studies our radiologists have to read. This system will serve as an initial screening tool, flagging images that require further review by our radiologists, thereby streamlining the diagnostic process.

Benefits to the Organization:

- Enhanced Efficiency: The AI system will quickly process routine chest X-rays, allowing radiologists to focus on more complex cases.
- Improved Diagnostic Accuracy: By providing a second level of analysis, the system can help reduce the likelihood of missed diagnoses.
- Reduced Workload: Automating the initial screening process will alleviate some of the pressures on our radiology staff, improve job satisfaction, and reduce burnout.
- Cost Savings: Over time, increased efficiency and accuracy can lead to reduced operational costs and better resource allocation.

The projected cost for developing and deploying this AI system is approximately $80,000. This estimate includes software development, salary costs for staff involved in the project, integration with existing systems, staff training, and initial maintenance. The project is estimated to take approximately three months to integrate fully and, at a satisfactory point, into the clinical workflow. We will start implementing the project small, starting with our senior radiologist, until it is deemed acceptable for our entire radiologist staff to implement into the system.

With over a decade of experience as a radiologic technologist and advanced training in computer science, I am uniquely positioned to lead this project. My understanding of radiology workflows and technical expertise ensures the solution will be effective and seamlessly integrated into our existing processes.

I look forward to discussing this proposal further and can answer any questions at your convenience.

Sincerely,

Bradley Pfahl

Bradley Pfahl BS RT(R)(MR)

Part B: Project Proposal Plan

Project Summary

This project is focused on using machine learning and image recognition software that fits the needs of a radiology department in a hospital setting. The software will assess chest X-rays done in the clinical setting and flag them if there is an abnormality. This is being created to address the increasing workload of radiologists in the Sacred Heart health network. They are a medium-sized health network that operates as a non-profit organization. They are facing a shortage of radiologists, and the current ones they have are getting an increasing amount of work. Fazekas et al. (2022) explored the application of artificial neural networks (ANNs) in healthcare, focusing specifically on radiology. They emphasize how radiologists' workload has risen and that they may be expected to interpret an image every 3-4 seconds of an 8-hour workday (Fazekas et al., 2022). This reinforces our need to relieve some of the workload in this profession.

We propose developing an automated image recognition system that utilizes convolutional neural networks (CNNs) to classify chest X-rays as normal or abnormal. This solution will serve as an initial screening tool, allowing radiologists to prioritize high-risk cases while reducing the time spent reviewing regular scans. The model will be trained on a large, well-annotated dataset of chest radiographs, leveraging deep learning techniques for feature extraction and classification. We will start by using the NIH chest X-ray dataset. This dataset is a great resource that allows us to start creating a model without processing our images into a dataset first. By automating the initial assessment of chest X-rays, this solution directly aligns with the organization’s need to reduce radiologists' workload, enabling them to focus on complex cases and improve overall efficiency in patient care.

Data Summary

-Data Source

The dataset for this project will be sourced from the publicly available and de-identified chest X-ray dataset, NIH Chest X-rays. If access is granted, an internal hospital dataset may be used, ensuring compliance with HIPAA regulations.

-Data Collection Method

The dataset will be collected through a publicly available dataset. It is available on Kaggle.com and is the NIH chest X-ray dataset. A future method is to collect the data from our health network. This will require more time to ensure that the data collected complies with HIPAA standards and is consistently done so it is compatible with the software.

Using publicly available datasets ensures a large, diverse dataset with standardized labeling, reducing the need for extensive manual annotation. If internal hospital data is included, it provides real-world clinical relevance, improving model performance on site-specific cases.

A significant limitation is the variability in image quality and annotation standards across datasets, which may introduce biases. Additionally, acquiring new labeled data is time-consuming and may require radiologist involvement for verification.

-Quality and Completeness of Data

The dataset will undergo preprocessing to ensure quality and consistency. All identifiable patient information will be removed to maintain HIPAA compliance. Images will be rescaled, normalized, and converted to a uniform format to ensure compatibility with the machine-learning model. Data augmentation techniques such as rotation, contrast adjustment, and noise addition may be applied to improve generalization.

-Precautions for Sensitive Data

Strict data protection protocols will be enforced to ensure compliance with HIPAA and institutional privacy regulations. Patient data will be de-identified before storage or processing, and access will be restricted to authorized personnel. Secure, encrypted storage solutions will be utilized for any locally stored datasets. Communication about data usage will be limited to necessary personnel, and model predictions will be carefully validated before clinical deployment to prevent incorrect diagnoses from affecting patient care.

