Dear Reviewers and AE,

We would like to express our appreciation for your insightful comments, constructive suggestions and valuable time for the paper. We sincerely appreciate the encouragement in your comments. We are happy to report that we have addressed the issues that the reviewers raised to the best of our capacity. We have revised the structure and addressed the editorial problems addressed from the reviewers' suggestions.

A color marked-up copy of our manuscript that highlights changes made to the original version is included. Red marks indicate major changes/revisions of paragraphs. Orange marks indicate minor grammatical/structure changes. Blue marks indicate new added paragraphs.

The responses to each comment in detail are as follows.

Response Reviewer #1:

1. In the beginning of Chapter 4 authors should define the D_{KL}.

We added the definition of D_{KL} in Line 141 as "Kullback–Leibler divergence"

2. In the Figure 1 authors can ommit some basic explanations (e.g. linear layer) as it is supposed the reader already understands BERT model which i more complex. Also in the figure not the same words are used.

We added the definition of linear layer in Equation 2 and Line 190. We also updated the figure and equations to have consistent wording.

3. In the results I am not sure what is NVDMb and NVDMo? Also, as accuracy is reported, the percentage of majority class should be mentioned (otherwise reader needs to calculate it from Table 4?).

We have moved the details of experiment settings from supplemental material to the main content. The definition of NVDMb and NVDMo now in the Line 234- 236, which is "1) original NVDM as described in [8] ("NVDMo" in the results); 2) NVDM with BERT representation ("NVDMb" in the results)."

4. Authors will publish annotated data. Could authors also publish automatically annotated data by their algorithm or publish their code. That would be useful for reproducibility and further comparitons.

We will publish the source code and the data used in this paper. Links to the code and data are added in Section 9 Software and Data

5. Authors sometimes start sentence with a formula (e.g. "p(x|z) is the generation"....) or reference (e.g. "[8] introduce a"...). I propose to reformat sentences in a way that they do not start like these.

Thanks for the suggestion, we have revised the manuscript accordingly

Response Reviewer #2:

1. Abstract should be rewritten. Now it seems to be slightly misleading because it is written that this research will develop "computational methods to support research on COVID-19 disinformation debunking and its social impact". In the abstract, it should be emphasized that the main focus of their research is to identify the topic of fake news, not to identify fake news. Furthermore, this abstract is missing an overview of research method and insight into the results.

Thanks for your suggestion, we have rewritten the abstract and parts of the introduction based on your suggestion.

2. In the introductory section authors describe the motivation, main goals and challenges of their research- The scientific contributions are clearly stated as well. My suggestion is to add one paragraph with concise descriptions of all experiments.

A concise description of experiments is added in Line 59-64

3. Section Dataset Structure is written with too many details. The first part of the Section related to Table 1, together with this table can be moved to the Supplementary materials and leaving only data about dataset statistics.

Thanks for your suggestion, we have now significantly reduced the content of Section Dataset Structure. We left Table 1 (with reduced rows) in the main content, because dataset building is one of the main contributions of this work, and the table is a clear way to present the structure of the data.

4. Furthermore, this second Section about the data structure can be a subsection of the section which describes the experiment.

Thanks for your suggestion, we now merged this section with Section 5 "COVID-19 Disinformation Analysis and Discussion"

5. The third Section about disinformation category labelling is also too extensive.

Thanks for your suggestion. We have revised the context and merged this section with Section 2 and renamed the sections as "Dataset and Annotation". Data annotation and the cleaning process are one of the most important steps in our work. This ensures the correctness of the category labeling work and affects the correctness of our experiment and

analysis process described in the later sections. We described this process within one page in the current version of the manuscript.

6. Section about related work needs to be extended with more references that are relevant for this research. I suggest the authors to include more publications that use BERT model in similar NLP tasks.

Thanks for the suggestion, we now enriched the related work section, and highlighted some of the work using BERT model.

Response editor Sanda Martinčić-Ipšić's queries:

1. Please ensure that your manuscript meets PLOS ONE's style requirements, including those for file naming.

We are using the PLOS ONE official Latex template and ensured the manuscript meets PLOS ONE's style requirements

2. In your Data Availability statement, you have not specified where the minimal data set underlying the results described in your manuscript can be found. PLOS defines a study's minimal data set as the underlying data used to reach the conclusions drawn in the manuscript and any additional data required to replicate the reported study findings in their entirety. All PLOS journals require that the minimal data set be made fully available. For more information about our data policy, please see http://journals.plos.org/plosone/s/data-availability.

The full dataset is publicly available at:
www.kaggle.com/dataset/fd97cd3b8f9b10c1600fd7bbb843a5c70d4c934ed83e74085c50b78d3db18443
The source code is publicly available at:
https://github.com/GateNLP/CANTM

We added one more section "Section 9 Software and data" in the updated manuscript reporting the availability of source code and dataset.

3. During our internal evaluation of the manuscript, we found significant text overlap between your submission and the following previously published work, of which you are an author: http://eprints.whiterose.ac.uk/164746/

The work http://eprints.whiterose.ac.uk/164746/ is a copy of our arxiv version of the paper (please note the link to the paper is arxiv). White Rose Research Online is a repository that automatically collects research outputs from University of Sheffield and two other universities. The collection includes an arxiv preprint.

We deeply appreciate your time reviewing our manuscript,

Thank you and best regards.

Xingyi Song