By: Tony Smith (s3972733@student.rmit.edu.au), Amay Iyer (s3970066@student.rmit.edu.au), Vikas Vikas (s3967625@student.rmit.edu.au), Tabrez Sheikh (s3989516@student.rmit.edu.au)

Proposal for a Sentiment Analysis Data Visualization Tool that Plots a Full Spectrum of Emotions

Introduction:

A sentiment analysis tool is used for mining opinions of users and customers. Organisations can use this tool to gain insights by collecting reviews, and comments from various websites such as Twitter, and Youtube to help better understand what users liked about their product. These organisations can then use this information to boost their popularity, and branding while gaining attention on their products. Sudden changes in sentiment towards a company can also be monitored and bring attention to special situations. [1] Sentiment analysis tools can also be helpful in gaining insights on customer complaints or issues before they become bigger problems. Having tools to proactively address these issues helps the customer, and the company at the same time. During times of crises, for example, a natural disaster, this tool can be implemented to pinpoint where to direct emergency services. Non-Government Organisations (NGOs) using sentiment analysis tools will have the ability to accurately engage with their target base. For example, highlighting workers' emotions on social media and discovering business malpractices being conducted at their company.

Problem Definition:

Current popular sentiment analysis tools generally use binary values for products. [2] For example, Youtube has a like or dislike while Tiktok and Instagram uses a heart for similar purposes. Not having the ability to measure sentiments on social media through texts can create ambiguity in emotions. For example, Youtube's likes and dislikes are supposed to be objective, but users can like a video as a joke versus truly liking the video. This can drive an organisation's key performance indicators and lead to false insights into a video. Businesses may also look at the number of comments on a social media post, news reports, and reviews as a metric with no ability to analyse if the sentiment is positive or negative. Also having the ability to measure sarcasm or idioms may play an important role in finding out how users truly feel about a product. The final issue with modern sentiment analysis tools is accurately visualising how people feel. Tools such as a bar graph or line graph using binary data do not show users' true sentiment as emotions are not binary, but more fluid. For example, a well-liked photo, that tells a powerful story, can create a rush of emotions from anger, sadness, fear, and also joy, but an organisation measuring how well this photo does may come up with a different reason on why users liked a photo. Reproducing the same style of photos will generate less interaction each time when metrics are binary as users become exhausted of the same style.

Significance:

According to GSMA, smartphone penetration in the Asia-Pacific region is going to increase from 76% in 2022 to close to 94% by 2030. [28] This is an increase of 18% of individuals who will be observing content on their phones. Although 18% is not a glamorous number, it should send a few trains of thought in the minds of Non-Government Organisations on how best to leverage the increased amount of traffic they will have access to. Any individual living today with a smartphone has a memory of an online media content that provoked strong feelings. For example, a video of a sea turtle getting a plastic straw pulled out of its nostril which led them to feel strongly about ocean pollution or a graphic image of a refugee child drowning in the ocean and swept onto the beach shore. If there is anything true about emotions, it is that we tend to identify strongly with issues that make us feel special, and that feeling is why we remember that content. Content consumption that provokes an instance of strong feelings and inaction, can mobilise a person to rally for a cause that made them feel passionate about a certain media content. In 2016, Facebook made a change in their user-interface that allowed its users to choose the way they want to react to content posted online. This was an interesting turn of events and completely different from the trajectory that Youtube decided to take (removing dislikes altogether). In the beginning, it was an interesting phenomenon to observe as the same story would be shared to different closed Facebook groups eliciting entirely opposing reactions. This came to a tipping point in 2018 when Cambridge Analytica, a consulting firm, was shown to farm reactions from people for the purpose of political advertising. [25] The reasons Cambridge Analytica was thrusted into the limelight as a scandal was not because of harvesting user data, but rather that the user data harvested was being used for application by any political faction. [26] If scrutinised from a FATE standpoint, organisations like Cambridge Analytica, followed all the necessary guidelines for psychological marketing that Facebook still does to this day, but when the public knows about the true societal consequences of their data being harvested, the tide shifts against their favour as most individuals do not like how their data is being used. Back in 2021, Apple rolled out a fresh privacy feature that allowed users to pick and choose which Applications they would like to allow their data access to. This was a major revenue killer for Facebook as their revenue depended on their users providing them with data. [27] This data would include but is not limited to emotional reactions, likes, dislikes, causes, etc. Clearly, as it has been shown by most public perception, people do not like their data being used to sell them products. Most are ambivalent about the tools that use their data when they are unable to control how that data is used, but when given the choice, they would like their data to be private. Such a predicament places NGOs in a tough position because firms like Cambridge Analytica, that followed most of the FATE principles to roll-out their sentiment analytics tools and got into trouble, might be hesitant to pursue any such processes as they too could fall into the trap. This is why it is critical for NGOs to have reliable tools that meticulously follow FATE principles, in addition to peer-reviewed scientific practices like replicability, to source their data, test it, and then deploy it. These NGOs rally for issues that make the public feel ‘strongly’ about issues, which in turn causes real-world impact by volunteers, politicians, and entrepreneurs alike; therefore, it follows that such NGOs have the best tools to achieve their organisational goals.

