Tone Analyzer

Aayush Agrawal[1], Shankhesh Amarkant Sonawane[2], Babeetha .S[3]
B.Tech Student[1, 2], Assistant Professor[3]
Department of Information Technology
SRM University, Ramapuram, Tamil Nadu, India

Abstract:
In this day-to-day life, we come across many messages, emails, tweets, feedbacks etc. but do they convey what the writer actually meant? Sometimes, we don't understand what the other person meant, i.e., was that person joyful or angry when he/she sent a message or when talking about the customer, were they satisfied with our service/product or not. So, thinking about these issues, we came up with a website which would offer the people a platform where they would be able to understand the other person better. This service uses linguistic analysis to detect and interpret emotions, social tendencies, and language style cues found in text. Tones detected within the General Purpose Endpoint include joy, fear, sadness, anger, disgust, analytical, confident, tentative, openness, conscientiousness, extraversion, agreeableness, and emotional range. We would try to understand a person by using his/her Tweets, an Online Review, Email message and a simple text.

1. INTRODUCTION

The Tone Analyzer service uses linguistic analysis to detect emotional, social, and language tones in written text. The service can analyze tone at both the document and sentence levels. You can use the service to understand how your written communications are perceived and then to improve the tone of your communications. Businesses can use the service to learn the tone of their customers' communications and to respond to each customer appropriately, or to understand and improve their customer conversations in general. You submit JSON, plain text, or HTML input that contains your written content to the service. The service accepts up to 128 KB of text, which are about 1000 sentences. The service returns JSON results that report the tone of your input. You can use these results to improve the perception and effectiveness of your communications, ensuring that your writing conveys the tone and style that you want for your intended audience. The following diagram shows the basic flow of calls to the service.
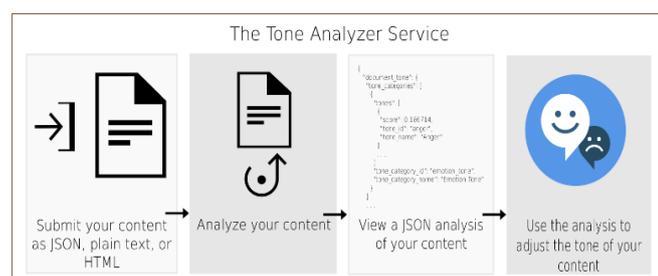
2. TONE ANALYZER ENDPOINTS

The service offers two endpoints:

- **General purpose endpoint**

Use the Tone Analyzer general purpose endpoint to analyze shorter web data, such as email messages or tweets, or longer documents, such as articles or blog posts. Monitor social media to understand what customers are saying about a brand and to determine whom to target with specific messaging. The endpoint accepts JSON, plain text, or HTML input. For more information about the method and the tones that it returns, see Using the general purpose endpoint. The general purpose lets you submit content to the service for analysis. The service returns overall and sentence-level analyses of the tone of the content.

- **Customer engagement endpoint**

Use the Tone Analyzer customer engagement endpoint to monitor customer service and support conversations. Escalate customer conversations when they turn sour or find opportunities to improve customer service scripts, dialog strategies, and customer journeys. The endpoint accepts JSON input. For more information about the method and the tones that it returns, see using the customer engagement endpoint. The Customer engagement analyzes conversations between customers and customer service agents. The service measures customer satisfaction and concerns, assesses agent performance, and lets you gauge how the interaction evolves.

3. USE CASES

Some interesting use cases of the service follow:

- *Social listening and audience monitoring:*
Monitor social media to understand what customers are saying about your brand in real-time. For example, you might determine that your customers in Chicago are sad after the Bulls lost or happy during the Taste of Chicago festival. (General purpose endpoint)

- *Personalized marketing:*
Determine whom to target with personalized messaging and when. For example, a travel company might target happy consumers with "treat yourself" messaging, sad consumers with "escape" messaging, and angry consumers with "relax" messaging. (General purpose endpoint)

- *Chat bots:*
Enable an automated agent to detect customer tones and, based on the tones identified, craft suitable responses. For example, you might respond to sadness with "I'm sorry you are upset about this problem" or to satisfaction with "I'm glad you are satisfied with our service." (Customer engagement endpoint)

- *Customer engagement monitoring and quality assurance:*
Monitor the overall tone of agent and customer communications, detect anomalies, and highlight opportunities

to train agents on how to better communicate. (Customer engagement endpoint)

4. PREDICTING CUSTOMER SATISFACTION IN SUPPORT FORUMS

IBM analyzed customer support forums at a software company that is focused on multiple industries. The company actively contributes to customer support forums. Users can give *Kudos* to answers that they find useful.

