

# Review of: "EOESGC: Predicting miRNA–disease Associations Based on Embedding of Embedding and Simplified Graph Convolutional Network"

Yuping Sun<sup>1</sup>

<sup>1</sup> Guangdong University of Technology

**Potential competing interests:** The author(s) declared that no potential competing interests exist.

Recommendation: major revision

Comments□

This paper proposes a new miRNA-disease associations prediction model (EOESGC) which integrates two existing graph embedding methods, embedding of embedding and simplifying graph convolutional networks, to generate the final miRNA (disease) embedding.

Advantages

①EOESGC creatively adopts EOE to calculate the miRNA embedding and the disease embedding instead of some other random walk based graph embedding methods. For a heterogeneous network, EOE can better capture the topology information of network.

②The employed SGC not only simplifies the structure of network (compared with GCN) and then improves the efficiency of the whole framework, but also maintains the superior predictive performance.

Limitations

①In conclusion, the paper states that the information of the graph can be better learned by using two embedding models. But the paper fails to elaborate the significance of two graph embedding methods, and doesn't give an appropriate ablation experiment to verify the opinion.

Questions about the method:

Regarding the formulas in the section of EOE, the paper does not explain specific values of above weight matrixes. The corresponding variables of miRNA embedding and disease embedding do not appear in the formula. In addition, according to the original EOE method,  $d_i(m_i)$  represents the  $i$ -th row of disease embedding (miRNA embedding).  $W_d, W_m, W_{dm}$  represents the weight of each edges of networks which is equal to the similarity or association. In this situation, we can minimize the loss function to update  $d_i(m_i)$  and further obtain the embedding results. In a word, the description of the EOE may need further modification.

Additional questions:

①The paper lacks many details about the setting of model parameters (e.g. the number of network layers, the number of feature dimensions, etc.)

②In the section of simplifying graph convolutional network, the description of the process of linking SGC

with MLP to output miRNA-disease association score is not clear enough. Besides, more explanation is needed on the generation of the miRNA and disease embedding and the physical meaning of the related parameters, etc.

③The paper has used some evaluation metrics like AUC, AUPR, F1 score to reflect the superior predictive performance of EOESGC in the abstract. But in the process of comparison, the paper only adopts the AUC of predictive results as the evaluation metric.

④Some expressions may be wrong. The title of table 1 is “The ROC curves of EOESGC and baseline”, but the content is the AUC of models. Besides, in the section of experimental approaches and evaluation criteria, "performance" has been misspelled as "ferformance" at the first sentence.

⑤Some sentences are repeated in this paper.

⑥In the last paragraph of introduction, the paper does not well express necessities of integrating embedding of embedding with SGC. In addition, the paper has mentioned that EOESGC "not only guarantees the positional information between related nodes, but also effectively aggregates the information between neighbors". These two contributions are similar, and it may be necessary to differentiate these two contributions.