Implementation

The development will follow the CRISP-DM (Cross-Industry Standard Process for Data Mining) methodology, which will be used to guide the machine learning project. This structured approach ensures thorough solution exploration, development, and deployment. The following steps outline how CRISP-DM will be applied to this project:

1. Explore:

- Data Collection and Preprocessing: Acquire a relevant chest X-ray (normal and abnormal) dataset, ensuring it is appropriately labeled. Preprocessing includes resizing images, normalizing, and handling missing or corrupt data.
- Exploratory Data Analysis (EDA): Investigate the dataset to understand patterns, check for any biases, and visualize key attributes to inform feature selection for the model.

2. Modify:

- Feature Engineering: Identify and extract relevant features from the images, such as texture, shape, and edges, which are essential for classification. Augment the dataset with transformations like rotation or flipping to improve model generalization.
- Data Augmentation: Apply techniques like flipping, rotating, or adjusting contrast to artificially increase the size of the training data, improving the model's robustness.

3. Model:

- Model Selection: We will implement Convolutional Neural Networks (CNNs), which are particularly well-suited for image recognition tasks. Yasaka et al. (2018) discuss the growing role of deep learning, particularly convolutional neural networks (CNNs), in radiology. CNNs are effective for image recognition because they can automatically extract important features from medical images without manual preprocessing.
- Training: Train the model on the processed chest X-ray dataset, adjusting hyperparameters to optimize performance, such as the learning rate, batch size, and number of epochs.
- Validation: Use cross-validation techniques to evaluate the model's performance and ensure it generalizes well to new, unseen data.

4. Assess:

- Performance Evaluation: Use metrics such as accuracy, precision, recall, and F1-score to evaluate the model's effectiveness in distinguishing normal from abnormal X-rays.
- Iterative Improvement: Fine-tune the model by testing different algorithms, adding more data, or using advanced techniques like transfer learning to improve classification accuracy.
- Deployment Readiness: Assess the solution's integration readiness with existing PACS systems and its ease of use for radiologists.

Timeline

March 24, 2025– Proposal Accepted

- Milestone: Formal approval of the project proposal and the project planning phase initiation.

April 7, 2025- Technical Proof of Concept Presented

- Milestone: Present a working proof of concept demonstrating the feasibility of the image recognition model using a sample dataset. The proof of concept will include preliminary model results and performance metrics.

April 24, 2025– Model Refinement and Dataset Finalization

- Milestone: Complete model training with the final dataset, including feature engineering and augmentation. Perform model evaluation using performance metrics such as accuracy, precision, and recall.

May 15, 2025– System Integration Testing

- Milestone: Begin integration testing of the machine learning model with the existing radiology workflow systems, including any necessary modifications for PACS integration. Conduct user acceptance testing (UAT) with a small group of radiologists.

June 16, 2025– Final Deliverables and Full Deployment

- Milestone: Deliver the final machine learning model, documentation, and user training materials. Deploy the solution in a live clinical setting and ensure ongoing model monitoring and support.

Sprint	Start	End	Tasks
1	March 24, 2025	April 7, 2025	Data collection, preprocessing, and EDA
2	April 7, 2025	April 24, 2025	Feature engineering, data augmentation, and model selection
3	April 24, 2025	May 15, 2025	Model training, hyperparameter tuning, and validation
4	May 15, 2025	June 16, 2025	System integration, UAT, final evaluations

Evaluation Plan

The machine learning model's performance will be evaluated using a combination of standard metrics for image classification, including accuracy, precision, recall, and F1-score. Given the medical application, recall (sensitivity) will be prioritized to minimize false negatives, ensuring that abnormal chest X-rays are not overlooked. Additionally, the ROC-AUC (Receiver Operating Characteristic - Area Under the Curve) score will assess the model's ability to distinguish between normal and abnormal cases.