Proposed Data Driven Solution:

Plutchik's Wheel of Emotions classifies emotions across 8 different clusters ('joy', 'trust', 'fear', 'surprise', 'sadness', 'disgust', 'anger', and 'anticipation') which are Eurocentric at best. [19] When we apply such a model of emotions to understand the spectrum of emotions of Indigenous people, it seems to fail. Therefore, in order to represent the emotional breadth of indigenous communities better, we propose to add a few more thereby making the full set of emotions as the following: 'joy', 'trust', 'fear', 'surprise', 'sadness', 'disgust', 'anger', 'anticipation', 'pride', 'shame', 'communal_joy', 'nature_connectedness', and 'ancestral_reverence'. [24] Emotions run deep and something that runs deeper is the distinction between the mind and the body. A deep sense of spirituality underlies Indigenous communities, a metric that is largely ignored in Eurocentric models of emotions like the Plutchik Wheel. Now that we have a proposal for an improvement, the Machine-Learning Clustering model that we feel would be best at exemplifying this range of emotion would be t-Distributed Stochastic Neighbour Embedding (TSNE). Now that we have our basis of the clusters that ground our model, we need to select the axes that would compose the three dimensions. For this, we chose Prototypical Emotion Recognition, Prototypical Emotion Knowledge, and Advanced Emotion Understanding as from a young age, children's emotions are structured across those three factors. [23] Now that we have an initial proposal for a model, let's go over how

Methodology (As per the FATE principles)**:**

1. **Start with the Basic Emotion Model (Transparency, Fairness):** We should begin by crafting a 3D visualisation of the Plutchik emotion wheel. We will assign initial values for Prototypical Emotion Recognition, Prototypical Emotion Knowledge, and Advanced Emotion Understanding for each of the 8 emotions recognized by Plutchik ('joy', 'trust', 'fear', 'surprise', 'sadness', 'disgust', 'anger', and 'anticipation'. Being transparent about these initial parameters is essential, as they lay the foundation for subsequent analyses and visualisations. Additionally, ensuring fairness in these initial values is also crucial, avoiding any inadvertent bias towards any particular emotion ensures we do not repeat the mistakes made by Meta.

2. **Expand the Emotional Spectrum (Ethics, Fairness):** After 1. We should ethically enhance the initial model to capture a broader array of emotions by incorporating additional ones so that the model will be a better representative of not just Eurocentric emotional pattern, but also indigenous and abstract ones such as pride, shame, communal_joy, nature_connectedness, and ancestral_reverence bringing us to a total of 13 emotions ('joy', 'trust', 'fear', 'surprise', 'sadness', 'disgust', 'anger', 'anticipation', 'pride', 'shame', 'communal_joy', 'nature_connectedness', and 'ancestral_reverence'). This step ensures fairness by treating a more comprehensive range of emotions equally in the model.

3. **Generate More Data (Accountability, Replicability):** Followed by 2, we should aim to increase the robustness of our visualisation model by generating extra data. Create additional samples randomly from the original set of emotions and the three

emotion-related value parameters. An accountable approach is required to maintain data integrity. Furthermore, this step increases the likelihood of replicability by producing a larger and more diverse dataset.

