

Analizando a Justiça de Grupo em Sistemas de Recomendação: Uma Avaliação de Estratégias de Filtragem e Agrupamentos de Usuários no Dataset MovieLens

Rafael V. M. Santos¹, Giovanni V. Comarela²

¹Departamento de Informática – Instituto Federal do Espírito Santo (IFES)
Caixa Postal 15.064 – 91.501-970 – Cachoeiro de Itapemirim – ES – Brazil

²Departamento de Informática – Universidade Federal do Espírito Santo (UFES)
Vitória, ES.

rafaelv@ifes.edu.br, gc@inf.ufes.br

Abstract. *This study investigates equity in recommendation systems using the MovieLens dataset, applying collaborative filtering strategies: ALS, KNN, and NMF. We assess injustice in different grouping configurations: Gender, Age, Ratings, and Agglomerative. The results indicate significant variations in injustice among the strategies, with the Agglomerative method standing out for presenting the highest levels of group injustice in most approaches. This analysis suggests the need for careful selection of the filtering strategy and grouping method to promote fairer and more inclusive recommendation systems, highlighting the importance of considering multiple dimensions of injustice in the design of these systems.*

Resumo. *Este estudo investiga a equidade em sistemas de recomendação utilizando o dataset MovieLens, aplicando estratégias de filtragem colaborativa: ALS, KNN e NMF. Avaliamos a injustiça em diferentes configurações de agrupamento: Gênero, Idade, Avaliações e Aglomerativo. Os resultados indicam variações significativas de injustiça entre as estratégias, com o método Aglomerativo destacando-se por apresentar os maiores níveis de injustiça do grupo na maioria das abordagens. Esta análise sugere a necessidade de uma seleção cuidadosa da estratégia de filtragem e do método de agrupamento para promover sistemas de recomendação mais justos e inclusivos, destacando a importância de considerar múltiplas dimensões de injustiça na concepção destes sistemas.*

1. Introdução

Em um cenário onde as interfaces digitais se tornam cada vez mais interativas e personalizadas, os sistemas de recomendação surgem como ferramentas essenciais, moldando de forma significativa as escolhas e interações dos usuários nas plataformas online. Quanto mais se conhece sobre o usuário, melhor será a qualidade dos itens recomendados a ele [Hazrati and Ricci 2024, Pereira et al. 2018].

Sistemas de recomendação fornecem sugestões de itens para seus usuários, sendo hoje incorporados a sites de comércio eletrônico, bibliotecas digitais e redes sociais [Liu et al. 2024, Souza et al. 2022]. E a disseminação desses sistemas destaca a necessidade urgente de avaliar e mitigar potenciais repercussões sociais adversas que possam

emergir, especialmente à medida que se integram profundamente em redes sociais e ambientes digitais.

Investigações recentes, como as realizadas por [Deldjoo et al. 2024], revelam inadvertidamente que os sistemas de recomendação podem intensificar vieses e desigualdades, criando assim disparidades no serviço oferecido a diferentes segmentos da população. Esta constatação destaca a complexidade das interações sociais mediadas por tais sistemas e a importância de promover práticas equitativas e justas.

A crescente dependência desses sistemas e sua influência nas dinâmicas de redes sociais exigem um exame detalhado de seus impactos, visando desenvolver soluções que assegurem um espaço digital equitativo e inclusivo. Destaca-se, portanto, a necessidade de transcender a tradicional métrica de precisão em sistemas de recomendação, incorporando a justiça como um parâmetro crucial na avaliação do impacto dessas tecnologias sobre a sociedade.

Este estudo propõe uma abordagem inovadora, integrando métricas de justiça em algoritmos de recomendação, com o objetivo de elucidar sobre as desigualdades presentes. Por meio de uma análise detalhada aplicada ao dataset MovieLens, consideramos variáveis sociodemográficas e comportamentais, como gênero, idade e frequência de avaliações de itens, além da utilização de métodos de agrupamento hierárquico para revelar grupos ocultos. A inclusão deste último método tem por objetivo a identificação de grupos potencialmente latentes, cujas características não são prontamente aparentes, visando uma abordagem mais holística e inclusiva no exame das disparidades de justiça social manifestas nas recomendações do sistema.

2. Trabalhos Relacionados

Esta seção inicia delineando as definições de justiça que fundamentam este estudo.

Justiça tornou-se um assunto de interesse ampliado no domínio da aprendizagem automática. Na sequência desta discussão, um sistema de recomendação será considerado justo se garantir uniformidade na qualidade do serviço (isto é, acurácia das previsões) para todos os indivíduos ou grupos de usuários [Zafar et al. 2017].

Prosseguiremos discutindo a relação entre as nossas métricas de justiça para sistemas de recomendação e aquelas introduzidas em estudos precedentes.

