Detection of Top Customer Intentions and Complaints
How PayPal uses RapidMiner to detect churn and identify customer service issues

The Customer

Founded in 1998 and acquired by eBay in 2002, PayPal is the faster, safer way to pay and get paid online, providing simpler ways to send and receive money around the world. With 143 million active accounts in 193 markets and 26 currencies around the world, PayPal enables global commerce, processing almost 8 million payments every day.

The Challenge

For Han-Sheong Lai, Director of Operational Excellence and Customer Advocacy at PayPal and Jiri Medlen, Senior Text Analytics Specialist at PayPal DT, driving customer satisfaction and reducing customer churn are never ending, challenging tasks. Han and Jiri knew that figuring out what drives product experience improvement without adequate knowledge of customer perspective and feedback is like “shooting in the dark,” hoping that opinion-based actions translate into tangible business improvements. By applying basic voice-of-the-customer-concepts and text analytics to customer feedback in over 60 countries worldwide, Han, Jiri and their team were able to identify, classify and count customers as “top promoters” and “top detractors,” according to their feedback verbatim. They identified the top complaint areas for internal PayPal product managers, specific to “product topics,” so they have knowledge of potentially actionable, specific issues to address proactively. One text mining technology Jiri and Han relied on is RapidMiner, predictive analytics and text mining software that analyzes historical and real-time customer data, and helps predict what issues might prompt customers to abandon the service and which specific customers expressed their intention to abandon using PayPal.

“In 2011, Eric Salvatierra, a PayPal VP, asked Jiri and me to explore using RapidMiner to systematically analyze PayPal’s vast quantity of customer feedback verbatim to provide product...
managers with a go-to place to read customer complaint verbatim, specific to their product areas. He also challenged us to cover multiple languages from our global customer base—English, of course, but also German, Portuguese, Spanish, Russian, Slavic from Eastern Europe, and Asian languages including Chinese, Japanese, and Korean. We soon discovered only RapidMiner could process text in all these languages, especially the Asian languages. It could parse prominent words and help us isolate, month to month, the top 50 detractors and promoters online, and the top complaints,” Han said.

While “frozen” customer accounts might seem like the most vocalized PayPal issue, RapidMiner helped quantify the top issue—customer password login problems. “Every month, we use RapidMiner for sentiment analysis of customer comments and feedback from around the world. It takes three days to analyze the text from over 150,000 customer post-contact comments in almost every language, including 50,000 tweets and some public Facebook posts. What we find is that out of millions of active PayPal users, a sizeable number of customers have password problems when trying to use PayPal every month. Using RapidMiner, we extracted the list of the most frequent words from the customer complaint verbatim dataset. We then found two top password or PayPal access ‘defects’ and identified the actionable, root causes, after discussing results with the relevant engineers. We also found, over time, that this spikes every November and December, when Holiday shopping begins, and customers would like to use PayPal or access their accounts, but can’t remember their password,” said Jiri.

“The Solution

“We then deployed the engineering fixes with the help of a cross-functional team and again used text analytics to monitor the improvements made, evidenced by any significant drop in customer complaint volume. Two to three weeks after the fixes, we saw that the password recovery success rate went up by almost 50 percent and password related complaint verbatim volume declined significantly,” Han added.

Jiri and his text analytics engineering team use RapidMiner with the Perl programming language, which provides initial text processing of customer feedback data, analyzes and parses subsets of keywords, and puts feedback into directories of complaints and compliments. In addition, they use a natural language processing (NLP) based software to process and auto- categorize verbatim with further grammatical and sentiment accuracy. To complement the NLP engine, they also use Apache Lucene, a high-performance text search engine library written in C++, for nearly any application that requires full-text search, across platforms.

“One of the great things about RapidMiner is its ability to process customer feedback in multiple languages. We tried using the R programming language, but it was far more difficult to deploy and it was not able to support some Asian languages. In fact, it was a pain. We also tried a big vendor's text analytics software module that was priced at about half a million dollars, which was not what we wanted to spend. While there is some learning curve, RapidMiner was a fraction of that cost and delivered real results in production mode within two to three months,” said Jiri.

“One of the great things about RapidMiner is its ability to process customer feedback in multiple different languages. We tried using the R programming language, but it was far more difficult to deploy and it was not able to support some Asian languages. In fact, it was a pain.”

-Jiri Medlen
Senior Text Analytics Specialist, PayPal DT
Han also piloted using RapidMiner to automatically determine customers’ intend-to-churn (ITC) sentiment, mining millions of feedback texts from customers surveyed after interacting with PayPal’s customer support. One challenge is ambiguous verbatim text, which might falsely present a customer as ITC, and not actually require follow-up contact. Han addressed this by using RapidMiner cross-validation tests, which increased the accuracy of the ITC categorization. Once the top ITC customers are more accurately flagged, PayPal can “surgically” follow-up with efforts to retain them as customers.

Han and Jiri have learned that technology is only one pillar of success. The other two are executive support, and disciplined innovation. “Business executives, who hold the power to allocate text analytics resources, are beginning to see and realize the potential benefits of text analytics to help better focus and solve business problems. Faced with thousands and millions of systematic, unstructured customer feedback texts, a company or team should quickly and effectively identify and monitor the biggest customer sentiments or complaints, as well as their fluctuations. The challenge is only a few truly understand what it takes to do and scale text analytics right, and deploy it as a competitive capability to make their products or services better.”

“Business executives, who hold the power to allocate text analytics resources, are beginning to see and realize the potential benefits of text analytics to help better focus and solve business problems.”

- Han-Sheong Lai
  Director of Operational Excellence and
  Customer Advocacy, PayPal