Goals
Predict customer satisfaction from the tone of the question and response. IBM assumed that an answer with Kudos meant that the user was satisfied.

Actions
•   Crawled the most recent 1000 threads from several forums, making sure to include the same number of responses with and without Kudos.
•   Analyzed both the questions and the responses.
•   Applied several state-of-the-art classifiers, such as naive Bayes, Support Vector Machine (SVM), and random forest, to predict whether an answer would receive Kudos.

Results
The service can predict Kudos with 66-percent accuracy. IBM found the following correlations between the tones that a forum response returns and whether those reply receives Kudos:
•   The more confident a response is, the more likely it is to earn Kudos (correlation of 0.23 between a high-value score on confidence and Kudos).
•   The more tentative a response is, the less likely it is to earn Kudos (negative correlation of -0.27 between a high-value score on tentative and Kudos).
•   The more open a response is, the more likely it is to earn Kudos (correlation of 0.10 between a high-value score on openness and Kudos).
•   The more organized, thoughtful, and thorough a response is, the more likely it is to earn Kudos (correlation of 0.10 between a high-value score on conscientiousness and Kudos).

5. PREDICTING CUSTOMER SATISFACTION IN TWITTER RESPONSES

Many companies are switching their customer support to Twitter. Twitter allows real-time answers, which helps to establish the brand as one with real people who care about their clients. IBM analyzed 333 customer support conversations on Twitter. The customers were satisfied with 240 of the conversations and not happy with 93 of the interactions. IBM measured satisfaction by reading through the conversations and labeling them. Responses were labeled "customer satisfied" when they solved the problem and the client seemed satisfied; they were labeled "customer not satisfied" when the problem was not addressed to the client's satisfaction.

Goals
Validate whether the tone of the conversations between the agent and the customer had any effect on overall customer satisfaction. Also identify the tone features that significantly impact customer satisfaction.

Actions
•   Stripped punctuation, mentions, and links from the tweets.
•   Split each interaction into customer tweets and support tweets.
•   Analyzed each side of the conversation with the Tone Analyzer service, and compared the results to find correlations.

Results
The service can predict customer satisfaction from the tone of the response with 67-percent accuracy. IBM identified the following correlations between the tone of customer service tweets and whether the client was satisfied with the response:
•   Customer support representatives who act disgusted in their tweets are less likely to satisfy the customer (negative correlation of -0.155 between a high-value score on disgust in a customer service tweet and customer satisfaction).
•   Customer support representatives who seem frustrated, angry, stressed, or insecure are less likely to satisfy the customer (negative correlation of -0.186 between a high-value score on emotional range in a customer service tweet and customer satisfaction).
IBM discovered the following correlations between the tone of customer tweets and whether the customer was satisfied with the response:
•   The angrier customers are, the less likely they are to be satisfied with the response (negative correlation of -0.198 between a high-value score on anger in a customer tweet and customer satisfaction).
•   The more disgusted customers are, the less likely they are to be satisfied with the response (negative correlation of -0.184 between a high-value score on disgust in a customer tweet and customer satisfaction).
•   The more organized, thoughtful, and thorough customers' tweets are, the more likely they are to be satisfied with the response (correlation of 0.177 between a high-value score on conscientiousness in a customer tweet and customer satisfaction).

6. PREDICTING TED TALK APPLAUSE

TED is a nonprofit organization that runs global conferences with the slogan "Ideas worth spreading." TED Talk speakers have 18 minutes to use innovative and engaging storytelling to address a wide range of topics within the research and practice of science and culture. Not all TED Talks are popular, and one way of measuring audience satisfaction with a talk is to measure the amount of applause it receives.

Goals
Discover which tone patterns in TED Talks lead to applause and which patterns do not. Also predict applause based on the tone of a sentence.

Actions
Sentences that received applause were already tagged in the data set.
•   Reviewed 1931 TED Talks.
•   Categorized as "applause text" a sentence tagged with "Applause." Also tagged the three sentences before the sentence with "applause text" and the three sentences after it with "non-applause text."
•   Analyzed both applause and non-applause text with the Tone Analyzer service.

- Based on the correlations that were found, created classifiers to predict applause in other TED Talks based on their tone.

Results

The service can predict applause with 75-percent accuracy. IBM found the following correlations between the tone of each set of sentences and whether those sentences received applause:

- The more disgusted a speaker is, the less likely they are to receive applause (negative correlation of -0.066 between a high-value score on disgust and applause).

- The more sadness a speaker expresses, the less likely they are to receive applause (negative correlation of -0.055 between a high-value score on sadness and applause).

- The more emotionless or impersonal a speaker seems, the less likely they are to receive applause (negative correlation of -0.29 between a high-value score on analytical and applause).

