

Sentiment Analysis: Concept, Analysis and Applications



Shashank Gupta · [Follow](#)

Published in Towards Data Science · 7 min read · Jan 7, 2018



2K



12



Sentiment analysis is contextual mining of text which identifies and extracts subjective information in source material, and helping a business to understand the social sentiment of their brand, product or service while monitoring online conversations. However, analysis of social media streams is usually restricted to just basic sentiment analysis and count based metrics. This is akin to just scratching the surface and missing out on those high value insights that are waiting to be discovered. So what should a brand do to capture that low hanging fruit?

With the recent advances in deep learning, the ability of algorithms to

analyse text has improved considerably. Creative use of advanced artificial intelligence techniques can be an effective tool for doing in-depth research. We believe it is important to classify incoming customer conversation about a brand based on following lines:

1. Key aspects of a brand's product and service that customers care about.
2. Users' underlying intentions and reactions concerning those aspects.

These basic concepts when used in combination, become a very important tool for analyzing millions of brand conversations with human level accuracy. In the post, we take the example of Uber and demonstrate how this works. Read On!

Text Classifier — The basic building blocks

Sentiment Analysis

Sentiment Analysis is the most common text classification tool that analyses an incoming message and tells whether the underlying sentiment is positive, negative or neutral. You can input a sentence of your choice and gauge the underlying sentiment by playing with the demo [here](#).

Top highlight

Intent Analysis

Intent analysis steps up the game by analyzing the user's intention behind a message and identifying whether it relates an opinion, news, marketing, complaint, suggestion, appreciation or query.