Equidade em Aprendizado de Máquina e Sistemas de Recomendação: A consciência sobre os impactos sociais negativos oriundos da aplicação de algoritmos de aprendizado de máquina em contextos decisórios tem crescido [Sudakov and Titov 2023, Emirhüseyinoğlu et al. 2023]. Como resposta, várias métricas e conceitos de justiça foram propostos para tarefas de aprendizado de máquina, englobando classificação [Dargan et al. 2024, Feng et al. 2023], regressão [Xue et al. 2024, Sikora et al. 2023], ordenação [Bhargava et al. 2022, Zehlike et al. 2022], e escolha de conjuntos [Vashistha et al. 2023, Celis et al. 2016]. Essas propostas se dividem em duas categorias principais: avaliações de justiça em nível individual e em nível de grupo [Dwork et al. 2011].

Em contraste com tarefas tradicionais de aprendizado, como classificação e regressão, a exploração de justiça em sistemas de recomendação permanece limitada. [Burke et al. 2018] salientaram que sistemas de recomendação que antecipam

preferências de usuário por itens devem assegurar equidade para ambos os lados: tanto para os usuários que recebem recomendações quanto para os itens recomendados. Inicialmente, trabalhos de [Kamishima and Akaho 2017, Kamishima et al. 2012, Kamishima et al. 2018] abordaram a equidade em nível de grupo, adaptando modelos de aprendizado para que as sugestões de itens fossem neutras em relação a atributos de usuários como etnia e gênero. Posteriormente, [Beutel et al. 2017] e [Yao and Huang 2017] definiram critérios de equidade em nível de grupo para sistemas de recomendação, fundamentados na acurácia das previsões para diferentes grupos de usuários ou itens.

Nós podemos resumir as contribuições deste trabalho da seguinte forma:

- Analisamos a justiça de grupo utilizando, dentre outros, o método de agrupamento aglomerativo (método hierárquico) para agrupar usuários. Isso difere de estudos anteriores [Liu et al. 2022, Rastegarpanah et al. 2019] que se concentraram no agrupamento de usuários com base em uma única característica sensível, como gênero, raça ou idade;
- Avaliamos três estratégias de filtragem colaborativa distintas em sistemas de recomendação. Assim como no estudo proposto por [Leonhardt et al. 2018], a escolha intencional de métodos variados, incluindo abordagens baseadas em modelo e memória, visa comparar e avaliar diferentes técnicas de filtragem colaborativa de forma abrangente. Todavia, introduzimos uma contribuição adicional ao incorporar duas estratégias baseadas em modelo para a análise comparativa de abordagens distintas no tratamento de dados faltantes.

3. Material e Métodos

Nesta seção, detalhamos a metodologia aplicada, incluindo a análise do dataset MovieLens e a utilização de técnicas de filtragem colaborativa para investigar medidas de justiça social em sistemas de recomendação.

3.1. Base de Dados

O estudo de caso utilizou o dataset MovieLens 1M¹[Harper and Konstan 2015], que contém cerca de 1 milhão de classificações de ≈ 4000 filmes feito por ≈ 6000 usuários, com avaliações em uma escala de 5 pontos. Consideramos para este trabalho os 300 usuários com o maior número de avaliações, além dos 1000 filmes mais avaliados.

3.2. Estratégias de Recomendação

Para as estimativas das classificações desconhecidas, foram testadas três diferentes estratégias de filtragem colaborativa em sistemas de recomendação, cada uma com características distintas. Especificamente, foram escolhidos métodos com abordagens diferenciadas para avaliar e comparar o comportamento dessas técnicas:

- **ALS** (*Alternating Least Squares*): Este método, baseado em modelo, é empregado na decomposição de matrizes em fatores, minimizando o erro quadrático de forma alternada ao fixar um fator por vez. Utiliza máscaras booleanas para ignorar diretamente os valores faltantes durante a otimização;

¹<https://github.com/ravarmes/recsys-rgrp-movielens/tree/main/data/MovieLens-1M>

- **NMF** (*Non-negative Matrix Factorization*): Também baseado em modelo, utiliza técnicas de otimização iterativa com restrição de não negatividade. Necessita do preenchimento dos valores faltantes com médias ou outras imputações antes da otimização, devido à impossibilidade de processá-los diretamente;
- **KNN** (*K-Nearest Neighbors*): Por outro lado, o KNN é um método baseado em memória, centrado na similaridade entre usuários ou itens. Para um dado usuário ou item, o sistema identifica os ‘k’ vizinhos mais próximos (sendo ‘k’ um número inteiro) e faz recomendações com base nas preferências desses vizinhos;

Esta revisão ressalta a escolha intencional de métodos com abordagens distintas, incorporando a diferenciação entre métodos baseados em modelo (ALS e NMF) e baseado em memória (KNN). Além disso, destaca a importância da análise do manejo de dados faltantes, particularmente relevante para avaliações de justiça nos resultados.

