

Implementação de Medidas de Justiça Social: Um Estudo de Caso no Contexto Cultural para Diferentes Estratégias em Sistemas de Recomendação

Implementation of Fairness Measures: A Case Study in the Cultural Context for Different Strategies in Recommendation Systems

Rafael Vargas Mesquita dos Santos¹ , Giovanni Ventorim Comarela² , Márcio Cesar Guimarães de Lima Junior¹ 

¹Departamento de Informática – Instituto Federal do Espírito Santo (IFES)
Cachoeiro de Itapemirim – ES – Brasil

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

rafaelv@ifes.edu.br, gc@inf.ufes.br, marciojr028@gmail.com

Abstract. This study analyzes equity in recommendation systems, specifically in a musical context, focusing on understanding disparities across different filtering strategies. Through the implementation of equity metrics, we investigate individual and group injustices among various recommendation approaches. The results reveal significant variations in the fairness applied by the examined strategies, highlighting the complexity of achieving equity in these systems. We conclude that a detailed analysis of justice in recommendation systems is crucial for identifying and understanding existing disparities, contributing to the future development of more equitable solutions.

Keywords. Recommender Systems, Fairness, Individual Fairness, Group Fairness

Resumo. Este estudo analisa a equidade em sistemas de recomendação, especificamente em um contexto musical, focando na compreensão das disparidades em diferentes estratégias de filtragem. Através da implementação de métricas de equidade, investigamos as injustiças individuais e de grupo entre várias abordagens de recomendação. Os resultados revelam variações significativas na justiça aplicada pelas estratégias examinadas, evidenciando a complexidade de alcançar equidade nesses sistemas. Concluímos que a análise detalhada da justiça em sistemas de recomendação é fundamental para identificar e entender as disparidades existentes, contribuindo para o desenvolvimento futuro de soluções mais equitativas.

Palavras-Chave. Sistema de Recomendação, Justiça, Justiça Individual, Justiça do Grupo

1. Introdução

Sistemas de recomendação desempenham um papel fundamental nas modernas plataformas digitais, orientando a navegação dos usuários através de vastas quantidades de conteúdo e influenciando uma ampla gama de escolhas cotidianas, desde seleções de entretenimento até decisões de compra. À medida que a ubiquidade desses sistemas se expande, torna-se imperativo considerar os efeitos sociais adversos que podem surgir, muitas vezes de maneira inadvertida, devido à natureza automatizada e aos algoritmos subjacentes que governam suas operações.

Pesquisas na vanguarda deste campo, como as conduzidas por Beutel et al. [Beutel et al. 2017], têm iluminado a problemática de como os sistemas de recomendação podem inadvertidamente perpetuar ou até exacerbar desequilíbrios sociais e preconceitos existentes. Essas investigações apontam para uma preocupante tendência de tais sistemas em fornecer uma qualidade de serviço variável, frequentemente em detrimento de certos usuários ou grupos sociais, destacando uma questão crítica de equidade na distribuição das recomendações.

Confrontados com a crescente dependência societal em sistemas de recomendação e reconhecendo seus potenciais impactos negativos, emerge como imperativo a busca por estratégias que mitiguem esses efeitos, promovendo assim um ecossistema digital mais justo e inclusivo. Este reconhecimento impulsiona a necessidade de repensar métricas tradicionais de sucesso de tais sistemas, tais como a precisão das sugestões, para incorporar considerações de equidade e justiça social.

A precisão, embora seja um indicador valioso do desempenho técnico de um algoritmo de recomendação, não captura completamente o espectro de impactos que tais sistemas podem ter sobre a sociedade. Assim, a equidade se apresenta como uma preocupação fundamental, destacando a importância de abordar as consequências desproporcionais das recomendações sobre grupos específicos.