To ensure robustness, the model will be tested on a holdout validation set and an independent test dataset to evaluate generalization. Cross-validation techniques will be employed to reduce overfitting. Beyond technical metrics, usability and real-world effectiveness will be measured by gathering feedback from radiologists on the model's predictions, focusing on its impact on workflow efficiency and decision-making. Our radiologists will give us valuable information on the performance and accuracy of the model that we can take forward to future iterations. Kromrey et al. (2024) investigate the effectiveness of AI-based software in detecting four major thoracic pathologies—fractures, pleural effusion, pulmonary nodules, and pneumonia—in chest radiographs. Their study at the University of Medicine Greifswald analyzed 1,506 patient X-rays and compared AI-generated diagnoses with radiologists' interpretations. The results demonstrated that the AI software had a higher detection rate than radiologists (18.5% vs. 11.1%), though it also produced a "nonnegligible number of false positives" (Kromrey et al., 2024). This demonstrates the importance of flagging these false positives when we do find them and adjusting the model to improve future iterations.

Resources and Costs

Resource	Description	Cost
Data Collection Tools	Software tools to gather and process medical imaging datasets for training. Used in Sprint 1 for data preprocessing and exploratory data analysis (EDA).	$2,500
Cloud Computing Services	Cloud services (e.g., AWS, Google Cloud) for model training and storage. Primarily used in Sprints 2 and 3 for model training, testing, and optimization.	$10,000
ML Development Software	Software tools and licenses for machine learning development. Used throughout Sprints 1-3 for model development and experimentation.	$3,000
Medical Imaging Dataset	We will be using our data in the health network and the dataset from the NIH, which is in the public domain.	$0
Data Scientist	Responsible for model development, feature engineering, and algorithm optimization. Active in Sprints 1-3.	$20,000
Machine Learning Developer	Implements model training pipelines, integrates AI models with PACS and optimizes deployment. Active in Sprints 2-4.	$15,000
Software Tester	Provides expert feedback, validates AI-generated results, and assists in performance evaluation. Active in Sprints 3-4.	$7,500
Radiologist	Conducts integration testing and user acceptance testing (UAT) and ensures reliability of AI-assisted diagnosis. Active in Sprint 2-4.	$15,000
Testing and Evaluation Tools	Tools for performance testing and system integration (software for integration with PACS). Used during Sprint 4 to ensure real-world readiness.	$4,000
Training and Documentation	Costs for user training and documentation development. Conducted in the final phase after Sprint 4.	$3,000
	Total	$80,000

Part C: Application

The completed application has been submitted alongside this document in a shared Google Drive folder, with view-only permissions to ensure privacy and security. The folder contains:

- Two Google Colab notebooks:
    - Model_Training_Source_Code.ipynb (source code and training notebook)
    - Chest_Xray_Prediction.ipynb (demonstration and user-friendly prediction notebook)
- The trained model file (cnn_xray_model.h5)
- Five sample chest X-ray images

The source code was developed using Jupyter Notebook within PyCharm, then finalized and uploaded to Google Colab for accessibility. Both notebooks are designed to run entirely within Colab's environment, requiring no special setup or local dependencies.

The dataset itself was stored locally during development. However, functionality and demonstration of the application are fully accessible in Colab. Evaluators can run the Chest_Xray_Prediction.ipynb notebook and interact with the model and sample images without the need to mount Google Drive or access additional files outside the provided folder.

If desired, users may download the NIH Chest X-ray dataset directly from Kaggle and update the file paths in the Model_Training_Source_Code.ipynb notebook to integrate their local copy of the dataset. A clear comment has been included in Section '3. Upload dataset' of that notebook indicating exactly where file path changes should be made.

This setup ensures a smooth, secure evaluation experience while also allowing for scalability or future integration of additional datasets.

Part D: Post-implementation Report

Solution Summary

The project addressed the problem of increasing workload and staffing shortages among radiologists, particularly in interpreting chest X-rays. A convolutional neural network (CNN) model was developed to classify chest X-ray images as normal or abnormal. The application automated preliminary image assessments, allowing radiologists to prioritize urgent cases and increase overall efficiency. The solution provided an accessible, deployable tool via Google Colab, enabling users to run the model and make predictions without requiring complex local setup. This project will continue to provide a helpful resource to our radiologists to streamline the interpretation of one of their most commonly ordered studies and will only improve with time as we receive more feedback and further train the model.

Data Summary

The raw data was sourced from the publicly available NIH Chest X-ray dataset hosted on Kaggle. This dataset contained de-identified patient information, compliant with HIPAA regulations, and included chest X-ray images and corresponding labels indicating medical findings. The data that was copied locally can manipulated in locally maintained drives.

