Dear Prof. JonathanWren and Prof. Alison Hutchins,

**RE**: BIOINF-2023-1850, “Multi-scale topology and position feature learning and relationship-aware graph reasoning for prediction of drug-related microbes”

Thank you very much for the constructive comments and the opportunity to resubmit the revised paper to Bioinformatics.

Enclosed is the revised version of the paper and point-by-point responses to the comments and suggestions.

Major revisions include:

* We added the description about the microbe and drug data which includes the microbe sequences, the drug-drug interactions, and the drug similarities.
* We conducted the ablation experiments about the topology features, the position features, and the relationship types, and we added the corresponding result analysis in the manuscript.
* Several parameters affected the prediction performance, such as the layer numbers of NFF and GFF, the step size of random walking, and the loss balance factor. We listed the experimental results in the supplementary file and added the analysis in the manuscript.
* In terms of case studies, we added the description about the relationship between the interested drug and each candidate microbe in the Supplement Table ST6.

We look forward to hearing from you. We greatly appreciate for your effort and help.

Yours sincerely,

Ping Xuan on behalf of all the authors,

Department of Computer Science, Shantou University, China.

Email: [pxuan@stu.edu.cn](mailto:pxuan@stu.edu.cn)

**Reviewer 1**

Comments：

Identifying the drug-related candidate microbes plays important role for exploring the therapeutic effects of drugs. The manuscript proposed a microbe-drug association prediction model to encode node neighborhood topologies with multiple scales and to perform graph reasoning by propagating multiple types of connections and the diverse node information. Learning the long-distance correlations among the nodes in the whole heterogeneous graph and considering the diverse connection relationships is interesting and novel. The manuscript compared the proposed method and five state-of-the-art methods, and the superior prediction performance of the new method was demonstrated by the comprehensive comparison experiments.

**Response** - We thank the reviewer for the valuable comments. According to the suggestions, we carefully revised the manuscript and provided the following responses. In the revised manuscript, the revisions are highlighted in blue.

1. The manuscript focuses on prediction of drug-related microbes, while the related research work including drug-target interaction prediction, and miRNA-disease association prediction can provide the valuable reference for it. Thus, the introduction part may include the brief description about the related research work.

**Response** **-** Following your suggestion, we have added the following description in the 3th paragraph of the introduction.

“Recently, the computational methods were proposed for predicting the drug-target interactions (Li *et al*. 2022), incRNA-miRNA interactions (Wang *et al*. 2022), miRNA-disease associations (Peng *et al*. 2022), metabolite-disease associations (Gao *et al*. 2023), and incRNA-disease associations (Wang *et al*. 2023).”

**Reference：**

Li F., Zhang Z., Guan J., and Zhou S. Effective drug-target interaction prediction with mutual interaction neural network. Bioinformatics. 2022;38(14):3582-3589.

Wang W., Zhang L., Sun J., et al. Predicting the potential human lncRNA-miRNA interactions based on graph convolution network with conditional random field. Briefings in Bioinformatics. 2022;23(6):bbac463.

Peng W, Che Z., Dai W, et al. Predicting miRNA-disease associations from miRNA-gene-disease heterogeneous network with multi-relational graph convolutional network model. IEEE/ACM Trans Comput Biol Bioinform. 2022 Jul 1. doi: 10.1109/TCBB.2022.3187739.

Gao H., Sun J., Wang Y., et al. Predicting metabolite-disease associations based on auto-encoder and non-negative matrix factorization. Briefings in Bioinformatics. 2023;24(5):bbad259.

Wang S., Hui C., Zhang T., et al. Graph Reasoning Method Based on Affinity Identification and Representation Decoupling for Predicting lncRNA-Disease Associations. Journal of Chemical Information and Modeling. 2023;63(21):6947-6958.

1. The dataset contains the associations among drugs and microbes, the drug similarities, and the attribute features of microbes in section 2.1. It was obtained from the previous prediction work. It will be better if the authors can describe which database the data was originally extracted from.

**Response** - We have added sources of drug interaction and microbe gene sequence data in section 2.1, respectively. The description is listed as follows.

