

# VALE QUANTO REVELA

Serasa Experian é a primeira instituição a utilizar no Brasil o SAS Text Miner para processamento de dados não estruturados, além de outras ferramentas proprietárias do grupo Experian, de maneira a continuar inovando e provendo os modelos matemáticos mais precisos do mercado **POR PAULA ZAIDAN**

**O** BRASIL NÃO SAIU ILESO DA CRISE mundial dos mercados de crédito no mundo, deflagrada pela quebra do banco Lehman Brothers no ano passado. No entanto, aos poucos, esse segmento se restabelece e os setores mais afetados são monitorados pelo governo para receber tratamento diferenciado, de acordo com suas necessidades específicas. A gestão de crédito e risco no Brasil tem sido uma das maiores preocupações de corporações no País, abrangendo desde a indústria, o comércio e o consumidor.

Enquanto as instituições financeiras discutem perfis de credores e devedores para avaliar riscos, bilhões de dados tratados de diversas formas saltam nas telas de computadores de analistas e consultores da

Serasa Experian, respondendo *on-line* e em tempo real a quatro milhões de consultas por dia, demandadas por 400 mil clientes diretos e indiretos.

Diante de volume tão vultoso de dados, a Serasa Experian avalia constantemente as melhores práticas de TI para atingir o estado da arte em seu principal negócio – a informação.

A Serasa-Experian é reconhecida globalmente por promover ideias e metodologias para análises de dados, criando valor para seus clientes por meio de soluções. A busca pelo melhor resultado de uma análise de crédito, no entanto, não se dá apenas pela mineração de dados estruturados.

## A EXPLOÇÃO DOS DADOS NÃO ESTRUTURADOS

Com a profusão de informações na Internet, dados são cada vez mais guardados e acessados de fontes externas como ferramentas de busca, redes sociais, blogs, serviços noticiosos e de assinaturas, wiki, fóruns, entre outras modalidades. De acordo com o Gartner, 80% dos negócios baseiam-se em informações não estruturadas. O mesmo instituto de pesquisa estima que o volume desses dados não estruturados dobre a cada três meses.

“Para a Serasa Experian, a inteligência dos dados é muito mais crítica porque a análise de informações está no cerne do nosso serviço”, explica Fabio Wendling Muniz de Andrade, gerente da área de Analytics & Data Intelligence. Andrade reconhece, porém, que o mercado ainda levará um tempo para amadurecer quanto ao uso de soluções que automatizem a inteligência analítica de negócios por meio da mineração de dados não estruturados, “mas queremos estar



Andrade, da Serasa Experian: sabedoria de quem enxerga oportunidades em dados não estruturados

prontos e continuar sendo um centro de excelência global, agora também para estes tipos de dados”.

#### A MINA DE OURO DAS PALAVRAS

Usuária da solução SAS Enterprise Miner, de mineração de dados estruturados, e de outras soluções do próprio grupo Experian, a Serasa Experian percebeu uma oportunidade no volume crescente de informações em forma de texto. “Com o aumento do volume de dados não estruturados, notícias de jornais, documentos da própria organização e principalmente da Internet, temos a necessidade de organizá-los e estruturá-los para que possam ser utilizados mais efetivamente para gerar soluções diferenciadas”, afirma Andrade ao explicar o

#### NÚMEROS

### O volume crescente dos Dados

**85%** de todos os dados armazenados estão no formato não estruturado (BUTLER GROUP)

**80%** dos negócios baseiam-se em informações não estruturadas (GARTNER GROUP)

Dados não estruturados **dobram** a cada **3 meses** (GARTNER GROUP)

**7 milhões** de páginas web são adicionadas a cada dia (GARTNER GROUP)

A dificuldade de encontrar informação custa **US\$ 6 milhões/ano** para uma empresa de mil funcionários. A falta de não ter o conhecimento nas oportunidades sai caro (IDC)

motivo da aquisição da solução SAS para mineração de textos.

Adquirido em meados de março, o SAS Text Miner – que otimiza o processo de extração de dados não estruturados de forma automatizada – atenderá o mercado latino-americano da Serasa Experian em três áreas iniciais: avaliação de crédito, fraudes e marketing. “Ainda neste ano, esperamos incorporar os ganhos que virão da utilização da informação não estruturada dentro dos nossos produtos”, prevê Andrade.

Um dos maiores desafios para os CEOs, CFOs e CIOs, atualmente, é a descentralização das informações, espalhadas em soluções tecnológicas de retenção de clientes, material de apoio

à decisão, ambiente heterogêneo de TI e uma infinidade de documentos fiscais e regulatórios. Segundo a IDC, a dificuldade de encontrar uma informação pode representar para uma empresa de mil funcionários o custo de US\$ 6 milhões de dólares por ano. E não ter o conhecimento das oportunidades vale ainda mais.