Este trabalho se dedica à implementação e análise de medidas específicas de justiça, tanto em nível individual quanto de grupo. Essas medidas serãometiculosa-mente aplicadas em um estudo de caso culturalmente contextualizado, envolvendo um conjunto de dados composto por 59 usuários e 10 músicas, resultando em um total de 590 avaliações. Este contexto específico serve como um laboratório para investigar as nuances e o comportamento das medidas de justiça de grupo quando aplicadas em conjunto com diferentes estratégias de filtragem em sistemas de recomendação.

O objetivo primário deste estudo é analisar e compreender o comportamento dessas métricas de justiça - particularmente a justiça de grupo - sob o prisma de várias estratégias de recomendação, incluindo filtragem colaborativa, baseada em conteúdo, e abordagens híbridas. Através desta análise, estudamos como diferentes abordagens de recomendação influenciam a equidade das sugestões fornecidas aos usuários, especialmente em um cenário que reflete diversidade cultural e preferências musicais. O intuito é contribuir para o campo de sistemas de recomendação, fornecendo evidências empíricas e orientações práticas para a concepção de sistemas mais justos e equitativos, que considerem adequadamente tanto a justiça individual quanto a justiça de grupo em suas recomendações.

2. Trabalhos Relacionados

A equidade tem emergido como uma preocupação primordial no campo do aprendizado de máquina, refletindo o reconhecimento crescente da importância de sistemas justos em diversos contextos sociais e econômicos. [Zafar et al. 2017] destacam que um sistema de recomendação é considerado justo quando proporciona uma qualidade de serviço consistente e imparcial a todos os usuários ou grupos de usuários. Esta perspectiva reforça a importância de desenvolver algoritmos que não apenas sejam eficientes e precisos, mas também promovam a equidade e a inclusão.

No âmbito do aprendizado de máquina e sistemas de recomendação, a conscientização acerca dos potenciais prejuízos sociais derivados da implementação indiscriminada de algoritmos em processos decisórios tem sido intensificada, conforme discutido por [Barocas and Selbst 2016]. A resposta da comunidade científica incluiu a proposição de várias noções e métricas de equidade aplicáveis em diversos aspectos do aprendizado de máquina, incluindo classificação [Zafar et al. 2017], regressão [Berk et al. 2017], e ordenação [Zehlike et al. 2022]. Estas iniciativas visam garantir que os sistemas automatizados atuem de maneira justa, sem perpetuar ou amplificar desigualdades existentes.

Existem, primordialmente, duas abordagens para avaliar a equidade em sistemas algorítmicos: uma focada na equidade a nível individual e outra centrada na equidade a nível de grupo [Dwork et al. 2012]. A equidade a nível individual procura garantir que decisões similares sejam tomadas para indivíduos similares, independentemente de suas características pessoais. Em contrapartida, a equidade a nível de grupo visa assegurar que diferentes grupos demográficos, definidos por atributos como gênero ou etnia, recebam tratamento equitativo.

Embora as questões de equidade tenham sido amplamente estudadas nas áreas de classificação e regressão, a investigação de equidade em sistemas de recomendação é relativamente incipiente. [Kamishima et al. 2018] foram pioneiros ao explorar conceitos de equidade a nível de grupo em sistemas de recomendação, propondo métodos para minimizar a influência de atributos sensíveis na sugestão de itens. [Beutel et al. 2017] estabeleceram critérios de equidade a nível de grupo para sistemas de recomendação, focando na acurácia das previsões para diversos conjuntos de usuários ou itens, e destacando a importância de considerar a diversidade e a representatividade na avaliação de sistemas.

