The Pillars of Text Analytics: Sentiment, Categorization, Effort, and Emotion

A CLARABRIDGE WHITEPAPER

CLARABRIDGE

The Pillars of Text Analytics: Sentiment, Categorization, Effort, and Emotion

A CLARABRIDGE WHITEPAPER

CLARABRIDGE

Contents

Introduction 2

Sentiment analysis. 3

Clarabridge Sentiment. 3

Eleven-Point Scale 3

Negation and Conditional Sentiment 4

Amplification 4

Context-Dependent Sentiment . . . 4

Exception Rules. 5

Categorization 6

Vertical and Horizontal Templates 6

Rules-Based Categorization. 6

Auto Classification 6

The Clarabridge Effort Score 7

Calculating Effort. 7

Business Applications 7

Emotion 9

Emotion as a Measurement of Customer Experience 9

Historical Frameworks for Analyzing Emotion 9

Clarabridge Emotion Analysis . . . 9

Emotion Category Templates . . . 10

Conclusion 11

Introduction

Most competitive businesses recognize the need to examine customer interaction data to interpret customer experience; however, companies may not fully understand how to analyze different types of data, the differences between available analytics solutions, or the importance of choosing the right one.

Text analytics provide a highly effective means of considering what customers are saying, but not all text analytics functions are equal or provide the same depth of analysis. By taking the time to understand the variations and recent advancements of this popular and highly promoted tool, businesses are equipped to identify the most effective solution.

In this paper, we outline the several factors that contribute to the Clarabridge solution and examine them closely. The report discusses common dimensions such as sentiment and categorization, but also several measures of analysis that are unique to Clarabridge, such as effort and emotion. In fact, **The Forrester Wave™: AI-Based Text Analytics Platforms, Q2 2018** reports that Clarabridge offers "highly differentiated emotion, effort, and intent analysis, while most of its competitors still mainly offer sentiment analysis." We see each of these dimensions as core and invaluable components of the text analytics engine of the platform, allowing users to effectively translate data-driven insights into business acumen. Text analytics as a process can be quite complex, but the following descriptions and definitions can provide practical context to help readers more fully understand the value of text analytics and how the Clarabridge platform is best positioned to help businesses understand the voice of the customer.

The Forrester Wave™: AI-Based Text Analytics Platforms, Q2 2018 reports that Clarabridge offers "highly differentiated emotion, effort, and intent analysis, while most of its competitors still mainly offer sentiment analysis."

Sentiment Analysis

One important component of text analytics is the ability to determine **sentiment**, the positivity or negativity expressed through the text. As with language processing, there are multiple ways to approach sentiment analysis. These approaches vary with respect to both efficiency and accuracy.

TRADITIONAL APPROACHES

Manual coding is the process by which a person reads a document and assigns it a sentiment value (usually "positive" or "negative"). This method is time-consuming, imprecise, and subject to human error, and it is impractical when dealing with large amounts of feedback.

"Bag of words" is a method that looks at the feedback document and assigns positive, negative, or neutral values to each word based on a predefined, built-in dictionary of words that bear sentiment. This method is prone to inaccuracy because it doesn't account for how the words relate to each other. There is also a risk of ambiguity because simply detecting "emotion" words may indicate whether a sentence is positive or negative but it does not indicate the topic about which that sentiment is directed.

Simplistic sentence-level scoring is how most solutions categorize and provide sentiment analysis, but this approach is imprecise because it does not dig deeply into the greater context of an expression. For example, for the statement "I liked the food, but the waiter was rude," sentence-level scoring would likely assign this sentence a positive sentiment value based on the first clause. It misses out on the complete picture due to its inability to separate and individually analyze each unique idea, thought, and phrase.

