Supplementary Information for

**DA-HGL: A Domain-Augmented Heterogeneous Graph Learning Framework for Protein Function Prediction**

**Sai Hu1, Wei Zhang2,3, Bihai Zhao2,3\***

1 School of Mathematics, Changsha University, Changsha, Hunan 410022, China

2 School of Computer Science and Engineering, Changsha University, Changsha, Hunan 410022, China

3 Hunan Provincial Key Laboratory of Industrial Internet Technology and Security, Changsha University, Changsha, Hunan 410022, China

Supplementary Table 1. Statistical Significance, Effect Size, and Mean Differences Between DA-HGL and Competing Methods

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Categories | Comparison | Yeast | | | Human | | |
| *p* | Cohen's d | Mean Difference | *p* | Cohen's d | Mean Difference |
| BP | zhang | 9.21E-12 | 19.2890 | 0.2828 | 1.92E-10 | 9.7182 | 0.196 |
| DSCP | 5.32E-11 | 13.1083 | 0.2107 | 5.51E-09 | 7.1421 | 0.1394 |
| NC | 2.11E-08 | 8.9402 | 0.1538 | 5.54E-08 | 2.9406 | 0.0678 |
| PHN | 0.00015 | 2.2103 | 0.0446 | 3.22E-06 | 1.8948 | 0.0512 |
| exp2GO | 1.03E-07 | 7.1649 | 0.1187 | 1.86E-08 | 3.5635 | 0.0895 |
| GrAPFI | 1.01E-12 | 19.8161 | 0.334 | 1.77E-10 | 15.0673 | 0.261 |
| Domain-PFP | 2.03E-06 | 3.7304 | 0.0713 | 6.67E-08 | 3.4354 | 0.0736 |
| MF | zhang | 1.07E-10 | 10.0761 | 0.3224 | 1.96E-11 | 18.8522 | 0.2557 |
| DSCP | 4.53E-08 | 5.996 | 0.168 | 2.85E-08 | 9.7289 | 0.1412 |
| NC | 1.73E-08 | 8.0614 | 0.243 | 2.49E-10 | 13.9959 | 0.205 |
| PHN | 5.17E-09 | 4.0151 | 0.1496 | 8.11E-10 | 11.5313 | 0.1772 |
| exp2GO | 2.70E-08 | 6.8209 | 0.1872 | 7.40E-09 | 10.6146 | 0.1698 |
| GrAPFI | 2.62E-08 | 5.1109 | 0.1704 | 1.45E-08 | 7.1241 | 0.0916 |
| Domain-PFP | 5.56E-05 | 3.348 | 0.1606 | 0.006638 | 1.237 | 0.019 |
| CC | zhang | 2.36E-12 | 8.8769 | 0.2867 | 7.91E-10 | 9.7187 | 0.2196 |
| DSCP | 4.43E-11 | 7.8639 | 0.2216 | 1.17E-07 | 5.1924 | 0.1308 |
| NC | 3.37E-08 | 3.9865 | 0.1413 | 1.08E-07 | 3.9749 | 0.0888 |
| PHN | 0.000129 | 1.4936 | 0.0458 | 5.20E-06 | 2.2966 | 0.0464 |
| exp2GO | 9.44E-10 | 5.2897 | 0.1461 | 2.38E-06 | 3.3057 | 0.0882 |
| GrAPFI | 1.81E-11 | 10.3745 | 0.3599 | 2.23E-10 | 10.5305 | 0.2288 |
| Domain-PFP | 1.09E-07 | 3.6622 | 0.0966 | 4.07E-05 | 1.934 | 0.0519 |

Supplementary Table 2. Wilcoxon Signed-Rank Test Results and FDR-Adjusted p-values for Performance Comparisons

