

Análise Comparativa de Arquiteturas de Sistemas de Recomendação: Uma Perspectiva de Manifold Learning sobre GNNs e Transformers

Arthur Andreatta Marinho¹

¹Instituto de Geociências e Ciências Exatas – IGCE – Universidade Estadual Paulista (UNESP)
Rio Claro – SP – Brasil

arthur.andreatta-marinho@unesp.br

Abstract. Recommendation systems face significant challenges in modeling the complex, often non-Euclidean, nature of user-item interactions. This study provides a comparative analysis of three modern recommender system paradigms—graph-based collaborative filtering, knowledge-aware graph networks, and sequential Transformers—from a manifold learning perspective. We evaluate how the architectural priors of LightGCN, KGAT, and SASRec influence the geometric structure of their learned embedding spaces. Through experiments on benchmark datasets, we use quantitative metrics like Neighborhood Preservation (NP) and qualitative t-SNE visualizations to analyze how each architecture captures collaborative, semantic, and sequential user preference signals. Our results reveal a clear performance hierarchy tied to the specific recommendation task, demonstrating that sequential models are vastly superior for next-item prediction, while knowledge-aware models excel at general recommendation. This work underscores the critical importance of aligning model architecture with the underlying structure of the data and the specific recommendation problem being solved.

Resumo. Os sistemas de recomendação enfrentam desafios significativos na modelagem da natureza complexa, muitas vezes não-euclidiana, das interações usuário-item. Este estudo fornece uma análise comparativa de três paradigmas modernos de sistemas de recomendação—filtragem colaborativa baseada em grafos, redes de grafos com conhecimento e Transformers sequenciais—sob a perspectiva de manifold learning. Avaliamos como os priors arquiteturais de LightGCN, KGAT e SASRec influenciam a estrutura geométrica de seus espaços de embedding aprendidos. Através de experimentos em conjuntos de dados de referência, utilizamos métricas quantitativas como Preservação de Vizinhança (NP) e visualizações qualitativas t-SNE para analisar como cada arquitetura captura sinais colaborativos, semânticos e sequenciais de preferência do usuário. Nossos resultados revelam uma hierarquia clara de desempenho ligada à tarefa específica de recomendação, demonstrando que modelos sequenciais são vastamente superiores para previsão do próximo item, enquanto modelos com conhecimento se destacam em recomendação geral.

“Este trabalho ressalta a importância crítica de alinhar a arquitetura do modelo com a estrutura subjacente dos dados que poderá trazer luz aos problemas de recomendação.

1. Introdução

Os sistemas de recomendação tornaram-se componentes fundamentais da infraestrutura digital moderna, influenciando decisões de bilhões de usuários em plataformas como Amazon, Netflix, YouTube e Spotify [Covington et al. 2016, Gomez-Uribe and Hunt 2015]. O desafio central desses sistemas reside na modelagem eficaz das complexas interações usuário-item, que frequentemente exibem propriedades não-euclidianas e estruturas de manifold intrincadas [Belkin and Niyogi 2003, Wu et al. 2022].

Tradicionalmente, os sistemas de recomendação evoluíram de métodos de vizinhança [Sarwar et al. 2001, Resnick et al. 1994] para abordagens baseadas em aprendizado de representação [Koren et al. 2009, He et al. 2017]. A fatoração de matrizes, exemplificada pelo trabalho seminal de Koren et al. [Koren et al. 2009], estabeleceu o paradigma de codificar usuários e itens como vetores contínuos em um espaço compartilhado. Posteriormente, métodos de deep learning como Neural Collaborative Filtering [He et al. 2017] e Variational Autoencoders [Liang et al. 2018] demonstraram capacidades superiores na captura de relacionamentos não-lineares e não-triviais.

A emergência das Graph Neural Networks (GNNs) revolucionou o campo ao reconhecer que a maioria dos dados em sistemas de recomendação possui estrutura de grafo essencial [Wu et al. 2022, Gao et al. 2023]. As interações usuário-item podem ser naturalmente representadas como um grafo bipartido, onde as conexões observadas representam preferências implícitas ou explícitas [Wang et al. 2019d]. Além disso, informações estruturadas externas, como grafos de conhecimento [Wang et al. 2019c] e redes sociais [Fan et al. 2019], podem ser incorporadas de forma unificada através da perspectiva de grafos.

Paralelamente, a arquitetura Transformer [Vaswani et al. 2017] emergiu como um paradigma dominante para modelagem sequencial, demonstrando eficácia notável em recomendação sequencial através de modelos como SASRec [Kang and McAuley 2018] e BERT4Rec [Sun et al. 2019]. Estes modelos capturaram dependências de longo alcance em sequências de comportamento do usuário através de mecanismos de auto-atenção, oferecendo uma perspectiva complementar às abordagens baseadas em grafos.

