

Responsible NLP Checklist

Paper title: *SCOUT: Selective Coupling via Optimal Unbalanced Transport for Interpretable Text Classification*

Authors: *Junhao Jia, Hanwen Zheng, Yueyi Wu, Huangwei Chen, Haishuai Wang, Jiajun Bu, Lei Wu*

How to read the checklist symbols:

- the authors responded 'yes'
- the authors responded 'no'
- ^{N/A} the authors indicated that the question does not apply to their work
- the authors did not respond to the checkbox question

For background on the checklist and guidance provided to the authors, see the [Responsible NLP Checklist](#) page at ACL Rolling Review.

A. Questions mandatory for all submissions.

- A1. Did you describe the limitations of your work?

This paper has a Limitations section.

- A2. Did you discuss any potential risks of your work?

Our work primarily focuses on the technical limitations of the proposed method, such as computational overhead and span generation strategies. We did not explicitly discuss potential broader societal risks or negative impacts in this paper.

B. Did you use or create scientific artifacts? (e.g. code, datasets, models)

- B4. Did you discuss the steps taken to check whether the data that was collected/used contains any information that names or uniquely identifies individual people or offensive content, and the steps taken to protect/anonymize it?

Our research utilizes standard public NLP benchmark datasets (e.g., IMDB, Yelp, DBPedia). Because these are widely used, pre-established datasets, we did not perform additional, specific checks for personally identifying information (PII) or offensive content.

- B6. Did you report relevant statistics like the number of examples, details of train/test/dev splits, etc. for the data that you used/created?

We explicitly reported detailed statistics for the datasets, including the total number of examples and the specific train/validation/test splits, in Appendix B (Datasets).

C. Did you run computational experiments?

- C2. Did you discuss the experimental setup, including hyperparameter search and best-found hyperparameter values?

We discussed the experimental setup, including baselines and evaluation metrics, in Section 4.1. We also provided the complete list of hyperparameters used for our SCOUT model in Appendix C (Table 4).

- C3. Did you report descriptive statistics about your results (e.g., error bars around results, summary statistics from sets of experiments), and is it transparent whether you are reporting the max, mean, etc. or just a single run?

We reported the mean and standard deviation for predictive performance and explanation quality

The Responsible NLP Checklist used at ACL Rolling Review is adopted from NAACL 2022, with the addition of ACL 2023 question on AI writing assistance and further refinements based on ARR practice. ACL 2026 used a subset of ARR checklist form.

metrics in Tables 1 and 2. Furthermore, in Appendix D.1, we explicitly stated that each configuration was executed with 3 independent random seeds to ensure statistical reliability.

D. Did you use human annotators (e.g., crowdworkers) or research with human subjects?

D1. Did you report the full text of instructions given to participants, including e.g., screenshots, disclaimers of any risks to participants or annotators, etc.?

No human participants or annotators were involved in this study, hence no instructions were provided or reported.

D2. Did you report information about how you recruited (e.g., crowdsourcing platform, students) and paid participants, and discuss if such payment is adequate given the participants' demographic (e.g., country of residence)?

We did not recruit any human workers, making recruitment and payment details inapplicable.

D3. Did you discuss whether and how consent was obtained from people whose data you're using/curating (e.g., did your instructions explain how the data would be used)?

Our research relies solely on existing, publicly available benchmark datasets; no new human-subject data was collected that would require participant consent.

D4. Was the data collection protocol approved (or determined exempt) by an ethics review board?

This computational research strictly utilizes existing public datasets and involves no human subjects research, exempting it from ethics review board approval.

E. Did you use AI assistants (e.g., ChatGPT, Copilot) in your research, coding, or writing?

E1. If you used AI assistants, did you include information about their use?

We used AI assistants (specifically ChatGPT) to assist in the linguistic polishing of the manuscript and to help generate routine code snippets for data visualization. All AI-generated suggestions were carefully reviewed and verified by the authors to ensure accuracy and scientific integrity.