6COSC023W – Final Project

Project Proposal (PP)

Person's Behaviour Analysis with text messages via NLP - BAWT

Student: Mohammad Sabiq Sabry (19219495)

Supervisor: Anastasia Angelopoulou

Degree: BSc (Hons) in Computer Science

School of Computer Science & Engineering
College of Design, Creative and Digital Industries
University of Westminster

Date: 10/11/2022

Contents…

1. Problem statement, project aims and objectives

Problem Statement

Recent years have seen a significant increase in the use of texts for conversation. Through WhatsApp, a variety of statistics have been shared. With more than 2 billion active users (Ruby, 2022), WhatsApp is the most used chat application. Everyone has utilised it extensively, especially certain business people and young people. Customers can examine the personal or business WhatsApp communication using a variety of analysing tools.

Genuine users want to check the tone of conversation for a variety of reasons. There are numerous use cases for this assignment, including the desire of the discern to examine their childs' chats and for people who have the disease Alexithymia, the inability to recognize and express emotions.(Cherney, 2021)

Most people tend to be more honest when communicating via text messages with everyone, because they do not have to deal with the awkward stage of having to present themselves in a real life situation. So it is vital to have a system where anyone with Alexithymia can desert one's behavioural pattern.(Lister-Landman, Domoff and Dubow, 2017)

Project Aim

Alexithymic individuals struggle to understand and express their own feelings, as well as to acknowledge and react to the emotions of others so to have a system which will show them how a person behaves will make it easier for one's life. This will also be helpful for the police to have an in-depth understanding of a person's behavioural pattern because in today's world there are so many crimes happening all over the world just via texting.

Project Objectives

There are 2 objectives I hope to achieve in this paper; one a full on analysis report of a person's behavioural pattern using Sentiment Analysis. Research shows many people use emojis/gifs to relate to their feelings but sometimes we as people misread it into something else so Sentiment Analysis will portray the positivity level of a person. This will help the people with Alexithymia to understand a person thoroughly without causing distress to one another and it will also help the policia manage the use of improper conduct via text messages.

And two, to propose a system via NLP that, by implementing the base-level implementation, prevents the user from sending unsuitable or improper messages to the participants. Many people in today's world make mistakes when messaging someone important; I hope to present a system where such mistakes can be avoided just by checking the tone of the message.

2. Background review: similar software/applications, or similar research

As of right now there are no other applications/methodologies that have a system to recognize behaviour in any sort via texting or even a system to detect obscene or vulgar content in a user's typed message and prevent the user from submitting it, but according to research there are analysis algorithms that can analyse a person's behavioural pattern by NLP.

There are few research gaps which calculate the average time a person spends on WhatsApp texting to the relevant person and it also calculates the average amount a user uses certain words. (MaartenGr, 2020) The below images are a representation of it:

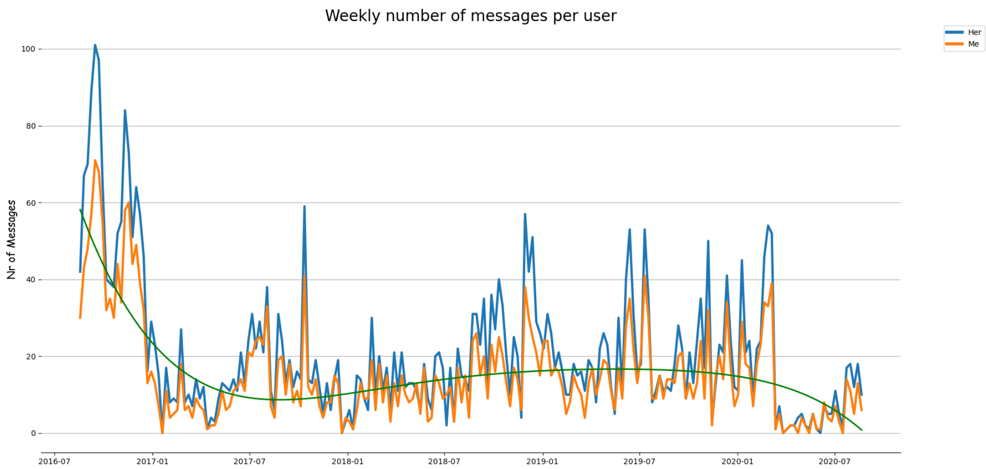

Fig1: Weekly Number of Messages (MaartenGr, 2020)

Fig2: Semantic Analysis of the most used words (MaartenGr, 2020)

