

download it into your computer. If you have an analytics tool that is capable of text mining, use that; if not, download RapidMiner (rapid-i.com) and install it. Also install the text analytics add-on for RapidMiner. Process the downloaded data using your text mining tool (i.e., convert the data into a structured form). Build models and assess the sentiment detection accuracy of several classification models (e.g., support vector machines, decision trees, neural networks, logistic regression, etc.). Write a detailed report where you explain your findings and your experiences.

Internet Exercises

1. Survey some text mining tools and vendors. Start with clearforest.com and megaputer.com. Also consult with dmreview.com and identify some text mining products and service providers that are not mentioned in this chapter.
2. Find recent cases of successful text mining and Web mining applications. Try text and Web mining software vendors and consultancy firms and look for cases or success stories. Prepare a report summarizing five new case studies.
3. Go to statsoft.com. Select Downloads and download at least three white papers on applications. Which of these applications may have used the data/text/Web mining techniques discussed in this chapter?
4. Go to sas.com. Download at least three white papers on applications. Which of these applications may have used the data/text/Web mining techniques discussed in this chapter?
5. Go to ibm.com. Download at least three white papers on applications. Which of these applications may have used the data/text/Web mining techniques discussed in this chapter?
6. Go to teradata.com. Download at least three white papers on applications. Which of these applications may have used the data/text/Web mining techniques discussed in this chapter?
7. Go to fairisaac.com. Download at least three white papers on applications. Which of these applications may have used the data/text/Web mining techniques discussed in this chapter?
8. Go to salfordsystems.com. Download at least three white papers on applications. Which of these applications may have used the data/text/Web mining techniques discussed in this chapter?
9. Go to clarabridge.com. Download at least three white papers on applications. Which of these applications may have used text mining in a creative way?
10. Go to kdnuggets.com. Explore the sections on applications as well as software. Find names of at least three additional packages for data mining and text mining.

End-of-Chapter Application Case

BBVA Seamlessly Monitors and Improves its Online Reputation

BBVA is a global group that offers individual and corporate customers a comprehensive range of financial and non-financial products and services. It enjoys a solid leadership position in the Spanish market, where it first began its activities over 150 years ago. It also has a leading franchise in South America; it is the largest financial institution in Mexico; one of the 15 largest U.S. commercial banks and one of the few large international groups operating in China and Turkey. BBVA employs approximately 104,000 people in over 30 countries around the world, and has more than 47 million customers and 900,000 shareholders.

Looking for tools to reduce reputational risks

BBVA is interested in knowing what existing clients—and possible new ones—think about it through social media. Therefore, the bank has implemented an automated consumer insight solution to monitor and measure the impact of brand perception online—whether this be customer comments on social media sites (Twitter, Facebook, forums, blogs, etc.), the voices of experts in online articles about BBVA and its competitors, or references to BBVA on news sites—to detect possible risks to its reputation or to possible business opportunities.

Insights derived from this analytical tool give BBVA the opportunity to address reputational challenges and continue to build on positive opinions. For example, the bank can now respond to negative (or positive) brand perception by focusing its communication strategies on particular Internet sites, countering—or backing up—the most outspoken authors on Twitter, boards and blogs.

Finding a way forward

In 2009, BBVA began monitoring the web with an IBM social media research asset called Corporate Brand Reputation Analysis (COBRA), as a pilot between IBM and the bank's Innovation department. This pilot proved highly successful for different areas of the bank, including the Communications, Brand & Reputation, Corporate Social Responsibility, Consumer Insight, and Online Banking departments.

The BBVA Communication department then decided to tackle a new project, deploying a single tool that would enable the entire group to analyze online mentions of BBVA and monitor the bank's brand perception in various online communities.

The bank decided to implement IBM Cognos Consumer Insight to unify all its branches worldwide and allow them to

use the same samples, models, and taxonomies. IBM Global Business Services is currently helping the bank to implement the solution, as well as design the focus of the analysis adapted to each country's requirements.

