Table S1 Performance of the proposed method on the Saccharomyces cerevisiae and Human datasets

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Dataset | Acc (%) | Sens (%) | Spec (%) | Prec (%) | MCC (%) | AUC (%) |
| Saccharomyces cerevisiae | 97.37±0.14 | 97.19±0.29 | 97.55±0.24 | 97.54±0.23 | 94.75±0.27 | 99.48 |
| Human | 99.71± 0.14 | 99.69± 0.19 | 99.72±0.21 | 99.73± 0.20 | 99.41± 0.30 | 99.9 |

Table S1 shows the performance of the hybrid classifier on the Saccharomyces cerevisiae, and Human datasets in terms of a accuracy is 97.37% and 99.71% with a standard deviation of 0.14 and 0.14 respectively.

Table S2 Comparision of accuracy (%) of the hybrid classifier with existing classifiers using Saccharomyces cerevisiae and Human datasets

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Dataset | DNN-XGB | DNN | XGB | RF | AB | SVM | NB | KNN |
| S.cerevisiae | 97.37 | 91.25 | 90.72 | 85.68 | 86.64 | 87.28 | 66.00 | 83.75 |
| Human | 99.74 | 97.49 | 97.94 | 96.44 | 96.65 | 93.84 | 71.23 | 90.81 |

Table S3 Comparision of t-statistic of hybrid classifier existing classifiers using Saccharomyces cerevisiae and Human datasets

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Algorithm  Dataset | DNN-XGB vs DNN  (t-value) | DNN-XGB vs XGB  (t-value) | DNN-XGB vs RF  (t-value) | DNN-XGB vs AB  (t-value) | DNN-XGB vs SVM  (t-value) | DNN-XGB vs NB  (t-value) | DNN-XGB vs KNN  (t-value) |
| S. cerevisiae | 13.98 | 20.20 | 38.27 | 31.62 | 27.82 | 51.75 | 52.06 |
| Human | 3.40 | 7.14 | 15.00 | 9.81 | 9.83 | 23.96 | 38.83 |

|  |  |
| --- | --- |
| SCcompare.png  Figure S1 Comparison of ROC curves of different classifiers using the S. cerevisiae dataset | SCcompare.png  Figure S2 Comparison of ROC curves of different classifiers using the S. cerevisiae (core subset) dataset |
| Humancompare.png  Figure S3 Comparison of ROC curves of different classifiers using the Human dataset | BHcompare.png  Figure S4 Comparison of ROC curves of different classifiers using the *Human- B. Anthracis* dataset |
| YHcompare.png  Figure S5 Comparison of ROC curves of different classifiers using the *Human- Y.pestis* dataset |  |

A comparative study between the performance of the DNN-XGB classifier with three deep neural network based methods, DeepPPI [25], EsnDNN [27], and Embedding+RNN [30] using *Saccharomyces cerevisiae* dataset, presented in Table S4. The proposed method provides around 2%-4.5% enhancement in accuracy compared to the existing methods. Further the proposed method is compared with existing method on *Human* dataset. The results are listed in Table S5. In the classical prediction methods the DCT+WT+GE+WSRC [7] produces the best accuracy and in deep learning based PPI prediction methods Embedding +RNN [30] yields the best accuracy. The proposed method achieves a 0.16% enhancement in prediction accuracy and 0.34% higher MCC as compared to the DCT+WT+GE+WSRC whereas an enhancement of 1.73% in terms of accuracy is observed as compared to Embedding + RNN.

Table S4: Comparison of prediction performance of the proposed method with existing state-of art-methods on *Saccharomyces cerevisiae* dataset

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Method** | **Acc (%)** | **Sens (%)** | **Prec (%)** | **MCC (%)** | **AUC(%)** |
| DeepPPI [25] | 92.58 | 90.95 | 94.41 | 85.41 | 97.54 |
| EsnDNN [27] | 95.29 | 95.12 | 95.45 | 90.59 | 97.00 |
| Embedding+RNN [30] | 92.59 | 91.40 | 93.65 | 85.20 | 97.40 |
| Proposed Method (DNN-XGB) | 97.37±0.14 | 97.19±0.29 | 97.55±0.24 | 94.75±0.27 | 99.48 |

Table S5: Comparison of prediction performance of the proposed method with other states of art methods on *Human* dataset

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Method** | **Acc (%)** | **Sens (%)** | **Prec (%)** | **MCC (%)** | **AUC(%)** |
| CT+SVM [4] | 83.9 | 84.8 | 84.21 | NA | NA |
| PCA+SVM [11] | 93.45 | 89.29 | 89.84 | 85.71 | 93.00 |
| pseudo-SMR+WSRC [15] | 96.71 | 93.49 | 99.67 | 93.60 | 96.87 |
| PSSM+RoF [16] | 98.95 | 98.60 | 99.20 | 97.89 | 98.87 |
| SNF+LPA [18] | 99.30 | 99.95 | 98.61 | 98.62 | 99.93 |
| CT+DNN [29] | 97.11 | 98.91 | NA | NA | 98.35 |
| AC+DNN [29] | 96.84 | 98.67 | NA | NA | 98.02 |
| LD+DNN [29] | 95.30 | 98.28 | NA | NA | 97.57 |
| MOS+DNN [29] | 94.34 | 98.28 | NA | NA | 97.73 |
| Embedding+RNN [30] | 97.98 | 96.85 | 98.90 | NA | 99.60 |
| AC+Autoencoder [31] | 97.19 | 98.06 | 96.27 | NA | NA |
| GE+WSRC [46] | 97.66 | 95.28 | 99.81 | 95.41 | 97.80 |
| LDA+RF [48] | 96.40 | 94.20 | NA | 92.8 | 98.20 |
| LDA+SVM [48] | 90.07 | 89.7 | NA | 81.3 | 95.20 |
| Proposed Method (DNN-XGB) | 99.71± 0.14 | 99.69± 0.19 | 99.73± 0.20 | 99.41± 0.30 | 99.9 |

Note: NA means not available

|  |  |
| --- | --- |
| SCraw.png  Figure S6 (a)  SCextracred.png  Figure S6 (b)  Figure S6 t-SNE plots of *Saccharomyces cerevisiae* dataset (a) raw input features (b) abstracted features present in the fusion layer | BHraw.png  Figure S7 (a)  BHextracred.png  Figure S7 (b)  Figure S7 t-SNE plots of *Saccharomyces cerevisiae (core subset)* dataset (a) raw input features (b) abstracted features present in the fusion layer |
| Humanraw.png  Figure S8 (a)  Humanextracred.png  Figure S8 (b)  Figure S8 t-SNE plots of *Human* dataset (a) raw input features (b) abstracted features present in the fusion layer | BHraw.png  Figure S9(a)  BHextracred.png  Figure S9 (b)  Figure S9 t-SNE plots of *Human- B.Anthracis* dataset (a) raw input features (b) abstracted features present in the fusion layer |
| BHraw.png  Figure 9 (a)  BHextracred.png  Figure 10 (b)  Figure S10 t-SNE plots of *Human- Y.pestis* dataset (a) raw input features (b) abstracted features present in the fusion layer |  |