PONTIFÍCIA UNIVERSIDADE CATÓLICA DE MINAS GERAIS

Programa de Pós-Graduação em Informática

Gabriel Bálbio Vieira Machado

EQNET– UMA ABORDAGEM PARA REDUZIR O VIÉS DE POPULARIDADE EM SISTEMAS DE RECOMENDAÇÃO DE FILTRAGEM COLABORATIVA

Belo Horizonte

2024

Gabriel Bálbio Vieira Machado

EQNET– UMA ABORDAGEM PARA REDUZIR O VIÉS DE POPULARIDADE EM SISTEMAS DE RECOMENDAÇÃO DE FILTRAGEM COLABORATIVA

Dissertação apresentada ao Programa de Pós-Graduação em Informática da Pontifícia Universidade Católica de Minas Gerais, como requisito parcial para obtenção do título de Mestre em Informática.

Orientador: Prof. Humberto Torres Marques Neto

Belo Horizonte

2024

FICHA CATALOGRÁFICA
Elaborada pela Biblioteca da Pontifícia Universidade Católica de Minas Gerais

Machado, Gabriel Bálbio Vieira
M149e EQNET: uma abordagem para reduzir o viés de popularidade em sistemas de recomendação de filtragem colaborativa / Gabriel Bálbio Vieira Machado. Belo Horizonte, 2024.
87 f. : il.

Orientador: Humberto Torres Marques Neto
Dissertação (Mestrado) – Pontifícia Universidade Católica de Minas Gerais. Programa de Pós-Graduação em Informática

1. Sistemas de recomendação (Filtragem da informação). 2. Redes complexas. 3. Interação homem-máquina. 4. Aprendizado do computador. 5. Algoritmos. 6. Recuperação da informação. I. Marques Neto, Humberto Torres. II. Pontifícia Universidade Católica de Minas Gerais. Programa de Pós-Graduação em Informática. III. Título.

CDU: 681.3.091

Ficha catalográfica elaborada por Fabiana Marques de Souza e Silva - CRB 6/2086

Gabriel Bálbio Vieira Machado

EQNET- UMA APLICAÇÃO MODULAR PARA REDUZIR O VIÉS DE POPULARIDADE EM SISTEMAS DE RECOMENDAÇÃO DE FILTRAGEM COLABORATIVA

Dissertação apresentada ao Programa de Pós-Graduação em Informatica como requisito parcial para qualificacao ao Grau de Mestre em Informatica pela Pontificia Universidade Catolica de Minas Gerais.

Prof. Dr. Humberto Torres Marques Neto –
PUC Minas

Prof. Dr. Wladmir Cardoso Brandao –
PUC Minas

Prof. Dr. Adriano César Machado Pereira –
UFMG

Belo Horizonte, 04/07/2024.

AGRADECIMENTOS

Gostaria de agradecer em primeiro lugar à minha família, em especial meus pais, Eduardo Souza Vieira Machado e Mércia de Meira Bálbio Vieira Machado, que me proporcionaram um lar, carinho e muito amor para que pudesse estudar e me aprimorar e ao meu orientador Prof. Dr. Humberto Torres Marques Neto, por me incentivar a iniciar esta jornada e acreditar em meu potencial. À minha irmã Júnia Vieira Machado (Juju), com quem sempre pude contar, e às minhas sobrinhas Júlia e Mariana, que fazem com que eu sempre busque ser o melhor de mim. À minha namorada Luísa Ribeiro que me incentivou muito, principalmente nos momentos finais do mestrado, fez com que eu acreditasse mais em mim e deixou essa jornada um pouco mais leve, assim como todos os meus amigos que fizeram parte direta ou indiretamente dessa jornada. E ao meu irmão Alexandre Vieira Machado, cuja ausência física é profundamente sentida, que sempre me estimulou a continuar a evoluir na minha vida acadêmica.

Muito obrigado!

"O êxito da vida não se mede pelo caminho que você conquistou, mas sim pelas dificuldades que superou no caminho."

Abraham Lincoln

RESUMO

Sistemas de recomendação desempenham um papel fundamental em plataformas digitais, facilitando usuários a ter experiências inovadoras ao ordenar e apresentar de maneira eficaz itens que se alinham com suas preferências. No entanto, esses sistemas frequentemente sofrem com o viés de popularidade, um fenômeno caracterizado pela inclinação do algoritmo em favorecer poucos itens com grande popularidade, resultando na sub-representação da grande maioria dos itens. Abordar esse viés e aprimorar a recomendação de itens de cauda longa é de suma importância. Neste trabalho, é proposto o EQNet, uma abordagem de reclassificação projetada para mitigar o viés de popularidade e melhorar a qualidade de um sistema de recomendação baseado em SVD. O EQNet utiliza as saídas do PageRank ou Contagem de Popularidade para reclassificar os itens, e sua eficácia é avaliada usando quatro métricas: popularidade média, porcentagem de itens de cauda longa, cobertura de itens de cauda longa e qualidade da recomendação. Para estabelecer uma base sólida, incorporamos o algoritmo de mitigação de viés amplamente reconhecido, o FA*IR, em nossos experimentos. Ao comparar o desempenho do EQNet com essa abordagem de ponta, é demonstrada a sua eficiência e é destacado o seu potencial para aprimorar métodos existentes de mitigação de viés de popularidade.

Palavras-chave: Sistemas de Recomendação, Baseados em Modelo, Filtro Colaborativo, Redes Complexas.

ABSTRACT

Recommendation systems play a pivotal role in digital platforms, facilitating novel user experiences by effectively sorting and presenting items that align with their preferences. However, these systems often suffer from popularity bias, a phenomenon characterized by the algorithm's inclination to favor a few popular items, resulting in the underrepresentation of the vast majority of items. Addressing this bias and enhancing the recommendation of long-tail items is of utmost importance. In this paper, we propose the EQNet, a re-ranking approach designed to mitigate popularity bias and improve the recommendation quality of an SVD-based recommendation system. EQNet leverages PageRank or Popularity Count outputs to re-rank items, and its effectiveness is evaluated using four metrics: average popularity, percentage of long-tailed items, coverage of long-tailed items, and recommendation quality. To establish a robust baseline, we incorporate the widely recognized bias mitigation algorithm FA*IR into our experimentation. By comparing the performance of EQNet against this state-of-the-art approach, we show the efficiency of EQNet and highlight its potential to enhance existing methods for mitigating popularity bias.

Keywords: Recommendation Systems, Collaborative Filtering, Model Based, Complex Network.

LISTA DE FIGURAS

LISTA DE TABELAS

SUMÁRIO

1 INTRODUÇÃO

Os sistemas de recomendação desempenham um papel crucial em plataformas virtuais, incluindo redes sociais, *marketplaces*, serviços de *streaming* e outras aplicações em rede. Reconhecida pela comunidade científica há mais de duas décadas, sua importância é inegável (HERLOCKER, 1999; RICCI, 2015; STEFANIDIS, 2021). Tal relevância advém da profusão de dados acessíveis atualmente e do valor agregado que esses dados proporcionam quando explorados com objetivos específicos, seja para análises mercadológicas, preditivas ou informativas (LINDEN, 2003). Como resultado, uma ampla gama de algoritmos, modelos e técnicas de processamento de dados foi desenvolvida para resolver os desafios recorrentes nos sistemas de recomendação, sempre com o intuito de aprimorar a qualidade das recomendações e a eficiência dos sistemas (SCHAFER, 2007; JALILI, 2018).

Os métodos de recomendação são normalmente categorizados com base nas características dos itens avaliados, incluindo abordagens item-item, usuário-usuário ou usuário-item e, nos modelos de filtragem empregados, como modelos de filtragem baseados em conteúdo, filtros colaborativos ou modelos de filtragem hibrida. Categorizando ainda mais a fundo, os modelos de filtragem colaborativa podem ser subdivididos de acordo com os métodos de predição adotados, como baseados em memória ou em modelos (JANNACH, 2010; Lü, 2011; RICCI, 2015; BOBADILLA, 2013). Cada uma dessas abordagens apresenta aplicações específicas e diferentes estratégias de otimização, destacando a necessidade de uma compreensão aprofundada das nuances de cada modelo para sua aplicação efetiva e eficiente em diferentes contextos computacionais. Essa diversidade de abordagens sugere a importância de uma análise cuidadosa para determinar a adequação de cada método às necessidades específicas de um determinado problema de recomendação, bem como a implementação de técnicas de otimização adequadas para melhorar o desempenho e a precisão dos sistemas de recomendação (AI, 2017; LINDEN, 2003).

Os algoritmos de filtragem colaborativa baseados em modelo desempenham um importante papel nas plataformas de grande escala devido à sua habilidade em lidar com volumes substanciais de dados de maneira eficiente, permitindo predições precisas mesmo em conjuntos de dados em constante crescimento (RANJBAR, 2015; YAO, 2014). No entanto, é essencial abordar cuidadosamente certos pontos de otimização, especialmente ao considerar os dois principais desafios enfrentados por esses algoritmos de recomendação: o problema de partida a frio e o viés de popularidade (SU; KHOSHGOFTAAR, 2009; ABDOLLAHPOURI, 2020).

Com redes informacionais cada vez maiores e mais complexas impulsionadas por sistemas de recomendação, otimizar o desempenho dos sistemas diante desses desafios emerge como um diferencial crucial tanto no âmbito mercadológico quanto acadêmico (ANDERSON, 2005; CASTELLS; VARGAS; WANG, 2011; TAYLOR, 2023). Esta evolução é evidenciada na sofisticação das estratégias de recomendação adotadas por plataformas líderes, como a Netflix, que emprega uma abordagem híbrida, combinando filtragem colaborativa e baseada em conteúdo para mitigar desafios como o problema da partida a frio. Embora muitos usuários possam não estar conscientes desses ajustes, a disparidade resultante nas recomendações pode levar a uma percepção de baixa qualidade no serviço (KOREN, 2009; YAO; HUANG, 2017; ABDOLLAHPOURI, 2020; RAHMAN, 2020). Além disso, estudos indicam que usuários frequentes tendem a valorizar ainda mais suas experiências de descoberta, buscando ativamente itens nichados que correspondam às suas preferências individuais e peculiaridades de perfil (BEDI, 2014; NEMATZADEH et al., 2018; YALCIN; BILGE, 2022). Essas considerações destacam a importância crítica de aprimorar a personalização e a diversidade das recomendações em sistemas de recomendação, visando a satisfação contínua e a fidelização do usuário.

Devido à sua importância, a mitigação do viés de popularidade em sistemas de recomendação é amplamente discutida na literatura. Diversas abordagens têm sido propostas para enfrentar esse desafio, cada uma com seus próprios méritos e focos específicos. Entre as técnicas recomendadas estão a reordenação das listas ou itens do portfólio (ADOMAVICIUS; KWON, 2009), modelos que ponderam a nota de recomendação dos itens com base na proximidade do seu *cluster* com outros itens (AI, 2020), ou ainda aqueles que consideram vínculos por informações implícitas (KISWANTO; NURJANAH; RISMALA, 2018). Embora variadas, todas essas abordagens compartilham o objetivo comum de buscar equilíbrio entre a diversidade das recomendações e a manutenção da qualidade. Nesse contexto, identificou-se uma oportunidade para explorar os valores de saída de algoritmos clássicos da literatura, como o PageRank (GOOGLE TEAM, 2011) e o Popularity Count (AWS Team, 2022). Estes algoritmos são conhecidos por sua aficiência em ranquear itens populares em uma rede e, portanto, foram escolhidos para ponderar os itens populares de uma lista de recomendação de maneira inversamente proporcional à sua nota.

Abordar o viés de popularidade é um desafio interessante, pois lida com o delicado equilíbrio entre popularidade e relevância. Remover recomendações impulsionadas pela popularidade pode levar a uma perda de qualidade, já que itens populares geralmente se alinham com as preferências do usuário (JANNACH, 2015; KOWALD; LACIC, 2022). Neste contexto, buscamos uma abordagem que permitisse mitigar o viés de popularidade sem comprometer a qualidade das recomendações. Para tanto, exploramos formas de identificar e reordenar os itens populares de acordo com critérios que consideram tanto a popularidade quanto a distância em relação aos interesses do usuário. Com base nessa

premissa, apresentamos o EQNet, uma abordagem de pós-processamento que aproveita o poder dos algoritmos de classificação para reclassificar listas de recomendação, garantindo que itens sub-representados e de alta qualidade recebam a atenção que merecem, sem comprometer a qualidade das recomendações.

1.1 Objetivos

Este trabalho tem como objetivo geral propor e avaliar uma abordagem para otimização de recomendações geradas por filtro colaborativo, visando reduzir o seu viés de popularidade sem impactar muito negativamente na qualidade das recomendações. A abordagem recebe informações intrínsecas de classificação dos itens com a finalidade de reordenar seus itens e atenuar o viés de popularidade. Para avaliar esta abordagem, foram realizados experimentos com um *dataset*s reais da Netflix e outro do MovieLens, comparando a performance do EQNet com outra abordagem de *baseline*.

Este estudo alcançou os seguintes objetivos específicos:

1. Desenvolver uma aplicação capaz de utilizar informações de ranqueamento de itens em um portfólio, visando estabelecer um processo sistemático de reordenação desses itens. Para isso, foi empregada uma abordagem baseada no algoritmo de Decomposição em Valores Singulares (SVD) em duas bases de dados distintas (Netflix e MovieLens), para gerar um sitema de recomendação de filmes.

2. Avaliar a performance da aplicação desenvolvida em comparação a um *baseline* estabelecido pela literatura. Esse processo de avaliação foi conduzido por meio de experimentos controlados, nos quais quatro métricas serão empregadas para medir e comparar o desempenho da aplicação nas duas bases de dados frente ao *baseline*.

3. Propor uma metodologia sistemática para a integração da aplicação desenvolvida em um sistema de recomendação baseado em filtro colaborativo. Esta metodologia esteve presente desde a adaptação da aplicação para o contexto específico do sistema até a definição de métricas adequadas para avaliar o desempenho do experimento.

4. Elucidar os procedimentos metodológicos necessários para a replicação dos experimentos em outros contextos. São detalhadas as etapas envolvidas na configuração do ambiente experimental, na coleta e preparação dos dados, na execução dos experimentos e na análise dos resultados obtidos. Especial atenção é dada à documentação dos processos e à transparência metodológica, visando assegurar a reprodutibilidade dos resultados por parte de outros pesquisadores.

Ao alcançar estes objetivos foi possível contribuir para o avanço do conhecimento no campo de sistemas de recomendação, ao propor uma abordagem inovadora para o

controle do viés de popularidade em recomendações baseadas em filtros colaborativos. O EQNet se mostrou capaz de mitigar o viés de popularidade sem impactar de forma negativa a qualidade das recomendações. Além disso, sua performance revelou-se significativamente competitiva quando comparada a um *baseline* amplamente reconhecido e tido como estado-da-arte em termos de *fairness*. Adicionalmente, sua integração em paralelo com o *baseline* resultou em melhorias substanciais em diversos cenários, superando os resultados obtidos por ambos os métodos de forma isolada. Diante disso, o EQNet não apenas oferece uma abordagem promissora para a gestão do viés de popularidade, mas também abre novas perspectivas para investigações futuras e aplicações práticas neste domínio.

1.2 Contribuições do Trabalho

Este trabalho apresenta contribuições no campo das Ciências da COmputação, especificamente no controle de viés de popularidade em sistemas de recomendação. As principais contribuições podem ser resumidas da seguinte forma:

1. **Desenvolvimento de uma abordagem para mitigar viés de popularidade:**

    - **Descrição:** Propomos uma nova abordagem para controle de viés de popularidade que combina técnicas de reanqueamento e reordenação para melhorar a capacidade exploratória da recomendação sem perdas significativas de qualidade.
    - **Impacto:** A abordagem tem o potencial de fornecer recomendações mais diversas com maior cobertura de portfólio sem perder a relevancia, aumentando a satisfação e o engajamento do usuário.

2. **Análise Comparativa com Abordagem no Estado-da-Arte:**

    - **Descrição:** Realizamos uma análise comparativa detalhada das diferentes formas de abordagem utilizando EQNet e o FA*IR, uma das abordagens no estado-da-arte utilizada para mitigar viés de popularidade e tornar as recomendações mais justas.
    - **Impacto:** Estas comparações fornecem insights valiosos sobre as vantagens e limitações de cada modelo, auxiliando pesquisadores e profissionais na escolha da técnica mais adequada para suas necessidades específicas.

3. **Implementação de um Sistema de Recomendação Prototipado:**

- **Descrição:** Desenvolvemos um protótipo funcional de sistema de recomendação, integrando a nova abordagem proposta e avaliamos o seu desempenho em um ambiente offline.
- **Impacto:** A implementação prática demonstra a viabilidade do algoritmo e serve como uma base para futuras pesquisas e desenvolvimento de sistemas de recomendação em diferentes contextos.

4. **Contribuições para a Literatura de Sistemas de Recomendação:**

- **Descrição:** Este trabalho contribui para a literatura existente, apresentando uma nova abordagem e resultados experimentais que podem ser utilizados como referência para estudos futuros.
- **Impacto:** As descobertas e metodologias apresentadas enriquecem o corpo de conhecimento na área de sistemas de recomendação, oferecendo novos caminhos para pesquisas subsequentes.

1.3 Organização do Trabalho

Para garantir uma estrutura coerente e neste trabalho, o presente documento é organizado em sete capítulos, cada um contribuindo para o entendimento, desenvolvimento e conclusão do estudo. O Capítulo 2 dedica-se à fundamentação teórica, explorando os fundamentos essenciais dos Sistemas de Recomendação e das Redes Complexas. Nesse contexto, serão definidos e discutidos conceitos básicos pertinentes, proporcionando uma base sólida para a compreensão dos temas abordados ao longo do trabalho.

No Capítulo 3, é apresentada uma revisão da literatura, enfocando os estudos correlatos relacionados às estratégias para mitigar o viés de popularidade em sistemas de recomendação. Além disso, são discutidas as técnicas que fazem uso de redes complexas como uma abordagem para aprimorar a eficácia desses sistemas. Essa revisão fornece uma visão panorâmica do estado atual da pesquisa nessa área, identificando lacunas e oportunidades para contribuições.

O Capítulo 4 introduz a abordagem proposta, denominada EQNet, destacando seus principais princípios, metodologias e diferenciais em relação às abordagens existentes. São detalhadas as etapas do desenvolvimento do EQNet, desde a concepção até a implementação prática, fornecendo uma compreensão de suas possíveis aplicações e potencialidades. Este capítulo serve como um guia detalhado para a compreensão da proposta e sua relevância no contexto atual da pesquisa em sistemas de recomendação.

No Capítulo 5, são apresentadas as métricas de avaliação utilizadas para medir o desempenho do EQNet, juntamente com uma descrição detalhada da aplicação experimen-

tal conduzida para validar sua eficácia e eficiência. Os resultados obtidos são analisados e discutidos criticamente no capítulo seguinte seguindo estes parâmetros, permitindo compreender um pouco mais de como o EQNet poderia se comportar em cenários reais de recomendação.

O Capítulo 6 é dedicado à exposição dos experimentos conduzidos com a abordagem do EQNet. Neste capítulo, são detalhados os procedimentos experimentais, incluindo a seleção de conjuntos de dados, configurações experimentais e métodos de avaliação. Os resultados obtidos serão apresentados, permitindo uma análise do desempenho do EQNet.

No Capítulo 7, é apresentada a conclusão deste estudo, em que os principais resultados, contribuições e limitações do trabalho são recapitulados. Além disso, algumas das possíveis direções para pesquisas futuras são propostas, identificando áreas promissoras para expansão e aprimoramento da abordagem.

2 REFERENCIAL TEÓRICO

O volume de informações presentes na internet é imenso e cresce de forma exponencial, como evidenciado por alguns estudos anteriores (DATA, 2017; BULAO, 2022). Enquanto essa abundância oferece inúmeras oportunidades para desenvolvedores e plataformas, ela também apresenta um desafio significativo para os usuários, que enfrentam uma sobrecarga de informação nos momentos de tomada de decisão. Nesse contexto, serviços eficientes de filtragem, priorização e recomendação de conteúdo tornam-se essenciais, e são amplamente fornecidos por diversas plataformas virtuais (GOOGLE TEAM, 2011; LINDEN, 2003; KOREN, 2009). Os sistemas de recomendação desempenham um papel crucial nesse cenário, sendo capazes de processar grandes volumes de informações sobre itens e usuários, fornecendo listas ranqueadas com base em correlações relevantes. Essas correlações podem ser estabelecidas por meio de características de itens previamente pesquisados ou comportamentos específicos dos usuários (RICCI, 2015).

É importante destacar que os dados utilizados por esses sistemas podem ser oriundos da própria plataforma ou consumidos de fontes terceirizadas, como dados de usuários do Facebook ou de instituições governamentais. Entretanto, dada a natureza sensível dos dados de usuários envolvidos, é fundamental que as pesquisas e desenvolvimentos em Sistemas de Recomendação estejam atentos às questões de privacidade e segurança de dados. Neste sentido, é imperativo observar os aspectos éticos e as regulamentações vigentes, como a Lei Geral de Proteção de Dados Pessoais (LGPD) no Brasil e a *General Data Protection Regulation* (GDPR) na União Europeia. No contexto deste estudo, é relevante ressaltar que o modelo proposto adere estritamente às diretrizes estabelecidas por essas legislações, utilizando bases de dados anonimizadas, conforme preconizado pelas normas (BRASIL, 2018; UNIÃO EUROPEIA, 2016).

Além das características de eficiência computacional e segurança, é de suma importância que as recomendações atendam às necessidades dos usuários, tanto em termos de sua afinidade com os interesses do usuário quanto em seu potencial de descoberta (STEFANIDIS, 2021; AI, 2020). Portanto, é crucial compreender os recursos e as distintas abordagens das metodologias de recomendação existentes, bem como entender como cada uma dessas características pode ser explorada em seus respectivos modelos. Os sistemas de recomendação são comumente classificados em quatro categorias principais,

relacionadas à maneira como realizam suas operações de ranqueamento: filtragem baseada em conteúdo, filtragem colaborativa, filtragem baseada em contexto e filtragem híbrida (RICCI, 2015; HARUNA et al., 2017), como ilustrado na Figura 1.

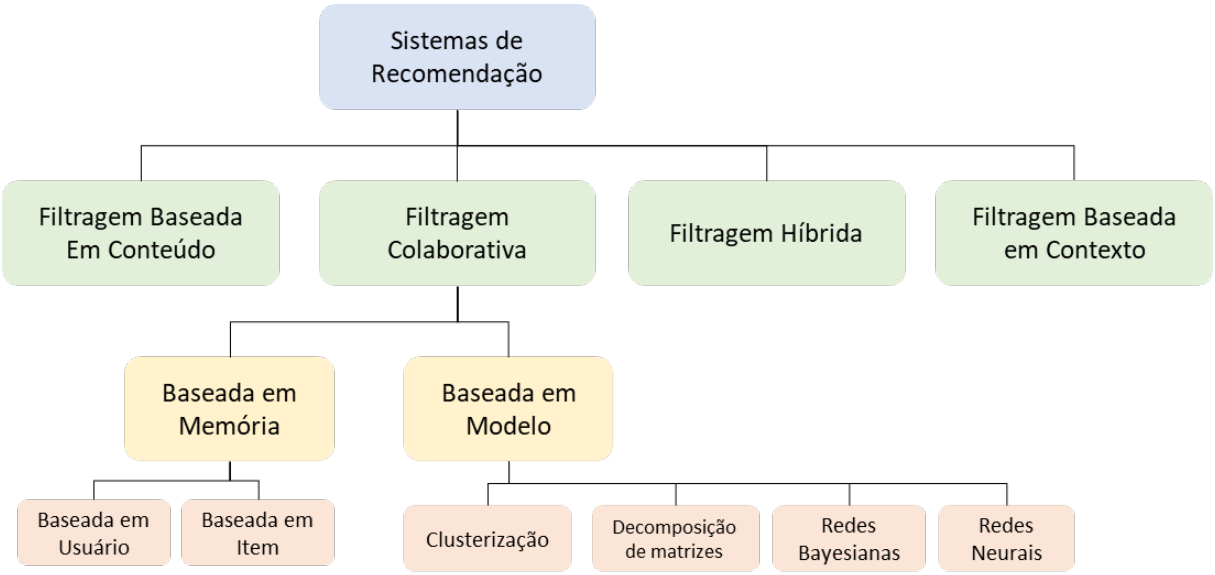

Figura 1 – Imagem representativa das principais tipos de sistemas de recomendação e suas possíveis ramificações

Fonte: Autor

Este capítulo estabelece a base teórica necessária para a compreensão dos sistemas de recomendação, abordando conceitos fundamentais que serão aplicados no desenvolvimento do EQNet. Ao explorar profundamente as técnicas de filtragem colaborativa e os desafios do viés de popularidade, criamos uma fundação sólida para a proposta deste trabalho, que visa otimizar recomendações sem comprometer a qualidade, conforme definido nos objetivos do trabalho.

2.1 Sistemas de Recomendação baseados em Conteúdo

Sistemas de recomendação baseados em filtros de conteúdo têm como objetivo sugerir itens que compartilham semelhanças com aqueles previamente selecionados pelo usuário. Esses sistemas abordam a questão como uma busca por itens afins, analisando o conteúdo e as descrições dos mesmos para identificar aqueles que possam interessar ao usuário. Nesse contexto, o processo envolve a filtragem e análise dos atributos dos itens, tais como palavras-chave, gênero de filme e estilo musical, com base nas interações passadas do usuário. Ao construir um perfil de interesses do usuário, esses sistemas consideram as preferências manifestadas por meio das interações anteriores com os conteúdos. A partir dessas informações, os itens são classificados e ranqueados em um catálogo, levando

em conta sua similaridade com os itens previamente consumidos e as avaliações atribuídas pelo usuário. A eficácia desse tipo de recomendação depende da qualidade do perfil de usuário construído e da precisão na identificação de itens relevantes. Esses sistemas representam uma abordagem importante para a personalização da experiência do usuário e são amplamente explorados em diversas áreas (LOPS; GEMMIS; SEMERARO, 2011; THANAPALASINGAM et al., 2018; WANG, 2018).

