**Literature Review – Data Science I Final Project**

*How do people react to shocking news? What are the effects?*

Shocking events have significant impacts on both the individual and societal level. Poutvaara and Ropponen, for instance, studied the effects of news of school shootings on cognitive performance of students. The research team found that news of school shootings led to worse performance on exams for male students, with test scores dropping by 4.3 points (Poutvaara and Ropponen 2018). A study by Koutra et. al. observed the effects of news retrieval and the “filter bubble,” or the phenomenon of individuals receiving news tailored to their ideological stances. Like Poutvaara and Ropponen, they focused on news of school shootings, and observe news retrieval tendencies before, during, and after the event. They found that respondents often consumed news that was more amenable to their ideological tendencies, and that only when news challenges existing views did individuals consume content outside of their “filter bubble” (Koutra et. al. 2015). This consumption bias is explored in further studies. Pinkerton and Zhou, for example, explore the psychological and physical effects of “negative” news on individuals, especially those who are predisposed towards “morbid curiosities” (Pinkerton and Zhou 2008). They found that those with higher likelihood of “morbid curiosities” are likely to react physically to reading about bad news, although their self-reported interest in the news may not differ significantly from those who do not have morbid curiosities (Ibid). Participants, regardless of level of morbid curiosity, preferred to read negative news stories rather than positive ones (Ibid).

The effects of news and methods of shocking news retrieval can also have more nefarious effects as well. Chen et. al. explored if and how internet platforms peddle harmful or extremist content. They employed a 1,000+ person survey, and observed exposure to YouTube channels that promote extremist views. It was found that small groups of individuals with pre-existing prejudices were consumers of some of the more ideologically extreme YouTube channels, and were more likely to subscribe to these channels as well, whereas non-subscribers were less likely to view hateful YouTube videos, and were also less likely to fall through YouTube “rabbit holes” (Chen et. al. 2023). Exposure to “bad” news also has the potential to degrade people’s emotional states. In a 2019 study with 63 respondents, De Hoog and Verboon explore the effects of news exposure on emotional states overtime, finding that exposure to negative news perceptions leads to more negative affects, but that these effects are stronger if the individual in question has personal connections to the matter at hand, and that personality traits themselves did not play as strong of a role (De Hoog and Verboon 2019).

*Sentiment Analysis*

Sentiment analysis is the use of methods such as natural language processing (NLP) and various text-as-data processing methods to observe trends, patterns, and beliefs in online reviews or comments (Hussein 2018). The use of sentiment analysis in relation to news is a well-documented, and it occupies a substantial niche in current literature. Deori et. al., for instance, used sentiment analysis to monitor trends on Indian news channels. They find that entertainment and politics were most popular categories for consumers, and that viewers were often unhappy with the news they were seeing (Deori et. al. 2021).

Sentiment analysis has also been used to observe comments on various websites, including YouTube. Komsah employed sentiment analysis in a 2021 study, which analyzed over 30,000 YouTube comments for the purpose of determining the accuracy of random-forest methods of machine learning (Komsah 2021). The use of methods like these in sentiment analysis is not a new phenomenon—machine learning methods have been an important aspect of YouTube comment sentiment analysis for years. Singh and Tiwari use multiple machine learning methods to analyze a corpus of over 1500 sentences (Singh and Tiwari 2021). They employed methods such as decision trees, random forest, Naïve-Bayes, support vector machine (SVM), and K nearest neighbor (KNN) in their research, all using the Scikit-Learn package in python (Ibid). Different methods tended to have their individual strengths: decision trees with n-grams gave the best F-scores in the study, while random forest and support vector machine have high accuracy (Ibid). [Mukwazvure](https://ieeexplore.ieee.org/author/37085701288) and Supreethi also used methods such as SVM and KNN in their 2015 study which used sentiment analysis to examine polarity of online news comments ([Mukwazvure](https://ieeexplore.ieee.org/author/37085701288) and Supreethi 2015). They found that when classifying comments as positive, negative, or neutral, SVM was more successful at identifying neutral articles than KNN (Ibid).

Thelwall et. al. go a bit deeper and explore not only the methodologies used for sentiment analysis, but the different types of people that comment on YouTube videos. They find that overall, over 20% of the comments observed were actually replies to other comments (Thelwall et. al. 2011). In addition, most of the comments made in response to others were negative comments, with the controversial topic of religion being the topic that generated the most replies (Ibid).

While a robust collection of literature on sentiment analysis on news reactions and on YouTube comments does exist, there are few if any studies on how sentiments change in relation to different types of news on YouTube. Our project will attempt to bridge this divide.