

Review comments on

Natural Language Processing for Analyzing Online Customer Reviews: A Survey, Taxonomy, and Open Research Challenges

Main considerations:

1. Abstract: Some statements are not right (or too strong): For example, the first sentence in the abstract, "In recent years, e-commerce platforms have replaced conventional marketplaces" is not 100% correct. Although E-commerce platforms have gained significant prominence and market share, they have NOT completely replacing conventional marketplaces. it is better to say e.g. ""In recent years, e-commerce platforms have become popular and transformed the way people buy and sell goods."
2. In Survey Methodology, how the keywords are used to search for papers needs to be explained e.g. when you search them you use A or B or C , alternatively A and B and C etc
3. In Taxonomy section, either rename the figure e.g. remove 'NLP in' or add '... NLP in analysing ...' . The reviewer also suggests to redesign the diagram. This taxonomy is very high level and some items overlap with each other in coverage. It's better to change Customer Feedback and Satisfaction to Customer Satisfaction/user experience enhancement(as customer reviews are customer feedbacks). It's better to change Review Analysis and Management to Fake Review Detection and Management as Review Analysis overlaps with Sentiment Analysis. Also sentiment analysis and user satisfaction overlaps to some extent as the purpose/result of sentiment analysis is to understand user satisfaction. It might be better to change Customer satisfaction to Customer Loyalty The reviewer felt that this Taxonomy is more about possible usage/applications of online customer reviews, rather than what NLP can do in this field. Typical applications of NLP in analysing customer reviews include sentiment analysis, recommendation, topic modeling, entity recognition, fake review detection, summarization, trend analysis etc.
4. In Sentiment Analysis and Opinion Mining subsection, it seems the paragraphs are organised in terms of application areas e.g. tourism (tripAdvisor, Amazon, Mobile) etc. It is recommended to give each application area a short subsection title to better guide the reader. If the paragraphs are organised in terms of analysis techniques e.g. deep learning, conventional machine learning etc, please group all deep-learning related papers into one paragraph and give it a subsection title. Do the same to group other technologies e.g. GA, Lexicon-based analysis, Word Clouds etc into different paragraphs.
5. In the subsection Customer Feedback and Satisfaction, some content overlaps with the coverage of sentiment analysis. It is better to redesign the coverage of Customer Feedback and Satisfaction and Sentiment Analysis and Opinion Mining subsections, making each subsection cover unique content.

6. In the Marketing and Brand Management section, it is recommended to add some subsection titles to the main points to guide the readers. This section is very long and contains no subsection titles, no figures, and no tables. The reader can easily get lost in the lengthy text.
7. In the Discussion section, it is recommended to add some subsection titles to the main points to guide the readers, similar to the ones used in the Open Challenges and Future Directions section.
8. In Discussion, some challenges need more explanations. For example, in Aspect-Based Sentiment Analysis, the 'ambiguity, user views' changing subtleties and complexities issues all need more explanations. For example, you may add 'Limitations' before the last paragraph in this Discussion section. You may add a subsection title 'The landscape of analysis techniques' before the paragraph "The changing landscape of techniques". You may also draw a diagram to illustrate the related techniques and their relationships.
9. The Open Challenges and Future Directions is well written.
10. In Conclusion, the main open challenges should be summarised and be consistent with those discussed in the Discussion section e.g. real-time, ethical issues, interpretability etc. The inconsistent summary of challenges mentioned in the Conclusion makes it harder to follow. The authors either re-group the challenges mentioned in the Discussion section into larger groups, or list all the challenges in the Conclusion section.
11. NOTE: *PeerJ* uses the APA ('Name. Year') style with an alphabetized reference list. The format of ALL in-text citations should be changed accordingly.

Minor writing style changes recommended:

- In Introduction, I suggest to Remove "their number and textual nature [3]. Traditional manual review analysis is impossible due to" [Note: keep ref [3]] [you have a similar but better sentence in the Aim paragraph]
- In Survey Methodology, finish the sentence with full stop '.': "... Fake Reviews were searched"
- In Discussion, Instead of "Despite advances like [118] and [148],", list the actual advances then followed by references.
- Expand ABSA (the use of the Acronym ABSA here does not save much space but hinders the readability of the paper).
- Instead of saying "[51] analyses", "[71] mines Twitter", list the author name first, then text, then references. E.g. change "[51] analyses" to "Roccabruna, Azzolin, and Riccardi analyse [51]"
- included in Section 3 => included in Section Taxonomy of ... [as you section titles are not numbered]