PS10

Audrey Hopewell

April 7, 2020

Algorithm	Optimal Values	Out-of-sample Performance
Trees	minsplit=49; minbucket=17; cp=0.0111	f1=0.8987161, gmean=0.6650725
Logistic regression	lambda=0.237; alpha=0.0901	f1=0.8859958, gmean=0.4725724
Neural network	size=9; decay=0.164; maxit=1000	f1=0.9093579, gmean=0.7526919
Naive Bayes	N/A	f1=0.8879585, gmean=0.7343107
kNN	k=29	f1=0.9009453, gmean=0.7430233
SVM	kernel=radial; cost=0.5; gamma=0.5	f1=0.9091085, gmean=0.7306514

The neural network algorithm has the best performance based on the F1 score; SVM performs well by this measure too. The neural network algorithm also does best based on the G-measure; kNN does fairly well by this measure as well. Overall, neural network seems to be the best-performing algorithm for this task.