Table S1. The details of the training data of RBP-24 dataset. For the intendent testing set, all RBPs have 500 positives and 500 negatives, which are directly downloaded from original study GraphProt.

|  |  |  |  |
| --- | --- | --- | --- |
| RBP | [Abbreviations](http://www.abbreviations.com/acronyms/NICKNAMES) | # of training positives | # of training negatives |
| ALKBH5 PAR-CLIP | ALKBH5 | 1197 | 1213 |
| C17ORF85 PAR-CLIP | C17ORF85 | 1860 | 1849 |
| C22ORF28 PAR-CLIP | C22ORF28 | 9369 | 9136 |
| CAPRIN1 PAR-CLIP | CAPRIN1 | 8140 | 7901 |
| Ago2 HITS-CLIP | Ago2 | 48095 | 44251 |
| ELAVL1 HITS-CLIP | ELAVL1H | 8595 | 8436 |
| SFRS1 HITS-CLIP | SFRS1 | 19438 | 17195 |
| HNRNPC iCLIP | HNRNPC | 21472 | 19794 |
| TDP43 iCLIP | TDP43 | 92031 | 75079 |
| TIA1 iCLIP | TIA1 | 18049 | 16135 |
| TIAL1 iCLIP | TIAL1 | 42332 | 36652 |
| Ago1-4 PAR-CLIP | Ago1-4 | 36902 | 31310 |
| ELAVL1 PAR-CLIP (B) | ELAVL1B | 9464 | 9283 |
| ELAVL1 PAR-CLIP (A) | ELAVL1A | 27275 | 23974 |
| EWSR1 PAR-CLIP | EWSR1 | 16292 | 14720 |
| FUS PAR-CLIP | FUS | 34581 | 31480 |
| ELAVL1 PAR-CLIP (C) | ELAVL1C | 125202 | 113686 |
| IGF2BP1-3 PAR-CLIP | IGF2BP1-3 | 8539 | 6838 |
| MOV10 PAR-CLIP | MOV10 | 13793 | 12987 |
| PUM2 PAR-CLIP | PUM2 | 9116 | 8227 |
| QKI PAR-CLIP | QKI | 10276 | 9142 |
| TAF15 PAR-CLIP | TAF15 | 7298 | 6606 |
| PTB HITS-CLIP | PTB | 44574 | 43700 |
| ZC3H7B PAR-CLIP | ZC3H7B | 20962 | 20018 |

