

# Case - Modelo preditivo para aprovação de crédito

Modelos de score de crédito calculam a probabilidade de inadimplência e são uma das principais ferramentas utilizadas por diversas empresas para aprovar ou negar um crédito.

O objetivo deste desafio é criar um modelo preditivo calculando a probabilidade de inadimplência de cada novo pedido de crédito.

## Dados:

Cada linha representa um cliente e as colunas representam os dados (informações) desses clientes.

A variável resposta é a coluna inadimplente, que indica se o tomador veio a se tornar inadimplente(1) ou não(0).

As variáveis da base de dados são descritas abaixo:

- idade: A idade do cliente..
- numero\_de\_dependentes: O número de pessoas dependentes do cliente.
- salario\_mensal: Salário mensal do cliente.
- numero\_emprestimos\_imobiliarios: Quantidade de empréstimos imobiliários que o cliente possui em aberto.
- numero\_vezes\_passou\_90\_dias: Número de vezes que o tomador passou mais de 90 dias em atraso.
- util\_linhas\_inseguras: Quanto que o cliente está usando, relativamente ao limite dele, de linhas de crédito que não são seguradas por qualquer bem do tomador e.g: imóveis, carros etc.
- vezes\_passou\_de\_30\_59\_dias: Número de vezes que o cliente atrasou, entre 30 e 59 dias, o pagamento de um empréstimo.
- razao\_debito: Razão entre as dívidas e o patrimônio do tomador. razão débito = Dívidas/Patrimônio
- numero\_linhas\_credito\_aberto: Número de empréstimos em aberto pelo cliente.
- numero\_de\_vezes\_que\_passou\_60\_89\_dias: Número de vezes que o cliente atrasou, entre 60 e 89 dias, o pagamento de um empréstimo.

**Obs:** Estes dados foram retirados de terceiros, portanto é possível que existam incoerências, o que é perfeitamente comum em dados reais.

## Objetivo:

**Construir qualquer modelo preditivo utilizando o arquivo treino.csv.**

Utilize este modelo para gerar as previsões na base teste.csv, inserindo uma nova coluna na tabela de dados do arquivo teste.csv que contenha as previsões e nomeie esta coluna

com o nome "inadimplente". Espera-se que não seja necessário mais que um dia para a resolução deste problema.

### **Avaliação:**

Serão avaliados o código, preferencialmente em python ou R, o desempenho do modelo e tudo que for necessário para entender a sua linha de raciocínio para a solução final.

Para isto, só será necessário enviar o código com **tudo** o que foi feito e a base de dados de teste **com a coluna de previsões**.