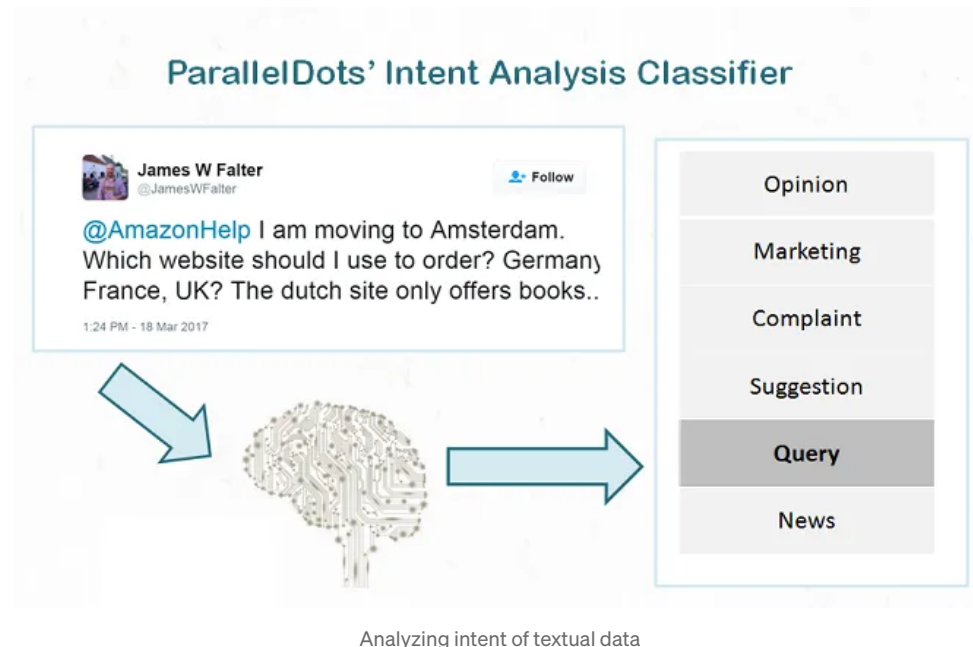
Search

Write

Sign up

Sign in

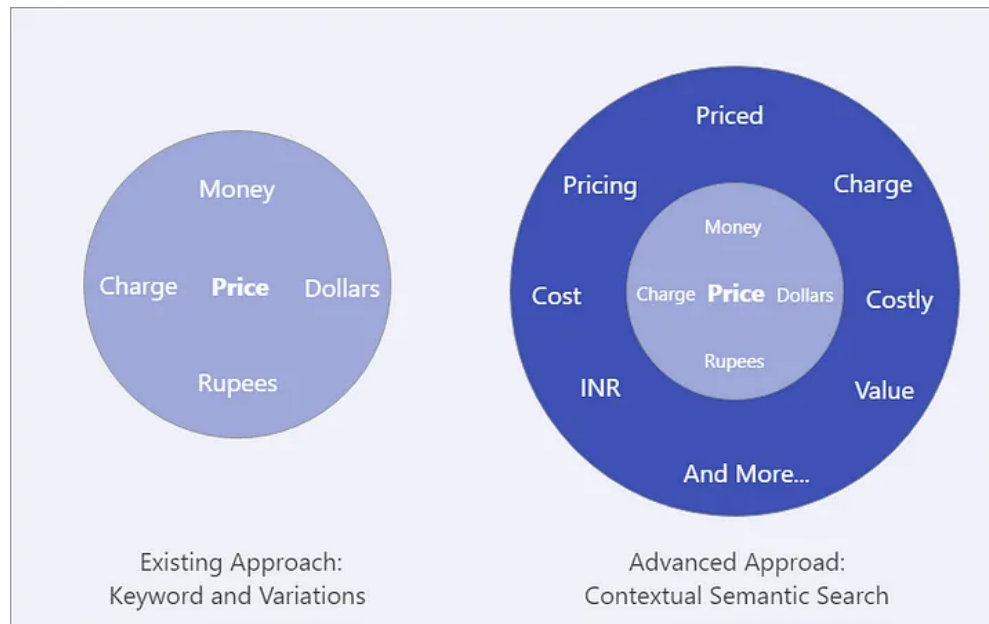




Contextual Semantic Search(CSS)

Now this is where things get really interesting. To derive actionable insights, it is important to understand what aspect of the brand is a user discussing about. For example: Amazon would want to segregate messages that related to: late deliveries, billing issues, promotion related queries, product reviews etc. On the other hand, Starbucks would want to classify messages based on whether they relate to staff behavior, new coffee flavors, hygiene feedback, online orders, store name and location etc. But how can one do that?

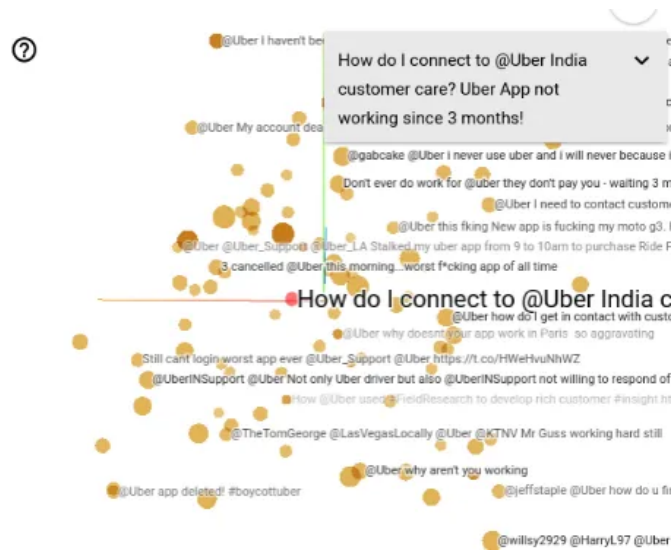
We introduce an intelligent smart search algorithm called **Contextual Semantic Search (a.k.a. CSS)**. The way CSS works is that it takes thousands of messages and a concept (like **Price**) as input and filters all the messages that closely match with the given concept. The graphic shown below demonstrates how CSS represents a major improvement over existing methods used by the industry.



Existing approach vs Contextual Semantic Search

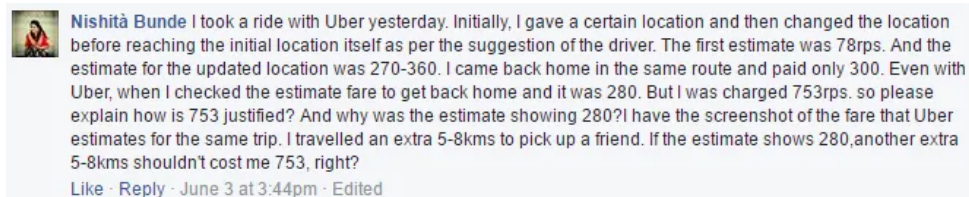
A conventional approach for filtering all **Price** related messages is to do a keyword search on **Price** and other closely related words like (pricing, charge, \$, paid). This method however is not very effective as it is almost impossible to think of all the relevant keywords and their variants that represent a particular concept. CSS on the other hand just takes the name of the concept (**Price**) as input and filters all the contextually similar even where the obvious variants of the concept keyword are not mentioned.

For the curious people, we would like to give a glimpse of how this works. An AI technique is used to convert every word into a specific point in the hyperspace and the distance between these points is used to identify messages where the context is similar to the concept we are exploring. A visualization of how this looks under the hood can be seen below:



Visualizing contextually related Tweets

Time to see CSS in action and how it works on comments related to Uber in the examples below:



Similarly, have a look at this tweet:



In both the cases above, the algorithm classifies these messages as being contextually related to the concept called **Price** even though the word **Price** is not mentioned in these messages.

