Sentiment Analysis

Analyzing the Polarity of Communication

INTRODUCTION

I. SENTIMENT ANALYSIS

WHAT IS SENTIMENT ANALYSIS? 3
CURRENT EXAMPLES OF SENTIMENT ANALYSIS 4
PROBLEMS WITH SENTIMENT ANALYSIS 4
METHODS OF DETERMINING SENTIMENT 5
SCALING SYSTEM 5
SUBJECTIVITY/OBJECTIVITY IDENTIFICATION 6
BALES INTERACTION PROCESS ANALYSIS (IPA) 6

II. USING SENTIMENT ANALYSIS: A WALKTHROUGH 8

1. IDENTIFYING A TARGET FOR ANALYSIS 8
2. CREATING A BALES IPA CHART 9
3. APPLYING YOUR ANALYSIS DEVICE 10
4. TABULATING THE RESULTS 13

III. CONDUCTING YOUR OWN ANALYSIS 15

PARAMETERS OF THIS TASK 15
GOALS AND TAKEAWAYS 15

IV. BIBLIOGRAPHY 16
Introduction

Greetings. The following guide is intended to familiarize you with Sentiment Analysis, a method for gauging the public’s opinion based on a devised criterion of responses.

Sentiment analysis is a fast growing subject in the technical communication field. With the increase in social media, online retail, and personal blogs and publications knowing where public sentiment is leaning has translated into a rapid evolution in sentiment analysis that can become a valuable skill. (Pringle)

After having completed this tutorial, you should be able to do the following:

1. Have a working understanding of sentiment analysis and understand several ways it is used to gauge public opinion.
2. Be able to design and implement an effective sentiment analysis study that will be able to measure the polarity of the public’s opinion on a form of online media.
3. (Optional) Be able to craft an automated analysis test using the tools described in section 3.

Please note that any websites displayed in the screenshots below may have changed over time and that all products and websites shown are trademarked or copyrighted by their respective owners.
I. Sentiment Analysis

This section is intended to give you firm foundation on the purpose of sentiment analysis, now it is used today, and some common methods it is used in.

What is Sentiment Analysis?

The purpose of sentiment analysis is to determine the attitude or inclination of a communicator through the contextual polarity of their speaking or writing. Their attitude may be reflected in their own judgment, emotional state of the subject, or the state of any emotional communication they are using to affect a reader or listener. Simply put, it is trying to determine a person’s state of mind on the subject they are communicating about. This information can be mined from texts, tweets, blogs, social media, news articles, or comments. (King)

Beginning in the 1950’s early methods of sentiment analysis or ‘opinion mining’ were focused on paper documents, usually of product reviews and movie reviews. As media evolved the public was able to express their opinion in a wider array of formats, and the Internet itself has exploded the way sentiment can be measured. (wiki)

As a visual example of what sentiment analysis is, look at figure 1.1 below. Let us pretend we are marketing analysts for Apple and we want to know what the feelings of the public are toward the iPad 2. Instead of ask people directly or conducting a survey, we will go to an online retailer such as Amazon and check the comments left by customers who have left their views on the product.

As you can see below, 746 comments have been left. In order to gain a meaningful understanding of the feelings of customers based on those comments, a sentiment analysis of these comments will be done, one that is much more detailed than a simple star rating. From this analysis, we will be able to know in what direction public opinion lays, the tone of the comments being left, and even the level of knowledge possessed by the customers based on sentiment analysis.
Current Examples of Sentiment Analysis

There are many examples of sentiment analysis being used in a variety of fields such as online retail, to all forms of blogging, and even in politics.

Consumer related businesses have become big users of sentiment analysis, and not just for customer product reviews. Sentiment analysis has begun to creep into business intelligence, customer service, and the management of brand reputation. (King 2011)

Sentiment analysis has even been used to help political strategists gauge public opinion on the Internet as this Yahoo News shows. (Weber) This Yahoo news article is a great example of measuring public opinion using public political queries.

The Obama administration has begun using sentiment analysis to predict responses to campaign messages and policy announcements. (Pringle 2011)

Problems with Sentiment Analysis

Like most scientific methods sentiment analysis is not without its problems.
The problem most often encountered with sentiment analysis is how subjective it is. (Liu) Interpreting the mood of a subject may vary from one person to another; a problem made even harder by the format the subject may be analyzed in. Just how easy would it be to determine the mood of a person about a certain subject if that person only has 140 characters or less to express it in their Twitter account?

