



Renovação de Apólices de Seguros: Um Estudo Comparativo usando Machine Learning

Rodriggo Santos de Souza Rabelo

Universidade Federal do Vale do São Francisco

(rodriggo.rabelo@discente.univasf.edu.br)

Thiago Magalhães Amaral

Universidade Federal do Vale do São Francisco

(thiago.magalhaes@univasf.edu.br)

CONTEXTUALIZAÇÃO: O mercado brasileiro de seguros apresenta crescimento consistente, acompanhado de uma concorrência cada vez mais acentuada. Nesse contexto, a fidelização de clientes assume papel central, pois a renovação de apólices garante previsibilidade de receita e contribui para a sustentabilidade financeira das corretoras. Manter clientes ativos, em muitos casos, torna-se tão estratégico quanto captar novos segurados. Com os avanços em ciência de dados, técnicas de aprendizado de máquina passaram a ser aplicadas para analisar grandes volumes de informação e apoiar decisões de negócio (Silva et al., 2024). Tais métodos permitem identificar padrões de comportamento e antecipar a probabilidade de renovação ou cancelamento (Hassonah et al., 2019). Entre os algoritmos disponíveis, destacam-se o *K-Nearest Neighbors* (KNN), que classifica registros com base em semelhança de perfis através de distâncias, e a *Árvore de Decisão*, que organiza os fatores em regras de interpretação clara. O uso dessas técnicas pode fornecer suporte valioso às corretoras na definição de estratégias de retenção de clientes.

OBJETIVO: Este estudo tem como objetivo aplicar e comparar o desempenho dos algoritmos KNN e *Árvore de Decisão* em uma base de clientes de uma corretora de seguros, com foco na previsão da renovação de apólices. Além de avaliar as métricas de



Anais do Simpósio Acadêmico de Engenharia de Produção (SAEPRO) da EEL-USP

IX SAEPRO – 25 e 26 de novembro de 2025

desempenho, busca-se discutir a utilidade prática dos resultados e seu potencial como ferramenta de apoio à gestão em corretoras.

MÉTODO: A pesquisa utilizou uma base de dados composta por 69 clientes de uma corretora de seguros, contendo informações de idade, tempo de relacionamento com a empresa, valor do prêmio anual, tipo de seguro contratado e indicação de aumento de prêmio em relação ao ano anterior. A variável alvo foi a renovação da apólice, codificada como 1 para clientes que renovaram e 0 para aqueles que não renovaram. O conjunto de dados foi preparado e analisado no software Orange Data Mining, no qual foram aplicadas etapas de limpeza dos dados, além de colocar ID para os clientes, visando a privacidade, normalização e codificação das variáveis categóricas. Em seguida, os dados foram divididos em 80% para treinamento e 20% para teste, adotando-se também validação cruzada em cinco partes (5-fold) para aumentar a robustez da análise. Foram empregados os algoritmos *K-Nearest Neighbors* (KNN) e Árvore de Decisão. O KNN classificou os clientes de acordo com a similaridade entre seus perfis, enquanto a Árvore de Decisão gerou regras que explicam o impacto de cada variável na decisão de renovação. A performance dos modelos foi comparada com base em métricas como Acurácia, Precisão, Recall, F1-score, AUC e MCC, identificando padrões de comportamento entre os clientes e a análise de quais fatores mais influenciam a decisão de renovar ou não a apólice, gerando insights aplicáveis à gestão estratégica de corretoras de seguros.

RESULTADOS E DISCUSSÕES: Os experimentos indicaram desempenho superior do KNN, que alcançou acurácia de 85%, F1-score de 0,91, precisão de 86,8% e recall de 95,5%. Tais valores evidenciam sua capacidade de identificar de forma consistente os clientes que optaram pela renovação. Já a Árvore de Decisão, apesar de apresentar acurácia inferior (68,6%) e F1-score de 0,79, demonstrou relevância por oferecer regras interpretáveis sobre os fatores que influenciam a tomada de decisão dos clientes de acordo com as variáveis estipuladas. A análise dos padrões permitiu observar que: Clientes com maior tempo de relacionamento apresentaram forte propensão à renovação; Clientes recentes que sofreram reajuste no prêmio tiveram maior tendência a não renovar; O valor do prêmio mostrou-se um fator secundário, influenciando apenas em casos específicos.



