A Summary of Clickbait Spoiling via Question Answering and Passage Retrieval

Clickbait Spoiling via Question Answering and Passage Retrieval is a paper published by Matthias Hagen, Maik Fröbe, Artur Jurk, and Martin Potthast studying and detailing their approach to the task of clickbait spoiling. They aimed to classify the type of spoiler required and create appropriate spoilers for the clickbait presented.

A corpus of identified clickbait with spoilers was required to create the models to identify clickbait and generate spoilers. Their corpus was primarily based on five social media accounts that manually spoil clickbait. Posts were manually reviewed to determine if the post was spoiled clickbait. A final corpus of 4,204 posts was the result. Question answering and passage retrieval models combined with their corpus was how they created their spoiler-type classifier and spoiler generator.

Question answering and passage retrieval models combined with their corpus was how they created their spoiler-type classifier and spoiler generator. They found that there are three types of spoilers: phrase spoilers, passage spoilers, and multipart spoilers. A spoiler-type classifier was able to be created achieving an accuracy of 80% using the DeBerta-large model. It was also the best for spoiler generation as it performed moderately well in creating a correct spoiler for both phrase and passage type spoilers (35-41%).

To evaluate spoiler type classification, it was assumed that a post has been correctly identified as clickbait, then they conducted 3 experiments: multi-class, one-vs-rest, and one-vs-one for the types of phrase and passage spoilers. All hyperparameters were optimized based on the validation set of their corpus. The RoBERTa and DeBERTa substantially outperformed (9-10 accuracy points) Naïve Bayes, SVM, and Logistic regression.

For evaluating spoiler generation quantitative and qualitative measures were introduced and thresholds were manually determined to define a spoiler as ‘correct’. To assess the quantitative measure three question-answering measures and one passage retrieval-oriented measure: BLEU-4, METEOR, BERTScore, and Precision@1. The question-answering measures used n-gram overlap and matching word pairs to determine similarity to ground truth. The passage retrieval measure was also used to determine whether the top-ranked passage contains the ground truth (all phrase spoilers and 98% of passage spoilers come from a single passage). A confidence threshold of 50% was set after the assessment of the results.

Mattias Hagen has 3455 citations on Google Scholar. I think his work here is important because it’s trying to address the ways social media manipulates its users. Clickbait spoiling is a task that helps users that don’t want to be manipulated by their curiosity to click a link that usually has a very mundane answer. This answer can often fit within the post that has the link in the first place. Work like this can be used to filter out clickbait posts from timelines and clickbait spoiling can immediately subvert the curiosity induced by a clickbait post. For me at least, I would love to see clickbait filtered out of my social media.