Avoid Mixed Sentiments on Social Media Analytics

For all their supposed benefits, social media analytics projects can be a struggle. There's still business value to be found, but it's a matter of knowing where—and how—to look.
Dimmed Prospects—But Still a Source of Enlightenment

**At first, social** media analytics was the brightest star in the big data firmament—a shining example of the business benefits that big data applications could provide. Collect tweets and Facebook posts, analyze the text for keywords and customer sentiment, then use the findings to plan social media marketing and outreach strategies and to react to complaints and criticisms before they go viral. That was the idea, at least. In practice, it turned out not to be so easy.

As a result, social media analytics has lost some of its luster. In a “hype cycle” report published in August 2014, Gartner included sentiment analysis, text analytics and content analytics in a list of technologies and applications that were sliding into what it calls the “trough of disillusionment” due to a failure to meet user expectations. Alta Plana consultant Seth Grimes, in an April 2014 article posted in his Breakthrough Analysis blog, quoted other industry experts on the challenges of analyzing tweets and other casual forms of text. “It’s still a bit of a crapshoot,” said one. Meanwhile, other big data analytics applications have started burning more brightly—analyzing machine data from sensors on the Internet of Things, for example.

But there’s still business value lurking in social media data, if it can be successfully uncovered. This guide offers insight and advice on how to get started. First, we catalog social media analytics challenges to be aware of. Next we look at different facets of the analytics process and their potential benefits. We close with a case study on one organization’s efforts to mine social media data for useful business information.

Craig Stedman
Executive Editor
SearchBusinessAnalytics
No Snap Judgment in Analyzing Social Media Data

Social media analytics presents one of the biggest of all big data challenges. The data is unstructured text that requires the use of specialized natural language processing techniques and tools to do the analysis—and it never stops coming. Together, those factors can make it a tall order for businesses to make sense of social media data.

But there are good reasons for trying. Knowing what customers and the public in general think of a company and its products can help corporate executives adjust business strategies and react to problems before they get out of hand. Social media analysis offers a broader and more real-time alternative to surveying customers—and the views of users on Twitter and other social networks are free for the taking. All an organization needs to do is figure out how to tap into and take advantage of the streams of data.

That’s where the problems can start. Simply having access to social media data doesn’t necessarily mean that a business is ready to derive insights from it. Users need to figure out things such as how to score sentiments and how to code algorithms so they can understand nuances like sarcasm and jokes. The notion that data speaks for itself is particularly out of touch when it comes to social media analytics.

“It’s a misnomer that all this data is at our fingertips now and big data has made it available,” said Sarah Biller, president of Capital Market Exchange, a Boston-based company that analyzes data from social networks and traditional media outlets to help investment managers evaluate the price of corporate bonds and assess their bond portfolios. Biller added that organizations need talented individuals in-house who can normalize and structure all the data being collected so it can be run through analytics engines.
SKILLS IN SHORT SUPPLY
Finding that talent is another issue. Nowhere is the shortage of data scientists more acute than in social media analytics. The field is relatively new, and few analytics professionals have the necessary experience and expertise to manage all the complexities of mining social media posts for useful business insights.

Jiri Medlen, a senior text analytics specialist at digital payment processor PayPal, said the whole idea of social media analysis is so new that even experienced data analysts often have trouble getting started on it. To make matters worse, he added, there are no commonly accepted best practices on how to quantify the sentiments expressed in text for analysis.

Medlen recommended that companies have specific business goals in mind before jumping into social media analytics. But even then, he said, the return on investment from social media sentiment analysis can be questionable.

For example, even when he and his team are able to identify negative feelings about PayPal voiced on social media, it may be too late to prevent the PayPal member who voiced them from closing his account. Knowing a customer’s sentiment doesn’t necessarily translate to meaningful action.

“The biggest question is what to do with this kind of data,” Medlen said. “How is this going to impact the bottom line of the company? We still have to answer the question of value.”

TECHNICAL DIFFICULTIES ON INTEGRATION
And while social media data can deepen analytical models and findings, it can also add a new layer of complexity that poses technical challenges some businesses may not be ready to handle. Speaking at the RapidMiner World 2014 user conference in Boston, Usama Fayyad, chief data officer at Barclays PLC and former CDO at Yahoo Inc., said the rapid streaming...
and unstructured nature of social data makes integrating it into analytics systems a big headache.

