

[SAS-DMTX41]: Text Analytics Using SAS Text Miner

Length : 2 Days
Delivery Method : Instructor-led (Classroom)

Course Overview

This course describes the functionality of SAS Text Miner software, which is a separately licensed component that is available for SAS Enterprise Miner. In this course, you learn to use SAS Text Miner to uncover underlying themes or concepts contained in large document collections, automatically group documents into topical clusters, classify documents into predefined categories, and integrate text data with structured data to enrich predictive modeling endeavors.

Audience Profile

Statisticians, business analysts, and market researchers who incorporate free-format textual information in their analyses; managers of large document collections who must organize and select documents using data mining; and students of data mining who want to learn about text mining

At Course Completion

- Convert documents stored in standard formats (Microsoft Word, Adobe PDF, and so on) into general purpose HTML or TXT formats
- Read documents from a variety of sources (web pages, flat files, data elements in a relational database, spreadsheet cells, and so on) into SAS tables
- Process textual data for text mining (for example, correct misspellings or recode acronyms and abbreviations)
- Convert unstructured text-based character data into structured numeric data
- Explore words and phrases in a document collection
- Query document collections using keywords (that is, identify documents having specific words or phrases)
- Identify topics or concepts that appear in a document collection
- Create user-influenced topic tables from scratch or by modifying machine generated topics or concepts using domain knowledge
- Use derived topic tables or pre-existing user-influenced topic tables (or both) to enhance information retrieval and document classification
- Cluster documents into homogeneous subgroups
- Classify documents into predefined categories.

Pre-Requisites

Before attending this course, you should.

AVANTUS TRAINING PTE LTD

80 Jurong East Street 21 #04-04 Devan Nair Institute Singapore 609607
Main Line: +65 6661 0888 | Fax: +65 6661 0886
Email: enquiries@AvantusTraining.com
www.AvantusTraining.com

- be acquainted with Microsoft Windows and Windows-based software
- have at least an introductory-level familiarity with basic statistics and regression modeling.

Previous SAS software experience, especially SAS Enterprise Miner experience, is helpful but not required.

This course addresses SAS Text Miner software.

This course uses SAS Text Miner 14.1 and SAS Enterprise Miner 14.1.

Course Outline

Module 1: Introduction to SAS Enterprise Miner and SAS Text Miner

Lessons

- Data mining and text mining
- Working with data sources
- Using SAS Enterprise Miner and SAS Text Miner

Module 2: Overview of Text Analytics

Lessons

- Using the Text Import node, adding a target variable, and comparing models
- A forensic linguistics application
- Information retrieval
- Text categorization

Module 3: Algorithmic and Methodological Considerations in Text Mining

Lessons

- Methods for parsing and quantifying text
- Dimension reduction with SVD

Module 4: Additional Ideas and Nodes

Lessons

- Some predictive modeling details
- Text Profile node
- High Performance (HP) Text Miner node

AVANTUS TRAINING PTE LTD

80 Jurong East Street 21 #04-04 Devan Nair Institute Singapore 609607

Main Line: +65 6661 0888 | Fax: +65 6661 0886

Email: enquiries@AvantusTraining.com

www.AvantusTraining.com