Machine Learning on Publicly Available Semantic Graph Data for Interpretable QSAR Modeling

Joseph D. Romano*, Yun Hao*, and Jason H. Moore†

Institute for Biomedical Informatics, University of Pennsylvania, Philadelphia, Pennsylvania 19104, United States

†*E-mail: jhmoore@upenn.edu*

**These authors contributed equally.*

This is a placeholder for the real abstract. Please write each sentence on its own line in the LaTeX source file for easier tracking in Git.

Keywords: Toxicology; Graph machine learning; QSAR; Artificial Intelligence.

1. Introduction

Testing. Here is a citation.[1]

2. Methods

2.1. *Data sources*

2.2. *Graph neural network models*

2.2.1. *Heterogeneous graph neural network architecture*

2.2.2. *Node prediction model*

2.2.3. *Edge prediction model*

2.3. *Baseline QSAR models*

3. Results

4. Discussion

5. Conclusions

6. Code availability

All source code pertaining to this study is available on GitHub at `https://github.com/JDRomano2/psb-gnn`, and in a 'frozen' version on FigShare at [XXX].

References

1. A. Cherkasov, E. N. Muratov, D. Fourches, A. Varnek, I. I. Baskin, M. Cronin, J. Dearden, P. Gramatica, Y. C. Martin, R. Todeschini *et al.*, Qsar modeling: where have you been? where are you going to?, *Journal of medicinal chemistry* **57**, 4977 (2014).