Este trabalho visa contribuir para o corpo de conhecimento existente não apenas adotando conceitos de equidade a nível de grupo, conforme sugerido por Beutel et al., mas também expandindo para considerar questões de equidade a nível individual. Ao fazê-lo, buscamos fornecer uma compreensão mais aprofundada e holística da justiça em sistemas de recomendação. Nossa abordagem inclui a implementação e avaliação de métricas de equidade específicas, adaptadas para o contexto de recomendação, e a análise dos impactos de diferentes estratégias de mitigação de viés na qualidade e justiça das recomendações fornecidas. Ao integrar considerações de equidade tanto a nível de grupo quanto individual, este estudo visa iluminar caminhos para o desenvolvimento de sistemas de recomendação mais justos e inclusivos, contribuindo assim para um uso mais ético e responsável da tecnologia na sociedade.

3. Material e Métodos

3.1. Base de Dados

O estudo de caso utilizou um conjunto de dados no contexto cultural¹, que contém 590 classificações de 10 músicas realizadas por 59 usuários, com avaliações em uma escala de 5 pontos [Harper and Konstan 2015].

Para as estimativas das classificações desconhecidas foram testadas cinco diferentes estratégias em sistemas de recomendação: filtragem colaborativa, filtragem baseada em conteúdo, filtragem híbrida ponderada, filtragem híbrida mista e filtragem híbrida de combinação sequencial.

Realizamos ainda, após as previsões calculadas pelos algoritmos de recomendação, um agrupamento aleatório dos usuários (50-50). Neste agrupamento, separamos cada grupo com a metade dos usuários.

Utilizamos o módulo do algoritmo apresentado na seção 3.2 para calcular as medidas de justiça social para o estudo de caso proposto.

Por fim, foram apresentados os resultados de medidas de justiça do grupo comparando cinco diferentes estratégias de recomendações.

3.2. Módulo do Algoritmo: cálculo de medidas sociais

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 $X \in \mathbb{R}^{n \times d}$ seja uma matriz de classificação parcialmente observada de n usuários e d 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 conhecidas 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 de usuário i , como a estimativa do erro quadrático médio sobre as classificações conhecidas do usuário i :

¹<https://github.com/ravarmes/srh-backend-fairness/blob/main/docs/fairness/Rpol-Rindy-Rgrp-DataSet-590Avaliacoes.xlsx>

$$\ell_i = \frac{\|P_{\Omega^i}(\hat{x}^i - x^i)\|_2^2}{|\Omega^i|} \quad (1)$$

Definiremos a injustiça individual como a variação das perdas do usuário:

$$R_{indv}(X, \hat{X}) = \frac{1}{n^2} \sum_{k=1}^n \sum_{l>k} (\ell_k - \ell_l)^2 \quad (2)$$

Justiça do grupo. Sendo I o conjunto de todos os usuários/itens e $G = \{G_1, G_2, \dots, G_g\}$ seja uma partição de usuários/itens dentro de g grupos, ou seja, $I = \bigcup_{i \in \{1, 2, \dots, g\}} G_i$. Definiremos a perda do grupo como a estimativa do erro quadrático médio sobre todas as classificações conhecidas no grupo i :

$$L_i = \frac{\|P_{\Omega_{G_i}}(\hat{X} - X)\|_2^2}{|\Omega_{G_i}|} \quad (3)$$

Para determinada partição G , a injustiça do grupo será a variação de todas as perdas do grupo:

$$R_{grp}(X, \hat{X}, G) = \frac{1}{g^2} \sum_{k=1}^g \sum_{l>k} (L_k - L_l)^2 \quad (4)$$

4. Resultados e Discussões

Nesta seção, mostramos um comparativo das medições de justiça individual e do grupo em diferentes estratégias de recomendação.

