

RELATÓRIO TÉCNICO – PREVISÃO DE INADIMPLÊNCIA

O problema central que a empresa enfrenta é prever com precisão quais clientes provavelmente se tornarão inadimplentes no próximo mês. A empresa busca minimizar as perdas financeiras relacionadas a essas inadimplências, oferecendo aconselhamento preventivo aos clientes.

O projeto visa criar um modelo que ajude a empresa a tomar decisões proativas, prevenindo o aumento da inadimplência e maximizando o retorno financeiro.

O problema é a identificação precoce de clientes que estão em risco de inadimplência. A empresa busca um modelo que permita encontrar um equilíbrio entre prever quem vai inadimplir sem gerar falsos positivos excessivos, o que levaria a ações desnecessárias.

O conjunto de dados foi fornecido pela empresa e incluía informações pessoais dos clientes, histórico de transações e dados sobre inadimplências passadas.

O dataset continha variáveis como idade, limite de crédito, histórico de pagamentos e valores de fatura.

Foram tratadas inconsistências nos dados, como valores ausentes e outliers. Realizamos imputação de dados faltantes e normalização de variáveis numéricas.

Também foram removidos dados duplicados e criadas novas features para ajudar a melhorar o desempenho dos modelos.

A análise exploratória mostrou correlações importantes, como a maior propensão à inadimplência entre clientes mais velhos e aqueles com limites de crédito mais baixos.

Clientes com atrasos recorrentes nos pagamentos também apresentaram maior probabilidade de inadimplência.

No entanto, identificamos inconsistências significativas nas colunas referentes (PAY_X), que indicam o comportamento de pagamento dos clientes. Essas inconsistências nos dados de pagamento limitaram a capacidade de confirmar com exatidão a hipótese 3, que propunha que o comportamento de pagamento dos últimos meses teria uma correlação direta com a inadimplência futura.

Modelos de machine learning, como Regressão Logística, Decision Tree e Random Forest, foram testados. A Random Forest mostrou o melhor desempenho.

A avaliação foi feita com métricas como precisão, recall, ROC AUC e acurácia.

O ajuste de hiper parâmetros, como número de árvores na Random Forest e profundidade máxima, melhorou a performance do modelo.

O threshold foi ajustado para maximizar a economia líquida da empresa, encontrando o melhor equilíbrio entre falsos positivos e falsos negativos.

O modelo final obteve um **ROC AUC de 0.77**, precisão de 51% e recall de 58%.

O modelo foi capaz de reduzir a inadimplência em 24,27%, gerando uma economia líquida de NT\$ 16.901.935,86.

Como entrega final, foi gerado um arquivo CSV contendo uma lista dos clientes com tendência a inadimplir no próximo mês, para que o time de consultoria possa trabalhar com eles e diminuir a inadimplência.

Como próximos passos, planeja-se testar o modelo XGBoost para comparação, além de ajustar hiper parâmetros periodicamente e monitorar a performance do modelo em produção.

RELATÓRIO DE NEGÓCIOS – PREVISÃO DE INADIMPLÊNCIA

Imagine poder prever com antecedência quais clientes do seu portfólio podem não conseguir pagar suas dívidas no próximo mês. Isso permitiria ações imediatas, como campanhas de aconselhamento ou negociação, mitigando perdas significativas.

Com o modelo preditivo de inadimplência, podemos antecipar cenários de risco, economizando milhões em potenciais perdas. Esta abordagem proativa não apenas melhora a saúde financeira da empresa, mas também fortalece o relacionamento com os clientes.

Nosso objetivo é proteger o caixa da empresa, identificando os clientes com maior probabilidade de inadimplência, enquanto evitamos intervenções em clientes que continuarão pagando normalmente. Assim, podemos otimizar os recursos direcionados para essas ações preventivas, economizando tempo e dinheiro.

Utilizamos um vasto conjunto de dados de clientes, analisando histórico de crédito e comportamento de pagamento. A análise dos dados forneceu uma base sólida para decisões mais informadas e precisas.

A análise inicial mostrou que alguns grupos, como clientes mais velhos ou com limites de crédito menores, apresentavam maior risco de inadimplência. Essa informação nos ajuda a concentrar esforços em perfis de maior risco.

Garantimos que todos os dados usados fossem confiáveis e consistentes. A limpeza dos dados foi crucial para garantir que as previsões fossem precisas, evitando distorções no modelo.

Clientes acima de 50 anos e com histórico de atrasos apresentam maior risco de inadimplência. Além disso, aqueles com limites de crédito mais altos são mais propensos a honrar seus compromissos. Esses insights ajudam a segmentar melhor as ações preventivas.

Após testar diferentes abordagens, optamos pelo modelo Random Forest, que apresentou o melhor equilíbrio entre precisão e capacidade de previsão. Ele nos permitirá identificar com mais eficácia os clientes em risco.

Utilizamos os algoritmos mais modernos e eficazes para garantir que as previsões sejam precisas e possam ser aplicadas rapidamente.

Ajustamos o modelo para garantir que ele traga o máximo de retorno financeiro, minimizando as intervenções desnecessárias em clientes que não inadimplirão.

A implementação do modelo resultará em uma economia estimada de NT\$ 16.901.935,86, reduzindo a inadimplência em 24,27%. Isso reflete uma grande oportunidade de melhorar a saúde financeira da empresa.

O modelo agora está em produção, oferecendo previsões semanais sobre os clientes mais propensos à inadimplência. As equipes de crédito poderão usar essas previsões para ações de mitigação.

Graças ao modelo preditivo, estamos agora em uma posição de vantagem. Podemos prever comportamentos críticos e tomar decisões mais inteligentes, garantindo que a empresa minimize suas perdas e maximize seu retorno financeiro.

O sucesso inicial nos deu uma base sólida. Agora, estamos focados em aprimorar o modelo, testar novos modelos e expandir suas capacidades, garantindo que ele continue a gerar valor de longo prazo.