

Problem 1:

Optimal value of K: 8

Accuracy: 0.9824561403508771

Precision: 0.9850746268656716

Recall: 0.9850746268656716

Figure 1: Accuracy, Precision, Recall values using SVM Classifier

Using a for loop, the optimal value of K that gives the highest accuracy was calculated, and as it can be seen that a value of 8 gives an accuracy of about 98 % and precision of about 98.5 %.

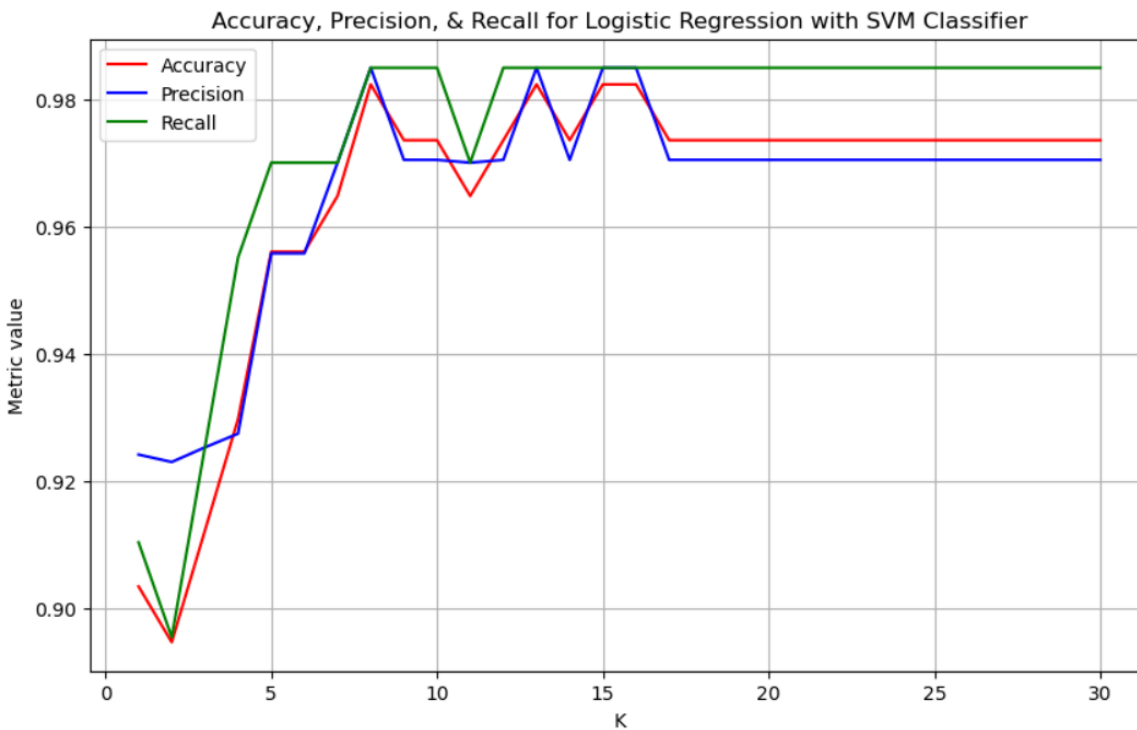


Figure 2: Plot of Accuracy, Precision, and Recall with the SVM Classifier

Optimal value of K: 8
Accuracy: 0.9736842105263158
Precision: 0.9705882352941176
Recall: 0.9850746268656716

