

**PONTIFÍCIA UNIVERSIDADE CATÓLICA**  
**CAMPUS MONTE ALEGRE**  
**CURSO DE CIÊNCIA DE DADOS E I.A.**

GABRIEL MACHADO  
JOÃO PEDRO GARCIA  
MATHEUS LOPES  
MIGUEL BATTENDIERI  
PEDRO VENDRAMETTO

**ANÁLISE DE RISCO:**  
**PREVISÃO DE INADIMPLÊNCIA EM CARTÕES DE CRÉDITO USANDO ABORDAGEM**  
**COMBINADA**

**sÃO PAULO**  
**2024**

## RESUMO

Prever contas com potencial de inadimplência de crédito antecipadamente é desafiador. As técnicas estatísticas tradicionais geralmente não conseguem lidar com grandes volumes de dados e a natureza dinâmica da fraude e do comportamento humano. Para resolver esse problema, pesquisas recentes se concentraram em abordagens baseadas em inteligência artificial e computacional. Neste trabalho, apresentamos e validamos uma abordagem heurística para minerar potenciais contas inadimplentes antecipadamente, onde a probabilidade de risco é pré-calculada a partir de todos os dados anteriores, e a probabilidade de risco para transações recentes é calculada assim que ocorrem. Além de nossa abordagem, também aplicamos um tratamento de aprendizado de máquina – Logistic Regression. Como resultado, descobrimos que essas abordagens aplicadas superam as abordagens mais avançadas existentes.

**Palavras-chave:** inadimplência; heurística; risco; aprendizado de máquina.

## LISTA DE ILUSTRAÇÕES

Quadro 1 – Planejamento e Metodologia	12
Gráfico 1 – Idade e Adinplência	12
Gráfico 2 – Casamento e Adinplência	11
Gráfico 3 – Gênero e Adinplência	12
Gráfico 4 – Correlação Colunas e Default	12
Gráfico 5 – Meses de maior pagamento	11

## SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO</b>	<b>5</b>
1.1	EXEMPLO DE TÍTULO DE SEÇÃO SECUNDÁRIA	5
<b>1.1.1</b>	<b>ANÁLISE EXPLORATÓRIA</b>	<b>6</b>
1.1.1.1	Exemplo de título de seção quartenária	6
<b>2</b>	<b>TREINO E VALIDAÇÃO DO MODELO</b>	<b>7</b>
2.1	RESULTADOS DA MATRIZ	8
<b>3</b>	<b>TABELA E ILUSTRAÇÕES</b>	<b>9</b>
<b>4</b>	<b>REFERÊNCIAS</b>	<b>11</b>
	<b>ANEXO A – Matriz de Confusão (TREINO)</b>	<b>12</b>
	<b>ANEXO B – Matriz de Confusão (TESTE)</b>	<b>13</b>

## 1 INTRODUÇÃO

De maneira geral, podemos nos referir à incapacidade de um cliente de pagar, inadimplência em um pagamento ou falência pessoal como potenciais problemas de não pagamento. No entanto, cada um desses cenários resulta de diferentes circunstâncias, como perda de emprego, problemas de saúde ou incapacidade de trabalhar. Às vezes, é um ato deliberado, quando o cliente sabe que não tem mais condições de usar o cartão de crédito, mas continua usando até que o banco cancele o cartão, caracterizando fraude, o que é difícil de prever e um grande problema para os credores.

Para resolver esse problema, as empresas de cartão de crédito tentam prever possíveis inadimplências ou avaliar a probabilidade de risco em um pagamento antecipadamente. Quanto mais cedo as contas potenciais de inadimplência forem detectadas, menores serão as perdas para os credores. Por isso, uma abordagem eficaz para prever inadimplência potencial é crucial para permitir que os credores tomem ações preventivas e também possam ajudar o cliente a evitar falências, minimizando as perdas.

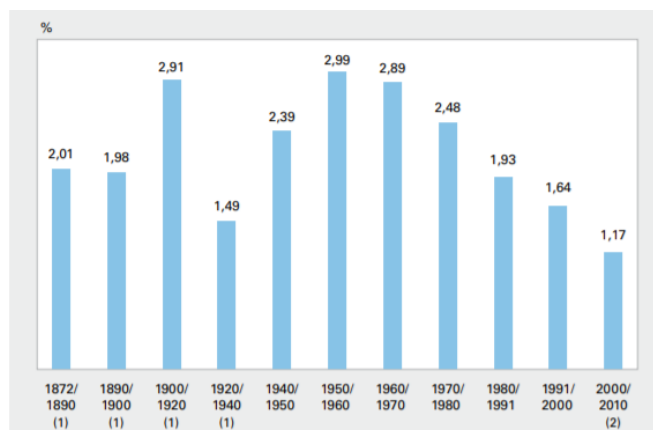
Analisar milhões de transações e fazer previsões com base nisso é demorado, consome muitos recursos e pode ser suscetível a erros devido a variáveis dinâmicas (como limite de saldo, renda, condições econômicas, etc.). Portanto, há uma necessidade de abordagens otimizadas que possam lidar com essas restrições.

