KNN

Set K=10 to find the result of different number of PCA bases being projected on:

Base	1	3	5	10	20	30
Acc	0.2809	0.5111	0.7588	0.9287	0.9693	0.9739
Base	40	60	70	80	90	100
Acc	0.9737	0.9727	0.9722	0.9718	0.9703	0.9701

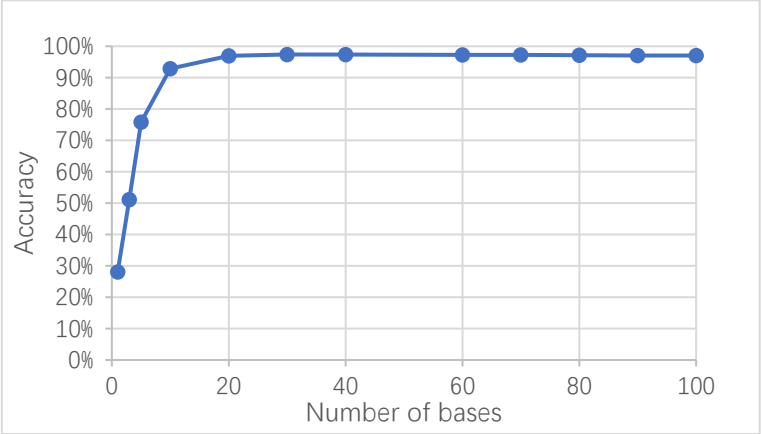

Figure 1 Accuracy with Different Number of Bases (K=10)

From the plot, it can conclude that the increase number of base will cause an increase in accuracy, but when the number of base is larger than 30, there will be no significant improvement in accuracy.

		K					
		5	10	20	30	50	100
Base	10	0.9275	0.9287	0.9280	0.9247	0.9202	0.9146
	20	0.9692	0.9693	0.9655	0.9606	0.9575	0.9515
	30	0.9750	0.9739	0.9710	0.9681	0.9624	0.9566
	40	0.9744	0.9737	0.9712	0.9675	0.9625	0.9559
	60	0.9743	0.9727	0.9693	0.9665	0.9608	0.9516
	80	0.9732	0.9718	0.9676	0.9642	0.9601	0.9502
	100	0.9726	0.9701	0.9657	0.9623	0.9575	0.9489

For the KNN method, the best test accuracy is 97.5%, with 30 bases and K=5.

SVM

With default setting of SVM, find the relationship between number of bases and accuracy.

Number of Bases	40	60	80	100
Accuracy	0.9362	0.9438	0.9447	0.9457

Set number of bases = 100,

		Box Constraint				
		0.01	0.1	1	5	10
Kernel	gaussian	0.1464	0.4736	0.8859	0.8894	0.8895
	linear	0.8469	0.8475	0.8476	0.8477	0.8477
	polynomial	0.9382	0.9347	0.9288	0.9103	0.8936

The polynomial kernel with small C has best result.

Set kernel as polynomial, and C=0.01, to find the relationship between number of bases.

Number of Bases	1	5	10	30	50
Accuracy	0.3133	0.7709	0.9382	0.9807	0.98

For SVM method, the best test accuracy is 98.07%, with 30 PCA bases, polynomial kernel and C=0.01.