

The Economic Times

Title : #Sentimental with Twitter

Author : Ashok Banerjee

Location :

Article Date : 08/13/2015

The availability of regular and real-time news shapes market sentiments. Enter news analytics

Consider a news headline posted on Twitter: Rupee soars & profit of Infosys dips.

Ask a machine to 'read' the news and 'tell' us what the news is all about. What appears to be a simple matter of reading for a human may get quite difficult for a machine unless it is 'trained' properly .

The headline includes two nouns ('rupee' and 'Infosys') and two verbs ('soars' and 'dips'). If the machine reads the verbs in the wrong places with respect to the nouns, the meaning of the headline completely changes. Thus, the algorithm should be developed in such way that a machine 'reads' and 'understands' the story in the same way a human does. This would involve several things that include associating every verb and adjective with the correct noun.

Welcome to the field of news analytics, tools that convert a text into sentiment score. It is known that stock markets worldwide are influenced by sentiments more than fundamental factors, and regular and real-time news shapes market sentiments. A discipline that measures the relevance, sentiment and novelty of news has emerged. This is news analytics.

Even a decade ago, financial news was confined to the print media and television. News-gathering has now become more technology-intensive, with crawlers that monitor large and cascading information on the world wide web as well as in databases. The web today is full of potentially investment-relevant text. There are three broad categories of news: pre-news (raw information that reporters access before they write the news), news (public information in print, on the web and from online news vendors), and social media (reviews, blogs, message boards). Analysts, mostly with the support of machine-learning language, now use real-time news to capture sentiments carried in the story at lightning speed and send trade signals to capture the alpha -the excess return of the fund relative to the return of the benchmark index.

Read: Eats, Shoots & Leaves

News analytics has seen the convergence of experts from diverse fields like finance, computer science, psychology and linguistics to decipher and automate news 'reading' and news extraction. Various news sources carry information on products, financial assets, customer feedback and product reviews. It is necessary for an algorithm to create a dictionary of its own to capture the sentiments ingrained in these stories.

The good news with financial reporting is that reporters use a 'limited' vocabulary . So, developing a working dictionary of relevant words is not difficult. The first step in extracting the sentiment of a news item is to find the 'relevance' of the news with regard to a noun in question.

For example, in our Twitter headline, 'Rupee soars & profit of Infosys dips', one has to first decide whether the story belongs to the company mentioned. The headline usually captures the essence of a news. So, the relevance of a company is searched in the headline. If the headline carries the name of a company or product, the story 'belongs' to that particular company or product.

A corporate action, for example, the announcement of earnings or a stock-split, is carried in multiple news channels. The novelty of a news report is highest for the channel that captures it first. Sentiment

score is finally obtained through the use of a proprietary algorithm. This, of course, requires handling challenges such as false positives.

This process of sentiment extraction supervised by experts works well in the case of structured news stories. However, the job gets complicated when one uses news or information from social media. Many consider sentiments expressed in social media as highly biased and, at times, unreliable. The ability to publish information in this media is easy. Which is why the entry barrier is low.

Business entities now use social media in a big way to promote their products or services, capture customer voice and improve one's corporate image. On stock message boards, valuable information is posted at times by CEOs. But on many occasions, message boards carry 'noise'.

Another problem is the use of non English texts. It is quite challenging for an algorithm to crawl through the maze of bilingual texts, symbols and wrongly spelt words to extract any meaning. Still, the importance of social media in our information-consuming and generating world can't be understated.

Since its inception in 2006, Twitter has grown immensely popular. It has also attracted researchers from various disciplines. While the use of social media data to gauge 'customer delight' about a product has been in vogue for quite some time, analysing social media news to predict stock returns is a recent phenomenon.

Tweet, Chirp, Squeal

Johan Bollen, Huina Mao and Xiao Jun Zeng in their research paper, 'Twitter Mood Predicts the Stock Market' (Journal of Computational Science, March 2011), measure collective mood states from Twitter feeds and calculate their correlation with the Dow Jones Industrial Average over time. Research also shows that even though people tweet more during non-trading hours, tweets during trading hours carry more 'market relevant' information.

The preliminary finding on the efficacy of social media posts in capturing stock market sentiment is encouraging. This is good news for news analysts.

The writer is professor of finance, IIM-Calcutta

Global research shows that even though people tweet more during non-trading hours, tweets during trading hours carry more 'market-relevant' information



Ashok Banerjee