|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | Noise 0.25 | | Noise 0.5 | | Noise 1 | | Noise 2 | | Noise 4 | |
|  | ACC | AU-ROC | ACC | AU-ROC | ACC | AU-ROC | ACC | AU-ROC | ACC | AU-ROC |
| Jaccard | **0.800** | **0.866** | 0.776 | 0.877 | **0.832** | **0.905** | **0.821** | **0.893** | **0.803** | **0.883** |
| Lin | 0.760 | 0.847 | **0.824** | **0.888** | 0.827 | 0.897 | 0.813 | 0.893 | 0.803 | 0.881 |
| Resnik | 0.755 | 0.846 | 0.824 | 0.887 | 0.805 | 0.884 | 0.810 | 0.892 | 0.799 | 0.882 |

Fig.: Experiment 6. Phenotype kernel matrices were calculated with different semantic similarity algorithms (Jaccard, Lin, Resnik). Different levels of phenotype noise were simulated by sampling without replacement (for Noise < 1) or by adding random HPO terms (for Noise > 1).