4.1. Configurações experimentais

Delineamos as configurações experimentais adotadas para investigar a equidade em sistemas de recomendação. A metodologia empregada é dividida em duas partes fundamentais: Agrupamento dos usuários e Estratégias de recomendação. Inicialmente, detalhamos como os usuários são agrupados com base em atributos sensíveis e preferências, visando explorar o impacto dessas divisões na avaliação da equidade das recomendações. Esta abordagem permite uma análise profunda de como diferentes grupos de usuários experimentam o sistema de recomendação, lançando luz sobre disparidades potenciais no serviço provido. Em seguida, examinamos as estratégias de recomendação implementadas para mitigar viéses e promover uma distribuição mais justa de recomendações. Especificamente, descrevemos os algoritmos utilizados, os critérios para avaliar a qualidade e a equidade das recomendações, e como essas estratégias são ajustadas para atender a diversos objetivos de equidade. Ao integrar essas duas dimensões, nosso estudo oferece insights valiosos sobre o design e a implementação de sistemas de recomendação que não apenas melhoram a experiência do usuário, mas também promovem a justiça e a inclusão.

4.1.1. Agrupamento dos usuários

No processo de configuração do experimento, foi essencial estabelecer grupos de usuários distintos para avaliar a eficácia das estratégias de recomendação sob diferentes condições. Para isso, os usuários participantes do estudo foram cuidadosamente divididos em dois grupos distintos, seguindo um critério de aleatoriedade. Essa divisão foi realizada de forma a garantir que cada grupo fosse composto por metade dos usuários, promovendo uma distribuição equitativa e minimizando qualquer viés pré-existente que pudesse influenciar os resultados do experimento.

A escolha por uma separação aleatória dos usuários visa replicar um cenário realista em que a distribuição de características, preferências e comportamentos dos usuários em sistemas de recomendação ocorre de maneira não uniforme e imprevisível. Ao adotar essa abordagem, pretendemos assegurar que os resultados obtidos refletem de forma mais fidedigna o desempenho das estratégias de recomendação em condições variadas, proporcionando uma base sólida para a análise comparativa entre os diferentes métodos aplicados.

Além disso, a divisão aleatória permite uma investigação mais precisa sobre o impacto das estratégias de recomendação em grupos heterogêneos de usuários. Essa heterogeneidade é crucial para compreender como diferentes perfis de usuários respondem às recomendações, permitindo a identificação de padrões de eficácia ou potenciais áreas de melhoria nas estratégias aplicadas. Dessa forma, a estruturação dos grupos de usuários constitui um passo fundamental para o sucesso do experimento, fornecendo um ambiente controlado e equilibrado para a avaliação das técnicas de recomendação.

4.1.2. Estratégias de recomendação

Foram exploradas cinco estratégias distintas de recomendação, cada uma representando um paradigma diferenciado na geração de sugestões personalizadas para os usuários:

- **Filtragem Colaborativa:** Esta abordagem baseia-se na premissa de que usuários com preferências similares no passado tendem a ter preferências similares no futuro. A filtragem colaborativa utiliza essas correlações entre usuários para recomendar itens, supondo que um item apreciado por um usuário será de interesse de outro usuário com um perfil semelhante [Su and Khoshgoftaar 2009].
- **Filtragem Baseada em Conteúdo:** Diferente da filtragem colaborativa, esta estratégia foca nas características dos itens para realizar recomendações. Utiliza um perfil de preferências do usuário, construído a partir das características dos itens com os quais o usuário interagiu previamente, para sugerir novos itens com atributos semelhantes [Lops et al. 2011].
- **Filtragem Híbrida Ponderada:** Esta estratégia busca combinar as forças tanto da filtragem colaborativa quanto da filtragem baseada em conteúdo, atribuindo pesos relativos a cada método para refletir sua importância na recomendação final. A ideia é aproveitar as vantagens complementares de ambos os métodos para proporcionar recomendações mais precisas e diversificadas [Burke 2002].

- **Filtragem Híbrida Mista:** Nesta abordagem, as recomendações geradas pelos métodos de filtragem colaborativa e baseada em conteúdo são integradas em uma única lista de recomendações. Diferentemente da filtragem híbrida ponderada, não se atribui pesos específicos a cada método, mas busca-se uma combinação equilibrada das sugestões de ambos [Burke 2002].
- **Filtragem Híbrida de Combinação Sequencial:** Esta estratégia aplica os métodos de filtragem em uma sequência definida, onde a saída de um método serve como entrada para o próximo. Esse encadeamento busca aproveitar as peculiaridades de cada método de forma sequencial para refinar progressivamente as recomendações [Burke 2007].