Nestes casos, os processos de recomendação podem ser subdivididos em três etapas (RICCI, 2015):

1) Análise de Conteúdo: É uma etapa crucial de pré-processamento quando a informação não está estruturada. Nessa fase, busca-se estruturar as informações relevantes sobre um determinado produto ou ação, a fim de prepará-las para alimentar os estágios subsequentes do processo. Por exemplo, considere-se um cenário no qual um item i, composto por n_i palavras-chave, está sendo comparado a outro item j com n_j palavras-chave. O coeficiente de semelhança c_{ij} entre esses dois itens pode ser calculado como a proporção relativa entre a interseção dos conjuntos de palavras-chave e a soma total de palavras-chave entre os dois itens:

$$c_{ij} = \frac{|n_i \cap n_j|}{|n_i + n_j|} \tag{2.1}$$

Esse cálculo fornece uma medida objetiva da similaridade entre os itens com base na sobreposição de suas palavras-chave, contribuindo assim para análises comparativas.

2) Vetorização: Os coeficientes entre os itens são submetidos à parametrização, representando uma etapa fundamental no processo analítico. Esta fase é caracterizada por um alto custo computacional, uma vez que demanda a comparação exaustiva de todos os pontos do quadro em análise (item por item), visando estabelecer a correlação entre eles e os pontos de interesse. Este processo intensivo requer recursos substanciais de processamento, pois envolve a avaliação de cada par de elementos, para tentar identificar padrões e correlações dentro do conjunto de dados.

3) Componente de Filtragem: Após o mapeamento, a matriz é filtrada por parâmetros de limitação para gerar as recomendações. Duas etapas muito comuns são a de limite mínimo, (onde se utilizam apenas itens que tenham um $c_{ij} > \mathrm{X}$), e/ou limite de listagem decrescente (onde se listam os itens da matriz de forma decrescente e seleciona-se apenas os top 10 da lista, por exemplo).

Após a fase de mapeamento, é essencial aplicar filtros à matriz resultante, seguindo parâmetros de limitação para produzir recomendações precisas e relevantes. Duas técnicas amplamente empregadas nesse contexto são o estabelecimento de um limite mínimo, onde somente itens com um coeficiente de similaridade c_{ij} acima de um valor X são considerados,

ou a imposição de um limite de listagem decrescente. Na abordagem do limite de listagem decrescente, os itens da matriz são ordenados de forma decrescente com base em seus valores de similaridade, e apenas os principais itens, por exemplo, os top 10 da lista, são selecionados para inclusão nas recomendações finais. Esses processos de filtragem buscam garantir que apenas as melhores sugestões sejam apresentadas aos usuários, contribuindo assim para a eficácia do sistema de recomendação.

À medida que o usuário interage com as recomendações, seu comportamento pode contribuir para um repositório de feedback. Esse repositório, inicialmente vazio ou com algumas definições explícitas fornecidas pelo usuário, registra tanto feedbacks positivos quanto negativos, como preferências ou aversões em relação aos itens recomendados. Esses feedbacks podem ser explícitos, como avaliações de produtos, ou implícitos, como o tempo gasto visualizando um determinado item. Esse ciclo de interações entre o usuário e as sugestões de itens resulta na geração de feedbacks relevantes, que podem ser integrados ao processo de aprendizado do algoritmo de recomendação. Essa retroalimentação contínua é essencial para melhorar a precisão e a personalização das recomendações oferecidas aos usuários.

2.2 Sistemas de Recomendação com Filtro Colaborativo

A filtragem colaborativa é uma técnica de recomendação que se fundamenta na ideia de que usuários com preferências semelhantes tendem a gostar dos mesmos itens. Em outras palavras, ela utiliza as opiniões e avaliações de um conjunto de usuários para prever as preferências de um usuário específico. Por meio da análise desses padrões de comportamento coletivo, os sistemas de filtragem colaborativa conseguem sugerir itens que possam ser do interesse do usuário-alvo, mesmo que estes não tenham sido previamente explorados por ele.

Similarmente aos filtros baseados em conteúdo, os sistemas de filtragem colaborativa visam resolver o desafio de recomendar itens relevantes aos usuários a partir de um vasto catálogo. Essa abordagem envolve a coleta e análise de dados de preferência dos usuários. Esses dados podem incluir avaliações explícitas, como classificações numéricas ou "curtidas", bem como informações implícitas, como histórico de compras, tempo de visualização e padrões de interação. Com base nesses dados, o sistema identifica usuários com gostos semelhantes e utiliza suas avaliações para fazer recomendações personalizadas para usuários individuais. (RICCI, 2015; ZHANG et al., 2019).

Um exemplo de sistema de filtragem colaborativa é o utilizado por plataformas de *streaming* de vídeo, como Netflix e Prime Video. Esses sistemas analisam o histórico de visualização e avaliações dos usuários para recomendar filmes e séries de TV que possam

corresponder aos seus gostos pessoais. Além disso, sistemas de recomendação baseados em filtragem colaborativa são amplamente aplicados em comércio eletrônico, mídias sociais, sistemas de música online, entre outros domínios, com o objetivo de aumentar a satisfação do usuário e impulsionar o engajamento (WANG et al., 2016; NANATH; AHMED, 2020).

A base para este fenômeno está na criação de agrupamentos (clusters) que reúnem usuários e itens semelhantes, juntamente com a parametrização dos comportamentos desses usuários. Devido à relevância desses agrupamentos, uma abordagem amplamente utilizada nesse tipo de algoritmo é a comparação entre o usuário e seu entorno próximo, a fim de identificar similaridades. Por meio dessa análise comparativa entre pares, o algoritmo é capaz de estabelecer parâmetros para a filtragem das listagens de itens (JANNACH, 2010).

Em suma, os sistemas de recomendação com filtragem colaborativa representam uma poderosa ferramenta para personalização e customização de experiências online. Ao analisar o comportamento coletivo dos usuários, esses sistemas são capazes de oferecer sugestões precisas e relevantes, contribuindo para a descoberta de novos conteúdos e maximizando a satisfação do usuário. Estes métodos de recomendação por filtro colaborativo podem ser divididos basicamente em dois tipo: os *métodos baseados em memória* e os *métodos baseados em modelo*. Estes métodos serão detalhado nos próximos tópicos.

2.3 Sistemas de Recomendação com Filtro Baseado em Contexto

Os Sistemas de Recomendação Baseados em Contexto (Context-Aware Recommender Systems, CARS) incorporam informações contextuais para fornecer recomendações mais precisas e relevantes, considerando fatores como localização, hora do dia, dispositivos utilizados, entre outros. Esta inclusão de informações contextuais pode melhorar consideravelmente a precisão das recomendações, uma vez que as preferências dos usuários podem variar conforme o contexto. Os CARS podem ser classificados de acordo com a maneira como lidam com as informações contextuais (HARUNA et al., 2017; CHAMPIRI; SHAHAMIRI; SALIM, 2015). As principais abordagens incluem:

1) Modelagem Explícita do Contexto: Nesta abordagem, o contexto é tratado como uma dimensão adicional nos dados de recomendação. Um modelo multidimensional é usado para capturar a interação entre usuários, itens e contexto.

2) Pré-Filtragem Contextual: O contexto é usado para filtrar os dados antes que o algoritmo de recomendação seja aplicado. Isso reduz o conjunto de dados a ser considerado, focando apenas nas interações relevantes ao contexto atual.

3) Pós-Filtragem Contextual: Após gerar uma lista de recomendações usando um algoritmo tradicional, a lista é ajustada com base no contexto. Isso pode envolver reordenar ou modificar as recomendações conforme o contexto atual.

A incorporação do contexto nos sistemas de recomendação não apenas enriquece a qualidade das recomendações, mas também abre novas possibilidades para a personalização e a satisfação do usuário (LIVNE et al., 2021). Isso se deve principalmente pelas abordagens contextuais permitem que as recomendações sejam mais dinâmicas e adaptativas, refletindo a complexidade das preferências dos usuários em diferentes situações. Ao modelar explicitamente o contexto, os CARS conseguem capturar nuances nas preferências dos usuários que sistemas tradicionais muitas vezes ignoram (YEBDRI et al., 2021).

2.3.1 Filtro Colaborativo Baseado em Memória

A identificação de usuários vizinhos que compartilham preferências é um aspecto crucial nos sistemas de recomendação, sendo geralmente baseada nos itens previamente avaliados pelo usuário em questão (ZHAO; SHANG, 2010). Uma vez estabelecida a vizinhança, uma variedade de algoritmos pode ser empregada para combinar suas preferências e gerar recomendações. Essas técnicas, devido à sua eficácia comprovada, encontram aplicação generalizada em diversos domínios.

Os algoritmos baseados em memória centrados no usuário operam a partir de uma representação parcial dos perfis dos usuários e uma série de pesos derivados da base de dados. Ao armazenar essas informações em memória, o algoritmo pode predizer as preferências com base na soma ponderada dos votos dos usuários que compartilham perfis semelhantes (BREESE, 1998). Tanto os itens quanto os usuários são avaliados quanto à sua similaridade nesse tipo de sistema. O algoritmo de k Vizinhos Mais Próximos (KNN) é frequentemente empregado para calcular essa similaridade, selecionando os k usuários mais similares para gerar predições (ADOMAVICIUS, 2005).

De maneira similar, a filtragem colaborativa baseada em itens calcula recomendações com base na similaridade entre os itens, em oposição à similaridade entre os usuários. Inicialmente, essa técnica realiza uma fase de construção, onde busca identificar a semelhança entre todos os pares de itens. Essa similaridade pode ter diversas formas, como a correlação entre as classificações ou o cosseno entre os vetores de classificação. Assim como nos modelos baseados em usuários, as métricas de similaridade se beneficiam de classificações normalizadas. (JANNACH, 2010).

Desta forma, esses algoritmos, conhecidos como algoritmos baseados em memória, realizam suas previsões de ranqueamento considerando a proximidade vetorial dos itens

em uma determinada coleção de avaliações de outros usuários. Esse modelo de recomendação é interessante, pois permite que os usuários explorem itens que compartilham características com aqueles que já foram buscados anteriormente. Trata-se de um método simples, facilmente justificável, eficiente e estável (RICCI, 2015).

Em um sistema de recomendação centrado no usuário, baseado em memória, a predição de uma nota r_{ui} atribuída por um usuário u a um item i é realizada utilizando as avaliações fornecidas por usuários similares a u, conhecidos como seus vizinhos mais próximos. Os k vizinhos mais próximos de u são identificados com base em uma medida de similaridade w_{uv} entre os usuários u e v. A partir do conjunto de vizinhos mais próximos $Ni(u)$ que já avaliaram o item i e são similares a u, é possível estimar a afinidade de u por i. Esta estimativa é calculada considerando que a similaridade entre usuários pode ser ponderada por pesos w_{uv}, os quais refletem a proximidade entre eles, e um fator de normalização h das avaliações para compensar as discrepâncias entre usuários mais e menos exigentes. A fórmula para estimar r_{ui} é dada por:

$$r_{ui} = \frac{\sum_{v \in N_i(u)} w_{uv} h(r_{vi})}{\sum_{v \in N_i(u)} |w_{uv}|} \tag{2.2}$$

Embora esta abordagem seja reconhecida por sua simplicidade e eficácia, é importante observar que ela demanda considerável espaço em memória para sua execução. Em contextos nos quais o volume de usuários e itens seja significativo, os limites de memória da infraestrutura podem ser ultrapassados. Por esse motivo, muitas plataformas optam por estratégias alternativas para gerar recomendações colaborativas, seja utilizando abordagens distintas ou combinando filtros baseados em memória com outras técnicas, como o método de filtro colaborativo baseado em modelo.

2.3.2 Filtro Colaborativo Baseado em Modelo

Os filtros colaborativos baseados em modelo visam descobrir padrões significativos a partir dos dados de interações dos usuários com a plataforma. Por meio da análise desses dados, esses filtros constroem modelos que capturam informações relevantes para o ranqueamento das recomendações, sem a necessidade de consultar toda a base de dados a cada solicitação. Essa característica torna-os eficientes, especialmente em ambientes com grandes conjuntos de dados, mantendo uma baixa demanda de memória (ADOMAVICIUS, 2005).

A construção desses modelos é realizada através da aplicação de técnicas de mineração de dados e aprendizado de máquina, com o objetivo de prever as preferências dos usuários para itens ainda não avaliados. Diversos algoritmos são empregados nesse contexto, como Redes Bayesianas, modelos de agrupamento, modelos semânticos laten-

tes (por exemplo, decomposição de valores singulares) e modelos baseados em processos de decisão de Markov (SU; KHOSHGOFTAAR, 2009). Esses modelos permitem que o sistema reconheça padrões complexos nos dados de treinamento e realize previsões inteligentes, tanto para dados de teste quanto para dados do mundo real, com base nos padrões aprendidos.

Uma vantagem significativa dos filtros colaborativos baseados em modelo é sua capacidade de superar algumas limitações dos algoritmos baseados em memória, tais como escalabilidade e generalização. Enquanto os algoritmos de memória dependem fortemente de consultas diretas à base de dados durante a fase de recomendação, os modelos baseados em modelo podem fornecer recomendações mais eficazes e personalizadas, mesmo em grandes sistemas de recomendação (BREESE, 1998; GONG, 2009). Essa capacidade de generalização e eficiência torna esses modelos uma área de pesquisa ativa e relevante na construção de sistemas de recomendação.

Um exemplo pertinente seria o caso de se tentar prever a afinidade de um usuário a um item j, com base nas avaliações dos i itens com características similares a j contidos no conjunto de avaliações do usuário $S(u)$. Nesse contexto, $dev_{j,i}$ denota o desvio médio entre os itens i e j. Por meio desse desvio, é possível formular uma predição para o item j em relação à u, expressa como a soma do desvio $dev_{j,i}$ e da avaliação do usuário ao item $i(U_i)$ (JANNACH, 2010).

$$dev_{j,i} = \sum_{(u_j,u_i)\in S_{j,i}(R)} \frac{u_j u_i}{|S_{j,i}(R)|} \tag{2.3}$$

$$pred(u,j) = \frac{\sum_{i\in S(u)-|j|}(dev_{j,i} + u_i) * |S_{j,i}(R)|}{\sum_{i\in S(u)-|j|} *|S_{j,i}(R)|} \tag{2.4}$$

Diversas metodologias têm sido empregadas para o desenvolvimento de algoritmos nesse contexto, tais como Fatoração de Matriz, Modelagem por Redes Complexas, Probabilidade Bayesiana e Redes Neurais (SCHAFER, 2007; YAO, 2014; AI, 2017). De maneira simplificada, esses modelos matemáticos visam interpretar as interações dos usuários com o sistema, sejam elas explícitas ou implícitas, a fim de gerar modelos comportamentais. Posteriormente, esses modelos são aplicados a outros usuários para antecipar suas futuras interações. Nesse perfil de algoritmo, o modelo geralmente é retroalimentado pelas interações dos usuários em relação às suas predições, permitindo, assim, um processo contínuo de aprendizado de máquina e aprimoramento na qualidade das sugestões ao longo do tempo (RICCI, 2015). Um exemplo de um algoritmo empregado com esse propósito é o SVD, que promove a desagregação de uma matriz de interações (item-item, usuário-usuário ou usuário-item) e, por meio dos parâmetros obtidos desta decomposição, gera modelos de sugestão (JANNACH, 2010).

Entretanto, as soluções de recomendação baseadas em modelo enfrentam desafios comuns, como em conjuntos de dados incipientes ou com usuários recentes, os quais podem apresentar condições para o problema conhecido como "partida a frio", no qual as informações disponíveis não são suficientes para gerar um modelo de recomendação eficaz (JANNACH, 2010). Outro desafio evidente é o viés de popularidade, que pode ser observado quando recomendações injustas são disseminadas, favorecendo poucos itens muito populares frente aos demais (Abdollahpouri, 2020). Tal prática pode influenciar o comportamento dos usuários e tornar a experiência de navegação menos atrativa para aqueles que frequentam o sistema com maior assiduidade, já que impacta muito o processo de descoberta (BEDI, 2014; SABOUNI, 2017).

2.3.3 Filtragem Híbrida

As técnicas de filtragem de informação mencionadas anteriormente desempenham um papel fundamental nos sistemas de recomendação, constituindo a base para descobrir relações entre diferentes itens e usuários. Quanto mais refinadas essas técnicas, mais precisas tendem a ser as predições e recomendações resultantes para o usuário. Entretanto, é importante reconhecer que essas abordagens não estão isentas de limitações, como discutido anteriormente.

A abordagem de filtragem híbrida surge como uma resposta para combinar as vantagens e mitigar as desvantagens das técnicas de filtragem mencionadas (KIM Q. LI, 2006). Diversas estratégias de filtragem híbrida foram propostas, cada uma operando de forma distinta na combinação dos componentes para gerar recomendações (BURKE, 2002; NGUYEN; ZHU, 2013; KIM et al., 2010; CANO; MORISIO, 2017). Destacam-se:

1) Abordagem Ponderada: Aqui, a filtragem baseada em conteúdo e a filtragem colaborativa são implementadas separadamente, com seus resultados combinados por meio de uma ponderação linear. É possível que uma normalização seja necessária nos resultados individuais antes da combinação, caso essas técnicas produzam valores em escalas diferentes.

2) Abordagem Mista: Nesta estratégia, as recomendações geradas por ambas as técnicas são combinadas no processo final, apresentando ao usuário uma lista integrada de recomendações provenientes de ambas as fontes.

3) Combinação sequencial: Esta abordagem inicia com a criação dos perfis de usuários pela filtragem baseada em conteúdo, os quais são posteriormente utilizados no cálculo da similaridade na filtragem colaborativa.

4) Estratégia de Comutação: Nesse caso, o sistema emprega critérios, como a confiança nos resultados, para alternar entre a filtragem baseada em conteúdo e a filtragem

colaborativa. A comutação pode ocorrer em momentos específicos, como nos pontos onde uma técnica compensa as deficiências da outra.

2.3.4 Viés de Popularidade e Diversidade

Controlar o viés de popularidade emerge como um dos principais desafios enfrentados pelos sistemas de recomendação, dada sua estreita associação com a exposição dos usuários a itens mais ou menos populares nas listagens de recomendação. Por exemplo, em plataformas de *streaming*, que oferecem milhares de filmes, a prevalência de alguns poucos blockbusters, amplamente consumidos e apreciados por um grande número de usuários, frequentemente resulta na dominação quase total das listas de recomendação por esses filmes. Tal fenômeno decorre do viés inerente aos dados de classificação, que tendem a favorecer os itens mais populares (STECK, 2011; MALDONADO, 2021).

Neste estudo, adotamos como critério de popularidade aqueles itens que acumulam 80% das interações com os usuários, considerando o restante como pertencente à chamada "cauda longa" (ANDERSON, 2008). Como uma proporção bastante reduzida do catálogo de itens normalmente concentra essa parcela de interações (tipicamente entre 5% e 10%), os algoritmos de recomendação acabam privilegiando esses itens em detrimento dos demais. Isso gera um ciclo vicioso em que os itens mais populares são continuamente recomendados, aumentando ainda mais sua popularidade e excluindo outros da atenção dos usuários. O Gráfico 1 ilustra a distribuição de popularidade no conhecido conjunto de dados MovieLens 1M, embora essa distribuição seja observada em diversos outros conjuntos de dados (ABDOLLAHPOURI M. MANSOURY; MOBASHER, 2019). Enquanto isso, os itens da cauda longa possuem conexões limitadas com a maioria dos usuários, o que os mantém à margem do interesse geral, a menos que medidas sejam tomadas para aumentar sua visibilidade e atratividade (BEDI, 2014).

Além disso, a relação entre a popularidade de um item e sua frequência de recomendação varia consideravelmente de acordo com a técnica de recomendação empregada. Em suas investigações, Abdollahpouri destaca que o método SVD mostrou uma tendência linear mais pronunciada e uma dispersão menor nessa correlação (ABDOLLAHPOURI, 2019). Com o intuito de simplificar a execução do experimento e permitir uma observação mais objetiva do desempenho da abordagem proposta, optou-se pela utilização de um algoritmo de recomendação similar neste estudo (LU, 2017; RENDLE et al., 2020; ANELLI et al., 2022).

Os itens de menor interação dos usuários, comumente referidos como "cauda longa", constituem uma parcela significativamente maior do catálogo em comparação com os itens populares. Essa dinâmica oferece uma ampla oportunidade para explorar a diversidade presente nesses itens mais específicos. Ao considerar a interseção dos usuários com esses

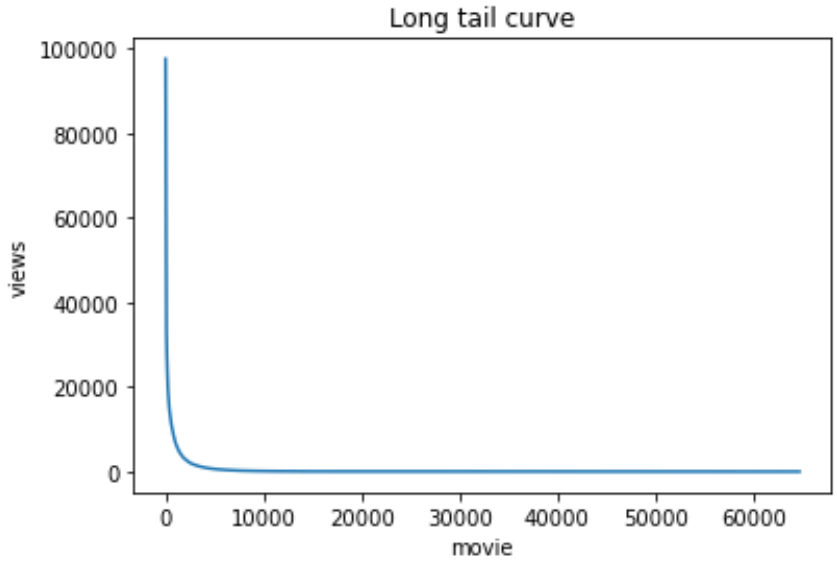

Figura 2 – Curva de filmes versus número de interação de usuário do *dataset* do MovieLens

Fonte: Autor

itens de cauda longa, é possível obter insights valiosos sobre os comportamentos de nicho, contribuindo para melhorias nas recomendações e no dimensionamento do consumo (ABDOLLAHPOURI M. MANSOURY; MOBASHER, 2019). Empresas como a Amazon têm uma parte substancial de sua receita derivada desses itens menos populares, evidenciando o impacto significativo que o viés de popularidade pode exercer sobre um negócio, caso não seja devidamente gerenciado (BRYNJOLFSSON; SMITH, 2006; ANDERSON, 2008). Além disso, à medida que as bases de dados aumentam em tamanho, as matrizes de interação entre usuário e item tendem a se tornar mais esparsas. Nesses cenários, é crucial adotar abordagens cuidadosas nas técnicas de recomendação para manter sua eficácia (SUN; XU, 2019; AI, 2020; JAVAID, 2021).

Nas últimas décadas, uma série de estudos têm abordado metodologias para explicar, medir e gerenciar o viés de popularidade. Estes trabalhos representam contribuições significativas para a identificação, caracterização e propostas de mitigação e manipulação dos vieses de popularidade em sistemas de recomendação (MARLIN, 2007; STECK, 2011; GOPALAN, 2015; JANNACH, 2015). Em um contexto mais recente, Cañamares e Castells investigaram o viés de popularidade em algoritmos de kNN, estabelecendo uma distinção formal entre avaliação com e sem viés (CAñAMARES; CASTELLS, 2017).

Os estudos sobre o viés de popularidade também revelam aspectos intrigantes sobre seu impacto na experiência do usuário, destacando a geração de recomendações conhecidas como "falsos positivos"e a repetição excessiva das mesmas sugestões para um mesmo grupo de usuários (CREMONESI, 2011; MALDONADO, 2021). Essas recomendações

equivocadas não apenas afetam adversamente as métricas de publicidade online e de mídia social, mas também levam a uma percepção negativa do serviço oferecido (INSTAGRAM BUSINESS TEAM, 2016; TOMAS, 2021). Como estas experiências negativas tendem a ser mais memoráveis aos usuários do que as positivas, a falta de diversidade e a repetição de recomendações podem ser profundamente prejudiciais para a viabilidade a longo prazo de uma plataforma (YIN, 2010).

Lembrando que o viés de popularidade não é necessariamente algo negativo que precisa ser eliminado, já que as recomendações se baseiam em grande parte nele, mas deve ser entendido e controlado. Desta maneira é possível perceber como não só gerenciar este tipo de viés pode ser positivo e melhorar a performance de plataformas como também, se deixado de lado pode causar grandes prejuízos. Por causa disso o grande interesse da comunidade científica em estudar este fenômeno (ABDOLLAHPOURI, 2020; ADOMAVICIUS; KWON, 2009; KISWANTO; NURJANAH; RISMALA, 2018).

No entanto, é importante ressaltar que o viés de popularidade não é intrinsecamente negativo e, de fato, é uma parte fundamental das recomendações, mas requer compreensão e controles adequados. Portanto, é crucial reconhecer que a gestão eficaz desse viés pode otimizar o desempenho das plataformas, enquanto a negligência em lidar com ele pode acarretar sérios prejuízos. Isso tem despertado um interesse considerável na comunidade científica, que tem se dedicado a investigar esse fenômeno complexo (ADOMAVICIUS; KWON, 2009; KISWANTO; NURJANAH; RISMALA, 2018; ABDOLLAHPOURI, 2020).

Além disso, a correlação entre a popularidade de um item e o número de vezes que é recomendado varia dependendo da técnica de recomendação aplicada. Estudos anteriores mostraram que o algoritmo SVD tem um comportamento mais linear em relação à popularidade de seus itens versus variação da recomendação (ABDOLLAHPOURI, 2019). Outro ponto interessante sobre recomendações por produto escalar, como o SVD, é que a similaridade obtida por estes meios simplifica a modelagem e o aprendizado, oferecendo uma melhor compatibilidade com outras áreas de pesquisa, como processamento de linguagem natural ou modelos de imagem (RENDLE et al., 2020). Portanto, o algoritmo escolhido para gerar a decomposição da matriz e criar as recomendações foi o SVD, popularizado por Simon Funk durante o Netflix Prize (KOREN, 2009; HUG, 2015).

2.4 O SVD

O algoritmo de Decomposição em Valores Singulares (SVD - Singular Value Decomposition) constitui uma técnica fundamental em diversas áreas, incluindo os sistemas de recomendaçãos, onde já é amplamente usado com sucesso a muitos anos (KOREN, 2009; CHAI et al., 2018; ANWAR; UMA; SRIVASTAVA, 2021; SHENGXI et al., 2024).

