



Big data: revolução no mercado de seguros

HENRIQUE HOELTGEBAUM & CÉSAR NEVES

A captura, armazenagem e processamento de uma grande quantidade de dados, seguidas por uma análise criteriosa dos mesmos para subsidiar a tomada de decisão com objetivos estratégicos, podem ser classificados como big data. O conhecimento e a utilização dessa ferramenta são essenciais para desenvolvimento e crescimento do mercado de seguros.

A primeira citação do termo “big data” apareceu em um artigo científico na NASA¹, em que havia um

problema de visualização gráfica em função do tamanho dos dados e capacidade computacional. De lá para cá, se popularizou e hoje é empregado em todos os fóruns de análise crítica de dados em grande dimensão. Contextualmente, essa metodologia se fomentou sobre o aumento exponencial da disponibilidade de informações que antes era escassa e restrita. Tal aumento se deve principalmente ao avanço da internet. Todas as atividades nela realizadas são armazenadas. Cada clique, cada busca, cada troca de e-mail ou postagem em alguma rede social deixa um rastro que pode se transformar em um conjunto de informações valioso, o qual se traduz em uma análise de comportamento do usuário que não se limita apenas à internet. Os celulares com geolocalização,

informações disponibilizadas em uma aplicação de emprego ou até mesmo registros médicos digitalizados são outros exemplos da amplitude de dados que deixamos disponíveis, afirmam os professores do departamento de economia da Universidade de Stanford Liran Einav e Jonathan D. Levin, em seu artigo científico “*The data revolution and economic analysis*”².

A profissão “cientista de dados”, figura como uma das principais profissões do século 21. Porém, qual a melhor formação acadêmica para esse profissional? Quais são as habilidades necessárias, de

¹ COX, M.; ELLSWORTH, D. *Application-controlled demand paging for out-of-core visualization*. Proceedings of the 8th conference on Visualization'97. Los Alamitos: IEEE Computer Society Press, 1997. p. 235-ff

² EINAV, Liran; LEVIN, Jonathan D. *The data revolution and economic analysis*. National Bureau of Economic Research, 2013.



O atuário é um profissional que já possui as qualificações analíticas necessárias

para calcular probabilidades, riscos e impactos financeiros em eventos futuros.

maneira a preparar um profissional qualificado para atuar nesse campo? Conforme artigo publicado na revista RISKSA³, uma pesquisa feita pela consultoria McKinsey mostrou que apenas atuários, estatísticos e analistas de pesquisa operacional possuem 100% do conhecimento analítico e talento para processamento e análise desse tipo de dado. Porém, a capacidade não se deve limitar a apenas essas formações: é importante uma equipe multidisciplinar para enriquecer ao máximo uma possível análise. Por exemplo, um grupo de médicos do Albert Einstein College of Medicine⁴ buscou a *expertise* de um pintor e artista para estruturar novas formas de visualização dessa onda de informações. Isso porque, ao se traba-

lhar com essa enorme quantidade de dados, a primeira dificuldade se materializa logo quando se executa uma tarefa tão simples quanto a visualização das variáveis envolvidas na análise em um gráfico.

E quanto ao atuário, onde essa formação fica inserida nesse novo contexto? Inicialmente, é importante contextualizar que big data é a taxonomia dada a uma grande quantidade de dados que serve como *input* para modelos preditivos. Conforme a entrevista à revista RISKSA, Peter Temple, presidente eleito da Sociedade de Atuária da África do Sul, relatou que o atuário é um profissional que já possui as qualificações analíticas necessárias para calcular probabilidades, riscos e impactos financeiros em eventos futuros. Por consequência, sempre teve domínio sobre esses modelos. Contudo, agora, tais modelos são utilizados em arcabouços que transcendem aplicações financeiras, como esportes e medicina, por exemplo. No entanto, há uma necessidade de qualificação e que tais profissionais estejam preparados para algumas mudanças de paradigmas.

No contexto de big data que focaremos neste artigo, a variável-resposta desses modelos preditivos são escores, que traduzem matematicamente qual o padrão de comportamento do usuário em análise. Esses escores são gerados via aplicação de aprendizado de máquina (“*machine learning*”) a uma grande massa de dados, os quais traduzem o comportamento de um público potencial de usuários. Na prática, essas aplicações se baseiam em converter uma grande quantidade de dados não estruturados em escores preditivos, frequentemente de uma maneira automatizada e, algumas vezes, em tempo real. Para Einav e Levin, as maiores empresas de tecnologia do mundo, como Apple, Amazon, Netflix e Google, lançam mão de alguns exemplos de utilização desses escores. A Apple usa a palavra ou sentença com escore mais alto e propõe as mesmas para o usuário, ao automaticamente completar uma frase. A Amazon recomenda itens que o modelo prevê que despertariam interesse por parte do usuário, com base em seu comportamento prévio. A Netflix sugere quais

3 Disponível em: <<http://www.risksa.com/big-data-revolution-means-jobs-for-actuaries/>>.

4 Disponível em: <http://www.nytimes.com/2015/03/29/sunday-review/learning-to-see-data.html?hp&action=click&pgt-type=Homepage&module=c-column-top-span-region®ion=c-column-top-span-region&WT.nav=c-column-top-span-region&_r=1>.

filmes o usuário poderia estar interessado em assistir. Já o Google indica produtos com base nos resultados das buscas e conteúdo dos e-mails contidos no Gmail.

Além disso, companhias de cartão de crédito usam modelos preditivos para implementar uma política de combate à fraude. Esse modelo deve ser capaz de avaliar, de uma maneira automatizada, se a transação em questão deve ser aprovada ou rejeitada, e ainda, qual deve ser investigada, baseada nos escores.

