More Predictive Models in Text Analytics

Dr. Goutam Chakraborty

Outline

- Describe other ways to handle prediction (document categorization) problems in SAS Text Miner.

- Demonstrate models such as regression, tree, neural net for document categorization problems using driver feedback data.

Other Ways (than Text Rule Builder) to Handle Document Classification Problems

- Use a traditional predictive model with input variables from Text Cluster nodes, Text Topic nodes, or both.

- Here are some of those variables that may be used as input in any predictive model:

  - Text Cluster node SVD variables
  - Text Cluster node segment variables
    - Level and Role must be changed to a nominal input
  - Text Topic node raw SVD variables
  - Text Topic node binary topic class labels
    - Role must be changed from segment to input variables

Procedure

- Follow handout titled "Demo with Driver Feedback"