

**Data Science
Academy**

www.datascienceacademy.com.br

Business Analytics

**Big Data e Detecção de Fraudes
Parte 2**

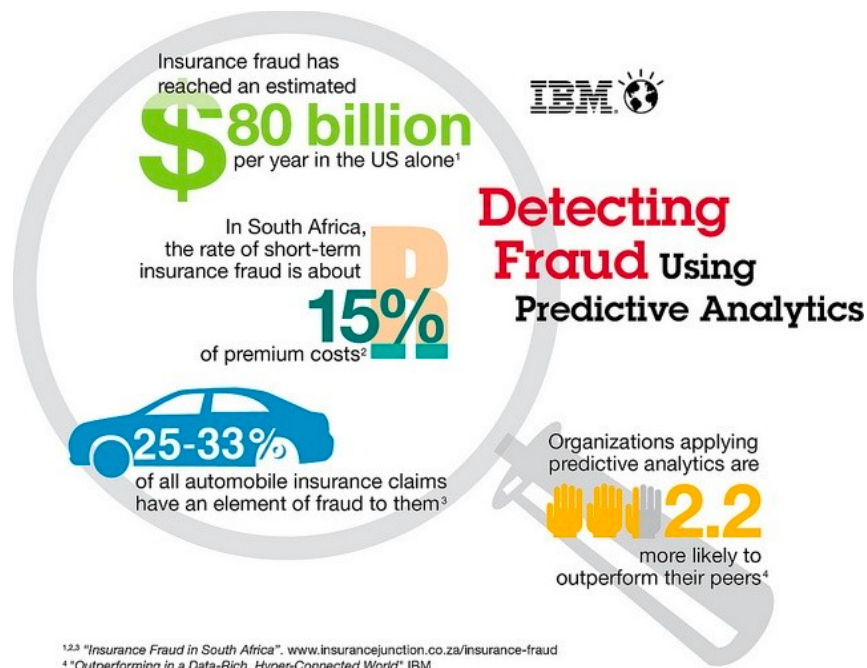


Uma outra razão pela qual a fraude se torna mais fácil de detectar quanto mais tempo passa é porque as melhores técnicas de detecção estão sendo desenvolvidas, estão ficando prontamente disponíveis e estão sendo implementadas e aplicadas por uma quantidade crescente de organizações. Um importante motor para melhorias no que diz respeito às técnicas de detecção é a crescente disponibilidade de dados. A informatização e digitalização de quase todos os aspectos da sociedade e da vida cotidiana leva a uma abundância de dados disponíveis. Estes dados, algo que você já deve conhecer bem a esta altura da Formação Cientista de Dados, são o Big Data, que podem ser explorados para uma série de propósitos, incluindo detecção de fraude, a um custo muito baixo.

A utilização do Big Data para análise de fraudes no Brasil ainda é mínima, o que claro leva a um oceano de oportunidades, para aqueles dispostos a desbravar esse mercado. É necessário que as empresas saibam lidar com o enorme volume de dados gerados, pois é certo que eles irão aumentar cada vez mais. Elas precisam olhar para o futuro e imaginar todos os dispositivos que estarão conectados por seus clientes. Calcula-se que até 2020, 30 bilhões de dispositivos estejam conectados, contra 10 bilhões em 2013. Quanto mais dados gerados, maior a possibilidade de se perder nessas informações e não identificar o que realmente interessa.

O primeiro passo para utilizar o Big Data na detecção e prevenção de fraudes, é identificar o perfil do cliente, utilizar as ferramentas necessárias e estar preparado para receber e analisar essas informações. A combinação ideal de várias fontes de dados e a melhoria da qualidade dos processos de análise dos mesmos são outros desafios para as companhias. É nesse contexto que a participação de um Cientista de Dados torna-se ainda mais essencial. Ele terá o conhecimento necessário para analisar os resultados gerados e trabalhar na prevenção à fraude com análises mais precisas.

Segundo as estimativas, pelo menos 10% dos pagamentos da companhia de seguros são para reivindicações fraudulentas, e a soma global destes pagamentos fraudulentos equivale a bilhões ou possivelmente trilhões de dólares. Embora a fraude de seguros não seja um problema novo, a gravidade do problema está aumentando e os perpetradores de fraudes de seguros estão se tornando cada vez mais sofisticados. Por exemplo, uma reivindicação de ferimento pessoal poderia potencialmente incluir reivindicações médicas falsificadas ou um acidente encenado. O Big Data Analytics, pode rapidamente procurar padrões em reivindicações históricas e identificar semelhanças ou levantar questões em uma nova reivindicação antes que o processo seja concluído.



Considere o seguinte exemplo. Uma companhia de seguros quer melhorar sua capacidade de tomar decisões em tempo real ao decidir como processar uma nova reivindicação. A despesa de custo da companhia incluindo pagamentos do litígio relacionados às reivindicações fraudulentas tem aumentado firmemente. A empresa tem políticas extensas para ajudar os seguradores a avaliarem a legitimidade dos sinistros, mas os analistas muitas vezes não tinham os dados no momento certo para tomar uma decisão informada.

A empresa implementou uma grande plataforma de Big Data Analytics para fornecer a integração e análise de dados de várias fontes. A plataforma incorpora o uso extensivo de dados de mídia social e streaming de dados para ajudar a fornecer uma visão em tempo real. Agentes de call center são capazes de ter uma visão muito mais profunda em possíveis padrões de comportamento e relações entre outros requerentes e prestadores de serviços quando uma chamada ocorre.

Um agente pode receber um alerta sobre uma nova reivindicação que indica que o requerente foi testemunha anterior sobre uma reclamação semelhante há seis meses. Depois de descobrir outros padrões incomuns de comportamento e apresentar essas informações ao reclamante, o processo de reivindicação pode ser interrompido antes que ele realmente comece a andar. Em outras situações, dados de mídia social podem indicar que as condições descritas em uma reivindicação não ocorreram no dia em questão. Por exemplo, um requerente indicou que seu carro foi atingido em uma enchente, mas a documentação das mídias sociais mostrou que o carro tinha sido realmente em outra cidade no dia em que ocorreu a inundação.



Fraude de seguros causam enormes prejuízos para as empresas e os executivos estão se movendo com o objetivo de incorporar Big Data Analytics e outras tecnologias avançadas para resolver o problema da fraude. As companhias de seguros não só sentem o impacto desses prejuízos, como também repassam esses custos em forma de taxas para os clientes honestos, que pagam mais pelo serviços. Soluções de Big Data Analytics para detecção de fraudes podem ajudar a evitar perdas de todos os lados!

Vamos trabalhar com Big Data no Mini-Projeto 4, na sequência deste capítulo.