- The more joyful, contented, and satisfied a speaker seems, the more likely they are to receive applause (correlation of 0.21 between a high-value score on joy and applause).

- The more organized, thoughtful, and thorough a speaker seems, the more likely they are to receive applause (correlation of 0.0964 between a high-value score on conscientiousness and applause).

- The more engaging, sociable, and outgoing a speaker seems, the more likely they are to receive applause (correlation of 0.0942 between a high-value score on extraversion and applause).

- The more caring, sympathetic, and trustworthy a speaker seems, the more likely they are to receive applause (correlation of 0.068 between a high-value score on agreeableness and applause).

- The more concerned, passionate, and fierce a speaker seems, the more likely they are to receive applause (correlation of 0.064 between a high-value score on emotional range and applause).

7. PREDICTING TWITTER RETWEETS AND LIKES

Establishing a brand on Twitter is becoming a requirement for companies to succeed. An essential part of establishing you or your company as worth following is creating tweets that garner many likes and retweets.

Goals

Find correlations between the tone of a tweet and whether that tweet is liked or retweeted.

Actions

- Crawled 5881 tweets from several business accounts on Twitter.

- Stripped punctuation, mentions, and links from the tweets.

- Analyzed each tweet with the Tone Analyzer service, and compared the results to find correlations.

Results

IBM found the following correlations between the tone of a tweet and whether it is liked:

- The more imaginative, open to change, or emotionally aware a tweet seems, the more likely it is to be liked (correlation of 0.02 between a high-value score on openness and likes).

- The more organized, thoughtful, or thorough a tweet seems, the more likely it is to be liked (correlation of 0.03 between a high-value score on consciousness and likes).

- The angrier, more stressed, or impulsive a tweet seems, the less likely it is to be retweeted (negative correlation of -0.02 between a high-value score on emotional range and likes).

IBM found the following correlations between the tone of a tweet and whether it is retweeted:

- The more imaginative, open to change, or emotionally aware a tweet seems, the more likely it is to be retweeted (correlation of 0.05 between a high-value score on openness and retweets).

- The more organized, thoughtful, or thorough a tweet seems, the more likely it is to be retweeted (correlation of 0.03 between a high-value score on consciousness and retweets).

- The angrier, more stressed, or impulsive a tweet seems, the less likely it is to be retweeted (negative correlation of -0.033 between a high-value score on emotional range and retweets).

8. PREDICTING ONLINE DATING MATCHES

Millions of people around the world use online dating to meet that special someone. People use online dating to find others who have much in common with them and to market themselves as potential partners.

Goals

Correlate the tone of an individual's profile with the tone of a potential match's profile. Also discover whether that correlation would predict match success.

Actions

- Crawled approximately 50,000 user profiles.

- Analyzed each profile with the Tone Analyzer service.

- Defined potential matches as those who communicated through the site.

- Compared the tone analysis of potential matches to find correlations.

- Developed a statistical model from the tone similarity of the profiles to predict whether two users would communicate. Then compared the model to multiple baselines that consider other attributes such as demographics.

Results

Tone similarity between profiles can make a 45-percent improvement in predicting whether two users will communicate as compared to predictors that dating websites regularly use. IBM discovered a strong overall correlation between tone similarity and the number of messages exchanged, as shown in the following image.

9. THE GENERAL-PURPOSE MODEL

The general purpose endpoint analyzes written content for a set of tones that are applicable to a broad range of uses. The Tone Analyzer service computes a scorecard that includes the following tones:

- **Emotional tone** is derived from IBM's work on emotion analysis, which is an ensemble framework that infers

emotions from a given text. To derive emotion scores from text, IBM use a stacked generalization-based ensemble framework; stacked generalization uses a high-level model to combine lower-level models to achieve greater predictive accuracy. Features such as n-grams (unigrams, bigrams, and trigrams), punctuation, emoticons, curse words, greetings (such as "hello," "hi," and "thanks"), and sentiment polarity are fed into machine-learning algorithms to classify emotion categories.

- **Social tone** consists of the Big Five personality dimensions of openness, agreeableness, and conscientiousness. For more information about these Big Five characteristics, see Personality models from the Personality Insights service.
- **Language tone** is calculated from linguistic analysis based on learned features.

10. MEASURING THE QUALITY OF THE SERVICE

IBM measured the quality of the Tone Analyzer service along each of the dimensions of tone mentioned in the previous section:

- *Emotional tone* categories were benchmarked against standard emotion data sets such as ISEAR and SEMEVAL. Results show that the average performance of the ensemble model (macro-average F1 score is around 41 percent and 68 percent, respectively, for the two data sets) is statistically better than the best reported accuracy of the state-of-the-art models (whose macro-average F1 scores are around 37 percent and 63 percent, respectively).

