

Trabalho Prático - Parte 1

Sistema de Recomendação para Moda

Heitor L. Werneck

24 de junho de 2021

1 Introdução

1.1 Contextualização

Um crescimento explosivo de quantidade de informação na Internet criou um desafio de sobrecarga de informações que impede o acesso a itens de interesse. Esse fato aumentou a necessidade de sistemas de recomendação. Sistemas de recomendação são sistemas de filtragem de informações que lidam com o problema de sobrecarga de informações Konstan and Riedl [2012], filtrando fragmentos de informações de uma grande quantidade de informações que é gerada dinamicamente por usuários utilizando um sistema de informação de acordo com seus interesses e preferências sobre os itens do sistema. Um sistema de recomendação tem a capacidade de prever se um determinado usuário preferiria um item ou não com base no histórico de consumo dos usuários.

Os sistemas de recomendação tem uma importância muito grande para grandes comércios, já que esses mesmo sistema tem a capacidade de influenciar em quantos itens são comprados Jannach and Jugovac [2019]. Os ganhos de lucro são observados em diversos comércios, como por exemplo: 35% de aumento de vendas em um varejista de DVDs Lee and Hosanagar [2014]; no eBay, Brovman et al. reportaram um aumento de 6% em termos de lucro na implantação de um novo método de recomendação comparado ao modelo linear que estava sendo utilizado; diversos exemplos de sucesso são encontrados na literatura Jannach and Jugovac [2019].

A importância das vendas online no espaço da moda de luxo tem crescido em um ritmo acelerado nos últimos anos, já que os consumidores das indústrias tradicionais agora esperam fácil acesso a uma rede mundial de marcas e varejistas. Para ter sucesso neste cenário, é necessário fornecer uma experiência de compra de moda sob medida, personalizada e confiável. Os sistemas de recomendação desempenham um papel importante na jornada do usuário, permitindo que os clientes descubram produtos que atendem ao seu estilo, complementam suas escolhas ou os desafiam com novas ideias ousadas Farfetch [2021].

Sistemas de recomendação de moda são tratados na literatura de diferentes perspectivas, como por exemplo: Wang et al. integra temas de moda e percepção humana sobre formas corporais personalizadas e conhecimento de designers

profissionais; Dong et al. [2020] apresenta um sistema de recomendação de roupas que utiliza informações do formato 3D do corpo. No geral a literatura sobre recomendação no domínio de moda é escassa em relação a literatura de outros domínios, porém recentemente há um crescimento de interesse sobre domínio He and McAuley [2016], Frejlichowski et al. [2016], Wakita et al. [2015], Kang et al. [2017], Zeng et al. [2013]. Existem trabalhos que propõem diversas técnicas de filtragem colaborativa, porém estes trabalhos divergem em mínimos pontos, como por exemplo no objetivo que varia muito de estudo para estudo, como por exemplo o estudo de Hu et al. que foca na criação de conjuntos de roupa (recomendação de um conjunto de itens). Porém, no geral há uma escassez de sistemas de recomendação que focam nesse domínio a partir de informações mais fundamentais da interação do usuário, as exceções são trabalhos como: Nguyen et al. que modela um método de learning to rank para recomendação personalizada de moda via feedback implícito (cliques, lista de desejos e compras), ainda é um método que pode ser generalizado para outros domínio e não há uma robusta avaliação quantitativa, comparando com métodos estado da arte, para comprovação de bons resultados.

1.2 Objetivo

Este trabalho propõe a criação de um sistema de recomendação competitivo para o domínio de moda que consiga explorar a larga gama de informações dos itens e usuários dadas pelo sistema de maneira eficaz e eficiente, explorando padrões ainda não explorados no domínio. A base de dados de moda foi retirada de um desafio da FARFETCH para o ECML PKDD 2021, no qual o algoritmo criado aqui também tem o objetivo de ser utilizado na competição.

2 Proposta

A proposta deste trabalho consiste no desenvolvimento de um sistema de recomendação inovador para o domínio de roupas de moda, no qual especificamente será utilizado a base de dados de desafio da FARFETCH ¹. Este trabalho espera contribuir com:

- A modelagem de um sistema de recomendação competitivo para o cenário de recomendação de roupas de moda;
- Identificação dos padrões de consumo de usuários no domínio respectivo;
- Avaliação do modelo proposto comprovando sua utilidade;
- Submissão dos resultados para o desafio da FARFETCH no ECML PKDD 2021.

