

Analysis of product feedback from platforms such as Reddit and GitHub

Supervisor: John (Al) Alphonsus Matthews (john.alphonsus.matthews@ut.ee)

Co-supervisor: Kuldar Taveter (kuldar.taveter@ut.ee)

Platforms such as Reddit and GitHub provide us with the opportunity to better understand user experiences of a variety of products. Through Application Programme Interfaces (APIs), we can extract information such as comments from these platforms, which in turn enable us to apply modern Machine Learning and Artificial Intelligence techniques. For example, allowing us to categorise information under themes such as how difficult, easy, and enjoyable a product is to use, and if users would recommend it or not.

This project offers the opportunity to agree with the supervisor on a specific set of products and what platforms we can study. Our research questions could be:

1. Is the platform we have chosen a useful source of information?
2. What can we recommend as improvements to a particular product?
3. What is working well for the product?

Python can be used for much of the analysis, using natural language packages such as nltk[1] and machine learning packages such as sklearn[2]. Other options are possible if the student prefers. This is an active area of research, with studies in similar fields available for reference [3,4].

References

- [1] Bird, S., Klein, E., & Loper, E. (2009). Natural language processing with Python: analyzing text with the natural language toolkit.
- [2] Pedregosa, F. et al., 2011. Scikit-learn: Machine learning in Python. *Journal of machine learning research*, 12(Oct), pp.2825–2830.
- [3] T. Iqbal, M. Khan, K. Taveter and N. Seyff, "Mining Reddit as a New Source for Software Requirements," 2021 IEEE 29th International Requirements Engineering Conference (RE), 2021, pp. 128-138, doi: 10.1109/RE51729.2021.00019.
- [4] Kim, Seoyun, Junyeop Cha, Dongjae Kim, and Eunil Park. "Understanding mental health issues in different subdomains of social networking services: computational analysis of text-based Reddit posts." *Journal of Medical Internet Research* 25 (2023): e49074.

Agent-based Models of the Media Landscape

Supervisor: John (Al) Alphonsus Matthews (john.alphonsus.matthews@ut.ee)

Co-supervisor: Kuldar Taveter (kuldar.taveter@ut.ee)

Agent-based modelling (ABM) approaches are used to model how information within the media landscape is communicated in society. For example, we can consider the likes of “Echo Chambers” and “Filter Bubbles” [1,2,3,4]. Modelling the media landscape is achieved by analysing the various actors and interactions that define it. In these modelling scenarios, the likes of journalists and consumers are considered agents or actors. Further, their various interactions include the transfer of information between one another.

To model the media scenarios, we will work primarily with the NetLogo software program [5,6]. We will discuss what we would like to simulate, including some clearly defined research questions. The goal is to design a system of agents and relationships, defined with environmental parameters to simulate different media and information transfer scenarios. This will include defining a probabilistic network of connections between various agents, including how people will align into different groups [7].

References

- [1] Geschke, Daniel, Jan Lorenz, and Peter Holtz. "The triple-filter bubble: Using agent-based modelling to test a meta-theoretical framework for the emergence of filter bubbles and echo chambers." *British Journal of Social Psychology* 58, no. 1 (2019): 129-149.
- [2] Sulis, Emilio, and Marcella Tambuscio. "Simulation of misinformation spreading processes in social networks: an application with netlogo." In *2020 IEEE 7th International Conference on Data Science and Advanced Analytics (DSAA)*, pp. 614-618. IEEE, 2020.
- [3] Tambuscio, M., Ruffo, G., Flammini, A. and Menczer, F., 2015, May. Fact-checking effect on viral hoaxes: A model of misinformation spread in social networks. In *Proceedings of the 24th international conference on World Wide Web* (pp. 977-982).
- [4] Tambuscio, Marcella, and Giancarlo Ruffo. "Fact-checking strategies to limit urban legends spreading in a segregated society." *Applied Network Science* 4, no. 1 (2019): 1-19.
- [5] Sterling, Leon, and Kuldar Taveter. *The art of agent-oriented modeling*. MIT press, 2009.
- [6] Sulis, Emilio, and Kuldar Taveter. "Agent-Based Business Process Simulation."
- [7] Tambuscio, Marcella, Diego FM Oliveira, Giovanni Luca Ciampaglia, and Giancarlo Ruffo. "Network segregation in a model of misinformation and fact-checking." *Journal of Computational Social Science* 1, no. 2 (2018): 261-275.