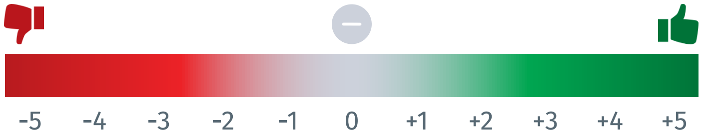
CLARABRIDGE SENTIMENT

Clarabridge overcomes the limitations of other sentiment analysis techniques by combining lexical and grammatical approaches. This system understands how sentiment is altered or intensified by taking into account grammatical constructs and function words. This gives Clarabridge the unique ability to accurately handle negation, conditional sentiment, context-specific sentiment, and exception rules.

ELEVEN-POINT SCALE

Clarabridge indexes its sentiment score on a normalized minus five (-5) to plus five (+5) scale. This provides significantly greater accuracy over a "positive," "negative," or "neutral" rating since we all know that there is a big difference between a good meal (+1) and the best food you've ever eaten (+5). For further analysis, users can perform sentiment filtering to isolate the degree of positivity or negativity and to conduct root cause analysis to determine drivers of the sentiment score.

The Clarabridge Sentiment Scale



NEGATION AND CONDITIONAL SENTIMENT

Negation and conditional sentiment are two ways in which relationships among the words in a sentence can impact overall sentiment.

I am not happy about the repair experience.

Negation: The word "happy" would typically carry positive sentiment; however, because it is paired with "not" in this sentence, the positive sentiment is negated. The Clarabridge sentiment score recognizes this important distinction.

Conditional sentiment: Similar to negation, conditional sentiment refers to statements that may seem positive but are neutralized by phrases such as "kind of" and "sort of." This type of sentiment is calculated using positional rules, which identify when these qualifying phrases are used together. Other examples of conditional sentiment include:

COULD + HAVE + BEEN + [positive] = negative

It could have been a good experience.

COULD + HAVE + BEEN + [bad/problem] = neutral

It could have been a problem, but ...

USED + TO + BE + [positive] = negative

The restaurant used to be cleaner.

AMPLIFICATION

A positive (or negative) phrase can be made more positive (or negative) by modifiers. Clarabridge recognizes the difference these modifiers make. For example, a sentence reading "The food was great" has a +2 positive sentiment score that jumps to a +3 positive sentiment score when written as "The food was so amazingly great."

CONTEXT-DEPENDENT SENTIMENT

Context-dependent sentiment refers to the ability to recognize that a phrase may be positive in certain contexts but negative in others. One example of this is the word "thin." In technology, "thin" things are generally considered good (think of a thin laptop). At the same time, in the hospitality industry "thin sheets" or "thin walls" are negative.

Clarabridge uses predefined rules and a taxonomy to quickly apply fine tuning to industry-specific words. Examples of industry templates include retail, automotive, high-tech, financial services, and restaurants. Meanwhile, horizontal templates include customer service, online experience, cries for help, and social media.

In individual situations, the Clarabridge platform also enables word-level tuning, which is the ability to increase or decrease the predefined sentiment value attached to a specific word based on the particular need. This process is easily done without the need for technical support or linguistic expertise. Sentiment tuning can be applied universally or for a specific project or subset of data only.

EXCEPTION RULES

Exception rules are predefined rules designed to account for different linguistic constructs that consistently change the sentiment of a word or phrase. Consider the following sentence:

The shirts were too orange.

The color "orange" would normally have a neutral sentiment value (a "0" on the 11-point scale); however, adding the word "too" changes the sentiment to a negative. This holds true often enough that it is valuable to have a rule within the tool stating that TOO + [neutral] = negative. Each sentiment exception rule can be turned on and off as needed and is completely customizable.

The Clarabridge platform includes over 500 built-in exception rules for English processing and enables users to easily add their own rules to meet their specific needs.