"Drugbank (Knox *et al*. 2023) provides the interactions among the drugs."

"The sequences of microbes were extracted from NCBI database, and then principal component analysis (PCA) was utilized to obtain their important features."

**Reference:**

Knox C., Wilson M., Klinger C.M., et al. DrugBank 6.0: the DrugBank Knowledgebase for 2024. Nucleic Acids Res. 2023 Nov 11. doi: 10.1093/nar/gkad976.

1. The authors utilized the similarities among the drugs to construct the heterogeneous network in section 2.1. Although the drug similarities were obtained from the previous prediction work, descripting the drug similarity calculation in your manuscript is necessary for understanding the network construction.

**Response** - In section 2.1, we modified the calculation description of drug similarity and the description is listed as follows.

"Drugbank (Knox *et al*. 2023) provides the interactions among the drugs. On the basis of the biological hypothesis that the drugs with similar treatment functions are more likely interact with the similar microbes, EGTMDA (Long *et al*. 2020) calculated the Gaussian kernel similarities of drugs based on their interactions. The structural similarity of two drugs was measured based on the common subgraphs within their chemical structures (Hattori *et al*. 2010). The final drug similarities were obtained by the weighted sum of the drug Gaussian kernel similarities and the drug structure similarities."

**Reference:**

Knox C., Wilson M., Klinger C.M., et al. DrugBank 6.0: the DrugBank Knowledgebase for 2024. Nucleic Acids Res. 2023 Nov 11. doi: 10.1093/nar/gkad976.

Long Y., Wu M., Liu Y., Kwoh C.-K., Luo J., and Li X. Ensembling graph attention networks for human microbe-drug association prediction. Bioinformatics, 36(Suppl-2):i779-i786, 12 2020b.

Hattori M., Tanaka N., Kanehisa M., and Goto S. SIMCOMP/SUBCOMP: chemical structure search servers for network analysis. Nucleic Acids Res, 38(Web Server issue): W652 - 656, Jul 2010.

1. The manuscript mentioned "When constructing microbe-microbe (or drug-drug) adjacent matrix, connecting edges are added between the microbe (or drug) nodes with a similarity not less than a threshold ε" in section 2.3. However, there is another symbol "ε" in equation 26 to denote the balance factor. I suggest to use different symbols to represent the different meanings.

**Response** - We have modified "ε" to "β" in section 2.3, and "β" indicates threshold.

1. A symbol △ is used to denote the cross entropy loss calculation in Equation (11). Though it is a common loss function, describing its formalization definition is still necessary.

**Response** - We have modified formula (11) as follows.

 (11)

1. The constructed prediction model exploited the topology and location features of the drug and microbe nodes. The ablation experiments about them should be added to evaluate their effectiveness for the better prediction performance.

**Response** - We have conducted research on topological and positional features, and increased experimental results and analysis in the 1th and 2th paragraphs of section 3.3. They are also listed below.

"We built the prediction model (NGMDA) without multi-scale topological feature learning and the one without position feature learning, respectively. Their AUCs decreased by 0.8% and 0.4%, and their AUPR decreased by 1% and 0.5%, respectively."

"Multi-scale topology and position features indicated the neighbors with multiple ranges and the location information of each node were important for the improved prediction performance."

Table 1. Results of the ablation studies.

|  |  |  |
| --- | --- | --- |
| Networks | Average AUC | Average AUPR |
| NGMDA | 0.944 | 0.728 |
| NGMDA w/o PTL | 0.935 | 0.713 |
| NGMDA w/o ES | 0.938 | 0.682 |
| NGMDA w/o NFF | 0.934 | 0.675 |
| NGMDA w/o GFF | 0.933 | 0.668 |
| NGMDA w/o Topo | 0.936 | 0.718 |
| NGMDA w/o Posi | 0.94 | 0.723 |
| NGMDA w/o Rel | 0.938 | 0.701 |

**Reviewer 2**

Comments:

A new prediction method was proposed to infer the potential microbe candidates for the drugs by learning the various features of the drug and microbe nodes from the neighbor view and from the whole heterogeneous graph view. An adaptive fusion strategy was designed to fuse the position and topology information of neighbor nodes based on heterogeneous graph neural networks. The graph feature fusion was also presented to encode the long-distance correlations among the nodes in the whole heterogeneous graph. The paper is well-organized and easy to follow, and the comparison experiments confirmed they could achieve superior prediction performance. Please according to the following suggestions to further improve its quality.

