Problem Set 10

Dillon Beake

April 15, 2024

1 Model Performance Summary

Table 1: Model Performance Summary

alg	accuracy	best_params
logit	0.8526025	lambda = 1e-10
tree	0.8684170	min_n = 10 , tree_depth = 15 , cost_complexity = 0.001
nnet	0.8347920	hidden_units = 1 , lambda = 1
knn	0.8433901	neighbors = 30
svm	0.8638108	cost = 13.6851686464717 , rbf_sigma = 0.0777665905051034

The DECISION TREE ALGORITHM achieved the highest accuracy performance for out-of-sample testing = 0.8684. NEURAL NETWORK ALGORITHM had the lowest, = 0.8348. When considering this range from highest to lowest, each algorithm performed fairly similarly.

Also, I did tune the parameters of SVM using 10 iterations and 20 iterations. I got the same accuracy for both, so I kept the 20 iterations. I did not try 30 iterations because of the time it took for the first two, and also I was a bit worried about my computer's computing powers.

The optimal hyper-parameters are also demonstrated by the table under the column (best-parameters).