BIOINFORMATION

Supplementary material:

Euclidian distances:

$$d=\sqrt{((e_1-o_1)^2 + (e_2-o_2)^2)}$$

where e_1, e_2 are expected outputs and o_1, o_2 are predicted outputs of two output nodes respectively. For example, maximum possible distance between the predicted and original outputs for the promoter can be considered when the predicted outputs is (1 -1) instead of (-1 1) and vice versa for a non-promoter. Therefore, value of maximum possible distance (max_d) between predicted and expected output comes out to be $2\sqrt{2}$. Class membership was thus defined as: $m=1-d/max_d$. Using this distance their actual membership for the class assigned to them was calculated. Mean value obtained for class memberships for both the classes was used as measure of robustness of classification.

Table 1: Detailed results of classification on training and validation sets (TP= number of true positives, TN= number of true negatives, FP= number of false positives, FN=number of false negatives)

	Size of data	TP	TN	FP	FN	Sensitivity	Specificity	Accuracy
Training Set	1797	663	590	310	234	73.91%	65.56%	69.73%
Validation Set	1195	424	357	243	171	71.26%	59.50%	65.36%