

# Mapping the Brain: An Introduction to Connectomics Creating Classifiers for the Identification of Gender in Brain Graph Data

Chris Micek, Addison Wright, Monica Rodriguez-Fernandez

January 21, 2016

## 1 Abstract

An abundance of MR brain graph data has recently been acquired via the m2g pipeline, which associate subjects with covariates such as age and gender. We focus specifically on the Southwest China Data set. However, these data have yet to be analyzed; moreover, feature importance in graph classification is not well established. Here we present an ensemble of MATLAB classifiers implemented to classify MR brain graph data from the Desikan atlas based on gender, as well as features deemed important by our linear classifier. We see that simply vectorizing the graphs and treating each one as a high dimensional point proves to be a powerful method when trying to classify gender. Possible future work may entail optimizing each classifier, attempting to classify based on non-binary covariates, extrapolating this methodology to other atlases and experimenting with other feature selections given a particular graph.

## 2 Results

After analysis the following was very clear: Weighted Graphs allow for increased classification rates; Treating the graphs as "full" matrices vs. "sparse" matrices considerably improves classification rates in MATLAB; Certain classifiers perform better given Binary or Weighted graphs. Binary Ranks: SVM, Voting, KNN, Trees, Linear, Adaboost  
Weighted: Trees and Adaboost (tie), Voting, Linear, KNN, SVM.

### 3 Figures

#### 3.1 Binary

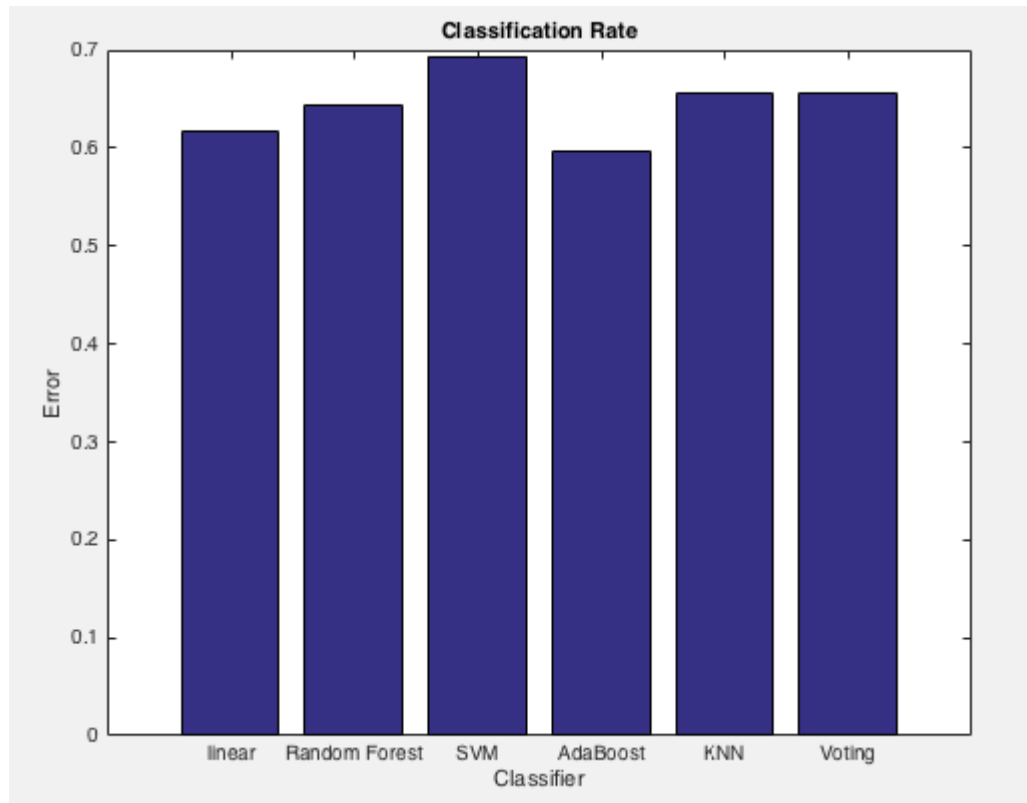


Figure 1: Classification rate for binary data.

