Protein Acetylation Sites with Complex-Valued Polynomial Model

Wenzheng BAO, Bin YANG

Frontiers of Computer Science, DOI: 10.1007/s11704-023-2640-9

Motivation

- Protein acetylation refers to a process of adding acetyl groups (CH3CO-) to lysine residues on protein chains. As one of the most commonly used protein post-translational modifications, lysine acetylation plays an important role in different organisms. In our study, we developed a human-specific method which uses a cascade classifier of complex-valued polynomial model (CVPM), combined with sequence and structural feature descriptors to solve the problem of imbalance between positive and negative samples.

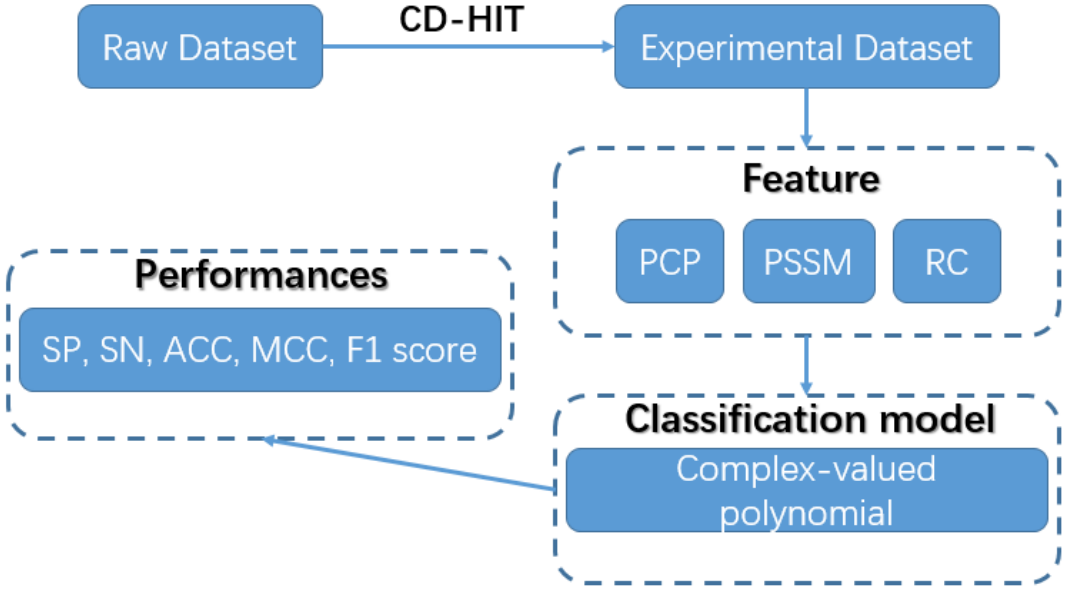

Fig 1 The Outlines of The Proposed Method.

Results

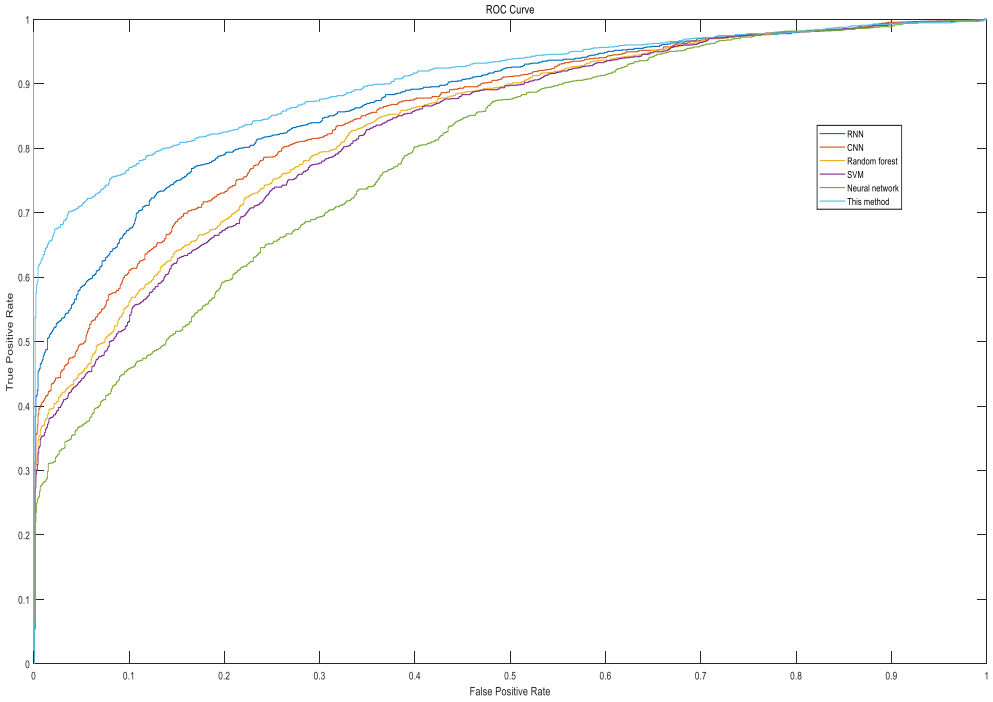

Fig 2 ROC Curves of Different Methods.

Table 2 The Performances of Different Classification Models.

Methods	S_n (%)	S_p (%)	*ACC*	*MCC*	*F*1
Neural network	75.28	55.38	65.33	0.3129	0.6847
SVM	60.62	62.88	61.75	0.2351	0.6131
Random forest	71.97	68.42	70.20	0.4042	0.7071
CNN	73.14	72.04	72.59	0.4519	0.7274
RNN	65.60	74.29	69.94	0.4004	0.6858
Proposed Algorithm	79.15	78.17	78.66	0.5733	0.7876