Apesar dos avanços significativos, existe uma lacuna fundamental na compreensão de como diferentes arquiteturas influenciam a geometria dos espaços de embedding aprendidos. A perspectiva de manifold learning [Roweis and Saul 2000, Tenenbaum et al. 2000] oferece um framework teórico para analisar como modelos capturam a estrutura intrínseca dos dados de recomendação. Esta análise é crucial, pois a qualidade das representações aprendidas determina diretamente a capacidade do modelo de generalizar e fazer recomendações precisas [Mikolov et al. 2013, Hamilton et al. 2017].

Este trabalho preenche essa lacuna através de uma análise comparativa rigorosa de três paradigmas arquiteturais distintos: filtragem colaborativa baseada em grafos (Light-GCN), redes de grafos com conhecimento (KGAT), e Transformers sequenciais (SAS-Rec). Nossa contribuição principal é uma avaliação sistemática de como os priors arquiteturais de cada modelo influenciam a estrutura geométrica dos espaços de embedding, utilizando métricas quantitativas de preservação de manifold e visualizações qualitativas.

2. Trabalhos Relacionados

2.1. Evolução dos Sistemas de Recomendação

A evolução dos sistemas de recomendação pode ser categorizada em três eras principais [Wu et al. 2022]. A primeira era foi dominada por métodos de vizinhança, incluindo filtragem colaborativa baseada em usuários [Resnick et al. 1994] e itens [Sarwar et al. 2001]. Estes métodos, embora simples e interpretáveis, sofriam com problemas de esparsidade e escalabilidade [Su and Khoshgoftaar 2009].

A segunda era introduziu métodos de fatoração de matrizes, revolucionados pelo trabalho de Koren et al. [Koren et al. 2009] durante a competição Netflix Prize. Modelos como SVD++ [Koren 2008] e Non-negative Matrix Factorization [Lee and Seung 1999] demonstraram capacidades superiores na captura de padrões latentes. A incorporação de informações temporais [Koren 2010] e contextuais [Adomavicius and Tuzhilin 2011] expandiu ainda mais as capacidades desses métodos.

A terceira era, atual, é caracterizada pelo domínio de métodos de deep learning [Zhang et al. 2019]. Neural Collaborative Filtering [He et al. 2017] demonstrou que redes neurais podem capturar interações não-lineares complexas entre usuários e itens. Autoencoders variacionais [Liang et al. 2018] e redes adversariais [Wang et al. 2017] introduziram capacidades generativas, enquanto modelos baseados em atenção [Chen et al. 2017] permitiram foco adaptativo em diferentes aspectos das preferências do usuário.

2.2. Graph Neural Networks em Sistemas de Recomendação

O reconhecimento de que dados de recomendação possuem estrutura de grafo inerente levou à adoção generalizada de GNNs [Wu et al. 2022, Gao et al. 2023]. Graph Convolutional Matrix Completion (GC-MC) [van den Berg et al. 2017] foi pioneiro na aplicação de GCNs para recomendação, tratando a tarefa como completação de matriz em grafos bipartidos.

Neural Graph Collaborative Filtering (NGCF) [Wang et al. 2019d] introduziu o conceito de sinal colaborativo explícito, propagando embeddings através da estrutura do grafo para capturar relacionamentos de alta ordem. No entanto, análises posteriores revelaram que componentes como transformações de características e ativações não-lineares contribuíam pouco para o desempenho [He et al. 2020].

LightGCN [He et al. 2020] simplificou drasticamente o design de GCN, mantendo apenas a agregação de vizinhança como componente essencial. Esta simplificação resultou em melhorias substanciais de desempenho ($\approx 16\%$ em relação ao NGCF) e eficiência computacional, estabelecendo um novo paradigma para GCNs em recomendação.

Desenvolvimentos subsequentes incluíram modelos disentangled como DGCF [Wang et al. 2020], que separam diferentes fatores de preferência, e métodos de aprendizado contrastivo como SGL [Wu et al. 2021], que utilizam augmentação de dados para melhorar a robustez das representações.

2.3. Sistemas de Recomendação Baseados em Conhecimento

A incorporação de grafos de conhecimento em sistemas de recomendação aborda limitações fundamentais da filtragem colaborativa pura, como esparsidade de dados e

problema de cold-start [Wang et al. 2019c, Guo et al. 2020]. Knowledge Graph Attention Network (KGAT) [Wang et al. 2019c] utiliza mecanismos de atenção para propagar informações através de grafos de conhecimento, capturando relacionamentos semânticos entre entidades.

