

## **Should I Stay or Should I Go? Predição de Rotatividade de Segurados no Mercado de Seguro por Distância Percorrida**

**WESLEY OLIVEIRA DE SOUZA**

FACULDADE DE ECONOMIA, ADMINISTRAÇÃO E CONTABILIDADE DA UNIVERSIDADE DE SÃO PAULO - FEA

**JOÃO VINÍCIUS DE FRANÇA CARVALHO**

FACULDADE DE ECONOMIA, ADMINISTRAÇÃO E CONTABILIDADE DA UNIVERSIDADE DE SÃO PAULO - FEA

## **Should I Stay or Should I Go? Predição de Rotatividade de Segurados no Mercado de Seguro por Distância Percorrida**

### **Introdução**

O uso da tecnologia telemática voltada para área de seguros é capaz de minimizar os efeitos adversos da falta de informação. Hoje, há duas principais categorias de produtos de uso ou comportamento do usuário: (i) pay-how-you-drive (PHYD), e; (ii) pay-as-you-drive (PAYD). O PHYD concentra-se principalmente no uso do comportamento de direção para avaliar os prêmios de risco. Já o seguro PAYD utiliza-se principalmente da distância percorrida para avaliar os prêmios de risco. Para as seguradoras, identificar os fatores explicativos do comportamento de evasão (churn) do usuário é relevante.

### **Problema de Pesquisa e Objetivo**

Este trabalho aborda o problema de churn rate no contexto do mercado de seguro por distância percorrida. O principal objetivo é avaliar a previsão de rotatividade de clientes no mercado de seguro por distância percorrida, utilizando diferentes técnicas de machine learning. Assim, é possível traçar estratégias de retenção para o período anterior ao cancelamento da apólice, como por exemplo, direcionar campanhas de marketing personalizadas, realizar ajustes no prêmio a posteriori, conhecidos como bonus-malus, além de reduzir custos operacionais relacionados a aquisição de novos clientes.

### **Fundamentação Teórica**

Churn é o evento em que um cliente decide conscientemente encerrar seu relacionamento com um prestador de serviço ou vice vice-versa. Há três principais tipos de churn: (i) voluntário, quando um cliente decide encerrar a relação de serviço com o fornecedor; (ii) involuntário, quando o provedor de serviço encerra o relacionamento com cliente, e; (iii) inevitável, quando o relacionamento é encerrado por razões exógenas, como quando um cliente está se mudando para um novo local ou quando uma região enfrenta um desastre natural. As taxas de perda de clientes são indicadores de negócios essenciais.

### **Metodologia**

Um dos pilares dos avanços das técnicas de machine learning é o método boosting, que melhora a performance preditiva dos modelos, convertendo vários aprendizados fracos em um único modelo de aprendizado forte. Para o presente trabalho, com o intuito de identificar qual o modelo preditivo de maior desempenho para à análise de churn rate, foram selecionados para comparação com o modelo XGBoost os modelos: redes neurais, random forest e regressão logística, considerados os modelos com maior nível de acurácia presentes na literatura sobre o assunto.

### **Análise dos Resultados**

A Regressão Logística desempenhou melhor na curva ROC e na especificidade, mas o modelo XGBoost superou as outras três técnicas pelo critério de acurácia. A variável comportamental de distância percorrida foi relevante em todos os modelos estimados: o seguro PAYD deve ser escolhido somente se o segurado tiver hábito de percorrer distâncias médias inferiores ao limite estabelecido pela seguradora para a concessão de desconto. Outras variáveis relevantes foram: (i) diferença de prêmios pagos entre o PAYD e um seguro ilimitado, (ii) histórico de sinistro do segurado, e; (iii) a idade do segurado.

### **Conclusão**

Ainda que os resultados não impliquem que o XGBoost seja o melhor modelo que possa ser aplicado a todos os casos, fazer sucessivos testes é fundamental para avaliar a adequação e capacidade preditiva. Uma possível extensão deste trabalho seria fazer avaliação similar para diferentes métodos (e.g., support vector machines, redes neurais recorrentes, inferência causal, bayesiana ou não) e em outros ramos de seguros para os quais variáveis comportamentais sejam determinantes. E a realização de experimentos pode ajudar a desenhar incentivos ótimos para evitar o cancelamento voluntário de apólices.

**Referências Bibliográficas**

Cappiello, A. (2020). The Digital (R)evolution of Insurance Business Models. *American Journal of Economics and Business Administration*, 12(1), 1-13. Cheng, J., Feng, F. Y., & Zeng, X. (2022). Pay-As-You-Drive Insurance: Modeling and Implications. *North American Actuarial Journal*, 1-19. Gao, G., Meng, S., & Wüthrich, M. V. (2022). What can we learn from telematics car driving data: a survey. *Insurance: Mathematics and Economics*, 104, 185-199.