¹<https://www.ffrecschallenge.com/ecmlpkdd2021/>

3 Plano e Metodologia de Trabalho

Para este projeto o cronograma e etapas de trabalho serão descritos a seguir, o modelo de processo em espiral será utilizado para realização das atividades.

1. Estudo sobre o domínio e sistemas de recomendação: Primeira tarefa que consiste na obtenção de um conhecimento prévio sobre o domínio para capacitar a criação de soluções inovadoras. Essa tarefa também inclui o estudo e pesquisa de sistemas de recomendação próprios para o domínio, assim como genéricos e outros que exploram potenciais padrões de loja de moda.
2. Obtenção da base de dados e preparação dos dados: Esta é a primeira tarefa em relação a implementação que será realizada e consiste simplesmente na coleta da base de dados do domínio especificado e a primeira preparação dos dados para serem pre-processados na etapa seguinte;
3. Pré-processamento da base de dados: Esta etapa é onde haverá o pre-processamento da base, que inicialmente será somente para preparação dos dados para experimentação posterior, que será revisitada posteriormente a criação dos modelos de recomendação para aumentar a utilidade do recomendador;
4. Propor e implementar métodos de recomendação: Nessa etapa o sistema de recomendação será elaborado e implementado, também sistemas de recomendação já existentes serão implementados, como por exemplo: Random (baseline); Popular; Item-kNN; User-kNN; SVD; NMF; SVD++ Koren [2008]; NCF He et al. [2017] e outros. Assim como a composição desses métodos através de *ensembles* pode ser explorada.
 - 4.1. Implementação da metodologia de treinamento: Essa é uma etapa, que normalmente não seria destacada com seu próprio item, ficando implícito no item anterior, porém o treinamento e afinação de um modelo vem se mostrando extremamente importante, pois é um ponto na experimentação no qual diversos modelos se tornam não reprodutíveis em uma pequena mudança, sendo ela fundamental para bom funcionamento de um modelo (principalmente em um baseado em redes neurais) Crane [2018], Dacrema et al. [2019]. Então, nesta etapa ocorrerá a criação e definição de metodologias para treinamento dos algoritmos, tanto para parâmetros quanto para hiperparâmetros, como por exemplo treinamento por minibatch, ou afinação de hiperparâmetros com k-fold cross-validation e entre outros;
5. Implementação da metodologia de avaliação e função objetivo: Nessa etapa a metodologia de avaliação e uma métrica objetivo do problema serão implementadas, que no caso será o *Mean Reciprocal Rank* (MRR)) e seu ambiente para execução será definido para que posteriormente recomendadores possam ser avaliados;

6. Execução dos recomendadores e análise dos resultados: Nesta etapa o sistema criado previamente será analisado frente a função objetivo do problema, assim como padrões e casos de sucesso obtidos pela solução criada para o problema, podendo ser por outras métricas que abordam a utilidade de diferentes maneiras, como por exemplo: NDCG, F1, Recall e outras.
7. Escrita do artigo: Nesta etapa o trabalho será formalizado e a proposta será escrita no formato de artigo acadêmico, contendo: introdução, referencial teórico, artigos relacionados, metodologia e conclusão.

3.1 Cronograma

Abaixo, na tabela 1, um cronograma estimado foi criado sobre a execução das atividades.

Atividade a ser realizada	Meses			
	5	6	7	8
1. Estudo sobre o domínio e sistemas de recomendação	X	X		
2. Obtenção da base de dados e preparação dos dados	X			
3. Pré-processamento da base de dados		X		
4. Propor e implementar métodos de recomendação		X	X	
4.1. Implementação da metodologia de treinamento		X	X	
5. Implementação da metodologia de avaliação e função objetivo			X	
6. Execução dos recomendadores e análise dos resultados			X	
7. Escrita do artigo			X	X

Tabela 1: Cronograma.

