April 2013

Ratings revisited: Textual analysis for better risk management

Andreas Kremer, Wolfgang Malzkorn, and Florian Strobel

Textual data—from analysts' reports to blog posts—can help banks enhance their credit-risk models.

Most banks use credit-rating models to help them make decisions about lending to companies. Such models are indeed a requirement for banks using Basel II's internal-ratings-based approach. But these models often have significant shortcomings. First, they are frequently backward-looking. Second, they rely on borrowers' formal financial reporting, which means that data are always at least 6 months old; toward the end of the fiscal year, data are nearly 18 months old. Third, qualitative assessments of borrowers are often simplistic. And finally, many banks rely on their credit-rating models to provide both a current snapshot and a longer-term view, with the result that they do neither well.

Textual information can help banks overcome some of these challenges and improve their credit-risk assessment, in particular their approach to qualitative assessment. This information includes professionally produced content such as analysts' reports and business journalism, as well as informal texts such as blogs and posts on social networks. Compared with the financial information available about small and midsize enterprises (SMEs) or corporates, the amount of textual content about companies is immense and provides a wealth of information. News articles describe the latest developments of companies; analysts' reports provide insightful analyses on companies' strategies, competitive positioning, and outlook; product ratings on online-shopping sites provide unfiltered views of customer satisfaction; and microblogs such as Twitter distribute the latest news (and sometimes gossip) with unprecedented speed.

Enormous quantities of textual information are available; this information offers companies a deep look at their health and performance, and it is notoriously difficult to use. But we have developed a prototype model that can identify and quantify sentiment within a trove of textual information, and it has performed extremely well in pinpointing default risks at a very early stage on a wide range of performance measures. We argue that if banks can put even a portion of this information to use in their systems, the accuracy, timeliness, and forward-looking character of their credit-risk-assessment systems would

all be improved. And textual analysis can also help banks in other areas, improving their traditional analyses of industries and sectors.

Challenges of textual data

The challenges of mining this information and separating the signal from the noise are substantial. To use textual data, banks must first face a practical challenge: computational capacity. The amount of text-based information available is already enormous, and it's getting bigger. Banks' computers would strain to read and analyze it in daily operations or even to batch process it for model development. A database with news articles on about 1,000 companies easily exceeds 20 GB, orders of magnitude more than a financial database on these companies. Storing this much data is not difficult, but any kind of statistical analysis becomes an "overnight job," even with optimized algorithms and systems.

Second, textual data are unstructured. While it is relatively easy to analyze financial data in a statistical way—figures become meaningful at a certain size and in relation to sample averages—texts are a priori meaningless to a computer. There are no standard or statistical procedures for a machine to analyze and interpret texts.

Third, texts are often ambiguous. In particular, the meaning of short messages in social media is difficult to interpret—even for humans. While complicated sentence structures can be taught to computers, the concept of sarcasm or irony is extremely difficult. In fact, almost all the semantic difficulties of written language pose immense problems for machines.

Sentiment analysis

Companies in other sectors have already begun to employ a new technique, sentiment analysis, that banks can also use to get around these obstacles. The basic idea is simple and elegant: textual information in any form (words, sentences, paragraphs, articles, or books) is assigned a "sentiment index"—a number that represents a kind and degree of opinion expressed by the writer, such as optimism, trust, skepticism, mistrust, pessimism, and so on. Gauging sentiment with an index makes it possible for machines to analyze the information; it can be converted, aggregated, and compared. And the index can be used with statistical analysis to build prediction models. Obviously, the difficulties come in the details of assigning the sentiment index.

At the core of the process is a lexicon that lists words or phrases that represent a certain kind of sentiment and, importantly, reflect the specific context in which the text appears. Phrases that a novel's heroine might use to describe her pessimism might not be the same as those—such as "legal proceedings," "owner disputes," or "financing problems"—

that indicate financial concerns about companies. The lexicon must be contextualized appropriately if it is to have the kind of accuracy banks need.[1]

Just as important as properly defining the lexicon is selecting and filtering data sources. A broad search will yield more potentially relevant articles to be analyzed, but it will also pull in much irrelevant material. That poses a problem when seeking information about companies with ambiguous or generic names. When stories about Berkshire Hathaway, the US multinational conglomerate, are wanted, stories about Anne Hathaway, the US film actress, are not, as one hedge fund found out to its cost. With smart search tags and additional text filters, these challenges can be overcome.

Applications and benefits

Sentiment analysis and the information it yields can improve banks' credit-rating models, and it can also help with two other important tasks.

In rating models, banks can use the sentiment index as an additional rating factor. Information gleaned from text searches is aggregated quarterly into a sentiment index for each company. After statistical analysis, the index is then integrated into the rating system at an appropriate weight. This can be particularly valuable in assessing new corporate customers for which banks typically have only limited information, most of it provided by the customer. A systematic screening of public information can reveal important additional insights. In emerging markets, where reliable customer data are scarce, the analysis of textual information can yield insights as well (exhibit).

Another use is in early-warning systems. Due to the timeliness of the information and the high level of automation that is possible, we foresee a great benefit in applying the sentiment index as an early indication of a company's troubles, either by itself or integrated into an early-warning system with other financial, industry, and behavioral factors. In such a system, the sentiment analysis would automatically screen all relevant news articles. These articles could be filtered based on defined thresholds and presented to the relevant credit officers for further action. The system would be valuable for banks and for regulators, which might use it to monitor public sentiment about the banks under their jurisdiction. If regulators could detect shifts in sentiment that portend danger for a bank, they would be able to move early to address the problem and bolster confidence. Text-based early-warning systems could be applied to SME portfolios of customers with revenues of €10 million or more and even to large commercial customers. With lexicons of local languages, text-based early-warning systems can also allow banks to monitor remote portfolios systematically. For example, Western European banks might use such systems to monitor portfolios in Eastern Europe; Malaysian and Singaporean banks might monitor customers in new markets such as China, Indonesia, and Vietnam.