4. **Prepare the Data (Transparency, Ethics):** Continuing from 3, we should combine our data into a pandas DataFrame. The 'emotion' column is categorical, therefore we should convert it into numerical data via dummy/indicator variables. This preprocessing step necessitates transparency for proper understanding and ethical considerations, ensuring respect for user privacy during data handling and transformation.

5. **Apply Dimensionality Reduction (Explainability):** We will use the t-SNE algorithm to reduce the dimensionality of our dataset. This technique, when properly explained, allows for high-dimensional data to be understood in a 3D format, improving the explainability of the model.

6. **Cluster the Data (Fairness, Explainability):** We will then apply KMeans clustering on our 3D data to create groups of similar data points. We will ensure fairness in this step by starting with a small number of clusters, allowing for equal opportunities for patterns to emerge from the data. Explainability is also key in this step to make the clustering process comprehensible.

7. **Fine-tune the Clustering (Transparency, Accountability):** Based on the patterns we observe, we should fine-tune the number of clusters to better capture nuances in our data. Being transparent about this adjustment, explaining the rationale behind the decision, and holding ourselves accountable for the resulting changes to the data structure is integral for a robust product.

8. **Create the Visualisation (Explainability, Ethics):** Finally, after meticulously following all the principles we will plot the 3D data points, their cluster labels, and corresponding emotions using a Python graphing library like Plotly. Creating an interactive visualisation that is both understandable (explainable) and respectful of the emotions it represents (ethics) is paramount.

Throughout this process, it is important to test the model iteratively on larger and larger datasets to evaluate its consistency and reproducibility. Following the FATE principles, we can ensure our sentiment analysis tool is not only effective but also ethically designed and considerate of user privacy and fairness.

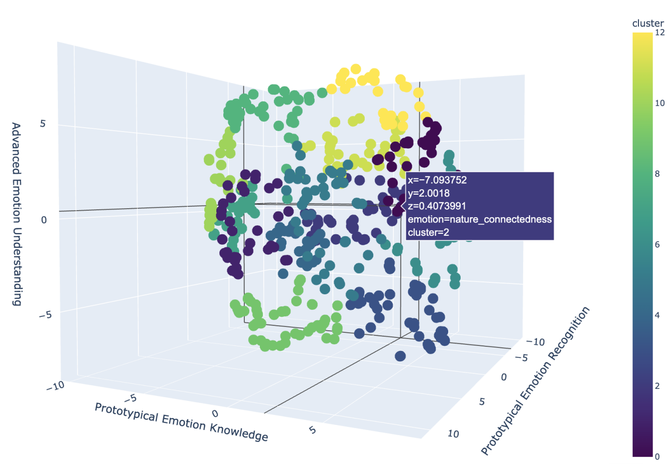
FATE Principles (as they apply to our project):

Educating people on data literacy and fostering trust in the data industry is a crucial aspect of the approach to improving models and systems with higher emotional awareness. This emphasis on data transparency will empower users to recognise their own emotional biases, leading to informed decision-making and effective strategies [16]. However, the lack of sentiment analysis tools that can plot emotional sentiments in a three-dimensional matrix restricts NGOs, impacting their ability to select and target specific emotions in marketing campaigns, which in turn affects their outreach and engagement efforts [17]. For instance, consider a video of a plastic straw getting pulled out of a sea-turtle's nostril. If they were to produce a YouTube video, it would not go viral due to YouTube's disabling of dislikes and lower search ranking [13]. If the video was posted on Facebook, it may gain virality, but this merely represents user interaction with the content, not active lobbying to regulate the way plastic is disposed of.

Applying the Fairness principle of the FATE framework to Facebook's solution reveals that the algorithm's bias towards virality driven by anger is unfair, disproportionately prioritising one emotion over others. Transparency is another significant principle, and while Facebook's API is publicly available, we cannot conclusively ascertain the extent of potential infringement on individual user privacy [15]. Moreover, the explainability principle suggests that spreading anger and outrage is the only clear action one can take to create viral content. However, such an emotionally charged approach can cloud rational judgement [19].