Uber: A deep dive analysis

Uber, the highest valued start-up in the world, has been a pioneer in the sharing economy. Being operational in more than 500 cities worldwide and serving a gigantic user base, Uber gets a lot of feedback, suggestions, and complaints by users. Often, social media is the most preferred medium to register such issues. The huge amount of incoming data makes analyzing, categorizing, and generating insights challenging undertaking.

We analyzed the online conversations happening on digital media about a few product themes: **Cancel, Payment, Price, Safety and Service**.

For a wide coverage of data sources, we took data from latest comments on Uber's official Facebook page, Tweets mentioning Uber and latest news articles around Uber. Here's a distribution of data points across all the channels:

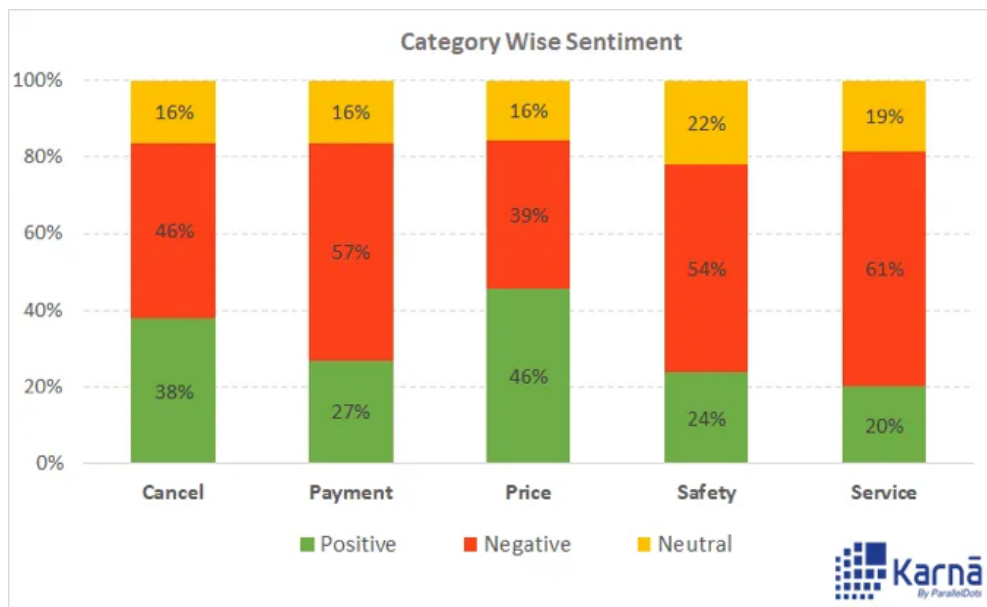
1. Facebook: 34,173 Comments
2. Twitter: 21,603 Tweets
3. News: 4,245 Articles

Analyzing sentiments of user conversations can give you an idea about overall brand perceptions. But, to dig deeper, it is important to further classify the data with the help of Contextual Semantic Search.

We ran the Contextual Semantic Search algorithm on the same dataset, taking the aforementioned categories in account (**Cancel, Payment, Price, Safety, and Service**).

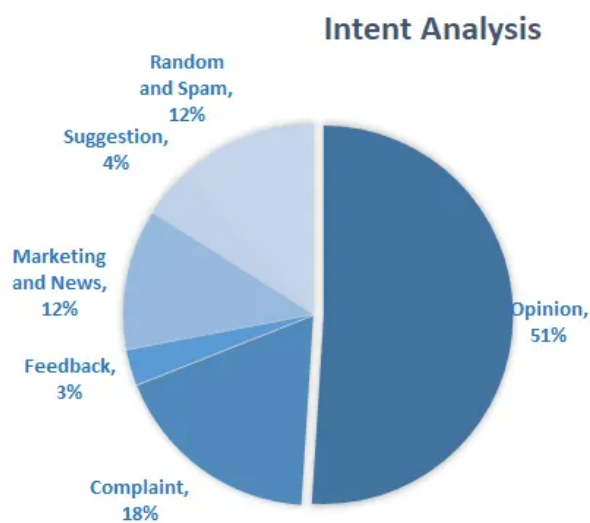
FACEBOOK

Sentiment Analysis



Breakdown of Sentiment for Categories

Noticeably, comments related to all the categories have a negative sentiment majorly, bar one. The number of positive comments related to **Price** have outnumbered the negative ones. To dig deeper, we analyzed intent of these comments. Facebook being a social platform, the comments are crowded *random content, news shares, marketing and promotional content and spam/junk/unrelated content*. Have a look at the intent analysis on the Facebook comments:



