Classifying Spam using URLs
Di Ai
diai@Stanford.edu

Model

1. **Linear SVM**
   - Train Cutoff: 50
   - Train Accuracy: 88.7%
   - Test Accuracy: 88.0%

2. **RBF SVM**
   - Train Cutoff: 50
   - Train Accuracy: 84.6%
   - Test Accuracy: 90.0%

3. **Random Forest**
   - Train Cutoff: 50
   - Train Accuracy: 90.4%
   - Test Accuracy: 91.1%

*Categorical variables were one hot encoded with a cutoff frequency threshold*