**Table 1.**

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | DIP | BioGRID | IntAct | STRING | HPRD | MINT | HIPPIE | BIND | Other | Negatome |
| LightGBM-PPI (Chen, Cheng et al., 2019) |  |  |  |  |  |  |  |  | Tr:7052/7052 T:12804/0 |  |
| GcForest-PPI (Yu, Bin et al., 2021) |  |  |  |  |  |  |  |  | Tr:7052/7052  T:220/0 |  |
| FWRF (Wang et al., 2018) | T:5594/5594 |  |  |  |  |  |  |  | Tr:1458/1458 |  |
| Profppikernel (Hamp and Rost, 2015) |  |  |  |  |  |  |  |  | Tr:4045/4045  T:1190/1190 |  |
| Goktepe et al. (Goktepe, Yunus Emre and Kodaz, Halife, 2018) | T:2399/2399 |  |  |  |  |  |  |  | Tr:3500/3500 |  |
| Hu et al. (Hu, Lun et al., 2022) |  |  |  | Tr:2450/2450 |  |  |  |  | Tr:17434/17434 |  |
| SSWRF (Wei, Zhi-Sen et al., 2016) |  |  |  |  |  |  |  |  | Tr:105/0  T:<200/0 |  |
| PrePPI (Zhang et al., 2012) |  |  |  |  |  |  |  |  | Tr:132559/132559  T:107194/2190000 |  |
| InterPred (Mirabello and Wallner, 2017) |  |  |  |  |  |  |  |  | Tr:111168/  14013121  T:268/268 |  |
| Dong et al. (Dong et al., 2019) |  | T:663/0 | T:663/0 | T:2440/0 |  |  |  |  | Tr:2026/50000 | Tr:1629 |
| Bryant et al. (Bryant et al., 2022) |  |  |  |  |  |  |  |  | Tr:216/3989  T:1461 | Tr:1964  T:1715 |
| Cluspro (Comeau et al., 2004) |  |  |  |  |  |  |  |  | Tr:48/0 |  |
| HADDOCK (de Vries et al., 2010) |  |  |  |  |  |  |  |  | T:15/0 |  |
| L3 (Kovács et al., 2019) |  |  |  |  |  |  |  |  | 1241/244 |  |
| SFCN (Li et al., 2018) |  |  |  |  |  |  |  |  | T:11693/0 |  |
| Sim (Chen et al., 2020) |  |  |  |  |  |  |  |  | T:326072/0 |  |
| Lei et al. (Lei and Ruan, 2013) |  | T:73929/0 |  |  |  |  |  |  | T:48167/0 |  |
| L3N (Yuen and Jansson, 2023) |  | T:31845/0 |  | T:20341/0 |  | T:11583/0 |  |  | T:26800/0 |  |
| PPI-MetaGO (Chen et al., 2019) |  |  |  |  |  |  |  |  | Tr:160124/352498 |  |
| Bandyopadhyay et al. (Bandyopadhyay, Sanghamitra and Mallick, Koushik, 2017) |  |  |  |  |  |  |  | T:10517/10517  T:750/0 | T:39358/3035624 |  |
| DeepPPI (Du et al., 2017) | T:5595/5594 |  |  |  |  |  |  | Tr:17257/48594 | T:19469/5720 |  |
| Struct2Graph (Baranwal et al., 2022) |  |  | T:12676/0 | T:446548/0 |  |  |  |  | Tr:4698/5036 |  |
| TAGPPI (Song et al., 2022) |  |  |  |  |  |  |  |  | Tr:38536/38536 |  |
| SGPPI (Huang et al., 2023) |  |  |  |  |  |  |  |  | Tr:2748/2748(?) |  |
| Sun et al. (Sun et al., 2017) |  |  |  |  |  |  |  |  | Tr:33052/32816  T:3493/3507 |  |
| D-SCRIPT (Sledzieski et al., 2021) |  |  |  |  |  |  |  |  | Tr:38345/383456  T:9587/95864 |  |
| DPPI (Hashemifar et al., 2018) |  |  |  |  |  |  |  |  | Tr:14292/142920  T:1588/15880 |  |
| PPI-RCC (Poot Velez et al., 2020) |  |  |  |  |  |  |  |  | Tr:171142/692  T:4819/692  T:2680/122 | 1506 |

The numbers in each cell correspond with the number of positive/negative instances used to train/test the model. Tr stands for training datasets, T for test datasets. The numerator accounts for the positive instances and the denominator the negative instances.