**Response** - We do appreciate the reviewer for the constructive suggestions and comments. We followed the suggestions closely and revised the manuscript carefully. In the revised manuscript, the revisions are highlighted in blue. The responses are also listed as below.

1. Section 2.3, 1st paragraph. The proposed prediction model established the connection edges among drugs. The weight of each edge was set to the similarity between two drugs. As far as I know, there are interactions among the drugs. What is the benefit of constructing the similarity connection edges instead of the interaction edges?

**Response** - Thank you for your insightful suggestions. When constructing drug-drug network, there are two reasons why we did not use interaction edges and instead used similarity connection edges. On the one hand, we have 5586 drug-drug interaction connections among 1373 drugs, which are relatively sparse. On the other hand, the similarity of drugs integrates the interactions between drugs and the chemical structure information of drugs. We consider that drug similarity provides richer information than interaction, so we construct similarity connecting edges instead of interaction connecting edges between drugs.

1. Section 3.2, 1st paragraph. The proposed NFF contained two layers of topology- and position-sensitive HGNNs and GFF had two layers of relationship-aware graph transformers. How did the authors choose these layer numbers?

**Response** - Considering too many layers can cause the over-smoothing problem and make nodes indistinguishable, the range of layers for NFF and GFF is {1, 2, 3}. We have added relevant descriptions in Section 3.2 "Parameter settings". The description is listed below, and the experimental results are added in the supplementary table ST2.

"For NFF and GFF, we fine-tuned the layer number within a range, {1, 2, 3}, and performed all the combinations of the layer number of NFF and GFF. As shown in the supplementary table ST2, the model gets the best performance when their layer numbers are 2."

Supplementary Table ST2. Prediction performance for the layer number of NFF and GFF.

|  |  |  |  |
| --- | --- | --- | --- |
| NFF | GFF | AUC | AUPR |
| 1 | 1 | 0.945 | 0.715 |
| 1 | 2 | 0.943 | 0.714 |
| 1 | 3 | 0.946 | 0.72 |
| 2 | 1 | 0.943 | 0.708 |
| 2 | 2 | 0.944 | 0.728 |
| 2 | 3 | 0.948 | 0.722 |
| 3 | 1 | 0.942 | 0.705 |
| 3 | 2 | 0.943 | 0.713 |
| 3 | 3 | 0.945 | 0.724 |

1. For the strategy based on random walk, will the steps of random walks affect the prediction performance? Conducting more experiments with different steps and listing the corresponding analysis could let the readers know it.

**Response** - We chose the random walk step size from {1, 2, 4, 8, 16, 32} to evaluate the performance of the model. We have added relevant description in Section 3.2 "Parameter settings". The description is listed as follows, and the corresponding results are listed in the supplementary table ST1.

"To assess the effect of random walk step size on the prediction performance, the step size was selected from {1, 2, 4, 8, 16, 32}. The model achieves the highest AUC (AUC=0.944) and AUPR (AUPR=0.728) when step size is 2 (supplementary table ST2)."

Supplementary Table ST1. Prediction performance for the step size of random walk.

|  |  |  |
| --- | --- | --- |
| Step size of random walk | AUC | AUPR |
| 1 | 0.942 | 0.722 |
| 2 | 0.944 | 0.728 |
| 4 | 0.94 | 0.718 |
| 8 | 0.941 | 0.717 |
| 16 | 0.938 | 0.71 |
| 32 | 0.94 | 0.72 |

1. Page 5 Section 3.2. The authors described the setting of hyper-parameters, but did not specify which strategy was used to obtain the near-optimal parameters.

**Response** - Thank you for your suggestion. We have added a description of the hyper-parameter selection strategy in section 3.2 parameter settings, and they are also listed below.