Intent analysis of Facebook comments

Intent analysis of Facebook comments

Thus, we removed all such irrelevant intent categories and reproduced the result:



Filtered Sentiment Analysis

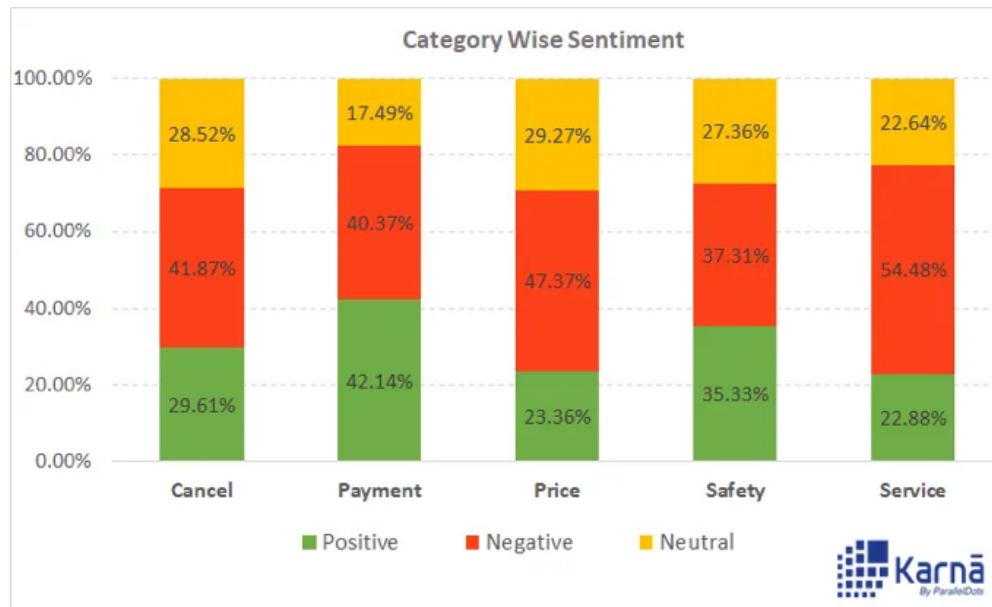
There is noticeable change in the sentiment attached to each category. Especially in **Price** related comments, where the number of positive comments has dropped from 46% to 29%.

This gives us a glimpse of how CSS can generate in-depth insights from digital media. A brand can thus analyze such Tweets and build upon the positive points from them or get feedback from the negative ones.

TWITTER

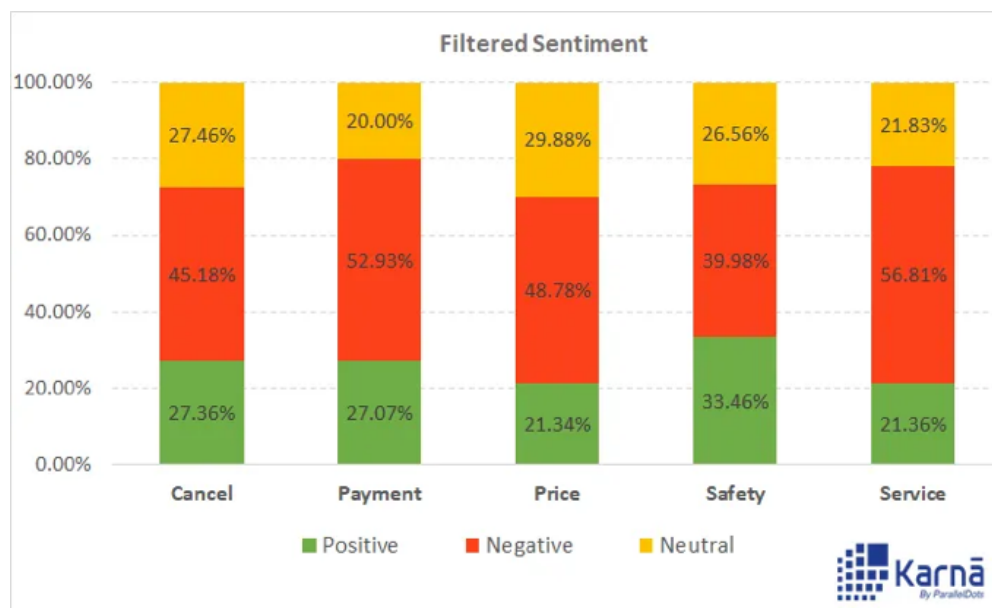
Sentiment Analysis

A similar analysis was done for crawled Tweets. In the initial analysis Payment and **Safety** related Tweets had a mixed sentiment.



