SAS® Sentiment Analysis

Automatically pinpoint sentiment from the Web and internal electronic documents to understand trends and develop effective strategies

What does SAS® Sentiment Analysis do?

SAS Sentiment Analysis automatically rates and classifies opinions expressed in electronic text. It collects text inputs from websites, social media outlets and internal file systems, and then puts them in a unified format to assess relevance to predefined topics. Reports identify trends or emotional changes, and an interactive workbench allows subject-matter experts to refine sentiment models. The solution automatically scores input documents as they're received, providing real-time updates about sentiment changes.

Why is SAS® Sentiment Analysis important?

By assessing sentiment in real time, the software provides quantified insights into the overall impressions people have of products, services and brands. Harnessing statistical techniques and linguistic rules, its superior results can quickly reveal consumer opinions, trends and potential problems – helping you evaluate competitive standing, adjust tactics and prevent problematic issues from escalating.

For whom is SAS® Sentiment Analysis designed?

SAS Sentiment Analysis is designed for marketing, public relations, service and development departments charged with enhancing the customer experience, evaluating new products or managing logistical operations for improvement.

In today's economy, it's vital to quickly understand opinions and experiences across channels – websites, communication centers, emails, forms, surveys, internal files and reports. But traditionally, the process of analyzing electronic text to discover positive and negative emotions has been very labor-intensive.

SAS Sentiment Analysis speeds this process by scoring input documents as they're received and providing real-time updates to the status of consumer sentiment. By uniting statistical learning with advanced linguistic methods, the solution hones in on the sentiment buried within textual data and provides accurate insight into consumer opinions. Results are based on an entire document and calculated separately for concepts (and their attributes and characteristics). This lends flexibility and thoroughness to your analysis.

Key Benefits

- **Evaluate sentiment and monitor changes over time.** The software automatically extracts sentiments in real time or over a period of time, with a unique combination of statistical modeling and rule-based natural language processing techniques. Built-in reports show patterns and detailed reactions.

- **Identify feedback sources to define new targets.** By actively monitoring internal collections (such as call centers and the Web) combined with social networking sites (using predefined templates for Twitter and Facebook), the software shows where you're being discussed and what is being said. It automatically extracts feedback as content is crawled, filtering out the most important concepts so you can pursue promising opportunities.

- **Continuously improve customer experience and competitive position.** The software searches for and evaluates internal and external content about your organization and competitors, identifying positive, negative, neutral and "no sentiment" texts – quantifying perceptions in the market.

- **Promote ongoing discovery with a closed-loop, integrated analysis environment.** With ongoing evaluations, you can refine models and adjust classifications to reflect emerging topics and new terms relevant to your customers, organization or industry.

Figure 1: Graphics and user-friendly reports readily describe sentiment insights.

Solution Overview

SAS Sentiment Analysis delivers superior results for understanding opinions because it harnesses the power of statistical techniques and linguistic rules to provide more accurate assessments of sentiment extracted from electronic text documents. It provides detailed feature and attribute evaluations through highly configurable linguistic rules.

Use the software to monitor and assess consumer sentiment through rich visualizations, reports and monitoring tools that make it easier to identify issues and potential problems, evaluate competitive standing and update measures. With this insight, organizational teams that monitor product capability, branding effectiveness and servicing gain feedback that's useful to marketing, research and development, as well as operations and logistics.

With these results, you can identify emerging issues before they become larger-scale problems. You can also monitor opinions after making adjustments and changes to process improvements, and gauge the impact of promotions and business communications.

Integrated with the SAS Business Analytics Framework, sentiment evaluations can be included within structured predictive modeling for insight into what will happen next. And by using SAS Business Intelligence, business users can examine sentiment results in their familiar reporting environment.

