



Analysis of tweets Jan-March 2011 in financial sector

Introduction

As an extended study, we stored almost 3 months worth of tweets covering the financial sector during the period January - March 2011. We wanted to look at how the various banks and financial institutions fared against each other, and what trends we could see in the tweets people were making.

Whilst sentiment analysis gives you an overview of the public perception of your brand or product, the strength of our technology is to go deeper than sentiment allows and show you **why** the public thinks the way it does. Our technology is able to identify the context of conversations, and how this relates to companies and products in particular, so that Business Intelligence can be gathered against ongoing marketing campaigns and other social media activities. This report shows how our service can be used to give an unparalleled insight into the public's perception of a brand or product, and how it can complement a standard sentiment analysis approach.

Analysis

In the following sections we will look at a number of banks or financial services, and what the analysis of their data shows. A number of word clouds are shown - these show the word that is the focus of the analysis in the centre, with words in concentric circles surrounding it. The closer these other words are to the centre, then the closer in context they are with the word under analysis.

Trending analysis over a number of days is also discussed - this is not presented as it is simply lists of words although these can be provided on request.

Example contextual analysis

We could look to examine a single topic using contextual analysis. For example, let's start with an investigation into words surrounding the concept of 'mortgage'.

If we start with analysis on a day towards the end of March, we see the analysis as shown in Figure 1. This shows the words found in context with mortgage and indicates that conversations were taking place that focussed on items such as refinance, rates, home reports, and words indicating a general discussion of what is required such as cash, money, unsecured, guaranteed and so on.

However let us say that we want to investigate things further. What if we wanted to know what words come up in context when we ask the question: What is the best mortgage? This analysis can be seen in Figure 2, where words that are found in context with **both** mortgage and best are shown. This indicates that the words 'cost' and 'payoff' are closest in context, which shows the public's main concerns when assessing mortgages against each other.

This highlights the advantages of our approach in that we can analyse **context within context**, to hone in on the conversations of particular interest.

We can take this still further however, since we can look at how the words used in context with 'mortgage' and also with 'best' change over time. For



Figure 1: Word cloud showing words in context with 'mortgage' from data in day in March



Figure 2: Word cloud showing words in context with 'mortgage' and also in context with 'best' from data in day in March



Figure 3: Word cloud showing words in context with ‘mortgage’ and also in context with ‘best’ from data in day in January.

example, we could look at a date earlier in the year in January, as shown in Figure 3.

This analysis shows ‘best’ coming up in close context with ‘bad’, and also words like ‘sharks’ and ‘unsecured’ coming up. This is from conversations discussing the warnings of mortgages that seem to be the best but in fact do not have good terms and should be avoided.

By tracking the changing conversations in this way over time, it is possible to see the effect of any marketing activity, or any other broadcast event that could have an effect on public opinion.

Again, the ability to go to this level of detail and to identify trending in the analysis sets our services apart from any other company in this space.

Trending in analysis

The above analysis shows the words that are important around certain contexts, and shows the word clouds that are produced. If we continue this analysis for words in context with both ‘best’ and ‘mortgage’ on a daily basis over the three month period, then we see that in the tweets towards the end of March, ‘refinance’ comes up a lot, indicating that this topic was discussed by the public, and that the ability to refinance was discussed in context with the best mortgage. Other words of interest towards the end of March were ‘rate’, and ‘equity’.