Para o experimento utilizamos o SVD proveniente da biblioteca Surprise do Python, sendo implementado para fornecer uma solução eficaz na geração de recomendações e testar a variação dos parâmetros de recomendação após o pós-processamento do EQNet (HUG, 2015). A essência do algoritmo reside na sua capacidade de decompor uma matriz de avaliações de usuários em produtos, possibilitando a representação dos dados de forma mais compacta e significativa. Considerando uma matriz R de dimensão $m * n$, onde m representa o número de usuários e n o número de itens, a decomposição SVD busca representar R como o produto de três matrizes:

$$R \approx U\Sigma v^t \tag{2.5}$$

Onde U é uma matriz $m*k$ cujas colunas representam os vetores de características latentes dos usuários. Σ é uma matriz diagonal $k * k$ contendo os valores singulares. V^T é a transposta de uma matriz $k * n$ cujas linhas representam os vetores de características latentes dos itens. O parâmetro k representa a dimensionalidade do espaço latente, determinando o número de características latentes consideradas na representação dos usuários e itens.

Desta forma, a recomendação de um item i para um usuário u é realizada através do produto interno entre o vetor de características latentes do usuário u e o vetor de características latentes do item i, conforme a equação:

$$r'_u i = q_i^T p_u \tag{2.6}$$

Assim, $r'_u i$ representa a estimativa da avaliação do usuário u para o item i, q_i denota o vetor de características latentes do item i e p_u denota o vetor de características latentes do usuário u. A obtenção dos vetores de características latentes p_u e q_i é realizada através da minimização da função de perda, que quantifica a discrepância entre as avaliações reais e as estimadas. Dessa forma, o algoritmo SVD, implementado na biblioteca Surprise, oferece uma abordagem robusta e eficiente na geração de recomendações.

2.5 Redes Complexas e PageRank

De acordo com a teoria das redes complexas, os sistemas complexos podem ser representados por grafos, nos quais os nós estão interconectados de maneira mutuamente relacionada. Essas redes apresentam características topológicas que não são observadas em redes simples, como as redes aleatórias, sendo frequentemente encontradas em grafos modelados a partir de sistemas reais (NASIRUZZAMAN; POTA, 2011). Essa abordagem permite a representação de uma diversidade de fenômenos do mundo real, proporcionando

soluções para uma ampla gama de problemas. Por exemplo, é viável modelar a infraestrutura física de uma rede de computadores, como a internet, em que os computadores conectados representam os nós e as conexões entre eles são representadas pelas arestas do grafo. Analogamente, outras situações podem ser modeladas, tais como a estrutura de páginas da World Wide Web (WWW), as relações sociais entre grupos de pessoas, redes organizacionais, redes neurais, entre outras (ALBERT; BARABáSI, 2002).

Os estudos relacionados às redes complexas tiveram início em meados da década de 1930, quando sociólogos buscavam compreender o comportamento da sociedade e as interações entre os indivíduos. Essas pesquisas visavam identificar características peculiares das redes sociais e seus elementos, como a centralidade (que se refere ao vértice mais central) e a conectividade (vértices com maior número de conexões) (BARABASI; ALBERT, 1999). Nas redes sociais, os indivíduos são representados como vértices, e as interações entre eles são representadas pelas arestas. A análise da centralidade e da conectividade era frequentemente empregada para identificar os indivíduos mais influentes ou aqueles que mantinham as melhores relações dentro da rede (FREEMAN, 1977, 1978).

Com o avanço contínuo da tecnologia da informação e a proliferação de computadores acessíveis, a capacidade de analisar dados em larga escala tem gerado uma transformação significativa no campo da ciência da computação. Anteriormente centradas em investigações que exploravam redes de pequena escala e as propriedades individuais de seus vértices e arestas, as pesquisas agora abrangem análises estatísticas em nível macro. Estudos contemporâneos frequentemente envolvem redes compostas por milhões ou até bilhões de nós, explorando uma variedade de relações, como centralidade, influência e popularidade. Essa mudança de paradigma revelou nuances que distinguem as redes do mundo real dos modelos de redes aleatórias, que, durante anos, foram considerados o principal modelo de redes (NEWMAN, 2003).

O PageRank, originalmente desenvolvido por Larry Page e utilizado no mecanismo de busca do Google, fundamenta-se na estrutura complexa de rede da web para calcular e atribuir pesos aos nós de forma proporcional à sua importância na rede (GOOGLE TEAM, 2011). Este algoritmo é particularmente interessante devido à sua capacidade de determinar a relevância de um nó com base na quantidade e qualidade das conexões que o apontam, utilizando cálculos de autovetores. Destaca-se que o PageRank é conhecido por privilegiar certos nós sobre outros, o que permite identificar quais têm maior influência sobre o sistema em questão. É relevante mencionar que todas as patentes associadas ao PageRank expiraram em 24 de setembro de 2019, o que amplia a liberdade para o uso de sistemas derivados deste algoritmo (PAGE, 1998).

Enquanto um autovetor, o PageRank representa uma heurística incremental estocástica. Nesse contexto, para manter a estabilidade na complexidade dos algoritmos,

é crucial identificar pontos de convergência ótimos, visando reduzir a variabilidade sem comprometer a eficiência do processo (YOU, 2017). Por exemplo, ao considerar uma rede V na qual o algoritmo PageRank está sendo aplicado para classificar o nó i e atribuir-lhe um valor x, a abordagem envolve calcular esse valor considerando todos os nós j pertencentes ao conjunto B_i, que interagem com i, e seus respectivos valores x, normalizados pelo número total de links D_j de j para i:

$$x_i = \sum_{j \in B_i} \frac{x_j}{D_j}, \forall i \in V \tag{2.7}$$

É importante notar que o algoritmo, em sua formulação padrão, não define um nó inicial para iniciar a avaliação, o que pode resultar em atribuições ligeiramente distintas para os nós em cada execução dentro de uma mesma rede. Além disso, durante o cálculo de x para i, o algoritmo faz transições para nós subsequentes de forma aleatória, repetindo iterativamente o processo até que todos os nós da rede tenham sido percorridos. Dependendo dos parâmetros definidos, essa iteração de classificação pode ser repetida n vezes, utilizando os valores de PageRank da iteração anterior para guiar as novas atribuições.

O PageRank ganhou destaque como um dos algoritmos mais influentes na classificação e análise de páginas da web, sendo a base do mecanismo de busca do Google. Além disso, sua capacidade de identificar e classificar itens populares em uma rede o torna uma ferramenta valiosa para diversas aplicações, incluindo recomendação de conteúdo, identificação de líderes de opinião e detecção de comunidades. Mais ainda, o PageRank aborda de maneira eficaz o problema do viés de popularidade, comumente encontrado em sistemas baseados em recomendação e classificação, fornecendo uma abordagem imparcial e robusta para avaliar a importância e relevância dos itens em uma rede, independentemente de sua popularidade inicial. Portanto, a inclusão do PageRank nesta dissertação se justifica não apenas por sua proeminência, mas também por sua capacidade de abordar problemas essenciais relacionados à análise e classificação de redes complexas.

Por sua capacidade de identificar e classificar itens populares em uma rede o PageRank já foi empregado de maneira eficaz para solucionar diversos desafios com Sistemas de Recomendação (NGUYEN, 2015; PARK W. LEE; LEE, 2019; AL_SULTANY; GHAIDAA, 2022). Portanto, o PageRank foi um dos algoritmos de ranqueamento escolhidos para integrar a abordagem descrita nesta dissertação e se justifica não apenas por sua proeminência, mas também por sua capacidade de abordar problemas essenciais relacionados à análise e classificação. Essa de aproveitar das capacidades de algoritmos de redes complexas para extrair informações intrínsecas de grupos de usuários foi inspirada em trabalhos anteriores, como o de Ai, que utilizou a centralidade de clusters para atenuar a influência dos nós mais dominantes em nós mais periféricos (AI, 2020).

Além do PageRank, utilizamos um outro algoritmo popular por ranquear os itens mais populares de um catálogo, chamado Popularity Count.

2.6 Popularity Count

O Popularity Count é um algoritmo utilizado e fornecido pela plataforma da AWS para permitir a compreensão das preferências e comportamentos dos consumidores, algo essencial para o sucesso de praticamente qualquer negócio. O Popularity Count permite identificar quais produtos ou serviços estão em alta demanda, facilitando a tomada de decisões estratégicas relacionadas a estoque, marketing e desenvolvimento de produtos (AWS Team, 2022). Além disso compreender como esses algoritmos funcionam e como podem ser otimizados contribui não apenas para avanços de mercado, mas também para o desenvolvimento teóricos do campo de recomendações. Este algoritmo é uma abordagem simples, porém eficaz, para identificar itens populares em um conjunto de dados. Sua lógica é direta: contabiliza o número de interações (como visualizações, compras ou avaliações) que cada item recebe e classifica os itens com base nesse número (JANNACH, 2010).

O funcionamento do algoritmo pode ser resumido em etapas simples:

1) Coleta de dados: O algoritmo requer acesso aos dados relevantes, como registros de interações de usuários com itens. 2) Contagem de interações: Para cada item, o algoritmo conta o número de interações recebidas ao longo de um período de tempo específico. 3) Classificação dos itens: Com base na contagem de interações, os itens são classificados em ordem decrescente de popularidade. 4) Recomendação: Os itens mais populares são recomendados aos usuários com base em suas preferências e comportamentos anteriores.

Entender como o EQNet performa com outros parâmetros de entrada vinculados à popularidade, diferentes do PageRank, também foi um ponto importante a ser avaliado. Para esta finalidade o algoritmo escolhido foi o Popularity Count também por ter sido aplicado método recente para mitigar efetivamente o Popularity Bias (STEFANIDIS, 2021). Em seu artigo, Borges aplicou a pontuação de Popularity Count para penalizar as pontuações atribuídas aos itens de acordo com a popularidade histórica por mitigar o viés e promover a diversidade nos resultados. Com n representando os usuários presentes em um cluster e $p(x_i)$ é a função de valor atribuído a uma determinada interação de um usuário i) com o item x com relação à sua recência (variando de 0 para nenhuma interação ou uma interação muito antiga para 1 em uma interação muito recente):

$$PopCount_i = \sum_{i=0}^{n} p(x_i) \tag{2.8}$$

Sua importância reside na capacidade de oferecer uma abordagem eficaz e direta para identificar padrões de popularidade em conjuntos de dados. Além disso, é o Popularity Count é versátil e apresenta variações que podem melhor atender outros padrões de bases de dados, como a possibilidade de reduzir a importância de iterações mais antigas (XIAO et al., 2020).

3 TRABALHOS RELACIONADOS

Diversas plataformas têm alcançado sucesso ao adotar estratégias eficazes para gerir os itens de cauda longo de seu portfólio, devido à sua importância para o refinamento do funcionamento de sistemas de recomendação (BRYNJOLFSSON; SMITH, 2006; ANDERSON, 2008; SUN; XU, 2019; DIDI et al., 2023). Grandes empresas de tecnologia, como Netflix, Amazon e outros *marketplaces*, obtêm cerca de 20% a 30% de sua receita por meio desses itens que não são tão facilmente encontrados em estabelecimentos de porte médio. Esse êxito se deve, em grande parte, ao profundo entendimento das preferências dos usuários e à eficiente gestão do viés de popularidade (ANDERSON, 2005).

Em 2009, Adomavicius abordou o desafio de gerenciar o viés de popularidade em sistemas de recomendação, desenvolvendo seis modelos de reclassificação para pontuação dos itens (ADOMAVICIUS; KWON, 2009). Neste estudo, o autor comparou a variação no aumento de itens de cauda longa nas listas de recomendações com a variação na precisão média das recomendações. Atualmente, compreende-se que a precisão representa apenas uma faceta da experiência do usuário em um sistema de recomendação. Assim, desde que não haja uma variação negativa significativa, a inclusão de itens de cauda longa tende a aprimorar a experiência global do usuário, mesmo que isso resulte em uma queda na precisão (KNIJNENBURG, 2012).

A preferência tendenciosa em direção a itens populares não só impacta as escolhas dos consumidores, mas também impede a visibilidade e o crescimento potencial de popularidade de itens menos populares, minando assim a equidade nas recomendações. As ramificações desse viés podem ser explicadas de forma abrangente ao examinar instâncias do mundo real. A prevalência do viés de popularidade leva à homogeneização do mercado, permitindo que poucos produtores de itens dominem o mercado (ABDOLLAHPOURI, 2019). Consequentemente, isso sufoca oportunidades para inovação e originalidade, restringindo a diversidade e limitando o escopo para novas ofertas.

Essa repetição de apenas alguns itens sendo recomendados ao mesmo usuário é muito cansativa e também representa um problema significativo de experiência (MALDONADO, 2021). Estudos psicológicos descrevem uma tendência no comportamento do usuário em que memórias negativas ligadas a uma experiência do usuário são mais fortes

e duradouras do que as boas (YIN, 2010), fazendo com que as experiências negativas decorrentes do viés de popularidade sejam muito devastadoras para uma plataforma a longo prazo. Portanto, esse tipo de viés pode também piorar a experiência do usuário e prejudicar a experiência geral que os usuários podem ter nela.

A correlação entre a popularidade de um item e o número de vezes que ele é recomendado varia dependendo da técnica de recomendação aplicada. Estudos anteriores mostraram que o algoritmo SVD tem um comportamento mais linear em relação à popularidade versus recomendação (ABDOLLAHPOURI, 2020). Além disso, o produto escalar pode ser uma melhor escolha padrão para combinar embeddings do que similaridades aprendidas usando MLP ou NeuMF, pois são mais relevantes para a indústria (LU, 2017; LI, 2019). A similaridade de produto escalar simplifica a modelagem e o aprendizado e proporciona um melhor alinhamento com outras áreas de pesquisa, como processamento de linguagem natural ou modelos de imagem (RENDLE et al., 2020).

3.1 Algoritmos de reordenação para redução de viés de popularidade

Para aferir os seus resultados, Adomavicius adota uma métrica de ganho de diversidade que é então comparada em paralelo com a perda de precisão ao longo da aplicação das cinco técnicas distintas de reclassificação avaliadas (por popularidade de item, predição reversa, avaliação média por item, apreciação absoluta por item e apreciação relativa por item) (ADOMAVICIUS; KWON, 2009). Nesse trabalho, a diversidade dos top N itens é descrita pela expressão $|\bigcup_{u \in U} L_N(u)|$ onde $L_N(u)|$ representa a lista dos top N itens que aparecem para o usuário u. Essas técnicas podem ser descritas matematicamente da seguinte forma:

- **Popularidade de Item**
  Classificação por popularidade do item, em ordem crescente, onde a popularidade é representada pelo número de classificações conhecidas que cada item tem:

$$rank_{ItemPop}(i) = |U(i)|, U(i) = \{u \in U | \exists R(u,i)\} \tag{3.1}$$

- **Predição Reversa**
  Classificação por valor de classificação previsto, em ordem crescente:

$$rank_{RevPred}(i) = R * (u,i) \tag{3.2}$$

- **Avaliação Média por Item**
  Classificação por uma média de todas as classificações conhecidas para cada item:

$$rank_{AvgRating}(i) = \overline{R(i)} = \frac{\sum_{u \in U(i)} R(u,i)}{|U(i)|} \tag{3.3}$$

- **Apreciação Absoluta por Item**
  Classificação por quantos usuários gostaram do item (ou seja, o avaliaram acima de T_H):

$$rank_{AbsLike}(i) = |U_H(i)|, U_H(i) = \{u \in U | R(u,i) \geq T_H\} \tag{3.4}$$

- **Apreciação Relativa por Item**
  Classificação pela porcentagem de usuários que curtiram um item (entre todos os usuários que o avaliaram).

$$rank_{RelLike}(i) = \frac{|U_H(i)|}{|U(i)|} \tag{3.5}$$

Existem também outras formas de avaliar a relevância dos itens além de sua popularidade direta, principalmente quando são levados em consideração alguns dos algoritmos de rede complexas como é o caso do PageRank. PageRank é um algoritmo amplamente conhecido para classificar páginas da web com base não apenas em suas características de popularidade explícita, mas também em suas conexões com outras páginas populares (GOOGLE TEAM, 2011). Alguns dos primeiros trabalhos relacionando o PageRank a sistemas de recomendação abordaram bases de filmes (MovieLens) e páginas da web, estruturando grafos a partir da interação entre usuários e itens para classificar nós e gerar listas de recomendações, semelhante à sua aplicação para o buscador Google (ZHANG; ZHANG; LI, 2008; SU; KHOSHGOFTAAR, 2009). Tendo como base este modelo, foi desenvolvida uma aplicação chamada EQNet que busca utilizar o PageRank como um de seus parâmetros de entrada para observar o princípio de redução de viés versus precisão.

Hoje temos várias maneiras de controlar o viés de popularidade por meio de algoritmos de reordenação (BEDI, 2014; SUN; XU, 2019; ABDOLLAHPOURI, 2020; MALDONADO, 2021). Neste contexto, dos algoritmos mais recentes, o FA*IR emerge como um *baseline* relevante para nossa investigação. O próximo passo desta dissertação é explorar detalhadamente o funcionamento do algoritmo FA*IR, destacando suas características distintivas e avaliando sua eficácia na mitigação do viés de popularidade em sistemas de recomendação. Ao adentrar na análise deste algoritmo específico, buscamos não apenas compreender suas nuances técnicas, mas também avaliar seu potencial para aprimorar a experiência do usuário e promover recomendações mais diversificadas e personalizadas.

3.2 Baseline - FA*IR

O algoritmo FA*IR foi desenvolvido com o propósito de mitigar o viés de popularidade em sistemas de recomendação e classificação de conteúdo. Desenvolvido com base em princípios de equidade e justiça, o FA*IR busca garantir que o processo de recomendação não favoreça desproporcionalmente itens ou conteúdos já populares, mas sim promova uma distribuição mais equitativa das recomendações (ZEHLIKE et al., 2017).

Em sua essência, o algoritmo FA*IR utiliza métodos avançados de análise de dados e probabilidade bayesiana aplicada para ponderar as influências de diferentes itens ou conteúdos, levando em consideração não apenas sua popularidade atual, mas também fatores como diversidade, relevância e impacto sobre os usuários. Dessa forma, busca-se promover uma recomendação mais justa e imparcial, que leve em conta as preferências individuais dos usuários sem reforçar injustiças pré-existentes.

O algoritmo FA*IR possui uma ampla gama de aplicações em diferentes domínios, incluindo sistemas de recomendação de produtos, conteúdos digitais, mídias sociais e muito mais. Em ambientes onde o viés de popularidade pode distorcer significativamente as recomendações, o FAIR se destaca como uma ferramenta poderosa para garantir uma experiência mais equitativa e personalizada para os usuários (YANG et al., 2023; MELUCCI, 2024).

Por exemplo, em plataformas de *streaming* de vídeo, o FA*IR pode ser utilizado para recomendar uma variedade mais diversificada de filmes e séries aos usuários, em vez de simplesmente priorizar os títulos mais populares ou amplamente promovidos. Da mesma forma, em redes sociais, o algoritmo pode ajudar a amplificar vozes e conteúdos menos conhecidos, contribuindo para uma maior inclusão e representatividade na disseminação de informações e ideias (BERNARD; BALOG, 2023).

A escolha do algoritmo FA*IR como *baseline* em pesquisas sobre controle do viés de popularidade é motivada por diversas razões. Primeiramente, o FA*IR representa uma abordagem consolidada e bem estabelecida no campo, com uma sólida fundamentação teórica e evidências empíricas de sua eficácia em promover a equidade e a imparcialidade nas recomendações.

Além disso, o FA*IR oferece uma estrutura flexível e adaptável, que pode ser facilmente ajustada e customizada para atender às necessidades específicas de diferentes contextos e cenários de aplicação. Isso o torna uma escolha atraente para pesquisadores interessados em explorar novas estratégias e técnicas para controlar o viés de popularidade em diferentes sistemas e plataformas.

3.3 Recomendações auxiliadas por análise de rede

Goldberg (1992) apresentou em seu trabalho uma abordagem inovadora no campo da filtragem colaborativa, destacando o Tapestry, um sistema experimental desenvolvido no Centro de Pesquisa da Xerox em Palo Alto. Sua pesquisa pioneira mostrou a relevância de incorporar decisões humanas no processo de filtragem, especialmente em ambientes complexos de comunicação, como o corporativo. O algoritmo proposto representou um marco ao fornecer soluções aplicáveis aos desafios crescentes enfrentados em redes de comunicação empresarial, oferecendo uma estrutura eficaz para o gerenciamento de e-mails corporativos (GOLDBERG, 1992).

A década de 2000 testemunhou uma expansão significativa no uso de filtros colaborativos, com Linden (2003) destacando sua eficácia no contexto da Amazon. Ao empregar tais filtros para recomendações baseadas em similaridades item-item, a Amazon estabeleceu um diferencial competitivo marcante em seu vasto mercado online. A capacidade desses filtros de atender às necessidades de uma ampla gama de usuários, cada um com perfis de consumo distintos, permitiu uma experiência de navegação simplificada e precisa em seu extenso catálogo de produtos (LINDEN, 2003).

Em 2003, Goh foi um dos precursores ao explorar a interseção entre recomendações e parâmetros de redes complexas. Sua análise do grau de centralidade de intermediação lançou luz sobre a influência de autores em redes sociais de publicações, como redes de coautoria. A introdução dessa métrica proporcionou insights valiosos que poderiam ser aproveitados para fornecer informações relevantes aos usuários, exemplificando a crescente interdisciplinaridade entre ciência da computação e ciências sociais (GOH, 2004).

Alguns dos primeiros estudos que vincularam o PageRank a sistemas de recomendação centraram-se em conjuntos de dados como o MovieLens e páginas da web, construindo grafos a partir da interação entre usuários e itens (ZHANG; ZHANG; LI, 2008; SU; KHOSHGOFTAAR, 2009). O objetivo era posicionar os nós da rede de modo a influenciar o ranqueamento dos itens nas listas de recomendação, em uma abordagem semelhante à sua aplicação no mecanismo de busca do Google. Desde então, diversas pesquisas exploraram outras aplicações do PageRank em sistemas de recomendação, modificando a estrutura dos grafos onde ele é aplicado, como em grafos direcionados, ou considerando sua interação com outros métodos de ranqueamento e categorização, como sistemas de tags (KIM; SADDIK, 2011; JIANG; WANG, 2010).

No ano de 2015, Sora desenvolveu um sistema de recomendação para identificar classes em sistemas de software, baseado no PageRank (SORA, 2015). Ao representar o sistema como uma rede complexa, as classes-chave podem ser mapeadas por meio de nós interconectados em grafos direcionados. Assim, o autor foi capaz de atribuir pesos

utilizando o algoritmo de PageRank, gerando recomendações rápidas e precisas para os usuários. Trabalhos mais recentes também têm explorado o uso dos valores de PageRank como um mecanismo de feedback adicional ou como um parâmetro para a clusterização, visando tornar as recomendações mais precisas ou reduzir vieses (FANANY, 2017; GUO, 2017; LU, 2017).

Ai, por exemplo, busca nas informações obtidas por meio de redes complexas novos parâmetros para otimizar filtros colaborativos. De acordo com o que foi observado por Ai, filtros colaborativos padrão buscam realizar ranqueamentos baseados apenas em pontos de similaridade, e isto gera uma perda de acuracidade, já que nós muito fortes tendem a gerar ruído, sendo recomendados a usuários que podem estar em uma "jornada´´ em um outro cluster da rede. Para aumentar sua acuracidade, um algoritmo que utiliza as avaliações da rede para construir grafos em clusters e aplicar o grau de centralidade dos nós como um fator de atenuação dos nós mais fortes quanto mais distante ele estiver do item avaliado, por meio da seguinte equação (AI, 2020):

Um dos avanços significativos na área da recomendação de sistemas é evidenciado por estudos recentes, como o de Ai (2020), que propõe uma abordagem inovadora ao explorar informações provenientes de redes complexas para aprimorar os parâmetros dos filtros colaborativos. Tradicionalmente, os filtros colaborativos baseiam-se exclusivamente na similaridade entre os pontos, o que muitas vezes resulta em perda de precisão. Isso ocorre principalmente devido à presença de nós muito influentes, que tendem a introduzir ruído nos ranqueamentos. Além disso, usuários com interesses específicos podem estar situados em clusters distintos da rede, o que torna essencial uma consideração mais ampla dos relacionamentos (AI, 2020).

Nesse contexto, Ai propõe um algoritmo que extrai informações da estrutura da rede para construir grafos em clusters e utiliza o grau de centralidade dos nós como um fator de atenuação para os nós mais influentes. Essa estratégia visa aprimorar a precisão das recomendações, especialmente quando os nós influentes estão distantes do item avaliado. A abordagem proposta por Ai mostrou resultados promissores, fornecendo insights valiosos para o desenvolvimento de sistemas de recomendação mais robustos e eficazes.

A formulação proposta pelo autor pode ser expressa pela equação abaixo, em que r'_{ui} representa a nota prevista do usuário u para o item i, r_i denota a média das classificações do item i, r_{uj} é a classificação do usuário u para o item j, s_{ji} a similaridade entre os itens j e i, ϕ um fator de atenuação e $S(j)$ indica a centralidade do nó j na rede:

$$r'_{ui} = \overline{r_i} + \frac{\sum_{j=1}^{n}[(r_{uj} + \overline{r_j}) * s_{ji} * \phi * S(j)]}{\sum_{j=1}^{n} s_{ji}} \tag{3.6}$$

Com base no mesmo racional, viu-se a oportunidade de testar o princípio de redução de viés ao se utilizar o PageRank com um dos parâmetros de entrada para o EQNet, tendo em vista a sua similaridade com a centralidade. E para isso a estrutura de rede complexa criado por Ai (descrita na Figura 3) foi utilizada como modelo ao determinar os itens como nós e a interação dos usuários como aspectos de interação entre os nós.

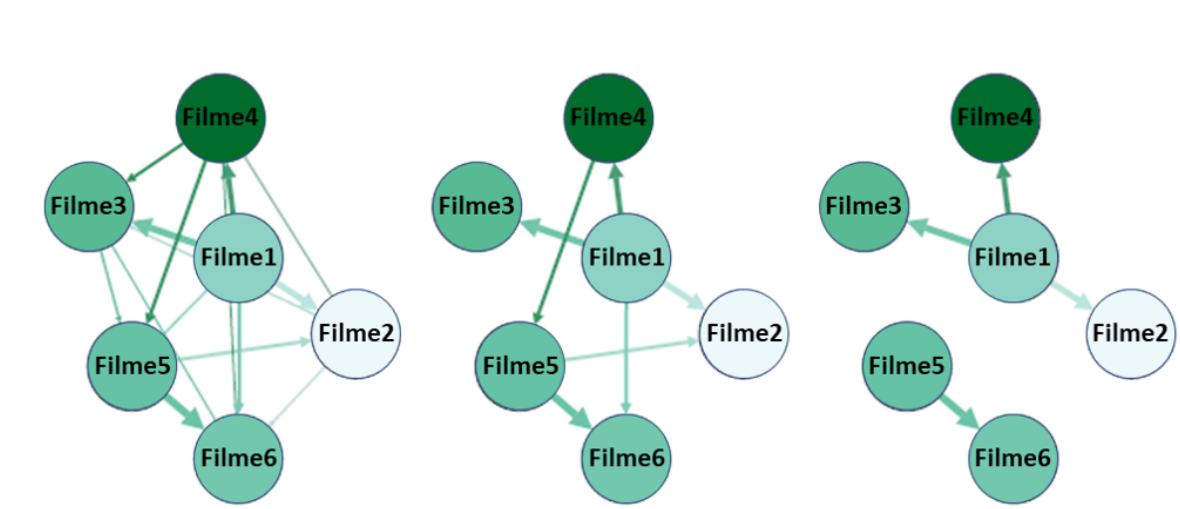

Figura 3 – Da esquerda para a direita, o peso das arestas válidas aumenta, reduzindo assim o número de ligações aparentes. A semelhança entre os nós é obtida a cada par de filmes. As semelhanças entre filmes são tratadas como ligação e influencia neste peso, construindo uma rede item-item. Esta semelhança varia entre um e zero e é distinta entre todos os nós.