## TABELA E ILUSTRAÇÕES

Quadro 1 – Planejamento e Metodologia

ETAPAS	AÇÕES
P	<b>Planejar e estudar</b> o Dataset por meio de um plano de ação após a identificação, reconhecimento das características e descoberta das causas principais do problema (inadimplência).
A	<b>Analisar e limpar</b> valores ausentes, caso existam, por valores apropriados.
D	<b>Dividir e converter</b> os dados em treinamento (80%) e teste (20%)
T	<b>Treinar e validar</b> um modelo de Regressão Logística e avaliar sua precisão, apresentando a matriz de confusão para o modelo e interpretações dos resultados obtidos.

Gráfico 1 – Idade e Adinplência

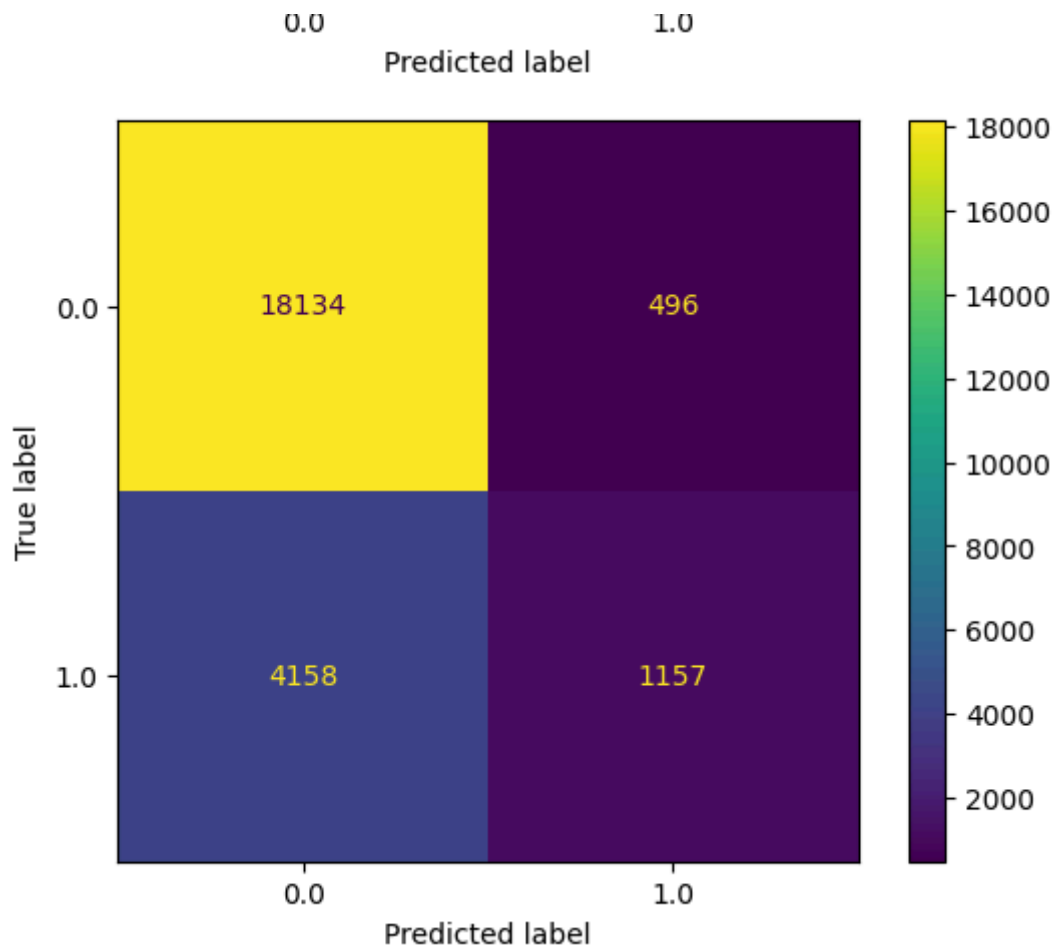


Fonte: IBGE (2011).

## REFERÊNCIAS

SCIKIT-LEARN. Documentação. Disponível em:  
<https://scikit-learn.org/0.21/documentation.html>. Acesso em: 20 set. 2024.

KAGGLE. Pesquisa: Default of Credit Card Clients. Disponível em:  
<https://www.kaggle.com/search?q=default+of+credit+card+clients+in%3Adatasets>.  
Acesso em: 24 set. 2024

**ANEXO A – Matriz de Confusão (TREINO)**



### ANEXO B – Matriz de Confusão (TESTE)

