On the Creation of the BeSiVa Algorithm to Predict Voter Support

By

Benjamin Joseph Rogers

Submitted to the graduate degree program in Political Science and the Graduate Faculty of the University of Kansas in partial fulfillment of the requirements for the degree of Master of Arts.

_____

Chairperson Dr. Paul Johnson

_____

Dr. Alesha Doan

_____

Dr. Mark Joslyn

Date Defended: December 16, 2014

The Thesis Committee for Benjamin Joseph Rogers

certifies that this is the approved version of the following thesis:

On the Creation of the BeSiVa Algorithm to Predict Voter Support

_____

Chairperson Dr. Paul Johnson

Date approved: December 16, 2014

Abstract

An algorithm was created for a political campaign in Fall 2014 to help find the best way to take voter data and use it to predict the likelihood of voter support. Survey workers collected individual surveys from potential voters in the 2014 Kansas Second House District race. Surveys were matched against data provided by the party in an attempt to predict candidate support.

This thesis compares different classifying methods that can narrow down relevant independent variables to the algorithm developed in part with input from the client, known as the Best Subset with Validation Algorithm (or BeSiVa). BeSiVa focuses on choosing independent variables based on prediction rates in a separated pseudo-test set. With a percent correctly predicted comparable to the other methods, BeSiVa dealt with missing data, created a more interpretable model with fewer predictors, and predicted support with a comparable accuracy to other methods on bootstrapped subsets of campaign data.

Acknowledgments

I'd like to thank my parents, Jean Memken and Michael Rogers, for their financial and emotional support throughout writing this thesis, as well as my sister Sarah and brother Jonathan. I also want to thank my thesis chair Dr. Paul Johnson and the members of my committee, Dr. Alesha Doan and Dr. Mark Joslyn. I'd also like to say thank you to Chris Petkus, who convinced me that I could actually write this. This thesis would have been impossible without the intervention of my employer Mark Sump, and everyone at the Wakefield Campaign who helped collect the data.

Table of Contents