Figure 3: Results for SVM Classifier with Linear Kernel

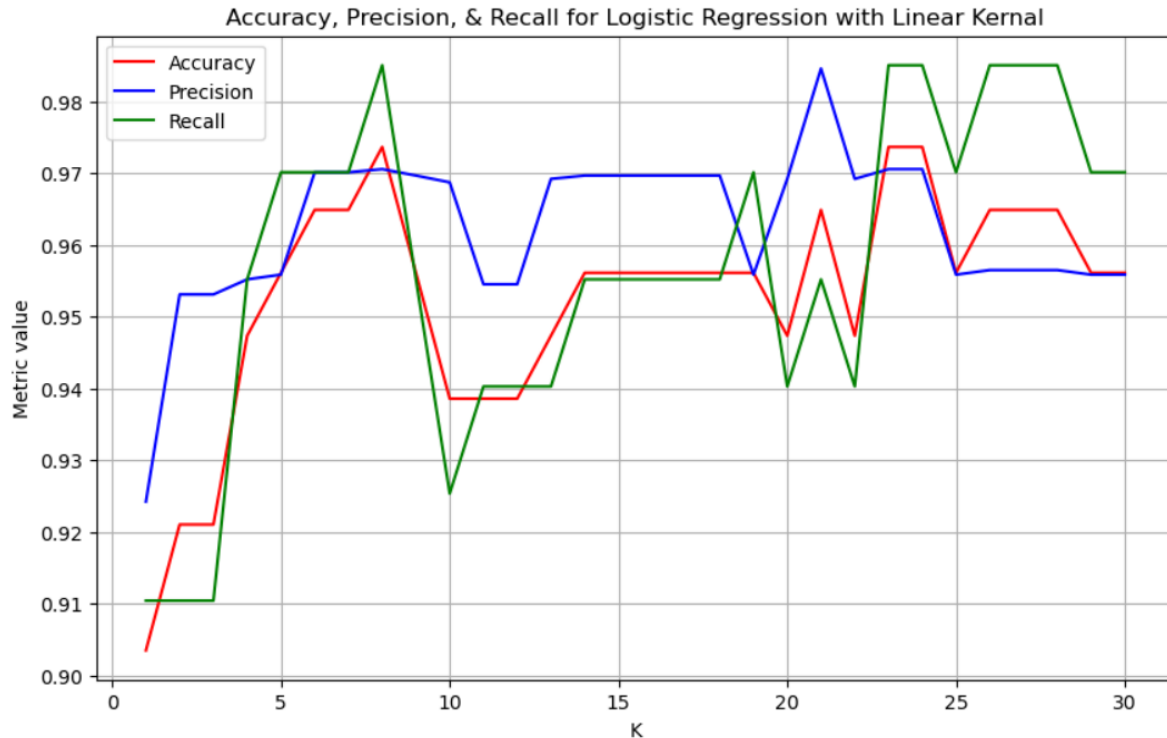


Figure 4: Plot for SVM Classifier with Linear Kernel

Figure 3 and Figure 4, both show results for the SVM Classifier with a “linear” kernel. The optimal value of 8 for k gave the highest accuracy value. These results are similar to the SVM Classifier model without a kernel specified. In comparison to the Logistic Regression results, the SVM Classifier with linear kernel gives more accurate results. The accuracy value for Logistic Regression was around 96 %.

Optimal value of K: 5
Accuracy: 0.956140350877193
Precision: 0.9305555555555556
Recall: 1.0

Figure 5: Results for SVM Classifier with Poly Kernel

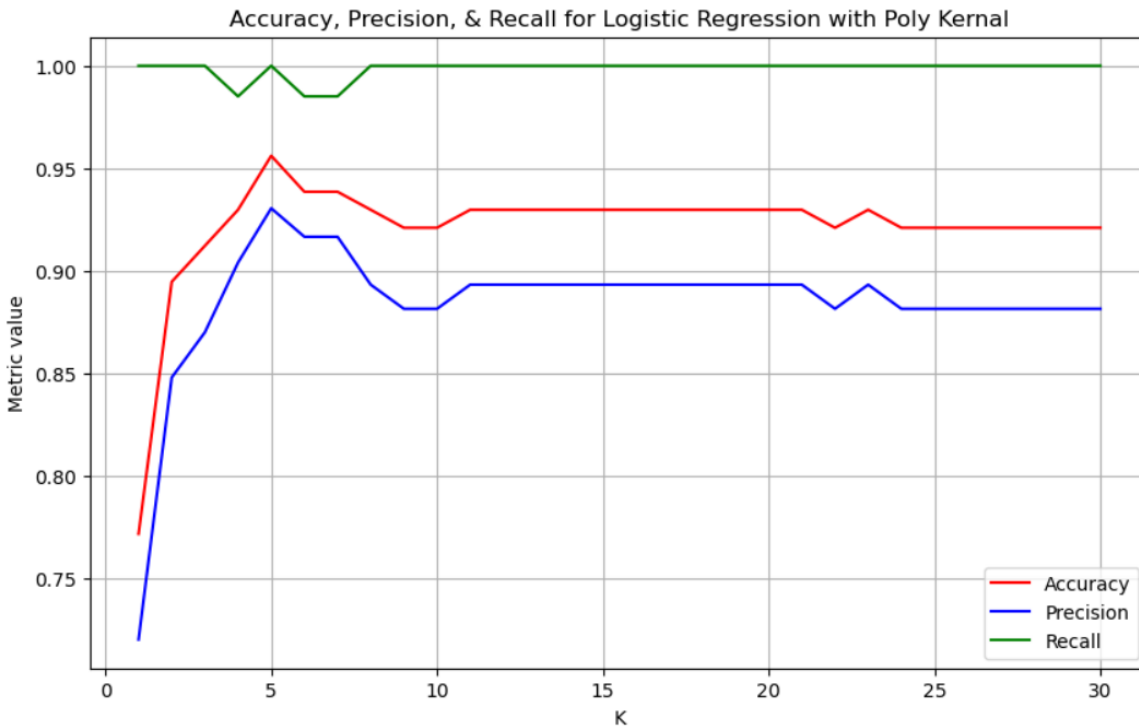


Figure 6: Plot for SVM Classifier with Poly Kernel

Figure 5 and Figure 6, both show results for the SVM Classifier with a “poly” kernel. The optimal value of 5 for k gave the highest accuracy value. These results are similar to the SVM Classifier model without a kernel specified. In comparison to the Logistic Regression results, the SVM Classifier with poly kernel gives more accurate results. The accuracy value for Logistic Regression was around 96 %. When compared against the SVM Classifier with linear kernel, this model is less effective with an average of 3 % less accurate. Another factor that was experimented in regard to the rbf kernel properties was the degree value. Three different values, 2, 4, and 5 were tested with and the value 4 gave the best results.