4 Problema

O problema de recomendação foi definido por Burke et al. como: Seja U o conjunto de todos os usuários e seja I o conjunto de todos os itens possíveis que podem ser recomendados, seja π uma função de utilidade que mede a utilidade do item i para o usuário u , ou seja, $\pi : U \times I \rightarrow R$, onde R é um conjunto totalmente ordenado, então para cada usuário u queremos um item $i \in I$ que maximiza a utilidade do usuário.

Neste trabalho o problema será mais especificamente prever qual produto que foi mostrado a um usuário levou a um clique, baseado nas interações que possuem rótulo de clique. Ou seja, os itens são subconjuntos de I específicos para cada usuário, e no caso esse subconjunto tem cardinalidade 6.

Também o problema de sistemas de recomendação vem sendo estudado por muitos anos e é um problema no qual desafios para diversos domínios são conhecidos, no caso de moda desafios próprios do domínio ainda não são sistematicamente descritos pelo melhor do meu conhecimento. Porém a grande parte dos recomendadores sofrem por alguns problemas bem conhecidos, que são desafios

na área, como por exemplo Khusro et al. [2016]: *cold start problem*, no qual um novo usuário ou item entra no sistema; explicabilidade, que é geralmente encontrado a falta desse mesmo conceito em recomendadores baseados em redes neurais; *grey sheep*, no qual um usuário não está de acordo com nenhum outro grupo (métodos de filtragem colaborativa normalmente sofrem desse problema); esparsidade, onde os usuários normalmente interagem em um sistema com um amplo catálogo de itens e a distinção entre interações sobre itens de usuários torna a matriz de interações esparsa, o que pode levar a recomendações com menos acurácia.

5 Metodologia de avaliação

A metodologia de avaliação deste trabalho inicialmente estará de acordo com a especificação dada no desafio da FARFETCH.

Nesse problema a tarefa de recomendação será a predição do próximo produto a ser consumido pelo usuário de uma lista de 6 produtos (que corresponde a uma requisição ou interação do usuário com o sistema). Essa tarefa tem como objetivo a maximização da utilidade para os usuários, que é definida como o *Reciprocal Rank*.

Como dito anteriormente a utilidade do usuário será definida como o *Reciprocal Rank*, e a utilidade média do recomendador será computada usando o *Mean Reciprocal Rank* (MRR). O MRR é definido na equação 1, onde n são as impressões dos usuários sobre uma lista de itens recomendados, e nessa equação dado a impressão i o $rank_i$ é o rank da primeira predição correta.

$$MRR = \frac{1}{|n|} \sum_{i=1}^{|n|} \frac{1}{rank_i} \quad (1)$$

A base de testes é dada no desafio e consiste em amostras de 1 impressão para cada usuário, que podem estar no treino ou não, como será visto posteriormente na análise da base de dados.

6 Base de Dados

A base de dados que será utilizada foi provida pelo desafio da FARFETCH², no qual uma larga amostragem de impressões e eventos de clique no sistema de recomendação da FARFETCH foi feito durante um período de 2 meses. No qual, no total tem-se 5000000 eventos de recomendação, 450000 produtos e 230000 usuários únicos. A base de dados possui dados reais de interações de usuários com um sistema real na plataforma da FARFETCH, que porém são todos anonimizados (como por exemplo: usuário, produto, categorias, preços, marcas e etc). Além do contexto do usuário, os produtos também possuem atributos que o descreve. É importante notar que essa base possui diferentes

²<https://www.ffrecschallenge.com/ecmlpkdd2021/>

amostras para fases definidas pelo desafio e neste trabalho será usado os dados da fase 1. Na fase 1 temos na verdade somente 4195182 eventos, 219035 usuários e 443150 produtos, que é uma base com esparsidade de 99.996749%.

Uma coisa que será possível observar nas subseções a seguir é que grande parte dos atributos são categóricos. Os atributos anonimizados são representados por padrão como uma cadeia de 64 caracteres, que podem, posteriormente na fase de pre-processamento serem transformados em outros atributos e também podem ser selecionados.