There is also an application called My Compass App(Assistive Technology at Easter Seals Crossroads, 2018) that helps people with Alexithymia to sort out their emotions individually by categorising their energy levels, but it does not prove to help the people affected with the cause to understand another person via texting or any other methods of communication; this application is solely for the purpose of showing an Alexithymic individual on what they are feeling. Below is a sample image taken from the application:

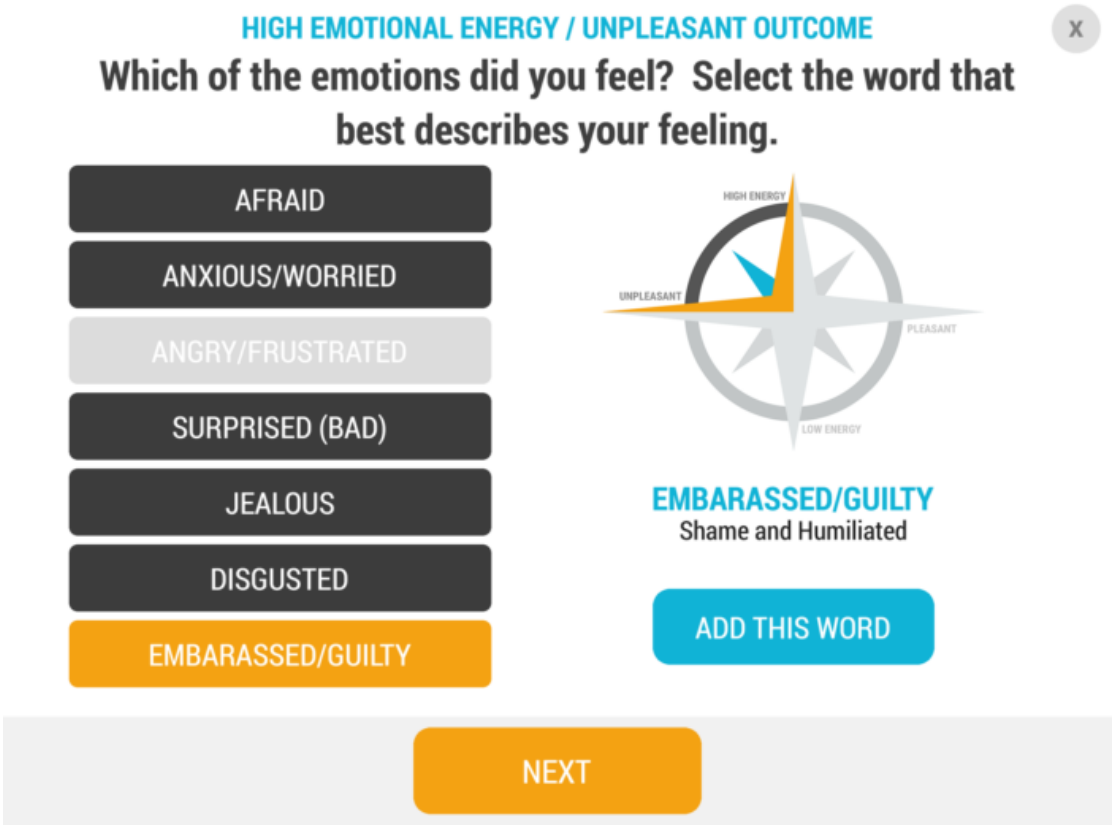

Fig3: Snapshot of My Compass App (Assistive Technology at Easter Seals Crossroads, 2018)

	Research Gap	My Compass App	BAWT
Emotion Based Analysis	No	Yes	Yes
Applies when texting (for eg: via WhatsApp)	Yes	No	Yes

Table1: Comparison Table

3. Tools and skills

The following tools and skills are required to complete the project research gap basing off the objectives and expected outcomes.

Software Requirements:

1. Operating System (Linux/macOS/Windows): In this case, macOS will be used because I am familiar with the OS but Linux and Windows will be used for research purposes.
2. Python: the language that will be employed to build the deep learning and machine learning models. Many projects that combine data science have used the general-purpose language Python.
3. Tensorflow/Pandas/Numpy/MatPlotLib/SkitLearn and such: Libraries that will be used to enable the development, testing, and training of models.
4. VSCode: the environment that will be used to support the project development.
5. Google Colab/Jupyter: Cloud development environment that will help in building/training/testing ML and Deep Learning/NLP models.
6. Google Docs/Canva/Instagantt/Draw.io: Tools that will be utilised to create figures, architectures, and documentations.
7. Google Drive/Github: To backup and save all the files including the code of the project.

Hardware Requirements:

1. Intel i7 Processor/M1 (or above): this is to perform high intensive resource tasks.
2. 8gb RAM and above: To manage datasets & development environments.
3. Disk Space of 40GB or more: To store all the data and research gap code.

