Text Analytics and Analyzing Patient Satisfaction in SAS Enterprise Miner 12.1

Divya Sridhar, Advisory Analytics, Deloitte & Touche LLP

Abstract

A survey is often the preferred way to obtain information and feedback for program improvement. At a bare minimum, survey research comprises of economics, statistics, and psychology in order to show how surveys measure and predict important aspects of the human condition. Through healthcare-related research papers written by experienced practitioners in the industry, a list of several factors that are important for patients during hospital visits were hypothesized. Through text mining using SAS Enterprise Miner 12.1, tangible aspects of hospital quality and patient care were measured using online survey responses and comments found in the English National Health Service website. Patient comments were used to analyze positive, negative, or mixed feedback and whether it correlates with research done on hospital quality.

The implications of this research are vital in comprehending our health care system and the factors that are important in satisfying patient needs. Analyzing survey responses can potentially reduce disparities in health care services provided to population subgroups in the United States. Starting with online survey responses can provide leading practices to understand the overall methodology and motivation needed to analyze the surveys that have been created in agencies like the Center for Medicare and Medicaid Services (CMS).

Methodology

CLASSIFY
- Leverage clinical sources to derive a taxonomy for hospital quality
- Present survey responses to build a sampling frame of patient opinions on hospital environment

ANALYZE
- Identify intangible characteristics by building target based term weights
- Measure significance of intangible characteristics for patients

EVALUATE
- Show the strength of transforming unstructured qualitative data to structured quantifiable data
- Understand the link between the text mining tool and pay-for-performance metrics in the healthcare industry

Overview of Methodology

SAS Enterprise Miner’s Text Feature is used to parse, filter, and transform patient survey data, into structured results that may be used to analyze patient satisfaction.

Inputs
- Patient Surveys
- Preprocessed Data

Outputs
- Parse Unstructured Survey Data
- Filter using Taxonomies
- Transform Unstructured Data into Structured Form
- Reduce and Combine Inputs
- Analyze and Extract Value
- Produce Results and Present

Text Mining Methodology

Copyright © 2016 Deloitte Development LLC. All rights reserved
The healthcare industry is approaching an exciting paradigm shift in patient engagement as we move away from being a patient-centric industry to a patient-driven one. After collecting recognized literature on the subject, we were able to build a taxonomy of weighted terms used to underscore qualities important to patients based on demographics (i.e., race, gender, age).

Concept linking is a way to find and display the terms that are highly associated with another term. The selected term is surrounded by terms that correlate strongest with one another.
Correlation Matrix: Hospital Qualities based on Race

The heat-map below shows the statistical relationship between hospital quality traits and race.

Scatterplot: Information, Communication, and Education (Doctors and Nurses) Importance in Males and Females

Measuring Quality Factors in Seniors vs. Young Adults

Seniors (over 60)

Young Adults (21-30)

Patients who are senior citizens tend to really value emotional support and involvement of friends and family. Young adults tend to value physical comfort (i.e., pain management techniques, medication, therapy recommendations, etc.) followed closely by information, communication, and education of both the doctors and nurses.

Data Source for Analysis: English National Health Service Database
**SAS Topic Miner**

Access to “test results” might be an underlying cause for the overall improvement of the assessment. Therefore negative forms of the comment data are looked at in order to evaluate if there were fewer negative comments about “test results” after the hospital visit. Topic Miner is used to extract the themes of relevance from the negative text data.

**SAS Assessment Node**

This node is used to compare various steps used to convert unstructured to structured data and explain the process flow of this methodology. Decision Trees and Regression models are just some of the assessments that may be performed.

**Snapshot of Raw Data for Sentiments Collection**

- **Positive Statement**: An excellent service in a clean efficient hospital
- **Negative Statement**: Terrible Treatment
- **Mixed Statement**: Good but staffing level was too low.

Sentiment analysis involves user-written rules efficiently capturing sentiment based on adjectives, adverbs, and verbs. These are used in the rules which define the positive or negative sentiment.

---

**Copyright © 2016 Deloitte Development LLC. All rights reserved**