A diversidade dessas abordagens reflete a complexidade do problema de recomendação e a necessidade de estratégias sofisticadas para abordar diferentes aspectos da personalização e precisão nas recomendações.

4.2. Resultados dos experimentos

Os resultados obtidos através dos experimentos conduzidos são fundamentais para o entendimento do impacto das estratégias de recomendação examinadas neste estudo. A análise desses resultados, detalhadamente apresentada na Tabela 1, fornece insights valiosos sobre a eficácia das diversas abordagens de filtragem sob a ótica das métricas de justiça individual e de grupo implementadas. Esses achados são cruciais para avaliar não apenas o desempenho técnico das estratégias de recomendação em termos de precisão e relevância, mas também sua capacidade de promover uma distribuição equitativa das recomendações entre os usuários.

Ao explorar os resultados, buscamos identificar padrões significativos e tendências que possam emergir das diferentes configurações experimentais. Isso inclui uma investigação detalhada sobre como as variáveis controladas, como o tipo de estratégia de filtragem utilizada, afetam os indicadores de justiça no contexto do nosso estudo de caso cultural. A compreensão dessas dinâmicas é essencial para propor melhorias nos sistemas de recomendação, visando otimizar tanto a satisfação do usuário quanto a equidade nas recomendações fornecidas.

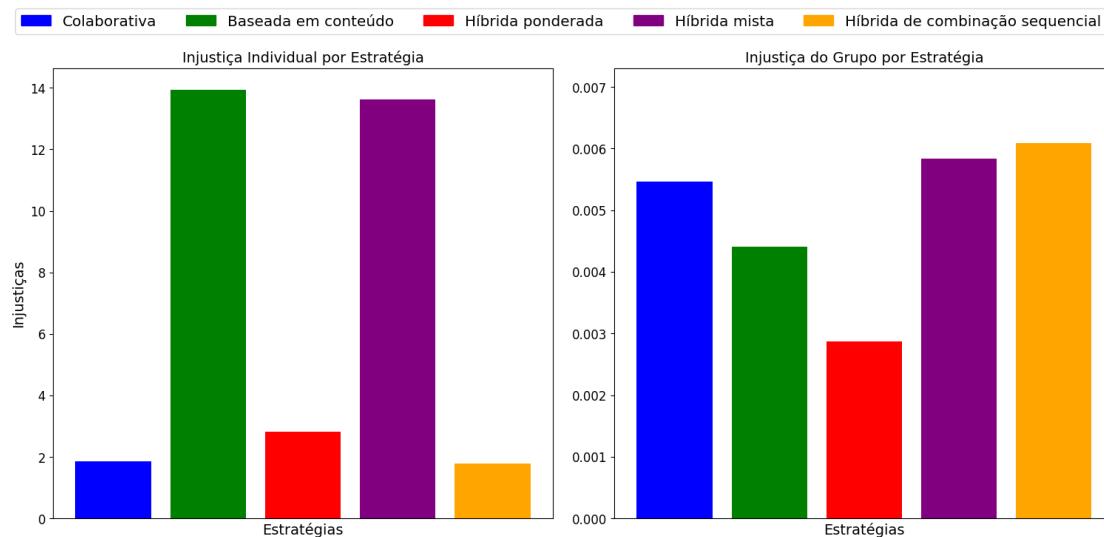
Adicionalmente, os resultados destacam as complexidades inerentes à implementação de medidas de equidade em sistemas de recomendação. Eles servem como um ponto de partida para discussões mais profundas sobre os desafios e as oportunidades na busca por um equilíbrio entre acurácia e justiça, contribuindo assim para o avanço da pesquisa na área e o desenvolvimento de práticas mais inclusivas e justas em tecnologias de recomendação.

