Problem Statement and Goals
Chest Scan

Team 16, Ace
Harrison Chiu
Hamza Issa
Ahmad Hamadi
Jared Paul
Gurnoor Bal

Table 1: Revision History

Date	Developer(s)	Change
21 Sept 2024	Harrison Chiu	Started basic info of doc
23 Sept 2024	Jared Paul	worked on Challenge Level and Appendix
23 Sept 2024	Gurnoor Bal	Worked on research focus, goals and stakeholders
23 Sept 2024	Hamza Issa	refined problem statement and worked on inputs

1 Problem Statement

This section describes the problem that this project aims to solve. It goes over the general problem, expected inputs and outputs of the project, the environment, and the stakeholders.

1.1 Problem

Chest X-rays are one of the most common diagnostic tools for detecting lung and heart conditions. However, interpreting them accurately requires expertise. Radiologists must examine each x-ray carefully. Sometimes they can miss subtle indicators of a disease. Since chest x-rays are taken for many different reasons, such discovering lung and cardiac conditions or ruling out diseases, there is an increasing volume of medical imaging data. This rising number of x-rays leads to significant delays in diagnosis, overwhelming radiologists which could cause them to make mistakes. As a result, the demand for more efficient and reliable

diagnostic support grows. A neural network designed to process chest x-rays can help radiologists detect conditions, reduce diagnostic errors, and provide faster assessments.

1.2   Inputs and Outputs

The high-level **inputs and outputs** of this problem are described as follows:
Inputs:

- Chest X-ray images sourced from publicly available datasets (i.e. ChexPert, MIMIC) for model training and testing.

- Existing academic literature on AI improvements to chest x-ray analysis and results accuracy.

Outputs:

- Tuned convolutional neural network model that attempts to accurately classify particular chest related defects from X-rays, such as cardiac or lung disease.

- Research paper discussing experiment approach, novel findings and measurable success metric such as precision, and accuracy.

- Large web application that enables interested researchers to experiment with our CNN model online.

1.3   Stakeholders

- **Radiologists**: Healthcare professionals who will use the AI system for diagnostic assistance. These professionals will need to learn how to use the system, and learn about the practicality of the research in regards to their professional needs.

- **Biomedical Researchers**: Researchers that have an interest in the intersection of AI and Healthcare, who will benefit from new research insights and potential conducting their own research in reference to ours.

- **Patients**: Individuals whose chest X-rays will be analyzed by the system. Patients will potentially receive a more hassle free service with accurate results as a result of this research.

- **AI Researchers and Engineers**: Individuals who are actively working in the AI space who can find practicality in the Chest-X ray CNN research as it applies to their particular field of interest.

- **Hospitals and Medical Centers**: Institutions that will implement this technology for faster and more accurate diagnoses. May need to conduct legal liability assessment based on the degree of accuracy of the research results.

1.4 Environment

- **Software**:Machine learning frameworks such as TensorFlow or PyTorch for model development, and natural language processing libraries for report generation.

- **Hardware**: High-performance GPUs or cloud-based systems for model training and deployment.

2 Goals

While the research aspect is the main focus, some degree of product development will still be pursued. The product component will serve as a proof of concept or MVP.

- **G1**: Conduct a comprehensive literature review to identify existing AI methods for chest X-ray interpretation. This includes gathering relevant research papers and datasets, understanding the techniques, and identifying gaps or opportunities for improvement.

- **G2**: Perform a market analysis and design study to determine which features are most valuable and practical for an AI-based chest X-ray interpretation system.

- **G3**: Train a machine learning model (either convolutional neural network or transformer-based) to automatically detect a targeted set of clinical findings from chest X-ray images. The focus will be on optimizing the training process to handle large datasets efficiently and improve model accuracy.

- **G4**: Develop a full-stack application that processes chest X-ray images, runs the model, and generates a list of findings.

- **G5**: Develop a method for converting the models findings into a structured radiology report.

- **G6**: Explore the possibility of generating free-form reports if sufficient training data is available.

3 Stretch Goals

- **Advanced Model Exploration**: Experiment with advanced techniques, such as fine-tuning transformer-based models for medical imaging.

- **Dataset Augmentation**: Investigate methods for augmenting the training dataset to improve model performance on underrepresented conditions. This may include elements of transfer learning techniques, depending on the availability of relevant datasets.

- **Free-Form Report Generation**:Extend the radiology report generation beyond structured outputs to free-form, natural language-based reporting.

