Sensitivity of Jury Trial Outcomes to Trial Factors

Joseph M. Abruzzo
Department of Psychology, Stanford University

ABSTRACT

This project presents a novel machine learning approach to feature impact analysis in the context of American criminal jury trials. Specifically, four data mining algorithms were used to rank nineteen characteristics of jury trial procedure according to their influence on trial outcomes. The algorithms were implemented using various packages in the statistical computing language R. Implications, limitations, and directions for future research are discussed.

MOTIVATION

- Previous research has focused on two types of analysis: behavioral studies and data set analysis:
  - Behavioral studies are limited in scope – focus mainly on juror / defendant demographics
  - Data set analysis has been very crude – rarely more complex than linear regression
- The motivation for this project is to extend the scope of analysis to trial procedure characteristics and update the method of analysis with more sophisticated machine learning methods

DATA SET

- Data set: *Evaluation of hung juries in Bronx County, New York, Los Angeles County, California, Maricopa County, Arizona, and Washington, DC, 2000-2001 (ICPSR version)*
- Acquired from National Archive of Criminal Justice Data at the University of Michigan

ANALYSIS

- Four methods were used to rank features according to their impact on trial outcomes
- LOGISTIC CLASSIFICATION WITH LASSO
  - An optimal regularization parameter was found via 5-fold CV (See *Figure 1*)
  - This process left a model with three nonzero coefficients that were used to rank feature impact
- PENALIZED SUPPORT VECTOR MACHINE
  - An optimal tuning parameter was found via 5-fold CV
  - This process left a model with fifteen features – these features were ranked in impact according to their resultant weights
- RANDOM FOREST & OAT
  - A random forest classifier was fit to the data set – features were then left out one at a time (OAT) and then ranked in impact by the increase in misclassification error that resulted
- NEURAL NETWORK
  - A recurrent neural network with 16 hidden nodes was fit to the data
  - Features were ranked according to their weights in the trained network

RESULTS

Table 1
Output of ranking methods and corresponding feature impact rankings

Rank	Logistic & LASSO	Penalized SVM	Random forest & OAT	Neural network
1	`jurySubQuest` [0.0897]	`witnessDefen` [-0.2106]	`witnessDefen` [0.0154]	`numJurors` [120.72]
2	`witnessDefen` [-0.0794]	`juryAwareSentence` [0.1019]	`defenExhibit` [0.0124]	`prosExhibit` [72.24]
3	`expWitnessPros` [0.0081]	`guidedConduct` [-0.0678]	`jurySubQuest` [0.0077]	`witnessPros` [66.23]
4	N/A	`trialLength` [0.0546]	`struckDefen` [0.0048]	`struckDefen` [56.90]
5	N/A	`struckPros` [-0.0308]	`numJurors` [0.0036]	`witnessDefen` [53.34]

FIGURES

Figure 1

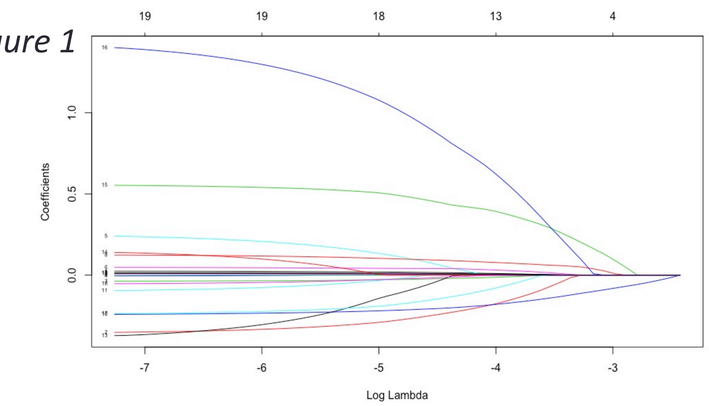
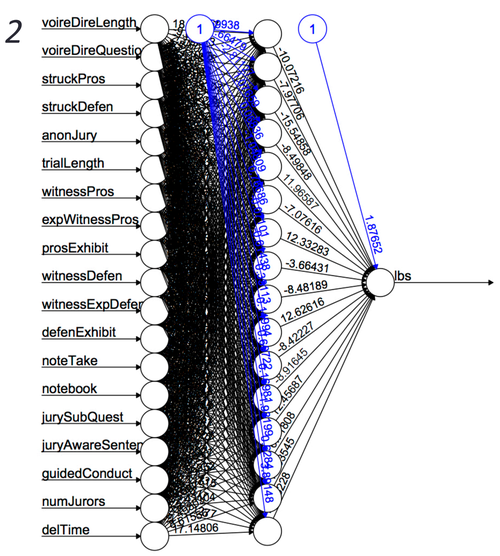


Figure 2



CONCLUSION

- The features with the most impact were sometimes unexpected
- The features least relevant to evidence strength sometimes ended up ranked high
- Results indicate profound influence of bias in jury trial process