Os métodos e abordagens analisados, como o FA*IR, fornecem um contexto valioso para a inovação proposta pelo EQNet. Este alinhamento teórico com as práticas existentes reforça a relevância e a aplicabilidade da nossa abordagem, alinhando-se com o objetivo de oferecer uma solução competitiva e eficiente.

4 ABORDAGEM PROPOSTA - EQNET

O EQNet é uma abordagem de reclassificação que utiliza valores de popularidade computados por algoritmos de classificação para reclassificar os itens de uma lista de recomendação. Entendemos que incluir valores de popularidade é crucial para este tipo de problema, já que a presença de viés de popularidade está relacionada com a distinção entre itens populares e cauda longa (ABDOLLAHPOURI, 2020; MALDONADO, 2021). A idéia de desenvolver uma abordagem que utiliza valores intrínsecos associados à popularidade dos item para sua reclassificação está alinhado com pesquisas existentes no campo e nos permite adotar uma abordagem para explorar possibilidades adicionais. Selecionamos os algoritmos Popularity Count e PageRank para avaliar a popularidade do item devido à sua eficiência em identificar e destacar itens relevantes dentro de uma coleção e sua importância na área. Ao empregar uma nova abordagem, conduzimos experimentos com esses algoritmos bem estabelecidos para explorar sua eficácia em mitigar o viés de popularidade.

Para calcular os valores de saída do Popularity Count foi aplicada a fórmula como descrita na Equação 3.6. Já, para o PageRank foi necessário desenvolver uma arquitetura de rede complexa com os dados disponíveis. Com base no modelo construído por (JIANG; WANG, 2010), foi possível adaptar uma rede com os filmes representado os nós e a jornada do usuário indo de um filme ao outro ao longo do tempo representando os vértices, como mostrado na Figura 5 do Tópico 5.3. Com esta estrutura de rede foi possível calcular um valor de PageRank para cada filme e próximo passo foi definir como usar estes valores.

Para efetuar o cálculo do PageRank, foi necessário construir uma rede que refletisse da melhor forma possível o comportamento dos usuários ao navegarem pelos itens do portfólio. Diante disso e seguindo parte do racional do trabalho de Sultany optou-se por adotar um grafo direcional, no qual os nós são representados pelos filmes disponíveis, enquanto as arestas direcionadas representam a sequência de interações dos usuários com os filmes, determinada pela ordem cronológica das avaliações realizadas por cada um dos usuários (AL_SULTANY; GHAIDAA, 2022). Portanto, conforme ilustrado na Figura 6, as distintas trajetórias dos usuários ao interagirem com os filmes delineiam a estrutura da rede a ser analisada. Os dados utilizados para construir essa estrutura de grafo são apresentados na Tabela 6.

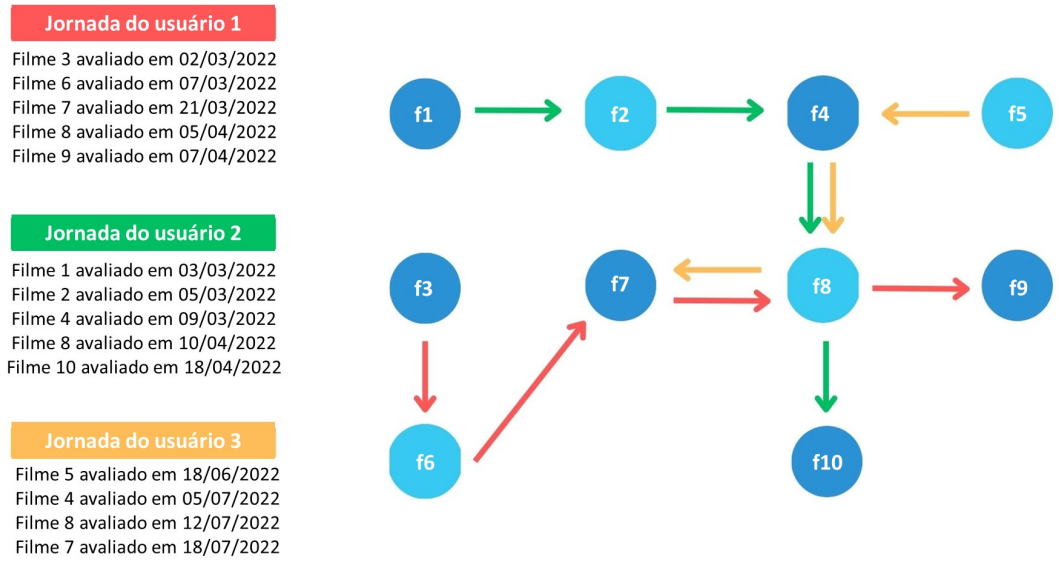

Figura 4 – Alguns nós da Rede Complexa para ilustrar a estrutura usada para calcular o PageRank usando os dados de filmes avaliados. Especificamente, neste snapshot, os usuários 1 2 e 3 avaliaram um set de 10 filmes hipotéticos ao longo de 2022.

A abordagem EQNet pode operar tanto com os valores de saída do PageRank quanto do Popularity Count como a variável principal de reclassificação α_i, necessitando de um processo de normalização prévio para comprimir os valores entre 0 e 1. Na Equação 4.1, apresentamos o processo de transformação realizado pelo EQNet, onde a nota da classificação original S_{ai} é convertida em uma nota S_{bi}, pronta para a reclassificação. Incluir um parâmetro λ nos permitiu um controle refinado sobre a intensidade do procedimento de reclassificação. Isto nos permitiu ajustar o grau em que as recomendações são influenciadas por considerações de equidade, garantindo que as recomendações resultantes equilibrem relevância e mitigação de viés.

A equação central do EQNet ficou definida da seguinte maneira:

$$S_{bi} = S_{ai} * (\frac{1 - \alpha_i * \lambda}{1 - \frac{\sum_{i=0}^{n} \alpha_i, i \in R}{n} * \lambda})^t \tag{4.1}$$

Com esta abordagem foi possível reequilibrar as listas de recomendação, especificamente diminuindo a proeminência de itens altamente populares e centrais, enquanto elevamos a relevância de itens de cauda longa. A Figura 5 ilustra como a recomendação padrão funciona após a classificação SVD, e a Figura 6 ilustra a dinâmica do processo de

reclassificação quando o EQNet é aplicado. Essencialmente, as alterações induzidas pela redução na popularidade de certos itens e o aumento em outros levam à substituição de itens populares previamente dominantes por itens de cauda longa que exibem uma forte afinidade do usuário, aumentando assim a diversidade e consequentemente reduzindo o viés de popularidade.

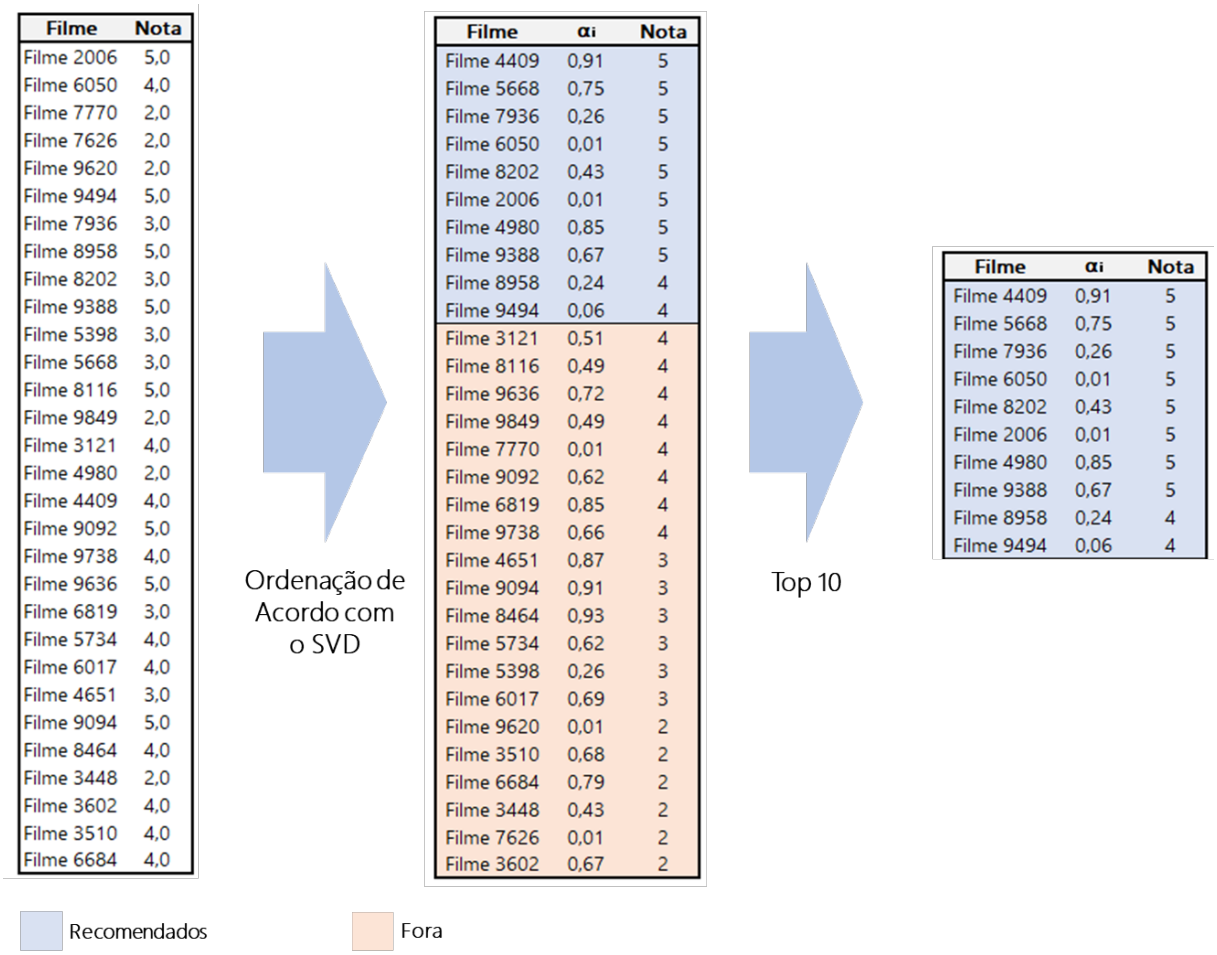

Figura 5 – Representação da recomendação padrão SVD de uma lista dos 10 melhores itens de um universo contendo 30 itens.

A descrição do EQNet neste capítulo apresenta os fundamentos e a metodologia da abordagem, destacando seu funcionamento. Com um foco claro na redução do viés de popularidade e na manutenção da qualidade das recomendações, o EQNet se posiciona como uma alternativa aos métodos de controle de viés de popularidade. Esta proposta está diretamente alinhada com o objetivo de buscar fornecer uma solução prática e eficaz para os desafios envolvendo viés de popularidade.

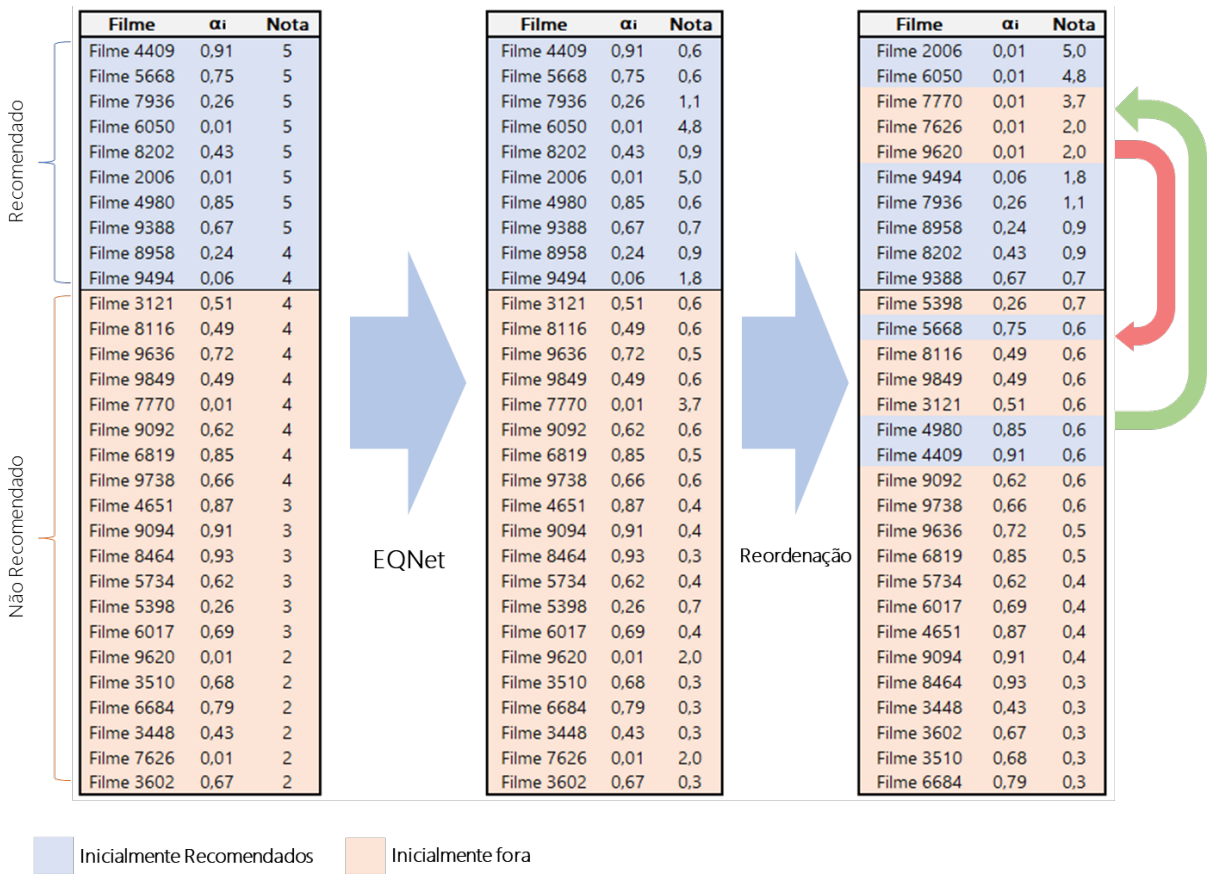

Figura 6 – Representação da abordagem EQNet para uma lista de recomendação dos top-10 de um universo com 30 itens já classificados pelo SVD. Isso envolve o processo de reclassificação usando o escore de popularidade na coluna (α_i) para ajustar sistematicamente os escores dos itens inversamente aos seus níveis de popularidade.

5 PROJETO E DESENVOLVIMENTO

Para a construção e execução dos algoritmos e da aplicação do EQNet foi utilizado um computador pessoal, com processador AMD Ryzen 5 3600x com 3.8GHz e 32GB de memória DDR4 com 3.6GHz, o que mostra sua capacidade de performar até em máquinas mais simples. Os códigos foram implementados em Python rodando na IDE do Jupyter Notebook em um sistema operacional Windows 10. Os parâmetros estocásticos de execução para o EQNet foram testados mantendo suas propriedades aleatórias, para não corroborar com a tese de que os resultados pudessem estar enviesados.

A aplicação proposta foi submetida a testes utilizando dois conjuntos de dados públicos e foi comparada ao *baseline*. O primeiro conjunto de dados é oriundo do MovieLens, o qual compreende 7.549.007 avaliações anônimas de aproximadamente 8.327 filmes, realizadas por 97.979 usuários registrados no sistema MovieLens (HARPER; KONSTAN, 2015; ACADEMIC TORRENTS, 2016). O segundo conjunto de dados utilizado foi o Prêmio Netflix, o qual foi empregado em diversos desafios de algoritmos de recomendação (KOREN, 2009; ACADEMIC TORRENTS, 2009). A partir deste segundo conjunto de dados, foram obtidas um total de 7.486.850 classificações, fornecidas por 436.544 usuários, referentes a 1.500 filmes distintos. A escolha desses conjuntos de dados se deu pela sua relevância e utilização frequente na literatura acadêmica e em desafios de pesquisa na área de recomendação de conteúdo.

Para aprimorar a qualidade da análise dos dados, adotou-se o procedimento de redução de dados proposto por Abdollahpouri et al. (2017), o qual consiste na exclusão de usuários com menos de 20 avaliações do conjunto de dados. Essa abordagem foi adotada pois observou-se que usuários mais ativos tendem a interagir com itens de cauda longa com maior frequência. Dessa forma, os usuários mantidos após a filtragem são aqueles com uma maior probabilidade de terem interagido com itens de cauda longa. Após a aplicação desse procedimento, o conjunto de dados do MovieLens reduziu-se a 55.900 usuários que avaliaram 8.315 filmes, totalizando 7.098.574 avaliações, o que representa uma redução de aproximadamente 6%. Ao aplicar os mesmos critérios ao conjunto de dados da Netflix, o número de usuários diminuiu para 121.470, os quais avaliaram 1.500 filmes, totalizando 5.388.586 avaliações, uma redução de cerca de 28%.

	user_id	Score	Date_x	movie_id	id	Date_y	Name
0	1488844	3	2005-09-06	1	75761	2003.0	Dinosaur Planet
1	822109	5	2005-05-13	1	342571	2003.0	Dinosaur Planet
2	885013	4	2005-10-19	1	352293	2003.0	Dinosaur Planet
3	30878	4	2005-12-26	1	262389	2003.0	Dinosaur Planet
4	823519	3	2004-05-03	1	342787	2003.0	Dinosaur Planet
...	...	...	...	...	...	...	...
2798699	651950	4	2005-06-28	500	315951	2002.0	Mail Call: The Best of Season 1
2798700	924510	3	2005-07-12	500	358415	2002.0	Mail Call: The Best of Season 1
2798701	965381	3	2005-08-19	500	364751	2002.0	Mail Call: The Best of Season 1
2798702	822391	1	2004-11-04	500	342612	2002.0	Mail Call: The Best of Season 1
2798703	2082054	3	2005-06-22	500	167686	2002.0	Mail Call: The Best of Season 1

2798704 rows × 7 columns

Tabela 1 – Recorte da tabela de filmes da Netflix contendo 500 filmes.

Com a aplicação e os dados definidos, o próximo passo foi definir o parâmetro λ para o experimento. Nesta etapa foram observados os maiores valores de α, que representa os valores de saída para os algoritmos de ranqueamento por popularidade (PageRank e Popularity Count), de modo que α*λ tivesse valores próximos a 0,1 para gerar variações de pelo menos 10% nas avaliações dos casos mais extremos de popularidade. Esta resultante entre α e λ permite administrar o quanto se deseja atenuar a importância de itens mais populares para forçar sua reordenação.

Para obter uma compreensão abrangente de cada método empregado em nossos experimentos, conduzimos o experimento utilizando dez valores distintos para o parâmetro λ (de 1 a 10), possibilitando uma compreensão mais abrangente. No caso do cenário de referência, avaliamos o impacto na qualidade das recomendações do algoritmo FA*IR ao testá-lo com diversos valores para a proporção do parâmetro de candidatos protegidos p e observamos sua influência na redução do viés de popularidade. Além disso, avaliamos o desempenho do EQNet como um fator de pós-processamento para o algoritmo FA*IR, investigando possíveis sinergias entre as duas soluções.

O algoritmo de recomendação utilizado para teste, foi escolhido o SVD. Por se tratar de um modelo clássico, presente em bibliotecas abertas para teste e eficiente para trabalhos em bases esparsas. A versão utilizada foi a da biblioteca Python Scipy, para decomposição de valores singulares de matrizes esparsa (VIRTANEN et al., 2020).

Ao realizar a recomendação pelo método de SVD, foi preciso transformar as listas em tabelas de usuários vs. filmes, contendo as notas que estes usuários deram. Para isto, os dados de filmes a serem avaliados no lote foram indexados da seguinte forma. 72% da base foi designada para treino e 18% para calibração e 10% para teste.

		Filmes									
		1	2	3	4	...	496	497	498	499	500
Usuários	1	0	0	0	0		0	2	0	0	5
	2	4	0	3	0		0	2	0	5	0
	3	0	0	0	0		0	5	0	0	0
	4	0	0	0	0		4	0	4	3	0
	5	0	0	5	0		0	0	0	4	0
	6	0	0	0	4		0	0	0	4	0
	...										
	26032	0	4	0	0		0	0	0	3	0
	26033	3	0	5	0		4	0	0	0	0
	26034	0	0	4	0		3	0	0	0	0
	26035	0	0	5	0		5	0	0	0	0
	26036	0	0	0	0		0	0	0	0	5
	26037	0	0	4	0		0	0	0	0	0
	26038	0	4	0	0		0	0	5	0	0

Tabela 2 – Matriz USUÁRIO vs. FILME preparada para ser decomposta pelo SVD e gerar aproximações

5.1 Métricas de avaliação

Para testar o modelo, quatro métricas e suas variações foram calculadas e armazenadas para cada valor de λ nos experimentos. Os dez experimentos foram compostos por cinco experimentos em cada uma das duas bases (Netflix e MovieLens) sendo eles: EQNet com PageRank, EQNet com Popularity Count, FA*IR, FA*IR + EQNet com PageRank e FA*IR + EQNet com Popularity Count. Para medir o viés de popularidade e o comportamento geral do sistema de recomendação simulado em relação ao *baseline*, quatro medidas foram tomadas para entender o impacto da abordagem no viés de popularidade e na recomendação como um todo:

- **Average Recommendation Popularity (ARP):** Esta métrica mede a média de popularidade dos itens em cada lista. Ela pode ser descrita pela fórmula 4.1, com U_t representando o número total de usuários, L_u o número total de itens em uma lista de recomendações e Φ o número total de vezes que o item i foi avaliado no treinamento (YIN et al., 2012).

$$ARP = \frac{1}{|U_t|} \sum_{u \in U_t} \frac{\sum_{i \in L_u} \Phi(i)}{|L_u|} \tag{5.1}$$

- **Porcentagem Média de Itens de Cauda Longa (APLT):** Conforme usado anteriormente por (ABDOLLAHPOURI, 2019), esta métrica é usada para medir o

valor médio percentual de itens de cauda longa presentes nas listas de recomendações. Nela, Γ representa o grupo de itens de cauda longa.

$$APLT = \frac{1}{|U_t|} \sum_{u \in U_t} \frac{|i, i \in (L_u \cap \Gamma)|}{|L_u|} \tag{5.2}$$

- **Cobertura Média de Itens de Cauda Longa (ACLT):** Outra métrica usada por (ABDOLLAHPOURI, 2019), para avaliar quanta exposição os itens de cauda longa estão recebendo por meio da recomendação. Ela possibilita uma visão complementar ao APLT, pois mostra a quão diversa é essa recomendação e expõe se a recomendação está listando sempre o mesmo conjunto de itens de cauda longa. Nele $1(i \in \Gamma)$ é igual a 1 em quando i está em Γ.

$$ACLT = \frac{1}{|U_t|} \sum_{u \in U_t} \sum_{i \in L_u} 1(i \in \Gamma) \tag{5.3}$$

- **Ganho cumulativo com desconto normalizado (NDCG):** É uma métrica utilizada para avaliar a qualidade das recomendações, uma vez que não basta a recomendação ter apenas um grau de acuraciade elevado e sim o quão relevante são as recomendações para os usuários. Ela pode ser traduzida como a razão entre o Ganho Acumulado com Desconto (DCG) e o Ganho Acumulado com Desconto Ideal (IDCG) (ABDOLLAHPOURI, 2019).

Além das métricas previamente mencionadas, foram realizados cálculos de duas outras métricas conhecidas como RMSE (Root Mean Square Error) e MAE (Mean Absolute Error). Essas métricas foram empregadas na avaliação das recomendações iniciais geradas pelo algoritmo de SVD (Singular Value Decomposition), visando assegurar a qualidade satisfatória dessas recomendações em um estágio inicial do processo de recomendação. Nestas métricas n é o número total de observações, y_i é o valor observado e y'_i é o valor previsto pelo modelo.

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(y_i - y'_i)^2}{n}} \tag{5.4}$$

$$ACLT = \frac{\sum_{i=1}^{n} |y_i - y'_i|}{n} \tag{5.5}$$

Utilizando essas métricas, foi possível observar como o EQNet impacta o resultado final da recomendação quando comparado e também trabalhando junto a outras abordagens da literatura. Além disso, coletando dados a cada repetição do EQNet, foi possível identificar comportamentos diferentes em cada uma das abordagens para cada perfil de *dataset* ao longo do experimento. Dessa forma, a estruturação do experimento pode ser observada na Figura 7.

