Hazard Analysis
Natural Language Processing for Mental Health Risk Prediction

Team 13, The Cognitive Care Crew
Jessica Dawson
Michael Breau
Matthew Curtis
Benjamin Chinnery
Yaruo Tian

Table 1: Revision History

Date	**Developer(s)**	**Change**
October 19, 2023	Matthew Curtis	Adding section 4 and SR 3, 4, and 5
October 19, 2023	Michael Breau	Adding section 1 and parts of FMEA
October 19, 2023	Jessica Dawson	Adding section 2 and parts of FMEA
October 20, 2023	Yaruo Tian	Adding section 3 and SR 1 and 2
October 20, 2023	Benjamin Chinnery	Adding roadmap and SR 6, 7, and 8
January 30, 2024	Yaruo Tian	Modifying and adding to section 3

Contents

1 Introduction

This document is used for the analysis of possible hazards that could affect the NLP Early Risk Detection of Mental Health Issues project. The definition of a hazard as used in this document is any potential cause for a loss to the interests of the stakeholders of the project. The purpose of this document is to aid with resolving potential risks for hazards before they become an issue.

2 Scope and Purpose of Hazard Analysis

The purpose of this hazard analysis is to identify potential risks to stakeholders and the failure modes of the software that lead to these risks being realized. The document only covers stakeholder losses caused by a failure of our software system as the focus is on guiding the design of the system, not defining all possible stakeholder losses. The document identifies potential hazards in the system and from these derives a set of security requirements that will be used to refine our system design.

3 System Boundaries and Components

The system that the hazard analysis with be conducted on consists of:

1. The system's security/ethical aspects which consist of the following:
    - Data ingestion component
    - Data processing component
    - Algorithm bias feature
    - Data protection component
    - Accuracy of model prediction component

    The major concern of this system is the security of user data provided by E-Risk. It must be known that data leaks from the system would be a potential hazard to these users. Data ingestion, processing, protection components of the system must be analyzed.

4 Critical Assumptions

- We will assume that any health care provider using our product would not intentionally misuse it.
- We will assume that any health care provider using our product would not use our product with malicious intent and only use it to help their patient.
- We will assume that the libraries and functions we utilize in our code work as expected such as NumPy, Pandas, and SkLearn.

5 Failure Mode and Effect Analysis

Table 2: Failure Mode and Effect Analysis Table

Component	**Failure Modes**	**Effects of Failure**	**Causes of Failure**	**Recommended Action**	**SR**	**Ref.**
General	Compromised sensitive data	Legal consequences and loss of trust	System security breach or compromised device from team member	The team should be careful to not expose sensitive data anywhere that could lead to it being seen by non-team members	SR1, SR2	H1-1
	Hardware Limitation (Crashes or not enough processing power)	Loss of trust from users, in the case of this project specifically: potential poor results at eRisk competition	Process causes device to crash, processes errors out unexpectedly, or the code is unable to complete in a time constraint that makes sense	The system should be thoroughly tested as well as account for edge cases that could possibly lead to one of the previously mentioned failures	SR3	H1-2
	An nlp pipeline breaks on deployment	Loss of time to try different models and approaches during the eRisk competition	Lack of comprehensive validation systems before deployment	Implementation of pipeline testing systems to ensure issues are caught early	SR4	H1-3

	Loss of code	Loss of time to try different models and approaches during the eRisk competition	Lost code was not backed-up	Robust backup systems implemented into development processes	SR5	H1-4
Risk Prediction	An incorrect diagnosis could result in harm to the health care professional giving the diagnosis or the person being diagnosed	Legal consequences and loss of trust	Upload of incorrect code that produces false results	The team through tools as well as themselves hold themselves accountable to their code as the wellbeing of others can be affected by it	SR4, SR8	H2-1
	Corrupted data causes incorrect model results	Model gives a wrong prediction, impacts success in the eRisk competition and could result in an incorrect diagnosis	Data is distorted in some way (entries duplicated, deleted, changed, etc.)	System will check data to ensure it is intact before it is given to the model	SR6, SR7	H2-2

6 Safety and Security Requirements

SR1. User data must not be shared or re-used in any system not part of this system

Rationale: As users have an expectation that their personal data will be handled with care, sharing it with other systems will not guarantee their safety as it will not be under our control.

Associated Hazards: H1-1

SR2. Sensitive user data is must not be present within the results generated

Rationale: This is to ensure legal compliance and uphold ethical and professional standards. In addition, exposing people's PPI could lead to unauthorized access, data breaches, or privacy violations.

Associated Hazards: H1-1

SR3. The system will be tested periodically to detect crashes and a potential lack of processing power

Rationale: As we build our system out we will be adding more features and increasing the complexity, which increases the processing power required. This may lead to crashes or issues with a lack of processing power that will need to periodically be checked. This is to check the hardware limitations.

Associated Hazards: H1-2

SR4. The system will run through a series of tests before deploying/pushing new builds

Rationale: This will be done to ensure that the system is still working as expected with our new changes and that nothing has broken before we deploy a new build. These tests will help us detect that and allow us to go back and fix any issues that arise.

Associated Hazards: H1-3, H2-1

SR5. The team will regularly push their code by committing to a repository while working on the code

Rationale: This needs to be done in order to avoid losing code/work. If something were to happen to the team members computer for whatever reason and the work that was done after the last commit was to be deleted/lost, frequent commits would help ensure minimal lose.

Associated Hazards: H1-4

SR6. The system will operate using cleaned data that does not contain duplicates

Rationale: It is important for the system to operate off clean and effective data in order to mitigate the chances of incorrect predictions.

Associated Hazards: H2-2

SR7. The system data will only operate off of verified data free from copying errors

Rationale: It is important for the system to work off of internally approved datasets that have come from the correct sources and do not contain data transfer errors in order to help ensure a better output.

Associated Hazards: H2-2

SR8. The system will be tested against plainly appearent data to guide and ensure prediction functionality

Rationale: The system should be able to output correct diagnosis for less nuanced dataset tests in order to help ensure consistent predication capabilites between system updates.

Associated Hazards: H2-1

7 Roadmap

This hazard analysis has been able to identify various threats to the safety and security of this project, that will need to be accounted for in order to help keep project progress on track and reach the desired milestones. The hope for the team is to be mindful of these hazards and gradually implement protections over the course of the development project with the hopes of meeting all of these requirements in the Revision 1 implementation. Due to the nature of the project being built to the specifications of the eRisk competition, the team must be mindful of our limitations when it comes to implementation, and we recognize that aspects of the project and any corresponding requirements will be out of our hands and may not be met. The requirements that we felt were most vital and achievable in our timeline were the safety and privacy concerns from SR1 and SR2, as that falls clearly under our personal responsibilities and practises.