

Description of program:

In this program, I used the sklearn library to test the accuracy of the SVM.SVC model with MNIST data of images of handwritten number 8 and 6. From this data set, I created 3 models: SVC with a linear kernel, SVC with a poly kernel, and SVC with a rbf kernel. From these I created a K-fold cross validation data set with 5 splits. I then ran each model with the training sets from each fold and used it to predict the labels for the X_testing dataset. After predicting labels for each value in the X_testing dataset, I computed the accuracies of each model's folds with the values in the y_testing dataset. After each fold's accuracy was collected, it is appended to a list of accuracies that is used to calculate the average accuracy of each model. Below is the table of all the accuracies that each model generated:

Accuracy Table			
	Linear Model	Poly Model	RBF Model
Fold 1	0.9871951219512195	0.995731707317073	0.9939024390243902
Fold 2	0.9859756097560975	0.9963414634146341	0.9957317073170732
Fold 3	0.9859756097560975	0.9969512195121951	0.9957317073170732
Fold 4	0.9884146341463415	0.9951219512195122	0.9951219512195122
Fold 5	0.9871951219512195	0.9945121951219512	0.9963414634146341
Average	0.9869512195121951	0.9957317073170732	0.9953658536585366

Conclusion:

With the calculations of the accuracies of each model, there are percentages that repeat amongst the models. In the 4th fold of both the Poly model and the RBF model, the accuracy is 0.9951219512195122 which either shows that the prediction function of the models is similar to each other, or the fold of dataset has issues. Furthermore, the average accuracy of each model is very close to predicting except for linear being off by a 0.008780487804878057 of a percent to the Poly model and 0.008414634146341471 of a percent to the RBF model. In conclusion, I believe that the SVM model for linear, poly, and RBF models are similar in their ability to predict on images.