Ethically, there should be stringent scrutiny of how NGOs leverage these models to advance their objectives, as the potential for misuse and financial exploitation of such emotion-driven tools exists. Reproducibility is equally crucial, especially given the known challenge in social sciences: The Replication Crisis. This refers to the difficulty of reproducing study results across large datasets [20] [21]. Given our engagement with sentimental data, ensuring robustness through iterative testing on increasingly larger datasets is integral to ascertain the model's fair classification of sentiments. A possible solution that aligns with these principles is offered by Affectiva, providing qualitative insights through actionable steps involving focus groups [12]. However, it poses financial constraints for average NGOs and does not provide free or accessible data for visualisation in a 3-Dimensional matrix [1] [2] [3] [17]. Thus, while working towards upholding the FATE principles, accessibility and financial feasibility remain key challenges.

Design Prototype:



If we meticulously follow the FATE principles and the incremental, iterative improvements laid out in the Methodology section, we should be able to cluster emotions along a 3-Dimensional axis on a chart like we have visualised above. To make the above visualisation, we adhered to the FATE principles (Fairness, Accountability, Transparency, Ethics, and Explain-ability):

1. **Fairness:** The code used a balanced approach to generate data for the range of emotions studied. Each emotion was assigned equal consideration in the data generation process, which eliminated potential bias in data representation.

2. **Accountability:** Accountability was embedded in the way the code created, manipulated, and used the data. Any issues or errors that arose from the processing could be traced back to the specific steps of the code. Additionally, the use of established and reliable libraries like pandas, numpy, and scikit-learn, further ensured accountability in the processing and analysis.

3. **Transparency:** Transparency was demonstrated in the clear, step-by-step structure of the code. Each section of the code had a clear purpose, and the overall process could be easily followed, understood, and scrutinised. Furthermore, the use of a scatter plot for visualisation ensured the results could be easily interpreted and understood, further promoting transparency.

4. **Ethics:** The code upheld ethical considerations by treating all emotions with equal importance, and by not employing any form of data manipulation that would prioritise one emotion over another. Moreover, the generation of additional data was randomised to avoid any intentional or unintentional bias, preserving the integrity of the results.

5. **Explainability:** The variable names were intuitive and easy to understand for any individual reading the code.. The visualisation also contributed to explainability by presenting the data and results in a manner that was visually intuitive and easy to interpret. The hovering text also adds to its explainability.

6. **Reproducibility:** The code set random seeds to ensure that the generation of random data was the same every time the code was run, which enabled reproducibility of the results. The use of libraries like pandas, numpy, and scikit-learn, which have standardised and well-documented functions, also ensured that the results could be reliably reproduced.

The visualisation is a 3-dimensional scatter plot that presents a visual representation of the clusters of emotions, based on their associated features. The visualisation serves as a transparent and interpretable way of understanding the relationships between different emotions, as similar emotions (in terms of the features) are grouped together in the same cluster. In the context of an NGO, this code and visualisation offer a replicable and transparent way of analysing emotional responses to various campaigns. This analysis can help understand which emotions are most prevalent and how they're interrelated, aiding in the development of more effective and ethically sound strategies for spreading awareness and spurring action.

Conclusion:

In light of our comprehensive investigation into emotional representation and sentiment analysis, a number of significant insights have emerged. These tools, while providing detailed scrutiny of user sentiments relating to a myriad of products, events, or situations and proactively mitigating arising problems, are currently limited in their capability to accurately represent user sentiment. This is due, in part, to the fact that emotions encompass a vast spectrum, and not just a binary perspective. The practice of binary classification of sentiment often results in a distorted representation, manipulated to increase popularity or attention, without truly understanding the consequential impact on an organisation's reputation or product. As a remedy, we advocate for a broader measurement of emotional range for a more nuanced perception of user sentiments. Underpinning this is the importance of leveraging machine learning and natural language processing methods, such as SVMs or naive bayes for sentiment categorization and tokenization and stemming for text preprocessing. Combined with web scraping and APIs, these methods can discern user emotions on popular platforms through textual analysis. Visualisation of these sentiments on a 3D graph with vibrant colours provides a genuine representation of customer sentiment, enabling organisations, including NGOs, to acquire a deeper understanding of their audience's sentiments and to facilitate more effective campaigns and initiatives.

