

Reflection Report on Natural Language Processing for Mental Health Risk Prediction

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

[Reflection is an important component of getting the full benefits from a learning experience. Besides the intrinsic benefits of reflection, this document will be used to help the TAs grade how well your team responded to feedback. In addition, several CEAB (Canadian Engineering Accreditation Board) Learning Outcomes (LOs) will be assessed based on your reflections. —TPLT]

1 Changes in Response to Feedback

[Summarize the changes made over the course of the project in response to feedback from TAs, the instructor, teammates, other teams, the project supervisor (if present), and from user testers. —TPLT]

[For those teams with an external supervisor, please highlight how the feedback from the supervisor shaped your project. In particular, you should highlight the supervisor's response to your Rev 0 demonstration to them. —TPLT]

1.1 SRS and Hazard Analysis

Our SRS has undergone many changes since the revision 0 submission. Some were from feedback from the TA such as reformatting document structure to be more clear, creating more readable functional requirements with labelling, adding a reflection section which we missed for rev 0, etc... We also made many changes directly because of our expanding of scope and knowledge of natural language processing pipelines increasing so we had a better idea of the functional requirements needed as well as our division into the three different NLP tasks.

1.2 Design and Design Documentation

1.3 VnV Plan and Report

2 Design Iteration (LO11)

[Explain how you arrived at your final design and implementation. How did the design evolve from the first version to the final version? —TPLT]

2.1 Task 1

2.2 Task 2

For our second natural language processing task related to the binary classification of anorexia our initial model was a simple iteration of grouping documents together and we decided that for the "training" the group with the most amount of positive users would be considered to be the group predicted to have anorexia. The initial model was not able to predict new users which was added later on. Another issue we ran into was that Bertopic has many different parameters but with a long runtime we were not able to test many of them but from the recommendation of one of our supervisors we used grid hyper parameter tuning which allowed us to determine the best parameters for many different fields to be tailored to the training data.

2.3 Task 3

3 Design Decisions (LO12)

[Reflect and justify your design decisions. How did limitations, assumptions, and constraints influence your decisions? —TPLT]

3.1 Task 1

3.2 Task 2

The primary limitations and constraints influencing the design of this task stemmed from the rules and guidelines of the competition. It was important for the team to try to find the best balance between achieving the highest accuracy possible, but also keeping a good handle on length of operation, resource allocation, and of course repeatability in order to optimize our performance scores. Although the team experimented with various models which boasted all sorts of situational expertise's, the team felt that the BERTopic model was able to strike the best balance between our allocated priorities. A major component of this task is that the model is meant to take information in sequential chunks in order to imitate real time posting behaviours, what made BERTopic stand out was its ability to update decisions on its users on the fly, without having

to retrain its database every time, cutting down on operation time as well as resource usage.

3.3 Task 3

4 Economic Considerations (LO23)

The target market for an Natural Language Processing system capable of detecting depression, anorexia and eating disorders is large, namely due to the growing acknowledgement and importance of the mental health problems. To market this product, targeting healthcare professionals, mental health organizations, clinics staff and/or individuals or relatives who seek support would be carried out. The cost factors of a sell-able version would include development, testing and on-going maintenance, which could fall into the moderate to more significant category depending on the complexity and scale of the system. The pricing, most likely would depend on the target market and the value promised. If we plan to further maintain and develop this product into something that could actually be used by real life users, then we could charge a monthly or yearly subscription fee per user registered.

To make a profit of units sold should be larger than the total production and marketing costs. If the project uses the open-source model and would not to be sold to the public, attracting users along the help of community engagement, collaborations with mental health experts, and making the system available for researchers and professionals will be possible. This audience is potentially the most big and diversified because it comprises healthcare providers, people with mental health disorders, researchers, and educators.

5 Reflection on Project Management (LO24)

[This question focuses on processes and tools used for project management. —TPLT]

5.1 How Does Your Project Management Compare to Your Development Plan

After completing the project, our team now understands how lacking in scope we were at the start of the project. Originally we were just planning on creating this one pre diction model but we ended up expanding our project to cover three different natural language processing tasks. In turn a lot of our development plan had to be updated to reflect these changes. For example our team roles had to be changed to suit the expansion of three tasks so we had to all learn the entire natural language processing workflow instead of having members specialize in single skill sets. Overall our team did follow the communication plan well which did not need any changes. The technology we planned on using did also expand quite a bit since we learned of new methods over the course of the project

and were constantly trying many different technologies which would have been impossible to predict at the start of the project.

5.2 What Went Well?

[What went well for your project management in terms of processes and technology? —TPLT]

5.3 What Went Wrong?

[What went wrong in terms of processes and technology? —TPLT]

5.4 What Would you Do Differently Next Time?

If given the opportunity to work on a similar project which would require the coordination of 2 independent groups, we feel that the best course of action would be to be more strict with our merging timelines in order to foster better results for everyone. Earlier stages of the project's development lost valuable time when there weren't clear boundaries and roles of how the teams would interact. Both teams were so focused on managing their immediate responsibilities that less emphasis was put on the bigger picture, so that when the time came for the systems to start integration both teams realized that valuable time and effort was misplaced either doing overlapped responsibilities, or creating systems that would require some reworking in order to work within the other team's design constraints. Overall, better communication would lead to stronger team synergy, and less misplaced effort that could be better directed at making a more efficient project.