Improving Quality Assessment of Online Reviews Using Formal Argumentation Theory and Knowledge Graphs

Anonymous authors

Abstract

In this extended abstract, we introduce a previously developed framework for the assessment of the quality of product reviews by using formal argumentation theory and machine learning.

Then, we outline current development on the use of knowledge graphs to better mine arguments, refine argumentation reasoning, and produce more meaningful explanations.

1. Previous Work

Review scores collect users' opinions in a simple and intuitive manner. However, review scores are also easily manipulable, hence they are often accompanied by explanations. A substantial amount of research has been devoted to ascertaining the quality of reviews, to identify the most useful and authentic scores through explanation analysis. In a previosu work of ours [authors, 2021], we advance the state of the art in review quality analysis. We introduce a rating system to identify review arguments and to define an appropriate weighted semantics through formal argumentation theory. We introduce an algorithm to construct a corresponding graph, based on a selection of weighted arguments, their semantic similarity, and the supported ratings. Such an algorithm identifies tokens in corpora of reviews, and then clusters them according to their similarity. Attacks are defined between tokens when they belong to conflicting reviews (i.e., to reviews which scores are different). Such attacks are weighted on the readability level of the reviews (since we use readability as a proxy for quality) and on the importance of the token in the review. Potential arguments are stronger when they come from the most readable reviews, and when their importance in the review is high.

We provide an algorithm to identify the model of such an argumentation graph, maximizing the overall weight of the admitted nodes and edges. We evaluate these contributions on the Amazon review dataset by McAuley et al. [McAuley et al., 2015], by comparing the results of our argumentation assessment with the upvotes received by the reviews. Also, we deepen the evaluation by crowdsourcing a multidimensional assessment of reviews and comparing it to the argumentation assessment. Lastly, we perform a user study to evaluate the explainability of our method. Our method achieves three goals: (1) it identifies reviews that are considered useful when looking at their number of upvotes; (2) when deepening the analysis on the quality of the reviews that are accepted on the basis of argumentation reasoning, we can observe that, in particular, they are considered as comprehensible and truthful; and (3) our user study shows that our approach provides a comprehensible explanation of review quality assessments.

2. Current Developments

One of the limitations of the aforementioned work is that all tokens are treated as potential arguments. Then, they are filtered according to their estimated strength and importance, and only the most important ones are considered when creating the argumentation graph. Yet, by linking identified entities with Wikidata [Vrandečić and Krötzsch, 2014] concepts, we aim at better identify arguments in reviews. By exploiting the existing patterns among such entities, we can then better identify reviews that are semantically relevant to each user and cluster them. This allows us understanding more precisely the rationale behind a given review score. In fact, while reviews refer to a given item, by identifying the arguments in the review text, we can more precisely identify the rationale behind the given score. This rationale refers both to the argument(s) and to the relevant aspect of the item that determined the score.

In turn, this improvement will help refining the resulting explanation for the given assessment of the review, so to be able to present an augmented (and richer) graph to the user who want to inspect the reasoning behind the assessment. In fact, currently, explanations only show relations (attacks) among reviews and their strength, but such extension can help us in better characterizing the resulting graph.

Lastly, while such an additional step requires some preprocessing in order to train an entity linker component, it will contribute to improve the computational speed, as it will allow directly identifying related arguments, instead of having to compute one-to-one semantic distances.

References

Anonymous authors. Anonymous title. In *Anonymous*, page 00. Anonymous, 2021.

Julian J. McAuley, Christopher Targett, Qinfeng Shi, and Anton van den Hengel. Image-based recommendations on styles and substitutes. In *Proceedings of SIGIR*, pages 43–52. ACM, 2015.

Denny Vrandečić and Markus Krötzsch. Wikidata: A free collaborative knowledgebase. *Commun. ACM*, 57(10):78–85, September 2014. ISSN 0001-0782. doi: 10.1145/2629489. URL https://doi.org/10.1145/2629489.