Fidelity to FATE principles (Fairness, Accountability, Transparency, Ethics, Explainability, and Reproducibility) is essential in sentiment analysis and emotional mapping. This commitment ensures the creation of unbiased models that value user privacy and uphold transparency in their operations. Accountability is due from organisations that disregard these principles, considering the significance of trust and responsibility in dealing with such influential tools. Our investigation also underscores the need for better Indigenous representation, understanding of online emotion and reaction proliferation, and complex emotional mapping. In closing, the shortfalls of current sentiment analysis tools exemplified in YouTube's binary rating and Facebook's reactions, accentuate the demand for a more sophisticated and innovative solution to address the challenges encountered by NGOs and other organisations in our data-centric world. With an emphasis on FATE principles—fairness, accountability, transparency, ethics, and explainability—this fresh approach can enable organisations to more effectively connect with their target audience, effect meaningful change, and make informed decisions. Crucially, a balanced rating system that equally represents all emotions is vital for a fair sentiment analysis tool. This would curb the undue prioritisation of certain emotions, such as anger, and enable organisations to access a broader range of emotional responses. Shared accountability among NGOs, users, and visualisation models, along with data literacy education and data transparency emphasis, can help build trust and ensure ethical and responsible data usage. This improved awareness can lead to more accurate sentiment analysis, mutually benefiting organisations and the communities they serve. The deficiency of sentiment analysis tools capable of 3D visualisation impairs NGOs' ability to effectively target specific emotions in their campaigns, thereby affecting their outreach and engagement efforts. A 3D matrix visualisation tool that offers enhanced explainability and presentation capabilities would better equip decision-makers to create informed strategies and fully utilise sentiment analysis. Addressing the ethical implications of sentiment analysis tools is equally important, ensuring that they do not exploit specific emotions or unfairly allocate resources. Through careful consideration of how NGOs utilise these tools and by promoting responsible use, this innovative sentiment analysis solution can drive positive impact for organisations and their served communities.

In summary, a novel sentiment analysis solution, incorporating FATE principles and advancements in visualisation technology, can revolutionise how NGOs and other organisations connect with their target audience and effect meaningful change. By overcoming current tool limitations and fostering a more emotionally resonant approach, this pioneering solution will play a crucial role in constructing a more just, equitable, and informed world.

Bibliography

[1] Gonçalves, P, Araújo, M, Benevenuto, F & Cha, M 2013, ‘Comparing and combining sentiment analysis methods’, *Proceedings of the first ACM conference on Online social networks - COSN ’13*, viewed 5 July 2019, <https://dl.acm.org/citation.cfm?id=2512951>

[2] Andrea, A, Ferri, F & Grifoni, P 2015, ‘Approaches, Tools and Applications for Sentiment Analysis Implementation’, *International Journal of Computer Applications*, vol. 125, no. 3, pp. 975–8887

[3] Moz Pro Pricing 2023, Moz website, accessed 23 March 2023, <https://moz.com/products/pro/pricing>

[4] The Ultimate SEMrush Guide 2023, Udemy website, accessed 23 March 2023, <https://www.udemy.com/course/ultimate-semrush-course/?utm_source=adwords&utm_medium=udemyads&utm_campaign=DSA_Search_la.EN_cc.AU_PP&utm_term=_._ag_137987506236_._ad_606065560737_._kw__._de_c_._dm__._pl__._ti_dsa-1705455363324_._li_9071354_._pd__._&matchtype=&gclid=Cj0KCQjwiZqhBhCJARIsACHHEH-utUTpz9nY6BSD7b_uAZS-y-m8_wGVCd3Dv5A8rCAMFf05H9jNQLkaAqUMEALw_wcB>

[5] Google Trends Tutorial for Beginners 2023, AnsonAlex website, accessed 23 March 2023, <https://ansonalex.com/tutorials/google-trends-tutorial/>

[6] The UX of rating systems, Medium website, accessed 29 March 2023, <https://uxdesign.cc/the-ux-of-rating-systems-bc4f9d424b90>

[7] Which Rating System is Best for Your User?, Medium website, accessed 29 March 2023, <https://medium.com/enjoyhq/which-rating-system-is-best-for-your-users-9fddeeb455b5>

[8] 5 stars vs. thumbs up/down-which rating system is right for your app?: Appcues Blog, appcues website, accessed 30 March 2023, <https://www.appcues.com/blog/rating-system-ux-star-thumbs>