[1] Tim Loughran and Bill McDonald, "When is a liability not a liability? Textual analysis, dictionaries, and 10-Ks," 2011, *Journal of Finance*, Volume 66, Number 1, pp. 35–65.

4

Exhibit Textual analysis has many applications in banks' risk-management systems

Example application

Early-warning system for SMEs[1]	• Conduct daily or weekly sentiment monitoring of SME customers (>€10 million turnover) • Integrate sentiment factor into early-warning system
Remote or international customer monitoring	• Automatically screen remote international portfolios • Conduct benchmarking with own risk assessment of local credit officers
Financial-market monitoring	• Screen news on banks and financial markets on a daily basis to identify sentiment shifts • Use in regions where credit-default-swap markets are not liquid and many counterparties are semipublic companies
Industry-trend monitoring	• Systematically analyze published expectations on industries/sub • Use sentiment index as input in portfolio models or economic-forecast models
Additional risk factor in emerging markets	• Include sentiment index in rating models in emerging markets such as Indonesia • Compensate for data scarcity with public information

[1]Small and midsize enterprises.

Additionally, banks could deploy these tools in industry-trend analysis. Econometric analysis has proved useful in predicting an industry or sector's development, but sentiment analysis can complement it to yield a better result. In the short to midterm, an automated and objective analysis of news articles on industry outlooks can be helpful. An analysis of the sentiment regarding certain industry trends can help credit portfolio managers detect early shifts in opinion and initiate remediation measures. Here the sentiment index could be one input in a portfolio model; such crowd-sourced opinion can be used to influence banks' decisions on where and how to grow.

In all these applications, sentiment analysis has distinct advantages over traditional risk measures. First, sentiment analysis accesses a completely new kind and source of information; it retrieves default factors such as extraordinary events that are hardly covered in today's rating models. Because the information it analyzes is brand new and

reflective of reality, it overcomes the time lag present in almost all current rating models. Another advantage is that some news articles, such as opinion pieces and commentaries, are forward looking. Adding such pieces to the sentiment index can give it a prospective bias that nicely counteracts the retrospective tendencies of traditional risk systems.

Two other benefits are also important. Sentiment analysis can be largely automated, including the selection and filtering of articles, the calculation of the sentiment index, and the production and delivery of reports to the people responsible for them. Additionally, sentiment analysis provides a kind of objective measurement of qualitative information. It can help risk managers challenge front-office assessments; similarly, it can be a means to enforce regular monitoring of company activities. This is particularly important in large, geographically diverse banks, helping the central risk-management group monitor the quality of the credit portfolio.

A prototype analysis

To demonstrate the power of this new approach, we developed a prototype model tailored to the specific needs of banks and their SME or corporate credit-risk models.[2] The system uses business journalism as its input data. It analyzes news articles with a proprietary lexicon, derived from a comprehensive analysis of a substantial body of content by industry and banking experts, and produces a sentiment index.

We validated the performance of the model by back-testing it with more than 400,000 articles on over 700 companies from different industries, all with at least $10 million in annual revenues. The sample includes listed and unlisted companies; both subsets include some companies that defaulted during the period in question. We assessed the performance of the model with the measures typically used by banks for this purpose: an accuracy ratio (derived from the number of predicted defaults that actually defaulted, which is an assessment of the predictive power of the model) and a score for false negatives (predicted defaults that did not happen).

The accuracy rate for the privately held companies in our sample (by far the more interesting and challenging credit decisions) ranges from 55 to 80 percent, depending on a progressively conservative definition of deterioration and default. The model's performance compares well with a good early-warning system at a leading bank but has the advantage of working from a new source of information that complements traditional financial information.

We also reviewed the performance of the lexicon. A review of a sample of articles revealed that three out of four had been classified correctly—that is, the number and kind of phrases

[2] At present, the prototype relies on a German-language lexicon, contextualized for Europe, and does not accommodate social media or other forms of content. Lexicons in English and other languages, contextualized for other parts of the world, as well as a social-media capability, are under development.

in the article, cross-referenced to the lexicon, determined an overall positive or negative sentiment for the article in the same way that an experienced reader would.

"Trumpet" picture analyses, named for the shape of their curves, are often used to depict the timeliness of default detection by rating models or individual variables. For the listed companies in our sample, defaults were detected as much as two years in advance of the bankruptcy filing. For our final assessment, we compared the sentiment index of listed companies with the performance of their shares. A correlation analysis found that, on average, more than 90 percent of all companies in the sample showed a positive correlation between the two. Of the remaining companies, in many cases, we found that share prices did not vary much because they are thinly traded. Note that the prototype is not a tool to predict stock-market movement; rather, the correlation demonstrates that the prototype gathers and assesses public information in much the same way as market participants.

● ● ●

Using textual data to perform sentiment analysis can help banks develop more accurate early-warning systems, improve their credit-rating models, and create better portfolio-management systems. It also has the potential to help banks better understand customer needs, improve customer satisfaction, and ultimately, shape long-term strategy. □