Data Requirements:

1. User Preference Profile Data: Datasets of personal/group chats from Kaggle, Amazon, Kelp will be used for testing/training purposes.

Skill Requirements:

1. The ability to create optimised Machine Learning/ Deep Learning/ Natural Language Processing models.
2. Research Writing Skills

Risk Mitigation Table

Risk Item	Severity	Frequency	Mitigation Plan
Losing all the code to all the development of the research gap	5	2	Keeping all the code all backed up on Github/Google Drive and on an external storage device.
Corruption of documents	4	4	Use a cloud based document approach and have a backup of it every week so that progress would not be lost.
Inability to explain the research gap worked on due to a sudden illness	4	2	Have a screen recorded demonstration along with a detailed documentation explanation of the work done.
Inability to complete all the deliverables within the given time	2	1	Work on the deliverables in a priority based way so that everything can be delivered on time without any delay.

Table2: Risk Mitigation Table

4. Initial list of requirements

Research Methodology:

Qualitative methods analyse contextualised, unstructured data. These time and money-consuming techniques frequently produce tiny sample numbers and difficult-to-replicate results. Natural language processing (NLP) integration can boost productivity by reducing time and expenses, expanding sample sizes, and enabling replication for validation. NLP techniques can help qualitative research advance.The same interpretation and identification can be used to assess new data that is added throughout time.

Development Methodology:

The Agile Development Life Cycle (BMC Blogs, 2020) was chosen as the research development method because iterative development is needed and it should be flexible and adaptable enough to gain a feedback based approach.

5. Challenges

This research gap, BAWT, is a rare domain so there is not enough research done on the topic of tracking behaviour of a person using Natural Language Processing. Hence why it's important to identify the features and external factors before an Alexithymic individual starts using the developed model/application.

"People who do have alexithymia may describe themselves as having difficulties with expressing emotions that are deemed socially appropriate, such as happiness on a joyous occasion. Others may furthermore have trouble identifying their emotions. Such individuals don't necessarily have apathy. They instead may not have as strong of emotions as their peers, and may have difficulties feeling empathy." - Kristeen Cherly (Cherney, 2021)

It's going to be a challenge to gather enough information/datasets on chats to make sure we can train the model enough before giving it a certain public image. Because generally people do not like to share such intimate text messages with the public so it's gonna take some time into training the model to work according to a certain individual's behaviour.

It is also going to be a challenge to work on new technologies like NLP because that is a new domain in today's world and to tackle it on the project is going to take some time but I deem it to be possible with the amount of time given. Using the identified factors to be considered, a suitable architecture needs to be implemented to forego the research gap.