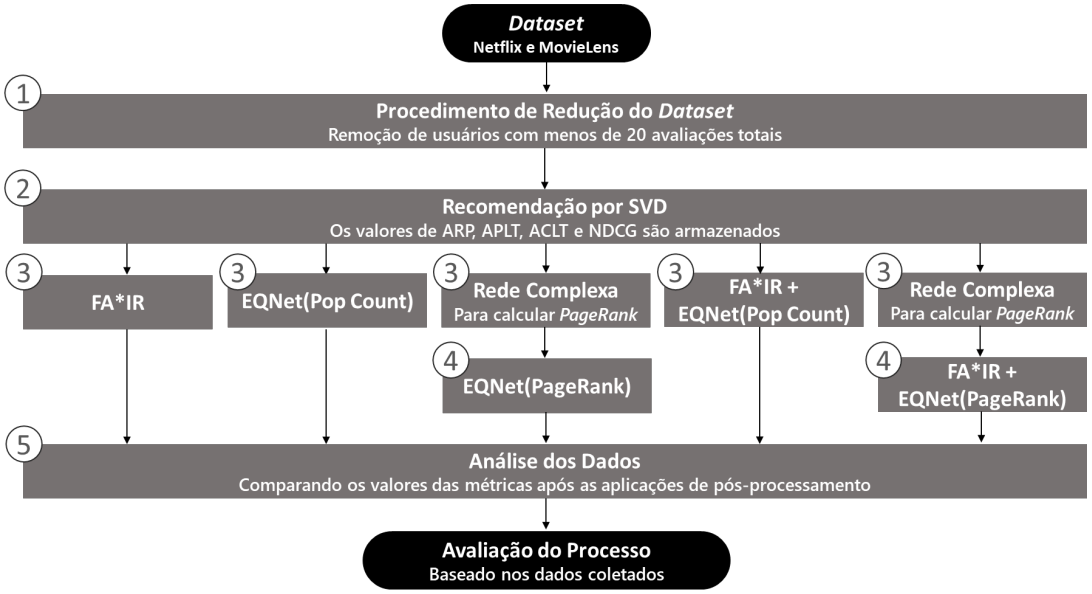

Figura 7 – Fluxograma sumarizado com as principais etapas experimentais do processo.

5.2 Aplicando o EQNet

O cálculo do valor de PageRank foi realizado para todos os nós da rede, adotando um fator de amortecimento de 0,9, inicialização aleatória e pesos iniciados com o valor unitário. Posteriormente, com os valores de PageRank obtidos para cada nó, procede-se à criação de uma tabela que consolida as avaliações e os valores de PageRank associados a cada filme, conforme apresentado na Tabela 5. É importante ressaltar que o PageRank é um atributo intrínseco de cada filme, não sofrendo variações em função do usuário, da data ou de qualquer outro atributo associado que não pertença ao filme em si.

O processo de valoração foi igualmente aplicado ao Popularity Count, conforme

	user_id	Score	Date_x	movie_id	id	Date_y	Name	pagerank
0	1488844	3	2005-09-06	1	75761	2003.0	Dinosaur Planet	0.001996
1	822109	5	2005-05-13	1	342571	2003.0	Dinosaur Planet	0.001996
2	885013	4	2005-10-19	1	352293	2003.0	Dinosaur Planet	0.001996
3	30878	4	2005-12-26	1	262389	2003.0	Dinosaur Planet	0.001996
4	823519	3	2004-05-03	1	342787	2003.0	Dinosaur Planet	0.001996
...	...	...	...	...	...	...	...	...
2798699	651950	4	2005-06-28	500	315951	2002.0	Mail Call: The Best of Season 1	0.001912
2798700	924510	3	2005-07-12	500	358415	2002.0	Mail Call: The Best of Season 1	0.001912
2798701	965381	3	2005-08-19	500	364751	2002.0	Mail Call: The Best of Season 1	0.001912
2798702	822391	1	2004-11-04	500	342612	2002.0	Mail Call: The Best of Season 1	0.001912
2798703	2082054	3	2005-06-22	500	167686	2002.0	Mail Call: The Best of Season 1	0.001912

2798704 rows × 8 columns

Tabela 3 – Tabela de atributos por filme incluindo os valores de PageRank

evidenciado na Tabela 6. Atribuindo valores tanto ao Popularity Count quanto ao PageRank para cada filme, assim o EQNet pode recalcular a pontuação dos filmes nas listas de recomendação do SVD, resultando na reordenação das recomendações. O algoritmo FAIR foi adotado como referência, otimizado para k = 10 (os 10 principais elementos a serem retornados), p = 0,95 (garantindo 95% de preservação dos elementos classificados) e alpha = 0,1 (indicando o nível de significância). Nos experimentos híbridos (FAIR + EQNet), optou-se por primeiramente reordenar utilizando o FA*IR, seguido pelo EQNet. Todos estas saídas de reordenação tiveram seus valores de ARP, ACLT, APLT e NDCG salvos e comparados com os valores originais.

	user_id	movie_id	score	date	movie_id_m	title	genres	pop_count
0	26299	1309	2.5	2005-05-07 02:09:45	1369	I Can't Sleep (J'ai pas sommeil) (1994)	Drama|Thriller	0.126194
1	21819	1309	3.0	1997-03-15 20:42:58	1369	I Can't Sleep (J'ai pas sommeil) (1994)	Drama|Thriller	0.126194
2	26956	1309	3.0	2005-03-22 06:58:14	1369	I Can't Sleep (J'ai pas sommeil) (1994)	Drama|Thriller	0.126194
3	34769	1309	2.0	1998-02-26 16:15:15	1369	I Can't Sleep (J'ai pas sommeil) (1994)	Drama|Thriller	0.126194
4	40040	1309	0.5	2003-06-28 02:47:29	1369	I Can't Sleep (J'ai pas sommeil) (1994)	Drama|Thriller	0.126194
...	...	...	...	...	...	...	...	...
5678854	48074	3201	3.0	2000-07-25 18:55:53	3376	Fantastic Night, The (Nuit fantastique, La) (1...	Romance	0.000034
5678855	46330	3083	3.0	2000-04-24 18:52:49	3255	League of Their Own, A (1992)	Comedy|Drama	0.000034
5678856	15808	7487	3.5	2004-05-15 18:56:36	8459	Heiress, The (1949)	Drama|Romance	0.000069
5678857	41476	6486	1.5	2017-02-28 10:54:16	6714	So Close (Chik Yeung Tin Sai) (2002)	Action|Comedy|Romance|Thriller	0.000034
5678858	38566	6637	3.0	2004-08-10 03:14:15	6866	Nine Dead Gay Guys (2003)	Comedy|Crime	0.000069

5678859 rows × 8 columns

Tabela 4 – Tabela de atributos por filme incluindo os valores de Popularity Count

O processo de desenvolvimento e implementação do EQNet, detalhado neste capítulo, demonstra a viabilidade técnica da nossa abordagem. A utilização de dados reais e a aplicação de métricas robustas garantem que o EQNet não apenas atenda aos requisitos experimentais, mas também esteja preparado para aplicações práticas. Este alinhamento

técnico com os objetivos do trabalho assegura que a abordagem seja replicada e validada em diferentes contextos, reforçando sua relevância e aplicabilidade.

6 RESULTADOS

Os resultados de cada experimento foram representados graficamente em três gráficos distintos, cada um comparando as variações do NDCG em relação ao ARP, ACLT e APLT, respectivamente, para ambas as bases de dados. Além disso, os dados foram tabulados para proporcionar uma visualização mais clara. Cada um dos dez experimentos é analisado e, posteriormente, uma análise é apresentada, visando sintetizar as evidências coletadas e comparar os resultados entre si.

Os gráficos gerados a partir dos resultados obtidos com a aplicação do EQNet permitem uma melhor visualização do comportamento das métricas de popularidade ao longo da variação de qualidade da recomendação. Cada experimento consistiu em nove iterações do processo de reclassificação, executadas para diferentes valores de λ, visando à obtenção de dados abrangentes que permitissem capturar a dinâmica da redução de viés aplicada, buscando, assim, encontrar um ponto de equilíbrio entre a maximização da variação e a minimização da perda de NDCG.

6.1 Experimento base com SVD

Após uma filtragem inicial para trazer apenas usuários mais relevantes da base do MovieLens (que tivessem avaliado ao menos 20 filmes), o conjunto de dados do ficou com 55.900 usuários, 8.315 filmes e um total de 7.098.574 avaliações. Aplicando os mesmos critérios para redução no número de usuários para a base da Netflix, o total de entradas foram reduzidas a 5.388.586 classificações dadas por 121.470 usuários a 1.500 filmes, representando uma redução de aproximadamente 28% da base. Inicialmente foi realizado um processo de SVD da biblioteca Suprise, para criar uma decomposição da matriz usuário/filme com seus valores default. Para garantir que a recomendação teria uma boa acuracidade e que as variações de performance não deveriam ao fato da recomendação ser fraca, medimos o RMSE da recomendação inicial para as duas bases. Para a base do MovieLens obtivemos um RMSE de 0.8421 e MAE de 0.6446 e a base Netflix um RMSE de 0.9066 e MAE de 0.7080.

Após uma filtragem inicial visando selecionar apenas os usuários mais relevantes da base do MovieLens, definidos como aqueles que tenham avaliado ao menos 20 filmes,

obtivemos um conjunto de dados composto por 55.900 usuários, 8.315 filmes e um total de 7.098.574 avaliações. Da mesma forma, aplicamos os mesmos critérios de seleção para redução do número de usuários na base da Netflix, resultando em um conjunto de dados contendo 5.388.586 avaliações atribuídas por 121.470 usuários a 1.500 filmes. Inicialmente, empregamos uma técnica de SVD utilizando a biblioteca Surprise, com valores padrões de entrada para facilitar a replicação do experimento, a fim de criar uma representação da matriz usuário/filme. Para assegurar que as recomendações alcançassem um nível adequado de acurácia e que eventuais variações de desempenho não fossem decorrentes de recomendações deficientes, calculamos o Root Mean Square Error (RMSE) e o Mean Absolute Error (MAE) para as recomendações iniciais em ambas as bases de dados. Para a base do MovieLens, os valores obtidos foram RMSE = 0.8421 e MAE = 0.6446, enquanto para a base da Netflix, os valores correspondentes foram RMSE = 0.9066 e MAE = 0.7080. É importante ressaltar que RMSE e MAE foram adotados como métricas principais para avaliar o desempenho dos modelos de recomendação.

Após a recomendação, os valores para as ARP, APLT, ACLT e NDCG foram coletados. Os resultados obtidos foram:

Dataset	**NDCG Original**	**ARP Original**	**APLT Original**	**ACLT Original**
MovieLens	0.226	0.155	0.506	5.059
Netflix	0.208	0.344	0.669	6.686

Tabela 5 – Tabela com os valores originais de NDCG, ARP, APLT e ACLT para a recomendação SVD usando as bases do MovieLens e Netflix.

No âmbito desta pesquisa, é importante considerar que as técnicas de fatoração de matrizes, apesar de sua eficácia comprovada, estão intrinsecamente ligadas à disponibilidade e qualidade dos dados de treinamento, que estaria em seu primeiro ciclo. Um desempenho inicialmente inferior em métricas como o NDCG pode ser atribuído à fase inicial de aprendizado do sistema de recomendação. Durante esse período, o modelo está se adaptando aos padrões de preferências dos usuários com base nos dados fornecidos. Quando os dados de treinamento são escassos ou não representativos o suficiente, o modelo pode falhar em capturar adequadamente a complexidade das preferências dos usuários, resultando em recomendações menos precisas e, consequentemente, em valores mais baixos de NDCG (KOREN, 2009). Adicionalmente, a avaliação de desempenho do modelo não se baseou na operação real de um sistema de recomendação, mas sim na comparação de recomendações com o comportamento de escolha natural dos usuários, sem a presença de um sistema de recomendação.

Além disso, dois pontos reforçam que o valor obtido foi satisfatorio para o modelo inicial. O primeiro é que o teste de eficiência do modelo se deu não por meio de um sistema de recomendação em funcionamento mas sim por comparar uma recomendação

sobre o comportamento natural de escolha dos usuários, sem a presença de um sistema de recomendação mais robusto. Segundo, os valores de NDCG estão alinhados com aqueles relatados em estudos no campo ao utilizar SVD como técnica de recomendação em teste provenientes de dados offline (FERRARI; CREMONESI, 2022; VALCARCE et al., 2018).

As métricas de ARP, APLT e ACLT derivadas das recomendações por SVD destacam uma distribuição razoável de itens de cauda longa nas recomendações produzidas. A importância de começar com valores aceitáveis de participação e cobertura de itens de cauda longa é crucial, pois isso reflete semelhanças com sistemas convencionais, além de uma variação mais justa diante desse contexto. A manipulação de uma amostra com uma alta concentração de itens populares seria facilmente reduzida de maneira significativa, o que ressalta a necessidade de garantir a representatividade das recomendações geradas.

Partindo deste ponto, foram realizados 10 experimentos com as bases de teste e os valores de NDCG, ARP, APLT e ACLT foram comparados com os originais. Os gráficos das variações a cada época foram plotados e podem ser vistos nas próximas seções do trabalho.

O *baseline* escolhido para o experimento, como explicado anteriormente, foi a aplicação conhecida como FA*IR que, de acordo com o as análises presentes no trabalho de Zehlike de 2017, mostrou consistentemente que este algoritmo supera abordagens convencionais em termos de equidade e eficiência, proporcionando recomendações que são tanto relevantes quanto justas. Além disso, o algoritmo continua a ser atualizado tendo inclusive uma grande atualização em 2022 para atuar com múltiplos grupos protegidos (ZEHLIKE et al., 2017, 2022). Em nosso experimento aplicamos o fair em uma listagem de 100 elementos para reordenar apenas os top 10 itens com um nível de significância de 0.1 e nove proporções de proteção p (10%, 20%, 30%, 40%, 50%, 60%, 70%, 80% e 90%), para acompanhar a variação na performance. Com isso, para cada um dos nove valores de p foi possível obter a variação nas métricas aferidas inicialmente.

Para cada um dos outros quatro métodos, também foram registrados os valores de nove iterações em relação ao valor inicial. Ao final de cada iteração, os valores de ARP, APLT, ACLT e NDCG foram comparados com os valores obtidos pela recomendação SVD sem otimização, utilizando-se esse valor relativo para gerar a análise. Por exemplo, a variação de APLT de uma iteração t seria $\Delta APLT_t = \frac{APLT_t}{APLT_0}$, e esse cálculo foi realizado para cada iteração a fim de registrar a variação total em relação ao valor inicial. Para as abordagens utilizando o EQNet, foram utilizados nove valores diferentes de λ para capturar estas diferentes iterações.

Em nosso experimento, os valores de todas as quatro métricas-chave (ARP, APLT, ACLT e NDCG) foram registrados ao longo das nove iterações em relação aos seus valores iniciais para cada um dos cinco métodos: FAIR, EQNet com PageRank, EQNet com Po-

pularity Count, FAIR seguido por EQNet com PageRank e FAIR seguido por EQNet com Popularity Count. Por meio dessa análise iterativa, obtivemos insights sobre o comportamento do trade-off entre a qualidade da recomendação e a redução de viés. Além disso, ao comparar o desempenho das abordagens FAIR seguidas por EQNet com seus equivalentes isolados, pudemos discernir e quantificar o impacto específico do método combinado na abordagem de viés e melhoria da qualidade da recomendação.

6.2 Análise de Resultados

Para obter uma visão abrangente do impacto do EQNet nos resultados do Sistemas de Recomendação, foi realizada uma análise das variações entre as métricas de recomendação originais e os resultados pós-processados. No contexto do ARP, variações negativas são favoráveis, pois indicam uma redução dos itens populares geralmente presentes nas listas de recomendação. Além disso, foram examinadas métricas como APLT e ACLT, que medem o envolvimento e a cobertura de itens de cauda longa, respectivamente. Aqui, uma variação positiva mais alta indica um melhor desempenho. Com o NDCG, o objetivo foi minimizar perdas enquanto maximiza ganhos em outras variações. Essas análises fornecem insights sobre o impacto do método na gestão de viés de popularidade e na qualidade das recomendações.

As Tabelas 8 e 9 mostram o desempenho médio comparativo de cada abordagem em termos de variação percentual, com foco em APLT e ACLT, bem como ARP ao utilizar os algoritmos EQNet e FA*IR. Os resultados indicam que o EQNet supera o FA*IR na redução do ARP e simultaneamente melhora APLT e ACLT com apenas uma pequena redução no Ganho Cumulativo Normalizado por Desconto (NDCG). Além disso, ao combinar o EQNet com o FAIR, os experimentos apresentam resultados ainda mais promissores, alcançando melhorias adicionais em APLT e ACLT, embora com um custo modesto para o NDCG. Essas descobertas destacam a eficácia do EQNet como uma ferramenta poderosa para mitigar viés e aumentar a justiça em sistemas de recomendação, com potencial para utilização complementar ao lado do FA*IR para alcançar desempenho superior na otimização de múltiplas métricas de viés.

Método	ARP var. (%)	APLT var. (%)	ACLT var. (%)	NDCG var. (%)
FA*IR	-32.73	18.23	17.78	-1.39
EQNet (PageRank)	-79.39	44.93	44.93	-2.01
EQNet (PopCount)	-57.95	65.77	65.77	-2.53
FA*IR + EQNet (PageRank)	-54.18	44.62	44.62	-1.66
FA*IR + EQNet (PopCount)	-13.34	85.70	85.70	-1.91

Tabela 6 – Tabela que descreve a variação percentual das métricas para cada método usando o banco de dados MovieLens.

Conforme evidenciado na Tabela 8, os resultados obtidos nos experimentos realizados com o *dataset* do MovieLens revelaram-se promissores ao comparar o desempenho do EQNet com o FA*IR. Observou-se uma notável redução na quantidade de itens populares nas listagens, conforme indicado pela variação da métrica ARP, acompanhada por melhorias significativas na abrangência e cobertura dos itens de cauda longa nas recomendações, como mostrado pelas variações de APLT e ACLT. É relevante notar que esses avanços foram alcançados com um custo muito baixo em termos de variação do NDCG.

A análise dos experimentos envolvendo a combinação de FA*IR com EQNet sugere que a variação de ARP diminui ligeiramente, enquanto APLT e ACLT se mantêm consistentes quando o EQNet utiliza o PageRank como entrada, e são ainda mais acentuados ao empregar o Popularity Count como entrada. Esses resultados indicam a capacidade do EQNet de complementar o desempenho do FAIR, gerando outputs de interesse. Além disso, observou-se que a perda de NDCG nessas circunstâncias é menor do que quando se aplica apenas o EQNet. Esse achado sugere uma sinergia potencial entre as abordagens, proporcionando recomendações mais eficazes e mantendo a qualidade geral do ranking.

Métodos	ARP var. (%)	APLT var. (%)	ACLT var. (%)	NDCG var. (%)
FA*IR	-31.54	17.23	16.88	-1.37
EQNet (Pagerank)	-33.20	7.71	7.71	-0.06
EQNet (PopCount)	-18.51	12.96	12.96	-1.73
FA*IR + EQNet (PageRank)	-52.90	15.42	15.42	-0.20
FA*IR + EQNet (PopCount)	-51.09	14.19	14.19	-1.45

Tabela 7 – Tabela que descreve a variação percentual das métricas para cada método usando o banco de dados Netflix.

Os resultados mostrados na Tabela 9, sobre os experimentos realizados com o *dataset* da Netflix, também se revelaram quanto à performance do EQNet. Observou-se uma redução menos ascentuada na quantidade de itens populares, na abrangência e cobertura dos itens de cauda longa nas recomendações. No entanto é importante ressaltar que a variação de NDCG para o EQNet utilizando PageRank foi muito menor que a observada no FA*IR, lhe dando uma eficiência muito superior. Além disso, neste experimento, o funcionamento do EQNet como uma abordagem complementar ao FA*IR se mostrou muito superior aos métodos isolados, principalmente quando observamos os resultados do FA*IR + EQNet com PageRank.

A análise dos experimentos envolvendo a combinação de FA*IR com EQNet neste experimento se sugere que a o EQNet potencializa todos os parâmetros de controle do viés de popularidade podendo até reduzir a perda da qualidade da recomendação gerada pelo FA*IR. Esta melhora da performance da recomendação após o processamento do EQNet aponta que alguns algoritmos de popularidade sejam mais indicados para alguns perfils de plataforma com distribuições específicas de usuário/item. Além disso, atingindo

valores mais baixo de NDCG para o EQNet que o FA*IR, o EQNet tem potencial para não só auxiliar na redução do viés de popularidade e aumentar a recomendação de itens de cauda longa, como também controlar a perda da qualidade do NDCG.

Os resultados apresentados na Tabela 9, provenientes dos experimentos conduzidos utilizando o *dataset* da Netflix, oferecem uma visão detalhada da performance do EQNet. Observou-se uma redução menos pronunciada na presença de itens populares, bem como na abrangência e cobertura dos itens de cauda longa nas recomendações. No entanto, é importante destacar que a variação do NDCG para o EQNet, empregando o algoritmo PageRank, revelou-se significativamente menor em comparação com o FA*IR, evidenciando uma eficácia de variação substancialmente superior. Além disso, neste estudo, a integração do EQNet como uma abordagem complementar ao FA*IR mostrou resultados superiores em relação aos métodos isolados, especialmente quando consideramos os desempenhos do FA*IR + EQNet com PageRank.

A análise dos experimentos que combinaram o FA*IR com o EQNet sugere que este último potencializa todos os parâmetros de controle do viés de popularidade, podendo inclusive mitigar a degradação na qualidade das recomendações geradas pelo FA*IR. Essa melhoria na performance da recomendação após o processamento do EQNet indica que determinados algoritmos de popularidade podem ser mais adequados para diferentes perfis de plataforma, com distribuições específicas de usuários e itens. Ademais, ao alcançar valores inferiores de NDCG em relação ao FA*IR, o EQNet mostra potencial não apenas para auxiliar na redução do viés de popularidade e aumentar a recomendação de itens de cauda longa, mas também para controlar a degradação na qualidade do NDCG.

A análise dos resultados das Tabela 8 e 9 sugerem que a abordagem combinada de FA*IR e EQNet se destaca na tarefa de recomendação, proporcionando uma distribuição mais equilibrada entre itens populares e de cauda longa nas listagens. Essa eficácia pode ser atribuída à capacidade do EQNet de capturar nuances e relações complexas entre os itens, complementando assim a abordagem baseada em filtragem colaborativa do FA*IR. Além disso, a utilização de diferentes métricas de entrada, como o PageRank e o Popularity Count, permite uma adaptação flexível às características específicas do conjunto de dados, otimizando ainda mais o desempenho do sistema de recomendação. Essas observações destacam a importância de abordagens híbridas e adaptativas, como o EQNet, na construção de sistemas de recomendação robustos e eficientes para diversos cenários de aplicação.

As Figuras 10, 11 e 12 mostram a relação entre os valores coletados de NDCG com os de ARP, APLT e ACLT para os mesmos valores de λ. O gráfico apresenta uma correlação direta entre a queda de ARP e aumento de APLT e ACLT com a queda de NDCG, indicando uma queda constante na qualidade da recomendação a medida que mais itens

de cauda longa entram para as listas de recomendação. Isto cria um padrão muito claro de troca onde a aplicação do método para redução do viés de popularidade e cobertura do catálogo está intrinsecamente ligada a uma menor qualidade de recomendação.

6.2.1 Dados de evolução de ARP por queda de NDCG

A análise da queda de ARP em relação à redução do NDCG desempenha um papel fundamental na compreensão e no controle do viés de popularidade em sistemas de recomendação. Enquanto o NDCG mede a qualidade global das recomendações, incorporando a relevância dos itens recomendados para os usuários, o ARP quantifica a popularidade média dos itens recomendados. Portanto, ao observar a queda de ARP frente à queda de NDCG, é possível identificar se a redução na qualidade das recomendações está associada a uma diversificação mais eficaz dos itens recomendados, ou se simplesmente reflete uma mudança na distribuição de popularidade dos itens. Inicialmente, foi conduzida uma análise utilizando os dados provenientes do *dataset* do MovieLens, para iniciar a compreender os padrões de evolução do ARP em relação ao NDCG.

A Figura 8 ilustra as variações de NDCG por variações de ARP dos dados obtidos a partir das cinco metodologias usando o *dataset* do MovieLens. O primeiro gráfico da análise, mostra um intervalo específico da abordagem do EQNet utilizando PageRank como algoritmo de ranqueamento, no qual a correlação linear entre os pontos sugere uma relação entre a redução do NDCG e a queda no ARP. A análise sugere que mesmo uma variação relativamente pequena de apenas 2% no NDCG resultou em uma redução substancial de 79.5% no ARP. Essa observação destaca a sensibilidade do ARP às alterações na qualidade das recomendações, evidenciando que é possível diversificar a popularidade dos itens com o EQNet sem grandes impactos na qualidade da recomendação.

O segundo gráfico da análise, mostra um intervalo específico da abordagem do EQNet utilizando Popularity Count como algoritmo de ranqueamento e ilustra a variação dos dados obtidos a partir do experimento utilizando EQNet com contagem de popularidade. Neste gráfico de pontos, observamos um intervalo específico da abordagem, onde a correlação praticamente linear entre os pontos sugere uma relação direta entre a redução do NDCG e a queda do ARP. É notável que, mesmo com uma variação relativamente pequena de apenas 2.5% no NDCG, foi possível alcançar uma redução de até 73% no ARP, uma performance um pouco inferior a do EQNet utilizando PageRank. Essa análise ressalta a eficácia do EQNet em controlar o viés de popularidade, mesmo utilizando outros parâmetros de popularidade para reordenar os itens da lista de recomendação.

