7 10 2018 001

Note on the Contribution of Twitter to Marketing Research (Technial note)

Technical note prepared by **Shahin RASOULIAN**,[1] **Yany GRÉGOIRE**,[2] and **Normand TURGEON**[3]

Instructions for use

This technical note was written as an optional companion document for use with the Twitter/United Airlines exercise "Twitter and Marketing Research: How High is United Airlines Flying?" It can also be read on its own to gain a greater understanding of how social media/Twitter metrics can be used or as a companion to other teaching cases.

Overview

Developments in artificial intelligence (AI) and big data analysis have revolutionized marketing research. AI uses many different types of algorithms to collect large volumes of data about customers and competitors from difference sources, including social media and websites, that can then be analyzed and translated into management actions and decisions. Marketers and researchers are starting to use these sophisticated tools to conduct marketing research to better understand customer intentions and attitudes.

Artificial intelligence in marketing

AI is a field of computer science focused on creating intelligent machines and algorithms that function like humans and are able to perform tasks that normally require human intelligence (e.g., visual perception, speech recognition, decision-making, data analysis, etc.).[4] The use of these machines has arguably simplified the construction, exploration, and analysis of big data.[5]

[1] Shahin Rasoulian is a data scientist at the National Bank of Canada.

[2] Yany Grégoire is Chairholder of HEC Montréal's Omer DeSerres Chair of Retailing.

[3] Normand Turgeon is a professor in HEC Montréal's Department of Marketing.

[4] Ming-Hui Huang and Roland T. Rust, "Artificial intelligence in service," *Journal of Service Research*, Vol. 21, No. 2, 2018, pp. 155–172, accessed June 16, 2018, http://journals.sagepub.com/doi/abs/10.1177/1094670517752459.

[5] Uthayasankar Sivarajah, Muhammad Mustafa Kamal, Zahir Irani, and Vishanth Weerakkody, "Critical analysis of big data challenges and analytical methods," *Journal of Business Research*, Vol. 70, January 2017, pp. 263–286, https://www.sciencedirect.com/science/article/pii/S014829631630488X (accessed June 16, 2018)

"Big data" is a term referring to extremely large datasets (containing over 10 million observations or terabytes of data) that can include both structured data, such as a company's transactions or accounting data, and unstructured data, such as texts and images.

Social media is a rapidly expanding data source for AI-powered marketing[1] that can provide companies with valuable information about their customers. The likes and comments of consumers can provide insights into which of the company's actions get positive and negative feedback, for example, enabling them to measure customer satisfaction, track major complaints, make changes, and create a better customer experience.[2] Social media can also give companies insights into competitors that can be used to measure and improve their market position.

Social media data collection

Social media data collection is done with application programming interfaces (API). Generally speaking, APIs are software intermediaries that allow applications to communicate with other applications, servers, or infrastructures. Most mobile and desktop applications are built on APIs, for example. When customers use mobile phone apps, they are using APIs to connect to servers and retrieve information. Many websites now provide developers and programmers with APIs to develop applications based on the content and capabilities of their own websites. Amazon and eBay APIs allow developers to use their existing retail infrastructure to create online stores.

In the case of social media marketing, APIs operate as gateways, allowing companies, marketers, and researchers to access information about public profiles and users that can be used for a variety of purposes. Today, many social media sites, including Facebook, Twitter, Instagram, Google Plus, and Meetup, provide API access to individuals. These APIs are either free or pay as you go.

As mentioned above, APIs are merely gateways or protocols. To obtain data from social media, researchers must develop data collection algorithms that can be used in programming environments, connect to the target API, and retrieve the required information. Python and R are among the most popular programming environments for connecting to APIs and collecting social media data.

Social media, Twitter, and marketing metrics

The choice of social media used for data collection depends on the goal of the research, but Twitter has received enormous attention from marketers and become an important brand

[1] Maria Teresa Pinheiro Melo Borges Tiago and José Manuel Cristovao Verıssimo, "Digital marketing and social media: Why bother," *Business Horizons*, Vol. 57, No. 6, 2014, pp. 703–708, https://www.researchgate.net/publication/265380521_Digital_Marketing_and_Social_Media_Why_Bother (accessed June 16, 2018).