- *Social tone* categories were tested against ground-truth data collected by administering personality surveys to more than 1500 people. The derived social scores were correlated with survey-based scores (p-value less than 0.05), with an average correlation coefficient of 0.33. The average Mean Absolute Error (MAE) between the inferred and actual scores was 0.12 for Big Five dimensions.

- *Language tone* was evaluated with an in-depth study of more than two hundred thousand sentences collected from sources such as debate forums, speeches, and social media. Of these sentences, IBM randomly selected 1330 sentences for analytical tone and 1000 sentences each for confident and tentative tones. IBM then submitted these sentence to the Tone Analyzer service and also asked humans to analyze them. Five annotators labeled each sentence, and IBM used the most prevalent of the five annotated results to determine the final labels.

- For *analytical tone*, humans labeled 915 of the 1330 sentences as analytical, 411 as non-analytical, and 4 as not understandable. By comparing the predicted label with these ground-truth labels, IBM found that its analytical tone detection received an F1 score of 0.7518.

- For *tentative tone*, humans labeled 292 of the 1000 sentences as tentative, 706 as non-tentative, and 2 as not understandable. By comparing the predicted label with the ground-truth labels, IBM found that its tentative tone detection received an F1 score of 0.6369.

- For *confident tone*, humans labeled 623 of the 1000 sentences as confident, 374 as non-confident, and 3 as not understandable. By comparing the predicted label with the ground-truth labels, IBM found that its confident tone detection received an F1 score of 0.7288. Overall, the differences between the predicted and ground-trust labels are not statistically significant. This indicates that the service performs well.

11. THE CUSTOMER-ENGAGEMENT MODEL

The customer engagement endpoint identifies tones from the customer-care domain. To select the tones to evaluate with the customer-engagement model, IBM first conducted a study to identify which dimensions of tone are perceived as important in the domain:

1. IBM selected a set of 53 tons from the tone dimensions used in marketing, the dimensions used to describe writing styles, and emotion and personality scales from psychology.
2. IBM asked workers on to rate the extent to which the 53 tone attributes describe a specific utterance in 1000 customer-care conversations. To simplify the rating task in the context of crowd-sourcing, IBM divided the 53 tones into four subsets. Human annotators needed only to rate a subset of the tones; IBM then determined ratings for all of the tones from aggregations of these results.
3. IBM performed factor analysis on a 53-by-53 correlation matrix and found at least seven significant factors (dimensions). IBM determined the names of factors to represent the most important concepts reflected by each of the dimensions.

These steps identified seven important tone dimensions for the customer-care domain: frustration, satisfaction, excitement, politeness, impoliteness, sadness, and sympathy.

12. HOW DOES IT WORK?

AUTHENITCATION
You authenticate to the Tone Analyzer API by providing the username and password that are provided in the service credentials for the service instance that you want to use. The API uses HTTP basic authentication. After creating an instance of the Tone Analyzer service, select **Service Credentials** from the navigation on the left side of its dashboard page to see the username and password that are associated with the instance.

REQUEST LOGGING
By default, all Watson services log requests and their results. Logging is done only to improve the services for future users. The logged data is not shared or made public. To prevent IBM from accessing your data for general service improvements.

RESPONSE HANDLING
The Tone Analyzer service uses standard HTTP response codes to indicate whether a method completed successfully. A 200-level response always indicates success. A 300-level response indicates the requested resource has not been modified. A 400-level response indicates some sort of input failure. And a 500-level response typically indicates an internal system error.

13. SUMMARY

The Tone Analyzer service uses linguistic analysis to detect emotional, social, and language tones in written text. The service can analyze tone at both the document and sentence levels. You can use the service to understand how your written communications are perceived and then to improve the tone of

your communications. Businesses can use the service to learn the tone of their customers' communications and to respond to each customer appropriately, or to understand and improve their customer conversations in general. You submit JSON, plain text, or HTML input that contains your written content to the service. The service accepts up to 128 KB of text, which is about 1000 sentences. The service returns JSON results that report the tone of your input. You can use these results to improve the perception and effectiveness of your communications, ensuring that your writing conveys the tone and style that you want for your intended audience. The following diagram shows the basic flow of calls to the service.

14. REFERENCES

[1]. Tone Patterns in the KikambaLexicon: An Analysis

[2]. Stress, Rhythm, Tone and Intonation, Ching Kang Liu National Taipei University

[3]. Tone and Intonation by MARY E. BECKMAN AND JENNIFER J. VENDITTI