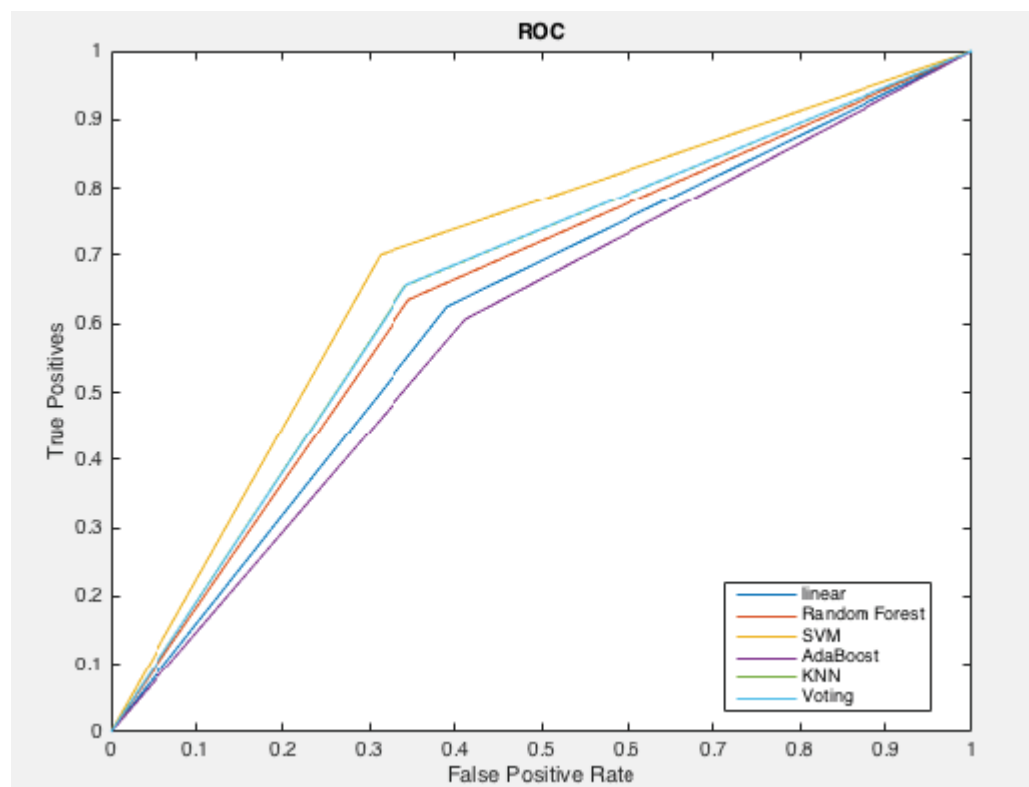


Figure 2: ROC curve for binary data.