Table S2. The details of the training data of RBP-47 dataset. In each dataset, the number of positives equals to the number of negatives, and only RBPs with at least 4000 training samples are kept. The 80% of data size of individual RBPs are used as training set, and the remaining 20% are used as testing set. The results of RNAcommender are taken from the original paper, ‘1’ means no prediction results for this RBP in the paper.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| RBPs | Data size | RNAcommender | Pse-SVM | iDeepE-L | iDeepE-G | iDeepE |
| MOV10 | 11168 | - | 0.86 | 0.85 | 0.86 | **0.88** |
| SRRM4 | 40610 | - | 0.67 | 0.71 | 0.78 | **0.79** |
| ZFP36 | 6918 | - | 0.86 | 0.85 | 0.86 | **0.88** |
| CAPRIN1 | 11220 | - | 0.75 | 0.74 | 0.75 | **0.78** |
| HNRNPU | 7578 | - | 0.74 | 0.72 | 0.75 | **0.77** |
| C22ORF28 | 9502 | - | 0.69 | 0.65 | 0.70 | **0.72** |
| RC3H1 | 4720 | - | 0.66 | 0.66 | 0.69 | **0.70** |
| LARP4B | 30132 | - | 0.76 | 0.75 | 0.80 | **0.81** |
| PABPC1 | 3716 | - | 0.60 | 0.66 | 0.65 | **0.68** |
| ATXN2 | 16680 | - | 0.79 | 0.76 | 0.81 | **0.83** |
| U2AF2 | 3524 | - | 0.66 | 0.62 | 0.63 | 0.65 |
| ADAR1 | 3536 | 0.70 | 0.73 | 0.79 | 0.80 | **0.82** |
| STAU1 | 5632 | 0.48 | 0.61 | 0.67 | 0.71 | **0.73** |
| HNRNPF | 7206 | 0.79 | 0.76 | 0.76 | 0.75 | **0.80** |
| TAF15  | 7140 | **0.90** | 0.81 | 0.78 | 0.79 | 0.81 |
| FXR2  | 16736 | **0.87** | 0.78 | 0.78 | 0.80 | 0.82 |
| LIN28B  | 24102 | **0.86** | 0.75 | 0.74 | 0.79 | 0.81 |
| HNRNPD  | 25258 | 0.61 | 0.66 | 0.68 | 0.72 | **0.74** |
| FMR1\_iso1  | 27078 | **0.86** | 0.77 | 0.77 | 0.82 | 0.83 |
| FMR1\_iso7  | 29166 | **0.77** | 0.68 | 0.70 | 0.75 | **0.77** |
| TIA1  | 31126 | **0.89** | 0.83 | 0.82 | 0.86 | 0.87 |
| TIAL1  | 40986 | **0.88** | 0.82 | 0.82 | 0.86 | 0.87 |
| AGO1  | 51144 | 0.82 | 0.87 | 0.86 | 0.90 | **0.91** |
| EWSR1  | 9944 | **0.91** | 0.82 | 0.83 | 0.81 | 0.85 |
| MSI1  | 17282 | 0.80 | 0.76 | 0.75 | 0.79 | **0.81** |
| LIN28A  | 20514 | **0.88** | 0.77 | 0.75 | 0.80 | 0.83 |
| EIF4A3  | 34816 | 0.65 | 0.70 | 0.75 | 0.80 | **0.81** |
| RBM47  | 29846 | 0.79 | 0.85 | 0.83 | 0.87 | **0.88** |
| FUS  | 12124 | **0.87** | 0.79 | 0.79 | 0.81 | 0.83 |
| AGO2  | 33218 | **0.85** | 0.77 | 0.76 | 0.81 | 0.82 |
| ELAVL1  | 41144 | 0.72 | 0.75 | 0.77 | 0.81 | **0.82** |
| DDX21  | 15080 | 0.67 | 0.65 | 0.66 | 0.70 | **0.71** |
| ZC3H7B  | 19904 | 0.82 | 0.83 | 0.83 | 0.85 | **0.87** |
| PCBP2  | 6000 | 0.78 | 0.85 | 0.76 | 0.83 | **0.84** |
| FXR1  | 5374 | **0.93** | 0.86 | 0.83 | 0.86 | 0.88 |
| YTHDF1  | 10638 | 0.81 | 0.77 | 0.75 | 0.79 | **0.82** |
| HNRNPC  | 7680 | **0.85** | 0.83 | 0.81 | 0.82 | **0.85** |
| RBM10  | 15950 | 0.72 | 0.67 | 0.68 | 0.73 | **0.74** |
| HNRNPH1  | 7774 | 0.72 | 0.82 | 0.80 | 0.83 | **0.86** |
| RBPMS  | 7530 | 0.86 | 0.81 | 0.82 | 0.84 | **0.87** |
| IGF2BP2  | 14824 | **0.81** | 0.75 | 0.67 | 0.68 | 0.70 |
| IGF2BP3  | 18288 | **0.75** | 0.72 | 0.65 | 0.67 | 0.67 |
| IGF2BP1  | 15024 | **0.79** | 0.71 | 0.63 | 0.67 | 0.67 |
| HNRNPA2B1  | 3522 | **0.82** | 0.76 | 0.77 | 0.78 | 0.80 |
| PUM2  | 5730 | 0.76 | 0.84 | **0.81** | 0.79 | **0.81** |
| YTHDF2  | 3374 | **0.85** | 0.81 | 0.82 | 0.78 | 0.82 |
| PUM1  | 6062 | 0.53 | 0.64 | 0.65 | 0.70 | **0.70** |

