

Tarefa 5 de AEDI

(22/12/2024)

Nome: _____ Matrícula: _____

Questão

Você trabalha como analista de dados em uma grande instituição bancária, responsável por desenvolver modelos preditivos que ajudem o banco a melhorar suas estratégias de retenção de clientes. Recentemente, o banco observou um aumento no número de clientes cancelando suas contas e deseja entender melhor os fatores que levam à saída dos clientes.

Sua equipe recebeu um conjunto de dados contendo informações sobre os clientes atuais e antigos, incluindo características como idade, renda, saldo na conta, tipo de produto bancário utilizado, entre outros. O objetivo é criar um modelo que seja capaz de prever se um cliente permanecerá no banco ou decidirá sair. Com essas informações, o banco poderá tomar decisões mais assertivas, como oferecer incentivos personalizados para reter clientes em risco.

Para ajudar nesse projeto, você deve construir e avaliar um modelo preditivo usando **Regressão Logística**, respondendo às seguintes demandas:

- a) Realize a modelagem para prever se um cliente irá deixar o banco ou não utilizando Regressão Logística. Escolha algumas características dos clientes (por exemplo, idade, renda e tipo de conta) e desenvolva o modelo. Explique os coeficientes estimados e interprete os resultados em termos da influência de cada característica na probabilidade de retenção do cliente.
- b) Valide os pressupostos da Regressão Logística (como a linearidade entre as variáveis independentes e o logit, ausência de multicolinearidade e balanceamento da variável dependente). Caso necessário, aplique técnicas como balanceamento de classes ou regularização. Justifique as escolhas feitas e interprete os resultados obtidos.
- c) Identifique e interprete as variáveis mais importantes para prever se um cliente permanecerá no banco. Utilize métodos como análise dos coeficientes do modelo ou a importância das variáveis. Explique como essas variáveis contribuem para a decisão e para o planejamento estratégico de retenção.

- d) Avalie o desempenho do modelo ajustado utilizando métricas apropriadas, como AUC, Curva ROC, precisão, sensibilidade e especificidade. Explique como essas métricas refletem a qualidade do modelo na previsão de retenção de clientes.

Conjunto de Dados

Utilize o conjunto de dados disponível em [Retenção de Clientes de Unidades Bancárias](#).