6.1 Análise das Interações dos Usuários

Os atributos dos objetos das interações são descritos a seguir na tabela 2 (Obs: os valores únicos foram obtidos pelo conjunto de dados de treino):

Na base obtida, somente o atributo `context_type` possui valores faltantes, que corresponde a 4% dos objetos. Ou seja, no geral essa base de dados não sofre de muitas irregularidades. No geral não há um alto número de diferentes categorizações para as interações do usuário, diferentemente dos produtos como será visto posteriormente. O número baixo de atribuições para o atributo 'week' evidência um outro fato da base já mencionado anteriormente, que é uma base de um curto periodo temporal que pode limitar a utilização e exploração dos padrões temporais de longo prazo ou de temporadas, fazendo atributos temporais não parecerem tão interessantes a primeiro instante, porém padrões temporais de curto período ainda podem existir.

6.2 Análise dos Produtos

Os atributos dos objetos de produtos são descritos a seguir na tabela 3:

Não existe nenhum dado faltante nesse conjunto de dados. Pela tabela podemos ver a quantidade de valores únicos para esse conjunto de dados em cada um de seus atributos, primeiro é possível notar de maneira comparativa que muito mais categorias existem para os produtos, do que para os usuários, que é o esperado, pois normalmente os itens são cadastrados por um administrador que deseja maxima efetividade e completude na descrição de seus produtos. Sobre o ponto de vista de mineração de dados, uma ampla gama de atributos pode abrir possibilidade para encontrar diferentes e diversos padrões no sistema, que podem ser valiosos, que porém podem atrapalhar se a amostragem de treinamento for pequena.

6.3 Análise Geral e Básica

Dado toda essa meta-análise da base, será mostrado nessa seção mais análises simples sobre os dados e comportamentos superficiais dos mesmos, só para dar uma ideia do comportamento da base nessa proposta. Esse é só uma amostragem de todas análises feitas, pois seria inviável colocar todas formalizadas aqui.

Um dos primeiros pontos a se observar é o comportamento de consumo de itens pelos usuários. O gráfico abaixo mostra a taxa de consumo dos itens,

Atributo	Escala	Cardinalidade	#Valores únicos	Descrição
query_id	Nominal	Discreta	584665	identificador único da impressão de recomendação
user_id	Nominal	Discreta	208393	identificador único do usuário
session_id	Nominal	Discreta	317426	identificador único da sessão de navegação
product_id	Nominal	Discreta	408263	identificador único do produto
page_type	Nominal	Discreta	5	tipo de página em que a recomendação foi mostrada
previous_page_type	Nominal	Discreta	23	tipo de página em que o usuário navegou antes de ver a recomendação
device_category	Nominal	Discreta	3	tipo do dispositivo
device_platform	Nominal	Discreta	2	plataforma do dispositivo
user_tier	Ordinal	Discreta	6	identificador do nível de usuário no programa de recompensa FARFETCH
user_country	Nominal	Discreta	196	identificador do país de onde o usuário está navegando
context_type*	Nominal	Discreta	4	uma string que identifica o tipo de contexto. Os valores possíveis são "designer_id", "category_id" e "product_id" (correspondendo às páginas de lista de designer, páginas de lista de categoria e páginas de produto, respectivamente)
context_value*	Nominal	Discreta	187906	o ID correspondente da entidade fornecida em context_type
is_click	Ordinal	Binária	2	1 se o produto foi clicado, 0 se não
product_price	Razão	Continua	1757598	preço normalizado do produto no momento dado
week	Ordinal	Discreta	8	número normalizado da semana em que o evento aconteceu
week_day	Ordinal	Discreta	7	dia da semana em que o evento aconteceu

Tabela 2: Descrição geral dos atributos das interações. "*" indica atributo que pode conter múltiplos valores.

é possível observar que, assim como a maioria dos domínios de recomendação, possui uma cauda longa entre usuários e catálogo de itens.

Também usuários nesse domínio de exibição de produtos para cliques podem recorrentemente clicar nos mesmos itens, como pode-se ver na tabela 4, porém a probabilidade é baixa como é possível observar na tabela abaixo que apresenta as porcentagens de frequências de clique de usuários em produtos, ou seja, nesse caso 96.8% das interações entre usuário e produto tiveram somente 1 clique.