Optimal value of K: 10
Accuracy: 0.9912280701754386
Precision: 0.9852941176470589
Recall: 1.0

Figure 7: Results for SVM Classifier with RBF Kernel

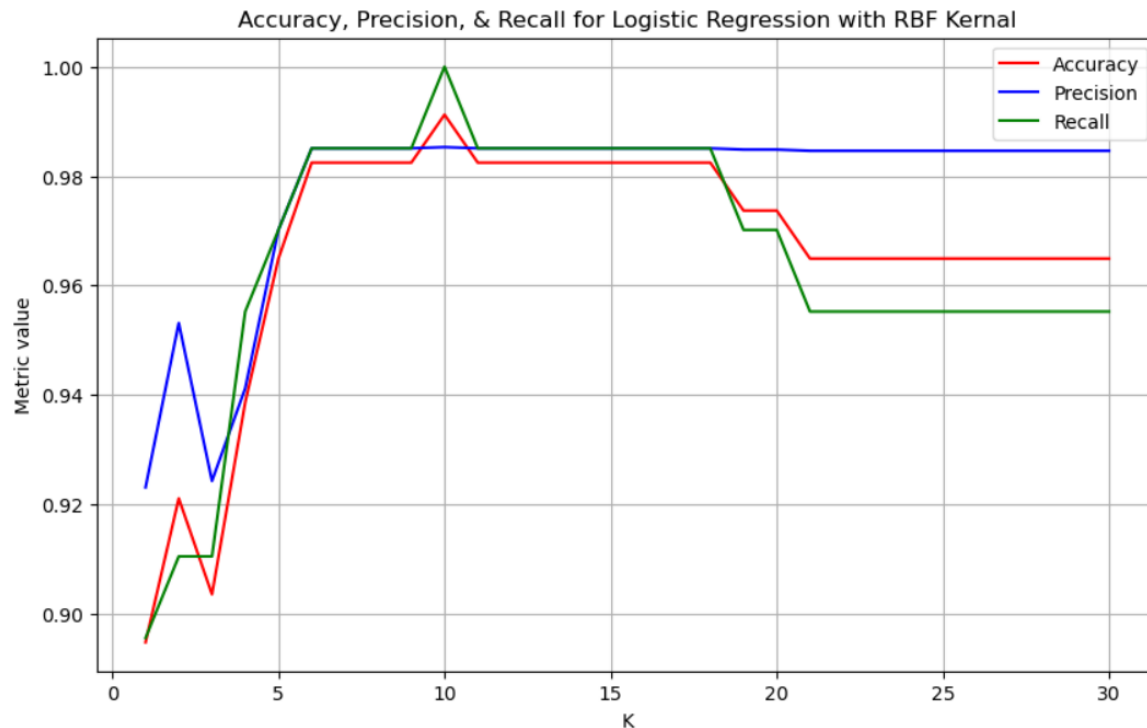


Figure 8: Plot for SVM Classifier with RBF Kernel

Figure 5 and Figure 6, both show results for the SVM Classifier with a “rbf” kernel. The optimal value of 10 for k gave the highest accuracy value. These results are similar to the SVM Classifier model without a kernel specified. In comparison to the Logistic Regression results, the SVM Classifier with poly kernel gives more accurate results. The accuracy value for Logistic Regression was around 96 %. When compared against the SVM Classifier with linear and poly kernels, this model was effective with an average of 3 - 4 % more accurate. Another factor that was experimented in regard to the rbf kernel properties was the gamma value. Three different values, 0.1, 0.01, and 0.001 were tested with and the value 0.01 gave the best results.

Problem 2:

Mean Squared Error	=	0.34800643643616763
Mean Squared Error (Linear)	=	0.27125614105684565
Mean Squared Error (Polynomial)	=	0.4602592505883177
Mean Squared Error (RBF)	=	0.26449921132586146

Figure 9: Mean Squared Errors for SVR Model

Unlike the previous problem, the accuracy, precision, and recall cannot be calculated for this problem because of the binary values (0,1) present in the dataset. Hence, to compare the accuracy of the different models of SVR with “linear”, “poly”, and “rbf” models, the `mean_squared_error()` function was used. As it can be seen in Figure 9, of the 3 different models, the SVR model with the RBF and the linear kernels were far more accurate than the model with poly kernel. Also, due to the fact that this data is being evaluated using `mean_squared_error()` function, it cannot be compared with the model computed in Homework # 1.