An effective graph embedding algorithm should preserve as much structural information as possible when transforming a knowledge graph into vector representations. To systematically evaluate the performance of different embedding algorithms, we designed a node classification task: predicting the category of nodes based on their learned vector representations. This task serves as a robust measure of how well the embeddings retain graph structural information, since classification accuracy directly reflects both the discriminative power of the learned representations and their fidelity to the underlying topological structure. comprehensive experimental study using the following methodology.

we divided the FPGDKG into training and test sets in a ratio of 7:3, employed the Random Forest [1] in the downstream task of node classification, and compared them using F1 scores in order to evaluate the performance of six different algorithms. Consequently, we selected the best-performing embedding algorithm and employed cosine similarity to calculate the case similarities relevant to user inputs.

Our comprehensive comparative analysis on the FPGDKG dataset included seven state-of-the-art embedding algorithms: DeepWalk [2], Node2Vec [3], SocioDim [4], RandNE [5], NetMF [6], and GLEE [7]. **Supplementary Table 2** shows that the GLEE excels in the node classification tasks, achieving a superior F1 score of up to 0.83, thus significantly outperforming competing algorithms. GLEE effectively transforms FPGDKG into a knowledge graph vector space while maintaining the basic structure of the data. By embedding user inputs alongside the FPGDKG dataset and using cosine similarity measures, we were able to ascertain cases resembling user-provided information, offering them valuable insights.

Supplementary Table 2. Graph embedding algorithm evaluation.

Model	Micro-F1	Macro-F1	Weighted-F1
DeepWalk [2]	0.59	0.27	0.50
Node2Vec [3]	0.62	0.29	0.53
SocioDim [4]	0.79	0.52	0.77
RandNE [5]	0.47	0.18	0.37
NetMF [6]	0.80	0.52	0.78
GLEE [7]	**0.83**	**0.57**	**0.82**

[1] L. Breiman, "Random Forests," *Machine Learning*, vol. 45, no. 1, pp. 5-32, 2001/10/01 2001, doi: 10.1023/A:1010933404324.

[2] B. Perozzi, R. Al-Rfou, and S. Skiena, "Deepwalk: Online learning of social representations," in *Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining*, 2014, pp. 701-710.

[3] A. Grover and J. Leskovec, "node2vec: Scalable feature learning for networks," in *Proceedings of the 22nd ACM SIGKDD international conference on Knowledge discovery and data mining*, 2016, pp. 855-864.

[4] L. Tang and H. Liu, "Relational learning via latent social dimensions," in *Proceedings of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining*, 2009, pp. 817-826.

[5] Z. Zhang, P. Cui, H. Li, X. Wang, and W. Zhu, "Billion-scale network embedding with iterative random projection," in *2018 IEEE international conference on data mining (ICDM)*, 2018: IEEE, pp. 787-796.

[6] J. Qiu, Y. Dong, H. Ma, J. Li, K. Wang, and J. Tang, "Network embedding as matrix factorization: Unifying deepwalk, line, pte, and node2vec," in *Proceedings of the eleventh ACM international conference on web search and data mining*, 2018, pp. 459-467.

[7] L. Torres, K. S. Chan, and T. Eliassi-Rad, "GLEE: geometric Laplacian eigenmap embedding," *Journal of Complex Networks*, vol. 8, no. 2, p. cnaa007, 2020.