Executive Summary

For our upcoming bid to Company X, our submission must include machine learning in order to achieve a high level of competitiveness against other bidders.

To demonstrate the power of machine learning, I have utilized the Madelon dataset. This dataset is highly non-linear, contains 500 features, and predicts a binary classification target. Based on these characteristics, I have selected two models that traditionally perform well: Logistic Regression (LR) and K Nearest Neighbors (KNN).

The work contained within this report outlines three different approaches (see summary below), each with selected parameters based on machine learning.

1. **LR using Standard Scaler** with little regularization, utilizing l2, known as the ridge model
2. **LR using Standard Scaler** with regularization (default gamma of 1), utilizing l1, known as the lasso model
3. **LR and KNN using SelectKBest** with regularization, LR utilizing l2, and a grid search to cross validate results

The conclusion of the work herein has confirmed the assertion that machine learning is necessary in order to present a competitive bid. Our model KNN using SelectKBest has selected 5 salient features (out of 500) to predict our target with an acceptable test score (>85%). The ability to select features through machine learning will allow our bid to seriously compete on price and allow for a shortened timeline, leaving a positive impression on our potential client and the ability to sell additional work in the future.