IST707 Data Analytics

HW1: Association Rule Mining

Due: 11:59pm, Sept 22th, 2019

Homework instructions

- Analyze *employ_attrition.csv* dataset provided
- Follow CRISP-DM process
  - Data exploration and apply descriptive statistics
  - Data preprocessing, cleaning, transformation: identify potential data quality issues and fix them
  - Run association rule mining algorithm using default settings
  - Fine tune the model by experimenting with different algorithm parameters
  - Output and present the most interesting and significant rules which could predict *"Attrition"*
  - Provide interpretations of the above chosen association rules and also discuss why you consider them interesting and significant
- Use Rmarkdown to structure your report and submit the html output
  - All the codes and relevant outputs (limit the size of outputs to only include those relevant contents and refrain from printing out excessive amount of irrelvant information or data)
  - Analysis writeup (interpretation and discussion of the results with the proper session titles and all the information useful to grade your work)
- Develop a Shiny app to host the analytics process and upload to shinyapps.io
  - Instruction of uploading R app to shinyapps.io (https://shiny.rstudio.com/articles/shinyapps.html)
  - Include the link in your html submission
  - Provide the appropriate control widgets in the app UI that allows app users to choose different values of model parameters (e.g. support and confidence)
  - Output top associate rules according to users' choices
  - Include one visualization to plot the association rules on 2D space defined by the association rule performance metrics

Grading rubrics

- Rmarkdown report (60%)
  - Include all the key data mining steps which are neatly structure in the report with both R codes and relevant otputs (using proper section titles) (40%)
  - In-depth interpretation of the analysis output (20%)
- R shiny app (40%)
  - A functioning web app that meets all the specification as in this instruction (30%)
  - Allow the maximal level of user interactions and return properly formatted analysis results to the browser (10%)

Submission instructions:

- Submit the Rmarkdown html report (which includes the shiny app link) to the blackboard
- Deadline: 11:59pm, Sunday, Sept 22th, 2019
- Late submission policy: Late submission will incur 20% penalty for every additional 24 hours' delay until all points are deducted