**References**

Bandyopadhyay, Sanghamitra, Mallick, Koushik, 2017. A New Feature Vector Based on Gene Ontology Terms for Protein-Protein Interaction Prediction. IEEE/ACM Trans. Comput. Biol. Bioinform. 14, 762–770. https://doi.org/10.1109/TCBB.2016.2555304

Baranwal, M., Magner, A., Saldinger, J., Turali-Emre, E.S., Elvati, P., Kozarekar, S., VanEpps, J.S., Kotov, N.A., Violi, A., Hero, A.O., 2022. Struct2Graph: a graph attention network for structure based predictions of protein-protein interactions. BMC Bioinformatics 23, 370. https://doi.org/10.1186/s12859-022-04910-9

Bryant, P., Pozzati, G., Elofsson, A., 2022. Improved prediction of protein-protein interactions using AlphaFold2. Nat. Commun. 13, 1265. https://doi.org/10.1038/s41467-022-28865-w

Chen, Cheng, Zhang, Qingmei, Ma, Qin, Yu, Bin, 2019. LightGBM-PPI: Predicting protein-protein interactions through LightGBM with multi-information fusion. Chemom. Intell. Lab. Syst. 191, 54–64. https://doi.org/10.1016/j.chemolab.2019.06.003

Chen, K.-H., Wang, T.-F., Hu, Y.-J., 2019. Protein-protein interaction prediction using a hybrid feature representation and a stacked generalization scheme. BMC Bioinformatics 20, 308. https://doi.org/10.1186/s12859-019-2907-1

Chen, Y., Wang, W., Liu, J., Feng, J., Gong, X., 2020. Protein Interface Complementarity and Gene Duplication Improve Link Prediction of Protein-Protein Interaction Network. Front. Genet. 11, 291. https://doi.org/10.3389/fgene.2020.00291

Comeau, S.R., Gatchell, D.W., Vajda, S., Camacho, C.J., 2004. ClusPro: an automated docking and discrimination method for the prediction of protein complexes. Bioinforma. Oxf. Engl. 20, 45–50. https://doi.org/10.1093/bioinformatics/btg371

de Vries, S.J., van Dijk, M., Bonvin, A.M.J.J., 2010. The HADDOCK web server for data-driven biomolecular docking. Nat. Protoc. 5, 883–897. https://doi.org/10.1038/nprot.2010.32

Dong, S., Lau, V., Song, R., Ierullo, M., Esteban, E., Wu, Y., Sivieng, T., Nahal, H., Gaudinier, A., Pasha, A., Oughtred, R., Dolinski, K., Tyers, M., Brady, S.M., Grene, R., Usadel, B., Provart, N.J., 2019. Proteome-wide, Structure-Based Prediction of Protein-Protein Interactions/New Molecular Interactions Viewer. Plant Physiol. 179, 1893–1907. https://doi.org/10.1104/pp.18.01216

Du, X., Sun, S., Hu, C., Yao, Y., Yan, Y., Zhang, Y., 2017. DeepPPI: Boosting Prediction of Protein-Protein Interactions with Deep Neural Networks. J. Chem. Inf. Model. 57, 1499–1510. https://doi.org/10.1021/acs.jcim.7b00028

Goktepe, Yunus Emre, Kodaz, Halife, 2018. Prediction of Protein-Protein Interactions Using An Effective Sequence Based Combined Method. Neurocomputing 303, 68–74. https://doi.org/10.1016/j.neucom.2018.03.062

Hamp, T., Rost, B., 2015. Evolutionary profiles improve protein-protein interaction prediction from sequence. Bioinforma. Oxf. Engl. 31, 1945–1950. https://doi.org/10.1093/bioinformatics/btv077

Hashemifar, S., Neyshabur, B., Khan, A.A., Xu, J., 2018. Predicting protein-protein interactions through sequence-based deep learning. Bioinforma. Oxf. Engl. 34, i802–i810. https://doi.org/10.1093/bioinformatics/bty573

Hu, Lun, Yang, Shicheng, Luo, Xin, Yuan, Huaqiang, Sedraoui, Khaled, Zhou, MengChu, 2022. A Distributed Framework for Large-scale Protein-protein Interaction Data Analysis and Prediction Using MapReduce. IEEECAA J. Autom. Sin. 9, 160–172. https://doi.org/10.1109/JAS.2021.1004198.

