Hazard Analysis
Software Engineering

Team 17, Team RAdiAIdance
Allison Cook
Ibrahim Issa
Mohaansh Pranjal
Nathaniel Hu
Tushar Aggarwal

October 21, 2023

Table 1: Revision History

Date	**Developer(s)**	**Change**
October 16th, 2023	Allison, Tushar	Added Introduction, Scope & Purpose of HA, System Boundaries & Components, Critical Assumptions
October 18th, 2023	Allison, Ibrahim, Mohaansh, Nathaniel, Tushar	Switch to Overleaf LateX, Added FMEA Table, Safety & Security requirements, updated Critical Assumptions
October 20th, 2023	Ibrahim, Mohaansh, Tushar	Added Safety & Security Requirements, Out-of-Scope Hazards and descriptions, Roadmap

Contents

1 Introduction

This document aims to outline the possible hazards and the associated hazard controls for the "AI for Chest X-Ray" project. In the context of the project, a hazard is any condition, event, or circumstance that could jeopardize the safety, reliability, or effectiveness of the system for diagnosing chest X-ray images. Hazard analysis is an essential step in the project's development lifecycle, focusing on the identification, assessment, and mitigation of potential hazards & risks and safeguarding the quality & reliability of the proposed solution.

1.0.1 List of Acronyms

Table 2: List of Acronyms

Acronym	Description
AI	Artificial Intelligence
FMEA	Failure Modes and Effects Analysis
HA	Hazard Analysis
IT	Information Systems
ROC	Receiver Operating Characteristic

2 Scope and Purpose of Hazard Analysis

The purpose of the hazard analysis is to identify the possible hazards associated with the system, the potential effects, and how to reduce the risk within the system through preventative design and actions. In our project, hazards can be linked to patient information being collected incorrectly, data leaks in patient records, issues with authorization, false positives or false negatives outputs in disease detection, and more as outlined below.

The scope of hazard analysis in the project focuses on ensuring the safety, effectiveness, and reliability of the AI-based diagnostic system for chest X-ray images. It extends to the entire ecosystem of the AI system, including the machine learning algorithms, user interface, data management, and the interaction between healthcare professionals, patients, and the AI system.

3 System Boundaries and Components

The system will involve the following components/functions:

- **Frontend**
  - A web interface for user interaction and user authentication
  - Image upload, retrieval and access

  - Display diagnostic notes & findings

- **Backend**
  - An AI Model for identifying the anomalies related to the selected diseases
  - Databases that will hold X-ray images and store patients' records
  - The host support system (i.e., Medical Institution/Diagnostic Centre IT Systems)

The system boundary includes the application, AI model, database, and physical host system. The host system and load times of the database are not controlled by the system. The host system is managed by the institution, hospital, or office. However, all components are to be considered in the hazard analysis as they can impact the system.

4 Critical Assumptions

The following are the critical assumptions for the project:

1. The hospital/work location will have a secure database where the generated diagnostic notes document will be stored after generation, a copy of the document will not be stored within the system, only the final diagnoses of the AI.
2. Data stored by the AI model to continue training during operation will have minimum identifying demographic information of the patient, to limit unnecessary data stored within the system.
3. The chest X-ray images provided will be of sufficient quality for the AI system to accurately detect abnormalities.
4. The system is assumed to be available for use at all times, with minimal downtime.
5. The system will be accessed by trained healthcare professionals and will have a certain level of competency in interpreting the diagnostic findings.

5 Failure Mode and Effect Analysis

The following is a detailed breakdown of the possible failure modes and effects analysis, or FMEA table. For each component, the possible failure modes, effects, causes, detection, controls, risk, recommended action, associated safety/security requirements and references are described.

Table 3: FMEA Worksheet

Component	Failure Modes	Effects of Failure	Causes of Failure	Detection	Controls	Risk	Recommended Action	Req.	Ref.
Web Interface - User Access Authentication	Fails to authenticate user	medical professional unable to login	Authentication error in web application	Manual testing		Low	Include alternative methods to authenticate user	PR0	H1.1
		unauthorized third party able to login					Include safeguards to prevent unauthorized parties from logging in	AR0, AR1, PR0, PR1	H1.2
Web Interface - Image Upload	Fails to upload chest x-ray image	user cannot upload chest x-ray image to web app	image upload error in web application	Manual testing		Low	Include alternative methods to upload chest x-ray images		H2
Web Interface - Display Diagnostic Findings	Fails to show diagnostic findings	user cannot view diagnostic findings of chest x-ray analysis	data access error in web application	Manual testing		Low	Include alternate methods for users to view diagnostic findings	AR1	H3.1
	Shows diagnostic findings by mistake	unauthorized user can view diagnostic findings of chest x-ray analysis					Include alternative methods to authenticate user	AR1	H3.2
Detect disease in chest x-ray image	Generates false positive	Healthy patient could be diagnosed, resulting in unnecessary treatment	Model mistakenly detects disease absent in a normal chest x-ray	Manual testing		High	Optimize chest x-ray analysis AI to minimize false positives	SR0, SR1	H4.1
	Generates false negative	Diseased patient undiagnosed, could escalate symptoms	Model fails to detect disease in an x-ray with the diseases				Optimize chest x-ray analysis AI to minimize false negatives	SR0, SR1	H4.2
AI Algorithm Training	Model Overfitting During Training	Model performs well on training data but poorly on new, unseen data, leading to inaccurate diagnoses	Overfitting due to complex model architecture	Monitoring validation data during training	Implement dropout and regularization techniques, fine-tune hyperparameters	High	Implement techniques to detect and prevent overfitting	SR0, SR1	H5
Data Storage	Data loss	Loss of patient data	Database server malfunction	Regular data backups	Database redundancy	Data loss risk	Implement robust data backup	IR0, SR2	H6.1
	Data corruption	Loss of patient images and records	Database server corruption	Data integrity checks	Regular data backups	Data loss leak	Implement data integrity checks		H6.2
Data Access	Data transfer failure	Inability to retrieve patient data	Network communication issues	Automated data transfer checks	Redundant data transfer paths	Data retrieval risk	Implement data transfer redundancy		H7
Data Security	Cyberattacks	Unauthorized access to patient data	Weak security measures	Intrusion detection	Enhanced cybersecurity measures	Data breach risk	Enhance cybersecurity	AR0, IR0, PR0, PR1, SR2	H8.1
	Unauthorized access	Data breach and patient privacy	Weak access control measures		Enhanced access controls		Enhanced access control		H8.2
Backend Server	Network failure	Disruption of connection to database	Network connectivity issues	Real-time monitoring	Redundant network connections	Operational disruption	Implement network redundancy		H9.1
	Server downtime	Unable to access to patient data	Server hardware failure		Redundant server systems		Implement server redundancy		H9.2