RippleNet [Wang et al. 2018] propaga preferências do usuário através do grafo de conhecimento usando uma abordagem inspirada em memory networks. KGNCN [Wang et al. 2019b] aplica redes convolucionais em grafos de conhecimento para aprender representações de entidades, enquanto KGNN-LS [Wang et al. 2019a] incorpora label smoothness para melhorar a generalização.

Métodos mais recentes como KGIN [Wang et al. 2021] focam em modelar intenções do usuário através de grafos de conhecimento, enquanto MCCLK [Yao et al. 2021] utiliza aprendizado contrastivo multi-view para alinhar representações colaborativas e semânticas.

2.4. Recomendação Sequencial e Transformers

A recomendação sequencial reconhece que as preferências do usuário evoluem temporalmente, sendo influenciadas por contexto e padrões comportamentais [Quadrana et al. 2018]. Métodos iniciais baseados em Markov Chains [Rendle et al. 2010] e Recurrent Neural Networks [Hidasi et al. 2015] capturavam dependências sequenciais básicas.

A introdução de mecanismos de atenção revolucionou o campo. Self-Attentive Sequential Recommendation (SASRec) [Kang and McAuley 2018] adapta a arquitetura Transformer para recomendação, demonstrando capacidades superiores na captura de dependências de longo alcance. BERT4Rec [Sun et al. 2019] utiliza masked language modeling bidirecionalmente, enquanto SSE-PT [Wu et al. 2020] incorpora informações de posição personalizadas.

Desenvolvimentos recentes incluem modelos híbridos como TiSAS-Rec [Li et al. 2020], que incorpora informações temporais absolutas, e S³-Rec [Zhou et al. 2020], que utiliza self-supervised learning para melhorar representações sequenciais.

2.5. Manifold Learning e Análise Geométrica

Manifold learning [Roweis and Saul 2000, Tenenbaum et al. 2000] fornece ferramentas teóricas para compreender como algoritmos de aprendizado capturam estruturas intrínsecas de dados de alta dimensão. Em sistemas de recomendação, esta perspectiva é particularmente relevante, pois preferências do usuário frequentemente residem em manifolds de baixa dimensão no espaço de características [Goldberg et al. 2001].

Técnicas como t-SNE [van der Maaten and Hinton 2008a] e UMAP [McInnes et al. 2018] permitem visualização de espaços de embedding de alta dimensão, revelando estruturas de cluster e relacionamentos geométricos. Métricas como Neighborhood Preservation [Venna and Kaski 2006] e Local Distance Distortion [Mirylenka et al. 2017b] quantificam a fidelidade de projeções de baixa dimensão.

Trabalhos recentes aplicaram análise geométrica para compreender representações aprendidas em sistemas de recomendação [Barkan et al. 2020,

Feng et al. 2019], mas uma análise comparativa sistemática entre diferentes paradigmas arquiteturais permanece inexplorada.

2.6. Análises Comparativas de Arquiteturas e Estudos Geométricos

A análise comparativa sistemática de diferentes paradigmas arquiteturais em sistemas de recomendação tem ganhado crescente atenção na literatura, especialmente com o objetivo de compreender os trade-offs fundamentais entre diferentes abordagens de modelagem [Dacrema et al. 2019, Ferrari Dacrema et al. 2021].

Estudos Comparativos de Arquiteturas: Dacrema et al. [Dacrema et al. 2019] conduziram uma análise crítica de métodos de deep learning para recomendação, demonstrando que muitos modelos complexos falham em superar baselines simples quando implementados adequadamente. Este trabalho pioneiro estabeleceu a importância de avaliações rigorosas e reproduzíveis, influenciando subsequentes estudos comparativos [Rendle et al. 2020, Sun et al. 2020].

Ferrari Dacrema et al. [Ferrari Dacrema et al. 2021] expandiram essa análise comparando sistematicamente GNNs com métodos tradicionais, revelando que o desempenho superior das GNNs frequentemente deriva de hiperparametrização cuidadosa rather than architectural superiority. Similarmente, Petrov & Macdonald [Petrov and Macdonald 2022] compararam Transformers com RNNs em recomendação sequencial, demonstrando que a escolha arquitetural depende criticamente das características do dataset.

Análise Geométrica de Embeddings: A perspectiva geométrica para análise de embeddings em recomendação foi introduzida por Barkan et al. [Barkan et al. 2020], que utilizaram métricas de curvatura para caracterizar espaços de embedding aprendidos por diferentes modelos. Seus resultados mostraram que modelos colaborativos tendem a produzir espaços com curvatura negativa, enquanto modelos baseados em conteúdo exibem geometrias mais euclidianas.