Table S3. Parameter selection for iDeepE on RBP-24 datasets

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| learning rate | weight\_decay | dropout | window size | shift size | AUC |
| 0.001 | 0.01 | 0.5 | 81 | 50 | 0.91592992 |
| 0.001 | 0.01 | 0.5 | 101 | 20 | 0.91742379 |
| 0.0001 | 0.0001 | 0.25 | 81 | 50 | 0.917788104 |
| 0.001 | 0.01 | 0.5 | 101 | 30 | 0.917798801 |
| 0.0001 | 0.01 | 0.5 | 81 | 50 | 0.918257797 |
| 0.001 | 0.01 | 0.5 | 151 | 50 | 0.918260741 |
| 0.0001 | 0.01 | 0.5 | 81 | 20 | 0.918301934 |
| 0.001 | 0.01 | 0.5 | 81 | 30 | 0.918343363 |
| 0.001 | 0.01 | 0.5 | 151 | 30 | 0.918362605 |
| 0.0001 | 0.001 | 0.5 | 81 | 30 | 0.918453411 |
| 0.001 | 0.01 | 0.5 | 81 | 20 | 0.918794747 |
| 0.0001 | 0.0001 | 0.5 | 81 | 50 | 0.919371399 |
| 0.0001 | 0.001 | 0.25 | 81 | 50 | 0.919461984 |
| 0.001 | 0.001 | 0.5 | 101 | 20 | 0.920100627 |
| 0.001 | 0.001 | 0.5 | 81 | 50 | 0.920146301 |
| 0.001 | 0.01 | 0.25 | 81 | 50 | 0.920167661 |
| 0.001 | 0.01 | 0.5 | 151 | 20 | 0.920356212 |
| 0.001 | 0.01 | 0.5 | 101 | 50 | 0.92043817 |
| 0.0001 | 0.0001 | 0.5 | 101 | 20 | 0.920507976 |
| 0.001 | 0.01 | 0.25 | 101 | 30 | 0.92053586 |
| 0.0001 | 0.01 | 0.5 | 101 | 20 | 0.92060177 |
| 0.0001 | 0.01 | 0.25 | 81 | 50 | 0.920624225 |
| 0.0001 | 0.001 | 0.25 | 81 | 20 | 0.920662676 |
| 0.0001 | 0.01 | 0.5 | 101 | 30 | 0.920794807 |
| 0.0001 | 0.0001 | 0.5 | 101 | 30 | 0.920810359 |
| 0.0001 | 0.001 | 0.25 | 101 | 30 | 0.920917155 |
| 0.0001 | 0.001 | 0.5 | 81 | 50 | 0.921054178 |
| 0.0001 | 0.01 | 0.25 | 81 | 30 | 0.921168072 |
| 0.001 | 0.01 | 0.25 | 81 | 30 | 0.921423714 |
| 0.0001 | 0.0001 | 0.25 | 101 | 30 | 0.921468826 |
| 0.001 | 0.01 | 0.25 | 81 | 20 | 0.921604201 |
| 0.0001 | 0.0001 | 0.5 | 81 | 20 | 0.921635037 |
| 0.0001 | 0.01 | 0.5 | 81 | 30 | 0.921646669 |
| 0.0001 | 0.0001 | 0.25 | 101 | 20 | 0.92174411 |
| 0.0001 | 0.0001 | 0.5 | 81 | 30 | 0.921754502 |
| 0.0001 | 0.0001 | 0.25 | 81 | 20 | 0.921871274 |
| 0.001 | 0.01 | 0.25 | 101 | 20 | 0.921953088 |
| 0.0001 | 0.001 | 0.5 | 101 | 20 | 0.92210856 |
| 0.001 | 0.0001 | 0.25 | 81 | 50 | 0.922293548 |
| 0.0001 | 0.001 | 0.5 | 81 | 20 | 0.922329077 |
| 0.001 | 0.0001 | 0.5 | 81 | 50 | 0.922353255 |
| 0.0001 | 0.001 | 0.5 | 101 | 30 | 0.922359981 |
| 0.0001 | 0.01 | 0.25 | 81 | 20 | 0.922368137 |
| 0.001 | 0.001 | 0.5 | 101 | 50 | 0.922440475 |
| 0.0001 | 0.001 | 0.25 | 81 | 30 | 0.922559852 |
| 0.0001 | 0.0001 | 0.25 | 81 | 30 | 0.922571535 |
| 0.001 | 0.001 | 0.5 | 81 | 30 | 0.922615792 |
| 0.0001 | 0.001 | 0.25 | 101 | 20 | 0.922657423 |
| 0.001 | 0.01 | 0.25 | 151 | 50 | 0.922876543 |
| 0.0001 | 0.01 | 0.25 | 101 | 30 | 0.923081873 |
| 0.0001 | 0.01 | 0.25 | 101 | 20 | 0.923222209 |
| 0.0001 | 0.0001 | 0.5 | 151 | 50 | 0.923367046 |
| 0.001 | 0.0001 | 0.25 | 101 | 20 | 0.923449456 |