### 3.2 Weighted

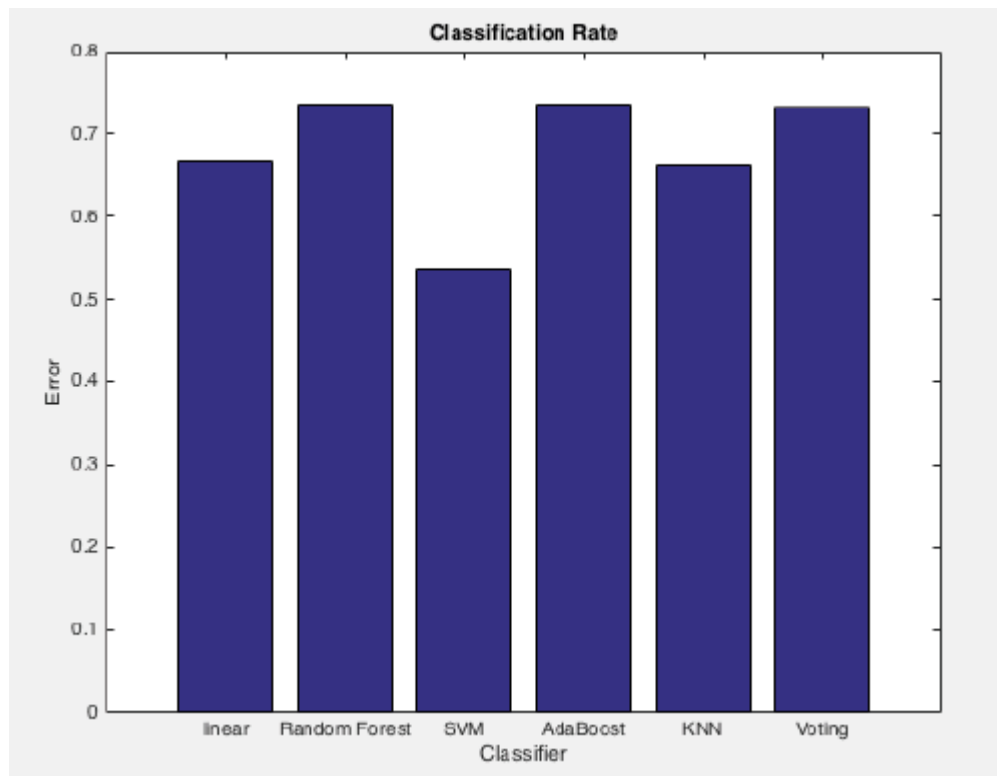


Figure 3: Classification rate for weighted data.

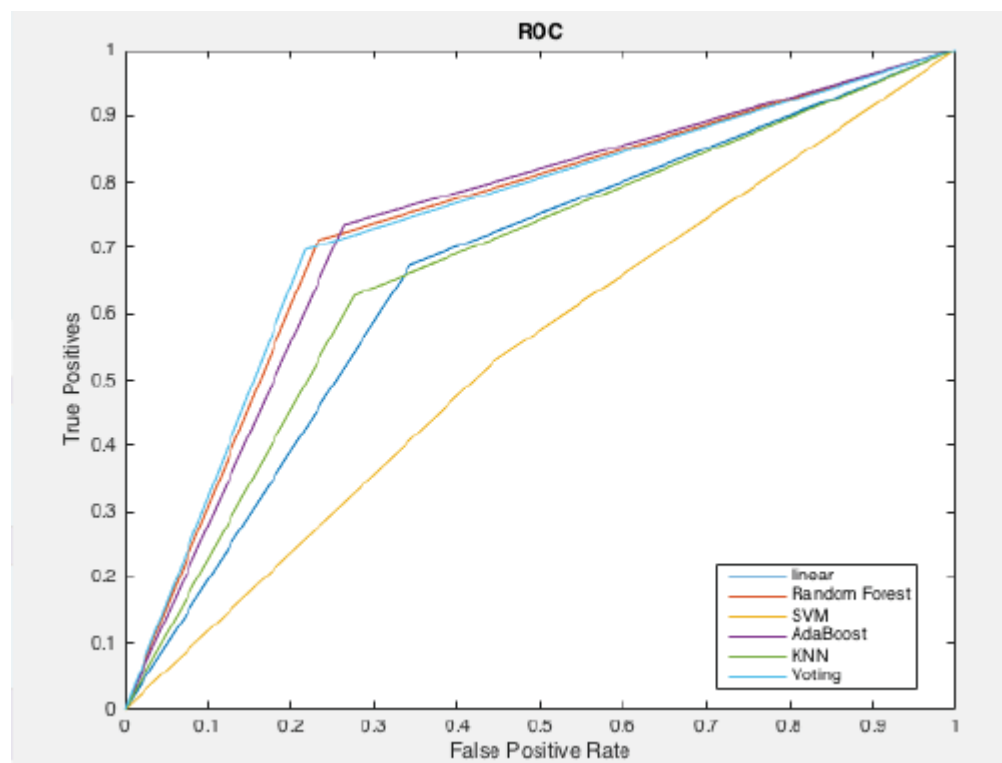


Figure 4: ROC curve for weighted data.