

Madelon Dataset

Accuracy Scores without feature selection:

Logistic Regression	Linear Discriminant Analysis	Decision Tree Classifier
0.590	0.577	0.753

neighbors	KNN
1	0.653
2	0.658
3	0.695
4	0.670

Accuracy Scores with feature selection

Forward selection with feature count that yields best accuracy score:

Since it takes so much computation time when upper limit of number of selections increased, I set upper limit of feature count to 20.

N = feature count

Accuracy Score	Linear Regression
N = 6	0.603

N = feature count

Accuracy Score	Linear Discriminant Analysis
N = 19	0.628

N = feature count

Accuracy Score	Decision Tree Classifier
N = 3	0.558

N = feature count

Accuracy Score	KNN (neighbors=1)
N = 3	0.558

N = feature count

Accuracy Score	KNN (neighbors=2)
N = 9	0.587

N = feature count

Accuracy Score	KNN (neighbors=3)
N = 10	0.587

N = feature count

Accuracy Score	KNN (neighbors=4)
N = 3	0.558

Backward selection metrics does not included here because it requires so much computation power to compute in reasonable time. Due to high number of features in Madelon Dataset computations could not complete. But algorithm implemented in project.