There are also technical challenges to sentiment analysis. (Liu) If I were writing a response that said “An IBM iSeires System Storage DS5020 Express is the equivalent to the x450 IBM server.” Most observers without a heavy technical background or a lot of investigation will not know if that is a complimentary comparison or a negative, so that sentence will also be a problem.

**Methods of Determining Sentiment**

As the types of media available to the population has evolved, so to have the way of determining public sentiment. The following are some quick explanations of common sentiment analysis you have probably seen before. Each example is followed by a brief explanation. The last example is a little longer because we will be seeing it later in this tutorial.

**Scaling System**

If you have ever shopped online you or seen a movie rating you are probably familiar with this system. This method assigns a rating system, usually one through four or five for determining the personal appreciation of the item or event the rater has experienced. Even the popular Siskel and Ebert ‘Thumbs Up’ and ‘Thumbs Down’ system qualifies although it does not leave much for degrees of appreciation. (Turney)

![Figure 1.2 Five star ratings.](image)
This is not the only scaling method for gauging public opinion of course, just an example of one.

Just like how reviewers and customers rate their experiences using this system sentiment analysis does so as well by analyzing a block of text and determining the positive or negative sentiment and assigning a rating to it. Once all the relevant material has be tallied, a conclusion can be drawn based on the mean or median scoring by those that analyzed it.

Subjectivity/Objectivity Identification

A subjective/objective system for determining sentiment is typically used one a form of documentation; specifically any sentence of interest in the document that can be determined to be either subjective or objective.

This method has proven difficult to implement because the meaning of a word or sentence may change if it is removed from any kind of context. For instance, a persons response to a quote removed from a newspaper may change depending of if that quote originally came from the Opinions section rather than the front page of a paper.

Bales Interaction Process Analysis (IPA)

The IPA method of analyzing sentiment is to identify and then record each instance of a group interaction based on the nature of the ongoing small group interaction. It is not used to measure the content of this interaction. (Bales)

Bales IPA is based on scoring interactions based on ‘units’ of interaction or communication, these scores are applied to a predetermined set of categories, and an analysis made based on the scores of each category. These units are typically made up from one simple sentence expressing one idea. Complex sentences expressing more than one idea are scored based on the number of independent clauses they may contain.

Fragments of sentences can be scored as communication, one ‘point’ each, but the problem with these comes from the interpretation of them. For instance, a statement of “What?” can range from simply seeking clarification to an expression of disbelief. These fragment sentences must then be taken in context in order to be categorized.

In cases of oral or physical studies, simple sounds like grunts or sighs can typically be categorized and even facial expressions if the observer feels they convey enough meaning to be categorized. (Bales)
For a further explanation of the Bales IPA chart, please click on the following link to launch a video tutorial.

http://relayfiles.njit.edu/da225/NJIT_Default/Flash%20(Medium)/Bales_Tutorial_-_Flash_(Medium)_-_20120324_10.05.47PM.html

Key:
A. Positive reactions  
B. Attempted answers  
C. Questions  
D. Negative Answers

a. Problems of orientation  
b. Problems of evaluation  
c. Problems of control  
d. Problems of decision  
e. Problems of tension-management  
f. Problems of integration

II. Using Sentiment Analysis: A Walkthrough

We will be walking through the creation of a sentiment analysis of a blog posting we select.

1. Identifying a Target for Analysis

As you now know, a sentiment analysis can be used on any item or topic that garners public responses. Manufacturers and designers monitor product reviews to know how the public feels about their products, political committees want to know public opinion about candidates and hot topic issues, etc. Now we have to find our own find our own topic or post to analyze.

   a. We will use this [website](http://onlinecolleges.net) sponsored by onlinecolleges.net to find a technical writing based blog posting for us to analyze. We will be looking for a topic that has these criteria:

      i. It applies to us as technical communicators.

      ii. For a meaningful analysis, this post ideally has more than a dozen outside responses.

      iii. There is a minimal number of repeated responses from the same user(s).

   b. For this walkthrough, we have decided to analyze the responses post [A Reverse Approach to Help Authoring: Writing Documentation Post-Release](http://example.com). As you can see, there are a number of posts available for use to perform an analysis on the responses to this topic.
2. Creating a Bales IPA Chart
   
a. Using the Bales IPA chart shown above in figure 1.3 as a guide, devise an IPA chart that you will use to gauge the overall attitude of the post you have selected. The chart shown above does not allow for an observer to make notations or keep a tally so transferring it to a better format is necessary.

