



# Maximização de vendas em Marketplace através de cash-back

Philippe D. Almeida- ICMC-USP <sup>1</sup>, Gleici Perdoná, FMRP-USP

## 1 Introdução

Plataformas de compras com 'cash-back', que na sua tradução significa 'dinheiro de volta', tem por objetivo impulsionar vendas de diversos segmentos com a prática de devolver ao cliente parte do dinheiro usado na compra do produto. Essas plataformas digitais (também conhecidas como aplicativos) tem sido uma ferramenta utilizada pela indústria como um canal alternativo para a oferta e impulsionamento de seus produtos [1].

Uma vez que o mantenedor da plataforma recebe a lista de produtos a serem disponibilizados para venda e a quantia a ser gasta com o 'dinheiro de volta', imediatamente surge a necessidade de otimizar o gasto desse dinheiro. No caso em questão, a melhor forma possível se torna a maximização das vendas com a menor quantia investida. Definir o preço ou o desconto certo para um produto é um problema antigo na teoria econômica, e determinar o 'cash-back' ideal se faz diretamente análogo. Os valores de cash-back, até então, têm sido determinados de forma manual com base em uma porcentagem do valor total do produto, e a performance avaliada pelo total de vendas. Uma vez determinado esse valor, ele é praticado igualmente para quaisquer usuário na plataforma. Portanto, é razoável pensar que ao invés de otimizar um valor de acordo com o que já foi feito e praticá-lo igualmente aos usuários, cada um desses usuários pode estar propenso à comprar um determinado produto de forma diferente. Por exemplo, o usuário 1 possui maior chance de comprar o produto 1 do que o usuário 2, logo, o valor de cash-back necessário para efetivação da compra para o usuário 2 deve ser maior do que para o usuário 1.

### 1.1 Proposta de solução

Saber quanto cada usuário estaria disposto a pagar por um determinado produto naquele momento dado o seu padrão de compras e o cenário atual do mercado.

Com a crescente melhoria da experiência do usuário em aplicativos de compras e sua popularização, esse projeto propõe uma solução para encontrar o melhor valor de 'cash-back' de um determinado

---

<sup>1</sup>philipedalmeida@usp.br

produto para cada usuário utilizando técnicas de Machine Learning. Os modelos de ML podem considerar um grande número de produtos e otimizar preços globalmente. O número e a natureza dos parâmetros e suas múltiplas fontes e canais permitem que eles tomem decisões usando vários critérios. Essa é uma tarefa impossível de se fazer manualmente, ou mesmo usando software básico.

Entender como os usuários da plataforma reagirão a diferentes estratégias de preços de produtos. Encontrar os melhores preços para um determinado cliente, considerando seus objetivos.

## **2 Materiais e metodologia**

### **2.1 Descrição das variáveis de estudo**

Existem 4 grupos principais de variáveis a serem estudadas e utilizadas. São esses os de cliente, produto, interação do cliente e produto dentro da plataforma, e concorrência.

- Usuário: gênero, idade, classe social, região de residência.
- Produto: categoria e valor.
- Concorrência: categoria e valor.
- Interação do cliente e produto dentro da plataforma: ativação da oferta, data de expiração da oferta, quantidade adquirida, valor de cash-back, ofertas visualizadas.

A partir dessas relações podemos tirar informações agregadas de cada usuário e item, ou globalmente, como a taxa de conversão de compra a partir da ativação, total de cash-back acumulado, tempo entre ativação e compra, compra na concorrência dentro da plataforma, total de visualizações por oferta, dentre outras.

### **2.2 Modelagem proposta**

Existem dois principais problemas que o nosso conjunto de dados possui:

- Nem todos os usuários compram todos os produtos.
- Um mesmo cliente não possui várias compras do mesmo produto com valores de cash-back diferentes.

Esses problemas trazem limitações no sentido de não se poder chegar a um valor ideal de cash-back para um item. E dado a natureza dos dados, sabemos que isso ocorre na maioria dos casos. Para contornar essa situação, se faz necessário criar uma representação de cada usuário e item de forma que, baseado no padrão de compra, tenhamos como medir similaridades. Por exemplo, usuários que compram os mesmos itens, em sua maioria, terão uma representação numérica em um espaço N-dimensional bem próximos. A mesma coisa vale para os itens que são comprados pelos mesmo usuários. Dessa forma será possível projetar o comportamento de um usuário relacionado a um item de acordo com outros usuários similares à ele e que já efetuaram a compra desse item.

Para tanto, existem técnicas já consolidadas para a tarefa descrita, e uma das mais famosas é a fatorização de matrizes não-negativas, que em [3] e [4] nos mostra como ela é utilizada em sistemas de recomendação. Da mesma forma que a Netflix projetava um valor de avaliação de um usuário para um filme/série/documentário, podemos projetar a demanda/gasto de um usuário em relação a um determinado produto, e assim construir as devidas representações.