O terceiro gráfico apresentado mostra a variação dos dados obtidos a partir do experimento utilizando o *baseline* FA*IR. Este gráfico de pontos também exibe um intervalo específico da abordagem, no qual a correlação aparentemente hiperbólica entre os pontos

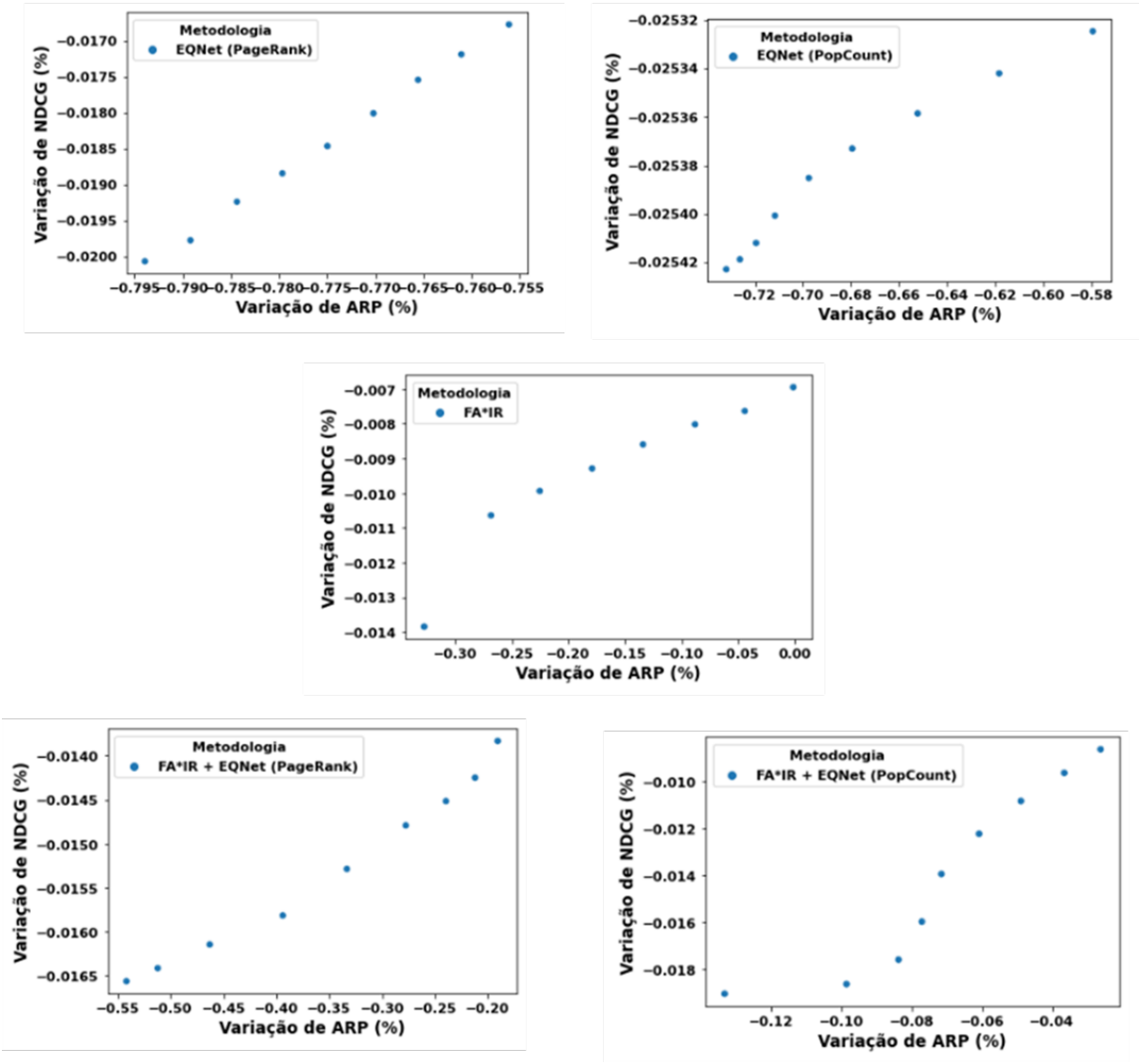

Figura 8 – Gráfico de variação de NDCG por variações de ARP para as cinco metodologias aplicadas, utilizando o *dataset* do MovieLens

sugere uma relação não linear entre a redução do NDCG e a queda do ARP. Como já identificado em trabalhos anteriores, mesmo com uma variação pequena de apenas 1% no NDCG, o FA*IR também consegue provocar grandes mudanças no ARP (25%). No entanto, a partir deste ponto de maior variação de ARP praticamente não conseguimos mais evoluir o ARP sem impactar fortemente o NDCG, o que ficou representado no último ponto a esquerda, com uma queda ascenturada em NDCG sem acompanhar esta proporcionalidade na variação de ARP.

No gráfico que representa o metodo FA*IR seguido pelo EQNet com PageRank, observa-se um intervalo específico da abordagem, onde a correlação praticamente linear entre os pontos sugere uma relação direta entre a redução do NDCG e a queda do ARP. Assim como no experimento utilizando apenas o EQNet com PageRank, mesmo com uma

variação pequena de apenas 1.65% no NDCG, foi possível alcançar uma redução de até 55% no ARP. Isso indica que é possível ter uma boa resposta na redução da popularidade média aplicando os dois métodos em sequência.

O quinto e último gráfico apresentado as variações de NDCG e a queda do ARP exibe os dados obtidos a partir do experimento utilizando o FA*IR seguido pelo EQNet com PageRank. Neste gráfico de pontos, também é mostrado um intervalo específico da abordagem, onde a correlação aparentemente linear entre os pontos sugere uma relação direta entre a redução de NDCG e a queda do ARP. Neste caso, mesmo com uma variação pequena de apenas 1.8% no NDCG, a redução de ARP ficou muito abaixo dos outros experimentos (12%). Isso sugere que, mesmo tendo uma resposta positiva, o método a ser aplicado após o FA*IR pode impactar na eficiência da redução da popularidade média.

A partir deste ponto, iniciaremos nossa análise utilizando os dados provenientes das análises realizadas no *dataset* da Netflix, caracterizada por uma maior dispersão dos dados e uma proporção mais significativa de usuários em relação aos filme, fornecendo assim um outro contexto para avaliar o desempenho da metodologia.

A Figura 9 exibe as variações de NDCG por variações de ARP dos dados obtidos a partir das cinco metodologias usando o *dataset* da Netflix. Notavelmente, o primeiro gráfico (da metodologia EQNet com PageRank) revela um intervalo específico onde a relação entre ARP e NDCG segue uma correlação parabólica. Este padrão sugere que há um ponto de inflexão onde é possível obter a melhor relação entre ARP e NDCG. Este padrão não apareceu no experimento usando a base do MovieLens, indicando que uma maior coleta de dados seja necessária em nos próximos experimentos. Este achado sugere que, identificar de qual lado da inflexão se está operando é de grande importância para ajustar adequadamente os parâmetros do sistema, visando otimizar tanto a diversificação das recomendações quanto a qualidade geral das mesmas.

O gráfico da parte do experimento utilizando EQNet com Popularity Count revela uma correlação hiperbólica entre ARP e NDCG, sugerindo que o recorte observado pertence ao final da curva, próximo à uma inflexão, semalhante ao observado na Figura 8 para a mesma metodologia. Mesmo neste estágio final da curva, foi possível alcançar uma redução de 50% em ARP com uma redução de apenas 2% em NDCG. Este é interessante, pois indica que a abordagem EQNet com Popularity Count conseguiu reduzir popularidade média das recomendações, mantendo uma degradação mínima na qualidade das mesmas e com um perfil de curva parecido com o que vimos com o EQNet utilizando PageRank.

O terceiro gráfico representa os dados de variação do experimento utilizando o *baseline* FA*IR. Neste gráfico, é possível observar um comportamento semelhante ao esperado de acordo com a literatura, com uma variação pequena de apenas 1.1% em NDCG, foi

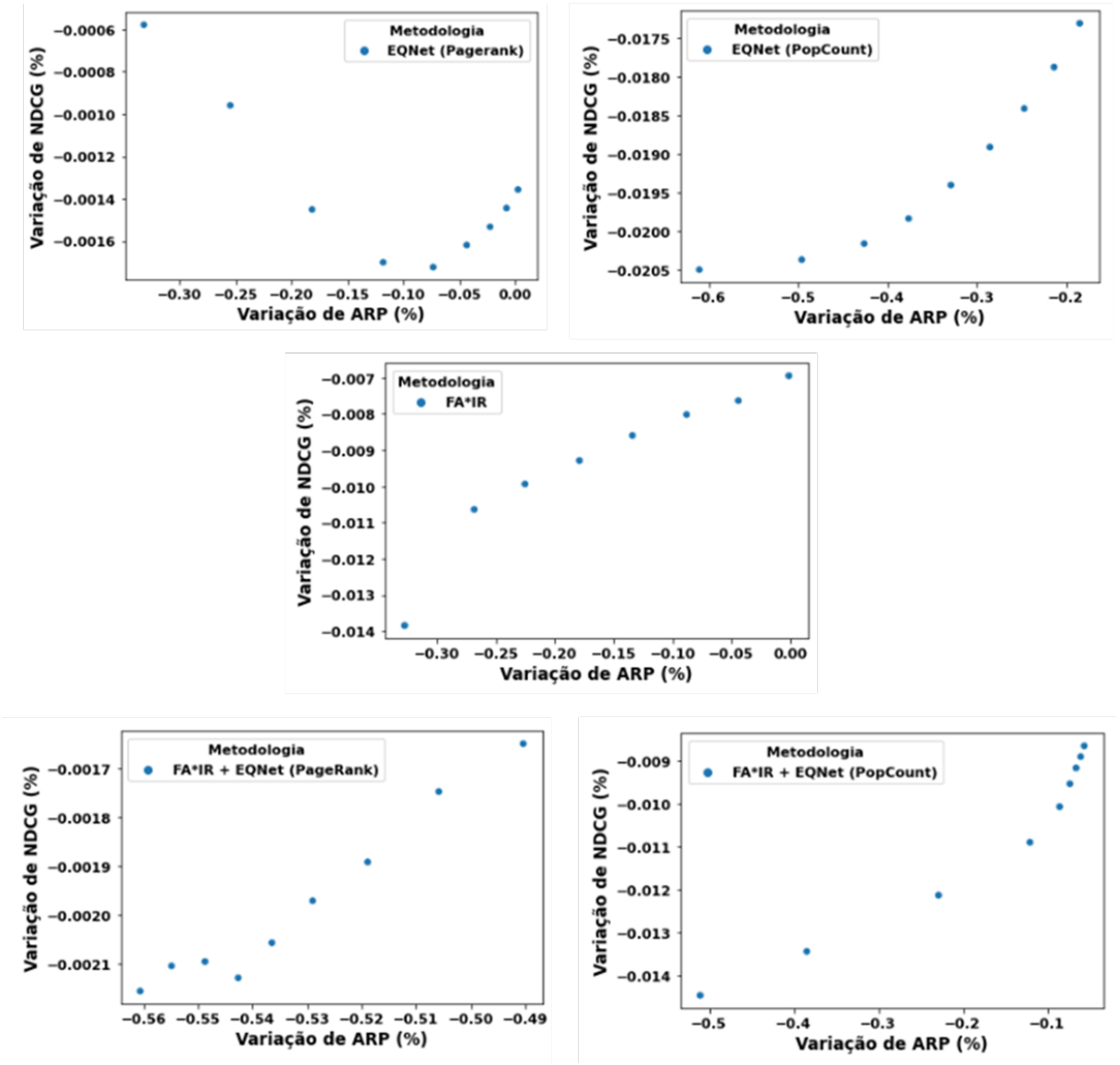

Figura 9 – Gráfico de variação de NDCG por variações de ARP para as cinco metodologias aplicadas, utilizando o *dataset* da Netflix

possível alcançar uma redução de até 25% em ARP (BERNARD; BALOG, 2023; MELUCCI, 2024). Novamente, a partir deste ponto de maior variação de ARP praticamente não conseguimos mais evoluir o ARP sem impactar fortemente o NDCG, o que ficou representado no último ponto a esquerda, com uma queda ascenturada em NDCG sem acompanhar esta proporcionalidade na variação de ARP.

No gráfico que representa o metodo FA*IR seguido pelo EQNet com PageRank, é possível obeservar uma correlação praticamente linear entre as variações de ARP e NDCG, sugerindo que com apenas uma pequena redução de 2.1% em NDCG, foi possível alcançar uma redução de até 55.5% em ARP. Esses resultados indicam que a combinação desses métodos conseguiu alcançar uma considerável redução na popularidade média das recomendações, de maneira similar ao experimento realizado com a base do MovieLens.

O quinto e último gráfico da figura 9 mostra exibe os dados coletados para a parte do experimento utilizando a metodologia do FA*IR seguido de EQNet com PageRank. Novamente, observamos uma correlação praticamente linear entre as duas métricas, sugerindo que uma variação mínima de apenas 1.4% em NDCG resultou em uma redução substancial de até 51% em ARP. Esse padrão reforça a eficácia dessa combinação de métodos em controlar o viés de popularidade, garantindo ao mesmo tempo uma qualidade aceitável nas recomendações. Este caso foi interessante pois performou de maneira mais eficiente que o mesmo método na base do MovieLens. Isto sugere que os parametros ou o método estão mais alinhados ao perfil da base, reforçando a importancia de se avaliarem estes pontos para a aplicação do EQNet.

6.2.2 Dados de evolução de APLT por queda de NDCG

A análise comparativa entre a queda do APLT e a redução do NDCG revela mais uma perspectiva sobre o controle do viés de popularidade nas recomendações. Enquanto a redução do NDCG indica a perda geral na qualidade das recomendações, a diminuição do APLT evidencia especificamente a representação dos itens de cauda longa. Esta abordagem conjunta permite uma compreensão mais aprofundada de como os algoritmos de recomendação estão lidando com o desafio da popularidade desigual dos itens. Ao iniciar a análise com os dados provenientes do *dataset* da MovieLens, que apresenta uma distribuição mais esparsa e uma maior proporção de usuários por filme em comparação com a Netflix, estamos em uma posição privilegiada para examinar a eficácia das técnicas em condições mais desafiadoras, fornecendo insights robustos sobre sua capacidade de adaptação e generalização.

A Figura 10 apresenta as variações de NDCG por variações de APLT dos dados obtidos a partir das cinco metodologias usando o *dataset* do MovieLens. No primeiro gráfico, mostrando os dados da metodologia EQNet com PageRank, a correlação levemente hiperbólica dos pontos sugere que a redução tende a reduzir a eficiência a medida que λ atenua menos os parâmetros de base do EQNet. Mesmo com uma variação pequena de apenas 2% no NDCG, foi possível alcançar uma melhoria de até 45% no APLT. Essa observação destaca a sensibilidade da relação entre essas métricas e a necessidade de uma análise detalhada para otimizar o equilíbrio entre a representatividade dos itens de cauda longa e a qualidade geral das recomendações.

O gráfico seguinte ilustra a variação da abordagem utilizando o EQNet com Popularity Count. Observa-se uma correlação praticamente linear entre a redução do NDCG e a evolução do APLT, com pouca variação entre os resultados encontrados e uma curva ascendente. Essa linearidade sugere que o intervalo de valores experimentados esteja situado após o ponto de inflexão, indicando que valores mais elevados de lambda ou outro

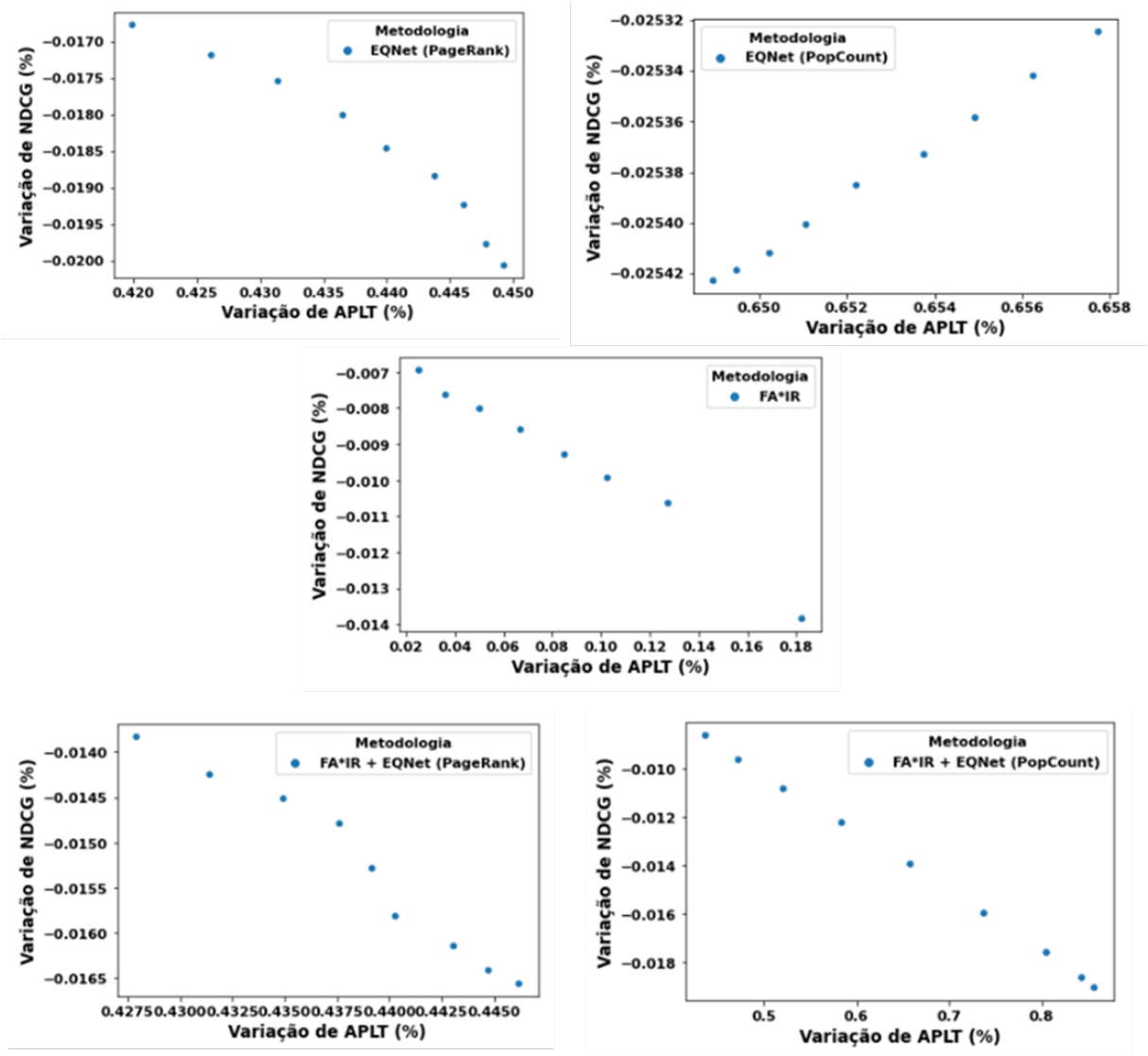

Figura 10 – Gráfico de variação de NDCG por variações de APLT para as cinco metodologias aplicadas, utilizando o *dataset* do MovieLens

fator de atenuação seriam mais indicados para otimizar a relação entre as métricas. Mas, mesmo com uma pequena variação de apenas 2.5% no NDCG, foi possível alcançar uma melhoria de até 65% no APLT, evidenciando a eficácia da abordagem na promoção da representatividade dos itens de cauda longa nas recomendações.

O terceiro gráfico representa os dados de variação do experimento utilizando o *baseline* FA*IR. A correlação linear dos pontos sugere que, com uma variação pequena de apenas 1.4% no NDCG, foi possível alcançar um aumento de até 18% no APLT. Isso indica que, mesmo utilizando uma abordagem mais tradicional como o FAIR, é possível obter melhorias significativas na representação dos itens de cauda longa, mostrando a importância de técnicas eficazes de controle do viés de popularidade para melhorar a diversidade e a relevância das recomendações em sistemas de recomendação.

No gráfico que representa o metodo FA*IR seguido pelo EQNet com PageRank, observamos uma correlação praticamente linear dos pontos sugere que mesmo com uma variação relativamente pequena de apenas 1.65% no NDCG, foi possível alcançar uma melhora de até 45.5% no APLT. Similar ao que foi analisado com relação ao ARP, podemos ver que a atuação das abordagens em paralelo tem o potencial de melhorar a performance dos parâmetros.

O quinto conjunto de dados exibido na figura representa a variação do experimento utilizando primeiro o FAIR seguido pelo EQNet com PageRank. A correlação dos pontos é praticamente linear e mostra que com uma variação pequena de apenas 1.8% no NDCG, foi possível alcançar um aumento de até 87% no APLT. Essa análise destaca o potencial significativo do EQNet em melhorar a representatividade dos itens de cauda longa quando combinado com abordagens estabelecidas como o FA*IR. A melhoria na inclusão dos itens de cauda longo neste caso foi bem acima do observado nos outros casos, ainda mais quando observamos a baixa perda de NDCG.

Passando para a Figura 11, com os dados provenientes das análises no *dataset* da Netflix, com maior esparcidade de dados e uma proporção mais elevada de usuários por filme, temos o primeiro gráfico representando a variação do experimento utilizando o EQ-Net com o algoritmo PageRank. O gráfico mostra um intervalo específico da abordagem, onde a correlação entre APLT e NDCG parece seguir uma tendência levemente hiperbólica. Essa observação sugere a necessidade de um estudo mais detalhado para identificar em qual lado da inflexão estamos trabalhando, visando alcançar a melhor relação entre APLT e NDCG. Neste caso, as variações de NDCG foram muito pequenas, da ordem de apenas 0.06%, mas mesmo assim a melhoria de APLT foi de 8%, uma escala mais de cem vezes maior que o impacto em qualidade.

No gráfico que representa a parte do experimento utilizando EQNet com Popularity Count é possível inferir uma correlação que se assemelha a uma função cúbica, elucidando uma relação mais complexa entre a variação do NDCG e a evolução do APLT que fora possível capturar nos outros experimentos. A amplitude das variações observadas, tanto positivas quanto negativas, indica que o intervalo de valores experimentados foi abrangente o suficiente para capturar uma ampla gama de possíveis comportamentos. Destaca-se que o resultado mais otimizado revelou uma melhora notável de até 14% no APLT, alcançado com uma redução relativamente pequena de apenas 1.75% no NDCG. Isso evidencia uqe nem sempre o melhor ponto de ARP estará relacionado com a melhor performance de APLT, ou das outras métricas, e cada parâmetro precisa ser ajustado para melhor alcançar os objetivos específicos da aplicação.

O terceiro gráfico, por sua vez, representa os dados do experimento utilizando o *baseline* FA*IR. Nele, é perceptível uma correlação linear entre a variação do NDCG e

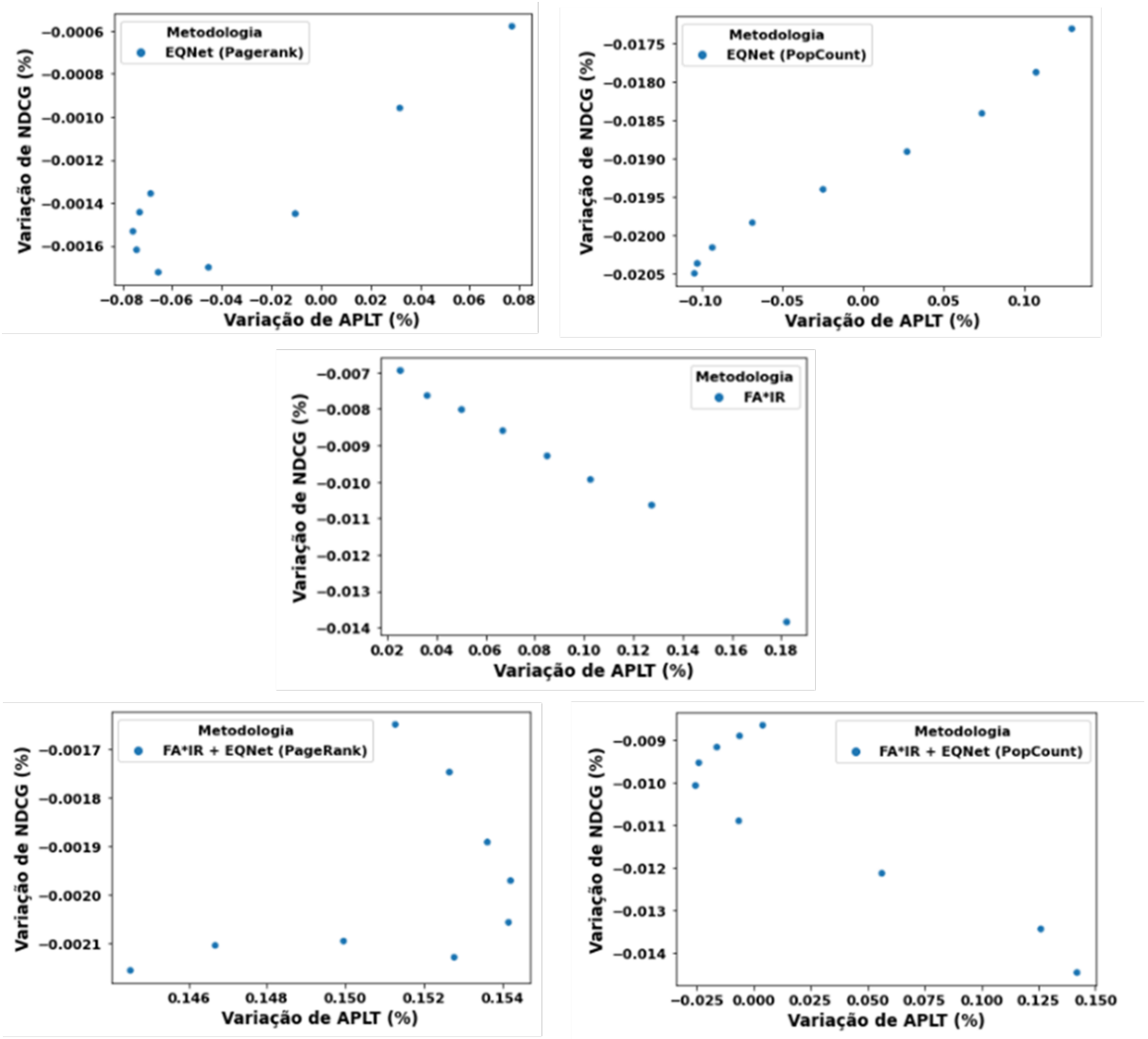

Figura 11 – Gráfico de variação de NDCG por variações de APLT para as cinco metodologias aplicadas, utilizando o *dataset* da Netflix

o aumento do APLT, sugerindo uma relação mais direta e previsível entre essas duas métricas. Intrigantemente, mesmo com uma variação mínima de apenas 1.4% no NDCG, foi possível alcançar um significativo aumento de até 18% no APLT. O valor parecido com o obtido no experimento do MovieLens sugere que as variações do volume de itens de cauda longa nas recomendações pelo FA*IR estariam menos suscetíveis ao perfil da base da plataforma.

O gráfico seguinte representa os dados coletados com o metodo FA*IR seguido pelo EQNet com PageRank. Nele observamos os resultados da variação do experimento utilizando a combinação de FA*IR seguido de EQNet com PageRank, com foco na evolução do APLT no *dataset* da Netflix. O gráfico exibe um intervalo específico da abordagem, onde se destaca uma relação notável entre a redução do NDCG e o aumento do APLT. Mesmo com uma pequena variação de apenas 0.2% no NDCG, foi possível alcançar uma

melhora significativa de até 15.4% no APLT. A disperção dos pontos chama a atenção neste experimento por ser bem menos ordenado que o observado no ARP, mas ainda sim seguir uma certa ordem.

No próximo gráfico, continuamos a análise dos dados de variação do experimento utilizando FA*IR seguido de EQNet com PageRank. Novamente, o gráfico de pontos representa um intervalo específico da abordagem, destacando uma correlação praticamente linear entre a variação do NDCG e o aumento do APLT. Aqui, observamos que uma pequena variação de apenas 1.4% no NDCG resultou em um aumento de até 14.3% no APLT. Apesar de menos significativo que o experimento com FA*IR seguido de EQNet com PageRank, um perfil similar de disperção de pontos pode ser observado. Além disso foi um caso onde a performance da metodologia em sequência ficou mais baixa do que a dos métodos isolados.