Também um outro ponto observado sobre a base de dados foi a presença de usuários e itens *cold start* no teste, isto é, a presença de usuários é de 9% e itens de 7%. Parte significativa do teste são esses casos, o que mostra a necessidade

Atributo	Escala	Cardinalidade	#Valores únicos	Descrição
product_id	Nominal	Discreta	443150	identificador do produto
gender	Nominal	Discreta	4	gênero do produto
main_colour	Nominal	Discreta	17	a cor principal do produto
second_colour	Nominal	Discreta	22	a segunda cor mais predominante do produto
season	Nominal	Discreta	7	temporada da moda
collection	Nominal	Discreta	9	identificador da coleção de moda
category_id_l1	Nominal	Discreta	43	primeiro nível da árvore de categorias de produtos
category_id_l2	Nominal	Discreta	394	segundo nível da árvore de categorias de produtos
category_id_l3	Nominal	Discreta	583	terceiro nível da árvore de categorias de produtos
brand_id	Nominal	Discreta	3399	identificador da marca do produto
season_year	Nominal	Discreta	20	ano da temporada da moda
start_online_date	Razão	Contínua	2150	número de dias online em relação a uma data de referência predefinida
attribute_values*	Nominal	Discreta	859	atributos do produto, como informações adicionais sobre o tipo de produto e padrões
material_values*	Nominal	Discreta	129	composição do material do produto

Tabela 3: Descrição geral dos atributos dos produtos. "*" indica atributo que pode conter múltiplos valores.

Quantidade de cliques	Porcentagem de pares usuário-produto
1	96.888511%
2	2.728706%
3	0.293456%
4	0.058175%
5	0.018416%
6	0.006712%
7	0.002754%
8	0.001377%
9	0.000516%
10	0.001205%
11	0.000172%

Tabela 4: Distribuição de frequência de cliques em produtos por usuários

de técnicas que tratam este problema para atingir um bom resultado. Também mensurar os ganhos sobre esse grupo de usuários é interessante.

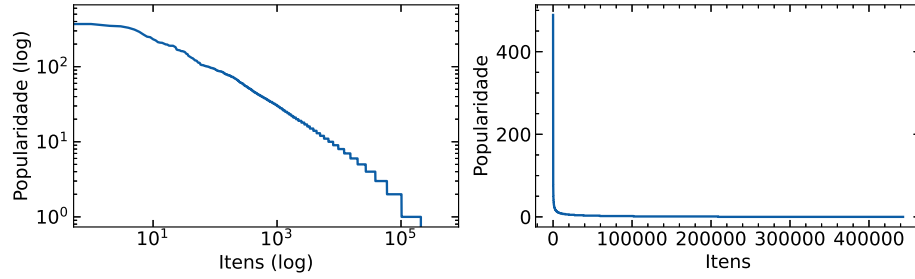


Figura 1: Popularidade dos itens.

7 Resultados Preliminares

Atualmente diversas proposta de recomendadores foram testadas como benchmark inicial, sem um treinamento e afinação de parâmetros robusto. Entre elas temos:

1. NeuMF: Que é um método baseado em redes neurais, que funde a fatorização generalizada de matrizes por meio de redes neurais com camadas de multilayer perceptron. Uma ilustração do modelo é apresentada abaixo:

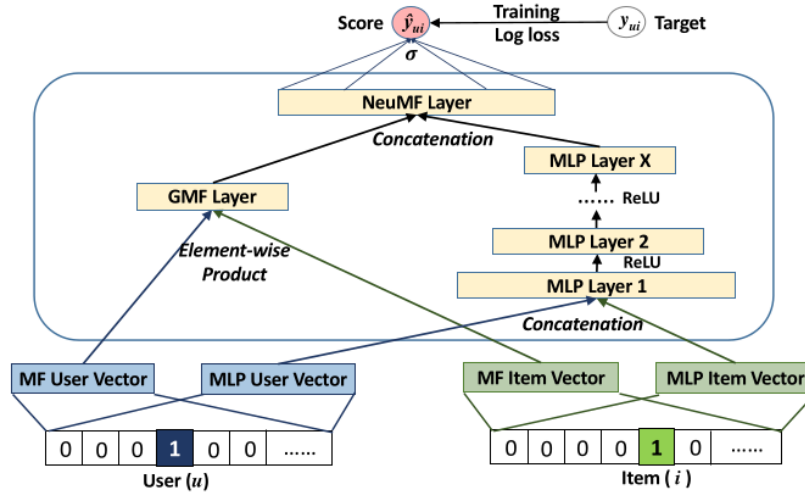


Figura 2: NeuMF.

