SuperLearner Package

Thomas Carpenito

1. Abstract
2. Introduction
   1. What is machine learning
      1. Mention the current state of machine learning techniques in programming.
      2. What industries currently use machine learning
      3. What is SuperLearner
         1. Brief
   2. Current Status of machine learning in Stata
      1. Cite and list current data adaptive packages
   3. What this project hopes to do
      1. Provide a set of tools for individuals who wish to use machine learning techniques
      2. Encourage others to develop machine learning techniques in STATA.
   4. What this package will not do
      1. Absolve statisticians of critically thinking about their models
   5. Who this package will benefit
      1. Epidemiologists
      2. Researchers
      3. Everyone!
3. SuperLearner
   1. Discrete SuperLearner
      1. Theory behind how the discrete SuperLearner works
   2. SuperLearner
      1. Theory of how SuperLearner works
   3. Caveats of SuperLearner
      1. Does not absolve mindless thinking
4. Package
   1. Syntax
      1. Mandatory Requirements
      2. Options
   2. Saved Results
      1. Depends on output
   3. Github and documentation
5. Examples
   1. Simulation
   2. Example for Discrete SuperLearner
      1. Model selection example
   3. Example for SuperLearner
   4. Example for Prediction
6. Discussion
   1. Difficulties with stata
   2. Future Directions
      1. Restate introduction
   3. What needs to be added to the package
      1. More parameter estimates for customization
7. References
   1. Theory papers
   2. Past packages I’ve used