Unique hybrid approach to sentiment analysis

Subject-matter and business expertise can be applied in addition to the rigor of statistical approaches to best classify textual information with SAS Sentiment Analysis. Subject-matter expertise alone is effective but not necessarily efficient, available or economical. Statistical models alone are less useful to identify the nuances of the features and attributes described in the text. Typically, a combination of statistical methods/machine learning with subject-matter/business expertise will create the most powerful and efficient sentiment models.

Multiple-level taxonomies

With SAS Sentiment Analysis, you can extend to multiple taxonomy levels to assess the expressed opinion of:

- The overall document.
- A concept in the document (e.g., a brand or product).
- An attribute or feature of the product, brand or concept.
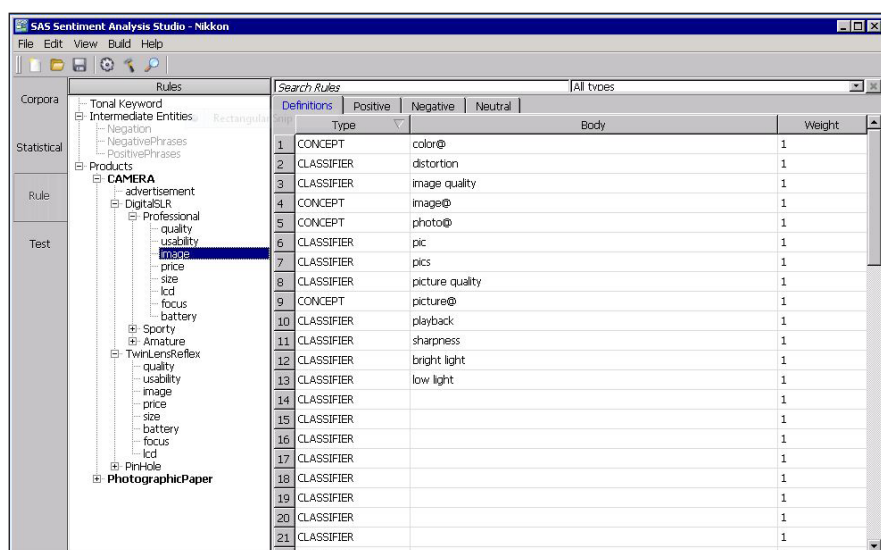- A characteristic of a feature (a subattribute).

Linguistic rules can be configured at each level – providing the detailed knowledge needed to understand what's at the core of the expressed sentiment. In turn, you can easily identify what actions are most relevant in any associated response.

Extended sentiment evaluation

At times, statements are made without emotional intent. SAS Sentiment Analysis enables you to specify these terms as unclassified to distinguish them from comments that are positive, negative or neutral. Unclassified categories can assist with model refinement – because this forces you to decide if there really is no expressed opinion or whether your model needs improvement to uncover the sentiment that exists. Another advantage is that recognizing unemotional statements – particularly if you are seeking to elicit sentiment – can help refine activity and communications.

Context examined for more accurate interpretations

SAS Sentiment Analysis searches and evaluates phrases, assigning weights to word and phrase matches to determine the optimal concept match for any particular document or source. The software also supports complex linguistic rules for one or several matches of a term, regular expressions and part-of-speech tags, along with

Figure 2: Multilevel taxonomies classify sentiment detail to pinpoint issues.

prebuilt, sophisticated Boolean operators expressing constraints, such as the distance and occurrence of concepts in relation to other words.

Dynamic sentiment analysis

Utilizing the power of your scalable hardware, the software can run in real time or in batch mode to scour intranet and Internet websites and to process document collections. Predefined APIs are available for popular social sites, and an interactive markup matcher makes extracting field data from XML or HTML easy. As a result, you can keep a pulse on the sentiments expressed about your organization.

A rules management system is included so you can create your own specialized sentiment rules that can combine the use of Boolean rules and other linguistic tools. The solution permits improvements and changes over time through multiple model generation within the same project. It also lets you edit and test model modifications before automating the server processing.

