

Supplementary Materials of the Paper “LibMultiLabel: A Library for Multi-label Classification”

Shao-Syuan Huang^{1,2}, Li-Chung Lin¹, Jie-Jyun Liu¹, Sheng-Wei Chen¹, Si-An Chen¹, Dongli He², Cheng-Yeh Li¹, Yu-Chen Lin¹, Yu-Jen Lin¹, Cheng-Hung Liu¹,
and Chih-Jen Lin^{1,2}

¹National Taiwan University

²Mohamed bin Zayed University of Artificial Intelligence

1 Additional Experimental Results

For the experiments in Table 1, we use two node classification datasets: Flickr (Tang and Liu, 2009) and BlogCatalog (Tang and Liu, 2009). Following the work of Lin et al. (2022), we use the graph representation model DeepWalk (Perozzi et al., 2014) to obtain the graph embeddings of these datasets.

We use LibMultiLabel to run all the experiments, following the settings in the work of Lin et al. (2022). For the thresholding and binary relevance methods, we train L2-regularized logistic regression models as the binary classifier for each class. We use the default hyperparameters for training, except for setting the stopping tolerance of training logistic regression to 0.0001. For the details of the default hyperparameters and the meaning of the tolerance, please refer to the documentation of LibMultiLabel¹ and LIBLINEAR (Fan et al., 2008).

The unrealistic setting also uses the binary relevance method. The difference is that during prediction, we take the true number of labels that have the largest decision values.

The experimental code is available at https://www.csie.ntu.edu.tw/~cjlin/papers/libmultilabel/LibMultiLabel_CIKM_code.zip

¹<https://www.csie.ntu.edu.tw/~cjlin/libmultilabel/>

References

- R.-E. Fan, K.-W. Chang, C.-J. Hsieh, X.-R. Wang, and C.-J. Lin. LIBLINEAR: a library for large linear classification. *Journal of Machine Learning Research*, 9:1871–1874, 2008. URL <http://www.csie.ntu.edu.tw/~cjlin/papers/liblinear.pdf>.
- L.-C. Lin, C.-H. Liu, C.-M. Chen, K.-C. Hsu, I.-F. Wu, M.-F. Tsai, and C.-J. Lin. On the use of unrealistic predictions in hundreds of papers evaluating graph representations. In *Proceedings of the Thirty-Sixth AAAI Conference on Artificial Intelligence (AAAI)*, pages 7479–7487, 2022. URL https://www.csie.ntu.edu.tw/~cjlin/papers/multilabel-embedding/multilabel_embedding.pdf.
- 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 (KDD)*, pages 701–710, 2014. doi: 10.1145/2623330.2623732.
- 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 (KDD)*, pages 817–826, 2009.