IBM Cognos Consumer Insight will allow BBVA to monitor the voices of current and potential clients on social media websites such as Twitter, Facebook and message boards, identify expert opinions about BBVA and its competitors on blogs, and control the presence of the bank in news channels to gain insights and detect possible reputational risks. All this new information will be distributed among the business departments of BBVA, enabling the bank to take a holistic view across all areas of its business.

Seamless focus on online reputation

The solution has now been rolled out in Spain, and BBVA's Online Communications team is already seeing its benefits.

"Huge amounts of data are being posted on Twitter every day, which makes it a great source of information for us," states the Online Communications Department of this bank. "To make effective use of this resource, we needed to find a way to capture, store and analyze the data in a better, faster and more detailed fashion. We believe that IBM Cognos Consumer Insight will help us to differentiate and categorize all the data we collect according to pre-established criteria, such as author, date, country and subject. This enables us to focus only on comments and news items that are actually relevant, whether in a positive, negative or neutral sense."

The content of the comments is subsequently analyzed using custom Spanish and English dictionaries, in order to identify whether the sentiments expressed are positive or negative. "What is great about this solution is that it helps us to focus our actions on the most important topics of online discussions and immediately plan the correct and most suitable reaction," adds the Department, "By building on what we accomplished in the initial COBRA project, the new solution enables BBVA to seamlessly monitor comments and postings, improve its decision-making processes, and thereby strengthen its online reputation."

"When BBVA detects a negative comment, a reputational risk arises," explains Miguel Iza Moreno, Business Analytics and Optimization Consultant at IBM Global Business Services. "Cognos Consumer Insight provides a reporting system which identifies the origin of a negative statement and BBVA sets up an internal protocol to decide how to react. This can happen through press releases, direct

communication with users or, in some cases, no action is deemed to be required; the solution also highlights those cases in which the negative comment is considered 'irrelevant' or 'harmless'. The same procedure applies to positive comments—the solution allows the bank to follow a standard and structured process, which, based on positive insights, enables it to strengthen its reputation.

"Following the successful deployment in Spain, BBVA will be able to easily replicate the Cognos Consumer Insight solution in other countries, providing a single solution that will help to consolidate and reaffirm the bank's reputation management strategy," says the Department.

Tangible Results

Starting with the COBRA pilot project, the solution delivered visible benefits during the first half of 2011. Positive feedback about the company increased by more than one percent while negative feedback was reduced by 1.5 percent—suggesting that hundreds of customers and stakeholders across Spain are already enjoying a more satisfying experience from BBVA. Moreover, global monitoring improved, providing greater reliability when comparing results between branches and countries. Similar benefits are expected from the Cognos Consumer Insight project, and the initial results are expected shortly.

"BBVA is already seeing a remarkable improvement in the way that information is gathered and analyzed, which we are sure will translate into the same kind of tangible benefits we saw from the COBRA pilot project," states the bank, "For the time being, we have already achieved what we needed the most: a single tool which unifies the online measuring of our business strategies, enabling more detailed, structured and controlled online data analysis."

QUESTIONS FOR THE END-OF-CHAPTER APPLICATION CASE

1. How did BBVA use text mining?
2. What were BBVA's challenges? How did BBVA overcome them with text mining and social media analysis?
3. In what other areas, in your opinion, can BBVA use text mining?

Source: IBM Customer Success Story, "BBVA seamlessly monitors and improves its online reputation" at http://www-01.ibm.com/software/success/cssdb.nsf/CS/STRD-8NUD29?OpenDocument&Site=corp&cty=en_us (accessed August 2013).

References

- Chun, H. W., Y. Tsuruoka, J. D. Kim, R. Shiba, N. Nagata, and T. Hishiki. (2006). "Extraction of Gene-Disease Relations from Medline Using Domain Dictionaries and Machine Learning." *Proceedings of the 11th Pacific Symposium on Biocomputing*, pp. 4–15.
- Cohen, K. B., and L. Hunter. (2008). "Getting Started in Text Mining." *PLoS Computational Biology*, Vol. 4, No. 1, pp. 1–10.
- Coussement, K., and D. Van Den Poel. (2008). "Improving Customer Complaint Management by Automatic Email