

Radial Basis Function Neural Networks in Protein Sequence Classification

Zarita Zainuddin and Maragatham Kumar

*School of Mathematical Sciences,
University Science Malaysia, 11800 USM Pulau Pinang*

ABSTRACT

Applications of neural networks in bioinformatics have expanded tremendously in recent years due to the capabilities of neural networks to solve biological problems. Neural networks have been implemented in numerous biological fields. In this paper, standard radial basis function and modular radial basis function neural networks are used to classify protein sequences to multiple classes. n -gram method is used to transform protein features to real values. A learning strategy known as the self-organized selection of centers is presented. In this strategy, a training algorithm based on subtractive clustering is used to train the network. The radial basis function created by the `newrb` function from Matlab uses gradient based iterative method as the learning strategy. The proposed method is implemented in the Matlab which creates a new network that undergo a hybrid learning process. The networks called SC/RBF (Subtractive Clustering–Radial Basis Function) and SC/Modular RBF (Subtractive Clustering-Modular Radial Basis Function) are used to test against the standard Radial Basis Function and modular Radial Basis Function in protein classification. Classification criteria consist of two heuristic rules are implemented to test on the classification performance rate. The real world problem that has been considered is classification of human protein sequences into ten different superfamilies which based on protein function groups. These human protein sequences are downloaded from Protein Information Resource (PIR) database.