However this changes as we look at data at the beginning of March. We no

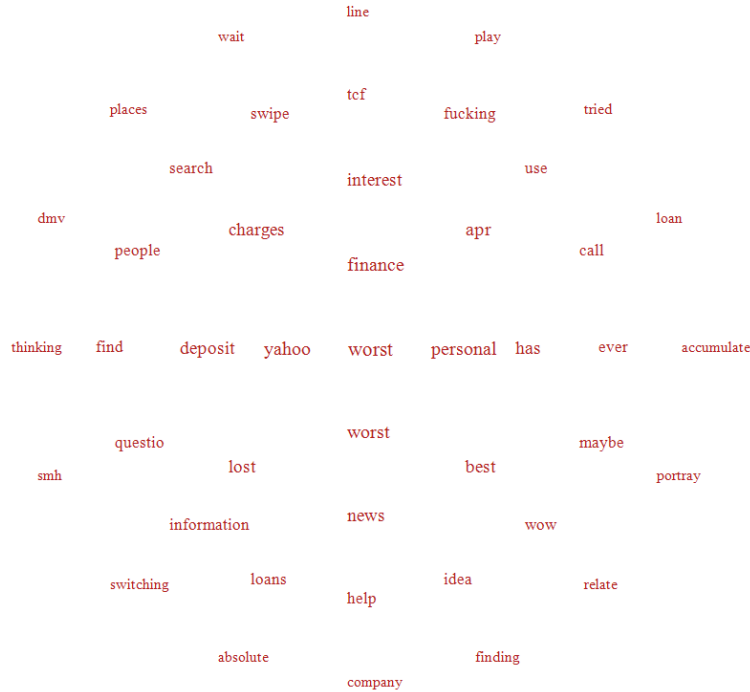


Figure 4: Word cloud showing words in context with ‘worst’ from data in day in January.

longer see ‘refinance’ but words like ‘rate’ and ‘interest’ continue to be seen.

In February we see the story changing again, with discussion of ‘analysts’ and certain place names or companies such as ‘Philadelphia’, ‘Newcastle’, ‘California’, ‘Australian’, ‘Morgan’ and ‘America’ coming through with a focus on issues more related to the health or otherwise of these companies.

In January there are other issues coming through such as ‘foreclosure’ and ‘bankruptcy’, perhaps reflecting the difficulties some households go through after the expense of Christmas.

The trending analysis can be seen in more detail on request and shows how trends in word usage for the concept of ‘best’ ‘mortgage’ come through strongly in our analysis.

Interestingly, if we repeat the analysis for ‘mortgage’ with this time the word ‘worst’ in context, we do not find many results, showing that the public does not (on Twitter at least) share their thoughts on what is the worst type of mortgage.

However if we search for ‘worst’ on its own, then we see a different picture emerging. If we looked at the middle of January (the same day as for the previous best mortgage analysis), then we see the contextual analysis as shown in Figure 4. This clearly shows a lot of tweets focussing on Yahoo’s personal finance service in a negative way.



Second trending example

Another example could be to look at a topic like 'bailout'. This is a rather emotive topic for the general public, and as you would expect the focus of their ire changes over time.

Initially, at the start of January, the main focus is on the UK government bailout and the fact that the money will not be returned until years later. Following this, there is discussion of the American bailouts, which then turns quickly to a discussion of Portugal, and a possible bailout there, with reference to the Euro zone and also Japan.

Towards the end of January, we see reference to Nick Clegg, the UK's Deputy Prime Minister, and statements that he has made on the government bailout, following which we see comments made by President Obama coming through in discussions made over a number of days. This continues with subsequent days focussing on the US Federal government's role and also the Republican party. Then a statement by Deutsche Bank Chief Executive Officer Josef Ackermann becomes prominent that discusses the bailout risk for hedge funds.

In February, this trend continues, with some topics that reach prominence remaining in context for a number of days. We start to see Greece mentioned due to its difficulties, and then Ireland and Libya too are seen in context with 'bailout'. For the rest of February and March we see these three countries, and Greece and Ireland in particular, continuing to be seen in context with 'bailout', as the discussion over the terms of any agreement is discussed on twitter regularly over this period. We also see Germany coming up in context, since it was the main country involved in the bailout deals for these countries.

In this way the larger news items can be tracked against key words and easily identified over time.