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Categories | Comparison | Yeast | | | Human | | |
| Wilcoxon | t-test  (FDR) | Wilcoxon  (FDR) | Wilcoxon | t-test  (FDR) | Wilcoxon  (FDR) |
| BP | zhang | 0.001953 | 3.22E-11 | 0.001953 | 0.001953 | 6.74E-10 | 0.001953 |
| DSCP | 0.001953 | 1.24E-10 | 0.001953 | 0.001953 | 1.29E-08 | 0.001953 |
| NC | 0.001953 | 3.69E-08 | 0.001953 | 0.001953 | 7.75E-08 | 0.001953 |
| PHN | 0.001953 | 0.000154 | 0.001953 | 0.001953 | 3.22E-06 | 0.001953 |
| exp2GO | 0.001953 | 1.44E-07 | 0.001953 | 0.001953 | 3.25E-08 | 0.001953 |
| GrAPFI | 0.001953 | 7.07E-12 | 0.001953 | 0.001953 | 6.74E-10 | 0.001953 |
| Domain-PFP | 0.001953 | 2.37E-06 | 0.001953 | 0.001953 | 7.78E-08 | 0.001953 |
| MF | zhang | 0.001953 | 7.52E-10 | 0.001953 | 0.001953 | 1.38E-10 | 0.002279 |
| DSCP | 0.001953 | 5.28E-08 | 0.001953 | 0.001953 | 3.32E-08 | 0.002279 |
| NC | 0.001953 | 3.78E-08 | 0.001953 | 0.001953 | 8.73E-10 | 0.002279 |
| PHN | 0.001953 | 1.81E-08 | 0.001953 | 0.001953 | 1.89E-09 | 0.002279 |
| exp2GO | 0.001953 | 3.78E-08 | 0.001953 | 0.001953 | 1.30E-08 | 0.002279 |
| GrAPFI | 0.001953 | 3.78E-08 | 0.001953 | 0.001953 | 2.03E-08 | 0.002279 |
| Domain-PFP | 0.001953 | 5.56E-05 | 0.001953 | 0.013672 | 0.006638 | 0.013672 |
| CC | zhang | 0.001953 | 1.65E-11 | 0.001953 | 0.001953 | 2.77E-09 | 0.001953 |
| DSCP | 0.001953 | 1.03E-10 | 0.001953 | 0.001953 | 2.05E-07 | 0.001953 |
| NC | 0.001953 | 4.72E-08 | 0.001953 | 0.001953 | 2.05E-07 | 0.001953 |
| PHN | 0.001953 | 0.000129 | 0.001953 | 0.001953 | 6.07E-06 | 0.001953 |
| exp2GO | 0.001953 | 1.65E-09 | 0.001953 | 0.001953 | 3.34E-06 | 0.001953 |
| GrAPFI | 0.001953 | 6.33E-11 | 0.001953 | 0.001953 | 1.56E-09 | 0.001953 |
| Domain-PFP | 0.001953 | 1.27E-07 | 0.001953 | 0.001953 | 4.07E-05 | 0.001953 |

Supplementary Table 3. Fmax scores (mean ± standard deviation) for disease-associated protein function prediction.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Methods | Diabetes | | | Parkinson | | |
| BP | MF | CC | BP | MF | CC |
| **DA-HGL** | **0.377±0.309** | **0.507±0.291** | **0.415±0.334** | **0.583±0.375** | **0.415±0.355** | **0.684±0.347** |
| zhang | 0.211±0.232 | 0.293±0.279 | 0.240±0.278 | 0.450±0.377 | 0.169±0.142 | 0.593±0.319 |
| DSCP | 0.314±0.276 | 0.479±0.313 | 0.393±0.289 | 0.498±0.357 | 0.388±0.365 | 0.612±0.304 |
| NC | 0.229±0.244 | 0.210±0.207 | 0.322±0.336 | 0.554±0.395 | 0.288±0.338 | 0.617±0.353 |
| PHN | 0.292±0.281 | 0.382±0.308 | 0.354±0.307 | 0.583±0.402 | 0.409±0.376 | 0.670±0.350 |
| exp2GO | 0.260±0.270 | 0.265±0.284 | 0.327±0.331 | 0.544±0.402 | 0.265±0.333 | 0.634±0.363 |
| GrAPFI | 0.185±0.250 | 0.374±0.337 | 0.159±0.244 | 0.088±0.187 | 0.243±0.354 | 0.070±0.125 |
| Domain-PFP | 0.293±0.230 | 0.401±0.263 | 0.365±0.312 | 0.600±0.343 | 0.546±0.301 | 0.625±0.229 |