   
   Note: The Bales IPA chart shown above can be customized to fit a situation, however for this instance we are going to use the standard IPA method for determining polarity.

   
b. Here is the chart we will use to record our findings for our sentiment analysis. As you can see it is similar to the Bales IPA chart above but has been shortened down to only four categories.

   
c. 

   Note: As previous studies have shown, the Bales IPA chart can be altered to better fit the study. (Coppola) It is highly recommended that you provide your reasons for altering format of the chart for future readers to understand the conclusions you draw from this.
<table>
<thead>
<tr>
<th>Category</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>1) Shows solidarity, raises other’s status, gives help or rewards response. Generally seems friendly.</td>
<td></td>
</tr>
<tr>
<td>2) Shows tension release such as jokes, laughs, shows satisfaction, and dramatizes in a positive way.</td>
<td></td>
</tr>
<tr>
<td>3) Agrees, shows passive acceptance, understands, concurs, complies.</td>
<td></td>
</tr>
<tr>
<td>4) Gives suggestions or direction implying autonomy for other.</td>
<td></td>
</tr>
<tr>
<td>5) Gives ones own opinion or evaluation, analysis, feelings, or wish on the subject.</td>
<td></td>
</tr>
<tr>
<td>6) Gives orientation, information, repeats or clarifies, confirms idea.</td>
<td></td>
</tr>
<tr>
<td>7) Asks for orientation, information, repetition of idea, or seeks confirmation.</td>
<td></td>
</tr>
<tr>
<td>8) Asks for opinion, evaluation, analysis, or expression of feeling on the matter.</td>
<td></td>
</tr>
<tr>
<td>9) Asks for suggestion, direction, and possible ways of action on a topic.</td>
<td></td>
</tr>
<tr>
<td>10) Disagrees, shows passive rejection, formality, withholds help.</td>
<td></td>
</tr>
<tr>
<td>11) Shows tension, asks for help, withdraws from the field.</td>
<td></td>
</tr>
<tr>
<td>12) Shows antagonism, deflates other’s status, defends or asserts self usually in a negative way.</td>
<td></td>
</tr>
</tbody>
</table>

3. **Applying Your Analysis Device**
   a. Now we will begin classifying the responses based on our Bales IPA-based chart above. After reading through each response, we will check our chart; attempt to match each complete idea to the category it best fits, keeping the count for that category.
   b. Here is an example of how we classified some of the comments that appear.

   Example 1:
We can see from the first sentence that Greg is praising the article so this we classified as a category 1.

The second paragraph contains several sentences with the first showing agreement with evidence he complies with the article. "We tell our customers...biggest key to success." Category 3. The second sentence we took to showing some concurrence with the author so also category 3.

The third paragraph is similar to the second in our opinion; agreement followed by compliance with the stated idea, so categories 1, 3, and 3.

The fourth paragraph is different from just agreeing with the author. It is a asking a question, but what kind? We decided he is asking for direct information, not an opinion or direction; so category 7 here.

Finally, the fifth paragraph, the responder is giving other readers additional information. While he is directing others to this site, is more as a way of presenting information, so we will say category 4 here.

Example 2:
This post of Tammy’s is a great example of a response riding the bottom end of the IPA chart without being rude.

The first sentence shows clear disagreement but in a polite and formal way. Category 10.

The second sentence shows her giving her evaluation of the problem. While in support of a disagreement, it is not antagonistic or inquisitive. Category 5.

The third sentence is a continuation of this evaluation but we thought it was more for clarification than an evaluation. Category 6.

The fourth sentence begins with a great tell, “I think...”. A great sign for category 5 and the rest of the sentence we feel carries on with this idea.

Remember above when we explained that IPA could consider simple sentences or even single word responses? Here we are! Now “Sorry!” cannot be interpreted as a question or as an attempt to answer one. Reading the entire response and taking it in context we believe it is part of the disagreement with the author, but since it is on the ‘nice’ end of the spectrum, we called this a category 10 disagreement.

c. Keep track of the count in anyway you wish, the important thing to do here is to try and remain as constant as you can in categorizing these communications. As you will come to see, this is where the unreliability of sentiment analysis comes into play.