3.3. Agrupamento dos Usuários

Os usuários foram agrupados por gênero, idade, número de avaliações de itens e por meio da técnica de agrupamento aglomerativo. Considerando que indivíduos de uma mesma rede tendem a formar grupos sociais estáveis [Coelho et al. 2023].

- Gênero: masculino e feminino;
- Idade: menores que 18 anos, de 18 até 24 anos, de 25 até 34 anos, de 35 até 44 anos, de 45 até 49 anos, de 50 até 55 anos, e maiores que 55 anos;
- Número de avaliações de itens (95-5): um grupo contendo 5% dos usuários com maior número de avaliações realizadas, e os demais usuários (95%) considerados em outro grupo. O grupo dos 5% seriam os usuários favorecidos, enquanto que o grupo de 95% são os usuários desfavorecidos;
- Agrupamento aglomerativo²: o número ideal de grupos foi determinado com base na análise dos escores de silhueta³. Utilizou-se um método hierárquico para agrupar os usuários em cinco grupos distintos, considerando características como gênero, idade e número de avaliações. O objetivo foi identificar grupos que não fossem imediatamente evidentes.

Na Tabela 1 temos a quantidade de usuários nos grupos dentro das configurações de agrupamentos.

3.4. Módulo do Algoritmo: Cálculo de Medidas Sociais

Nesta subseção apresentaremos o módulo do algoritmo elaborado para calcular as medidas de justiça social do estudo de caso proposto. É pertinente mencionar que todas as implementações das medidas de equidade utilizadas no algoritmo de justiça proposto foram baseadas no trabalho de [Rastegarpanah et al. 2019], fornecendo uma base sólida para nossa abordagem no tratamento da justiça social em sistemas de recomendação.

A implementação de todos os códigos da análise de justiça de grupo no conjunto de dados MovieLens está disponível no repositório [ravarmes/recsys-rgrp-movielens](https://github.com/ravarmes/recsys-rgrp-movielens)⁴ no GitHub.

²<https://scikit-learn.org/stable/modules/generated/sklearn.cluster.AgglomerativeClustering.html>

³https://scikit-learn.org/stable/modules/generated/sklearn.metrics.silhouette_score.html

⁴<https://github.com/ravarmes/recsys-rgrp-movielens>

Tabela 1. Detalhamento do agrupamento de usuários

Agrupamento	Grupos	Quantidade
Gênero	Masculino	240
	Feminino	60
Idade	Abaixo de 18 anos	5
	De 18 até 24 anos	53
	De 25 até 34 anos	142
	De 35 até 44 anos	58
	De 45 até 49 anos	24
	De 50 até 55 anos	12
Avaliações	Acima de 55 anos	6
	Favorecidos	15
	Desfavorecidos	285
	Aglomerativo	Grupo 1
Grupo 2		69
Grupo 3		42
Grupo 4		25
Grupo 5		18

Tendo em vista todas as especificações e discussões da seção anterior, iremos definir formalmente as métricas que especificam as funções objetivo associadas à justiça individual e justiça do grupo.

Começaremos apresentando a configuração do sistema, notação e a definição do problema. Suponha que $X \in \mathbb{R}^{n \times m}$ seja uma matriz de classificação parcialmente observada de n usuários e m itens, de modo que o elemento x_{ij} denota a classificação dada pelo usuário i para o item j . Seja Ω o conjunto de índices de classificações conhecidas em X . Além disso, Ω_i denota os índices de classificações de itens conhecidos para o usuário i , e Ω_j denota os índices de classificações de usuários conhecidos para o item j .

Para uma matriz A , $P_\Omega(A)$ é uma matriz cujos elementos em $(i, j) \in \Omega$ são a_{ij} e zero nas outras posições. Da mesma forma, para um vetor a , $P_{\Omega_j}(a)$ é um vetor cujos elementos em $i \in \Omega_j$ são os elementos correspondentes de a e zero em outras posições. Ao longo do artigo, denotamos a coluna j de A pelo vetor a_j e a linha i de A pelo vetor a^i . Todos os vetores são vetores de coluna.

Dado um sistema de recomendação tradicional é gerada uma matriz estimada de recomendações $\hat{X} = [\hat{X}_{ij}]_{n \times m}$. Neste problema de recomendação supomos usuários em um conjunto $\{u_1, u_2, \dots, u_n\}$ e itens em um conjunto $\{v_1, v_2, \dots, v_m\}$.