Supplementary Table 4. Statistical significance and effect sizes from T-tests for disease-associated protein predictions.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Categories | Comparison | Diabetes | | | Parkinson | | |
| *p* | Cohen's d | Mean Difference | *p* | Cohen's d | Mean Difference |
| BP | zhang | 2.34E-05 | 0.6122 | 0.1660 | 0.0006 | 0.3583 | 0.1333 |
| DSCP | 0.0193 | 0.2148 | 0.0625 | 0.0136 | 0.2362 | 0.0855 |
| NC | 2.63E-06 | 0.5353 | 0.1479 | 0.2652 | 0.0767 | 0.0292 |
| PHN | 0.0137 | 0.2906 | 0.0851 | 0.9954 | 0.0005 | 0.0002 |
| exp2GO | 0.0002 | 0.4060 | 0.1170 | 0.2184 | 0.1019 | 0.0392 |
| GrAPFI | 1.61E-06 | 0.6878 | 0.1920 | 2.04E-09 | 1.6909 | 0.4956 |
| Domain-PFP | 0.0047 | 0.3120 | 0.0844 | 0.706 | -0.0473 | -0.0168 |
| MF | zhang | 0.0002 | 0.7583 | 0.2138 | 0.0066 | 0.9338 | 0.2453 |
| DSCP | 0.3869 | 0.0920 | 0.0275 | 0.7589 | 0.0774 | 0.0271 |
| NC | 8.47E-09 | 1.1863 | 0.2965 | 0.2779 | 0.3763 | 0.1267 |
| PHN | 0.0087 | 0.4192 | 0.1242 | 0.9589 | 0.0155 | 0.0055 |
| exp2GO | 7.54E-07 | 0.8492 | 0.2415 | 0.2392 | 0.4485 | 0.1499 |
| GrAPFI | 0.0013 | 0.4254 | 0.1325 | 0.0058 | 0.4975 | 0.1714 |
| Domain-PFP | 0.0465 | 0.3830 | 0.1052 | 0.2452 | -0.4108 | -0.1314 |
| CC | zhang | 9.41E-05 | 0.5791 | 0.1759 | 0.0008 | 0.2773 | 0.0913 |
| DSCP | 0.5373 | 0.0742 | 0.0229 | 0.0170 | 0.2241 | 0.0722 |
| NC | 0.0087 | 0.2818 | 0.0933 | 0.0179 | 0.1936 | 0.0669 |
| PHN | 0.1773 | 0.1936 | 0.0614 | 0.6280 | 0.0409 | 0.0141 |
| exp2GO | 0.0883 | 0.2699 | 0.0887 | 0.1333 | 0.1439 | 0.0504 |
| GrAPFI | 7.79E-06 | 0.8869 | 0.2566 | 3.1E-12 | 2.3897 | 0.6148 |
| Domain-PFP | 0.4330 | 0.1592 | 0.0509 | 0.3478 | 0.2039 | 0.0592 |

Supplementary Table 5. Wilcoxon signed-rank test results and FDR-adjusted p-values for disease-associated protein predictions.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Categories | Comparison | Diabetes | | | Parkinson | | |
| Wilcoxon | t-test  (FDR) | Wilcoxon  (FDR) | Wilcoxon | t-test  (FDR) | Wilcoxon  (FDR) |
| BP | zhang | 5.47E-05 | 1.63E-05 | 3.81E-05 | 0.00226 | 0.00022 | 0.00078 |
| DSCP | 0.01928 | 0.0219 | 0.0219 | 0.03162 | 0.01043 | 0.02434 |
| NC | 9.19E-06 | 7.57E-08 | 2.65E-07 | 0.37126 | 0.04866 | 0.08516 |
| PHN | 0.01595 | 0.00021 | 0.0003 | 0.99545 | 0.16581 | 0.19345 |
| exp2GO | 0.00034 | 0.00005 | 0.00009 | 0.37126 | 0.16151 | 0.19345 |
| GrAPFI | 9.19E-06 | 2.77E-08 | 1.94E-07 | 1.43E-08 | 1.74E-07 | 1.22E-06 |
| Domain-PFP | 0.00656 | 0.00521 | 0.00608 | 0.82370 | 0.96718 | 0.96718 |
| MF | zhang | 0.00041 | 0.00029 | 0.00068 | 0.02323 | 0.01477 | 0.0517 |
| DSCP | 0.38688 | 0.21497 | 0.21497 | 0.88538 | 0.67912 | 0.67912 |
| NC | 5.93E-08 | 1.33E-07 | 9.28E-07 | 0.38905 | 0.03313 | 0.0773 |
| PHN | 0.01213 | 0.00237 | 0.00332 | 0.95894 | 0.27737 | 0.3236 |
| exp2GO | 2.64E-06 | 1.84E-06 | 6.45E-06 | 0.38905 | 0.05508 | 0.09638 |
| GrAPFI | 0.0022 | 0.00063 | 0.0011 | 0.02323 | 0.00098 | 0.00687 |
| Domain-PFP | 0.0543 | 0.10666 | 0.12443 | 0.38905 | 0.12393 | 0.1735 |
| CC | zhang | 0.00033 | 0.00011 | 0.00038 | 0.00277 | 0.00179 | 0.00627 |
| DSCP | 0.53734 | 0.47271 | 0.47271 | 0.03124 | 0.01231 | 0.02872 |
| NC | 0.02028 | 0.00179 | 0.00417 | 0.03124 | 0.01754 | 0.03069 |
| PHN | 0.24821 | 0.11843 | 0.20725 | 0.62805 | 0.02763 | 0.03869 |
| exp2GO | 0.1546 | 0.15136 | 0.21191 | 0.18666 | 0.13522 | 0.13522 |
| GrAPFI | 5.45E-05 | 9.29E-06 | 6.50E-05 | 2.17E-11 | 1.58E-07 | 1.11E-06 |
| Domain-PFP | 0.50515 | 0.42357 | 0.47271 | 0.4058 | 0.12843 | 0.13522 |