| 0.0001 | 0.001 | 0.5 | 151 | 20 | 0.923501198 |
| 0.001 | 0.001 | 0.5 | 101 | 30 | 0.92353602 |
| 0.0001 | 0.0001 | 0.5 | 151 | 20 | 0.923549028 |
| 0.001 | 0.001 | 0.5 | 81 | 20 | 0.923597983 |
| 0.0001 | 0.001 | 0.5 | 151 | 50 | 0.923639938 |
| 0.001 | 0.001 | 0.25 | 101 | 30 | 0.92384987 |
| 0.001 | 0.001 | 0.25 | 81 | 30 | 0.923866715 |
| 0.001 | 0.0001 | 0.25 | 101 | 30 | 0.924121694 |
| 0.001 | 0.001 | 0.25 | 81 | 50 | 0.924372076 |
| 0.001 | 0.0001 | 0.5 | 101 | 20 | 0.924411164 |
| 0.001 | 0.01 | 0.25 | 101 | 50 | 0.924533676 |
| 0.0001 | 0.01 | 0.5 | 151 | 30 | 0.924553316 |
| 0.0001 | 0.0001 | 0.5 | 101 | 50 | 0.924670855 |
| 0.001 | 0.001 | 0.5 | 151 | 50 | 0.924725111 |
| 0.0001 | 0.01 | 0.5 | 151 | 50 | 0.924732016 |
| 0.0001 | 0.01 | 0.5 | 101 | 50 | 0.924805297 |
| 0.001 | 0.001 | 0.25 | 101 | 20 | 0.92503193 |
| 0.0001 | 0.0001 | 0.25 | 101 | 50 | 0.925175611 |
| 0.001 | 0.0001 | 0.5 | 81 | 30 | 0.925235101 |
| 0.001 | 0.001 | 0.5 | 151 | 30 | 0.925276632 |
| 0.0001 | 0.001 | 0.25 | 101 | 50 | 0.925684529 |
| 0.001 | 0.0001 | 0.5 | 101 | 50 | 0.925769295 |
| 0.0001 | 0.001 | 0.5 | 101 | 50 | 0.925883198 |
| 0.001 | 0.0001 | 0.5 | 101 | 30 | 0.925888765 |
| 0.001 | 0.001 | 0.25 | 81 | 20 | 0.926076107 |
| 0.001 | 0.001 | 0.25 | 151 | 50 | 0.926219882 |
| 0.001 | 0.0001 | 0.5 | 81 | 20 | 0.926340309 |
| 0.001 | 0.0001 | 0.5 | 151 | 50 | 0.926480347 |
| 0.0001 | 0.0001 | 0.5 | 151 | 30 | 0.92663563 |
| 0.0001 | 0.001 | 0.25 | 151 | 50 | 0.926663003 |
| 0.0001 | 0.01 | 0.25 | 101 | 50 | 0.926908516 |
| 0.001 | 0.0001 | 0.25 | 81 | 30 | 0.926952054 |
| 0.0001 | 0.01 | 0.25 | 151 | 50 | 0.927153123 |
| 0.001 | 0.0001 | 0.25 | 101 | 50 | 0.927323808 |
| 0.0001 | 0.001 | 0.5 | 151 | 30 | 0.927410826 |
| 0.0001 | 0.0001 | 0.25 | 151 | 30 | 0.927430691 |
| 0.0001 | 0.0001 | 0.25 | 151 | 50 | 0.927520312 |
| 0.0001 | 0.001 | 0.25 | 151 | 20 | 0.92775604 |
| 0.001 | 0.0001 | 0.25 | 81 | 20 | 0.927757847 |
| 0.0001 | 0.01 | 0.5 | 151 | 20 | 0.927878595 |
| 0.0001 | 0.01 | 0.25 | 151 | 30 | 0.927995002 |
| 0.001 | 0.001 | 0.5 | 151 | 20 | 0.928175655 |
| 0.0001 | 0.001 | 0.25 | 151 | 30 | 0.928391935 |
| 0.001 | 0.0001 | 0.25 | 151 | 50 | 0.928544134 |
| 0.001 | 0.0001 | 0.5 | 151 | 30 | 0.928565375 |
| 0.0001 | 0.0001 | 0.25 | 151 | 20 | 0.928863056 |
| 0.001 | 0.01 | 0.25 | 151 | 20 | 0.929564805 |
| 0.0001 | 0.01 | 0.25 | 151 | 20 | 0.929954593 |
| 0.001 | 0.01 | 0.25 | 151 | 30 | 0.930326625 |
| 0.001 | 0.001 | 0.25 | 151 | 20 | 0.930441024 |
| 0.001 | 0.001 | 0.25 | 151 | 30 | 0.93122751 |
| 0.001 | 0.001 | 0.25 | 101 | 50 | 0.931302047 |
| 0.001 | 0.0001 | 0.5 | 151 | 20 | 0.932314743 |
| 0.001 | 0.0001 | 0.25 | 151 | 20 | 0.933717728 |
| 0.001 | 0.0001 | 0.25 | 151 | 30 | 0.934292645 |