Just as above, it would be possible to examine any of these issues in greater detail by examining, for example, the context of Ireland within this bailout context. This would give an insight into the words that are found in context with Ireland only in the context of 'bailout' - allowing a focus to be given to this topic only. This could be useful, since (for example) Ireland were involved in the Six Nations rugby tournament during this period (which is sponsored by RBS), and we would not want the large number of tweets that were made during this tournament to influence our analysis of the purely financial impact of the 'bailout' provision that Ireland arranged during this time. This again highlights the importance of our approach in being able to **analyse in context**.

Context and competitor analysis

One issue that is hugely important and that our technology is ideal for is the comparison of competitors against each other, particularly when that is done in context. What do we mean by this? Well, whilst a summary count of how many positive and negative tweets have been made about Company A versus Company B is useful, it does not give you insight into **why** one company is doing well, or another one poorly.

In addition, what if it is difficult to describe what is being analysed in terms of sentiment? For example, let's say we wanted to know which banks were close in context to words such as: interest, rate, payment, payoff, return? Banks closer in context implies more conversations about the bank that impinge on

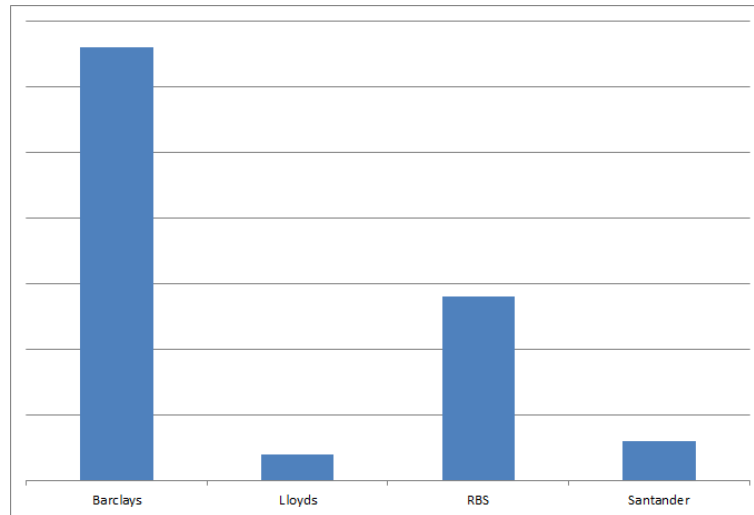


Figure 5: Bar chart showing banks context against key words in single day in January.

these terms. We can examine a couple of examples of this using the figures for discussion in our analysis.

Figure 5 shows the analysis of the banks in context against the key words outlined above. Here we can see that Barclays are closest in context on this day, with RBS their closest rival, and Lloyds and Santander lagging far behind. We can identify whether these ratings reflect market share or whether they are linked to a particular issue that was discussed that day.

We can take this further by analysing the same words on a day in February. We can see this analysis in Figure 6 and here we see a different story emerging. RBS have dropped out of the analysis, with Santander now the closest in context. Lloyds are close to Santander which is similar to the analysis in Figure 5, however Barclays have dropped far behind. We also see that Northern Rock are discussed in context on this day.

This form of analysis shows how competing banks can be measured against each other in terms of how close they are to key concepts that are important to them. It shows how changes in public perception can be identified, and tracked over time. This can be fed into internal metrics in order to improve performance of marketing campaigns or other such activity.

Royal Bank of Scotland

A further example of this can be seen with the Royal Bank of Scotland. If we search for references to 'rbos' we do not find many tweets at all. However we do see a large number of tweets coming through for 'rbs'. There are problems with this which we will discuss below.