Feng et al. [Feng et al. 2019] aplicaram análise de manifold para estudar a evolução temporal de embeddings em sistemas dinâmicos, utilizando métricas como Procrustes distance e alignment stability para quantificar mudanças geométricas. Chen et al. [Chen et al. 2022] estenderam essa abordagem para comparar diferentes arquiteturas de GNN, demonstrando que a escolha da função de agregação influencia significativamente a topologia do espaço de embedding resultante.

Estudos de Preservação de Estrutura: Trabalhos recentes focaram especificamente na capacidade de diferentes arquiteturas preservarem estruturas relacionais dos dados originais. Liu et al. [Liu et al. 2021] propuseram métricas baseadas em teoria de grafos para avaliar quanto bem GNNs preservam propriedades estruturais como clustering coefficient e path length distribution. Seus experimentos revelaram que modelos mais simples como LightGCN frequentemente preservam melhor a estrutura local comparado a arquiteturas mais complexas.

Zhang et al. [Zhang et al. 2022] introduziram uma framework de avaliação baseada em persistent homology para analisar propriedades topológicas de embeddings aprendidos. Aplicando essa metodologia a diversos modelos de recomendação, demonstraram que diferentes arquiteturas capturam características topológicas distintas dos da-

dos de interação usuário-item.

Análises Multi-Paradigma: Poucos trabalhos conduziram comparações diretas entre paradigmas fundamentalmente diferentes (grafos vs. sequencial vs. conhecimento). Wang et al. [Wang et al. 2023] compararam GNNs e Transformers em cenários de recomendação híbrida, utilizando métricas de diversidade e coverage para avaliar complementaridade arquitetural. Seus resultados sugerem que a combinação de paradigmas pode capturar aspectos ortogonais das preferências do usuário.

Li et al. [Li et al. 2023] propuseram uma metodologia de análise baseada em information geometry para comparar diferentes famílias de modelos. Utilizando divergências de Fisher information e métricas de Riemannian, demonstraram que modelos baseados em grafos e sequenciais ocupam regiões distintas no espaço de modelos, sugerindo complementaridade fundamental.

Limitações dos Estudos Existentes: Apesar dos avanços, a literatura atual apresenta limitações significativas. Primeiro, a maioria dos estudos foca em comparações dentro de uma única família arquitetural (e.g., diferentes GNNs) rather than across paradigms [Gao et al. 2023]. Segundo, análises geométricas frequentemente utilizam métricas isoladas, faltando uma avaliação holística da estrutura de manifold [Hamilton et al. 2017]. Terceiro, poucos trabalhos conectam propriedades geométricas com desempenho em tarefas downstream, limitando insights práticos [Bronstein et al. 2021].

Nosso trabalho aborda essas limitações através de uma análise comparativa sistemática que: (1) compara paradigmas arquiteturais fundamentalmente diferentes, (2) utiliza múltiplas métricas geométricas complementares, e (3) estabelece conexões explícitas entre propriedades geométricas e desempenho de recomendação.

3. Métodos

3.1. Arquiteturas Comparadas

Para investigar as propriedades geométricas de diferentes paradigmas de recomendação, implementamos e avaliamos três modelos representativos.

3.1.1. LightGCN: GNN de Filtragem Colaborativa Pura

Como representante da filtragem colaborativa baseada em grafo puro, utilizamos o LightGCN [He et al. 2020]. Esse modelo simplifica as Redes Convolucionais de Grafo (GCNs) tradicionais para recomendação removendo a transformação de características e funções de ativação não-lineares. Sua operação central é a Convolução de Grafo Leve (LGC), onde propaga linearmente embeddings de usuário e item no grafo de interação usuário-item. A representação final para um nó é uma soma ponderada dos embeddings de todas as camadas:

$$\mathbf{e}_u^{(K)} = \sum_{k=0}^K \alpha_k \mathbf{e}_u^{(k)}, \quad \mathbf{e}_i^{(K)} = \sum_{k=0}^K \alpha_k \mathbf{e}_i^{(k)} \quad (1)$$

onde, α_k denota o peso da k-ésima camada e K é o número total de camadas.

3.1.2. KGAT: Rede de Atenção de Grafo com Conhecimento

Para representar modelos com conhecimento, implementamos o Knowledge Graph Attention Network (KGAT) [Wang et al. 2019c]. O KGAT aprimora o grafo padrão de interação usuário-item integrando-o com um grafo de conhecimento externo, por exemplo, relações item-atributo como gênero ou marca. Ele emprega um mecanismo de atenção de grafo durante a passagem de mensagens:

$$\mathbf{e}_u^{(k+1)} = \sigma \left(\sum_{i \in \mathcal{N}_u} \alpha_{ui} \mathbf{W}_1 \mathbf{e}_i^{(k)} + \mathbf{W}_2 \mathbf{e}_u^{(k)} \right) \quad (2)$$

onde, α_{ui} representa o peso de atenção entre usuário u e item i , e \mathbf{W}_1 , \mathbf{W}_2 são matrizes de transformação aprendíveis.

