Reproducibility check Study for [PAPER]

First Author
Affiliation / Address line 1
Affiliation / Address line 2
Affiliation / Address line 3
email@domain

Second Author
Affiliation / Address line 1
Affiliation / Address line 2
Affiliation / Address line 3
email@domain

1 Introduction

Should explain the context of the paper. It should contain the following subsections:

1.1 Task / Research Question Description

What is the task the paper is trying to solve or what is the research question they are trying to answer?

1.2 Motivation & Limitations of existing work

Have others tried to solve the same task or answer a similar research question? What are they trying to do differently and why? What were the limitations or shortcomings of prior work?

1.3 Proposed Approach

Briefly describe the core contribution of the paper's proposed approach.

1.4 Likely challenges and mitigations

What is hard about this task / research question? What are your contingency plans if the reproduction turns out to be harder than expected or experiments do not go as planned?

2 Related Work

Include 3-4 sentence descriptions of no less than 4 relevant papers (as applicable). Also mention how your work differs from these. Note that prior work should be properly cited in References, e.g., when you use the BERT model (**?**) you could cite it in this way; if you want to refer to the authors of a certain paper, you should use `citet`, e.g., "**?** proposed the BERT model." See https://acl-org.github.io/ACLPUB/formatting.html for instructions.

3 Experiments

3.1 Datasets

Please list which datasets you used, whether or not you have access them, and whether or not they are publicly available with the same preprocessing and train / dev / tests as the previous work you will be comparing to (if applicable). If you plan to collect your own dataset for evaluating robustness, please describe clearly the data plan (the data source, how you plan to collect it, how you would preprocess it for the task, etc.).

3.2 Implementation

Please provide a link to a repo of your reimplementation (if applicable) and appropriately cite any resources you have used.

3.3 Results

Provide a table comparing your results to the published results.

3.4 Discussion

Discuss any issues you faced. Do your results differ from the published ones? If yes, why do you think that is? Did you do a sensitivity analysis (e.g. multiple runs with different random seeds)?

3.5 Resources

Discuss the cost of your reproduction in terms of resources: computation, time, people, development effort, communication with the authors (if applicable).

3.6 Error Analysis

Perform an error analysis on the model. Include at least 2-3 instances where the model fails. Discuss the error analysis in the paper – what other analyses could the authors have ran? If you were able to perform additional error analyses, report it here.

4 Robustness Study

Explain your approach for Evaluating the Model Robustness. Describe what robustness analysis you have performed. Provide sufficient details about your perturbation data, how you created it, how you used it as a robustness benchmark to evaluate the model, in what metrics, etc.

4.1 Results of Robustness Evaluation

Describe the evaluation results of your reproduced model on the robustness benchmark that you created. Include at least 2 examples where the model performs well and 2 examples where it fails (i.e., being not robust). Provide sufficient analysis and your thoughts on the observations.

4.2 Discussion

Provide any further discussion here, e.g., what challenges did you face when performing the analysis, and what could have been done if you will have more time on this project? Imagine you are writing this report to future researchers; be sure to include "generalizable insights" (e.g., broadly speaking, any tips or advice you'd like to share for researchers trying to analyze the robustness of an NLP model).

5 Workload Clarification

Describe how the team divides the workload in this checkpoint. Note that each team member should contribute roughly the same amount of work to this assignment.

6 Conclusion

Is the paper reproducible?