

Problem Statement and Goals

Team 13
Software Architecture for Natural Language
Processing/AI applied to Mental Health

Michael Breau,
Matthew Curtis,
Benjamin Chinnery,
Jessica Dawson,
Yaruo Tian

September 24, 2023

Table 1: Revision History

Date	Developer(s)	Change
September 24, 2023	Benjamin Chinnery, Yaruo Tian	Initial Release

1 Problem Statement

1.1 Problem

As society grows to become more conscious of the ever-growing issues surrounding mental health, the demand for support services only seems to be increasing, especially in a post pandemic world, where more and more people are no longer satisfied with just having a higher visibility on their own mental health issues, they wish to actively work on themselves using support services. Unfortunately, many Canadians find themselves in remote communities where they are unable to receive the support they require. In response a joint effort from Dr. Mosser's past teams, along with support from UQAM and several psychologists from around the world have developed methods and approaches to assist psychiatrists with diagnosing mental health issues.

In this capstone project, the team will be designing and validating a Natural Language Processing architecture based off the works and methods described prior. This NLP would be capable of recognizing and diagnosing behaviour

which displays signs of pathological gambling. Our team’s goal is to successfully replicate and hopefully improve upon the results of the original studies, to eventually create an NLP architecture that would be capable of assisting mental health professionals in their field.

1.2 Inputs and Outputs

At a high level, the inputs for the system’s initialization would be the data provided from the Early Risk Prediction on the Internet (eRisk) initiative, this data would be used alongside the prior studies done by the team to train the Natural Language Processor. Once the system has been trained, it would require text communications of people speaking to mental health services, to analyze these texts and produce an output.

This output would be using its natural language processing abilities to assist psychiatrists by pointing out behaviour of people who display warning signs of Mental Health issues such as gambling addiction.

1.3 Stakeholders

A major external stakeholder for this project would be the CLEF eRisk Organizers. They are responsible for providing project guidelines, datasets and performs evaluation of the project. Meeting their specifications and requirements is essential for the project’s success. Secondly, the dataset involves potential ”pathological gamblers” which are the researched communities. These individuals are also considered the external stakeholders.

Few important internal stakeholders to note are ourselves, the student developers of the project, Dr. Sebastien Mosser, researchers of UQAM (University of Quebec in Montreal), and few psychologists from around the world. These individuals have a direct interest in the project’s operation and success.

1.4 Environment

1.4.1 Hardware

- The product will be supported on desktops and laptops with connections to the servers that CLEF eRisk 2023 provides to us.

1.4.2 Software

- The product will be supported on Windows and macOS.

2 Goals

2.1 Ease of Use

In terms of operations, the product should be intuitive and require no further instructions from the external stakeholders such as CLEF eRisk Organizers.

2.2 Natural Language Processing

The product should use Natural Language Processing techniques to accurately determine which users have committed pathological gambling.

2.3 Training on eRisk dataset

The product should be able to process and analyze the test dataset and server data of pathological gamblers and non-pathological gamblers provided by eRisk.

2.4 Safety of Data

The product should never leak personal data of any user being analyzed.

3 Stretch Goals

3.1 Detection accuracy

The detected pathological gamblers of the product matches with the actual pathological gamblers 100% of the time.

3.2 Detect other disorders

In addition of pathological gambling, the product could also perform other tasks provided by CLEF eRisk (TBA).

3.3 Support for public usage

The product could have an user interface that allows general public could enter their own inputs and able to generate an accurate result.