[2] Clay M. Voorhees, Paul W. Fombelle, Yany Grégoire, Sterling Bone, Anders Gustafsson, Rui Sousa, and Travis Walkowiak, "Service encounters, experiences and the customer journey: Defining the field and a call to expand our lens," *Journal of Business Research*, Vol. 79, 2017, pp. 269–280, https://www.sciencedirect.com/science/article/pii/S0148296317301364 (accessed June 20, 2018).

management tool. It provides an effective way for customers to interact with companies through mentions, retweets, likes, and comments.[1] These features convey valuable information for companies, including consumer attitudes, engagement, preferences, reviews, and reactions to products and services.

Twitter's API allows users to collect the following information: a company's tweets and replies;[2] messages (or mentions) from followers using "at" signs (@); comments from users using hashtags (#); dates and sources of posts; number of likes for each post; and number of retweets for each post. It's worth noting that companies are no different from individual Twitter users and that all of the above-mentioned information can be retrieved for all Twitter users.

Twitter features and related marketing metrics are explained in greater detail below:

Twitter posts: Company posts can be of two types: tweets (i.e., messages or content containing text, photos, GIFs, and/or video) and replies (i.e., responses to customer complaints, questions, or likes). Company replies don't provide researchers with much information aside from how actively the company responds to customers. Tweets, on the other hand, can be valuable sources of information, with the number of likes and retweets translated into customer preferences and level of engagement with the company.

Number of likes: The number of favourites or likes that a company's tweet receives indicates the level of customer engagement with the company. Customer engagement can be defined as customers' cognitive and affective commitment to an active relationship with a company, with higher engagement indicating greater participation and connection with the organization's offerings.[3] Engaged customers can contribute to a company's reputation and long-term financial performance. Number of likes can also indicate the social or marketing preferences of customers, helping companies to generate content and adjust offerings tailored to their followers' interests.

It should be noted, however, that the total number of followers or app users does not indicate the engagement of users or the popularity of applications. This is because many accounts and profiles are not active in online environments.

Number of retweets: When a follower shares the tweet of a company (or another user) on their profile page, it is seen by all their followers and is known as a retweet. As with the number of likes, this metric is a proxy measure of the volume of customer engagement, indicating the interests and preferences of followers. The number of retweets can thus be expected to have a strong positive correlation with the number of likes received by a tweet. However, the absolute number of retweets is generally less than that of likes since it is easier and more common for users to simply click the "heart" (like) button than to retweet.

While there is a correlation between the number of likes and retweets, the impact of retweets can be far greater than that of likes, extending far beyond the current followers of a company. Retweets share

[1] Tiago and Cristovao Verıssimo, op. cit.

[2] Twitter can be used by both individuals and companies. Here, we focus on its use by companies.

[3] Shiri D. Vivek, Sharon E. Beatty, and Robert M. Morgan, "Customer engagement: Exploring customer relationships beyond purchase," *Journal of Marketing Theory and Practice*, Vol. 20, No. 2, 2012, pp. 122–146.

content that communicates company information to users who do not follow the company and can act as platforms for spreading word-of-mouth (WOM) advertising about the firm. Companies thus try to create highly shareable content that will inspire followers to communicate the company's brand image to other users.

Impactful Tweets: Given the above, it is in a company's best interests to generate shareable content that will improve its brand image, making the best possible use of free brand management tools and WOM advertising.

Mentions and hashtags: Sentiment analysis

Mentions and hashtags are user-generated content that can communicate thoughts and attitudes of followers and customers about specific brands or companies. This content can be extremely helpful in measuring customer satisfaction.