During development:

- The dataset's labels were simplified to binary categories: "Normal" and "Abnormal."
- The dataset was split into training, validation, and testing subsets using an 80/20 split.
- Images were preprocessed in real-time using Keras' ImageDataGenerator:
  - Images were resized to 128x128 pixels.
  - Pixel values were normalized to the 0-1 range.
  - Images were loaded and fed in small batches to conserve memory.

This preprocessing approach ensured consistent data handling throughout the design, development, and testing phases.

Machine Learning

For each employed method (at least one is required), provide the following:

- Identify the method and what it does (the "what").
- Describe how the method was developed (the "how").
- Justify the selection and development of the method (the "why").

A Convolutional Neural Network (CNN) was employed to classify chest X-rays.

The CNN model automatically learned spatial hierarchies of features within grayscale chest X-ray images, identifying patterns associated with normal and abnormal findings.

- The model architecture included:
  - Two convolutional layers followed by max pooling.
  - A fully connected dense layer with dropout regularization.
  - A final sigmoid activation output layer for binary classification.
- The model was trained using binary_crossentropy loss and the Adam optimizer.
- ModelCheckpoint was implemented to save the best model based on validation accuracy.

CNNs were chosen due to their proven ability to detect spatial features in image data, making them ideal for medical imaging applications. The architecture balanced complexity and efficiency, providing reliable classification without overfitting.

Validation

For each employed method described in the section above, provide the following:

- A proper validation method (typically a model performance metric).
- Results of the validation method *or* a future to obtain those results.

Validation of the model was performed using:

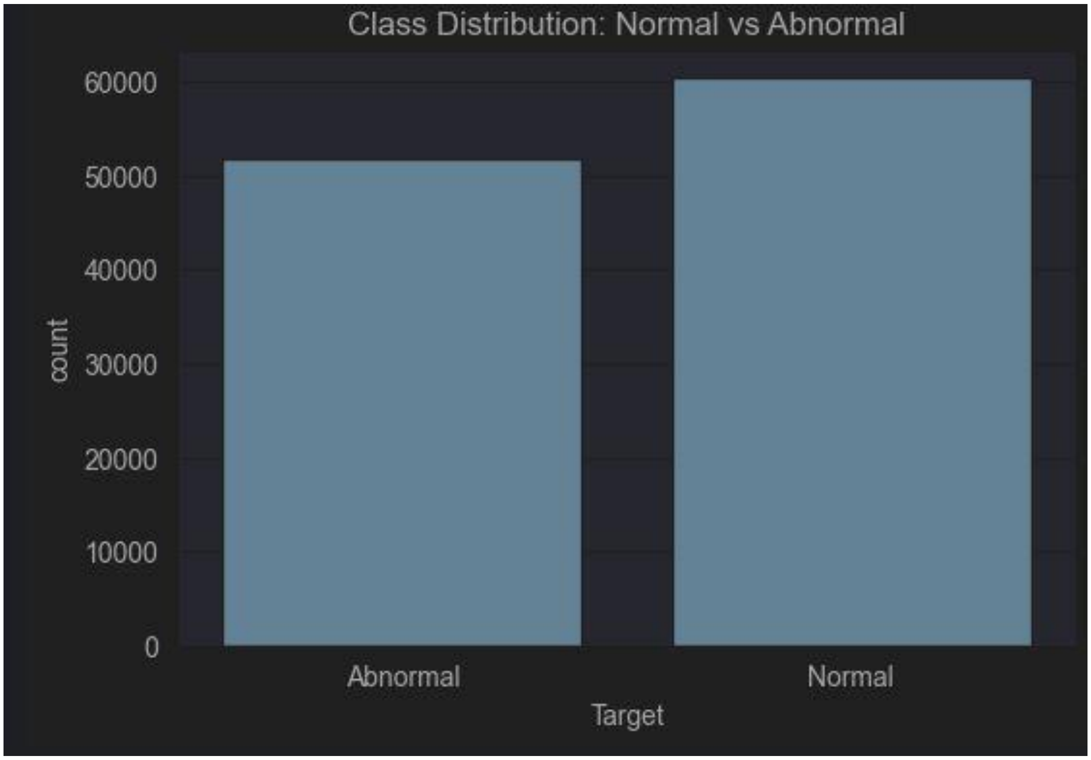
- Accuracy
- Precision
- Recall
- F1-Score
- Confusion Matrix

The validation was conducted on the testing dataset, which contained unseen images. Training and validation loss/accuracy were plotted across epochs to assess model performance and learning behavior visually.

