Project 2 Team 2 Report

<Yuanhang Lin, Pallavi,Raiturkar, Hudanyun Sheng, Weihuang Xu>

1. Member roles

Member	RVM	SVM	GPR	Cross validation Confusion matrix	Report writing
Y. Lin			1		
P. Raiturkar		1			
H. Sheng				1	1
W. Xu	1				1

2. Experiments and results

In this project, we used one-vs-all method to train Gaussian kernel based SVM, RVM and GPR. Since the data had 5 classes, there were 10 binary classifiers for each algorithm. The predicted label was calculated from voting based on the results from those 10 binary classifiers. In order to identify the input samples that are not from any class (non-class sample), we calculated the average Mahalanobis distance between samples and class center in each class and set threshold as 1.5 times the average Mahalanobis distance. For example, if one test sample is labeled as class 3 after voting, the Mahalanobis distance between this test sample and the center of class 3 is calculated. If the Mahalanobis distance is larger than the threshold (1.5 times the average Mahalanobis distance of class 3), it would be labeled as non-class. Otherwise, it would be labeled as class 3. We did 5-fold cross-validation on each algorithm. The training data was randomly divided into 5 groups at beginning and used for cross-validation. The confusion matrices from each validation set was reported and the number of support vectors and relevant vectors was also reported. In order to show the non-class results in confusion matrix. We labeled the non-classes data as class 6 in the confusion matrix.

2.1. Relevance Vector Machines (RVM)

The 1st fold of the cross validation get the best accuracy, so the parameters learned from 1st fold is chosen, the number of Relevance Vectors for the chosen fold for each pair of two-class classifiers are shown below:

Table.1 The number of relevance vectors for each binary classifier from 1st fold for RVM.

Two-class classifier	Number of Relevance Vectors	Two-class classifier	Number of Relevance Vectors
(1, 2)	23	**(2, 4)**	38
(1, 3)	26	**(2, 5)**	16
(1, 4)	21	**(3, 4)**	26
(1, 5)	16	**(3, 5)**	34
(2, 3)	18	**(4, 5)**	30