Initially at the start of January, we see a number of references to the Chief of RBS Stephen Hester and his bonus payment that at that stage had not been fully approved. However we also see a large number of Portuguese words coming

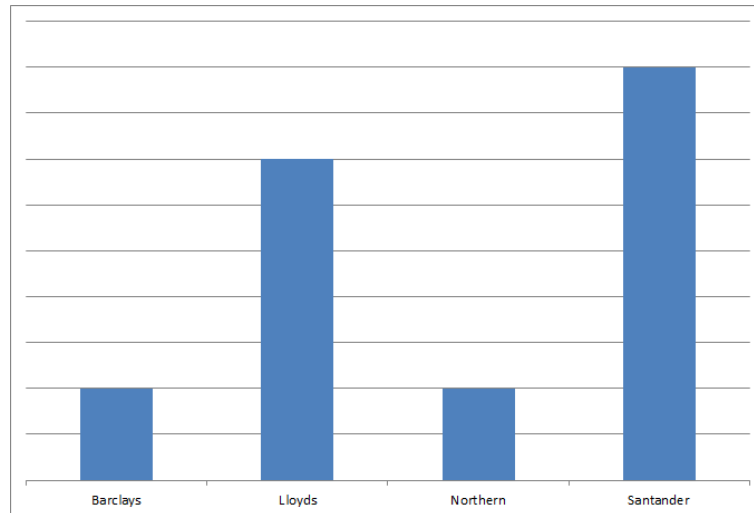


Figure 6: Bar chart showing banks context against key words in single day in February.

up in context with ‘rbs’ and if we investigate this a little further we see that there is a Brazilian TV station called RBS TV. In addition, in American football, RBs stands for ‘Running backs’ and is often used in tweets as an acronym as shorthand since it is an important position in the game.

This has three problems. The first is at the consumer side, in that anyone searching solely for RBS will not only receive tweets to do with the bank but also other topics that will certainly not be of interest to them. If they were searching for information on the bank they may give up and search for another bank instead.

The second problem is that the brand message of RBS whether promoted by the public or by media campaigns will be diluted due to the other meanings of ‘rbs’ as noted above.

The third problem is that the analysis of social media data will therefore require an additional step to remove this irrelevant data from the analysis. This will cost money and time for most approaches.

We would therefore recommend that RBS refer to themselves as RBOS as much as possible since this would differentiate themselves on twitter and other social media. This would obviously take some time to permeate public consciousness but would result in a clearer brand than is currently the case as shown above.

The above problem could make analysis of trends around RBS hard, but in our case, since our analysis is contextual, it becomes easier for us if we simply add in an additional context which would rule out any references to the TV station and also any sporting references as well.

As an example, we can focus on Stephen Hester within the context of ‘rbs’ and see how this affects the analysis.

We see that the majority of days do not have any analysis at all for him, indicating that no tweets were made on these days on this topic, although there



is some discussion at the beginning of January about the amount of money that Hester is likely to receive as a bonus.

The next mention is in the middle of February, where we see a number of words coming up in context: Hicks and Gillett. This is in relation to Hester's role as Chief of RBS in effectively forcing the sale of Liverpool Football Club who were owned by Hicks and Gillett thanks to loans made by RBS. Hester's decision to call in the loan and not allow any refinancing forced the sale of the club against the wishes of Hicks and Gillett.

Following this over a ten day period in March Hester's bonus is discussed as it was finally approved by UK Financial Investments that manages the 83 percent UK taxpayer stake in RBS.

Discussion

This white paper highlights the additional insight that can be gained by the use of our contextual analysis techniques.

We can analyse context within context, meaning that discussions that are not relevant to the topic under investigation will be ignored automatically.

We can show how word use changes over time, and how that can affect a brand or product. We can show trends in discussions around a particular topic, for a brand or for a combination of topic and brand. This gives unparalleled analysis for your brand which most importantly gives you insight into **why** the general public feel the way they do.

We have also shown how competitor analysis can be undertaken for a particular context, which is depending solely on your company's own profile, and that would be difficult if not impossible using sentiment analysis techniques. This can be tied into a company's internal Business Intelligence metrics, so as to give a quantifiable feedback on marketing campaigns (whether digital or otherwise), and therefore allow the ROI to be determined effectively through social media for the first time.

Contact

If you would like to discuss the contents of this whitepaper, or to see how this analysis could be used to benefit your own company, and the pricing level for the service you require, then please contact Andrew at andrew@brand-aura.com.