 |

Table S4. The impact of different sequence similarity cutoff 80%, 70%, 60% and 50% on the performance of iDeepE.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | 80% |  | 70.00% | 60.00% | 50.00% |
| RBP | iDeepE | iDeepE-G | iDeepE-L | iDeepE | iDeepE-G | iDeepE-L | iDeepE | iDeepE-G | iDeepE-L | iDeepE | iDeepE-G | iDeepE-L |
| CAPRIN1 | 0.685 | 0.657 | 0.661 | 0.688 | 0.641 | 0.671 | 0.620 | 0.607 | 0.596 | 0.568 | 0.603 | 0.546 |
| RBPMS | 0.784 | 0.760 | 0.745 | 0.761 | 0.739 | 0.724 | 0.709 | 0.674 | 0.704 | 0.572 | 0.586 | 0.538 |
| YTHDF1 | 0.717 | 0.706 | 0.684 | 0.713 | 0.703 | 0.678 | 0.668 | 0.661 | 0.639 | 0.641 | 0.597 | 0.619 |
| ZC3H7B | 0.787 | 0.791 | 0.731 | 0.797 | 0.784 | 0.771 | 0.759 | 0.750 | 0.723 | 0.604 | 0.617 | 0.574 |
| FXR2 | 0.744 | 0.744 | 0.681 | 0.735 | 0.723 | 0.673 | 0.713 | 0.675 | 0.700 | 0.522 | 0.586 | 0.489 |
| HNRNPC | 0.762 | 0.734 | 0.729 | 0.756 | 0.740 | 0.719 | 0.719 | 0.722 | 0.669 | 0.617 | 0.607 | 0.581 |
| PABPC1 | 0.601 | 0.602 | 0.554 | 0.607 | 0.562 | 0.611 | 0.524 | 0.524 | 0.523 | 0.239 | 0.380 | 0.325 |
| RC3H1 | 0.582 | 0.537 | 0.595 | 0.573 | 0.607 | 0.533 | 0.515 | 0.501 | 0.543 | 0.436 | 0.381 | 0.568 |
| DDX21 | 0.629 | 0.619 | 0.601 | 0.619 | 0.609 | 0.588 | 0.615 | 0.603 | 0.589 | 0.570 | 0.552 | 0.528 |
| YTHDF2 | 0.738 | 0.734 | 0.706 | 0.726 | 0.721 | 0.693 | 0.727 | 0.682 | 0.719 | 0.615 | 0.623 | 0.560 |
| HNRNPU | 0.691 | 0.674 | 0.650 | 0.688 | 0.677 | 0.627 | 0.654 | 0.632 | 0.624 | 0.527 | 0.549 | 0.504 |
| TAF15 | 0.728 | 0.739 | 0.672 | 0.710 | 0.702 | 0.658 | 0.698 | 0.678 | 0.669 | 0.670 | 0.686 | 0.596 |
| IGF2BP1 | 0.704 | 0.691 | 0.682 | 0.699 | 0.680 | 0.672 | 0.627 | 0.606 | 0.620 | 0.439 | 0.456 | 0.496 |
| TIA1 | 0.805 | 0.793 | 0.763 | 0.786 | 0.779 | 0.725 | 0.748 | 0.726 | 0.727 | 0.590 | 0.608 | 0.575 |
| TIAL1 | 0.765 | 0.757 | 0.727 | 0.763 | 0.752 | 0.722 | 0.755 | 0.744 | 0.722 | 0.718 | 0.683 | 0.665 |
| SRRM4 | 0.614 | 0.614 | 0.568 | 0.615 | 0.610 | 0.572 | 0.581 | 0.572 | 0.555 | 0.572 | 0.473 | 0.599 |
| ELAVL1 | 0.747 | 0.740 | 0.704 | 0.752 | 0.741 | 0.713 | 0.764 | 0.747 | 0.721 | 0.595 | 0.560 | 0.597 |
| LARP4B | 0.690 | 0.673 | 0.640 | 0.680 | 0.651 | 0.658 | 0.639 | 0.608 | 0.609 | 0.552 | 0.490 | 0.576 |
| FMR1\_iso7 | 0.642 | 0.622 | 0.613 | 0.635 | 0.610 | 0.607 | 0.633 | 0.610 | 0.618 | 0.604 | 0.578 | 0.584 |
| MSI1 | 0.728 | 0.711 | 0.687 | 0.702 | 0.697 | 0.664 | 0.675 | 0.657 | 0.643 | 0.530 | 0.531 | 0.531 |
| RBM47 | 0.846 | 0.840 | 0.831 | 0.835 | 0.826 | 0.807 | 0.807 | 0.798 | 0.784 | 0.670 | 0.635 | 0.629 |
| LIN28B | 0.704 | 0.685 | 0.661 | 0.687 | 0.677 | 0.656 | 0.654 | 0.641 | 0.630 | 0.512 | 0.477 | 0.564 |