6. Project timeline

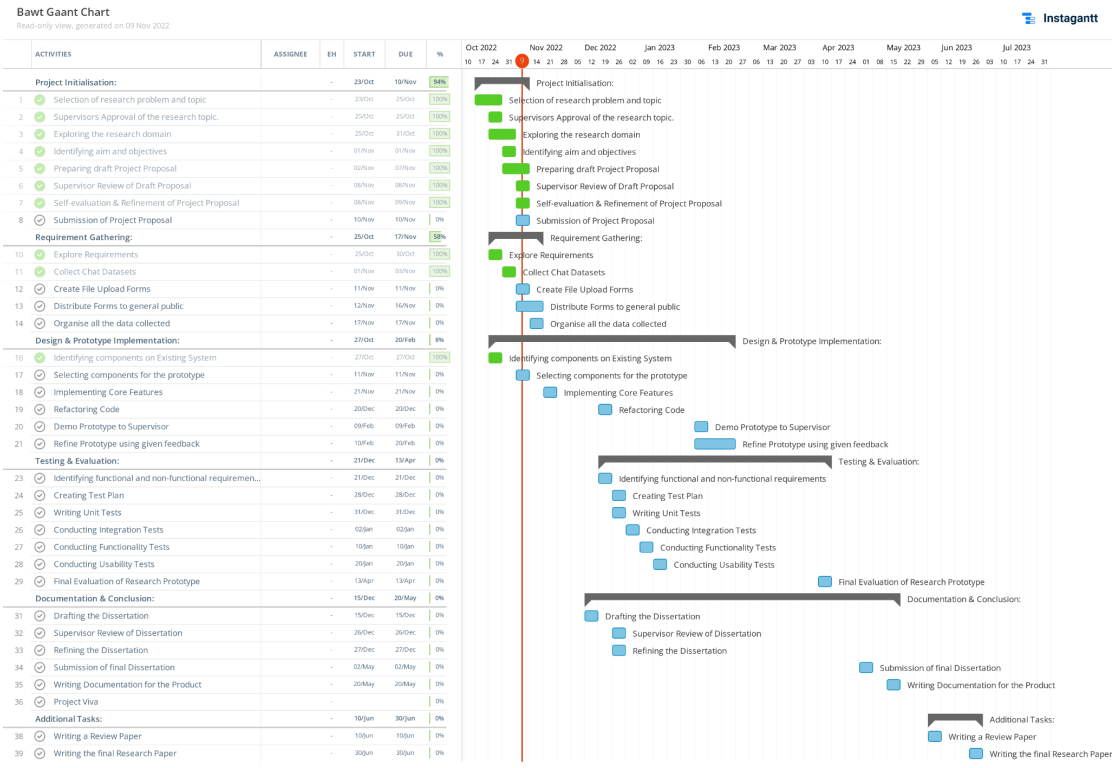

Fig4: BAWT Gantt Chart (Bawt Gantt Chart.pdf, 2022)

Deliverables:

Deliverable	Date
Project Proposal	November 10, 2022
Ethics Form	November 10, 2022
Project Specifications Design and Prototype (PSDP)	February 9, 2023
Project Progress Update	March 1, 2023
Final Year Project Submission	May 2, 2023

Table3: Deliverables

7. References / bibliography

BMC Blogs. (2020). *Agile vs Waterfall SDLCs: What's The Difference?* [online] Available at: https://www.bmc.com/blogs/agile-vs-waterfall/ [Accessed 9 Nov. 2022].

Abram, M.D. (2020). *Methods to Integrate Natural Language Processing Into Qualitative Research - Marissa D. Abram, Karen T. Mancini, R. David Parker, 2020*. [online] International Journal of Qualitative Methods. Available at: https://journals.sagepub.com/doi/10.1177/1609406920984608 [Accessed 9 Nov. 2022].

Assistive Technology at Easter Seals Crossroads. (2018). *My Emotional Compass App*. [online] Available at: https://www.eastersealstech.com/2018/09/25/my-emotional-compass-app/ [Accessed 9 Nov. 2022].

MaartenGr (2020). *GitHub - MaartenGr/soan: Social Analysis based on Whatsapp data*. [online] GitHub. Available at: https://github.com/MaartenGr/soan [Accessed 9 Nov. 2022].

Lister-Landman, K.M., Domoff, S.E. and Dubow, E.F. (2017). The role of compulsive texting in adolescents' academic functioning. *Psychology of Popular Media Culture*, 6(4), pp.311–325. doi:10.1037/ppm0000100.

Silva, A.C.N. da, Branco Vasco, A. and Watson, J.C. (2018). Alexithymia and therapeutic alliance: a multiple case study comparing good and poor outcome cases. *Research in Psychotherapy: Psychopathology, Process and Outcome*, [online] 21(2). doi:10.4081/ripppo.2018.313.

Hobson, H., Hobson, H., Brewer, R. and Bird, G. (2019). *The Role of Language in Alexithymia: Moving Towards a Multiroute Model of Alexithymia*. [online] ResearchGate. Available at: https://www.researchgate.net/publication/330936662_The_Role_of_Language_in_Alexithymia_Moving_Towards_a_Multiroute_Model_of_Alexithymia [Accessed 9 Nov. 2022].

Serani, D. (2014). *The Emotional Blindness of Alexithymia*. [online] Scientific American Blog Network. Available at: https://blogs.scientificamerican.com/mind-guest-blog/the-emotional-blindness-of-alexithymi

a/ [Accessed 9 Nov. 2022].

The Berkeley Well-Being Institute. (2019). *Alexithymia: Definition, Symptoms, & Examples*. [online] Available at: https://www.berkeleywellbeing.com/alexithymia.html [Accessed 9 Nov. 2022].

Brent, D. and ryanlight (2021). *The Emotional Mystery of Alexithymia - Beatanxiety*. [online] Beatanxiety. Available at: https://beatanxiety.me/the-emotional-mystery-of-alexithymia/ [Accessed 9 Nov. 2022].

Ruby, D. (2022). *Whatsapp Statistics 2022 — How Many People Use Whatsapp*. [online] demandsage. Available at: https://www.demandsage.com/whatsapp-statistics/ [Accessed 9 Nov. 2022].

Bawt Gannt Chart.pdf (2022). *Bawt Gannt Chart.pdf*. [online] Google Docs. Available at: https://drive.google.com/file/d/1uVq9iNuUdM6f-_PQQy1I3l0R4u3gsHXy/view [Accessed 9 Nov. 2022].

fin