Category wise sentiment analysis

To understand real user opinions, complaints and suggestions, we have to again filter the the unrelated Tweets(Spam, junk, marketing, news and random):



Filtered sentiment

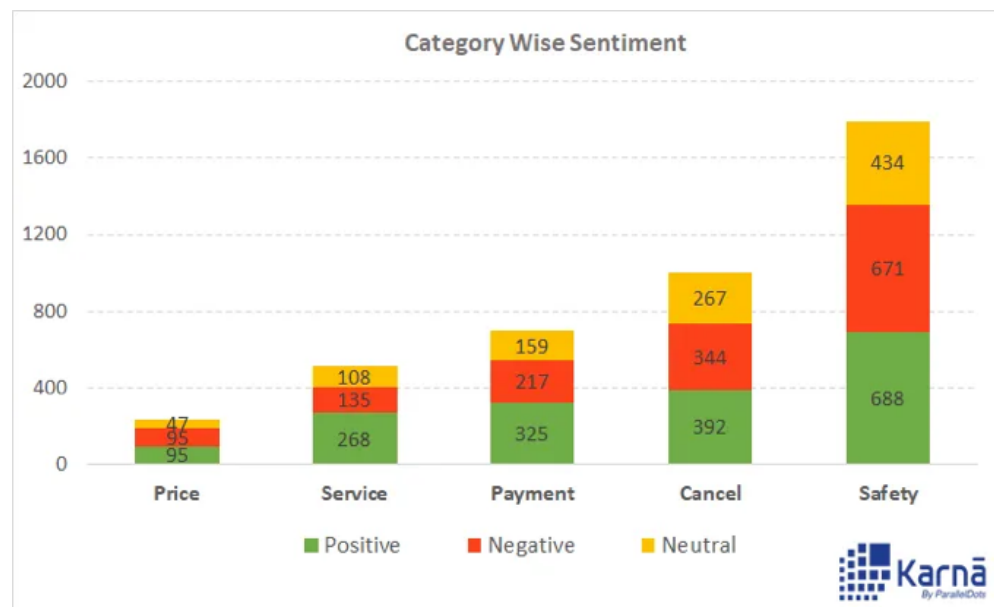
There is a remarkable reduction in number of positive **Payment** related Tweets. Also, there is a significant drop in the number of positive Tweets for the category **Safety**(and related keywords.)

Additionally, **Cancel**, **Payment** and **Service** (and related words) are the most talked about topics in the comments on Twitter. It seems that people talked most about drivers cancelling their ride and the cancellation fee charged to them. Have a look at this Tweet:



Brand like Uber can rely on such insights and act upon the most critical topics. For example, **Service** related Tweets carried the lowest percentage of positive Tweets and highest percentage of Negative ones. Uber can thus analyze such Tweets and act upon them to improve the service quality.

NEWS



Sentiment Analysis for News headlines

Understandably so, Safety has been the most talked about topic in the news.

Interestingly, news sentiment is positive overall and individually in each category as well.

We classified news based on their popularity score as well. The popularity score is attributed to the share count of the article on different social media channels. Here's a list of top news articles:

1. [Uber C.E.O. to Leave Trump Advisory Council After Criticism](#)
2. [#DeleteUber: Users angry at Trump Muslim ban scrap app](#)
3. [Uber Employees Hate Their Own Corporate Culture, Too](#)
4. [Every time we take an Uber we're spreading its social poison](#)
5. [Furious customers are deleting the Uber app after drivers went to JFK airport during a protest and strike](#)

Conclusion

The age of getting meaningful insights from social media data has now arrived with the advance in technology. The Uber case study gives you a glimpse of the power of Contextual Semantic Search. It's time for your organization to move beyond overall sentiment and count based metrics. Companies have been leveraging the power of data lately, but to get the deepest of the information, you have to leverage the power of AI, Deep learning and intelligent classifiers like Contextual Semantic Search and [Sentiment Analysis](#). At [Karna](#), you can contact us to license our technology or get a customized dashboard for generating meaningful insights from digital media. You can check the demo [here](#).

[ParallelDots AI APIs](#), is a Deep Learning powered web service by [ParallelDots Inc](#), that can comprehend a huge amount of unstructured text and visual content to empower your products. You can check out some of our text analysis [APIs](#) and reach out to us by filling this form [here](#) or write to us at apis@paralleldots.com.