Note: Be sure to disregard any posts created by the posts owner or moderator. We are interested in outside opinions only. Bloggers often respond to comments left on their posts. This is a good way encourage feedback in readers and build a fan base, but it is not something that should be measured here. Responses to these author posts are another matter and since they are reader posts, should still be counted.
4. Tabulating the Results

a. After reading through and classifying the responses to our chosen post, we can take a look at the results of our classification.

<table>
<thead>
<tr>
<th>Category</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>1) Shows solidarity, raises other's status, gives help or rewards response. Generally seems friendly.</td>
<td>9</td>
</tr>
<tr>
<td>2) Shows tension release such as jokes, laughs, shows satisfaction, and dramatizes in a positive way.</td>
<td>6</td>
</tr>
<tr>
<td>3) Agrees, shows passive acceptance, understands, concurs, complies.</td>
<td>21</td>
</tr>
<tr>
<td>4) Gives suggestions or direction implying autonomy for other.</td>
<td>20</td>
</tr>
<tr>
<td>5) Gives ones own opinion or evaluation, analysis, feelings, or wish on the subject.</td>
<td>94</td>
</tr>
<tr>
<td>6) Gives orientation, information, repeats or clarifies, confirms idea.</td>
<td>72</td>
</tr>
<tr>
<td>7) Asks for orientation, information, repetition of idea, or seeks confirmation.</td>
<td>10</td>
</tr>
<tr>
<td>8) Asks for opinion, evaluation, analysis, or expression of feeling on the matter.</td>
<td>5</td>
</tr>
<tr>
<td>9) Asks for suggestion, direction, or possible ways of action on a topic.</td>
<td>3</td>
</tr>
<tr>
<td>10) Disagrees, shows passive rejection, formality, withholds help.</td>
<td>11</td>
</tr>
<tr>
<td>11) Shows tension, asks for help, withdraws from the field.</td>
<td>0</td>
</tr>
<tr>
<td>12) Shows antagonism, deflates other's status, defends or asserts self usually in a negative way.</td>
<td>0</td>
</tr>
</tbody>
</table>

b. From these results we are able to generalize the public opinion or response to this article based on the number of responses. The following conclusion is based on our own interpretation of the tally above. This is one person's interpretation and your opinion may me slightly different. Remember, sentiment is subjective to the observer's own views.

**Conclusion from this analysis:** Overall this seemed to be a well-liked article that many readers agreed with and responded to. It also inspired readers to give their
own ideas about the topic that may aid future readers in understanding this topic as well. Some questions about specific ideas were given but responses to these where generally helpful and supportive with clearly given information or evaluations provided. There was some disagreement about the topic but it was worded politely and usually had some supporting information or personal experience to go along with it.

For a blog targeting technical communicators and a topic specifically for technical writers, it succeeded in presenting an informative topic that many found valuable and also inspired a lively debate that many participated.
III. Conducting Your Own Analysis

Now it is your turn. We learn best by doing and in order to fully grasp the concept, I recommend making the attempt to performing your own sentiment analysis. The following section will explain the best way to make use of this guide to help you perform an analysis of a topic of your choice.

Parameters of this task

Using the process modeled above in part 2, create a sentiment analysis of a blog post whose topic involves a tool or idea related to technical communication.

1. Choose a topic from any of the websites listed in this website, or you may select your own resource but I recommend a subject related technical communication in some way.
2. Create a Bales IPA based sentiment analysis chart you can use to analyze your post.
3. Categorize the responses to your chosen post and keep a count of the categories these responses match in the chart you created.

   Note: Remember not to include posts that were created by the author of the topic.

4. Analyze your results and write a brief description of how you interpret them and what you what you would recommend to the author of the post based upon them.

Goals and Takeaways

Our goal in performing this assignment is to have a further understanding of sentiment analysis and how it is used to gauge public opinion. If you never read about this topic again if you realize the potential importance of the comments and responses section of a blog or topic we will consider this time spent as a success.

After completing this assignment should result in the following deliverables:

- A Bales IPA based sentiment analysis chart with categories addressing the responses to your chosen topic. These categories should include a count of the number of responses to the topic that fit each group.
- An analysis of what your results translate into for the topic including your interpretation of the results and if you believe the post was positively/negatively received and what kind of dialogue it generated, if any.
IV. Bibliography


Liu, Bing. "Sentiment Analysis: A Multi-Faceted Problem.”


Ogneva, Maria. “How Companies Can Use Sentiment Analysis to Improve Their Business.”