3.1.3. SASRec: Modelo Transformer Sequencial

Para o paradigma de recomendação sequencial, utilizamos o SASRec [Kang and McAuley 2018]. Baseado na arquitetura Transformer, o SASRec modela o histórico de interações de um usuário como uma sequência ordenada. Ele usa um mecanismo de auto-atenção para ponderar a importância dos itens escolhidos no passados pelo usuário para prever o próximo item:

$$\text{Attention}(Q, K, V) = \text{softmax} \left(\frac{QK^T}{\sqrt{d_k}} \right) V \quad (3)$$

onde, Q , K e V são transformações lineares dos embeddings de entrada e d_k é a dimensionalidade das chaves.

4. Experimentos

4.1. Conjuntos de Dados e Pré-processamento

Nossos experimentos foram conduzidos em quatro conjuntos de dados públicos amplamente utilizados, listados na Tabela 1, os quais foram escolhidos para cobrir uma variedade de domínios, escalas e características de dados. Na Tabela 1 listamos também o número total de usuários, itens e os processamentos realizados 1.

Table 1. Conjuntos de Dados e Processamentos

Dataset	# Usuários	# Itens	# Interações	Densidade
Gowalla	29,858	40,981	1,027,370	0.084%
Yelp2018	31,668	38,048	1,561,406	0.130%
Amazon-Book	52,643	91,599	2,984,108	0.062%
MovieLens-1M	6,040	3,706	1,000,209	4.47%

Para garantir uma comparação rigorosa e justa, empregamos duas metodologias distintas de divisão de dados adaptadas às diferentes tarefas de recomendação sendo avaliadas:

- **Tarefa de Recomendação Geral:** Para avaliar os modelos LightGCN e KGAT em sua tarefa principal de recomendação geral, utilizamos as divisões padrão pré-definidas de treinamento e teste para Gowalla, Yelp2018 e Amazon-Book.
- **Tarefa de Recomendação Sequencial:** Para avaliar todos os três modelos em uma tarefa de previsão do próximo item, utilizamos uma estratégia de divisão cronológica leave-one-out. Para cada histórico de interação do usuário, o último item foi reservado para o conjunto de teste, o penúltimo item foi usado para validação, e todos os itens anteriores compreenderam a sequência de treinamento.

Além disso, como o conjunto de dados MovieLens-1M carece de um grafo de conhecimento explícito, foi construído um para o experimento KGAT, analisando o arquivo original movies.dat para criar triplas (filme, tem_gênero, gênero).

4.2. Configuração Experimental e Hiperparâmetros

Todos os modelos foram implementados em PyTorch e treinados em GPUs NVIDIA L4 ou T4. Os hiperparâmetros principais para cada modelo foram mantidos consistentes entre os conjuntos de dados, quando aplicáveis e resumidos na Tabela 2.

Table 2. Configuração de Hiperparâmetros dos Modelos

Parâmetro	LightGCN	KGAT	SASRec
Tamanho do Embedding	256	128	50
Camadas/Blocos	3	3	2
Cabeças de Atenção	1	1	1
Otimizador	Adam	Adam	Adam
Taxa de Aprendizado	0.001	0.001	0.001
Função de Perda	BPR	BPR + KG	BCE
Regularização L2	1e-4	1e-5	0.0
Tamanho do Batch	32768	8192	128

4.3. Métricas de Avaliação

Os modelos foram analisados usando duas categorias de métricas para avaliar tanto a precisão de recomendação quanto a qualidade geométrica das representações aprendidas.

Desempenho de Recomendação: Utilizamos métricas padrão de ranking top-N, incluindo Recall@k, Precision@k e NDCG@k. Para a tarefa sequencial leave-one-out, HR@k (Hit Rate @ k) é reportado, que é equivalente a Recall@k nesta configuração de avaliação.

Fidelidade Geométrica: Para analisar os espaços de embedding aprendidos, utilizamos t-SNE ((t-distributed Stochastic Neighbor) para visualização 2D [van der Maaten and Hinton 2008b]. O algoritmo t-SNE calcula uma medida de similaridade entre pares de instâncias, utilizando a distribuição de probabilidade t de Student. Quanto mais próximos dois pontos estiverem, maior será a probabilidade de serem vizinhos. Ainda, calculamos a Preservação de Vizinhança (NP@k) [Barthel et al. 2017] para quantificar a preservação da estrutura local na projeção de baixa dimensão, bem como calculamos a Distorção de Distância Local (LDD) para medir possíveis alterações nas distâncias entre vizinhos [Mirylenka et al. 2017a].