A coleta e análise automática dessas informações fornecem respostas acuradas, além de eliminar a dificuldade do processo manual de mineração de dados em texto. “O SAS Enterprise Miner prevê exatamente isso. A solução extrai informações de qualquer lugar, seja em um banco de dados, planilha ou até texto com o Text Miner, que podem ser ana-

## Aplicação na retenção de clientes

Apesar do mercado ainda não ter a cultura de adoção de soluções de negócios analíticas baseadas em mineração de texto, como o SAS Text Miner, a Serasa sai na frente e é pioneira na adoção da tecnologia entre usuários corporativos no Brasil.

Nos meios acadêmicos, a solução já integra pesquisas. Marcelo Schiessl, professor da Universidade de Brasília e especialista da Caixa Econômica Federal (CEF), utilizou-a em sua tese de mestrado. “Ele desenvolveu o projeto baseado nas reclamações que chegavam à central de atendimento. Por meio do Text Miner, identificou a possibilidade de não só automatizar a garimpagem das informa-

ções, como reduzir o tempo de resposta”, conta André Novo.

O objetivo era extrair fatos relacionais de dados não estruturados e combiná-los com dados estruturados de inteligência de negócio, uma vez que 85% de todos os dados armazenados estão no formato textual, segundo estudo do Butler Group. A solução para o Serviço de Atendimento ao Usuário da Caixa Econômica Federal foi combinar o data mining SAS Enterprise Miner com o SAS Text Miner.

O desenvolvimento da solução ocorreu em quatro etapas: estender a utilização do SAS Enterprise Miner para dados textuais; extrair informação de textos nos docu-

mentos do SAC com o SAS Text Miner; estruturar em formato adequado à mineração; e integrar esta informação à estrutura relacional para utilizá-la com soluções de inteligência de negócios.

“Assim, o banco poderia aproveitar os recursos de informação de forma mais eficiente e ainda identificar pontos fracos segundo a visão de seus clientes. Essas novas informações também poderiam ser utilizadas para gerar negócios e aproveitar as oportunidades descobertas no texto”, conclui Marcelo Schiessl. Além do SAC da CEF, a informação textual disponível na Internet poderia ser estruturada e utilizada pela área de Inteligência Competitiva.

lisadas de acordo com o interesse do negócio”, detalha André Novo, gerente de contas do SAS para o Brasil.

Fabio de Andrade, da Serasa Experian, explica que o maior desafio de sua área é desenvolver constantemente os melhores modelos estatísticos que suportem tanto os produtos de classificação de risco de crédito de empresas e consumidores, combinado com os serviços de consultoria para acelerar o ROI de nossos clientes. “Somos um birô de informação e dentro das nossas competências-chave está o tratamento da informação. A inovação nesta área é um dos nossos fatores críticos de sucesso”, observa Andrade.

### EXTRATOS PARA EXCELÊNCIA

No mundo corporativo, após o desenvolvimento acentuado das bases de dados estruturados e das modelagens, descobriu-se que nem todos os cruzamentos identificam informações ou conhecimento valiosos para a inteligência competitiva. Os gerentes de informações precisam de um conhecimento de mercado possível de se obter somente com novas modelagens, coleta ou desdobramento de dados não estruturados.

No processo tradicional, a análise de informações de uma organização implica longo levantamento de dados, seja com base em documentos armazenados na Internet ou em dados internos da companhia, gerando um volume muito grande de informação não estruturada. Os analistas continuamente buscam compreender o que ocorre ao redor da organização, para desenvolver uma interpretação compartilhada que poderá servir como contexto para a tomada de decisão.

Essa tentativa esbarra na necessidade de se retirar a equivalência dos fatos, determinar o significado das mudanças e isolar as porções

## A Serasa Experian

Presente em todas as capitais e principais cidades do País, totalizando **140 pontos** estratégicos, a Serasa Experian conta com aproximadamente **2.600 profissionais** e a retaguarda de um amplo centro de telemática. Como maior banco de dados da América Latina sobre consumidores, empresas e grupos econômicos, a companhia participa da maioria das decisões de crédito e de negócios tomadas no Brasil, respondendo *on-line/real-time*, a **4 milhões** de consultas por dia, demandadas por **400 mil clientes** diretos e indiretos.

de eventos para exame mais detalhado. Esse filtro gera matéria-prima para ser transformada em significado e depois em ação. Assim, as organizações que desenvolvem conhecimento operam com modelos de interpretação.

“O Text Miner, com o apoio de ferramentas da própria Experian, automatizará a análise de informações baseadas em texto permitindo que elas sejam utilizadas em modelos estatísticos. E a solução contribuirá para que aumentemos nossa liderança no mercado, bem como a excelência na prestação de serviços e produtos mais qualificados”, afirma Andrade.

A Serasa espera também atingir resultados operacionais como a otimização de trabalho de analistas de crédito, responsáveis pela avaliação do grande volume de informações sobre empresas publicadas em jornais e na Internet. “A solução identifica os termos chave no texto que possuem valor informacional e vão possibilitar maior eficiência na análise das informações”. ■

#### ONLINE

Saiba mais sobre a Serasa Experian:  
[www.sas.com/br/sascom/online/serasa.pdf](http://www.sas.com/br/sascom/online/serasa.pdf)  
Valorize o seu BI com Text Analytics:  
[www.sas.com/events/cm/149049](http://www.sas.com/events/cm/149049)

