Question 2.1
I work for a political campaign. A Classification Model would be appropriate for many different situations, but one good example is in determining whether or not someone would potentially contribute to the campaign. Potential predictor could include income level, previous political giving history, age, occupation, or education levels.

Question 2.2.1
Please review the included 2.2.1.R file along with this write-up, which includes the code I used and comments on what each section is doing.

Using the attached code, I modified the C value from a range of .00001 to 1,000,000. Changes to this value changes the models tradeoff between minimizing classification error and maximizing margin. I chose a C value of 100,000 because it resulted in the minimum classification error with the highest margin. Margin for this model is 1.986 and accuracy is 86.39%.

The equation for the classifier at this C value (rounded to 3 decimals to save space) is:

$$(-0.004 * X1) + (-0.088 * X2) + (0.130 * X3) + (-0.083 * X4) + (0.988 * X5) +$$
$$(0.031 * X6) + (-0.056 * X7) + (-0.037 * X8) + (0.022 * X9) + (0.019 * X10) + .081 = 0$$

Because changes in C values over a large range (between .001 and 100,000) have minimal impact on the model's output, we can guess that the data is likely well separated and we don't have to make much tradeoff between a large margin and avoiding mistakes.

Question 2.2.2
I tried a few different kernels, but found the highest accuracy with the radial basis function kernel ("rbfdot"). As shown in the attached 2.2.2.R file, prediction accuracy is over 99.5% with this kernel with a C value of 10000. This is almost certainly overfitting to a large degree. We are using a non-linear kernel which increases the model's ability to separate linearly non-separable features. Additionally, by using a high C value, our model is favoring minimizing error over margin increasing the likelihood of overfitting. Most importantly, by testing accuracy on the same dataset we used for training we have no way of validating that the model is not overfit to the training data. It would likely perform worse on a new data set.

Question 2.2.3