Huang, Y., Wuchty, S., Zhou, Y., Zhang, Z., 2023. SGPPI: structure-aware prediction of protein-protein interactions in rigorous conditions with graph convolutional network. Brief. Bioinform. 24, bbad020. https://doi.org/10.1093/bib/bbad020

Kovács, I.A., Luck, K., Spirohn, K., Wang, Y., Pollis, C., Schlabach, S., Bian, W., Kim, D.-K., Kishore, N., Hao, T., Calderwood, M.A., Vidal, M., Barabási, A.-L., 2019. Network-based prediction of protein interactions. Nat. Commun. 10, 1240. https://doi.org/10.1038/s41467-019-09177-y

Lei, C., Ruan, J., 2013. A novel link prediction algorithm for reconstructing protein-protein interaction networks by topological similarity. Bioinforma. Oxf. Engl. 29, 355–364. https://doi.org/10.1093/bioinformatics/bts688

Li, S., Huang, J., Zhang, Z., Liu, J., Huang, T., Chen, H., 2018. Similarity-based future common neighbors model for link prediction in complex networks. Sci. Rep. 8, 17014. https://doi.org/10.1038/s41598-018-35423-2

Mirabello, C., Wallner, B., 2017. InterPred: A pipeline to identify and model protein-protein interactions. Proteins 85, 1159–1170. https://doi.org/10.1002/prot.25280

Poot Velez, A.H., Fontove, F., Del Rio, G., 2020. Protein-Protein Interactions Efficiently Modeled by Residue Cluster Classes. Int. J. Mol. Sci. 21, 4787. https://doi.org/10.3390/ijms21134787

Sledzieski, S., Singh, R., Cowen, L., Berger, B., 2021. D-SCRIPT translates genome to phenome with sequence-based, structure-aware, genome-scale predictions of protein-protein interactions. Cell Syst. 12, 969-982.e6. https://doi.org/10.1016/j.cels.2021.08.010

Song, B., Luo, Xiaoyan, Luo, Xiaoli, Liu, Y., Niu, Z., Zeng, X., 2022. Learning spatial structures of proteins improves protein-protein interaction prediction. Brief. Bioinform. 23, bbab558. https://doi.org/10.1093/bib/bbab558

Sun, T., Zhou, B., Lai, L., Pei, J., 2017. Sequence-based prediction of protein protein interaction using a deep-learning algorithm. BMC Bioinformatics 18, 277. https://doi.org/10.1186/s12859-017-1700-2

Wang, L., You, Z.-H., Xia, S.-X., Chen, X., Yan, X., Zhou, Y., Liu, F., 2018. An improved efficient rotation forest algorithm to predict the interactions among proteins. Soft Comput. 22, 3373–3381. https://doi.org/10.1007/s00500-017-2582-y

Wei, Zhi-Sen, Han, Ke, Yang, Jing-Yu, Shen, Hong-Bin, Yu, Dong-Jun, 2016. Protein–protein interaction sites prediction by ensembling SVM and sample-weighted random forests. Neurocomputing 193, 201–212. https://doi.org/10.1016/j.neucom.2016.02.022

Yu, Bin, Chen, Cheng, Wang, Xiolin, Yu, Zhaomin, Ma, Anjun, Liu, Bingqiang, 2021. Prediction of protein–protein interactions based on elastic net and deep forest. Expert Syst. Appl. 176, 114876. https://doi.org/10.1016/j.eswa.2021.114876

Yuen, H.Y., Jansson, J., 2023. Normalized L3-based link prediction in protein-protein interaction networks. BMC Bioinformatics 24, 59. https://doi.org/10.1186/s12859-023-05178-3

Zhang, Q.C., Petrey, D., Deng, L., Qiang, L., Shi, Y., Thu, C.A., Bisikirska, B., Lefebvre, C., Accili, D., Hunter, T., Maniatis, T., Califano, A., Honig, B., 2012. Structure-based prediction of protein-protein interactions on a genome-wide scale. Nature 490, 556–560. https://doi.org/10.1038/nature11503