Acompanhe a Tabela 1, que resume os resultados obtidos, seguida de uma análise detalhada dos mesmos, discutindo suas implicações para o design e a avaliação de sistemas de recomendação mais equitativos e eficazes.

A Figura 1 apresenta uma comparação entre a injustiça individual e a injustiça do grupo para diferentes estratégias de filtragem, destacando as variações nas métricas de equidade. A escala diferenciada no eixo Y para a injustiça do grupo permite uma visualização mais detalhada das diferenças entre as estratégias.

Tabela 1. R_{indv} e R_{grp} considerando o agrupamento aleatório (50-50)

Estratégia	R_{indv}	R_{grp}
Filtragem colaborativa	1,84960	0,00547
Filtragem baseada em conteúdo	13,94690	0,00441
Filtragem híbrida ponderada	2,80900	0,00287
Filtragem híbrida mista	13,61730	0,00584
Filtragem híbrida de combinação sequencial	1,78670	0,00609

**Figura 1.** Análise das medidas de injustiças nas diferentes estratégias

4.2.1. Análise geral dos resultados

A partir dos dados apresentados na Tabela 1, podemos extrair algumas conclusões importantes sobre a injustiça individual (R_{indv}) e a injustiça do grupo (R_{grp}) em diferentes estratégias de filtragem, considerando um agrupamento aleatório (50-50).

1. **Filtragem colaborativa** apresentou um valor moderado de injustiça individual (1,84960) e a menor injustiça de grupo (0,00547) entre as estratégias testadas, sugerindo um equilíbrio razoável entre tratar os indivíduos de forma justa e manter a equidade entre os grupos.
2. **Filtragem baseada em conteúdo** mostrou a maior injustiça individual (13,94690), mas também uma injustiça de grupo relativamente baixa (0,00441). Isso indica que, enquanto a abordagem pode ser mais tendenciosa em nível individual, ela mantém uma consistência de tratamento entre grupos.
3. **Filtragem híbrida ponderada** revelou uma injustiça individual significativamente menor (2,80900) comparada à filtragem baseada em conteúdo, com a menor injustiça de grupo (0,00287) dentre todas as estratégias. Essa abordagem parece oferecer um melhor equilíbrio entre justiça individual e de grupo.
4. **Filtragem híbrida mista** registrou uma alta injustiça individual (13,61730) similar à filtragem baseada em conteúdo, mas com a maior injustiça de grupo

(0,00584), indicando que essa estratégia pode não ser ideal para promover equidade tanto em nível individual quanto de grupo.

5. **Filtragem híbrida de combinação sequencial** teve uma injustiça individual (1,78670) comparável à filtragem colaborativa e uma injustiça de grupo (0,00609) ligeiramente maior. Isso sugere que, apesar de ser justa no nível individual, pode introduzir ligeiras discrepâncias na equidade entre grupos.

4.2.2. Análise da injustiça individual versus injustiça do grupo

Uma estratégia de filtragem em sistemas de recomendação que apresenta um valor alto para a injustiça individual (R_{indv}) e baixo para a injustiça do grupo (R_{grp}), como nas filtragens baseada em conteúdo e híbrida mista, sugere um cenário que reflete os desafios envolvidos na personalização de recomendações para usuários individuais versus a otimização para o bem-estar coletivo de um grupo.

Um valor alto para injustiça individual (R_{indv}) indica que há uma variação significativa nas perdas experimentadas pelos usuários individuais. Isso pode ocorrer por várias razões:

- Diversidade de preferências: Os usuários podem ter gostos e preferências altamente diversificados;
- Viés de dados: Caso o algoritmo de recomendação seja treinado em um conjunto de dados que não representa adequadamente a diversidade do conjunto de usuários;
- Limitações do modelo: Algumas estratégias de filtragem podem ser inherentemente melhores em capturar as preferências de certos tipos de usuários.