Categorization

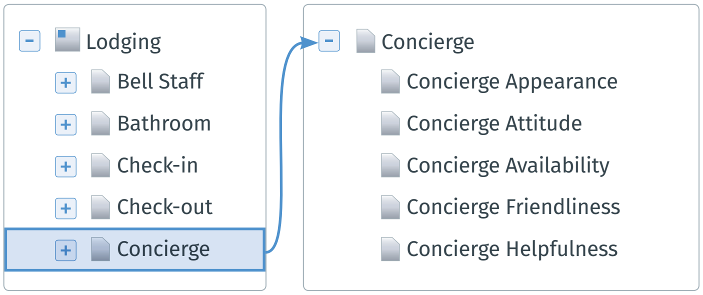
Clarabridge uses both linguistic and statistical techniques to group the data into related buckets. The goal of this process is to categorize as much of the data as possible (known as “high recall”) as accurately as possible (known as “precision”). This method enables easy reporting and promotes discovery of data-driven insights. Clarabridge conducts categorization with unique features including templates, rules-based categorization, and auto classification.

VERTICAL AND HORIZONTAL TEMPLATES

Beyond sharpening sentiment analysis, as discussed earlier, templates provide additional benefits such as helping to categorize textual data more quickly. One or multiple templates can be applied to the same data set, or no template at all, thereby allowing deep-dive analysis in specific areas such as the online shopping experience for cell phones. Furthermore, these templates can be extended and customized based on business needs, and users can build their own models as well.

For example, a Lodging template would classify items related to lodging such as bell staff, bathroom, check-in/out, and concierge. Within each sub-group, the content is further refined to highlight additional categories. All of this organization is done automatically and without manual intervention.

Example: Lodging Categorization Template



RULES-BASED CATEGORIZATION

In addition to offering out-of-the-box templates, Clarabridge allows users to add specific categorization rules. All extracted entities and relationships are displayed so that business users can quickly refine the categorization model. Words can be dragged and dropped into the definition of the category, and additional words are included based on the word stem. For example, if a category was defined to incorporate the word “clean,” the solution would also identify instances of “cleaning,” “cleaned,” and “cleaner” within the data set. It also accounts for misspelled instances of a word such as “claen.”

AUTO CLASSIFICATION

After the built-in templates and rules-based engines categorize the data, there will still be text that doesn’t naturally fit into the existing categories. Auto classification is designed to handle this leftover text and creates categories that can be described as the “other” bucket. By leveraging advanced statistical algorithms, auto classification builds category trees for the leftover data. This capability allows users to load millions of different pieces of text into the system, which Clarabridge will quickly sort. The platform’s ability to build an outline without prompting allows users to understand the initially uncategorized data so that they can see what topics appear and modify existing categories if needed.

The Clarabridge Effort Score

The Clarabridge Effort Score provides an innovative way to examine how much work companies ask their customers to put forth by looking at unstructured customer feedback. Some organizations include a structured effort question in a survey that asks the customer to rate how easy it was to do business with the organization on a numeric scale. However, this approach limits analysis of customer effort to just those survey responses. On the other hand, the Clarabridge Effort Score is derived directly from text, making it a unifying metric that empowers analysis across all text data sources.

CALCULATING EFFORT

The Clarabridge Effort Score is automatically calculated when omni-channel data is ingested and processed through the Clarabridge Natural Language Processing. This AI-powered feature is built into the Clarabridge platform and automatically analyzes effort, allowing businesses to eliminate time spent on tuning and to begin deriving insights immediately. If desired, users can adjust the calculation and identification thresholds of the Clarabridge Effort Score to reflect how effort is measured in a particular industry.

The Clarabridge Effort Score is calculated by a machine learning algorithm that evaluates individual sentences for significant words, phrases, and linguistic features that are commonly found in expressions of effort. The algorithm assigns a whole, non-zero value between -5 (very hard) and +5 (very easy) or null (when no effort indicators are expressed). The scoring system differs from that of Clarabridge's sentiment score, which will

The Clarabridge Effort Scale

assign a zero rather than a null value when no sentiment-bearing words are detected. This null value for effort is not included in averages so as not to dilute the other values; the numeric values are then aggregated in reporting across topics, attributes, and categories.

BUSINESS APPLICATIONS

Measuring effort helps businesses quickly understand issues and develop better design solutions that ultimately make it easier for customers to interact with them.