Justiça Individual. Para cada usuário i , definimos ℓ_i , a perda do usuário para i , como a estimativa do erro quadrático médio sobre as avaliações conhecidas do usuário i . E a injustiça individual R_{indv} como a variação das perdas dos usuários.

$$\ell_i = \frac{\|P_{\Omega^i}(\hat{\mathbf{x}}^i - \mathbf{x}^i)\|_2^2}{|\Omega^i|} \quad R_{indv}(X, \hat{X}) = \frac{1}{n^2} \sum_{k=1}^n \sum_{l>k} (\ell_k - \ell_l)^2 \quad (1)$$

Justiça de Grupo. Seja I o conjunto de todos os usuários/itens e $G = \{G_1, G_2, \dots, G_g\}$ uma partição de usuários/itens em g grupos, ou seja, $I = \bigcup_{i \in \{1, 2, \dots, g\}} G_i$. Definimos a perda do grupo L_i como a estimativa do erro quadrático médio sobre todas as avaliações conhecidas no grupo i . E dada uma partição G , a injustiça do grupo R_{grp} é a variação de todas as perdas do grupo:

$$L_i = \frac{\|P_{\Omega_{G_i}}(\hat{X} - X)\|_2^2}{|\Omega_{G_i}|} \quad R_{grp}(X, \hat{X}, G) = \frac{1}{g^2} \sum_{k=1}^g \sum_{l>k} (L_k - L_l)^2 \quad (2)$$

4. Resultados e Discussões

Nesta seção, mostramos um comparativo das medições de justiça do grupo em diferentes estratégias de filtragem colaborativa, bem como, diversas configurações de agrupamento de usuários.

Na sequência, apresentamos os resultados obtidos em nosso experimento, os quais são essenciais para a compreensão do impacto das diferentes estratégias de filtragem colaborativa analisadas. Os dados, detalhados por estratégia de filtragem e tipo de agrupamento utilizado, são sumarizados na Tabela 2. Este conjunto de resultados permite uma avaliação comparativa do desempenho das estratégias em termos da medida de injustiça de grupo R_{grp} .

Tabela 2. Injustiça de Grupo (R_{grp}) por estratégia e agrupamento

Estratégia	Agrupamento	R_{grp}
ALS	Gênero	0,00426530
	Idade	0,00170270
	Avaliações	0,00129660
	Aglomerativo	0,00615080
NMF	Gênero	0,00301780
	Idade	0,00161200
	Avaliações	0,00400160
	Aglomerativo	0,00487470
KNN	Gênero	0,00053500
	Idade	0,00767110
	Avaliações	0,00195270
	Aglomerativo	0,00304140

A Figura 1 compara as métricas de injustiça de grupo para diferentes configurações de agrupamento, destacadas por estratégia de filtragem. Os grupos são representados por barras, sendo um gráfico por estratégia, facilitando a identificação de quais promovem maior equidade. Esta análise visual ajuda a revelar as diferenças significativas no impacto de cada estratégia sobre a injustiça nos grupos avaliados.

Pela Figura 1 o método Aglomerativo se destaca por elevar os níveis de injustiça em todas as estratégias, reforçando a importância de escolhas criteriosas em estratégias de filtragem e agrupamento para reduzir tais desigualdades.

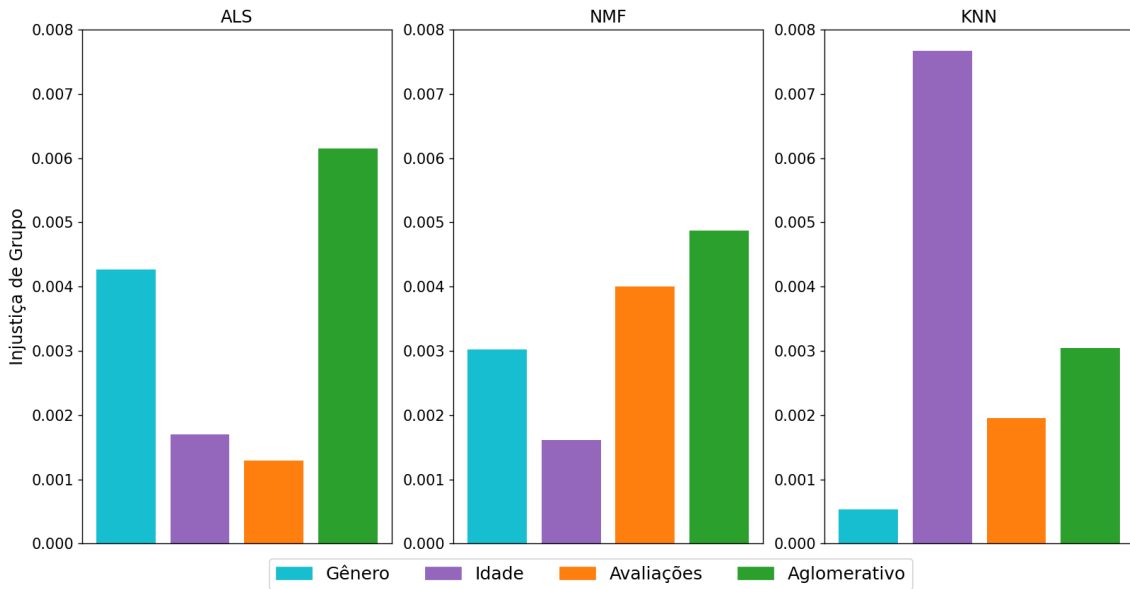


Figura 1. Injustiça de grupo R_{grp} por estratégias de filtragem e agrupamentos

A Figura 2 detalha um pouco mais sobre a injustiça do grupo, mostrando as perdas de grupo L_i em cada um dos grupos dentro da configuração de agrupamentos.

