



# Recomendação de Estabelecimentos Comerciais



Andressa Ribeiro  
e-mail: [andressaribeiro@usp.br](mailto:andressaribeiro@usp.br)



# SOBRE O PROBLEMA



Como a VR pode recomendar os estabelecimentos certos para usuários que aceitaram receber recomendações, otimizando o uso dos cartões e a satisfação?

## Dados usados

- Transações (histórico de uso)
- Trabalhadores
- Estabelecimentos
- Dados combinados e filtrados para usuários com aceite

## Engenharia de Features

- Características de usuário (demográficas, comportamento passado).
- Características de estabelecimento (categorias, tipo, etc.).
- Features temporais (ex: "dias desde última transação", "taxa média de transações passadas do usuário").

**Por quê?** Modelo híbrido ideal para recomendação, combinando padrões de interação (filtragem colaborativa) com informações de conteúdo (neste caso, estabelecimentos e trabalhadores).

## OTIMIZAÇÃO

- O modelo foi treinado utilizando `sample_weight` (peso de amostra) baseado no **lucro normalizado**. Isso garante que o algoritmo priorize aprender e recomendar interações que geram maior retorno financeiro para a VR.

## VALIDAÇÃO ROBUSTA

- Divisão temporal dos dados (`data_corte`) para treino/teste: Simula um cenário de aplicação real, avaliando a capacidade do modelo de prever o comportamento e o lucro em dados futuros.

# ASSERTIVIDADE



## TREINO: 80% DA BASE

205.938 transações

Antes de 2025-04-05

Total de trabalhadores no teste: 5467

Trabalhadores que o modelo NÃO viu no treino (cold-start): 77

% Cold-start: 1,41%

## TESTE: 20% DA BASE

51.485 transações

A partir de 2025-04-05

Total de ECs únicos no teste: 4631

ECs no teste que o modelo NÃO viu no treino (cold-start): 853

% Cold-start: 18.42%

Se um usuário tem, em média, 9,4 transações no conjunto de teste, a **chance de acertar aleatoriamente**:  $5 * (9,417 / 4.631) \approx 1,02\%$ .

Com o modelo, **a chance é 13,75%**! Além disso, o **AUC é de 95,73%**, ou seja, possui excelente capacidade de ranquear corretamente itens relevantes acima de irrelevantes.

Supondo modelo que atribua aleatoriamente 5 estabelecimentos como baseline e considerando APENAS acertos (trabalhadores que compraram de fato no estabelecimento predito):

**K = 5**  
(em 5.390 usuários avaliados)

**LUCRO (R\$)**

**ANTES**

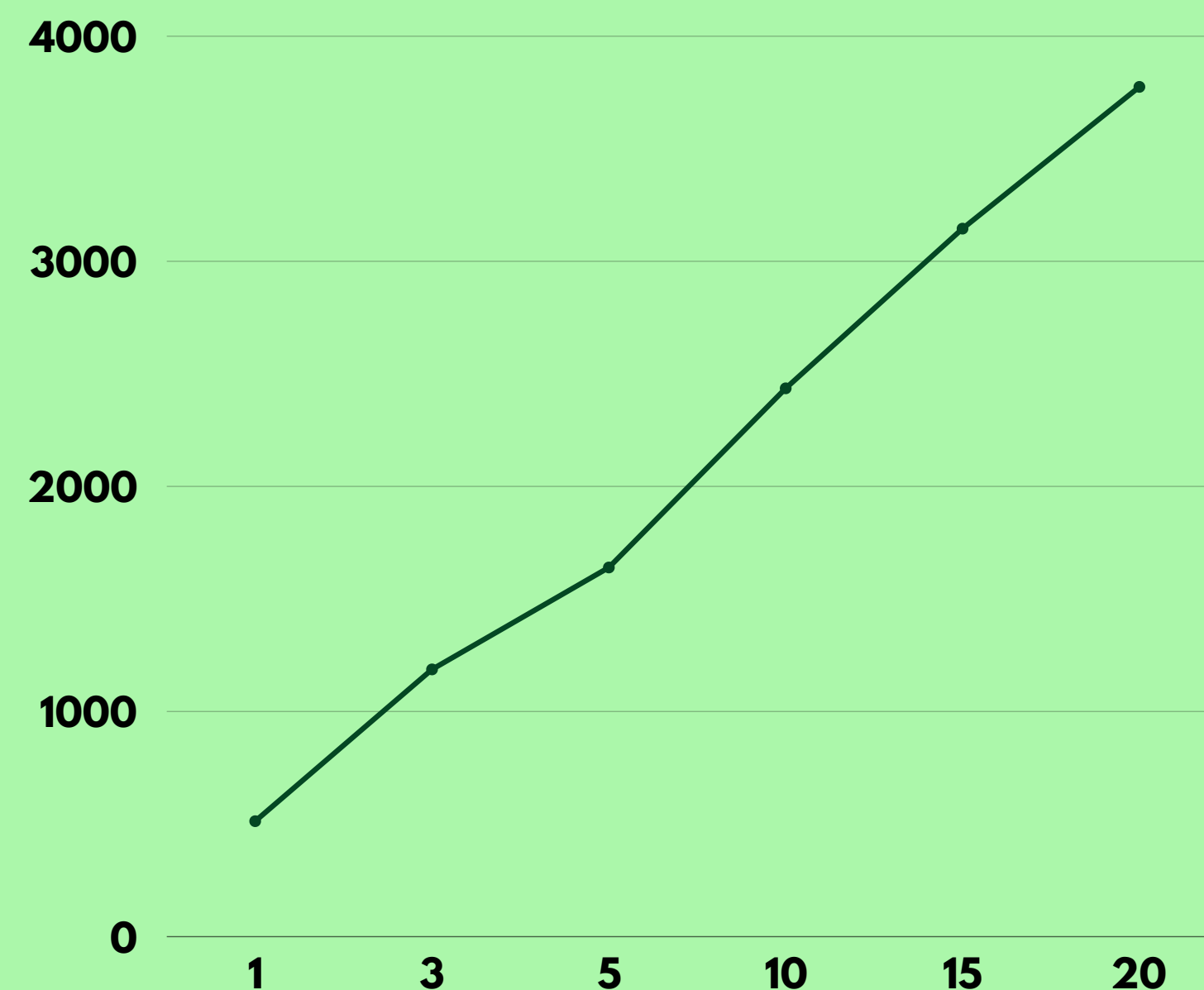
**2,86**

**DEPOIS**

**1.639,87**

**+56K%**

## Curva de Lucro Real do Modelo por K Recomendações



# THANK YOU!



**Andressa Ribeiro**  
e-mail: [andressaribeiro@usp.br](mailto:andressaribeiro@usp.br)