"The proposed model has some hyper-parameters including the balance factor of loss (ε), the layer numbers of NFF and that of GFF, and the steps of random walking. We firstly establish the variation range for each hyper-parameter, and then select the value which obtains the best performance for the model as the final value of the hyper-parameter."

1. Section 3.4, about the statistical experiment. The paired wilcoxon tests were conducted to demonstrate the prediction performance of the proposed model is significantly higher than the compared methods. Clarifying how the statistical experiment was conducted would be easier for the researchers to follow this sort of work.

**Response** - We have added a description of the implement details of the paired Wilcoxon test in the 2th paragraph of section 3.4, and the details are as follows.

"To observe whether NGMDA’s prediction performance is significantly higher than each compared method, the statistical test was conducted. NGMDA has 1373 AUCs (AUPRs) for the 1373 drugs, and the compared methods also have 1373 AUCs (AUPRs) for these drugs. The paired wilcoxon test was executed on NGMDA’s AUCs (AUPRs) and the AUCs (AUPRs) of the compared methods (Table 3). The results indicated NGMDA obtained the significantly higher prediction performance than all the compared methods."

**Reviewer 3**

Comments:

The paper focused on the microbe-drug association prediction problem, which could facilitate the exploration of the therapeutic effects of drugs. The position and topological features of the microbe and drug nodes were encoded by the position-sensitive and topology-sensitive heterogeneous graph learning. The multiple types of relationships among the microbe and drug nodes were embedded in the graph transformer. There are some comments for the authors.

**Response** - We greatly appreciate for your valuable suggestions. We have carefully revised the manuscript. The responses are given as below. Revisions are highlighted by blue color in the revised manuscript.

1. The features of the microbe nodes and drug nodes were obtained from the previous prediction work. The authors should add the description about these features.

**Response** - Thank you very much for your suggestion. The features of microbes include their gene sequences, and that of drugs include drug interactions, Gaussian kernel drug similarity, and drug structural similarity. We have added the corresponding description in section 2.1 "Dataset".

"The sequences of microbes were extracted from NCBI database, and then principal component analysis (PCA) was utilized to obtain their important features."

"Drugbank (Knox *et al*. 2023) provides the interactions among the drugs. Gaussian kernel drug similarity was calculated using the interaction profiles between drugs. The structural similarity of two drugs was measured based on the common subgraphs within their chemical structures (Hattori *et al*. 2010). The final drug similarities were obtained by the weighted sum of the drug Gaussian kernel similarities and the drug structure similarities."

**Reference:**

Knox C., Wilson M., Klinger C.M., et al. DrugBank 6.0: the DrugBank Knowledgebase for 2024. Nucleic Acids Res. 2023 Nov 11. doi: 10.1093/nar/gkad976. Epub ahead of print.

Hattori M., Tanaka N., Kanehisa M., and Goto S. SIMCOMP/SUBCOMP: chemical structure search servers for network analysis. Nucleic Acids Res, 38(Web Server issue): W652 - 656, Jul 2010.

1. The prediction model was constructed based on the position- and topology-sensitive graph neural networks and relationship-aware graph transformer. What type of techniques do you use to avoid the overfitting problem of the model?

**Response** - We have adapted early stop and dropout strategies to avoid overfitting of the model. Early stop is a cross validation strategy. When the performance on the validation set deteriorates, the training of the model is immediately terminated. Meanwhile, we set dropout to 0.5 in the fully connected neural network of section 2.7 "Representation integration and optimization".

1. The final loss of the prediction model was calculated by the weighted sum of the ES loss and the association prediction loss. It used a hyperparameter to balance the contributions from these two items for the association prediction. Could the authors please add an experiment to illustrate the impact of the hyperparameter on the prediction performance?

**Response** - We used a hyper-parameter to balance the losses of embedding enhancement strategy and association prediction. We have analyzed its impact on predictive performance in Table ST3 and added the following description in section 3.2 "Parameter settings".