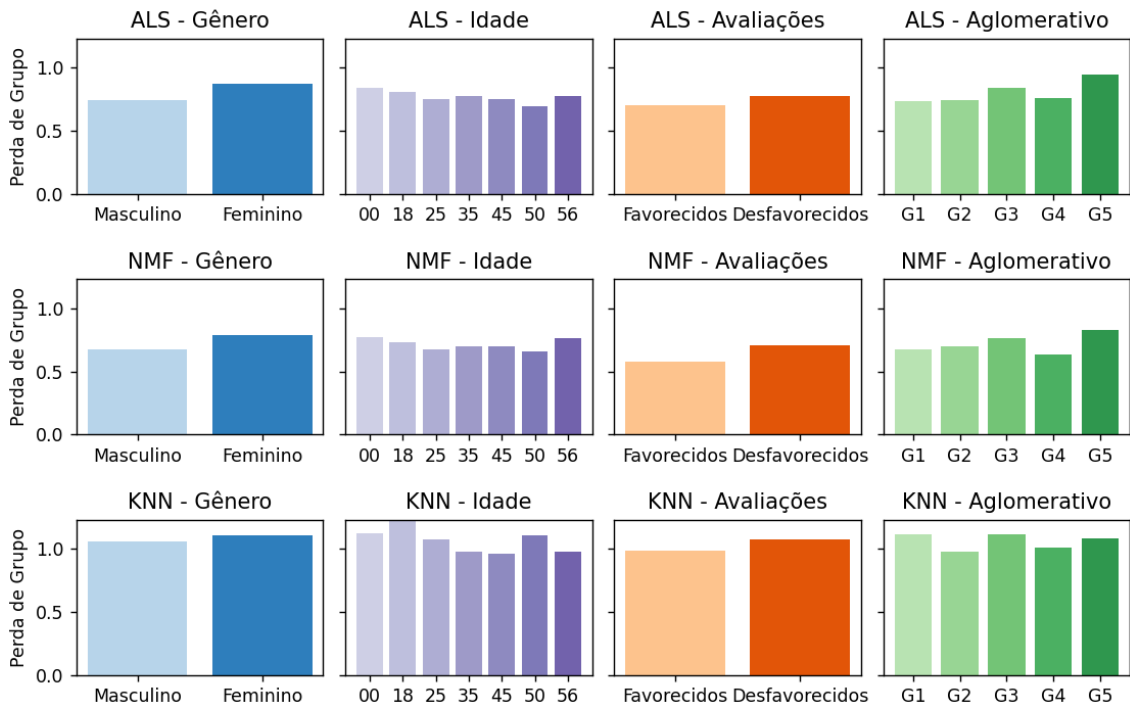


Figura 2. Perdas de grupo L_i por estratégias de filtragem e agrupamentos

A Figura 3 oferece uma visão detalhada da distribuição da injustiça do grupo para as estratégias de filtragem colaborativa ALS, KNN e NMF, utilizando diferentes configurações de agrupamento. Através dos gráficos de caixa, é possível observar as variações, medianas, e a presença de outliers nas medições de injustiça, proporcionando uma análise comparativa entre as estratégias. Esta representação visual destaca as

diferenças na performance de cada estratégia de filtragem em termos de equidade, permitindo uma avaliação mais precisa da eficácia de cada abordagem no contexto de sistemas de recomendação mais justos e inclusivos.

