

Review Paper PEERJ-86499

# **Feature extraction from consumer reviews using enhanced rules**

July 3, 2023

## **General Review**

### **Positives Issues**

- The paper presents a novel approach to improve the performance of explicit feature extraction from product review documents. The author's objective is to identify and extract features with opinions by utilizing sequential pattern rules on five datasets. Notably, this study introduces a new set of enhanced pattern rules, which complement existing rules from previous studies, addressing gaps that were previously overlooked.
- The proposed approach improves performance by considering previously unaddressed aspects. The experimental results highlight the success of the enhanced pattern rules in accurately extracting features with opinions from review documents. Moreover, the study expands the understanding of feature extraction by demonstrating that features and opinions can extend beyond nouns and adjectives, respectively, as long as they reflect the relevant product or service attributes.

### **Negative Issues**

- One notable concern regarding the paper is the lack of comparison with other methods for evaluating the results. While the proposed approach demonstrates promising performance in explicit feature extraction from product review documents, it would have been beneficial to include a comparison with other widely used methods such as BERT or alternative feature extraction techniques like Tf-Idf. By incorporating such comparisons, the paper could have provided a more comprehensive analysis of the proposed approach's effectiveness in relation to existing state-of-the-art methods. This omission raises questions about the approach's comparative advantages and limits the readers' understanding of its true performance in relation to other established techniques. Consequently,

the inclusion of additional comparisons would greatly enhance the paper’s credibility and allow for a more comprehensive evaluation of its proposed method.

- One notable limitation of the paper is the lack of sufficient emphasis on the new feature extraction approach as the primary contribution throughout the Results and Discussion section. The paper predominantly presents the details and evaluation of the proposed method in the final paragraph and in Table 4. While the development of a new feature extraction approach is undoubtedly a valuable contribution, it would have been beneficial to integrate its significance and novelty more prominently throughout the discussion. By providing a more comprehensive and continuous exploration of the novel approach’s advantages, limitations, and potential implications, the paper could have better showcased its unique contribution to the field of sentiment analysis. The limited focus on the new approach within the Results and Discussion section might lead to a diminished understanding of its true value and could potentially hinder the paper’s overall impact and recognition within the research community.

## Detailed Review

### Abstract

- The abstract lacked at least one very important feature that should be part of the abstract:
  1. A notable drawback of the paper is the absence of a clear mention of the new set of rules for feature extraction in the abstract. While the development of these rules is discussed in the last paragraph of the Results and Discussion section and presented in Table 4, their significance is not adequately conveyed in the abstract. This omission limits the readers’ understanding of the unique contribution made by the research. To enhance the clarity and impact of the paper, it is crucial to explicitly highlight the new rules in the abstract, providing a concise explanation of their significance and enabling readers to gauge the relevance of the study from the outset.

### Proposed Methodology

- One notable gap in the paper is the limited explanation of the dataset used and the lack of evaluation on a broader range of datasets, particularly those with different characteristics or data sources. While the paper may mention the use of datasets, it lacks a comprehensive description of their origin, size, and composition. This information is crucial for readers to understand the representativeness and generalizability of the results.

Additionally, the impact of applying the proposed new approach to datasets beyond the scope of the study is not adequately addressed. The paper primarily focuses on a specific dataset, such as Amazon customer reviews, which may have its own unique characteristics and pre-processing steps. It would be beneficial for the authors to explore the application of their approach to more diverse and original datasets, without extensive pre-processing, to assess its effectiveness and performance in different contexts.

By incorporating a more detailed explanation of the dataset used and conducting evaluations on a wider range of datasets, the authors can enhance the robustness and generalizability of their proposed approach. This would allow readers to better understand the potential limitations and applicability of the method in various real-world scenarios, leading to a more comprehensive and impactful research contribution.

- In the Proposed Methodology section, an important aspect that should have been addressed is the absence of a comparison with other methods to evaluate the results obtained. Although the proposed approach shows promising performance in explicit feature extraction from product review documents, the lack of comparison with widely used methods like BERT or alternative techniques such as Tf-Idf is a notable concern. By including such comparisons, the paper could have provided a more comprehensive analysis of how the proposed approach fares against established state-of-the-art methods. This omission raises doubts about the relative advantages of the approach and limits readers' understanding of its true performance compared to existing techniques. Therefore, incorporating additional comparisons would significantly enhance the paper's credibility and allow for a more thorough evaluation of the proposed method.
- One critical suggestion for improvement is the inclusion of a separate table or a footnote in Table 3 to explain the abbreviations used. Currently, the table contains abbreviations that may be difficult for readers to comprehend without additional context. By providing a dedicated table or a footnote, the authors can ensure clarity and facilitate a better understanding of the abbreviations utilized in Table 3. This will enhance the accessibility of the data presented and enable readers to interpret the results accurately. Including this explanatory information will greatly improve the overall readability and usability of the table.

## Results and Discussion

- A significant drawback in the Results and Discussion section is the lack of practical implications and guidance on how the findings of the paper can be utilized to enhance decision-making in consumer reviews. While the results may provide valuable insights and performance metrics of the

proposed approach, it is essential for the authors to bridge the gap between the research and its real-world application.

To address this gap, the authors should provide a discussion on the practical implications of their findings. They could explore how the extracted features and opinions can be leveraged to make more informed decisions in the context of consumer reviews. This could involve highlighting potential applications, such as sentiment analysis for product improvement, identifying key features that drive customer satisfaction, or enabling personalized recommendations based on customer preferences.

Additionally, the authors could discuss the limitations and challenges of implementing their approach in real-world scenarios. This would allow readers to gain a better understanding of the practical considerations and potential hurdles when utilizing the findings in decision-making processes.

By providing insights into the practical applications of the research findings and addressing the challenges associated with implementation, the Results and Discussion section can offer more value to readers and facilitate the transfer of knowledge from academia to practical decision-making contexts in the field of consumer reviews.