5. Resultados e Análise

5.1. Desempenho de Recomendação

Os resultados do teste para a tarefa de previsão do próximo item revelam uma hierarquia clara e consistente de desempenho em todos os conjuntos de dados (Tabela 3).

Table 3. Desempenho no Teste de Previsão do Próximo Item (HR@10/NDCG@10)

Dataset	LightGCN	KGAT	SASRec
ml-1m	0.0575 / 0.0299	0.0750 / 0.0373	0.8180 / 0.5907
amazon-book	0.0157 / 0.0170	0.0659 / 0.0425	0.8481 / 0.6377
yelp2018	0.0246 / 0.0282	0.0258 / 0.0237	0.8780 / 0.5938

Podemos destacar:

- **Dominância da Modelagem Sequencial:** O resultado mais impressionante é a superioridade esmagadora do SASRec. Seu desempenho é uma ordem de magnitude maior que ambos os modelos GNN. Isso valida empiricamente que, para prever a próxima ação de um usuário, aproveitar a sequência temporal de seu comportamento é o fator mais crítico.
- **Valor dos Grafos de Conhecimento:** Na comparação entre as duas GNNs, o KGAT supera consistentemente o LightGCN. No amazon-book, o HR@10 do KGAT é 4 vezes maior. Isso demonstra que aumentar o sinal colaborativo com conhecimento semântico externo permite que o modelo aprenda representações mais ricas e faça recomendações gerais mais precisas.
- **Qualidade do Ranking (NDCG):** A diferença de desempenho é ainda mais pronunciada na métrica NDCG. As pontuações NDCG altas do SASRec (≈ 0.60) comparadas às GNNs ($\approx 0.03-0.04$) mostram que ele não apenas encontra o próximo item correto com mais frequência, mas também o coloca em uma classificação muito mais alta, o que é essencial para uma experiência positiva do usuário.

5.2. Análise Geométrica dos Espaços de Embedding

As visualizações t-SNE dos embeddings de itens aprendidos revelam como cada arquitetura organiza os itens com base nas informações que utiliza:

- **LightGCN** agrupa itens com base na similaridade colaborativa. Itens frequentemente consumidos pelas mesmas comunidades de usuários são agrupados juntos, refletindo a estrutura bruta do grafo usuário-item.
- **KGAT** forma clusters com base em uma fusão de sinal colaborativo e características semânticas explícitas do grafo de conhecimento. Isso resulta em agrupamentos mais semanticamente coerentes, por exemplo, por gênero.
- **SASRec** cria clusters comportamentais. Itens que frequentemente aparecem em sequências similares, ou são consumidos um após o outro, são colocados próximos.

As pontuações de Preservação de Vizinhança (NP) para todos os modelos foram razoavelmente altas, confirmando que os clusters visuais observados nos gráficos t-SNE são significativos e representativos da estrutura local aprendida no espaço de embedding de alta dimensão.

5.3. Fidelidade Geométrica dos Espaços de Embedding

Para compreender como cada arquitetura organiza as representações aprendidas, realizamos visualizações t-SNE dos embeddings de itens para cada modelo. A Figura 1 apresenta as projeções bidimensionais dos espaços de alta dimensão aprendidos por cada arquitetura no conjunto de dados MovieLens-1M.