Extensive native language support, part of an integrated framework

SAS Sentiment Analysis addresses all types of electronic collections, and is readily extendable to natively address multiple languages, so you can evaluate sentiment from across the globe. SAS Sentiment Analysis is part of an integrated analytic framework. Because of this, you can discover patterns for materials with completely unknown topics; integrate your information with SAS categorization and extraction models to extract the sentiment; and use SAS business intelligence capabilities to extend findings to stakeholders throughout the organization.

Key Features

Statistics and linguistics combined to provide more accurate sentiment analysis results

- Provides a choice of approaches to sentiment analysis:
  - Statistical modeling: Provides predefined default parameters – that can also be configured – to identify the document sentiment from text.
  - Linguistic rules: Lets subject-matter experts define the elements to be examined for sentiment assessment.
  - Hybrid approach: Provides the unique ability to use both statistical rigor and linguistic rules to define sentiment models driving more detailed sentiment evaluations.
- Ability to import and/or create concepts for evaluations.

Context of features examined for accurate interpretations

- Supports complex linguistic rules for one or several matches of a term, regular expressions, part-of-speech tags and more.
- Includes prebuilt tasks to simplify linguistic pattern identification. It offers:
  - Predicate rules to define semantic relationships between concepts.
  - Operators to locate related information with greater precision, such as co-referencing.
  - Identification of intermediate concepts that contain rules referenced by other concepts (shortens the rule-writing process).
  - Case-insensitive matching, so you can match both uppercase and lowercase terms.

Dynamic sentiment analysis

- Permits improvements and/or changes over time with multiple model generation within the same project, and the ability to edit and test model modifications.
- Allows subject-matter experts to refine model performance with an easy-to-use point-and-click workbench.
- Monitoring of results over time will inform model analysis refinement, and comparisons to benchmark training sets can be established to allow for continuous improvement.

Easy-to-use interface for model development

- A mechanism to directly upload sentiment analysis models to the server reduces manual model deployment.
- Keyboard shortcuts are available for interactivity, as well as enhanced search functionality assisting with model development activity.
- Project wizard allows you to configure your project while it is being defined.
- Panes in the interface simplify tasks and display more information, such as rule evaluations and search results.

Interactive workbench for model refinement

- Graphics depicting sentiment are displayed to readily identify the resulting classification – as positive, negative, neutral or unclassified.
- Point-and-click exploration of classified text, drilling to detail as desired.
- Enhanced report formatting improves results display.
- Ability to add new concepts/entities to capture desired topics.
- Word cloud report, based on defined concepts, automatically illustrates extracted noun phrases.
- Some workbench functionality is available through Web services APIs so documents can be programmatically added to different projects, exported, managed, searched, etc.

Updates on Web postings, reviews and opinions

- High-performance, multithreaded crawler that can be deployed in a distributed and/or grid mode to maximize processing and support extremely large-scale crawls for both internal file system and Internet crawls.
- Powerful linguistic technology is built in to extract URLs from Java scripts.
- Crawler plugins are available for popular social media sites, including Google, Facebook, Twitter, Bing, BoardReader, Flickr, LinkedIn and Yahoo! scripts.

Key Features (continued)

- Crawls can be interrupted and resumed or incrementally updated. The transversal depth of crawls can also be eliminated.
- Markup matcher provides a point-and-click interface to simplify the extraction of fielded data from HTML or XML, specific to a site, and a stream or batch process to update new documents. Edit and test matches for both XPath and regular expression rules.

Multiple languages supported

- Full suite of 30 native languages (plus dialects) available.
- Supports Arabic, Chinese (both Simplified and Traditional), Czech, Danish, Dutch, English (US/UK), Farsi, Finnish, French (French/Canadian), German (New/Old), Greek, Hebrew, Hindi, Hungarian, Indonesian, Italian, Japanese, Korean, Norwegian (Nynorsk/Bokmål), Polish, Portuguese (Portugal/Brazil), Romanian, Russian, Slovak, Spanish (Spain/South America), Swedish, Thai, Turkish. Ukrainian and Vietnamese.