Mentions: Mentions, also known as at-signs (@), direct public posts to specific companies (or Twitter users).[1] These posts may contain positive, negative, or neutral messages such as customer complaints, likes, or questions. Since these posts directly address specific companies, the companies are expected to respond to them, but most companies lack the resources to respond to all such posts. The percentage of replies could thus be used as a proxy measure of a company's responsiveness.

Hashtags: Like mentions, hashtags (#) link posts to specific companies or subjects and may be either positive or negative.[2] Unlike with mentions, however, target companies are not expected to respond to hashtags. Twitter users and followers use hashtags to generate content, and they can be used to generate WOM advertising for companies.

The marketing value of hashtags and mentions lies in the connotations and implications of texts that carry information about the feelings, attitudes, and expectations of customers. Marketers and researchers extract business value from text analytics, which includes a wide range of analysis options from content classification to sentiment analysis. The analysis methodology based on voice recognition, texts, and other linguistic data is called natural language processing (NLP), a branch of AI. Sentiment analysis is one of the most valuable and popular methods for analyzing user-generated content.

Sentiment analysis: Sentiment analysis quantifies the mood of a text or tweet by comparing the number of positive and negative words,[3] subtracting the negative from the positive to arrive at a

[1] Shaha Al-Otaibi, Allulo Alnassar, Asma Alshahrani, Amany Al-Mubarak, Sara Albugami, Nada Almutiri, and Aisha Albugami, "Customer satisfaction measurement using sentiment analysis," *International Journal of Advanced Computer Science and Applications*, Vol. 9, No. 2, 2018, pp. 106–117, http://thesai.org/Publications/ViewPaper?Volume=9&Issue=2&Code=IJACSA&SerialNo=16 (accessed June 20, 2018).

[2] Ibid.

[3] Nadia F.F. da Silva, Eduardo R. Hruschka, and Estephan R. Hruschka Jr., "Tweet sentiment analysis with classifier ensembles," *Decision Support Systems*, Vol. 66, October 2014, pp. 170–179, https://www.sciencedirect.com/science/article/pii/S0167923614001997 (accessed June 20, 2018).

sentiment score. This score will depend on the machine learning algorithm used[1] and can range from -1 to +1, with "0" indicating a neutral statement. The sentiment score indicates the writer's level of satisfaction with the company.

Serious service crises: Time series of customer satisfaction

Calculating the average customer satisfaction level produces an overall index. Researchers and marketers also find it helpful to measure changes in satisfaction over time and in different periods. One way to do this is to carry out a time series analysis of the ups and downs of customer satisfaction and thus the company's performance over time. This is similar to what is often done for the number of likes and retweets. Another method is to link changes in customer satisfaction to specific events, decisions, announcements, and actions. Marketers can then use this data to measure the extent to which various factors affect customer satisfaction.

Researchers may pay particular attention to the impact of highly publicized service crises, comparing customer satisfaction before the event to that following the crisis. The usual window for this study is one week before and one week after the crisis.

Negative user-generated content

Just as positive tweets can be widely shared, enhancing the company's image, negative content can also be shared, harming the corporate image. Widely shared negative content can help marketers identify weaknesses in areas such as service delivery. Companies can then work to address those weaknesses, improve customer satisfaction, and reduce the number of negative comments.

Benchmarking

While a company's metrics can offer invaluable insights into its performance, weaknesses, and customer satisfaction, they are more meaningful when compared to those of competitors or the market average. This is why marketers do benchmarking, using the performance of competitors or market averages as reference points. Since information about all Twitter users is available to the public, companies can collect useful information about competitors.

Conclusion

To give students a better understanding of these concepts and allow them to apply them to a real-life situation, we have created an exercise based on the Twitter account of United Airlines (Twitter and Marketing Research: How High is United Flying?)
2018-10-26

[1] Natural-language processing (NLP) and machine learning (ML) are two separate areas of artificial intelligence (AI). ML is mostly concerned with statistical computation algorithms. For sentiment analysis in this study, we use an algorithm that is the combination of both (NLP-ML). This algorithm computes linguistic features.