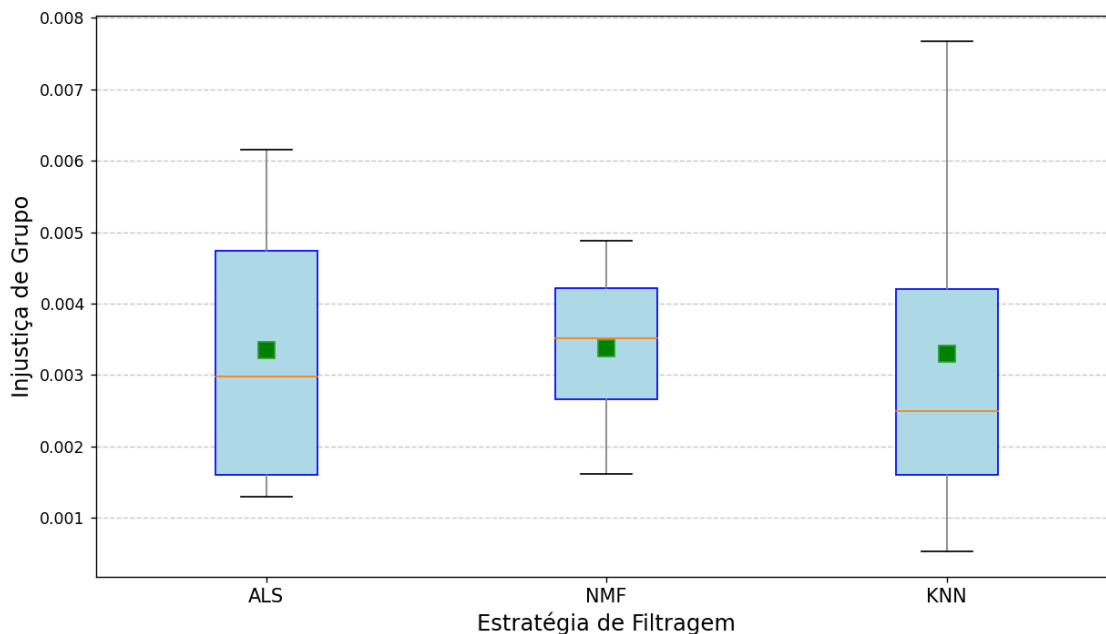


Figura 3. Distribuição da injustiça por estratégia de filtragem

Temos na Figura 3 variações notáveis na injustiça do grupo entre as estratégias de filtragem colaborativa ALS, KNN e NMF, com a KNN apresentando a maior variação e outliers que indicam disparidades significativas. A presença de outliers na estratégia KNN enfatiza a necessidade de investigar e ajustar as configurações que conduzem a altas injustiças. Os resultados sugerem que uma abordagem mais refinada é necessária para desenvolver sistemas de recomendação que sejam equitativos e inclusivos, atendendo às necessidades de todos os usuários de forma justa.

A Figura 4 apresenta uma análise comparativa da injustiça de grupo em sistemas de recomendação através de dois gráficos de radar lado a lado, oferecendo perspectivas distintas sobre as estratégias de filtragem colaborativa (ALS, KNN, NMF) e suas interações com diferentes agrupamentos de usuários (Gênero, Idade, Avaliações, Aglomerativo). O primeiro gráfico ilustra como cada estratégia de filtragem impacta os níveis de injustiça para os diversos agrupamentos, permitindo uma visão direta das variações entre as estratégias. Inversamente, o segundo gráfico destaca a influência de cada agrupamento de usuários nas estratégias de filtragem, evidenciando a sensibilidade de cada estratégia a diferentes agrupamentos. Juntos, esses gráficos de radar proporcionam uma visão holística e detalhada sobre a dinâmica entre estratégias de filtragem e agrupamentos de usuários, enfatizando a complexidade da promoção da equidade em sistemas de recomendação.

Na Figura 4 desdobramos a análise de injustiça de grupo em uma dupla perspectiva visual através de gráficos de radar, ampliando a compreensão sobre como as estratégias de filtragem colaborativa (ALS, KNN, NMF) e agrupamentos de usuários (Gênero, Idade, Avaliações, Aglomerativo) interagem entre si. No primeiro gráfico, observa-se uma

