

Análise de inadimplentes em uma cooperativa do ramo agrícola utilizando mineração de dados

Jaisson Duarte¹, Edimar Manica^{1*}

*Orientador

¹Instituto Federal de Educação, Ciência e Tecnologia do Rio Grande do Sul (IFRS) - *Campus Ibirubá*.
Ibirubá, RS, Brasil.

O Brasil possui cerca de 62 milhões de inadimplentes, segundo levantamento feito pela empresa Serasa no ano de 2021. São pessoas que não conseguiram honrar seus compromissos financeiros, comprometendo seu próprio crédito e gerando prejuízos às empresas que, por sua vez, muitas vezes não conseguem repassar os valores financeiros aos seus fornecedores. As cooperativas do ramo agropecuário estão inseridas nesse cenário, pois geram movimentações financeiras através da comercialização de produtos e serviços aos seus associados e clientes e sofrem pelas mesmas dificuldades das demais empresas. A cooperativa alvo deste trabalho possui mais de 100 mil clientes e um crescimento em torno de 4% ao ano. Atualmente, o setor financeiro avalia em média apenas 5 mil clientes ao ano, a fim de permitir ou negar a concessão de crédito, o que gera uma sobrecarga de trabalho e possíveis falhas humanas. Avaliar os clientes é fundamental para amenizar a inadimplência futura. Clientes não avaliados possuem maiores chances de se tornarem inadimplentes, em comparação com os que foram submetidos a uma análise de crédito efetiva. Este resumo descreve um Trabalho de Conclusão de Curso de Ciência da Computação que visa criar um modelo de classificação para identificar os clientes com maiores chances de inadimplência. Pretende-se, por meio da análise de dados históricos e utilizando-se de algoritmos de classificação disponíveis em uma plataforma de mineração de dados, obter o melhor algoritmo seguindo determinadas métricas, desenvolver o modelo mais eficaz e disponibilizar uma ferramenta que seja capaz de identificar se um novo cliente deve ser analisado ou não. Desta forma, o propósito final é auxiliar os analistas de crédito em priorizar os clientes que serão avaliados na concessão de crédito. Até o presente momento, este trabalho criou uma base de dados histórica compreendendo o período de 2015 a 2021, com dados anonimizados, totalizando cerca de novecentos mil registros. Além disso, foram executados mais de cinquenta testes, e até o momento o algoritmo mais eficaz foi o *DecisionStump*. Entretanto ainda se faz necessário mais experimentos e análises para definir o melhor algoritmo para a tarefa alvo deste trabalho e assim obter o modelo mais apropriado.

Palavras-chave: Inadimplência; Mineração de Dados; Classificação.