### 2.2.1 Fatorização de matriz não-negativa com deep learning

A fatoração de matriz não-negativa (NMF), que visa fatorar uma matriz em duas matrizes não-negativas cujo produto reconstrói a matriz de dados original, tem mostrado aplicações extensas em muitos domínios, como processamento de sinal, aprendizado de máquina, mineração de texto e assim por diante. Ou seja, para reconstruir, aproximadamente, a matrix de usuário-item  $X$  onde em cada linha temos os  $m$  clientes, nas colunas os  $k$  itens e cada entrada sendo o número de vezes que o usuário comprou aquele item, decompomos  $X_{m \times k} = C_{m \times n} \cdot P_{n \times k}$ . Aqui,  $n$  é a dimensão arbitrária para a qual cada usuário e item serão representadas. Portanto, para cada combinação linear entre as linhas e colunas das matrizes  $C$  e  $P$  teremos uma aproximação da quantidade de itens que o usuário comprou representada em  $X$ . Daí sairá a projeção daquele valor onde não houve compra de um item por um usuário.

Apesar do seu grande sucesso, existe um problema intrínseco. O modelo básico de NMF sugere que os sinais podem ser eficientemente aproximados como uma combinação linear de átomos de dicionário. No entanto, a suposição de reconstrução linear, que é explicitamente imposta por muitos trabalhos existentes, não é válida para muitos sinais complicados [2]. Por isso a utilização de redes neurais com a técnica de 'gradient descent' será utilizada para reconstruir os valores originais e fazer aproximações com combinações não-lineares. Dessa forma, algumas suposições não precisarão ser respeitadas e obteremos uma representação mais robusta.

### 2.2.2 Regressão

Uma vez construída as representações de cada usuário-item, será utilizado e testado vários algoritmos na tarefa de regressão, como Xgboost ou um Muti-layer Perceptron para prever a quantidade de venda juntamente com as demais informações de usuário, produto, concorrência e principalmente o cash-back aplicado. Uma vez modelado, a principal tarefa será calcular o efeito do valor de cash-back, primeiramente com as outras variáveis fixas e se possível com as interações.

### 2.2.3 Proposta de avaliação e impacto do modelo

Uma vez que a demanda por um produto pode ser estimada em função do valor de cash-back, se torna possível avaliar o seu efeito. Para isso, o cálculo de elasticidade da demanda pode ser aplicado individualmente e globalmente.

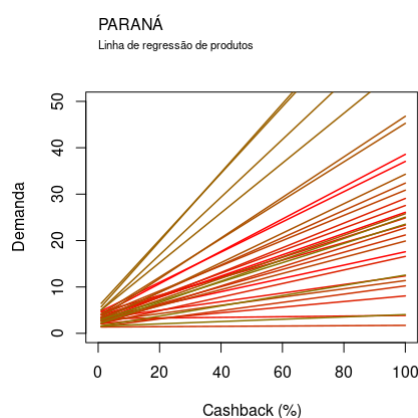
A demanda elástica ocorre quando o preço ou outros fatores têm um grande efeito sobre a quantidade que os consumidores desejam comprar. A fórmula da demanda elástica é a variação percentual na quantidade demandada dividida pela variação percentual no preço.

A partir do cálculo de elasticidade será possível determinar se o valor sugerido pelo modelo será de fato efetivo, através de técnicas de otimização inteira. Essas técnicas permitem otimizar a alocação de dinheiro disponível para o incentivo de forma a impactar itens que possuem maior

sensibilidade a preço, diferentemente de outros que são menos sensíveis, levando em consideração algumas restrições como estoque disponível.

### 3 Considerações-Resultados Parciais

Uma outra abordagem que pode ser aplicada inicialmente, e que se torna menos complexa a princípio, é a de modelar a elasticidade do preço de cada categoria de produto. Sendo assim, o preço ótimo será aplicado igualmente a todos os usuários. Dessa forma, pode-se controlar a margem global e a demanda predita de cada um desses produtos de forma mais direta. Caso contrário, podemos utilizar o mesmo procedimento de otimização mas com uma diferença no cálculo da demanda dado um preço, que nesse caso seria ponderada pela probabilidade de cada motorista carregar o determinado produto dado aquele preço.



A figura acima ilustra as retas de regressão para diferentes tipos de produtos. Quanto maior a inclinação, maior sensibilidade a preço.

Há a necessidade de se aplicar testes A/B para confirmar a efetividade do algoritmo, bem como desafios de controlar o efeito de outras promoções concorrentes.

### Referências

- [1] Ballestar, M., Grau-Carles, P. & Sainz, J. Customer segmentation in e-commerce: Applications to the cashback business model. *Journal Of Business Research*. **88** pp. 407-414 (2018)
- [2] Zhang, H., Liu, H., Song, R. & Sun, F. Nonlinear non-negative matrix factorization using deep learning. *2016 International Joint Conference On Neural Networks (IJCNN)*. pp. 477-482 (2016)
- [3] Ekstrand, M., Riedl, J., Konstan, J. & Others Collaborative filtering recommender systems. *Foundations And Trends® In Human-Computer Interaction*. **4**, 81-173 (2011)
- [4] Kim, H., Ji, A., Yeon, C. & Jo, G. A user-item predictive model for collaborative filtering recommendation. *International Conference On User Modeling*. pp. 324-328 (2007)