**Table 1.** Cross validation on the 1-step reaction sequence-likeness (enzymatic-reaction-likeness).

| Chemical | Diff-common L1SVM | | Diff-only L1SVM | | Baseline | | Random | |
|---|---|---|---|---|---|---|---|---|
| fingerprints/descriptors | AUC | AUPR | AUC | AUPR | AUC | AUPR | AUC | AUPR |
| CDK extended | 0.6917 | 0.0603 | 0.6742 | 0.0468 | 0.6161 | 0.0289 | 0.5000 | 0.0199 |
| CDK graph-only | 0.7161 | 0.0602 | 0.6960 | 0.0432 | 0.6255 | 0.0228 | 0.5000 | 0.0199 |
| CDK hybridization | 0.7003 | 0.0489 | 0.6761 | 0.0403 | 0.6023 | 0.0272 | 0.5000 | 0.0199 |
| E-state | 0.6088 | 0.0371 | 0.5567 | 0.0297 | 0.5085 | 0.0169 | 0.5000 | 0.0199 |
| KlekotaRoth | 0.6031 | 0.0347 | 0.5683 | 0.0304 | 0.5203 | 0.0042 | 0.5000 | 0.0199 |
| MACCS | 0.6837 | 0.0489 | 0.6582 | 0.0342 | 0.5914 | 0.0189 | 0.5000 | 0.0199 |
| PubChem | 0.7170 | 0.0531 | 0.7026 | 0.0422 | 0.6752 | 0.0307 | 0.5000 | 0.0199 |
| NS-descriptor | 0.8858 | 0.2134 | 0.8429 | 0.0968 | 0.6566 | 0.0446 | 0.5000 | 0.0199 |
| KCF-S descriptor | 0.9659 | 0.3943 | 0.9610 | 0.2801 | 0.6945 | 0.0755 | 0.5000 | 0.0199 |

**Table 2.** Cross validation on the 2-step reaction sequence-likeness prediction (with 1 intermediate compound).

| Chemical | Diff-common L1SVM | | Diff-only L1SVM | | Baseline | | Random | |
|---|---|---|---|---|---|---|---|---|
| fingerprints/descriptors | AUC | AUPR | AUC | AUPR | AUC | AUPR | AUC | AUPR |
| CDK extended | 0.7747 | 0.1730 | 0.7178 | 0.1352 | 0.4815 | 0.0576 | 0.5000 | 0.0665 |
| CDK graph-only | 0.7429 | 0.1711 | 0.7510 | 0.1626 | 0.4969 | 0.0567 | 0.5000 | 0.0665 |
| CDK hybridization | 0.7871 | 0.1879 | 0.6971 | 0.1271 | 0.4631 | 0.0554 | 0.5000 | 0.0665 |
| E-state | 0.6318 | 0.0900 | 0.5389 | 0.0563 | 0.3941 | 0.0407 | 0.5000 | 0.0665 |
| KlekotaRoth | 0.6462 | 0.1080 | 0.5309 | 0.0807 | 0.4949 | 0.0181 | 0.5000 | 0.0665 |
| MACCS | 0.7474 | 0.1418 | 0.6634 | 0.1152 | 0.4465 | 0.0502 | 0.5000 | 0.0665 |
| PubChem | 0.7674 | 0.1589 | 0.7270 | 0.1357 | 0.5732 | 0.0710 | 0.5000 | 0.0665 |
| NS-descriptor | 0.8898 | 0.2937 | 0.8673 | 0.2651 | 0.6187 | 0.0937 | 0.5000 | 0.0665 |
| KCF-S descriptor | 0.9411 | 0.4493 | 0.9419 | 0.4473 | 0.6621 | 0.0635 | 0.5000 | 0.0665 |

**Table 3.** Cross validation on the 3-step reaction sequence-likeness prediction (with 2 intermediate compounds).

| Chemical | Diff-common L1SVM | | Diff-only L1SVM | | Baseline | | Random | |
|---|---|---|---|---|---|---|---|---|
| fingerprints/descriptors | AUC | AUPR | AUC | AUPR | AUC | AUPR | AUC | AUPR |
| CDK extended | 0.8103 | 0.1436 | 0.7542 | 0.0959 | 0.5474 | 0.0368 | 0.5000 | 0.0367 |
| CDK graph-only | 0.8363 | 0.1508 | 0.7785 | 0.0949 | 0.5621 | 0.0369 | 0.5000 | 0.0367 |
| CDK hybridization | 0.8293 | 0.1558 | 0.7631 | 0.0973 | 0.5226 | 0.0345 | 0.5000 | 0.0367 |
| E-state | 0.6192 | 0.0515 | 0.5450 | 0.0495 | 0.4125 | 0.0225 | 0.5000 | 0.0367 |
| KlekotaRoth | 0.6395 | 0.0605 | 0.5463 | 0.0539 | 0.5165 | 0.0116 | 0.5000 | 0.0367 |
| MACCS | 0.7608 | 0.0986 | 0.6770 | 0.0713 | 0.4959 | 0.0309 | 0.5000 | 0.0367 |
| PubChem | 0.8097 | 0.1239 | 0.7656 | 0.0910 | 0.6365 | 0.0489 | 0.5000 | 0.0367 |
| NS-descriptor | 0.9284 | 0.2638 | 0.9028 | 0.1989 | 0.7069 | 0.0807 | 0.5000 | 0.0367 |
| KCF-S descriptor | 0.9624 | 0.4232 | 0.9585 | 0.4094 | 0.6621 | 0.0635 | 0.5000 | 0.0367 |

**Table 4.** Cross validation on the 4-step reaction sequence-likeness prediction (with 3 intermediate compounds).

| Chemical | Diff-common L1SVM | | Diff-only L1SVM | | Baseline | | Random | |
|---|---|---|---|---|---|---|---|---|
| fingerprints/descriptors | AUC | AUPR | AUC | AUPR | AUC | AUPR | AUC | AUPR |
| CDK extended | 0.8577 | 0.1062 | 0.7867 | 0.0649 | 0.5863 | 0.0172 | 0.5000 | 0.0156 |
| CDK graph-only | 0.8611 | 0.0988 | 0.7957 | 0.0482 | 0.5961 | 0.0175 | 0.5000 | 0.0156 |
| CDK hybridization | 0.8581 | 0.1083 | 0.7869 | 0.0583 | 0.5490 | 0.0157 | 0.5000 | 0.0156 |
| E-state | 0.6053 | 0.0235 | 0.5452 | 0.0230 | 0.4336 | 0.0097 | 0.5000 | 0.0156 |
| KlekotaRoth | 0.6270 | 0.0273 | 0.5519 | 0.0248 | 0.5104 | 0.0057 | 0.5000 | 0.0156 |
| MACCS | 0.7663 | 0.0582 | 0.6898 | 0.0351 | 0.5187 | 0.0141 | 0.5000 | 0.0156 |
| PubChem | 0.8536 | 0.0818 | 0.7962 | 0.0481 | 0.6590 | 0.0234 | 0.5000 | 0.0156 |
| NS-descriptor | 0.9535 | 0.2058 | 0.9304 | 0.1341 | 0.7521 | 0.0436 | 0.5000 | 0.0156 |
| KCF-S descriptor | 0.9772 | 0.3283 | 0.9837 | 0.3202 | 0.7039 | 0.0315 | 0.5000 | 0.0156 |