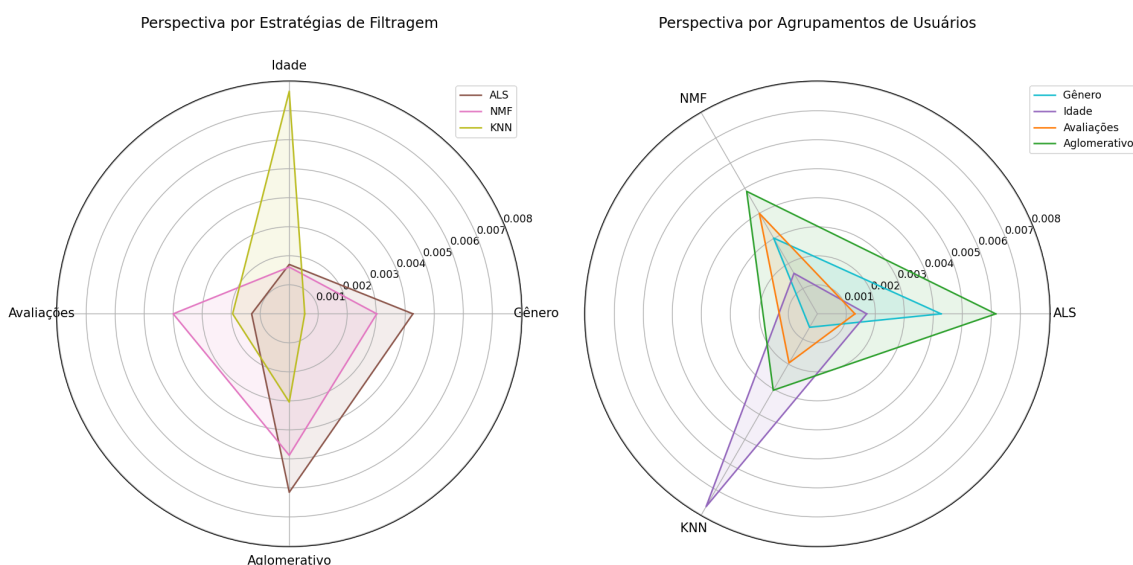


Figura 4. Comparação de injustiça por estratégia de filtragem

representação circular das estratégias, que permite identificar o equilíbrio ou desequilíbrio na distribuição da injustiça entre os agrupamentos, com destaque para a consistência das estratégias ALS e NMF em contraste com a variabilidade da KNN. Invertendo a lente no segundo gráfico, a influência dos agrupamentos sobre as estratégias é explorada, revelando como certos agrupamentos podem ser mais desafiadores em termos de manutenção da equidade. O agrupamento Aglomerativo resultou nas maiores injustiças de grupo, mantendo altos índices independente da estratégia. Já os agrupamentos por Gênero, Idade e Avaliação se mostraram mais sensíveis as estratégias de filtragem. Destaque para a Idade, com as maiores variações.

5. Conclusão

A análise dos resultados oferece insights valiosos sobre a interação entre estratégias de filtragem colaborativa e configurações de agrupamento. Primeiramente, a variabilidade na injustiça do grupo entre as estratégias de filtragem (ALS, KNN e NMF) ressalta a importância da seleção criteriosa da estratégia para promover equidade nas recomendações. Observa-se que diferentes estratégias apresentam desempenhos distintos em relação à injustiça do grupo, com o KNN mostrando uma resposta diferenciada em certos agrupamentos, como Gênero e Idade, destacando-se na análise comparativa.

Notavelmente, a abordagem de agrupamento utilizando o método Aglomerativo demonstrou os maiores níveis de injustiça do grupo em grande parte das estratégias de filtragem examinadas. Este achado sublinha a complexidade inerente aos sistemas de recomendação e a importância de uma avaliação multidimensional que considere as correlações entre todas as variáveis envolvidas.

Os resultados apontam para a necessidade de uma abordagem mais personalizada no design de sistemas de recomendação. A integração de considerações de justiça social desde a concepção até a implementação de tais sistemas é fundamental. Recomenda-se aos desenvolvedores e pesquisadores que explorem estratégias de filtragem e métodos de agrupamento que se alinhem aos objetivos de equidade, visando a criação de sistemas de

recomendação que não apenas sejam eficientes, mas também justos e inclusivos.

Essas conclusões ressaltam a complexidade inerente à busca por sistemas de recomendação equitativos e apontam para áreas promissoras de pesquisa futura. Um caminho a ser explorado é a realização de análises multivariadas para investigar as interações entre as diversas variáveis que caracterizam os usuários, contribuindo assim para o aprimoramento da justiça e inclusão nas plataformas digitais.