Social media data can pose technical challenges some businesses may not be ready to handle.

Fayyad said he and his team at Yahoo used to think the standard 25 characteristics that the Internet company collected on every user qualified as big data—really big data. That information included things like age, demographics and search history, data used mainly to develop profiles of Yahoo users for targeted marketing. But then the analysts started to look at how they might be able to utilize data from the Facebook, Twitter and LinkedIn accounts of users—and that truly introduced them to the challenges of working with big data, Fayyad said.

Early tests of social media analytics applications showed promise, he added, but coping with the complexity of the unstructured data proved difficult. “We really don’t know how to deal with variety,” Fayyad said. “You can enhance any data set with all this and it makes it better. But it’s a monster.” —Ed Burns
Social Media Analytics Spans a Spectrum of Uses

One of the notable aspects of the online social networking phenomenon is that the full breadth of the interactions between participants can be documented. Systems capture who the interacting parties are and what they’re communicating to one another, along with statistics about the level of their social media activity.

And of course, in any scenario in which transaction histories can be recorded, there’s an opportunity to analyze those histories for business purposes—in this case, to monitor social networking sites for comments about a company and its products, and to measure the effectiveness of an organization’s social media strategy and outreach efforts in influencing things such as brand recognition and customer sentiment.

The term social media analytics is a catch-all phrase that’s used to describe any type of reporting and analysis associated with the business impact of social networks. But there are different facets of social media analytics that offer different kinds of insights.

Let’s consider three of them: reporting on the quantifiable aspects of social media interactions, analyzing social media content, and modeling and understanding the behavior of specific individuals within a social network.

Run the Numbers
Some examples of the statistical data that can be collected from social networks and then analyzed include audience distribution, number of impressions for posts, mobile-device interactions and responses by users, such as Twitter retweets and click-through numbers for embedded URLs. These metrics can be broken down by different dimensions, including time of day, geographic location, browser type and corporate domains. In addition to providing
useful information themselves, they’re the foundation upon which the other types of social media analytics metrics are built.

Content analytics then attempts to discern actionable information about the messages that are being posted by the users of social media sites. The types of analytics that can be done include identifying designated keywords in social media posts, monitoring for posts that refer to specific products or corporate brands, tracking customer sentiment based on positive or negative references to a company, and detecting problems that could pose a threat to a company’s reputation and revenues.

These types of analyses blend emerging text analytics tools and techniques with more traditional reporting approaches to scan social media content for specific language and to apply semantic analysis concepts to the collected text. Instead of just looking at the mechanics of social media interactions, such as the number of followers, you’re now looking at how the content of those interactions could affect the way your organization is viewed by people—and, ultimately, its business performance.

GETTING TO KNOW PEOPLE
That brings us to the third facet of social media analytics: examining the characteristics of the individual entities within social networks and how they’re connected to one another. Again, these analyses start with fundamental ideas, such as correlating demographic information with social media users. For example, an organization might collect data about the people following a particular Twitter account—sex, age, location, educational attainment, even things such as annual income or predisposition to purchasing particular kinds of products.

Such information allows companies to seek to relate individual or aggregate user profiles to
possible actions in response to specific types of social media content or promotional offers. Social media data analysis could point, say, to unmarried males between the ages of 18 and 34 who are likely to respond positively to a discount offer broadcast on a company’s Twitter feed or Facebook page (e.g., “Use code OFF15 for 15% off until 6:00PM EDT!”). Ongoing analysis of the uptake on the offer, and comments posted about it, could help the organization refine the message and how it’s communicated to improve the response rate and the online feedback.

Analysis of that sort feeds into a more complex layer: looking at how different participants in social media communities interact with each other. Doing so can help identify influential people—for example, Twitter users with a large group of followers or who post information, such as links to product reviews in a blog, that is likely to influence the opinions of others. Such people could then be singled out for closer monitoring to help spot trends or discussion threads that a company needs to respond to, either to reinforce positive comments or to answer negative ones.

