Southern New Hampshire University

DAT 640 Practical R Activity Seven Guidelines and Rubric: Random Forest Analysis

Using the randomForest package, follow the instructions in the Tutorial Example in Section 12.4 of the textbook to build a random forest model using Rattle. Visit this site and select one of the following data sets:

- MASS **Boston** (Target: medv [median value of owner-occupied homes in \$1000s.])
- MASS **birthwt** (Target: low [indicator of birth weight less than 2.5 kg] **(Note: In the Rattle Data tab, set the variable X to Ident and the variable bwt to Ignore. Remember to Execute after you make these changes.)**
- Boot **urine** (Target: r [Indicator of the presence of calcium oxalate crystals.])
- Cluster **plantTrait** (Target: any of the binary factors)
- Ecdat **HDMA** (Target: deny [mortgage application denied?])

Download the data set and read it into Rattle or R. Run a decision tree on the data set. This tree may be run in R or Rattle. Provide a summary of the results. Next, complete a random forest analysis on the set and provide a summary of the results. You will be providing a narrative of your evaluation of your models and the differences between the two approaches in the discussion post.

Guidelines for Submission: Your submission should follow these formatting guidelines: double spacing, 12-point Times New Roman font, one-inch margins, and citations, if any, in APA format when appropriate.

Critical Elements	Exemplary (100%)	Proficient (90%)	Needs Improvement (70%)	Not Evident (0%)	Value
Demonstration of R Command Execution	Includes demonstration of all R command executions in the form of screenshots or command output listings	Includes demonstration of most of R command executions in the form of screenshots or command output listings	Includes limited demonstration of R command executions in the form of screenshots or command output listings	Does not include demonstration of R command executions in the form of screenshots or command output listings	15
Required Elements	Meets "Proficient" criteria and provides additional analyses or relevant supporting scholarly material	Submission includes all the required elements of the analysis	Submission includes 50% or more, but not all, of the required elements of the analysis	Submission includes less than 50% of the required elements of the analysis	15
Model Computation	Meets "Proficient" criteria and includes assessment of alternative settings/tuning procedures	Demonstrates accurate utilization of the R/Rattle packages per the model specifications	Demonstrates utilization of the R/Rattle packages with inaccuracies or does not adhere to the model specifications	Does not demonstrate utilization of the R/Rattle packages	30
Model Analysis	Meets "Proficient" criteria and draws insightful conclusions that are thoroughly defended with evidence and examples	Provides in-depth analysis that demonstrates complete understanding of concepts and draws informed conclusions that are justified with evidence	Provides an analysis that demonstrates a general understanding of concepts or draws logical conclusions, but does not defend with evidence	Does not provide in-depth analysis	30

Southern New Hampshire University

Articulation of Response	Submission is free of errors related to citations, grammar, spelling, syntax, and organization and is presented in a professional and easy-to-read format	Submission has no major errors related to citations, grammar, spelling, syntax, or organization	Submission has major errors related to citations, grammar, spelling, syntax, or organization that negatively impact readability and articulation of main ideas	Submission has critical errors related to citations, grammar, spelling, syntax, or organization that prevent understanding of ideas	10
				Earned Total	**100%**