6.2.3 Dados de evolução de ACLT por queda de NDCG

A avaliação da evolução do ACLT em relação à redução do NDCG permite discernir a capacidade do sistema em equilibrar a relevância dos itens recomendados com a inclusão de itens de cauda longa, que são menos populares e quanto do portfolio de itens de cauda longa está sendo realmente contemplado. Iniciaremos esta análise com dados provenientes do *dataset* da MovieLens, que apresenta uma distribuição mais uniforme de usuários por filme em comparação com outras bases de dados.

A primeira visualização que exploraremos neste cenário é o primeiro gráfico apresentado na Figura 12, representando a variação do experimento utilizando o EQNet com PageRank. Este gráfico revela uma relação praticamente linear, mas com disperção, entre a variação do NDCG e ACLT. Os dados registrados mostram que, no experimento, uma pequena variação de apenas 2.0% no NDCG resultou em uma queda substancial de até -47.5% no ACLT. Isso sugere que o EQNet utilizando as o PageRank reduz a cobertura de portfólio dos itens de LT.

A análise do segundo gráfico, que exibe os dados de variação do experimento utilizando o EQNet com Popularity Count, revela uma correlação de proporcionalidade direta entre a redução do NDCG e o aumento do ACLT. Esta observação sugere que os valores experimentados estão localizados além do ponto de inflexão, onde a relação entre as grandezas é mais linear, já que os pontos capturados nos outros experimentos indicam um comportamento de proporcionalidade inversa. Isso indica que valores mais elevados de lambda ou outro fator de atenuação podem ser mais indicados para otimizar a relação entre a cobertura de itens de cauda longa e a qualidade geral das recomendações. Apesar disso, mesmo com uma variação relativamente pequena de apenas 2.5% no NDCG, foi possível alcançar uma melhoria significativa de até 78% no ACLT. Isto indica que o

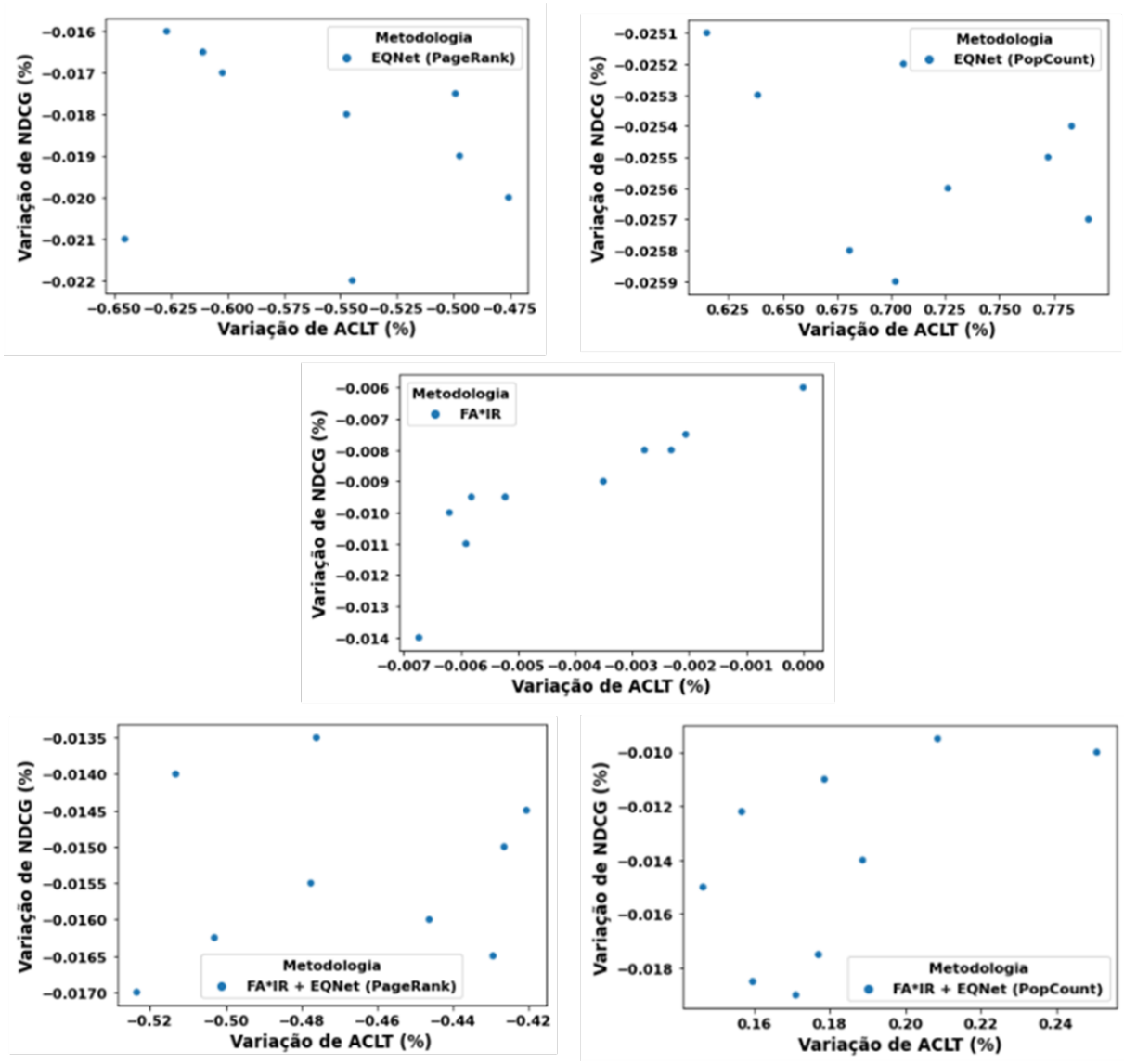

Figura 12 – Gráfico de variação de NDCG por variações de ACLT para as cinco metodologias aplicadas, utilizando o *dataset* do MovieLens

EQNet com inputs vindos do Popularity Count podem ser mais indicados para aumento de cobertura de portfólio.

O gráfico seguinte exibe os dados de variação do experimento utilizando o FA*IR, observa-se uma correlação linear entre a redução do NDCG e o aumento do ACLT. A linearidade dos pontos sugere que mesmo pequenas variações no NDCG podem resultar em melhorias pequenas, porém consistentes, no ACLT. Por exemplo, uma variação de apenas 1.4% no NDCG levou a uma queda muito baixa de -0.7% no ACLT. Esses resultados estão alinhados com os outros experimentos e mostram que o FA*IR tende a não atuar sobre a cobertura de itens de cauda longa do portfolio.

O quarto gráfico ilustra os dados de variação do experimento utilizando FA*IR

seguido de EQNet com PageRank no *dataset* MovieLens. Nela é possível observar uma relação entre a redução do NDCG e o aumento do ACLT. Notavelmente, mesmo com uma variação mínima de apenas 1.65% no NDCG, foi possível alcançar uma redução de até 52% no ACLT. Este indício reforça o potencial do PageRank em restringir as recomendações do portfólio.

Por fim, temos os dados de variação do experimento utilizando FA*IR seguido de EQNet com PageRank. Nele também observamos uma correlação praticamente linear entre a redução do NDCG e o aumento do ACLT. A linearidade dos pontos sugere que mesmo pequenas variações no NDCG podem resultar em melhorias substanciais no ACLT. Por exemplo, uma variação de apenas 1.9% no NDCG levou a um aumento de até 25% no ACLT. Esses resultados estão alinhados com os observados em outras variações de métrica, onde o Popularity Count consegue potencializar a cobertura de portfólio.

A seguir, passamos para as análises com os dados provenientes do *dataset* da Netflix, caracterizados por uma distribuição mais esparça de avaliações e uma maior proporção de usuários por filme. Ao examinarmos o primeiro gráfico apresentado na Figura 13 (que exibe os dados de variação do experimento utilizando EQNet com PageRank) torna-se evidente uma correlação aparentemente direta entre a redução do NDCG e ACLT. O gráfico revela um intervalo no qual os valores experimentados estão além do ponto de inflexão, sugerindo uma relação mais linear entre as grandezas. Essa observação indica que valores mais elevados de lambda ou outros fatores de atenuação podem ser mais indicados para otimizar a relação entre a cobertura de itens de cauda longa e a qualidade geral das recomendações. Apesar da redução de ACLT registrado ser elevado (-60%), destaca-se que a redução correspondente do NDCG foi de apenas 0,25%.

O segundo gráfico exibe os dados de variação do experimento utilizando EQNet com Popularity Count, com uma correlação praticamente linear entre a evolução de ACLT e a redução de NDCG. No entanto, observamos novamente um intervalo ascendente da curva. Isso sugere, novamente, que os valores experimentados estão localizados além do ponto de inflexão, onde a relação entre as grandezas é mais linear. Esta ocorrência prevalente no EQNet acompanhado do Popularity Count sugere que este método pode precisar de um fator de atenuação mais expressivo que o aplicado para o EQNet com PageRank. Apesar disso, mesmo com uma variação modesta de apenas 1.75% de NDCG, foi possível alcançar uma melhora significativa de até 17% de ACLT, destacando a eficácia do EQNet com Popularity Count na promoção de uma maior diversidade nas recomendações.

Na terceira figura temos os dados de variação do experimento utilizando o *baseline* FA*IR, revela uma correlação linear dos pontos. Nesse contexto, com uma variação relativamente pequena de apenas 1.4% de NDCG, o valor de ACLT se manteve praticamente estável, caindo apenas 0.6%. Esses resultados evidenciam novamente a característica do

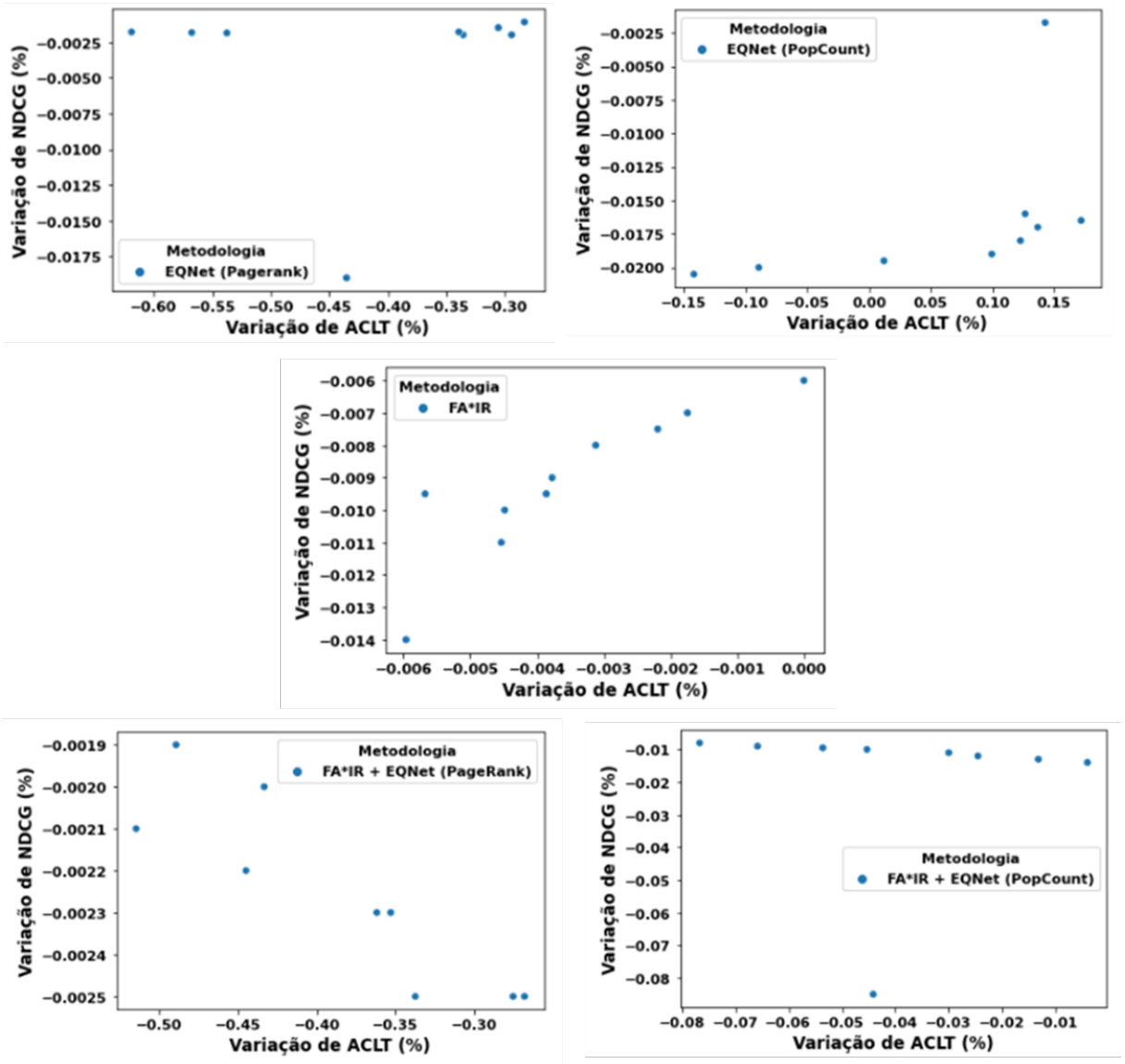

Figura 13 – Gráfico de variação de NDCG por variações de ACLT para as cinco metodologias aplicadas, utilizando o *dataset* da Netflix

FA*IR de promover ganhos em participação de itens de cauda longa, sem grandes alterações de cobertura de portfólio.

No gráfico seguinte temos o experimento utilizando FA*IR seguido de EQNet com PageRank no *dataset* da Netflix. Nele podemos observar uma disperção de pontos onde a ocorrência de maior eficiência em relação à evolução do ACLT pela redução de NDCG resultou uma redução mínima de apenas 0.25% no NDCG, com uma queda significativa de até -30% no ACLT. Essa observação evidencia a capacidade do EQNet com PageRank de atenuar o viés de popularidade quanto a participação de itens, mas reduzindo a cobertura de portfólio.

O último gráfico apresenta os dados de variação do experimento utilizando o FA*IR

seguido de EQNet com PageRank. Nele, podemos observar novamente uma amplitude considerável nas variações, capturando tanto variações de ACLT positivas quanto negativas. Isso sugere que o intervalo de valores experimentados foi abrangente o suficiente para capturar uma ampla gama de possíveis comportamentos na relação entre a redução do NDCG e ACLT. Neste experimento, o NDCG se manteve praticamente constante entre 1% e 1,5%, mas gerando uma queda de até -8% em ACLT. Essa observação reforça o indício de que o FA*IR ajuda a balancear o EQNet, fazendo com que a variação de ACLT frente ao NDCG se de de maneira mais lenta e, portanto, sem constar com variações positivas de ACLT no intervalo observado.

Em suma, os dados capturados durante os experimentos e as análises subsequentes permitiram uma profunda investigação sobre a eficácia do EQNet em conjunto com diferentes algoritmos de ranqueamento em ambientes diversos. A avaliação detalhada do desempenho do EQNet, tanto em isolamento quanto em colaboração com o FA*IR, revelou nuances significativas em sua capacidade de controlar o viés de popularidade e promover uma recomendação mais diversificada. Os resultados obtidos durante os experimentos dão fortes indícios sobre a influência dos algoritmos de ranqueamento, como PageRank e Popularity Count, na performance geral do EQNet, oferecendo uma visão valiosa sobre como esses componentes interagem em diferentes contextos. Além disso, a análise das interações entre o EQNet e o FA*IR forneceu insights sobre pontos específicos que podem ser aprimorados para otimizar ainda mais a eficácia dessa abordagem em futuras iterações.

A análise dos resultados obtidos com a aplicação do EQNet, conforme apresentado neste capítulo, confirma a eficácia da nossa abordagem na mitigação do viés de popularidade. Os dados demonstram que o EQNet não só mantém a qualidade das recomendações, como também melhora significativamente a diversidade dos itens recomendados. Estes achados validam os objetivos do trabalho, mostrando que é possível equilibrar qualidade e popularidade nas recomendações.

7 CONCLUSÃO

Este estudo aborda o problema do viés de popularidade em sistemas de recomendação e seus desafios. Uma revisão de literatura foi realizada para apresentar os trabalhos mais representativos no campo de sistemas de recomendação, discutindo suas limitações e similaridades com este trabalho. A partir desse conhecimento, propõe-se um arcabouço baseado em uma abordagem de reordenação e redes complexas para complementar o sistema de recomendação, explorando sua robustez, eficácia e limitações.

As Figuras de 38 e 39 apresentam compilados dos gráficos experimentais de variação de ARP (Average Recommendation Popularity), APLT (Average Percentage of Long-Tail items) e ACLT (Average Coverage of Long-Tail items) para as bases de dados do MovieLens e Netflix. Observa-se nos gráficos apresentados na Figura 14, referentes ao MovieLens, que o método mais bem-sucedido para lidar com a popularidade média, indicado pela análise de ARP, foi o EQNet em conjunto com o algoritmo PageRank. Para abordar o percentual médio de itens de cauda longa nas recomendações e a cobertura do portfólio, as variações de APLT e ACLT apontam para o método FA*IR em conjunto do EQNet com Popularity Count. Nos gráficos apresentados na Figura 15, referentes à Netflix, observa-se que tanto para tratar a popularidade média quanto o percentual médio de itens de cauda longa nas recomendações e a cobertura do portfólio, o método mais eficaz foi o FA*IR, seguido do EQNet com PageRank.

Em todos os experimentos o EQNet se mostrou eficiente em reduzir o viés de popularidade baixa perda de qualidade de recomendação. O EQNet trabalhando com o PageRank obteve resultados superiores de variação em APLT e ACLT por perda de NDCG. Quando comparado ao *baseline*, o EQNet com PageRank mostrou ser capaz de performar de maneira mais eficiente em situações onde as bases tem uma relação de item por usuário menor e, portanto, um viés menos evidente. Em situações com o viés mais aparente, o EQNet foi capaz de incluir itens nas listagens de maneira mais eficiente, mas com um menor ganho na cobertura de catálogo.

As conclusões deste estudo destacam as contribuições significativas do EQNet para a área de sistemas de recomendação, especialmente na redução do viés de popularidade. Ao atingir os objetivos delineados, demonstramos a viabilidade e a eficácia da nossa

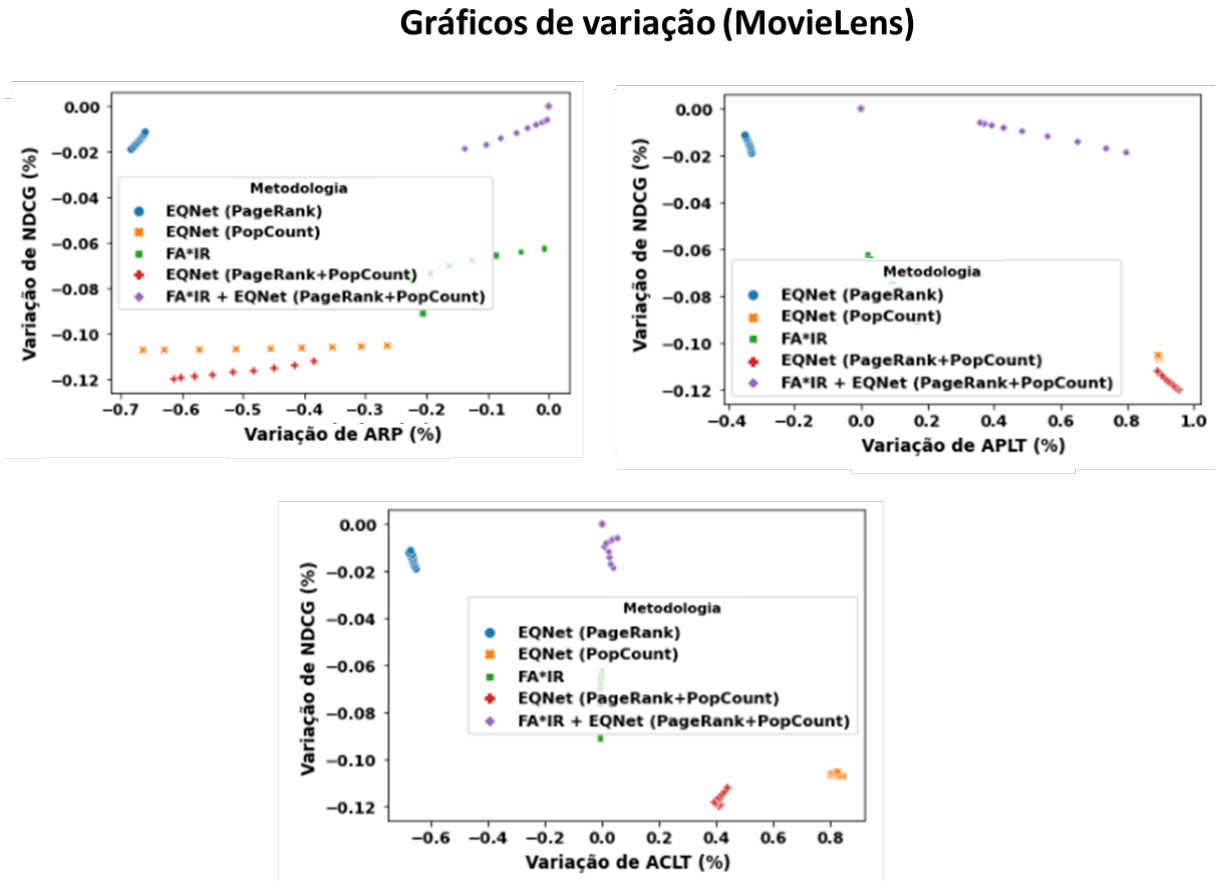

Figura 14 – Gráficos com a variação média de NDCG por ARP, APLT e ACLT nos experimentos utilizando a base do MovieLens.

abordagem, oferecendo uma solução inovadora que pode ser adotada e expandida em futuras pesquisas e aplicações práticas.

7.1 Considerações Finais

Filtros colaborativos representam uma abordagem essencial na busca e filtragem de informações relevantes para os usuários, contribuindo significativamente para a melhoria das experiências digitais. Sua versatilidade é evidenciada pela diversidade de metodologias aplicáveis, resultado do constante desenvolvimento na área de Ciência da Computação. Neste estudo, as conclusões foram alcançadas por meio da análise de métricas como NDCG, ARP, APLT e ACLT, em dez experimentos realizados em bases de dados abertas (Netflix e MovieLens). O EQNet, objeto de investigação neste trabalho, mostrou eficácia na redução do viés de popularidade nos dados experimentais, superando métodos do estado da arte da literatura ao gerar ganhos de diversidade sem comprometer a precisão das recomendações.

As previsões inicialmente geradas por um algoritmo de SVD foram submetidas a

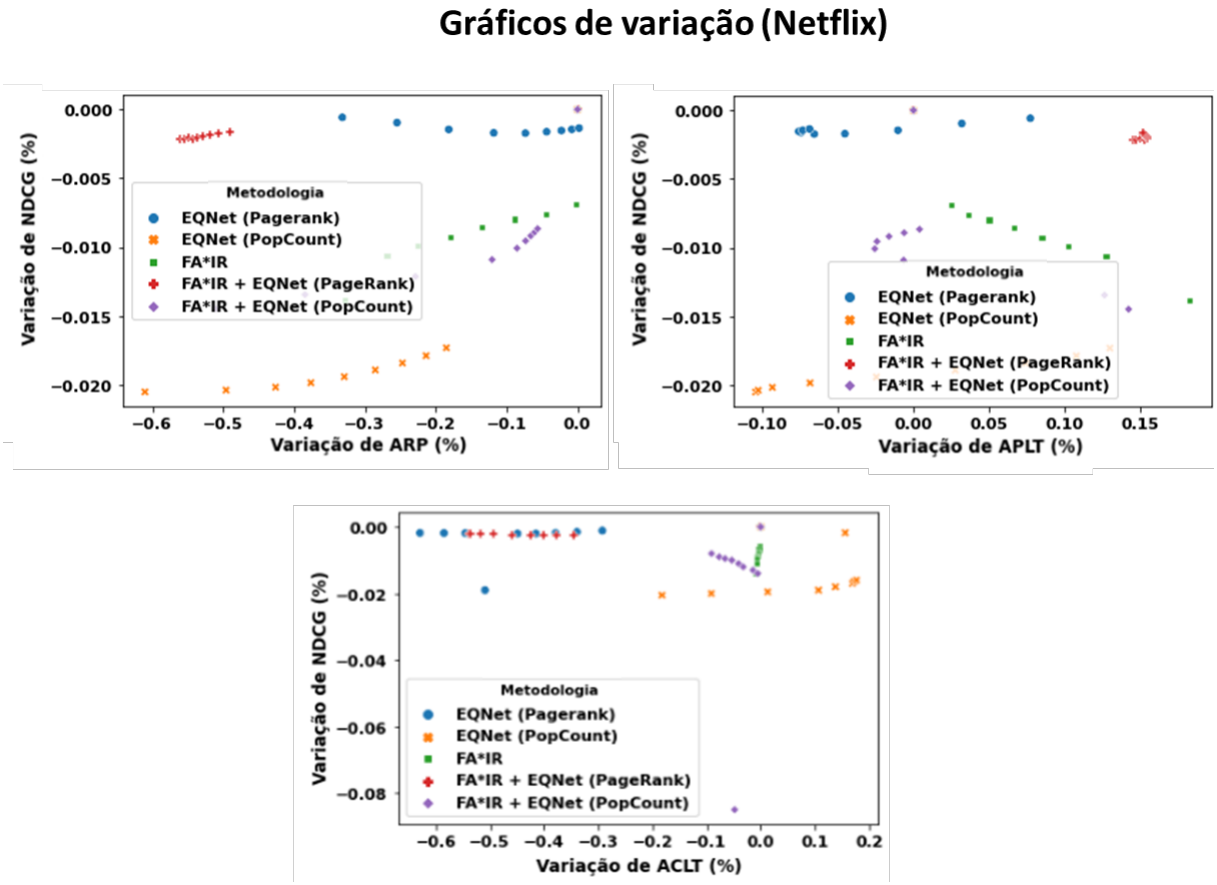

Figura 15 – Gráficos com a variação média de NDCG por ARP, APLT e ACLT nos experimentos utilizando a base da Netflix.

uma reclassificação utilizando o algoritmo FA*IR (ZEHLIKE et al., 2017). Além disso, duas instâncias do EQNet foram aplicadas, cada uma utilizando parâmetros diferentes, como a contagem de popularidade e o algoritmo PageRank e estas também foram utilizadas em conjunto com o FA*IR para avaliar a funcionalidade em conjunto das abordagens. Isto permitiu uma investigação detalhada sobre os efeitos das entradas e das características do banco de dados sobre o desempenho do EQNet.