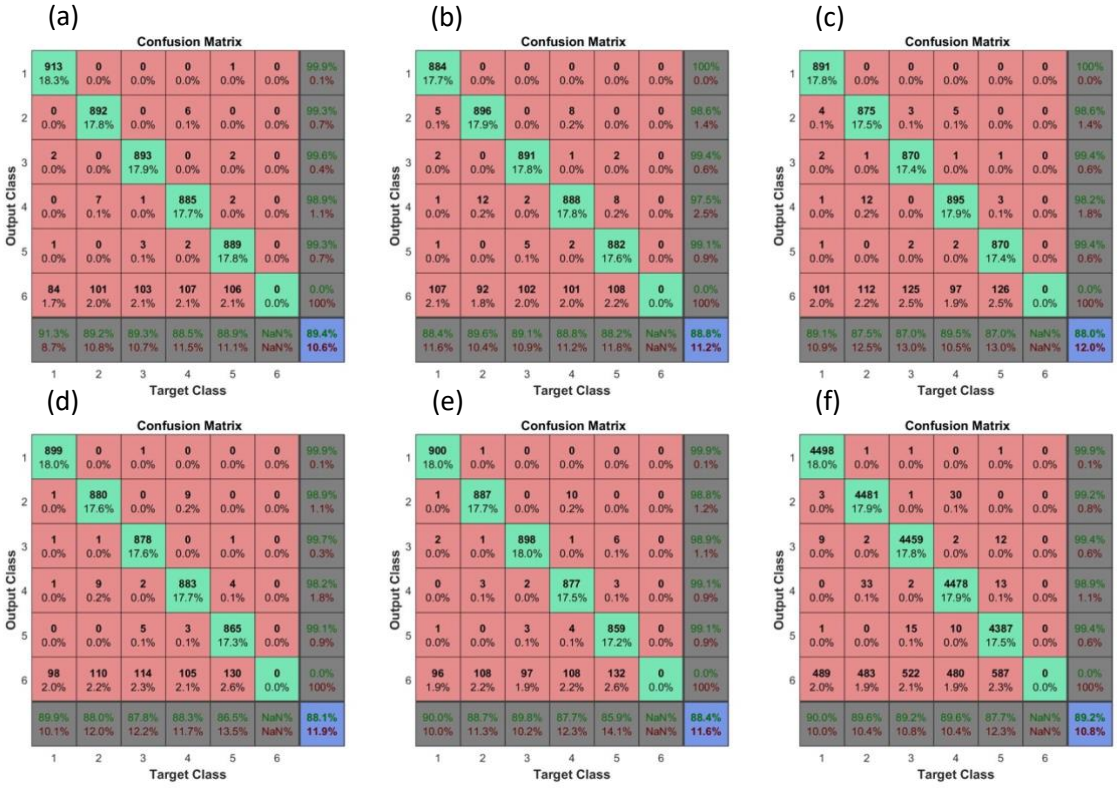

Fig.1. Confusion matrix for RVM classifier on validation set 1 (a), validation set 2 (b), validation set 3 (c), validation set 4 (d), validation set 5 (e), and overall results (f).

2.2. Support Vector Machines (SVM)

The 1st fold of the cross validation get the best accuracy, so the parameters learned from 1st fold is chosen, the number of Support Vectors for the chosen fold for each pair of two-class classifiers are shown below:

Table.2 The number of support vectors for each binary classifier in 1st fold for SVM.

Two-class classifier	Number of Support Vectors	Two-class classifier	Number of Support Vectors
(1, 2)	367	**(2, 4)**	1041
(1, 3)	426	**(2, 5)**	278
(1, 4)	318	**(3, 4)**	393
(1, 5)	303	**(3, 5)**	651
(2, 3)	302	**(4, 5)**	606

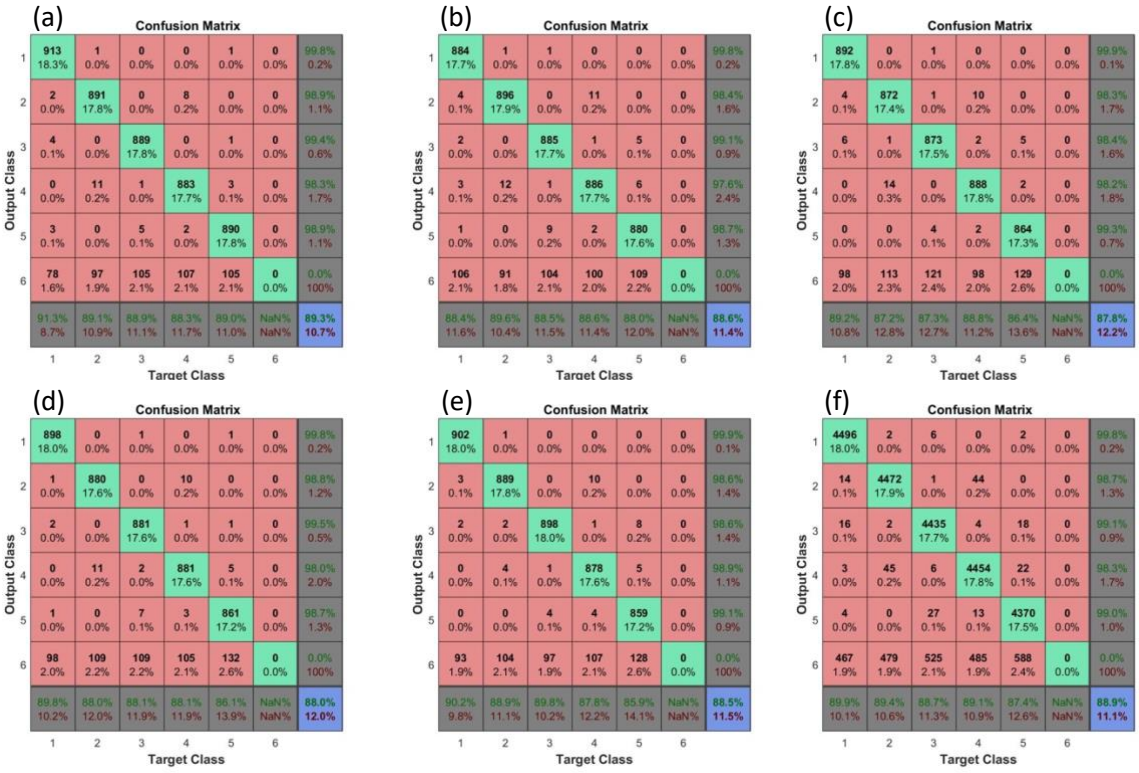

Fig.2. Confusion matrix for SVM classifier on validation set 1 (a), validation set 2 (b), validation set 3 (c), validation set 4 (d), validation set 5 (e), and overall results (f).