Examples of how companies can apply insights revealed by the Clarabridge Effort Score include:

- Finding points of high friction and customer confusion
- Discovering drivers of channel hopping
- Creating roadmaps to remove or alleviate drivers of high-effortt experiences
- Determining product flaws, website issues, and opportunities for process improvements
- Developing more-intuitive products and user interfaces
- Identifying and marketing competitive advantages

- Integrating findings with sentiment analysis to identify emerging trends that inform the development of empathetic solutions
- Combining results with emotion analysis to design solutions based on how they want customers to feel

Effort is interesting in isolation but can be more valuable when analyzed in conjunction with emotion, sentiment, satisfaction, and other KPIs. For example, effort analysis aids in discovering areas of customer friction efficiently while emotion analysis can then help explain how these difficulties made the customer feel. Together, they can be used to inform empathetic solution design and promote more customer-centric business decisions. Sentiment and satisfaction scores can be used to track trends over time and monitor the impact of customer experience initiatives, programs, or product changes.

Emotion

In addition to sentiment and effort, emotion is another useful tool when it comes to analyzing the customer experience. Emotion holds significance as a concept that pertains to how customers feel, how we should intentionally design experiences, and how to promote customer loyalty in the business world.

EMOTION AS A MEASUREMENT OF CUSTOMER EXPERIENCE

Analyzing emotions affords a unique lens into experiences in that these emotions explain how customers feel about their engagements with companies. Emotions are fluid and complex; they can change quickly or linger for long periods of time. In fact, many distinct emotions may contribute to a single satisfaction score. Examining and understanding these emotions help analysts empathize with their constituents and think about how specific actions or policies might result in certain feelings; however, emotions are not strictly useful for retrospective analysis.

Companies often aspire to evoke certain emotions from customers as a result of an encounter with a brand or an interaction with an agent. Understanding actual customer emotions and comparing them with desired emotions can help identify opportunities for business improvement and more customer-centric offerings.

Research has also shown that emotion is among the leading indicators of loyalty. Customers who are frustrated, confused, or angry are unlikely to spend more money with a business. In contrast, individuals who are pleased, delighted, or happy with an interaction are likely to recommend the organization to their peers. Businesses that wish to design experiences that encourage a particular response must analyze emotions to do so successfully.

HISTORICAL FRAMEWORKS FOR ANALYZING EMOTION

Emotions have been a key focus of academic research and inquiry for many centuries. Psychologists and philosophers have tried to classify the full spectrum of human emotions and the factors that affect them.

One of the most commonly referenced emotional frameworks was articulated by Robert Plutchik in 1980[1]. His wheel of emotions is oriented around eight core emotions that are depicted as vectors of expressions ranging from mild to intense (i.e., from annoyance to anger to rage). According to his work, each emotion also has a corresponding opposite.

Another useful reference model is W. Gerrod Parrott's 2001 paradigm[2]. It includes over 100 emotions that are tied to the six core emotions of love, joy, surprise, anger, sadness, and fear. Parrott asserts that each of these also has a secondary and tertiary corresponding expression.

CLARABRIDGE EMOTION ANALYSIS

Clarabridge drew inspiration for its emotion analysis primarily from the EARL (Emotion Annotation and Representation Language) framework[3]. This model was published in 2006 by the Human-Machine Interaction Network on Emotion

[1] Robert Plutchik, *Emotion: Theory, research, and experience: Vol. 1. Theories of emotion, 1,* (New York, Academic, 1980).
[2] W. Parrott, *Emotions in Social Psychology. Key Readings in Social Psychology,* (Philadelphia, Psychology Press, 2001).
[3] M. Schröder, H. Pirker, M. Lamolle, "First suggestions for an emotion annotation and representation language", In: *Proceedings of LREC'06 Workshop on Corpora for Research on Emotion and Affect,* (Genoa, Italy, 2006): 88–92.

(HUMAINE), now known as the Association for the Advancement of Affective Computing (AAAC), and functions specifically for representing and annotating emotions in technological contexts.

Not all emotions are equally valuable when thinking about customer experience. Certain emotions may be more or less actionable than others. For example, anger is less specific than confusion or frustration. Other emotions such as grief or remorse are simply not as relevant for customer experience management as they might be when examining the human psyche. In this context, emotion analysis must be as specific and actionable as possible.

