

A.

Metric	Basic Model Testing Performance (Sequence + Structural Features)					
	Gradient Boosting Classifier	Random Forest Classifier	Extra Trees Classifier	Logistic Regression	Support Vector Machine	K Nearest Neighbor
Accuracy	0.83	0.84	0.83	0.79	0.81	0.76
Precision	0.87	0.90	0.88	0.90	0.90	0.89
Recall	0.89	0.88	0.88	0.79	0.82	0.76
F1	0.88	0.89	0.88	0.84	0.86	0.82
AUC	0.79	0.82	0.80	0.79	0.80	0.77
Average Precision-Recall	0.93	0.93	0.92	0.91	0.92	0.90

Supplementary Figure S18. (A) In Table A, we display the prediction performance of the "Basic" model across all classifiers trained on the top 25 sequence and structural features identified using the hybrid feature selection technique. The table summarizes the algorithms' best scores on the testing dataset for each criterion (accuracy, precision, recall, F1, AUC-ROC, Average Precision-Recall).