

NILTON KAZUYUKI UEDA

POSTECH

DATA ANALYTICS

FRAMEWORK DE BIG DATA

# AULA 05

## SUMÁRIO

O QUE VEM POR AÍ? .....	3
HANDS-ON .....	4
SAIBA MAIS .....	5
O QUE VOCÊ VIU NESTA AULA? .....	9
REFERÊNCIAS .....	10

EMANDA

## O QUE VEM POR AÍ?

Durante esta aula, você será introduzido(a) aos conceitos essenciais dos sistemas de recomendação e entenderá como eles são aplicados em diversas áreas, como comércio eletrônico, entretenimento e mídia. Você descobrirá como esses sistemas utilizam algoritmos e técnicas avançadas para analisar dados de usuários, identificar padrões e sugerir itens relevantes e personalizados.

Ao longo da disciplina, exploraremos diferentes tipos de sistemas de recomendação, como filtragem colaborativa, recomendações baseadas em conteúdo e sistemas híbridos. Além disso, aprenderemos sobre as métricas de avaliação utilizadas para medir a eficácia desses sistemas e as melhores práticas para o desenvolvimento e implementação de recomendações.

Prepare-se para desvendar a arte por trás dos sistemas de recomendação e compreender como eles moldam nossas experiências on-line. Com essa aula introdutória, você estará preparado(a) para explorar novas oportunidades e aplicar os princípios dos sistemas de recomendação em seus próprios projetos, personalizando interações e agregando valor aos usuários.

Não perca a oportunidade de mergulhar no mundo dos sistemas de recomendação e descobrir como eles influenciam nossas escolhas e preferências.

## HANDS-ON

Entenderemos mais sobre os principais conceitos de sistemas de recomendação e suas aplicações reais, portanto, vamos praticar sobre os conceitos desta aula. Vamos lá?!

EMEND

## SAIBA MAIS

### O QUE SÃO ALGORITMOS DE RECOMENDAÇÃO?

De uma forma geral, os sistemas de recomendação são algoritmos destinados a sugerir itens relevantes para os usuários (itens como filmes para assistir, textos para ler, produtos para comprar ou similares, a depender dos setores).

Os sistemas de recomendação são realmente críticos em alguns setores, pois podem gerar uma grande quantidade de receita quando são eficientes ou ser uma maneira de se destacar significativamente dos concorrentes. Como prova da importância dos sistemas de recomendação, podemos citar que, há alguns anos, a Netflix organizou um desafio (o “prêmio Netflix”) onde o objetivo era produzir um sistema de recomendação com desempenho superior ao seu próprio algoritmo com um prêmio de 1 milhão de dólares para vencer.

Pontos importantes ao construir nossos sistemas de recomendação:

- Se você tiver um grande banco de dados e fizer recomendações on-line a partir dele, a melhor maneira seria dividir esse problema em 2 subproblemas: **1)** escolher os principais **N** candidatos e **2)** classificá-los.
- Como você mede a qualidade do seu modelo? Em conjunto com as métricas de qualidade padrão, existem algumas métricas que são usadas especialmente para problemas de recomendação. São elas: Recall e Precision, Average Recall e Average Precision.
- Se estiver resolvendo problemas de recomendação com algoritmos de classificação, você deve pensar em gerar amostras negativas. Se um usuário comprou um item recomendado, você não deve adicioná-lo como amostra positiva e outros como amostras negativas.
- Pense na pontuação on-line e na pontuação offline da qualidade do seu algoritmo. Um modelo de treinamento apenas em dados históricos pode levar a recomendações primitivas porque o algoritmo não saberá sobre novas tendências e preferências.

## **TIPOS DE ALGORITMOS DE RECOMENDAÇÃO**

Os três principais tipos de algoritmos de recomendação são:

### **Filtragem colaborativa**

A filtragem colaborativa é uma das abordagens mais populares para sistemas de recomendação. Ela se baseia no princípio de que usuários com preferências semelhantes em itens também têm probabilidades semelhantes de gostar de outros itens. O algoritmo utiliza dados históricos de interações entre usuários e itens (por exemplo, avaliações, compras, cliques) para identificar padrões de similaridade entre eles. Com base nessas similaridades, o sistema recomenda itens que outros usuários com gostos similares apreciaram.

Existem duas formas principais de filtragem colaborativa: a abordagem baseada em usuário, que recomenda itens com base nos interesses de usuários semelhantes, e a abordagem baseada em item, que recomenda itens com base em características semelhantes aos itens previamente apreciados pelo usuário.

### **Filtragem baseada em conteúdo**

A filtragem baseada em conteúdo utiliza características e atributos dos itens para fazer recomendações. O algoritmo analisa informações descritivas sobre os itens e as combina com o perfil de preferências do usuário. Por exemplo, em um sistema de recomendação de filmes, o algoritmo pode considerar gênero, diretor, elenco e sinopse para sugerir filmes que possam agradar ao usuário com base em filmes que ele já assistiu e gostou.

Essa abordagem tende a ser mais personalizada, pois se concentra nas preferências específicas do usuário em vez de depender exclusivamente das opiniões de outros usuários.