| ATXN2 | 0.731 | 0.714 | 0.695 | 0.715 | 0.695 | 0.672 | 0.690 | 0.651 | 0.681 | 0.576 | 0.552 | 0.566 |
| EIF4A3 | 0.672 | 0.642 | 0.657 | 0.668 | 0.654 | 0.635 | 0.641 | 0.627 | 0.619 | 0.640 | 0.593 | 0.639 |
| LIN28A | 0.730 | 0.710 | 0.696 | 0.723 | 0.708 | 0.686 | 0.672 | 0.678 | 0.631 | 0.596 | 0.592 | 0.584 |
| FUS | 0.771 | 0.731 | 0.762 | 0.756 | 0.732 | 0.725 | 0.744 | 0.722 | 0.720 | 0.664 | 0.647 | 0.616 |
| PUM1 | 0.554 | 0.536 | 0.559 | 0.541 | 0.525 | 0.546 | 0.578 | 0.560 | 0.557 | 0.565 | 0.514 | 0.620 |
| AGO2 | 0.721 | 0.705 | 0.686 | 0.713 | 0.690 | 0.693 | 0.654 | 0.635 | 0.632 | 0.546 | 0.530 | 0.531 |
| PUM2 | 0.854 | 0.833 | 0.830 | 0.827 | 0.822 | 0.788 | 0.785 | 0.780 | 0.760 | 0.582 | 0.575 | 0.604 |
| FMR1\_iso1 | 0.718 | 0.702 | 0.681 | 0.703 | 0.694 | 0.669 | 0.678 | 0.673 | 0.634 | 0.595 | 0.578 | 0.566 |
| ADAR1 | 0.749 | 0.735 | 0.733 | 0.742 | 0.729 | 0.716 | 0.739 | 0.698 | 0.710 | 0.740 | 0.711 | 0.692 |
| ZFP36 | 0.786 | 0.786 | 0.737 | 0.801 | 0.789 | 0.782 | 0.775 | 0.762 | 0.741 | 0.591 | 0.600 | 0.570 |
| HNRNPA2B1 | 0.744 | 0.707 | 0.722 | 0.723 | 0.695 | 0.690 | 0.671 | 0.636 | 0.639 | 0.537 | 0.497 | 0.563 |
| PCBP2 | 0.738 | 0.717 | 0.696 | 0.698 | 0.687 | 0.640 | 0.717 | 0.703 | 0.652 | 0.790 | 0.745 | 0.753 |
| U2AF2 | 0.491 | 0.495 | 0.491 | 0.517 | 0.527 | 0.494 | 0.538 | 0.492 | 0.558 | 0.358 | 0.361 | 0.400 |
| AGO1 | 0.788 | 0.817 | 0.726 | 0.798 | 0.814 | 0.746 | 0.789 | 0.778 | 0.758 | 0.561 | 0.637 | 0.535 |
| RBM10 | 0.657 | 0.643 | 0.616 | 0.629 | 0.605 | 0.612 | 0.580 | 0.567 | 0.555 | 0.467 | 0.529 | 0.470 |
| C22ORF28 | 0.564 | 0.555 | 0.531 | 0.567 | 0.572 | 0.537 | 0.546 | 0.545 | 0.539 | 0.580 | 0.508 | 0.600 |
| HNRNPH1 | 0.801 | 0.790 | 0.744 | 0.801 | 0.790 | 0.743 | 0.740 | 0.718 | 0.701 | 0.598 | 0.547 | 0.591 |
| HNRNPF | 0.777 | 0.761 | 0.733 | 0.732 | 0.730 | 0.690 | 0.707 | 0.690 | 0.658 | 0.460 | 0.505 | 0.456 |
| STAU1 | 0.581 | 0.550 | 0.585 | 0.568 | 0.539 | 0.573 | 0.590 | 0.551 | 0.573 | 0.502 | 0.554 | 0.459 |
| FXR1 | 0.857 | 0.831 | 0.849 | 0.850 | 0.829 | 0.829 | 0.822 | 0.806 | 0.804 | 0.568 | 0.510 | 0.607 |
| HNRNPD | 0.604 | 0.592 | 0.588 | 0.598 | 0.586 | 0.583 | 0.626 | 0.621 | 0.600 | 0.673 | 0.626 | 0.662 |
| IGF2BP3 | 0.745 | 0.734 | 0.718 | 0.730 | 0.719 | 0.699 | 0.692 | 0.664 | 0.666 | 0.579 | 0.577 | 0.567 |
| MOV10 | 0.851 | 0.834 | 0.825 | 0.837 | 0.812 | 0.811 | 0.783 | 0.766 | 0.758 | 0.601 | 0.611 | 0.583 |
| IGF2BP2 | 0.747 | 0.751 | 0.706 | 0.754 | 0.732 | 0.731 | 0.693 | 0.676 | 0.674 | 0.605 | 0.602 | 0.564 |
| EWSR1 | 0.786 | 0.775 | 0.757 | 0.782 | 0.764 | 0.761 | 0.772 | 0.755 | 0.737 | 0.597 | 0.605 | 0.522 |
| Mean | 0.717 | 0.704 | 0.685 | 0.708 | 0.695 | 0.676 | 0.681 | 0.663 | 0.657 | 0.573 | 0.565 | 0.565 |