"The balance factor ε regulates the importance of the loss of embedding enhancement strategy and that of the association prediction loss, and it was chosen from the range of {0, 0.1, …, 0.5}. The supplementary table ST3 demonstrated the corresponding results and ε was set to 0.2 finally."

Supplementary Table ST3. Prediction performance for the balance factor of loss.

|  |  |  |
| --- | --- | --- |
| ε | AUC | AUPR |
| 0 | 0.938 | 0.682 |
| 0.1 | 0.946 | 0.712 |
| 0.2 | 0.944 | 0.728 |
| 0.3 | 0.943 | 0.712 |
| 0.4 | 0.940 | 0.703 |
| 0.5 | 0.941 | 0.7 |

1. The authors compared their model with 5 prediction models, and they have described the primary techniques which were utilized to construct these models. Could the authors please describe which type of data about the microbes and drugs was exploited by each of the compared methods?

**Response** - NGMDA has been compared with five mainstream microbe-drug association prediction methods, including GCNMDA (Long *et al*. 2020), EGATMDA (Long *et al*. 2020), GACNNMDA (Ma *et al*. 2023), GSAMDA (Tan *et al*. 2022), and SCSMDA (Tian *et al*. 2022). The data description used in the comparison method has been added to section 3.4 "Comparison with other methods" and listed below.

* GCNMDA (Long *et al*. 2020). It established a microbe-drug heterogeneous network and integrated multiple kinds of similarities. These similarities were measured based on the chemical structures of drugs, the Gaussian interaction profiles of drugs (microbes), and the microbe sequences. The prediction model was constructed based on GCN and CRF.
* EGATMDA (Long *et al*. 2020). It constructed a microbe-disease-drug network and then inferred the microbe-drug associations by a hierarchical attention mechanism.
* GACNNMDA (Ma *et al*. 2023). The multiple microbe-drug heterogeneous networks were constructed based on the Gaussian interaction and Hamming interaction profiles of drugs (microbes). The potential microbe-drug associations were identified by the convolutional neural networks.
* GSAMDA (Tan *et al*. 2022). The model calculated the drug (microbe) similarities based on the Gaussian interaction profiles and Hamming interaction profiles of drugs (or microbes), and learned the node features by the graph attention networks and sparse auto-encoder.
* SCSMDA (Tian *et al*. 2022). The model constructed the microbe-drug networks based on the microbe gene sequence information, the Gaussian kernel interaction profiles of drugs (or microbes), and the chemical structures of drugs. It learned the features of the microbe and drug nodes by graph contrastive learning.

**Reference**:

Long Y., Wu M., Kwoh C.-K., Luo J., and Li X. Predicting human microbedrug associations via graph convolutional network with conditional random field. Bioinformatics. 2020;36(19):4918-4927.

Long Y., Wu M., Liu Y., Kwoh C.-K., Luo J., and Li X. Ensembling graph attention networks for human microbe-drug association prediction. Bioinformatics, 36(Suppl-2):i779-i786, 12 2020b.

Ma Q., Tan Y.-Q., and Wang L. GACNNMDA: a computational model for predicting potential human microbe-drug associations based on graph attention network and CNN-based classifier. BMC bioinformatics 24, 35(2023).

Tan Y.-Q., Zou J., Kuang L., Wang X.-Y., Zeng B., Zhang Z., and Wang L. GSAMDA: a computational model for predicting potential microbe-drug associations based on graph attention network and sparse autoencoder. BMC bioinformatics 23, 492 (2022).

Tian Z., Yu Y., Fang H.-C., Xie W.-X., and Guo M.-Z. Predicting microbedrug associations with structure-enhanced contrastive learning and self-paced negative sampling strategy. Briefings in Bioinformatics 2023;bbac634.

1. The model integrated the connection relationship types among the nodes when it learned the node features from the whole heterogeneous graph based on relationship-aware graph transformer. The prediction performance of the model with the relationship types and that of the model without the relationship types should be compared to show the effectiveness of considering the relationship types.

**Response** - Thank you for your insightful comments. We conducted ablation experiments on relationship types and added the following analysis in the 1th paragraph and 2th paragraph of section 3.3, respectively.