### **Sistemas híbridos**

Os sistemas híbridos combinam as abordagens de filtragem colaborativa e filtragem baseada em conteúdo para melhorar a qualidade das recomendações. A ideia é que ao combinar as duas técnicas, é possível mitigar as limitações de cada uma individualmente.

Os algoritmos híbridos podem ser desenvolvidos de várias maneiras, como: utilizando filtragem colaborativa para inicialmente fazer recomendações amplas com

base em tendências populares e, em seguida, refinando essas sugestões com a filtragem baseada em conteúdo para personalizá-las de acordo com as preferências do usuário. Outra abordagem comum é atribuir pesos diferentes às duas técnicas, dependendo do contexto da recomendação.

Esses três tipos de algoritmos de recomendação desempenham um papel essencial em diversas aplicações, desde sistemas de entretenimento até plataformas de e-commerce, tornando-se uma parte fundamental da experiência do usuário em muitos serviços digitais.

## **CASOS DE USO REAL DE SISTEMAS DE RECOMENDAÇÃO**

### **Filtragem colaborativa. Exemplo: Netflix.**

O serviço de streaming de vídeo Netflix utiliza a filtragem colaborativa para fazer recomendações personalizadas de filmes e séries para seus usuários. Com base nas avaliações e no histórico de visualização de cada usuário, o algoritmo identifica padrões de comportamento semelhantes entre diferentes usuários. Por exemplo, se um usuário A e um usuário B têm avaliações semelhantes para um conjunto de filmes, o sistema pode inferir que eles têm gostos similares. Em seguida, o algoritmo pode recomendar filmes que o usuário A assistiu e gostou a outros usuários que demonstraram preferências semelhantes, como o usuário B.

### **Filtragem baseada em conteúdo. Exemplo: Spotify.**

O serviço de streaming de música Spotify utiliza a filtragem baseada em conteúdo para fazer recomendações de músicas e playlists personalizadas para seus usuários. O algoritmo analisa os atributos musicais das faixas que o usuário já ouviu e apreciou, como gênero, artista, ritmo e letra. Com base nesses atributos, o sistema pode sugerir novas músicas e playlists que se alinhem com o gosto musical específico do usuário. Por exemplo, se um usuário ouviu muitas músicas de artistas pop e playlists com "vibração de festa", o sistema pode recomendar outras músicas pop animadas que se adequem ao seu perfil musical.

### **Sistemas híbridos. Exemplo: Amazon.**

A plataforma de e-commerce Amazon utiliza um sistema híbrido para fazer recomendações de produtos para seus usuários. O algoritmo combina a filtragem colaborativa, que leva em conta o histórico de compras e interações dos usuários, com a filtragem baseada em conteúdo, que analisa as características dos produtos. Por exemplo, se um usuário comprou um smartphone com especificações técnicas específicas (filtragem baseada em conteúdo), o sistema pode recomendar acessórios ou produtos relacionados que outros usuários que compraram smartphones semelhantes também adquiriram (filtragem colaborativa).



## O QUE VOCÊ VIU NESTA AULA?

Esta aula proporcionou uma compreensão da arte por trás dos sistemas de recomendação e seu impacto nas experiências on-line. Ao adquirir esse conhecimento, você está apto e apta a explorar novas oportunidades e aplicar os princípios dos sistemas de recomendação em diferentes projetos, proporcionando interações personalizadas e experiências relevantes para os usuários.



## REFERÊNCIAS

CAVANI, C. **TensorFlow: Recomendação com ALS (Collaborative Filtering)**, 2015. Disponível em <<https://cirocavani.github.io/post/tensorflow-recomendacao-com-als-collaborative-filtering/>>. Acesso em: 07 ago. 2023.

LIAO, K. **Prototyping a Recommender System Step by Step Part 1: KNN Item-Based Collaborative Filtering**, 2018. Disponível em <<https://towardsdatascience.com/prototyping-a-recommender-system-step-by-step-part-1-knn-item-based-collaborative-filtering-637969614ea>>. Acesso em: 07 ago. 2023.

LIAO, K. **Prototyping a Recommender System Step by Step Part 2: Alternating Least Square (ALS) Matrix Factorization in Collaborative Filtering**, 2018. Disponível em: <<https://towardsdatascience.com/prototyping-a-recommender-system-step-by-step-part-2-alternating-least-square-als-matrix-4a76c58714a1>>. Acesso em: 07 ago. 2023.

TENSORFLOW Documentação Oficial. [s.d.]. Disponível em <<https://www.tensorflow.org/guide>>. Acesso em: 07 ago. 2023.

## **PALAVRAS-CHAVE**

Recomendação. Python. Predição. Machine Learning. ALS. Tensorflow. KNN. SVM. Random Forest.

EMSE

The background is a dark blue field filled with numerous small, light blue dots. Overlaid on this are several large, wavy, translucent lines in shades of blue and yellow. These lines flow from the left side towards the right, creating a sense of motion. Scattered throughout the composition are various geometric shapes: a thin vertical line, a circle containing the number '7', a small circle, an 'X' mark, a small circle, and a hexagon in the bottom right corner.

POSTECH