Similarmente, no mercado de seguros, é comum quantificar o risco de uma variável a que um indivíduo esteja exposto. Através desse valor, tem-se uma ideia de qual seria, por exemplo, a probabilidade de esse sujeito vir a sofrer um acidente. Já no âmbito de planos de saúde, o escore é tipicamente a soma ponderada de indicadores da saúde do indivíduo, que podem identificar se o mesmo possui alguma condição crônica diferenciada, sendo os pesos escolhidos por meio de uma análise estatística. Trabalhos empíricos recentes nos Estados Unidos^{5,6} estudaram a escolha do tipo de seguro ou plano de saúde, utilizando os escores preditivos para resumir a heterogeneidade individual de uma maneira parcimoniosa. Os escores fornecem

5 EINAV, Liran et al. Selection on moral hazard in health insurance. *National Bureau of Economic Research*, 2011.

6 BUNDORF, M. Kate; LEVIN, Jonathan D.; MAHONEY, Neale. Pricing and welfare in health plan choice. *National Bureau of Economic Research*, 2008.



um quantitativo útil para, digamos, mensurar se indivíduos que se expõem mais a risco sistematicamente aderem a planos de seguro com maior cobertura e, ainda, se os preços cobrados pelas seguradoras no mercado se ajustam para o custo provável desse tipo de indivíduo.

Porém, as aplicações no mercado de seguro vão muito além. Elucidamos a seguir algumas das principais discutidas no trabalho institucional da empresa StackIQ⁷. Por exemplo, na prevenção de riscos, atualmente, a relação da seguradora com seus associados é descentralizada e virtual.

7 Capitalizing on Big Data Analytics for the Insurance Industry. Disponível em: <http://cdn.hubspot.net/hub/173001/file-18488782-pdf/docs/stackiq_insuranceind_wpp_fpdf>.

Entretanto, uma análise criteriosa na infinidade de dados disponíveis sobre o comportamento do indivíduo serve como um bom *proxy*. É possível o cruzamento dessas informações para, a partir do comportamento do usuário, quantificar melhor os riscos. Assim, um segurador pode avaliar o risco associado a um seguro residencial analisando dados de satélite da referida propriedade, padrões de clima e estatísticas de emprego regionais.

Outra aplicação que pode ser feita a partir da análise de comportamento do usuário é a customização individualizada de produtos para ele. A habilidade de oferecer ao cliente as apólices que necessita ao prêmio mais competitivo é uma excelente vantagem para as seguradoras. Com base nos escores dos modelos

preditivos aplicados ao comportamento do usuário, a seguradora irá estruturar os produtos e prêmios para os clientes, customizados, tendo como ponto de partida suas necessidades e fatores de risco. Por exemplo, algumas seguradoras americanas iniciaram a coleta de dados a partir de sensores nos carros dos clientes que gravam a média de quilômetros percorridos, velocidade média, horário do dia que o usuário costuma dirigir e o quão brusco aperta o freio. Esses dados são comparados com outros dados agregados para determinar a melhor taxa para cada motorista, com foco em seus hábitos, histórico e grau de risco.

Essa ampla coleta de dados, incluindo aqueles de cliques na web, atividades de redes sociais e informações contábeis, é responsável por gerar informações estratégicas adicionais para a entidade.

Com esse tipo de dado em mãos, é possível sugerir produtos adicionais ao consumidor que vão ao encontro das necessidades e orçamento do mesmo, realizando *cross selling* e *upselling*. Esse tipo de aplicação pode também estudar os hábitos do cliente para avaliar o risco e sugerir alteração de hábitos para reduzi-lo.

Com um maior número de dados armazenados e processados, até mesmo os riscos tomados pelas seguradoras são calculados de forma mais acurada. Isso possibilita que o erro em torno das estimativas dos parâmetros que compõem o cálculo do capital baseado em risco seja reduzido. Consequentemente, as análises de risco ficam mais precisas, podendo-se minimizar, dessa forma, o capital em risco que deve ser mantido disponível, assumindo-se a utilização de um modelo interno para seu cálculo.

Dessa forma, fica claro que a análise preditiva de uma grande quantidade de dados veio para ficar no mercado de seguros, possibilitando, através da pesquisa de todas as informações disponíveis, o estudo do perfil do consumidor, quer seja para lançamento de novos produtos customizados, quer seja para mensuração dos riscos das apólices vigentes. Ambas as ações acarretarão crescimento do mercado segurador. Ademais, com a captura e armazenamento de todos os dados relevantes, que devem ser feitos em tempo real, as seguradoras poderão mensurar de forma mais ajustada sua solvência e corrigir eventuais distorções, a fim de evitar perda financeira dos acionistas. Com essa “revolução” que está por vir, o desafio é a qualificação adequada dos profissionais e futuros profissionais, para que os resultados apareçam de forma mais efetiva. ●

HENRIQUE HOELTGEBAUM

Estatístico e mestre em Engenharia de Produção da Pontifícia Universidade Católica do Rio de Janeiro (PUC-Rio).

hnhelfer@hotmail.com

CÉSAR NEVES

Analista técnico da Superintendência de Seguros Privados (SUSEP), membro do Centro de Pesquisas Aplicadas em Risco (CEPAR), professor do Departamento de Estatística e Atuária da Universidade do Estado do Rio de Janeiro (UERJ), professor da Escola Superior Nacional de Seguros (ESNS) e doutorando em Engenharia Elétrica na Pontifícia Universidade Católica do Rio de Janeiro (PUC-Rio).

cesar.neves@susep.gov.br
cesar@ime.uerj.br