[9] What Does Removing Likes Mean for the Future of Social Media?, Digital Stand website, accessed 29 March 2023, <https://www.digitalstand.com/blog/removing-likes-from-social-media>

[10] Collaborative Filtering with Likes and Dislikes, davidcel website, accessed 29 March 2023, <https://davidcel.is/articles/collaborative-filtering-with-likes-and-dislikes/>

[11] Designing the user experience of a rating system, Medium website, accessed 29 March 2023, <https://uxdesign.cc/designing-the-user-experience-of-a-rating-system-2c6a4d33bb11>

[12] Affectiva Media Analytics for Qualitative Research, Affectiva website, accessed 29 March 2023, <https://www.affectiva.com/product/affectiva-media-analytics-for-qualitative-research/>

[13] Five points for anger, one for a 'like': How Facebook's formula fostered rage and misinformation, The Washington Post website, accessed 29 March 2023, <https://www.washingtonpost.com/technology/2021/10/26/facebook-angry-emoji-algorithm/>

[14] Facebook formula gave anger five times weight of likes, documents show, The Hill website, accessed 29 March 2023, <https://thehill.com/policy/technology/578548-facebook-formula-gave-anger-five-times-weight-of-likes-documents-show/>

[15] Facebook Reaction Object API's, accessed 6th Apr 6, 2023<https://developers.facebook.com/docs/graph-api/reference/v16.0/object/reactions>

[16] Data Transparency, Bias, and AI, Mozilla Distilled, accessed Apr 6, 2023

<https://blog.mozilla.org/en/mozilla/next-steps-on-trustworthy-ai-transparency-bias-and-better-data-governance/>

[17] Visual Representation: Implications for Decision Making, Sage Publications; Lurie, Nicholas H, accessed Apr 6, 2023

<https://journals.sagepub.com/doi/10.1509/jmkg.71.1.160>

[18] The Psychology of Rating System, Himanshu Khanna, accessed Apr 6, 2023

<https://medium.com/hackernoon/the-psychology-of-rating-systems-3103e26fddd8>

[19] Putting some emotion into your design: Plutchik's Wheel of Emotions, April 2020,

<https://www.interaction-design.org/literature/article/putting-some-emotion-into-your-design-plutchik-s-wheel-of-emotions>

[20] Hunter P. The reproducibility "crisis": Reaction to replication crisis should not stifle innovation. EMBO Rep. 2017 Sep;18(9):1493-1496. doi: 10.15252/embr.201744876. Epub 2017 Aug 9. PMID: 28794201; PMCID: PMC5579390.

<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5579390/#:~:text=One%20of%20the%20largest%20meta,costs%20of%20drug%20development%201.>

[21] Replication Crisis: What is wrong and how to fix it

<https://fantasticanachronism.com/2020/09/11/whats-wrong-with-social-science-and-how-to-fix-it/>

[22] Estimating the Reproducibility of Psychological Sciences

<https://psych.hanover.edu/classes/cognition/Papers/Science-2015--.pdf>

[23] Prototypical Emotion Recognition, Prototypical Emotion Knowledge, and Advanced Emotion Understanding https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5679467/

[24] “SHARING THE “MAGIC”: EXPLORING THE PHENOMENON OF COMMUNI COMMUNITY AND N Y AND NATURE CONNEC TURE CONNECTEDNESS IN ANCESTRAL TEDNESS IN ANCESTRAL SKILLS EDUCATION” Graham, 2021
https://scholarworks.umt.edu/cgi/viewcontent.cgi?article=12880&context=etd

[25] Chan, 2020
https://www.businessinsider.com/cambridge-analytica-whistleblower-christopher-wylie-facebook-data-2019-10

[26] Cadwalladr, 2020
https://web.archive.org/web/20200118093155/https://www.theguardian.com/technology/2017/may/07/the-great-british-brexit-robbery-hijacked-democracy

[27] Koetsier, 2022
https://www.forbes.com/sites/johnkoetsier/2022/02/19/apples-att-burned-facebook-bad-googles-privacy-sandbox-is-a-kiss-in-comparison/?sh=2972afb0382d

[28] GSMA, 2023
https://www.gsma.com/mobileeconomy/wp-content/uploads/2023/03/270223-The-Mobile-Economy-2023.pdf