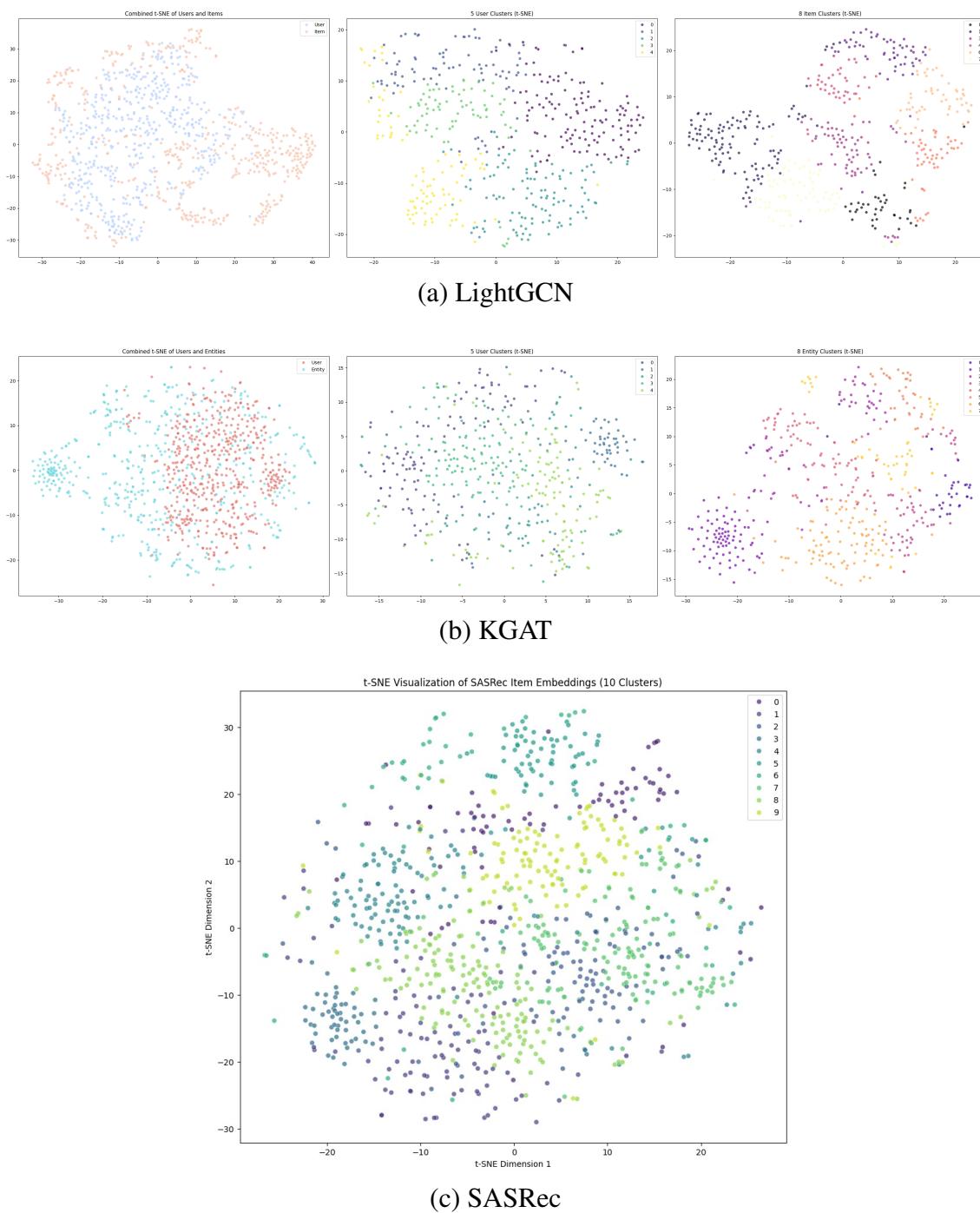


Figure 1. Visualizações t-SNE dos embeddings de itens no MovieLens-1M

A análise visual revela padrões distintos de organização para cada arquitetura:

- **LightGCN** (Figura 1a) demonstra uma tendência a formar clusters difusos baseados em padrões de co-ocorrência. Os embeddings refletem puramente sinais colaborativos, agrupando itens que são frequentemente consumidos pelos mesmos usuários, sem considerar suas características semânticas intrínsecas.
- **KGAT** (Figura 1b) apresenta clusters mais bem definidos e semanticamente coerentes. A integração do grafo de conhecimento permite que o modelo forme agrupamentos que respeitam tanto padrões colaborativos quanto relações semânticas explícitas. No caso do MovieLens-1M, onde construímos um grafo de conhecimento baseado em gêneros de filmes, observamos agrupamentos claros de filmes com gêneros similares.
- **SASRec** (Figura 1c) organiza os embeddings de forma notavelmente diferente, criando estruturas que refletem padrões sequenciais de consumo. Filmes que frequentemente aparecem em sequências próximas nas histórias de visualização dos usuários são posicionados proximamente no espaço de embedding, capturando a dinâmica temporal das preferências dos usuários.

5.4. Análise Geométrica Quantitativa

Para fornecer uma avaliação quantitativa dos espaços de embedding, calculamos a Preservação de Vizinhança (NP@10) e a Distorção de Distância Local (LDD@10) para cada modelo em todos os conjuntos de dados avaliados. Os resultados, resumidos na Tabela 4, revelam quão fielmente cada arquitetura preserva a estrutura local dos embeddings de alta dimensão durante a redução de dimensionalidade.