2. Popular: Popular é um modelo tradicional de baseline de sistemas de recomendação, esse método irá rankear os itens de acordo com o mais popular. A popularidade foi definida nesse trabalho como o número de diferentes usuários que clicaram no item.

3. SVD (Singular Value Decomposition): O SVD é um método de fatorização de matrizes que é utilizado em sistemas de recomendação para filtragem colaborativa.
4. BilinearNet: É um baseline simples de aprendizado de embeddings com redes neurais, que não faz uso de nenhuma camada e aprende as representações dos usuários e itens.
5. Random: Um modelo simples, para definir o lower bound do problema, que recomenda itens de maneira aleatória.
6. SVD++: É um método que minimiza uma função objetivo, que inclui um viés geral e outros 2 para os usuários e os itens, utilizando Stochastic Gradient Descent (SGD) assim como regularização com L2.
7. Item-kNN: Método que mede a correlação dos itens e depois usa grupos desses itens para identificar o conjunto de itens a ser recomendado.
8. User-kNN: Método que diferentemente do Item-kNN, ao invés de medir a similaridade entre itens, mede a similaridade entre usuários para após isso obter os itens que devem ser recomendados.

Esses algoritmos e outros que não tiveram bons resultados foram executados, a tabela abaixo apresenta o MRR de cada um desses métodos:

Métodos	MRR
NeuMF	0.4315
Popular	0.4213
SVD	0.4142
BilinearNet	0.4139
Random	0.4119
SVD++	0.3873
Item-kNN	Overflow de memória
User-kNN	Overflow de memória

Tabela 5: Resultados de MRR.

Esses resultados ainda não são definitivos, problemas podem ter sido feitos nessa etapa do trabalho, que é muito inicial, porém podemos ver que o aleatório não se diferencia muito dos outros algoritmos. O que se destaca é o NeuMF com bons resultados, em segundo o Popular que conseguiu resultados consideráveis comparado aos outros algoritmos, demonstrando que nessa base os outros modelos com um pouco mais de complexidade talvez não sejam suficientes. O SVD++ que é um baseline normalmente competitivo teve resultados ruins, demonstrando a não trivialidade do domínio.

Os métodos apresentam baixa diferenciação de resultados em relação ao aleatório, demonstrando que os métodos atuais ainda não estão bons, assim como o problema é difícil.

Os usuários e itens cold start podem ter grande impacto nesses resultados, o que necessita de maior investigação em trabalhos futuros.

Outras investigações de inserção de contexto em métodos foram óvadas, porém ainda não foi possível obter um resultado satisfatório.

8 Conclusão

Com este trabalho inicial foi possível definir formalmente o problema assim como demonstrar a importância do seu desenvolvimento. O desafio a ser solucionado foi formalizado e a metodologia de avaliação já foi possível ser definida. Assim como um entendimento melhor sobre a base de dados a ser utilizada foi dado.

Já foi possível observar alguns resultados preliminares, que parecem evidenciar a baixa capacidade de aprendizado dos métodos. Trabalhos futuros também precisam ser realizados, como por exemplo: avaliar melhor e aprimorar a metodologia de treinamento dos algoritmos; inserir contexto dos usuários e dos itens na recomendação personalizada; adicionar mecanismos para recomendação de itens para usuários cold start; explorar ideias de seleção de features ou redução de dimensionalidade após a criação de métodos que suportam features externas; reduzir o problema de overfitting nos modelos de redes neurais; modelagem do problema de predição como uma tarefa de prever o próximo item na sequência.