4 Research Focus

Given the emphasis on the research aspect of this project, the primary research focus will be on analyzing existing work in the field of AI-based chest X-ray diagnosis. This will include:

Literature Review: Collecting and reviewing papers that discuss methods for X-ray image classification, model training on large datasets, and report generation using natural language processing.
State-of-the-Art Analysis: Identifying the most promising approaches from recent research, focusing on methods that could be applied or extended for practical use in clinical settings.
Gap Analysis: Determining areas where current research falls short or where further investigation is needed, such as specific challenges in training models or handling edge cases in medical diagnostics.

This will enable our team to conduct our own research that builds upon our recent findings, and produce a model and findings that reflect our particular research goals.

5 Challenge Level and Extras

This project falls into the **research-focused, advanced challenge level**. The complexity arises from the need to thoroughly review and synthesize state-of-the-art AI techniques in medical imaging while also building an MVP that translates research findings into a functional system. The deliverables emphasize academic inquiry, but with practical application as a secondary goal.

Extras:

- **Research Paper Synthesis**: Write a research paper on our development. Our research will build on top of existing research. The goal is to improve existing methods and innovate better ways to apply neural networks in chest x-rays.

- **Usability Testing (Optional)**: If the MVP is developed to a sufficient level, testing with radiologists may be explored.

- **Documentation**: Create both research and user-facing documentation that outlines the methodologies used, as well as any findings on model performance.

Appendix — Reflection

1. **What went well while writing this deliverable?**

   - We took the time to organize the document in a well-structured manner with a clear separation of sections. This organization made it easier for the group to assign sections and tackle each part systematically. The use of existing datasets (ChexPert, MIMIC) provided a solid foundation for defining inputs and goals. Collaboration among team members also contributed to smooth progress. Another one thing that really went well while writing this deliverable was defining the stakeholders and outlining the inputs and outputs. We found it surprisingly easy to articulate these aspects, which really helped clarify our project's scope and goals. Another positive was how well we delegated roles and tasks within our team.

   - Everyone had a clear focus, and that made us work efficiently together. This collaboration not only boosted our productivity but also ensured we all contribute meaningfully to a well-structured deliverable

   - As a group we were adamant to ensure regular communication and accountability. At the onset of this deliverable, we ensured that each group member was made responsible for a particular section, and was likewise responsible for reviewing another section completed by a peer. This process of checks and balances ensured that as a group we were tackling the deliverable to the best of our ability and in an efficient manner. This ultimately enabled our group to review everything several times and ensure we were submitting a document that best represented our current problem statement.

2. **What pain points did you experience during this deliverable, and how did you resolve them?**

   - A difficult aspect of this deliverable was really learning about the problem itself. Our team members each did indepth research into existing research publications to help us reach consensus regarding a particular problem that we hope to tackle. This was particularly difficult due to our limited experience with research articles, specifically at the intersection of AI and healthcare literature. However we ultimately managed to accomplish this by slowly working through a number of research articles, which helped improve our confidence for eventually settling on a particular topic of research in the realm of Chest-X ray CNNs.

   - Additionally, determining the goals for this deliverable were particularly challenging as our capstone project is research that is aimed at providing some innovation in the intersection of Chest x-ray and AI. As a result of this, we found it difficult to be confident in particular

goals as we were unsure what results we would ultimately find, or the potential effectiveness of using such results by various stakeholders. Fortunately, we decided on goals that we felt best represented our hopes/outcomes of this research that we are striving to achieve, rather than particular metrics to meet.

3. **How did you and your team adjust the scope of your goals to ensure they are suitable for a Capstone project (not overly ambitious but also of appropriate complexity for a senior design project)?**

We planned to explore state-of-the-art solutions and build upon existing research in the field, by reviewing several existing research papers in the area. This strategy aimed to refine our objectives and align them with the expectations of a senior design project. Stretch goals and advanced features were set aside as optional to ensure the project remains manageable within the time frame. By focusing on innovative ideas, we hoped to develop solid concepts that could ultimately be presented in a published paper.

For instance, originally our goals were quite ambitious. We had initially planned to build a full-stack web application for radiologists and other stakeholders to use with our machine learning model. This application would include an admin portal, security reviews, and a comprehensive design. However, as we delved deeper into the research and reviewed existing papers on the intersection of chest X-rays and AI, we realized that conducting innovative research while simultaneously developing a large-scale web application would be too large of an undertaking.

As a result, we refocused our goals to prioritize producing innovative research and developing a Minimum Viable Product (MVP) web application. This approach allows stakeholders to use our research findings and model themselves, rather than providing a fully-fledged, large-scale web application.