Anais do Simpósio Acadêmico de Engenharia de Produção (SAEPRO) da EEL-USP

IX SAEPRO – 25 e 26 de novembro de 2025

Essas observações estão alinhadas à realidade do setor, no qual a fidelidade do cliente é um diferencial estratégico, e novos segurados tendem a ser mais sensíveis a alterações de preço.

CONSIDERAÇÕES FINAIS: O estudo demonstrou que o KNN se destacou pela maior precisão preditiva, enquanto a Árvore de Decisão ofereceu maior clareza interpretativa. Essa complementaridade sugere que a escolha do algoritmo deve considerar não apenas a performance numérica, mas também a necessidade de explicação e transparência nos resultados. Do ponto de vista prático, o uso de algoritmos de classificação pode auxiliar a corretora a identificar clientes com maior risco de cancelamento e adotar medidas preventivas de fidelização, como campanhas segmentadas, benefícios personalizados ou ajustes no atendimento. Para estudos futuros, recomenda-se ampliar a base de dados e incluir variáveis de perfil socioeconômico, de modo a aumentar o poder explicativo dos modelos. Também seria pertinente testar algoritmos adicionais, como *Random Forest* e *Support Vector Machine* (SVM), além de modelos lineares, a exemplo da regressão logística, a fim de verificar se os dados apresentam comportamento suficientemente linear para serem bem representados por técnicas mais simples. A incorporação dessas análises pode ampliar a robustez e a aplicabilidade dos resultados para o setor de seguros.

PALAVRAS-CHAVE: Aprendizado de Máquina; KNN; Árvore de Decisão; Seguros.

REFERÊNCIAS:

AGGER. **Mercado de seguros 2025: entenda o crescimento previsto.** 2025. Disponível em: <https://www.agger.com.br/blog/mercado-de-seguros-no-brasil/>. Acesso em: 28 set. 2025.

CONFEDERAÇÃO NACIONAL DAS SEGURADORAS (CNseg). **Setor segurador cresce 6% no 1º tri de 2025, apesar do cenário econômico.** 2025. Disponível em:



Anais do Simpósio Acadêmico de Engenharia de Produção (SAEPRO) da EEL-USP

IX SAEPRO – 25 e 26 de novembro de 2025

<https://cnseg.org.br/noticias/setor-segurador-cresce-6-no-1-tri-de-2025-apesar-do-cenario-economico>. Acesso em: 28 set. 2025.

DELOITTE BRASIL. **Perspectivas para o setor de seguros em 2025**. 2025. Disponível em: <https://www.deloitte.com/br/pt/Industries/insurance/perspectives/perspectivas-setor-seguros.html>. Acesso em: 28 set. 2025.

FARIAS, L. R. Fidelização de consumidores em companhias seguradoras: uma análise com Equações Estruturais. **Revista GESeC**, v. 14, n. 1, p. 1-19, 2023. Disponível em: <https://ojs.revistagesec.org.br/secretariado/article/view/1966>. Acesso em: 28 set. 2025.

HASSONAH, M. A. et al. A comparative study for customer churn prediction in Customer Relationship Management using Decision Tree and K-Nearest Neighbor algorithms. **International Journal of Advanced Computer Science and Applications**, v. 10, n. 2, p. 55-61, 2019. Disponível em: https://www.researchgate.net/publication/340890936_Churn_Prediction_A_Comparative_Study_Using_KNN_and_Decision_Trees. Acesso em: 21 out. 2025.

SILVA, Maria Clara Monteiro et al. BIG DATA E INTELIGÊNCIA ARTIFICIAL NA GESTÃO EMPRESARIAL: DESAFIOS, OPORTUNIDADES E TENDÊNCIAS. **Revista Esfera Acadêmica Tecnologia**, [S.l.], v. 9, n. 3, 2024. Disponível em: https://multivix.edu.br/wp-content/uploads/2025/08/REVISTA-REVISTA_ESFERA_TECNOLOGIA_V09_N03_ARTIGO02.pdf Acesso em: 21 out. 2025.

SOUZA, B. M.; et al. Análise de fidelização de clientes no segmento de seguros. **Revista de Administração Mackenzie**, v. 16, n. 3, p. 205-228, 2015. Disponível em: <https://editorarevistas.mackenzie.br/index.php/rmec/article/view/5918/6070>. Acesso em: 28 set. 2025.