 |

Table S5. The number of training samples for 12 shared RBPs between RBP-24 and RBP-47. For RBP-47, only 80% of the data is used for training.

|  |  |  |
| --- | --- | --- |
| [RBP](http://www.abbreviations.com/acronyms/NICKNAMES) | # of training in RBP-24 | # of training in RBP-47 |
| C22ORF28 | 18505 | 9502 |
| Ago2 | 92346 | 33218 |
| HNRNPC | 41266 | 7680 |
| TIA1 | 34184 | 31126 |
| TIAL1 | 78984 | 40986 |
| EWSR1 | 31102 | 9944 |
| FUS | 66061 | 12124 |
| ELAVL1C | 238888 | 41144 |
| MOV10 | 26780 | 11168 |
| PUM2 | 17343 | 5730 |
| TAF15 | 13904 | 7140 |
| ZC3H7B | 40980 | 19904 |

Table S6. The overlap of 12 shared RBPs between RBP-24 and RBP-47, here we calculate the percent of training sequences in RBP-47 overlapping with the training sequences in RBP-24. We decide whether they are overlapped or not based on the coordinate on genome

|  |  |  |
| --- | --- | --- |
| [RBP](http://www.abbreviations.com/acronyms/NICKNAMES) | % of overlapped positives | % of overlapped negatives |
| C22ORF28 | 92.2 | 4.1 |
| Ago2 | 97.1 | 15.3 |
| HNRNPC | 52.1 | 8.6 |
| TIA1 | 61.9 | 8.3 |
| TIAL1 | 64.9 | 14.4 |
| EWSR1 | 94.5 | 6.3 |
| FUS | 94.2 | 11.2 |
| ELAVL1C | 36.5 | 3.8 |
| MOV10 | 94.5 | 4.7 |
| PUM2 | 85.7 | 4.0 |
| TAF15 | 91.6 | 3.3 |
| ZC3H7B | 87.7 | 8.4 |

Table S7. The overlap of 5 RBPs between negative samples in RBP-47 and binding sites derived from eCLIP, we decide whether they are overlapped or not based on the coordinate on genome .

