Text Analytics and Sentiment Mining Using SAS

Big data: it's unstructured, it's coming at you fast, and there's a lot of it. In fact, the majority of big data is unstructured and text oriented, thanks to the proliferation of online sources such as blogs, e-mails, and social media. While the amount of textual data are increasing rapidly, businesses' ability to summarize, understand, and make sense of such data for making better business decisions remain challenging. No marketing or customer intelligence program can be effective today without thoroughly understanding how to analyze textual data. Emphasizing practical skills as well as providing theoretical knowledge, this hands-on course takes a comprehensive look at how to organize, manage, and mine textual data for extracting insightful information from large collections of documents and using such information for improving business operations and performance.

Speaker: Goutam Chakraborty, PhD

Has been a professor of marketing at Oklahoma State University for more than 17 years. During that time he has taught data mining and CRM applications, database marketing, new product development, marketing research, digital business strategy, Web business strategy, electronic commerce, interactive marketing, and product and pricing management. Chakraborty founded the SAS and OSU Data Mining Certificate program.

Schedule: 12-13 June, 2014

Venue: SAS Institute Sdn Bhd, Menara CIMB KL Sentral

Fee: RM7,500/pax
Objective

- gain a deep understanding of tools and techniques of text analytics and sentiment mining from statistical and NLP (natural language processing) perspectives
- import text data into SAS Text Miner from different sources and different formats
- create clusters from text data to understand customer segments
- derive topics from text data to better understand customer conversation
- create rules from text data to make predictions
- combine text data with numeric data to build better models
- create statistical, rule-based, and hybrid models for understanding and predicting customer sentiments
- create statistical and rule-based models to categorize documents into a specified taxonomy.

Target Audience

Business analysts, web analysts, BI professionals, customer intelligence professionals, data analysts, market researchers, marketing analysts, social media analysts, text analysts, and data miners who want to learn how to effectively use text data to generate customer insights and to understand and predict customer sentiments.

Pre requisite

Some experience with SAS and SAS Enterprise Miner is useful, but it is not mandatory. No experience with text analysis is necessary.

Course Outline

Text Analytics and Sentiment Mining

- history and roots of text analytics
- SAS tools and components for text analytics
- application areas of Text Analytics

Getting Textual Data into SAS Enterprise Miner and SAS Text Miner

- introduction to SAS Enterprise Miner and SAS Text Miner
- demonstration of the Text Import node for Web crawling
- demonstration of the File Import node for reading Excel files
- demonstration of XML mapper for reading XML files
- self-study: a brief introduction to SAS IR Studio
- hands-on exercises to enhance learning all of the above concepts

Text Parsing and Using Term by Document Matrix

- bag-of-words versus NLP approach in parsing
- tokenization, lemmatization, POS Tags, phrase and entity recognition
- weights and transformations used in handling term-by-document matrix
- demonstration of text parsing, text filtering, text search, and concept links
- hands-on exercises to enhance learning all of the above concepts
Clustering and Topic Extraction

- different types of similarity metrics and methods used in clustering
- different clustering algorithms available in SAS
- SVD and LSI for clustering textual data
- differences between the Text Cluster node and the Text Topic node
- demonstration of Text Cluster and Text Topic extraction
- hands-on exercises to enhance learning all of the above concepts

Predictive Modeling and Text Rule Builder

- predictive modeling essentials for text and numeric data
- demonstration of Text Rule Builder node for model building using text data
- demonstration of the Text Cluster node and the Text Topic node for model building using text data
- demonstration of model building using both numeric and textual data
- hands-on exercises to enhance learning all of the above concepts

Sentiment Analysis and Opinion Mining

- basics of sentiment analysis
- architecture and types of models in SAS Sentiment Analysis Studio
- demonstration of statistical and rule-based model building in SAS Sentiment Analysis Studio
- hands-on exercises to enhance learning all of the above concepts

Content Categorization, Concept Extraction, Wrap-up, and Takeaways

- basics of content categorization
- architecture and types of rules supported in SAS Content Categorization Studio
- demonstration of different models and rules in SAS Content Categorization Studio
- hands-on exercises to enhance learning all of the above concepts
- self-study: LITI definitions
- course wrap-up and takeaways

Software Addressed

This course addresses SAS Text Analytics, SAS Text Miner, SAS Enterprise Miner software. The demonstrations and exercises in this course primarily use SAS Enterprise Miner, SAS Text Miner, SAS Sentiment Analysis Studio, and SAS Content Categorization Studio.