2.3. Gaussian Process Regression (GPR)

The 1st fold of the cross validation get the best accuracy, so the parameters learned from 1st fold is chosen to get the overall confusion matrix.

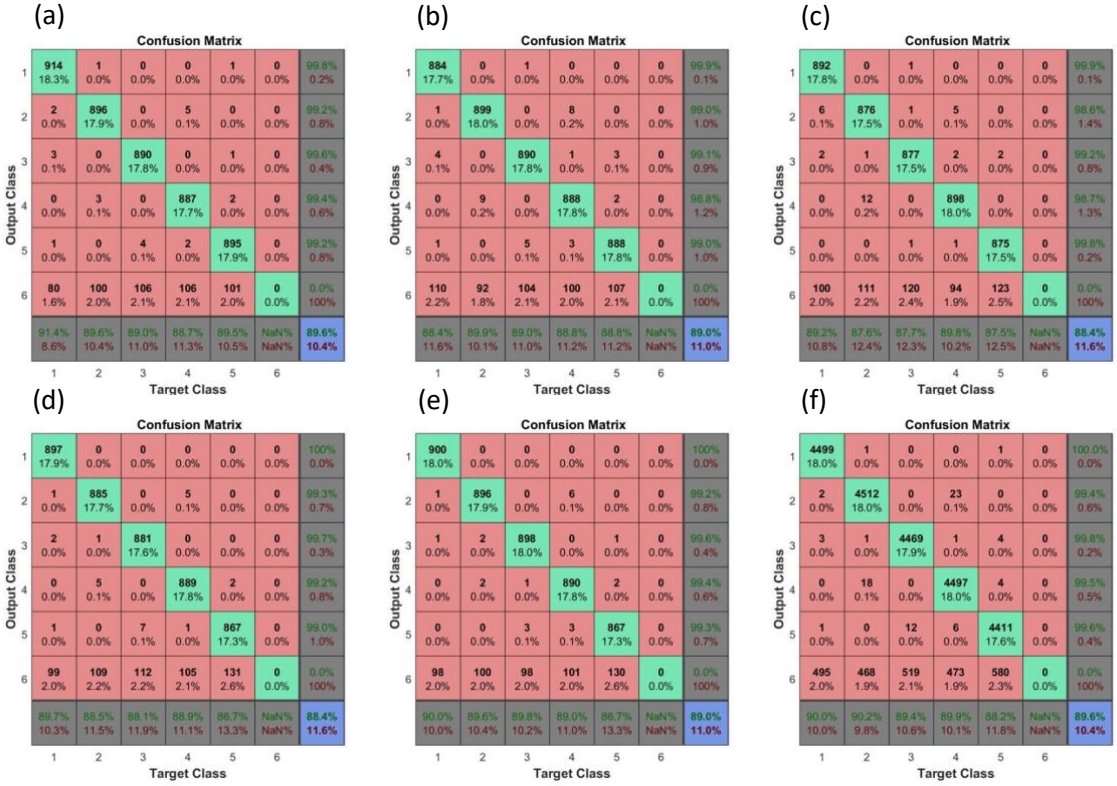

Fig.3. Confusion matrix for GPR classifier on validation set 1 (a), validation set 2 (b), validation set 3 (c), validation set 4 (d), validation set 5 (e), and overall results (f).

3. Observations: For all three algorithms, the results shown are those with Gaussian kernel used. We can see that the GPR method gets a little better result compared to other two algorithms. What is more, the total training procedure took SVM an average time of 421.85 seconds, took RVM an average time of 1428.62 seconds, while took GPR and average time of 631.75 seconds. When both using 100 max iterations, the SVM is much faster than RVM, with the performance only a little bit lower.