"After the relationship type integration was eliminated from the prediction model, its AUC and AUPR decreased by 0.6% and 2.7%. "

"The experimental results also demonstrated the relationship type integration is helpful for improving the prediction performance."

Table 1. Results of ablation experiment.

|  |  |  |
| --- | --- | --- |
| Networks | Average AUC | Average AUPR |
| NGMDA | 0.944 | 0.728 |
| NGMDA w/o PTL | 0.935 | 0.713 |
| NGMDA w/o ES | 0.938 | 0.682 |
| NGMDA w/o NFF | 0.934 | 0.675 |
| NGMDA w/o GFF | 0.933 | 0.668 |
| NGMDA w/o Topo | 0.936 | 0.718 |
| NGMDA w/o Posi | 0.94 | 0.723 |
| NGMDA w/o Rel | 0.938 | 0.701 |

1. Some of the drug-associated candidate microbes were supported by the literature in Table 4. What is the exact reason for determining the associations between microbes and drugs through the literature?

**Response** - The common reasons for the association between microbes and drugs include antibacterial activity, inhibiting biofilm formation, killing effectiveness, and drug resistance. Among them, antibacterial activity refers to the ability of drugs to inhibit or kill microbes. We have added descriptions of the reasons for the association between microbes and Ciprofloxacin, and listed in the supplementary table ST6.

Table ST6. Candidate microbes of the drug Ciprofloxacin.

|  |  |  |
| --- | --- | --- |
| Microbe name | Evidence | Description |
| Candida albicans | PMID: 31471074 | Antibacterial activity |
| Pseudomonas aeruginosa | aBiofilm, MDAD | Inhibiting biofilm formation |
| Staphylococcus aureus | aBiofilm, MDAD | Inhibiting biofilm formation |
| Escherichia coli | aBiofilm, MDAD | Inhibiting biofilm formation |
| Streptococcus mutans | PMID: 30468214 | Killing effectiveness |
| Staphylococcus epidermis | PMID: 10632381 | Antibacterial activity |
| Staphylococcus epidermidis | PMID: 28481197 | Antibacterial activity |
| Salmonella enterica | PMID: 26933017 | Ciprofloxacin susceptibility |
| Vibrio harveyi | PMID: 27247095 | Resistance to ciprofloxacin |
| Enterococcus faecalis | PMID: 27790716 | Resistance to ciprofloxacin |
| Human immunodeficiency virus 1 | PMID: 9566552 | Inhibiting microbial activity |
| Streptococcus sanguis | PMID: 11347679 | Inhibiting microbial activity |
| Stenotrophomonas maltophilia | aBiofilm, MDAD | Inhibiting biofilm formation |
| Listeria monocytogenes | PMID: 28355096 | Resistance to ciprofloxacin |
| Burkholderia cenocepacia | PMID: 27799222 | Resistance to ciprofloxacin |
| Streptococcus pneumoniae | PMID: 26100702 | Resistance to ciprofloxacin |
| Serratia marcescens | PMID: 23751969 | Resistance to ciprofloxacin |

**Reviewer 4**

Comments:

Prediction of the associations between drugs and microbes could provide reliable candidates for the biologists to discover the real associations. The authors presented a new prediction methods and it has several novelties: 1) the designed t-step random walk to reveal the topological neighborhoods with multiple scales. 2) the embedding enhancement strategy to form node embeddings and the more discriminative node distributions. 3) the heterogeneous graph feature fusion to encode the long-distance correlations among the nodes in the entire heterogeneous graph. The authors conducted the comparison experiments and the ablation studies to show the improved prediction performance and the effectiveness of the innovations of their method. The method also provided the potential microbe candidates for each of drugs. I hope the authors revise the manuscript according to my suggestions as follows.

**Response** - Thank you very much for your valuable suggestions. According to your suggestion, we have made changes accordingly, and changes are highlighted in blue in the revised version.

1. Section "Microbe-drug heterogeneous graph". The authors built a microbe-drug heterogeneous graph, and the association matrix was obtained by eliminating the similarity connections whose similarity values are less than a threshold. How do the authors choose the threshold?