Referências

- Beutel, A., Chi, E. H., Cheng, Z., Pham, H., and Anderson, J. (2017). Beyond globally optimal: Focused learning for improved recommendations. In *Proceedings of the 26th International Conference on WWW 2017, Perth, Australia, April 3-7, 2017*.
- Bhargava, A., Bansal, A., and Goyal, V. (2022). Machine learning–based detection and sorting of multiple vegetables and fruits. *Food Analytical Methods*, 15:228–242.
- Burke, R., Sonboli, N., and Ordonez-Gauger, A. (2018). Balanced neighborhoods for multi-sided fairness in recommendation. In *FAT*.
- Celis, L. E., Deshpande, A., Kathuria, T., and Vishnoi, N. K. (2016). How to be fair and diverse? *CoRR*, abs/1610.07183.
- Coelho, N. L. L., Figueiredo, T. F., and Figueiredo, R. (2023). Uso de programação linear inteira para geração e análise de agrupamentos de políticos da câmara dos deputados. In *BraSNAM 2023 - XII Brazilian Workshop on Social Network Analysis and Mining*, João Pessoa, PB, Brasil. Sociedade Brasileira de Computação.
- Dargan, S., Kumar, M., and et al., A. M. (2024). Handwriting-based gender classification using machine learning techniques. *Multimedia Tools and Applications*, 83:19871–19895.
- Deldjoo, Y., Jannach, D., Bellogin, A., Difonzo, A., and Zanzonelli, D. (2024). Fairness in recommender systems: research landscape and future directions. *User Modeling and User-Adapted Interaction*, 34(1):59–108.
- Dwork, C., Hardt, M., Pitassi, T., Reingold, O., and Zemel, R. S. (2011). Fairness through awareness. *CoRR*, abs/1104.3913.
- Emirhüseyinoğlu, G., Shahhosseini, M., Hu, G., et al. (2023). Validation of scenario generation for decision-making using machine learning prediction models. *Optim Lett.*
- Feng, X., Cai, Y., and Xin, R. (2023). Optimizing diabetes classification with a machine learning-based framework. *BMC Bioinformatics*, 24:428.
- Harper, F. M. and Konstan, J. A. (2015). The movielens datasets: History and context. *ACM Transactions on Interactive Intelligent Systems (TiiS)*, 5(4):Article 19.
- Hazrati, N. and Ricci, F. (2024). Choice models and recommender systems effects on users’ choices. *User Model User-Adap Inter*, 34(1):109–145.
- Kamishima, T. and Akaho, S. (2017). Considerations on recommendation independence for a find-good-items task. In *Workshop on Responsible Recommendation*.
- Kamishima, T., Akaho, S., and Asoh, H. (2012). Enhancement of the neutrality in recommendation. In *In Proc. of the 2nd Workshop on Human Decision Making in Recommender Systems*, pages 8–14.

- Kamishima, T., Akaho, S., Asoh, H., and Sakuma, J. (2018). Recommendation independence. In Friedler, S. A. and Wilson, C., editors, *Proceedings of the 1st Conference on Fairness, Accountability and Transparency*, volume 81 of *Proceedings of Machine Learning Research*, pages 187–201. PMLR.
- Leonhardt, J., Anand, A., and Khosla, M. (2018). User fairness in recommender systems. *CoRR*, abs/1807.06349.
- Liu, S., Ge, Y., Xu, S., Zhang, Y., and Marian, A. (2022). Fairness-aware federated matrix factorization. In *Proceedings of the 16th ACM Conference on Recommender Systems*, RecSys '22, page 168–178, New York, NY, USA. Association for Computing Machinery.
- Liu, S., Sun, J., Deng, X., and et al. (2024). Towards platform profit-aware fairness in personalized recommendation. *Int. J. Mach. Learn. & Cyber*.
- Pereira, F. S. F., Linhares, C. D. G., Ponciano, J. R., Gama, J., de Amo, S., and Oliveira, G. M. B. (2018). That's my jam! uma análise temporal sobre a evolução das preferências dos usuários em uma rede social de músicas. In *BraSNAM 2018 - VII Brazilian Workshop on Social Network Analysis and Mining*, Porto Alegre, RS, Brasil. Sociedade Brasileira de Computação.
- Rastegarpanah, B., Gummadi, K. P., and Crovella, M. (2019). Fighting fire with fire: Using antidote data to improve polarization and fairness of recommender systems. In *Proceedings of the Twelfth ACM International Conference on Web Search and Data Mining*, WSDM '19. ACM.
- Sikora, A., Zhang, T., and et al., D. M. (2023). Machine learning vs. traditional regression analysis for fluid overload prediction in the icu. *Scientific Reports*, 13:19654.
- Souza, E., Lichtnow, D., and Gasparini, I. (2022). Estratégia de pós-processamento aplicada a um sistema de recomendação de artigos para a melhora da diversidade. In *BraSNAM 2022 - XI Brazilian Workshop on Social Network Analysis and Mining*, pages 216–221, Porto Alegre, RS, Brasil. Sociedade Brasileira de Computação.
- Sudakov, V. A. and Titov, Y. P. (2023). Pandemic forecasting by machine learning in a decision support problem. *Math Models Comput Simul*, 15:520–528.
- Vashistha, R., Noor, Z., and et al., S. D. (2023). Application of statistical machine learning in biomarker selection. *Scientific Reports*, 13:18331.
- Xue, J., Yao, W., and Xiang, S. (2024). Machine learning embedded em algorithms for semiparametric mixture regression models. *Computational Statistics*.
- Yao, S. and Huang, B. (2017). Beyond parity: Fairness objectives for collaborative filtering. *CoRR*, abs/1705.08804.
- Zafar, M., Valera, I., Rodriguez, M., and Gummadi, K. P. (2017). Fairness beyond disparate treatment & disparate impact: Learning classification without disparate mistreatment. In *WWW '17: Proceedings of the 26th International Conference on World Wide Web*, pages 1171–1180.
- Zehlike, M., Sühr, T., Baeza-Yates, R., Bonchi, F., Castillo, C., and Hajian, S. (2022). Fair top-k ranking with multiple protected groups. *Processing & Management*, 59(1):102707.