Each of these different parts of the social media analytics process provides some level of insight into corporate perception and the effectiveness of a social media strategy. Integrated together, they can be even more powerful in pointing the way toward opportunities for improving both strategic and tactical approaches to the social media environment—ideally resulting in more revenue and higher profits. —David Loshin
Sentiment Analysis Helps Investors Place Bond Bets

The ongoing civil war in eastern Ukraine provides an example of a situation that social media analytics could be applied to—for example, by investment portfolio managers looking to evaluate risks to their corporate bond holdings in companies that might be affected by the conflict. Further, the political crisis demonstrates why social media sentiment analysis might be a difficult exercise for organizations to successfully pull off.

There’s no shortage of news organizations and commentators offering information about the Ukrainian crisis that might be relevant to bond investment teams, both through traditional channels and posts on Twitter and other social media outlets. The challenge, though, is to do systematic sentiment analysis of news stories and social media feeds that assess new developments and point to their possible ramifications for different types of investments. It could be safe to leave money in the bonds of defense companies but not those of natural gas suppliers, depending on the details of what’s happening on the ground.

Those are the kinds of questions that Capital Market Exchange grapples with on a daily basis. The Boston-based analytical services provider was founded three years ago to provide global bond investors with insights derived from social media and business news posts. Sarah Biller, the company’s president, said the key to making the process work is not only to look at whether a post is positive or negative, but to also understand more deeply why it was written in the first place.

“In the investment community, [sentiment analysis] is more often not actionable,” Biller said. “We went with the idea that if we could ascertain the second layer of sentiment—not just what’s positive and negative, but what drove experts to reflect positively or negatively—we could model that.”
A CURATED SELECTION OF DATA

Instead of looking at the entire stream of Twitter data on a topic like the strife in Ukraine, or every business news publication, Capital Market Exchange gathers information from a curated list of what it considers to be influencers in the bond market. That helps the company put their comments in context, Biller said. It then feeds the data into a homegrown analytics engine, which is partly based on the R programming language. Customers get access to the analytics output through a Web-based dashboard.

Biller said the company had to build its own system for analyzing the data because there weren’t any commercially available options that fit its needs. One problem common on any social media analytics project is how to score posts that may include sarcasm, irony or other counterintuitive forms of language. That problem is compounded in the investment world, where commentators often use unique jargon that off-the-shelf sentiment analysis systems don’t understand.

For example, Biller said investors typically want to stay away from “rich bonds,” ones that are trading at higher prices than financial analysts think they should be. But, she added, most commercial social media analytics engines are likely to interpret the word rich as a positive reference.

“There are tons of off-the-shelf sentiment products, but we found that we couldn’t use them because of the specific word choices the investment community uses,” Biller said. It’s a must, she noted, to “make sure the technologies are not picking up incorrect nuance.”

LOOKS DO MATTER

Another important step for Capital Market Exchange is visualizing the data. That may not seem like a core piece of social media analytics, but the company’s clients need to be able...
to look at the information in the dashboard and quickly discern what it shows. “You have to display the data in a way that makes sense to the investor,” Biller said. For example, the dashboard includes “a lot of color and blinking lights” to help make things clear for users.

Social media sentiment analysis isn’t an exact science, and Biller said clients use her company’s findings only as one piece of information in their decision-making processes—or, at least, they’d be wise to. “We aren’t perfect,” she acknowledged. “I wish we had a crystal ball, but we don’t.”

In the end, how successful Capital Market Exchange is at analyzing the data it collects will be judged on the basis of how its clients do financially by using the findings, according to Biller. “It’s a black-and-white world we work in: Do we make investors money or not? Ultimately, we want to give them enough information that they can better anticipate what’s going to happen tomorrow.” —Ed Burns
ED BURNS is site editor of SearchBusinessAnalytics; in that position, he covers business intelligence, analytics and data visualization topics. Prior to joining the site, he was a news reporter for SearchHealthIT. Email him at eburns@techtarget.com and follow him on Twitter: @EdBurnsTT.

DAVID LOSHIN is president of Knowledge Integrity Inc., a consulting and development services company that works with clients on big data, business intelligence and data management projects. He also is the author or co-author of various books, including Using Information to Develop a Culture of Customer Centricity. Email him at loshin@knowledge-integrity.com.