It is also useful to think of emotions as independent of and orthogonal to sentiment and effort and to recognize that emotions are not binary. Intuitively, we might think that some emotions are "good" and others are "bad"; however, for customer experience, it's essential to separate the concepts of emotion and sentiment. Worry is not always bad, and surprise is not always good. The intersection of different semantic dimensions (i.e., sentiment, emotion, and effort) affords new insights into data that would otherwise be obscured when each is viewed as independent and binary. For example, a customer may be positively surprised at an unexpectedly low price or negatively surprised at a high one. A customer may worry about the safety of a product or express that he or she is no longer worried about those former safety concerns. For this reason, Clarabridge recommends analyzing emotions in a sentiment-insensitive manner.

As mentioned earlier, emotions are fluid. A single experience may encapsulate a handful of different emotions. A trip to the coffee shop may begin with anticipation before turning to annoyance when seeing a long line. The same customer may experience relief when the line moves quickly, worry when observing an incompetent barista, and joy when receiving the drink. None of these specific emotions accurately characterizes the entirety of the experience. An organization will benefit from analyzing each emotional moment separately so that it can make informed decisions about each step of the customer journey.

EMOTION CATEGORY TEMPLATES

Clarabridge offers two out-of-the-box category models, or taxonomies, for analyzing emotions at both a basic and expanded level. Each model is hierarchical and constructed from keywords and phrases that align with an emotion. These key terms also include language-specific idioms and expressions.

The Emotions (Basic) Model, which is most closely aligned to the Parrott framework, is organized into 10 emotions: Happiness, Love, Sadness, Anger, Fear & Worry, Confusion, Disgust, Surprise, Anticipation, and Embarrassment. These categories are largely sentiment-sensitive, meaning that the antonymous terms are excluded from each topic.

The Emotions (Expanded) Model, which is most closely aligned to the EARL framework, is organized into 45 emotions including Despair, Envy, Shame, Anxiety, Appreciation, Doubt, Pride, Regret, Relief, and many more. These categories are sentiment-insensitive, meaning that both the affirmative and negative expressions are included in each topic.

Conclusion

Clarabridge approaches text analytics differently than alternative solutions and leverages a variety of components supported by very sophisticated technology. When it comes to identifying actionable insights, this matters because **better information leads to better decisions**.

The Clarabridge solution is designed to maximize the accuracy of Natural Language Processing and categorization as well as sentiment, effort, and emotion analysis. The detailed linguistic and technological methodology outlined in this report allows Clarabridge users to avoid false negatives (or false positives) when deriving information from their data and to employ an omni-channel analytics solution that integrates findings from multiple sources. Businesses that rely on accurate, comprehensive insights can act confidently to **improve the customer experience**.

Both the accuracy and intricacy of Clarabridge text analytics provide a framework for its entire suite of product offerings. Clarabridge technology empowers organizations to process the growing expanse of customer feedback and interactions accurately, efficiently, and purposefully to **extract operational insights** and empower customer-centric decisions across the organization.

About Clarabridge

Clarabridge helps the world's leading brands understand the true voice of their customers. Companies leverage Clarabridge's omni-channel capabilities to listen to all customer interactions and conversations, analyze this data via a best-in-class AI-powered text analytics engine, and get actionable insights to make mission-critical decisions.

ADDITIONAL STRENGTHS INCLUDE:

- A world-class analytics platform with beautiful, interactive and easily customizable dashboards that can be tailored for every role in the organization, advanced predictive algorithms and sophisticated case management workflows
- A patented, best-in-class Natural Language Processing (NLP) platform specifically designed for Customer Experience Analytics that combines the latest AI and Machine learning technologies and offers accurate and nuanced topic, emotions, effort and intent detection
- Clarabridge was named a Leader in Customer Feedback Management Platforms and Text Analytics, as well as a Strong Performer in Speech Analytics by Forrester Research