A abordagem proposta neste trabalho explora uma nova maneira de se reduzir o viés de popularidade nos Sistemas de Recomendação com base em parâmetros de redes complexas, como PageRank, como sua principal contribuição. A aplicação desta abordagem é mostrada com PageRank e Popularity Count como parâmetros de entrada, para comparação, e todo o fluxo experimental também está documentado. Por fim, a evolução das métricas de NDCG, ARP, APLT e ACLT foram avaliadas através das épocas de processamento do EQNet e os resultados finais foram explicitados e debatidos bem como algumas aplicações, melhorias e trabalhos futuros para melhor compreender todo o seu potencial são propostos.

A principal contribuição deste trabalho reside na proposição de uma abordagem

inovadora para mitigar o viés de popularidade em Sistemas de Recomendação, utilizando parâmetros derivados de redes complexas. A implementação desta abordagem é mostrada utilizando tanto o PageRank quanto o Popularity Count como parâmetros de entrada, permitindo uma comparação direta entre eles. O fluxo experimental é documentado detalhadamente para garantir a replicabilidade dos resultados. Por fim, os resultados são analisados e discutidos em profundidade, incluindo possíveis aplicações, melhorias e direções para pesquisas futuras, visando explorar todo o potencial do EQNet.

Devido à sua eficiência, simplicidade e baixa complexidade computacional, evidenciada pelo seu funcionamento baseado em operações de multiplicação vetorial, o EQNet mostrou-se um módulo de extrema utilidade para mitigar o viés de popularidade. O procedimento metodológico adotado durante os experimentos pode requerer adaptações para ser aplicado a outros tipos de métodos de recomendação; no entanto, ressalta-se que a implementação e a calibração dos parâmetros do EQNet permanecem independentes de tais ajustes. Tratando-se de uma ferramenta de reclassificação que utiliza parâmetros e métricas associadas à popularidade e à estrutura de rede para equilibrar o conjunto de recomendações, especula-se sobre a viabilidade de combinar diferentes variantes do EQNet para abordar cenários mais complexos.

7.2 Trabalhos Futuros

O EQNet mostrou uma eficácia notável na mitigação do viés de popularidade, fundamentando-se principalmente na utilização de algoritmos reconhecidos por rankear itens relevantes de grandes conjuntos de dados. No entanto, a análise pós-experimental revelou lacunas e questões que carecem de esclarecimento. Estas demandam uma reflexão mais aprofundada, visando abordar aspectos não contemplados pelo experimento realizado. Neste contexto, é imprescindível uma investigação adicional para compreender integralmente a eficácia e as limitações do EQNet, possibilitando assim contribuições mais significativas para a área de estudo em questão.

- **Como o EQNet funciona com outros arquétipos de RSs?**
  Estudos adicionais e mais abrangentes sobre sistemas de recomendação (RSs), são imperativos para expandir a análise para além do SVD, considerando uma gama mais ampla de RSs. Independentemente da complexidade do sistema em questão, que pode variar desde uma simples decomposição de matrizes até um processo avançado de aprendizado de máquina, é essencial que métricas apropriadas sejam empregadas para avaliar a eficácia e a utilidade das recomendações geradas. As métricas utilizadas neste trabalho, bem como o Root Mean Square Error (RMSE) e o Normalized Mean Absolute Error (NMAE) seguem como sugestões a serem apli-

cadas, mas outras métricas também podem ser interessantes de serem usadas.

- **Como o EQNet se comporta em outras etapas do fluxo de recomendação?**
  Seria interessante realizar uma avaliação do desempenho do EQNet em fases de pré-processamento, bem como em processos envolvendo múltiplas camadas, como é encontrado em algoritmos de redes neurais profundas. Investigar a aplicabilidade do EQNet em diversas camadas pode gerar insights significativos sobre como ele influencia os resultados finais dos experimentos.

- **Que outra estrutura de grafo pode ser construída para estruturar o PageRank e outras métricas de redes complexas?**
  É importante saber quais outras configurações e parâmetros de redes complexas podem influenciar nos resultados do EQNet. Essas variáveis têm o poder de extrair informações intrínsecas cruciais da estrutura da rede do sistema, exercendo influência direta sobre os resultados obtidos. Logo, uma investigação das diversas combinações de configurações e parâmetros seria muito interessante para um entendimento mais completo do funcionamento e potencialidades do EQNet.

- **Quais outros algoritmos de ranqueamento podem ser usados para produzir resultados satisfatórios com o EQNet?**
  Na busca por aprimorar a eficácia do EQNet, é crucial explorar outros algoritmos preditores de popularidade que possam ter suas saídas utilizadas como entrada para a abordagem. Tal estudo permitirá uma análise mais abrangente e precisa do desempenho do EQNet, contribuindo para uma compreensão mais completa de suas capacidades e limitações. Este processo de avaliação multifacetada é essencial para o refinamento contínuo e a otimização do modelo.

- **Como o EQNet performa em um ambiente de teste em tempo real?**
  Para testar a eficácia EQNet em ambientes reais é interessante explorar o seu comportamento em um ambiente real com recomendações e dados acontecendo em tempo real. Tal estudo permitirá uma análise mais precisa do desempenho do EQNet, contribuindo para uma compreensão mais completa de suas capacidades em ambientes reais de aplicação.

Caso tenha qualquer dúvida ou necessite de mais informações sobre como replicar os experimentos descritos nesta dissertação, por favor, não hesite em entrar em contato comigo através do e-mail: gabriel.balbio@sga.pucminas.br. Ficarei feliz em ajudar e fornecer qualquer suporte necessário.

REFERÊNCIAS

ABDOLLAHPOURI, e. a. H. Managing popularity bias in recommender systems with personalized re-ranking. In: FLAIR - International Florida Artificial Intelligence Conference. [S.l.: s.n.], 2019. p. 242 – 251.

ABDOLLAHPOURI, e. a. H. The connection between popularity bias, calibration, and fairness in recommendation. **RecSys 2020**, Sep. 2020.

ABDOLLAHPOURI M. MANSOURY, R. B. H.; MOBASHER, B. The unfairness of popularity bias in recommendation. **Workshop on Recommendation in Multistakeholder Environments**, Sep 2019.

ACADEMIC TORRENTS. Netflix Dataset. 2009. Data retrieved from Academic Torrents, `https://academictorrents.com/details/9b13183dc4d60676b773c9e2cd6de5e5542cee9a`.

ACADEMIC TORRENTS. MovieLens Dataset. 2016. Data retrieved from Academic Torrents, `https://academictorrents.com/details/296054417b4d8eeeb4c7b1c842570bf792ee4d14`.

ADOMAVICIUS, A. T. G. Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions. **IEEE Trans. Knowl. Data Eng.**, v. 17, n. 6, p. 734–749, Jun. 2005.

ADOMAVICIUS, G.; KWON, Y. Toward more diverse recommendations: Item re-ranking methods for recommender systems. **Workshop on Information Technologies and Systems**, Phoenix, AZ, USA, Dec 2009.

AI, e. a. J. Online-rating prediction based on an improved opinion spreading approach. **2017 29th Chinese Control And Decision Conference (CCDC)**, p. 1457–1460, 2017.

AI, e. a. J. Decentralized collaborative filtering algorithms based on complex network modeling and degree centrality. **IEEE Access**, v. 8, p. 151242 – 151249, Aug. 2020.

ALBERT, R.; BARABáSI, A. L. Statistical mechanics of complex networks. **Reviews of Modern Physics**, May. 2002.

AL_SULTANY, G.; GHAIDAA, A. Enhancing recommendation system using adapted personalized pagerank algorithm. In: 2022 5th International Conference on Engineering Technology and its Applications (IICETA). [S.l.: s.n.], 2022. p. 1–5.

ANDERSON, C. A methodology for estimating amazon's long tail sales. **The Long Tail**, Aug 2005. Disponível em: <https://longtail.typepad.com/the$_{l}ong_{t}ail$/2005/08/$a_{m}ethodology_{f}.html$>.

ANDERSON, C. The Long Tail: Why the Future of Business is Selling Less of More. 1. ed. [S.l.]: Hyperion, 2008. 15–27 p.

ANELLI, V. W. et al. Top-n recommendation algorithms: A quest for the state-of-the-art. 03 2022.

ANWAR, T.; UMA, V.; SRIVASTAVA, G. Rec-cfsvd++: Implementing recommendation system using collaborative filtering and singular value decomposition (svd)++. International Journal of Information Technology & Decision Making, v. 20, n. 04, p. 1075–1093, 2021. Disponível em: <https://doi.org/10.1142/S0219622021500310>.

AWS Team. Aws documentation - popularity-count recipe. NVIDIA Corporation, 2022. Disponível em: <https://docs.aws.amazon.com/personalize/latest/dg/native-recipe-popularity.html>.

BARABASI, A.-L.; ALBERT, R. Emergence of scaling in random networks. Science, American Association for the Advancement of Science (AAAS), v. 286, n. 5439, p. 509–512, out. 1999. ISSN 1095-9203. Disponível em: <http://dx.doi.org/10.1126/science.286.5439.509>.

BEDI, e. a. P. Using novelty score of unseen items to handle popularity bias in recommender systems. In: 2014 International Conference on Contemporary Computing and Informatics (IC3I). [S.l.: s.n.], 2014. p. 934–939.

BERNARD, N.; BALOG, K. A systematic review of fairness, accountability, transparency and ethics in information retrieval. ACM Comput. Surv., Association for Computing Machinery, New York, NY, USA, dec 2023. ISSN 0360-0300. Just Accepted. Disponível em: <https://doi.org/10.1145/3637211>.

BOBADILLA, e. a. J. Recommender systems survey. **Knowl.-Based Syst**, v. 46, p. 09–132, Jul. 2013.

BRASIL. Lei Geral de Proteção de Dados Pessoais (LGPD). [S.l.]: Secretaria-Geral, 2018.

BREESE, e. a. J. Empirical analysis of predictive algorithms for collaborative filtering. Microsoft Research, p. 43–52, 1998.

BRYNJOLFSSON, Y. J. E.; SMITH, M. From niches to riches: Anatomy of the long tail. [S.l.: s.n.], 2006.

BULAO, J. How Much Data Is Created Every Day in 2022? set. 2022. `https://techjury.net/blog/how-much-data-is-created-every-day/#gref`.

BURKE, R. Hybrid recommender systems: Survey and experiments. User Modeling and User-Adapted Interaction, Springer, v. 12, n. 3-5, p. 451–498, 2002.

CANO, E.; MORISIO, M. Hybrid recommender systems: A systematic literature review. Intelligent Data Analysis, IOS Press, v. 21, n. 6, p. 1487–1524, nov 2017.

CASTELLS, P.; VARGAS, S.; WANG, J. Novelty and diversity metrics for recommender systems: Choice, discovery and relevance. Proceedings of International Workshop on Diversity in Document Retrieval (DDR), 01 2011.

CAñAMARES, R.; CASTELLS, P. A probabilistic reformulation of memory-based collaborative filtering: Implications on popularity biases. p. 215 – 224, 2017.

CHAI, Z. et al. Recommendation system based on singular value decomposition and multi-objective immune optimization. IEEE Access, PP, p. 1–1, 06 2018.

CHAMPIRI, Z. D.; SHAHAMIRI, S. R.; SALIM, S. S. B. A systematic review of scholar context-aware recommender systems. Expert Systems with Applications, v. 42, n. 3, p. 1743–1758, 2015. ISSN 0957-4174. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0957417414005569>.

CREMONESI, e. a. P. Comparative evaluation of recommender system quality. In: Human-Computer Interaction – INTERACT 2011. [S.l.: s.n.], 2011. p. 1927–1932.

DATA, I. B. The exponential growth of data. Feb. 2017. Disponível em: <https://insidebigdata.com/2017/02/16/the-exponential-growth-of-data/>.

DIDI, T. et al. Promoting tail item recommendations in e-commerce. Association for Computing Machinery, New York, NY, USA, p. 194–203, 2023. Disponível em: <https://doi.org/10.1145/3565472.3592968>.

FANANY, e. a. M. I. Recommender system improvement cases through implicit feedbacks from social network. 2017.

FERRARI, N. F. D. M.; CREMONESI, P. Offline evaluation of recommender systems in a user interface with multiple carousels. Mathematical Problems in Engineering, p. 1–13, 09 2022.

FREEMAN, L. C. A set of measures of centrality based on betweenness. Sociometry, [American Sociological Association, Sage Publications, Inc.], v. 40, n. 1, p. 35–41, 1977. ISSN 00380431. Disponível em: <http://www.jstor.org/stable/3033543>.

FREEMAN, L. C. Centrality in social networks conceptual clarification. Social Networks, v. 1, n. 3, p. 215–239, 1978. ISSN 0378-8733. Disponível em: <https://www.sciencedirect.com/science/article/pii/0378873378900217>.

GOH, e. a. K. Betweenness centrality correlation in social networks. **Physical review. E, Statistical physics, plasmas, fluids, and related interdisciplinary topics**, v. 67, n. 1, Jan. 2004.

GOLDBERG, e. a. D. Using collaborative filtering to weave an information tapestry. **Communication of the ACM**, v. 35, n. 12, p. 61–70, Dec. 1992.

GONG, e. a. S. Combining memory-based and model-based collaborative filtering in recommender system. Proceedings of the 2009 Pacific-Asia Conference on Circuits, Communications and System, PACCS 2009, p. 690–693, 05 2009.

GOOGLE TEAM. Facts about google and competition. Nov. 2011. Disponível em: <https://web.archive.org/web/20111104131332/https://www.google.com/competition/howgooglesearchworks.html>.

GOPALAN, e. a. P. Scalable recommendation with poisson factorization. In: **31st Conference on Uncertainty in Artificial Intelligence (UAI'15)**. [S.l.: s.n.], 2015. p. 326 – 335.

GUO, e. a. S. A pagerank-based collaborative filtering recommendation approach in digital libraries. In: 50th anniversary conference of the american society for cybernetics. [S.l.: s.n.], 2017.

HARPER, F. M.; KONSTAN, J. A. The movielens datasets: History and context. ACM Transactions on Interactive Intelligent Systems, 2015.

HARUNA, K. et al. Context-aware recommender system: A review of recent developmental process and future research direction. Applied Sciences, v. 7, p. 1211, 12 2017.

HERLOCKER, J. A. K. J. L. An algorithmic framework for performing collaborative filtering. SIGIR '99: Proceedings of the 22nd Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, Association for Computing Machinery, New York, NY, USA, p. 230–237, 1999.

HUG, N. surprise Documentation. 2015. `https://surprise.readthedocs.io/en/stable/matrix_factorization.html`. Accessed: 2023-07-31.

INSTAGRAM BUSINESS TEAM. Increasing Website Conversions with Instagram. 2016. Disponível em: <https://business.instagram.com/blog/increasing-website-conversions-with-instagram>.

JALILI, e. a. M. Evaluating collaborative filtering recommender algorithm: A survey. **IEEE Access**, v. 6, p. 74003–74024, Jan. 2018.

JANNACH, D. What recommenders recommend: An analysis of recommendation biases and possible countermeasures. User-Modeling and User-Adapted Interaction, v. 25, p. 427 – 491, 2015.

JANNACH, e. a. D. Recommender systems, an introduction. p. 13 – 288, 2010.

JAVAID, e. a. N. Community search in a multi-attributed graph using collaborative similarity measure and node filtering. 2021.

JIANG, F.; WANG, Z. Pagerank-based collaborative filtering recommendation. p. 597 – 604, 2010.

KIM, H.; SADDIK, A. Personalized pagerank vectors for tag recommendations: Inside folkrank. 2011.

KIM, H. N. et al. Collaborative filtering based on collaborative tagging for enhancing the quality of recommendation. Electronic Commerce Research and Applications, v. 9, n. 1, p. 73–83, 2010.

KIM Q. LI, C. P. e. a. B. A new approach for combining content-based and collaborative filters. Intell Inf Syst 27, p. 79–91, Jul 2006.

KISWANTO, D.; NURJANAH, D.; RISMALA, R. Fairness aware regularization on a learning-to-rank recommender system for controlling popularity bias in e-commerce domain. In: 2018 International Conference on Information Technology Systems and Innovation (ICITSI). [S.l.: s.n.], 2018. p. 16–21.

KNIJNENBURG, e. a. B. P. Explaining the user experience of recommender systems. User Modeling and User-Adapted Interaction, p. 441–504, 2012.

KOREN, Y. The bellkor solution to the netflix grand prize. ago. 2009.

KOWALD, D.; LACIC, E. Popularity bias in collaborative filtering-based multimedia recommender systems. Springer International Publishing, Cham, p. 1–11, 2022.

LI, e. a. K. Deep probabilistic matrix factorization framework for online collaborative filtering. 2019.

LINDEN, e. a. G. Amazon.com recommendations item-to-item collaborative filtering. **IEEE Internet Computing**, IEEE, University of Washington, USA, v. 7, n. 1, p. 76–80, 2003.

LIVNE, A. et al. Evolving context-aware recommender systems with users in mind. Expert Systems with Applications, v. 189, p. 116042, 10 2021.

LOPS, P.; GEMMIS, M. de; SEMERARO, G. Content-based recommender systems: State of the art and trends. In: _____. Recommender Systems Handbook. Boston, MA: Springer US, 2011. p. 73–105. ISBN 978-0-387-85820-3. Disponível em: <https://doi.org/10.1007/978-0-387-85820-3_3>.

LU, e. a. J. Trust-enhanced matrix factorization using pagerank for recommender system. 2017.

Lü, T. Z. L. Link prediction in complex networks: A survey. **Physica A: Statistical Mechanics and its Applications**, v. 390, n. 6, p. 1150–1170, Mar. 2011.

MALDONADO, e. a. E. M. Popularity bias in false-positive metrics for recommender systems evaluation. **ACM Transactions on Information Systems**, v. 39, n. 3, May. 2021.

MARLIN, e. a. B. M. Collaborative filtering and the missing at random assumption. In: **23rd Conference on Uncertainty in Artificial Intelligence (UAI'07)**. [S.l.: s.n.], 2007. p. 267 – 275.

MELUCCI, M. On the trade-off between ranking effectiveness and fairness. Expert Systems with Applications, v. 241, p. 122709, 2024. ISSN 0957-4174. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0957417423032116>.

NANATH, K.; AHMED, M. User-centric evaluation of recommender systems: A literature review. International Journal of Business Information Systems, v. 1, p. 1, 01 2020.

NASIRUZZAMAN, A.; POTA, H. Critical node identification of smart power system using complex network framework based centrality approach. **Annual North-American Power Symposium**, Aug 2011.

NEMATZADEH, A. et al. How algorithmic popularity bias hinders or promotes quality. Scientific Reports, v. 8, 10 2018.

NEWMAN, M. The structure and function of complex networks. **SIAM Review**, p. 167–256, Mar 2003.

NGUYEN, e. a. P. An evaluation of simrank and personalized pagerank to build a recommender system for the web of data. 2015.

NGUYEN, J.; ZHU, M. Content-boosted matrix factorization techniques for recommender systems. Statistical Analysis and Data Mining: The ASA Data Science Journal, Wiley, v. 6, n. 4, p. 286–301, abr. 2013. ISSN 1932-1872. Disponível em: <http://dx.doi.org/10.1002/sam.11184>.

L. Page. Method for node ranking in a linked database. Jan. 1998. Disponível em: <https://patents.google.com/patent/US7058628B1/en>.

PARK W. LEE, B. C. S.; LEE, S. A survey on personalized pagerank computation algorithms. IEEE Access, PP, p. 1–1, 11 2019.

RAHMAN, W. The netflix prize — how even ai leaders can trip up. Jan. 2020. Disponível em: <https://towardsdatascience.com/the-netflix-prize-how-even-ai-leaders-can-trip-up-5c1f38e95c9f>.

RANJBAR, e. a. M. An imputation-based matrix factorization method for improving accuracy of collaborative filtering systems. **Engineering Applications of Artificial Intelligence**, v. 46, p. 58–66, Nov. 2015.

RENDLE, S. et al. Neural collaborative filtering vs. matrix factorization revisited. RecSys '20: Proceedings of the 14th ACM Conference on Recommender Systems, Jun 2020.

RICCI, e. a. F. Recommender systems handbook. Springer, p. 37 – 118, 2015.

SABOUNI, e. a. S. A preliminary radicalisation framework based on social engineering techniques. 2017.

SCHAFER, e. a. J. B. Collaborative filtering recommender system. **The Adaptive Web**, Springer, v. 17, n. 6, p. 291–324, 2007.

SHENGXI, F. et al. Multi-behavior recommendation with svd graph neural networks. Expert Systems with Applications, v. 249, p. 123575, 2024. ISSN 0957-4174. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0957417424004408>.

SORA, I. A pagerank based recommender system for identifying key classes in software systems. May. 2015.

STECK, H. Item popularity and recommendation accuracy. In: **5th ACM Conference on Recommender Systems (RecSys'11)**. [S.l.: s.n.], 2011. p. 125 – 132.

STEFANIDIS, R. B. . K. On mitigating popularity bias in recommendations via variational autoencoders. 36th Annual ACM Symposium on Applied Computing, Mar 2021.

SU, X.; KHOSHGOFTAAR, M. A survey of collaborative filtering techniques. **Advances in Artificial Intelligence**, v. 2009, Aug. 2009.

SUN, C.; XU, Y. Topic model-based recommender system for long-tailed products agains popularity bias. 2019.

TAYLOR, P. Volume of data/information created, captured, copied, and consumed worldwide from 2010 to 2020, with forecasts from 2021 to 2025. Aug 2023. www.statista.com/statistics/871513/worldwide-data-created/. Accessed: 2023-09-03.

THANAPALASINGAM, T. et al. Ontology-based recommendation of editorial products. In: _____. The Semantic Web – ISWC 2018. Springer International Publishing, 2018. p. 341–358. ISBN 9783030006686. Disponível em: <http://dx.doi.org/10.1007/978-3-030-00668-6$_2$1>.

TOMAS, D. What Are Social Media Ads? 2021. Disponível em: <http://cyberclick.net/numericalblogen/what-exactly-are-social-ads-types-and-examples-of-advertising-on-social-media>.

UNIÃO EUROPEIA. General Data Protection Regulation. Intersoft Consulting, 2016. Disponível em: <gdpr-info.eu>.

VALCARCE, D. et al. On the robustness and discriminative power of information retrieval metrics for top-n recommendation. In: Proceedings of the 12th ACM Conference on Recommender Systems. New York, NY, USA: Association for Computing Machinery, 2018. (RecSys '18), p. 260–268. ISBN 9781450359016. Disponível em: <https://doi.org/10.1145/3240323.3240347>.

VIRTANEN, P. et al. SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. Nature Methods, v. 17, p. 261–272, 2020.

WANG, e. a. D. A content-based recommender system for computer science publications. **Knowledge-Based Systems**, v. 157, p. 1–9, Oct. 2018.

WANG, X. et al. Social recommendation with strong and weak ties. In: Proceedings of the 25th ACM International on Conference on Information and Knowledge Management. New York, NY, USA: Association for Computing Machinery, 2016. (CIKM '16), p. 5–14. ISBN 9781450340731. Disponível em: <https://doi.org/10.1145/2983323.2983701>.

XIAO, C. et al. Time sensitivity-based popularity prediction for online promotion on twitter. Information Sciences, v. 525, p. 82–92, 2020. ISSN 0020-0255. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0020025520302371>.

YALCIN, E.; BILGE, A. Evaluating unfairness of popularity bias in recommender systems: A comprehensive user-centric analysis. Information Processing Management, v. 59, n. 6, p. 103100, 2022. ISSN 0306-4573. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0306457322002011>.

YANG, T. et al. Fara: Future-aware ranking algorithm for fairness optimization. In: Proceedings of the 32nd ACM International Conference on

Information and Knowledge Management. ACM, 2023. Disponível em: <http://dx.doi.org/10.1145/3583780.3614877>.

YAO, e. a. Z. Collaborative filtering recommendation algorithms research based on influence and complex network. In: 2014 11th International Conference on Service Systems and Service Management (ICSSSM). [S.l.]: IEEE, 2014.

YAO, S.; HUANG, B. Beyond parity: Fairness objectives for collaborative filtering. In: 31st International Conference on Neural Information Processing Systems. Red Hook, NY, USA: Curran Associates Inc., 2017. (NIPS'17), p. 2925–2934. ISBN 9781510860964.

YEBDRI, Z. et al. Context-aware recommender system using trust network. Computing, v. 103, 09 2021.

YIN, e. a. D. Are bad reviews always stronger than good? asymmetric negativity bias in the formation of online consumer trust. In: **31st International Conference on Information Systems (ICIS'10)**. [S.l.: s.n.], 2010. p. 1 – 18.

YIN, H. et al. Challenging the long tail recommendation. Proc. VLDB Endow., VLDB Endowment, v. 5, n. 9, p. 896–907, may 2012. ISSN 2150-8097. Disponível em: <https://doi.org/10.14778/2311906.2311916>.

YOU, e. a. K. Distributed algorithms for computation of centrality measures in complex networks. **IEEE TRANSACTIONS ON AUTOMATIC CONTROL**, v. 82, n. 5, p. 2080–2094, May. 2017.

ZEHLIKE, M. et al. Fa*ir: A fair top-k ranking algorithm. Association for Computing Machinery, New York, NY, USA, p. 1569–1578, 2017. Disponível em: <https://doi.org/10.1145/3132847.3132938>.

ZEHLIKE, M. et al. Fair top-k ranking with multiple protected groups. Information Processing Management, v. 59, n. 1, p. 102707, 2022. ISSN 0306-4573. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0306457321001916>.

ZHANG, L.; ZHANG, K.; LI, C. A topical pagerank based algorithm for recommender systems. In: Proceedings of the 31st Annual International ACM SIGIR Conference on Research and Development in Information Retrieval. New York, NY, USA: Association for Computing Machinery, 2008. (SIGIR '08), p. 713–714. ISBN 9781605581644. Disponível em: <https://doi.org/10.1145/1390334.1390465>.

ZHANG, S. et al. Deep learning based recommender system: A survey and new perspectives. ACM Comput. Surv., Association for Computing Machinery, New York, NY, USA, v. 52, n. 1, feb 2019. ISSN 0360-0300. Disponível em: <https://doi.org/10.1145/3285029>.

ZHAO, Z.; SHANG, M. S. User-based collaborative-filtering recommendation algorithms on hadoop. In: International Workshop on Knowledge Discovery and Data Mining. [S.l.: s.n.], 2010.