Farishah Nahrin
EMSE 6301.M02
April 10, 2023

<div align="center">ACL Paper Summary</div>

The title of the paper is Clickbait Spoiling via Question Answering and Passage Retrieval. The authors of this paper include Matthias Hagen, Maik Fröbe, Artur Jurk, Martin Potthast. Based on initial findings, it seems that Maik Fröbe and Artur Jurk do not have Google Scholar profiles. Artur Jurk is credited on Google Scholar for a few published papers but does not have a profile. Maik Fröbe also does not have a Google Scholar profile, he is a PhD student who specializes in Information Retrieval and NLP. Although he doesn't have a profile, Maik Fröbe has contributed to many papers on Google Scholar. Matthias Hagen is affiliated with Friedrich Schiller University Jena, an institution in Germany. And Martin Potthast is affiliated with the Leipzig University in Germany. Both Hagen and Potthast have listed informational retrieval and natural language processing as their interests/specialization on Google Scholar.

This paper explores finding a way to create a 'spoiler' for clickbait posts, such as from tweets and headlines. The authors defined clickbait as a way to "inappropriately entice" readers to visit the source's page. The authors believe that clickbait is inappropriate because the words or phrases that are usually left out of a title could have, as the authors put, "easily" included in the post. This is the main problem addressed by the authors; they are not keen towards headlines that stimulate interest, but rather headers should be a summary with more of an informative tone. The act of clickbaiting and its tendency to mislead readers, inspired the authors to create a way to spoil a clickbait with a few words, in order to get the gist of a headline or post containing clickbait.

Some of the popular prior work relating to preventing clickbait dates back to 2014. Upworthy, a website that aims to create viral headlines and posts to endorse pollical and social stories, became one of the major sources of clickbait on Facebook. Since they were so successful in promoting their clickbaited posts, Facebook released statements in 2014 and 2016 regarding changing their news recommendation algorithms, in order to try to reduce the amount of clickbait headlines that are shown to their users. In 2014, Facebook tried to hide or remove the posts based on the click-through-rate. As in, if a user clicked on an article and spend time on it, then it was not considered clickbait. On the other hand, if the user came back to their news feed after clicking on the article, they would lower the rank for that story. They also factored the number of shares, likes, and comments to factor in the rank. Then in 2016, Facebook stated that they built a system, which implemented NLP techniques, that compared the

clickbait titles to their corpus of phrases that are commonly used in clickbait headlines, in order to rank the articles on a person's news feed.

Some other previous work includes viewing clickbait as a question and answering problem, which in NLP, is when the system answers questions that are asked in a natural language by humans. So, Hagen et al., needed models to start their evaluation of clickbait headlines. According to Rajpurkar et al., they developed SQuAD (Stanford Question Answering Dataset), which contained a little over 100,00 questions and answers based on five hundred plus Wikipedia articles, where a little over 90% were factual related text, such as numbers, verb phrases, adjectives). The second model that has been used as a question and answering benchmark is TriviaQA. According to Joshi et al., they gathered about 95,000 question and answer pairs regarding difficult trivia questions, which could be useful when comparing clickbait. According to Guo et al, neural network models have been used on *passage* retrieval techniques successfully.

Hagen et al., created the Webis Clickbait Spoiling Corpus 2022, which contained 5,000 clickbait posts and the spoilers. The corpus comprised of these accounts: r/savedyouaclick, @HuffPoSpoilers, @SavedYouAClick, @UpworthySpoiler, and @StopClickBaitOfficial (from Twitter, Reddit, and Facebook), where Hagen and his team manually reviewed and labeled the spoilers for 5,000 posts. They observed there was two main types of spoilers: phrasing spoilers (single word or phrase) and passage spoilers (sentences describing the spoiler).

Hagen et al., tried various number of techniques in order to find a way to best classify the spoiler type. First, to try to find to find phrase spoilers, they used the question and answering methods that was used on the SQuAD data and then fine-tuned it on their Webis corpus, using methods like ALBERT, DeBERTa, etc, which are all transformer models. Then to find passage spoilers, they used other types of term-based retrieval methods and language models, like MonoBERT, etc. They found that transformer-based classifiers were more effective than feature based ones, such as DeBERTa and RoBERTa. They conducted both of spoiler types, in order to see if it can extract spoiler phrases or spoiler passages better, from the linked documents.

Based on the above study results, they selected six more models for hyperparameter tuning against DeBERTa. And DeBERTa was still found to be the most effective. They found that "type-agnostic question answering-based spoiling" was the most effective way to extract a spoiler. This means that when the models analyze the article to figure out the most important information, they find the spoiler as the phrase that "answers the specific question" about the article. So if a headline said, "This food will change your life!" The spoiler in this method of spoiling would convert this headline to a question and then have the spoiler be the answer.

Clickbait tweet   **Spoiler**

Above the Law ✔ @atlblog

The Surprising Way Recent Law
School Graduates Are Getting Their   "Networking."
First Job bit.ly/2CMMPxf

Figure 1: Examples of clickbait tweets and spoilers for them extracted from the respective linked web page (Hagen, et al.).

Matthias Hagen and Martin Potthast are the only two authors with a Google Scholar profile. Potthast has the most citations (10137) compared to Hagen (3462). I believe their work is important because clickbait can be harmful sometimes. If a person does not read an entire article and bases their information from headlines, then that could cause the spread of false information. The research that Hagen et al., did can show social media platforms a new approach to combatting clickbait. Social media platforms could potentially use this research to adopt a technology to spoil an article to their users before they click on it, as a feature (Twitter has recently added a misinformation label feature underneath certain misleading tweets). This research can also help users understand the seriousness of how important it is to feed social media users or users, in general, with the truth.

References:

Matthias Hagen, Maik Fröbe, Artur Jurk, and Martin Potthast. 2022. Clickbait Spoiling via Question Answering and Passage Retrieval. In Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), pages 7025–7036, Dublin, Ireland. Association for Computational Linguistics.

Pranav Rajpurkar, Robin Jia, and Percy Liang. 2018. Know What You Don't Know: Unanswerable Questions for SQuAD. In Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics, ACL 2018, 2: Short Papers, pages 784– 789. Association for Computational Linguistics.

Mandar Joshi, Eunsol Choi, Daniel S. Weld, and Luke Zettlemoyer. 2017. TriviaQA: A Large Scale Distantly Supervised Challenge Dataset for Reading Comprehension. In Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics, ACL 2017, Vancouver, Canada, July 30 - August 4, Volume 1: Long Papers, pages 1601–1611. Association for Computational Linguistics

Jiafeng Guo, Yixing Fan, Liang Pang, Liu Yang, Qingyao Ai, Hamed Zamani, Chen Wu, W. Bruce Croft, and Xueqi Cheng. 2020. A Deep Look Into Neural Ranking Models for Information Retrieval. Inf. Process. Manag., 57(6):102067.