Um valor baixo para injustiça do grupo (R_{grp}) sugere que o sistema de recomendação é eficaz em minimizar o erro quadrático médio sobre todas as classificações conhecidas no grupo. Esse cenário pode surgir nas seguintes condições:

- Compensação de erros: Os erros nas recomendações para alguns usuários podem ser compensados por recomendações precisas para outros usuários;
- Foco em preferências comuns: O sistema pode ser eficaz em identificar e recomendar itens que são geralmente populares ou bem aceitos pela maioria dos usuários.

5. Conclusão

Em suma, essas conclusões apontam para a complexidade de equilibrar a justiça em níveis individual e de grupo em sistemas de recomendação, com a filtragem híbrida ponderada emergindo como a estratégia mais equilibrada entre as avaliadas. Estes resultados destacam a importância de considerar múltiplas dimensões de justiça ao projetar e avaliar sistemas de recomendação.

A alta injustiça individual combinada com a baixa injustiça do grupo, nas filtragens baseada em conteúdo e híbrida mista, sugere que, embora o sistema de recomendação atenda bem às necessidades gerais do grupo, falha em proporcionar experiências personalizadas de alta qualidade para todos os usuários individualmente. Isso

destaca o desafio de equilibrar a personalização individual com a otimização para o coletivo.

6. Agradecimentos

Os autores gostariam de expressar sua profunda gratidão ao Instituto Federal do Espírito Santo (Ifes) pelo aporte financeiro e apoio contínuo ao longo do desenvolvimento deste trabalho. Este suporte foi fundamental para a realização da pesquisa e os resultados alcançados.

Referências

- [Barocas and Selbst 2016] Barocas, S. and Selbst, A. D. (2016). Big data's disparate impact. *California Law Review*, 104(3):671–732.
- [Berk et al. 2017] Berk, R., Heidari, H., Jabbari, S., Joseph, M., Kearns, M. J., Morgenstern, J., Neel, S., and Roth, A. (2017). A convex framework for fair regression. *CoRR*, abs/1706.02409.
- [Beutel et al. 2017] 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*.
- [Burke 2002] Burke, R. (2002). Hybrid recommender systems: Survey and experiments. *User modeling and user-adapted interaction*, 12(4):331–370.
- [Burke 2007] Burke, R. (2007). Hybrid web recommender systems. In *The adaptive web*, pages 377–408. Springer.
- [Dwork et al. 2012] Dwork, C., Hardt, M., Pitassi, T., Reingold, O., and Zemel, R. (2012). Fairness through awareness. In *Proceedings of the 3rd Innovations in Theoretical Computer Science Conference*, ITCS '12, page 214–226, New York, NY, USA. Association for Computing Machinery.
- [Harper and Konstan 2015] Harper, F. M. and Konstan, J. A. (2015). The movielens datasets: History and context. *Acm transactions on interactive intelligent systems (tiis)*, 5(4):1–19.
- [Kamishima et al. 2018] 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.
- [Lops et al. 2011] Lops, P., de Gemmis, M., and Semeraro, G. (2011). Content-based recommender systems: State of the art and trends. *Recommender systems handbook*, pages 73–105.
- [Su and Khoshgoftaar 2009] Su, X. and Khoshgoftaar, T. M. (2009). A survey of collaborative filtering techniques. *Advances in artificial intelligence*, 2009.

- [Zafar et al. 2017] Zafar, M. B., Valera, I., Gomez Rodriguez, M., and Gummadi, K. P. (2017). Fairness beyond disparate treatment & disparate impact: Learning classification without disparate mistreatment. In *Proceedings of the 26th International Conference on World Wide Web*, WWW '17, page 1171–1180, Republic and Canton of Geneva, CHE. International World Wide Web Conferences Steering Committee.
- [Zehlike et al. 2022] 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.