Decision Letter (etrij-2023-0364)

From: etrij@etri.re.kr

To: chrisjihee@etri.re.kr, isj@etri.re.kr, ohwoog@etri.re.kr, nash@jbnu.ac.kr

CC:

Subject: ETRI Journal - Decision on Manuscript ID etrij-2023-0364 [email ref: DL-SW-2-a]

Body: 10-Nov-2023

Dear Mr. Ryu:

Manuscript ID etrij-2023-0364 entitled "Improvement of Korean Morphological Analysis System Through Transformer-based Re-ranking" which you submitted to ETRI Journal, has been reviewed. The comments of the reviewer(s) are included at the bottom of this letter.

The reviewer(s) have recommended some minor revisions to your manuscript. Therefore, I invite you to respond to the reviewer(s)' comments and revise your manuscript.

There are two ways to submit your revised manuscript. You may use the link below to submit your revision online with no need to enter log in details:

*** PLEASE NOTE: This is a two-step process. After clicking on the link, you will be directed to a webpage to confirm. ***

https://mc.manuscriptcentral.com/etrij?URL_MASK=9842646f72f549b8a785a34b063c88e2

Alternatively log into https://mc.manuscriptcentral.com/etrij and enter your Author Center. You will find your manuscript title listed under "Manuscripts with Decisions." Under "Actions," click on "Create a Revision." Your manuscript number has been appended to denote a revision. Please DO NOT upload your revised manuscripts as a new submission.

You will be unable to make your revisions on the originally submitted version of the manuscript. Instead, revise your manuscript using a word processing program and save it on your computer. Please also highlight the changes to your manuscript within the document by using bold or colored text.

Once the revised manuscript is prepared, you can upload it and submit it through your Author Center.

When submitting your revised manuscript, you will be able to respond to the comments made by the reviewer(s) in the space provided. You can use this space to document any changes you make to the original manuscript. In order to expedite the processing of the revised manuscript, please be as specific as possible in your response to the reviewer(s).

IMPORTANT: Your original files are available to you when you upload your revised manuscript. Please delete any redundant files before completing the submission.

Because we are trying to facilitate timely publication of manuscripts submitted to ETRI Journal, your revised manuscript should be uploaded by 4-Dec-2023. If it is not possible for you to submit your revision in a reasonable amount of time, we may have to consider your paper as a new submission. If you feel that you will be unable to submit your revision within the time allowed please contact the Editorial Office (etrij@etri.re.kr) to discuss the possibility of extending the revision time.

Once again, thank you for submitting your manuscript to ETRI Journal and I look forward to receiving your revision.

Sincerely,
Editor, ETRI Journal

Editor Comments to Author:

Section Editor: 1
Comments to the Author:
(There are no comments.)

Editor: 2
Comments to the Author:
(There are no comments.)

Reviewer(s)' Comments to Author:

Reviewer: 1

Comments to the Author
Following are the comments:

1. The Authors have mentioned three corpus in papers i.e. Sejong, UCorpus and Everyone corpus. The

Authors should mention detailed information of Sejong corpus (i.e. the characteristics, figures (in no's) and how they are using the corpus). They should mention information of all three corpus in separate paragraph after Introduction section. (Also mention the purpose of each corpus in paper). They can also mention the count of sentences/words using from all three corpus in tabular form.

2. In Related work section, The Author have mention the various morphological methods in reference form. They can elaborate and give a brief overview of past methods in tabular form.

3. The Authors should add some more information on existing deep learning methods for morphological analysis and how they are different from their own proposed method.

4. The Author should show the step by step process of his own proposed method of morphological analysis Re-ranking BERT method) in block diagram form.

5. The Author should provide some more information on three pre-trained language models, KPF-BERT, ETRI-ELECTRA, and ETRI-RoBERTa, which they are using to fine-tune
their re-ranking model.

6. The Author should define evaluation parameters "eojeol" accuracy and "morpheme F1" score.

7. The Authors should mention the training parameters separately in tabular form along with tool/software information.

8. Please elaborate this line with reasons in section 5.5--"The entire morphological analysis model, including the re-ranking model, is not suitable for real-time processing."

Reviewer: 2

Comments to the Author
The study "Improvement of Korean Morphological Analysis System Through Transformer-based Re-ranking" is interesting but there are several major concerns that need to be addressed.

The abstract is not proper. More theoretical details are given without mentioning the quantified value (results) performance improvement with Transformer-based Re-ranking and how it's achieved? Also, the benefit of achieving a high performance as a conclusive ending in the abstract is missing. Please remove unnecessary details from the abstract.

The figures are not proper. The fonts are either too small or sometimes not distinguished.

The English language and grammar are not good. It is advised to please use small sentences and proofread text from some native English language speaker /writer.

Deep learning is used but the models are not properly explained with diagrams.
The flow of study is not proper. The transition from re-ranking model to experimental results is not up to mark. There has to be a flow between different sections.

The re-ranking performance should explain the need for Lattice + Transformer (En)? and why transformers are used? How performance is increased using Lattice + Transformer (En). What does the cost of performance, mean with the same computational resources? The Lattice + Transformer (En) model needs to be explained more in detail with diagrams.

The results need to have more rigor and experimental findings have to be enough to support the findings and concept.

Reviewer: 3

Comments to the Author
Very interesting paper.

Overall, it's well structured and written. Here are some comments.

1. Table 3 shows that syllable-based system shows better performance in written and spoken. It shows way better performance than dictionary-based(written) even with written data set. Is there any specific reason why the syllable-based shows better performance? It might be better to describe some insights for this.

2. Table 4 shows that Dictionary-based rerank shows better performance in all cases. However, the syllable-based shows a competitive performance in UC+EC (written and spoken). Can you explain or conjecture why that happened?

3. Page 9: 5th line from the left-bottom: These performance improvements.... : it's very difficult to agree this with Table 4 for UC+EC (written and spoken). The differences with syllable-based is marginal at best.

4. Table 5: Sejong data shows poor performance in other approaches and proposed with rerank shows the better performance (from Table 3 and 4). However, Table 5 does not show any huge improvement compared to other approaches. Did you try with other dataset? Not sure how much improvement it will show with others.

Date Sent: 10-Nov-2023