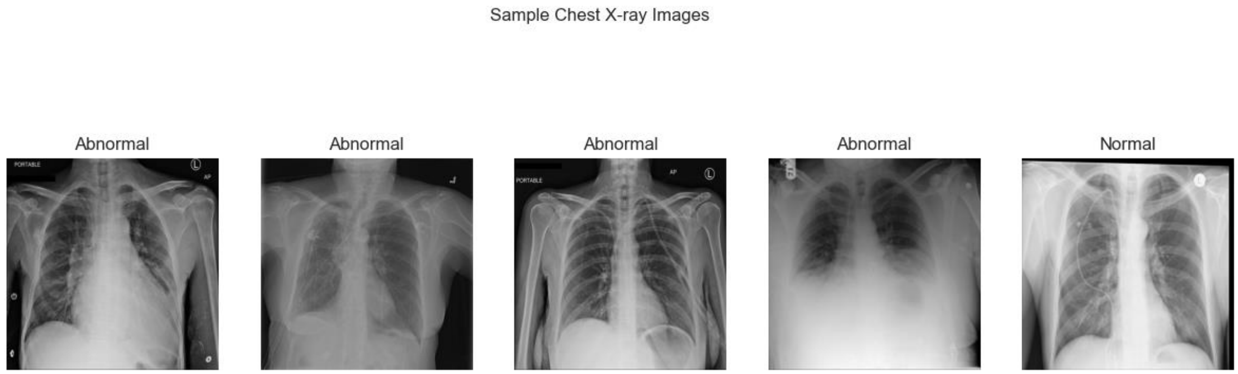
Visualizations

The following visualizations were included and embedded within the Google Colab notebook in the Model_Training_Source_Code.ipynb:

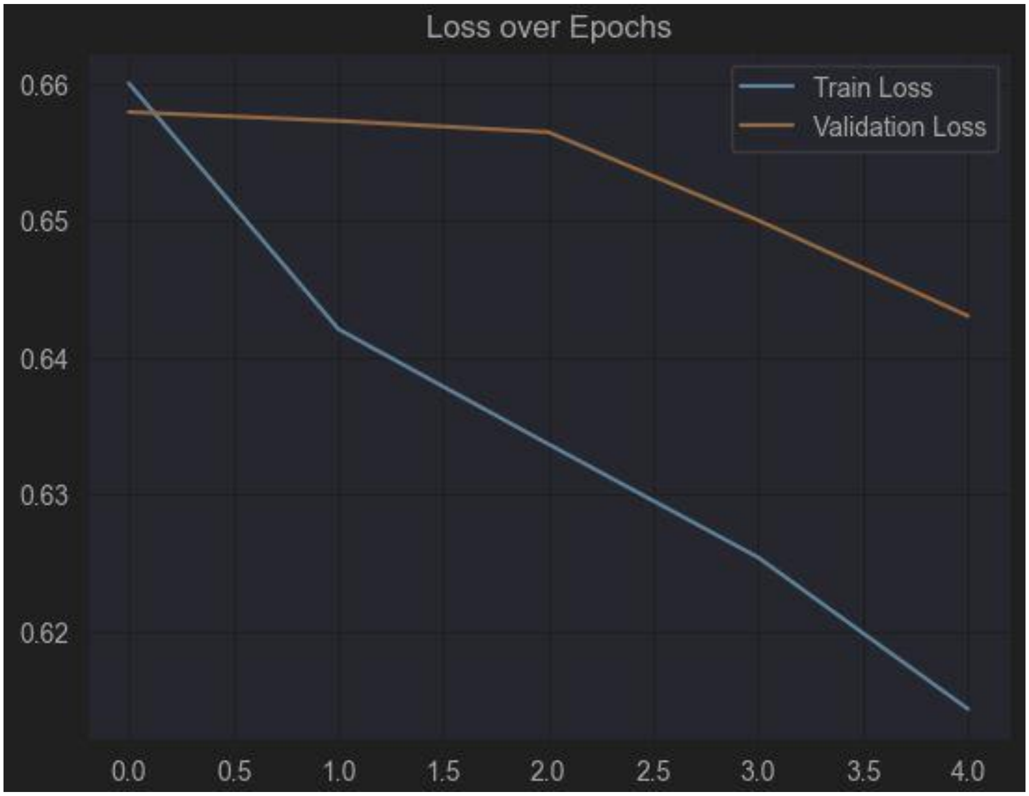
1. Class distribution plot showing the balance between normal and abnormal cases.

