Report.pdf

Shane Zabel

KNN.R

---

1. *Next carry out the experiments using a k-Nearest Neighbor classifier. Choose 5 different values of k (3, 5, 7, 9, and 11) and carry out 10 experiments each using different samples of training and test data.*

Table Containing Average Accuracies Over 10 Experiments for Each k-value

k	Average Accuracy of 10 Experiments
3	0.6779221
5	0.7077922
7	0.7038961
9	0.6922078
11	0.7038961

---

2. Write a short paragraph on which method works the best for this dataset.

As seen in the table below the SVM method with a linear kernel worked best for this dataset in the experiments performed with an average accuracy of 79.2% against the test data. Naïve Bayesian was second with an average accuracy of 78.2%. SVM methods were next with KNN methods performing worst.

Table Containing Average Accuracies Over 10 Experiments for Each Method

Metod	Average Accuracy of 10 Experiments
SVM - Linear	0.7922078
naïve Bayesian	0.7818181
SVM - Radial Basis	0.7675325
SVM - Sigmoid	0.7415584
SVM - Polynomial	0.7233766
KNN - k=5	0.7077922
KNN - k=11	0.7038961
KNN - k=7	0.7038961
KNN - k=9	0.6922078
KNN - k=3	0.6779221