CSE517: Project Proposal

Team Cookies: Author 1, Author 2, Author 3

Citation and Hypotheses:

To verify the decision-focused Summarization proposed in this paper [1], which is used to summarize relevant information for a decision, outperforms text-only summarization methods and model-based explanation methods in decision faithfulness and representativeness.

Usage of Data from Paper:
The paper uses the yelp dataset, freely available online at .

Whether existing code will be used:
We will use the code provided by the paper, which is available at .

Discussion of feasibility:
We will follow the methodology described in the paper, which is essentially as below.

Given an input text $X = \{x_s\}_{s=1}^{s=S}$, where S is the number of sentences, we wish to select a subset of sentences $\tilde{X} \subset X$ to support making a decision y. Thus, the training set, $D_{train} = \{(X_i, y_i)\}$, will be processed in such a way as to find defining traits in a text that lead to a certain decision y. The yelp dataset is defined as follows: for each restuarant, define X as the text body of the first k reviews and let y be the average rating of the first t reviews, where $t > k$ s.t. the problem is to predict future ratings. In the paper, the research group used $k = 10, t = 50$ s.t. their task was to find sentences from a given restaurant's first 10 reviews that supported a prediction of its future rating after 50 reviews.

At this level of abstraction of our project to reproduce the paper, feasibility seems high, given that the yelp data they used is publicly available online, and their documentation for their code is on Github and seems to be well-documented. Possible areas to look into at this level of abstraction include choosing different amounts of data from the yelp dataset and different t and k values to see their effect on the model's performance.

Taking a step down the abstraction ladder, we see that the authors of the paper split the entire yelp dataset of 18,112 restaurants into training, validation, and test datasets. [Insert few sentences about Section 3.1 from the paper and insert the paper's provided "Algorithm 1"]

[Insert paragraph about section 3.2 on evaluation metrics]

[Insert blurb about how we feel about feasibility at this lower level of abstraction and mention not being able to use human evaluation (Amazon mechanical turk) for our reproduction.]

References

[1] Chao-Chun Hsu and Chenhao Tan. Decision-focused summarization. In *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*, pages 117–132, Online and Punta Cana, Dominican Republic, November 2021. Association for Computational Linguistics.