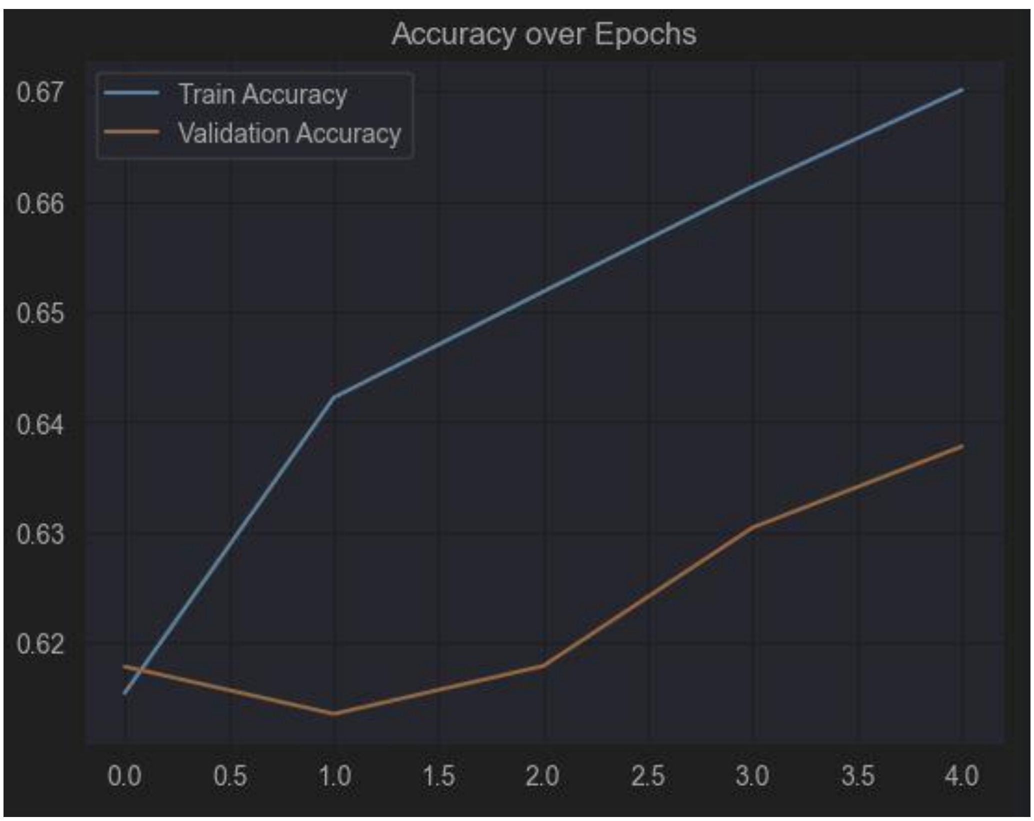


2. Sample chest X-ray images with corresponding labels.

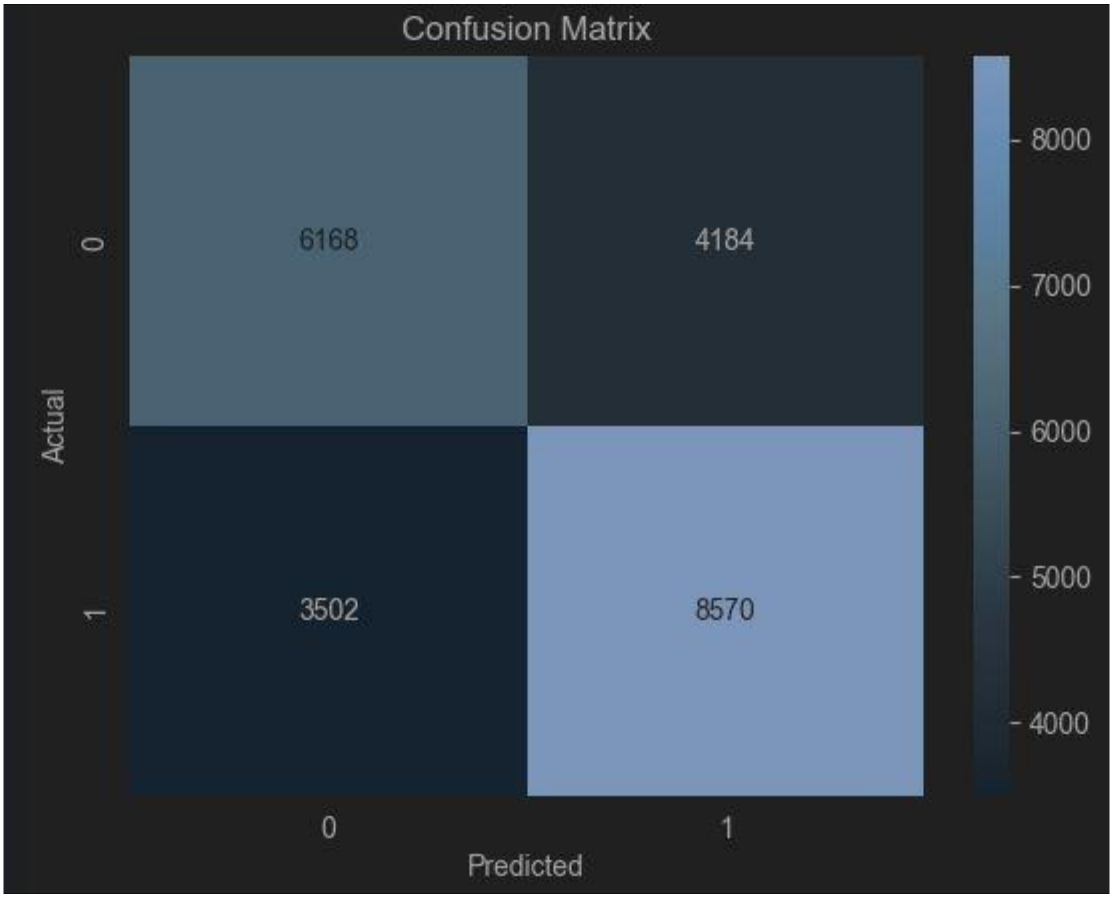


3. Training and validation accuracy/loss graphs.





4. Confusion matrix heatmap illustrating model performance.



User Guide

1. Open the provided link that will take you to a shared Google Drive.
2. Here, you will see two Google Colab files, five sample images, and the model as a .h5 file.
3. If prompted, you will need to sign in or create a Google account.
4. For demonstration purposes, open Chest_Xray_Prediction.ipynb (the source training code is included in the other Colab file titled Model_Training_Source_Code.ipynb)
5. In Colab, go to Runtime > Run All - This will run all notebook cells sequentially, all necessary libraries will be installed during this process.
6. When prompted to upload model select browse and upload the .h5 file titled cnn_xray_model.h5 from the shared Google Drive folder.
7. Once this has finished uploading, you will be prompted in the next code block to browse and upload a chest X-ray image. Here, the user can upload their image for the model or go into the folder to

upload one of the sample images included. These sample images were randomly taken from the project testing set.

8. The application will finish running the rest of the code blocks, and the processed image will appear at the end with a prediction of "Normal" or "Abnormal" along with a probability score.