Referências

- Yuri M Brovman, Marie Jacob, Natraj Srinivasan, Stephen Neola, Daniel Galron, Ryan Snyder, and Paul Wang. Optimizing similar item recommendations in a semi-structured marketplace to maximize conversion. In *Proceedings of the 10th ACM Conference on Recommender Systems*, pages 199–202, 2016.
- Robin Burke, Alexander Felfernig, and Mehmet H Göker. Recommender systems: An overview. *Ai Magazine*, 32(3):13–18, 2011.
- Matt Crane. Questionable answers in question answering research: Reproducibility and variability of published results. *Transactions of the Association for Computational Linguistics*, 6:241–252, 2018.
- Maurizio Ferrari Dacrema, Paolo Cremonesi, and Dietmar Jannach. Are we really making much progress? a worrying analysis of recent neural recommendation approaches. In *Proceedings of the 13th ACM Conference on Recommender Systems*, pages 101–109, 2019.
- Min Dong, Xianyi Zeng, Ludovic Koehl, and Junjie Zhang. An interactive knowledge-based recommender system for fashion product design in the big data environment. *Information Sciences*, 540:469–488, 2020.
- Farfetch, 2021. URL <https://www.ffrecschallenge.com/ecmlpkdd2021/>.

- Dariusz Frejlichowski, Piotr Czapiewski, and Radosław Hofman. Finding similar clothes based on semantic description for the purpose of fashion recommender system. In *Asian Conference on Intelligent Information and Database Systems*, pages 13–22. Springer, 2016.
- Ruining He and Julian McAuley. Ups and downs: Modeling the visual evolution of fashion trends with one-class collaborative filtering. In *proceedings of the 25th international conference on world wide web*, pages 507–517, 2016.
- Xiangnan He, Lizi Liao, Hanwang Zhang, Liqiang Nie, Xia Hu, and Tat-Seng Chua. Neural collaborative filtering. In *Proceedings of the 26th international conference on world wide web*, pages 173–182, 2017.
- Yang Hu, Xi Yi, and Larry S Davis. Collaborative fashion recommendation: A functional tensor factorization approach. In *Proceedings of the 23rd ACM international conference on Multimedia*, pages 129–138, 2015.
- Dietmar Jannach and Michael Jugovac. Measuring the business value of recommender systems. *ACM Transactions on Management Information Systems (TMIS)*, 10(4):1–23, 2019.
- Wang-Cheng Kang, Chen Fang, Zhaowen Wang, and Julian McAuley. Visually-aware fashion recommendation and design with generative image models. In *2017 IEEE International Conference on Data Mining (ICDM)*, pages 207–216. IEEE, 2017.
- Shah Khusro, Zafar Ali, and Irfan Ullah. Recommender systems: issues, challenges, and research opportunities. In *Information Science and Applications (ICISA) 2016*, pages 1179–1189. Springer, 2016.
- Joseph A Konstan and John Riedl. Recommender systems: from algorithms to user experience. *User modeling and user-adapted interaction*, 22(1):101–123, 2012.
- Yehuda Koren. Factorization meets the neighborhood: a multifaceted collaborative filtering model. In *Proceedings of the 14th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 426–434, 2008.
- Dokyun Lee and Kartik Hosanagar. Impact of recommender systems on sales volume and diversity. ., 2014.
- Hai Thanh Nguyen, Thomas Almenningen, Martin Havig, Herman Schistad, Anders Kofod-Petersen, Helge Langseth, and Heri Ramampiaro. Learning to rank for personalised fashion recommender systems via implicit feedback. In *Mining Intelligence and Knowledge Exploration*, pages 51–61. Springer, 2014.
- Yuka Wakita, Kenta Oku, Hung-Hsuan Huang, and Kyoji Kawagoe. A fashion-brand recommender system using brand association rules and features. In *2015 IIAI 4th International Congress on Advanced Applied Informatics*, pages 719–720. IEEE, 2015.

- LC Wang, XY Zeng, Ludovic Koehl, and Yan Chen. Intelligent fashion recommender system: Fuzzy logic in personalized garment design. *IEEE Transactions on Human-Machine Systems*, 45(1):95–109, 2014.
- Xianyi Zeng, Ludovic Koehl, Lichuan Wang, and Yan Chen. An intelligent recommender system for personalized fashion design. In *2013 Joint IFSA World Congress and NAFIPS Annual Meeting (IFSA/NAFIPS)*, pages 760–765. IEEE, 2013.