5.1 Hazards Out of Scope

The following hazards are considered to be out of scope for this project's proposed solution:

- **Compromised Host System:** If the host system used to run the application is compromised. Authorized access to the web application helps mitigate this hazard but most of the risk is beyond our control.
- **Malicious Cyberattacks:** The project acknowledges the importance of cybersecurity and implements security measures to protect user data. However, specific hazards related to highly sophisticated and malicious cyberattacks are considered out of scope due to their unpredictable and evolving nature.
- **Power Outages:** The project relies on access to power and certain technologies to fully function. Power outages are external events that are beyond the scope of the system.

6 Safety and Security Requirements

6.1 Access Requirements

AR0 The x-ray images should only be accessible to authorized users. Authorized users include doctors and IT staff responsible for storing medical data for the medical institution.

Rationale: This is to keep health records confidential and accessible to only those who have permission to view them, such as a doctor or nurse.

Fit Criterion: Users with an authorized username and password will be able to access the X-ray image.

AR1 The system should allow access to generated reports only by medical professionals and the patient to whom the report belongs.

Rationale: Similarly to NF-AR0 this is to keep health records confidential and accessible to only those who have permission to view them, such as a doctor or nurse.

Fit Criterion: Users with an authorized username and password will be able to access the generated report.

6.2 Integrity Requirements

IR0 The system will encrypt all stored data

Rationale: This is to help ensure that if the system is attacked any data is not easily collected.

Fit Criterion: All data is not stored in plain language.

6.3 Privacy Requirements

PR0 Only authorized users, doctors, will have access to patients information.

Rationale: Similarly to Access requirements this is to keep records confidential and accessible to only those that have permission to view.

Fit Criterion:

PR1 The system should maintain the privacy of the patient's personal and medical information.

Rationale: This is to follow the medical practices of keeping patient privacy.

Fit Criterion: Patient information is not openly accessible.

6.4 Safety Requirements

SR0 Accuracy of the algorithm

Description: The system will show accurate findings based on the area under the ROC curve threshold for all the diseases identified.

Rationale: False negatives need to be eliminated and inaccurate findings can result in poor health of the patient.

Fit Criterion: The area under the ROC curve for each disease after testing the model is greater than the recommended threshold for getting accurate results.

SR1 Algorithm Testing and Validation

Description: Continuous testing and validation of the machine learning algorithm to ensure it meets safety and accuracy standards.

Rationale: Testing is an efficient way to discover faults and improve performance.

Fit Criterion: The system passes industry-standard tests for the algorithm's accuracy.

SR2 Data Encryption

Description: Ensure data encryption during data transfers to prevent unauthorized access.

Rationale: Encryption is important for the security and privacy of data, and hence the security of patients.

Fit Criterion: The system uses a secure encryption algorithm like SHA-2 or better.

SR3 User Authentication and Access Control

Description: Ensure that only authorized users, such as healthcare professionals, can access the system and patient data.

Rationale: Unauthorized use could expose confidential personal information and threaten the safety of patients.

Fit Criterion: Users cannot access the system without logging in with their credentials.

7 Roadmap

After careful consideration and reassessment, the team realized there were many new requirements for us to take into consideration that were not initially apparent when writing the Software Requirement Specification. Ideally, we will aim to implement every safety requirement, but realistically when taking into account time and resource constraints, those requirements that are strictly necessary for system functionality may be the only ones that get implemented.

To be implemented during the capstone timeline:

- SR0, SR1, SR2, SR3

To be implemented in the future:

- **Algorithm Optimization:** Continuous efforts to enhance the efficiency & accuracy of chest X-ray analysis through algorithm optimization.
- **Audit Log Maintenance:** Maintain an audit log of all application activities. Incorporate an activity logger within the application framework.
- **Security Audits and Vulnerability Assessments:** Regular audits and assessments to maintain a secure environment and protect patient data.
- **Security Patches and Updates:** Regularly roll out security patches and updates to fix known vulnerabilities. Form a dedicated security updates team to monitor, identify, and rectify vulnerabilities.