



Peer Review Report


PEER REVIEW REPORT FOR:

Jabeen, S. (2024). Decoding consumer sentiments: Advanced NLP techniques for analyzing smartphone reviews. *Revista de Administração Contemporânea*, 28(4), e240102. <https://doi.org/10.1590/1982-7849rac2024240102.en>

HOW TO CITE THIS PEER REVIEW REPORT:

Jabeen, S., & Antunes, J. J. M. (2024). Peer review report for: Decoding consumer sentiments: Advanced NLP techniques for analyzing smartphone reviews. *Revista de Administração Contemporânea*, 28(4), e240102. *Zenodo*. <https://doi.org/10.5281/zenodo.13829600>

REVIEWERS:

-  Jorge Junio Moreira Antunes (Universidade Federal do Rio de Janeiro, Brazil)
The other reviewer did not authorize the disclosure of his/her report.

ROUND 1

Reviewer 1 report

Reviewer: Jorge Junio Moreira Antunes
Date review returned: July 17, 2024
Recommendation: Minor revision

Comments to the authors

The paper entitled "Decoding Consumer Sentiments: Advanced NLP Techniques for Analyzing Smartphones Reviews" is a very interesting work, with a highly relevant theme for consumers and companies. However, there are some issues to be resolved before this paper is ready for publishing. The authors can find these issues below.

* NLP is a field of artificial intelligence known to require substantial computational time. The authors should detail the computational setup utilized and the computational time necessary for the deep learning fine-tuning technique.

- * The codes should be available to readers in a public repository such as GitHub.
- * There are some formatting problems throughout the text, e.g., in the section "Methodology Applied in This Research."
- * The authors should provide some mathematical detailing about BERT and Sentence Transformers.
- * Figures 1, 2, 3, and 4 should be consolidated into a single figure. Also, the bar plot should have labels indicating the number of reviews per category.
- * All datasets are significantly unbalanced, raising concerns about the metrics used in this work. It is clear from the results that the model struggles to predict neutral responses. I suggest that the authors perform undersampling or/and oversampling on the original sample to mitigate this problem.
- * The authors should report other classification metrics in an appendix.
- * Figure 5 is important and representative, but it is small and difficult to read. This figure should be improved.
- * The section "Analysis of Results" should be rewritten without the topic structure. Additionally, this section should include the results for the topic modeling and should be positioned after the topic modeling section.
- * The topic modeling used KMeans with seven clusters. Qualitatively, the results for the selected topics are quite representative, but a sensitivity analysis for the number of clusters should be reported by the authors. It is not clear why the authors selected seven clusters.
- * Table 3 is represented in the text as an image and should be converted to a proper table.

Additional Questions:

Does the manuscript contain new and significant information to justify publication?: Yes

Does the Abstract (Summary) clearly and accurately describe the content of the article?: Yes

Is the problem significant and concisely stated?: Yes

Are the methods described comprehensively?: Yes

Are the interpretations and conclusions justified by the results?: Yes

Is adequate reference made to other work in the field?: Yes

Is the language acceptable?: Yes

Does the article have data and / or materials that could be made publicly available by the authors?: Yes

Please state any conflict(s) of interest that you have in relation to the review of this paper (state "none" if this is not applicable):

Rating:

Interest: 1. Excellent

Quality: 2. Good

Originality: 1. Excellent

Overall: 2. Good

Reviewer 2 report

Reviewer 2 for this round chose not to disclose his/her review report.

Authors' Responses

I am very thankful for your valuable suggestions. Your comments have greatly helped me improve my research paper. Below is a list of changes I made, which you can easily review as I have used track changes during the editing process.

Response to reviewer 1:

Comment 1: This comment enhanced the transparency of my work, and as per instructions, I added the computational setup and computational time on page 12 under the heading 'Methodology Applied in This Research.'

Comment 2: Author uploaded the datasets and code used in this research to a public GitHub repository, including the title and description of the study.

Comment 3: Along with the document indicating changes, I also submitted a changed document titled 'Revised Paper RAC,' formatted entirely according to the journal's instructions.

Comment 4: In the code used in this research, two mathematical functions were applied, which I explained at the end of page 13.

Comment 5: Four figures have been combined into a single figure, with the number of reviews for each dataset and each category also indicated. This can be found in Figure 1 on page 16.

Comment 7: I did my best to make the text in the diagram 5 which is now diagram 2 as visible as possible.

Comment 8: I rewrote the 'Analysis of Results' section according to the reviewer's instructions. Please refer to pages 19 and 20.

Comment 9: As per the reviewer's instructions, the selection criteria for the seven clusters are explained in detail on page 20. However, reporting the sensitivity analysis is not possible for two reasons: (1) the length of the research paper already exceeds the specified limit, and (2) explaining the other cluster numbers that do not yield quality results would make the findings confusing.

Comment 10: Table 3 has been converted into a Word table. Please refer to page 22.

Comment 11: Conflicts of interest are mentioned. Since there are no conflicts, 'None' is stated.

Response to Reviewer 2:

The authors' responses to the comments of Reviewer 1 for this round were omitted from this report, since the reviewer did not authorize the disclosure of his/her report.

Response to both reviewers:

Both reviewers mentioned the issue of the unbalanced dataset. However, the purpose of this research paper is to analyze consumer thoughts and emotions, so the analysis was conducted on the original datasets without under- or oversampling. Although three different datasets were used to address this imbalance, it reflects a real-world scenario where class imbalance is natural and unavoidable. The model performs adequately even with the imbalance, which is why I prefer to continue with the original datasets. The reasons for using these datasets are explained in detail on pages 14 and 15.

ROUND 2

Reviewer 1 report

Reviewer 1 There was no reviewers' evaluation for this round.

Reviewer 2 report

Reviewer 2 for this round chose not to disclose his/her review report.

Authors' Responses

Revisor: 1

There was no reviewer's evaluation for this round. For this reason, there is no authors' response.

Revisor: 2

The authors' responses to the comments of Reviewer 2 for this round were omitted from this report, since the reviewer did not authorize the disclosure of his/her report.

Disclaimer: The content of the Peer Review Report is the full copy of reviewers and authors' reports. Typing and punctuation errors are not edited. Only comments that violate the journal's ethical policies such as derogatory or defamatory comments will be edited (omitted) from the report. In these cases, it will be clearly stated that parts of the report were edited. Check [RAC's policies](#).