    a. Note: There occasionally is a warning that appears with this execution; if run again, the warning usually disappears, but it has no impact on functionality. It is mainly caused by the file being .h5 and not the newer .keras file format. Google Colab was having trouble working with the .keras file for the model, so I reverted to the stable and more compatible .h5 format.

Reference Page

The following were my sources, along with resources from my coursework at WGU:

Fazekas, S., Budai, B. K., Stollmayer, R., Kaposi, P. N., & Bérczi, V. (2022). Artificial intelligence and neural networks in radiology: Basics that all radiology residents should know. *IMAGING*, *14*(2), 73–81. https://doi.org/10.1556/1647.2022.00104

Kromrey, M. L., Steiner, L., Schön, F., Gamain, J., Roller, C., & Malsch, C. (2024). Navigating the spectrum: Assessing the concordance of ML-based AI findings with radiology in chest X-rays in clinical settings. *Healthcare*, *12*(22), 2225. https://doi.org/10.3390/healthcare12222225

TensorFlow/Keras Documentation. (2024). *Convolutional Neural Network (CNN) guide*. Retrieved from https://www.tensorflow.org/tutorials/images/cnn

Yasaka, K., Akai, H., Kunimatsu, A., Kiryu, S., & Abe, O. (2018). Deep learning with convolutional neural network in radiology. *Japanese Journal of Radiology*, *36*(4), 257–272. https://doi.org/10.1007/s11604-018-0726-3