|  |  |
| --- | --- |
| [RBP](http://www.abbreviations.com/acronyms/NICKNAMES) | % of overlap between negative samples in RBP-47 and binding sites in eCLIP |
| HNRNPC | 47.5 |
| TIA1 | 61.3 |
| EWSR1 | 28.7 |
| PUM2 | 44.0 |
| TAF15 | 55.6 |

Algorithm 1. Pseudo code of iDeepE for predicting RBP binding sites on RNAs.

|  |
| --- |
| Inputs: training RNA sequences, test RNA sequences.Outputs: binding probabilities P for test RNA sequences. |
| Do1. Pad RNA sequences into the same length, and convert them into one-hot encoded matrix
2. Train a global CNN on the one-hot encoded matrix and saved the trained model M1.
3. Do the same as step 1 for test RNA sequences, and use the trained model M1 to predict binding probabilities P1 for those test RNA sequences
4. Break the RNA sequences into overlapped subsequences with window size 101 and overlapped shift S = 50, and convert the subsequences into one-hot matrix
5. Train a local CNN on those one-hot matrix and save the trained model M2.
6. Do the same as step 4 for test RNA sequences, and use the trained model M2 to predict binding probabilities P2 for those test RNA sequences.
7. Calculate final binding probabilities: P = (P1 + P2)/2.

End |



Figure S1. The impact of different parameters on iDeepE for RBP-24 dataset. In total, there are 108 trials for the 5 parameters. In each sub figures, x-axis is a trial corresponding to the same combination of other 4 parameters.



Figure S2. The AUCs of iDeepE across 24 RBPs on non-redundant RBP-24, in which the sequences in testing set have no sequence similarity over 80%, 70% and 60% to any sequences in training set, respectively.



Figure S3. AUCs of omiXcore on testing set in RBP-24. The predicted binding scores are calculated by omiXcore.



Figure S4. The enriched binding motifs detected by iDeepE using AME for individual RBPs.



Figure S5. The AUCs of iDeepE across 47 RBPs on non-redundant RBP-47, in which the sequences in testing set have no sequence similarity over 80% to any sequences in training set.



Figure S6. The AUCs of iDeepE across 47 RBPs using shuffled positive sequences as negative samples, where positive sequences are the same as RBP-47, but the negative sequences are those shuffled sequences of positive sequences in training set.



**Figure S7.** The cross-dataset validation performance of iDeepE-1 on non-redundant RBP-47, in which the sequences in testing set have no sequence similarity over 80% to any sequences in training set.

**Network architectures**

The layers of the three models CNNs, CNN-LSTM.

CNN (

 (layer1): Sequential (

 (0): Conv2d(6, 16, kernel\_size=(4, 10), stride=(1, 1))

 (1): BatchNorm2d(16, eps=1e-05, momentum=0.1, affine=True)

 (2): ReLU ()

 (3): MaxPool2d (size=(1, 3), stride=(1, 1), dilation=(1, 1))

 )

 (layer2): Sequential (

 (0): Conv2d(16, 16, kernel\_size=(1, 10), stride=(1, 1))

 (1): BatchNorm2d(16, eps=1e-05, momentum=0.1, affine=True)

 (2): ReLU ()

 (3): MaxPool2d (size=(1, 3), stride=(1, 1), dilation=(1, 1))

 )

 (drop1): Dropout (p = 0.25)

 (fc1): Linear (1360 -> 200)

 (drop2): Dropout (p = 0.25)

 (relu1): ReLU ()

 (fc2): Linear (200 -> 2)

)

CNN\_LSTM (

 (layer1): Sequential (

 (0): Conv2d(6, 16, kernel\_size=(4, 10), stride=(1, 1))

 (1): BatchNorm2d(16, eps=1e-05, momentum=0.1, affine=True)

 (2): ReLU ()

 (3): MaxPool2d (size=(1, 3), stride=(1, 1), dilation=(1, 1))

 )

 (downsample): Conv2d(16, 1, kernel\_size=(1, 10), stride=(1, 1))

 (layer2): LSTM(87, 200, num\_layers=2, batch\_first=True, bidirectional=True)

 (drop1): Dropout (p = 0.5)

 (fc1): Linear (400 -> 200)

 (drop2): Dropout (p = 0.25)

 (relu1): ReLU ()

 (fc2): Linear (200 -> 2)

)