**Response** - We added the selection process of similarity threshold β in section 3.2 "Parameter settings", and listed below.

"The drug (microbe) similarity threshold, β, was selected from {0.5, 0.6, 0.7, 0.8, 0.9}, and it was set to 0.9 in our experiment (supplementary table ST4)."

Supplementary Table ST4. Prediction performance for the threshold of similarity.

|  |  |  |
| --- | --- | --- |
| β | AUC | AUPR |
| 0.5 | 0.918 | 0.674 |
| 0.6 | 0.923 | 0.703 |
| 0.7 | 0.939 | 0.691 |
| 0.8 | 0.941 | 0.712 |
| 0.9 | 0.944 | 0.728 |

1. Formula (15) calculated the importance of the position and topological features of vj to vi. Please explain the advance of calculating the importance.

**Response** - Thank you very much for your valuable suggestions. We have added the benefits of calculating topological and positional importance in the 3th paragraph of section 2.5.

"The multiple neighbors of node vi have their various topological neighborhoods and positions, so these neighbors have different importance for vi’s feature learning. Therefore, the importance of each neighbor node for vi was calculated before the vi’s features were updated."

1. Near formula (15). There is a sentence "The parameter balances the contributions between the position and topology representations". What is the exact value of τ in the experiment?

**Response** - The description of the exact values τ was added in section 3.2 and listed below.

"Parameter τ is utilized to balance the importance of the topology and position features, and τ varies from 0 to 1 with a step size of 0.2. The supplementary table ST5 indicates τ value of 0.4 is more favorable for the prediction performance of the model."

Supplementary Table ST5. Prediction performance for the balance factor of position and topology representation.

|  |  |  |
| --- | --- | --- |
| τ | AUC | AUPR |
| 0 | 0.941 | 0.723 |
| 0.2 | 0.943 | 0.721 |
| 0.4 | 0.944 | 0.728 |
| 0.6 | 0.942 | 0.717 |
| 0.8 | 0.939 | 0.718 |
| 1 | 0.937 | 0.714 |

1. Section "Comparison with other methods". The authors performed 5-fold cross-validation to estimate the performance of the proposed prediction method and the compared methods. The average AUC and AUPR on 1373 drugs were regarded as the prediction performance. Please give more details about the average AUC and AUPR calculation.

**Response** - Based on your suggestion, we have added details of the average AUC and AUPR in the 2th paragraph of section 3.1. The description is also listed as follows.

"The TPRs, FPRs, precisions, and recalls of each drug were calculated at different threshold θ, we calculated the average AUCs and average AUPRs of 1373 drugs for each fold. The five fold AUCs (or AUPRs) were averaged as the final AUC (or AUPR)."

1. Section "Comparison with other methods". The proposed method, NGMDA, was compared with the five state-of-the-art microbe-drug association prediction methods, including GCNMDA, EGATMDA, GACNNMDA, GSAMDA, and SCSMDA. Do you use the same data separation for training these prediction models?

**Response** - We have added the following description in the 1th paragraph of section 3.4.

"NGMDA and five compared methods were trained and tested by using the same data separation during five-fold cross validation."

1. Section "Case studies on three drugs". The authors selected top ranked 20 candidate microbes related to 3 drugs for further analysis. Please describe what kind of positive samples and negative samples were used to train the model which were utilized for case studies.

**Response** - Thank you for your suggestion. The description of the selection of positive and negative examples has been added in the 1th paragraph of section 3.5, and listed as follows.

"All the known microbe-drug associations and the randomly selected equal number of unobserved microbe-drug associations were utilized to train the model for case studies."

1. The constructed microbe-drug heterogeneous network includes the microbe similarities that were calculated by the cosine similarity measurement. It is recommended to provide the similarities as a supplementary file. The file can be downloaded by the researchers for the subsequent research.

**Response** - Microbe similarity data has been provided as the supplementary file SF1, and we have added the following explanation in section 2.2 "Calculation of microbe similarity".

"The microbe similarities were listed in the supplementary file SF1."