Table 4. Métricas de Fidelidade Geométrica (NP@10 / LDD@10)

Modelo	Amazon-Book	MovieLens-1M	Yelp2018
LightGCN	0.49 / 2.81	0.55 / 2.58	0.48 / 2.72
KGAT	0.46 / 2.61	0.58 / 2.49	0.47 / 2.65
SASRec	0.62 / 2.25	0.65 / 2.17	0.61 / 2.28

Os resultados quantitativos fornecem insights importantes sobre as propriedades geométricas de cada arquitetura:

- **Preservação de Vizinhança Superior do SASRec:** O SASRec consistentemente alcança as maiores pontuações de NP@10 em todos os conjuntos de dados (0.61-0.65), indicando que suas representações aprendidas mantêm melhor as relações de vizinhança local quando projetadas em espaços de menor dimensão. Isso sugere que o mecanismo de auto-atenção cria embeddings com estrutura local mais coerente.
- **Menor Distorção de Distância no SASRec:** As pontuações LDD@10 mais baixas do SASRec (2.17-2.28) demonstram que ele preserva melhor as distâncias relativas entre itens vizinhos. Isso é particularmente importante para tarefas de recomendação, onde a proximidade no espaço de embedding deve refletir similaridade real entre itens.

- **Trade-off entre Conhecimento e Geometria:** Interessantemente, embora o KGAT supere o LightGCN em métricas de recomendação, ele apresenta pontuações de NP ligeiramente inferiores em alguns casos. Isso sugere que a incorporação de conhecimento externo, embora benéfica para a precisão da recomendação, pode introduzir alguma complexidade geométrica adicional no espaço de embedding.

A análise geométrica quantitativa corrobora nossas observações visuais e fornece evidências adicionais de que diferentes arquiteturas impõem estruturas geométricas distintas em seus espaços latentes. A superior preservação de estrutura local do SASRec pode contribuir para seu desempenho excepcional em tarefas de previsão sequencial, onde manter relações temporais precisas entre itens é crucial.

6. Conclusão e Trabalhos Futuros

Neste trabalho apresentamos uma análise comparativa de três paradigmas de recomendação sob uma perspectiva de manifold learning. Nossos experimentos produziram uma conclusão clara: a arquitetura de modelo ótima está fundamentalmente ligada à tarefa específica de recomendação. Demonstramos empiricamente que, para previsão do próximo item, a informação sequencial capturada por um modelo baseado em Transformer como o SASRec é melhor, levando a um desempenho superior. Para recomendação geral, mostramos que aumentar o grafo usuário-item com conhecimento externo (como no KGAT) fornece uma vantagem significativa sobre abordagens de filtragem colaborativa pura (como o LightGCN). Finalmente, nossa análise geométrica confirmou visualmente que cada arquitetura impõe uma estrutura topológica única no espaço de embedding, refletindo o tipo de informação—colaborativa, semântica ou sequencial que é usada para modelar. Nossos resultados mostram que modelos baseados em grafo e sequenciais capturam aspectos distintos, mas complementares da preferência do usuário. Esse resultado aponta para o desenvolvimento de modelos híbridos que fundem essas abordagens. Propomos uma arquitetura que combina uma GNN com um Transformer. Nesse modelo, uma GNN como, por exemplo, o LightGCN, primeiro gera embeddings cientes da estrutura a partir do grafo global usuário-item. Esses embeddings enriquecidos então serviriam como entrada para um modelo sequencial como, por exemplo, o SASRec, permitindo que ele aprenda padrões temporais a partir de representações que já estão cientes do contexto colaborativo de cada item. Esse modelo híbrido poderia potencialmente capturar tanto o perfil de gosto de longo prazo de um usuário quanto sua intenção de curto prazo, levando a um sistema de recomendação mais robusto e preciso.

References

- Adomavicius, G. and Tuzhilin, A. (2011). Context-aware recommender systems. In *Recommender systems handbook*, pages 217–253. Springer.
- Barkan, O., Caciularu, A., Katz, A., and Koenigstein, N. (2020). Geometric understanding of deep learning embeddings for recommender systems. In *Proceedings of the 13th International Conference on Web Search and Data Mining*, pages 35–43.
- Barthel, K. U., Hezel, N., Jung, K., and Schall, K. (2017). Improved evaluation and generation of grid layouts using distance preservation quality and linear assignment sorting. In *SSDBM ’17: Proceedings of the 29th International Conference on Scientific and Statistical Database Management*, pages 1–12.
- Belkin, M. and Niyogi, P. (2003). Laplacian eigenmaps for dimensionality reduction and data representation. *Neural computation*, 15(6):1373–1396.
- Bronstein, M. M., Bruna, J., Cohen, T., and Veličković, P. (2021). Geometric deep learning: Grids, groups, graphs, geodesics, and gauges.
- Chen, J., Zhang, H., He, X., Nie, L., Liu, W., and Chua, T.-S. (2017). Attentive collaborative filtering: Multimedia recommendation with item-and component-level attention. In *Proceedings of the 40th International ACM SIGIR conference on Research and Development in Information Retrieval*, pages 335–344.
- Chen, L., Wu, L., Hong, R., Zhang, K., and Wang, M. (2022). Geometric analysis of graph neural networks for collaborative filtering. In *Proceedings of the 45th International ACM SIGIR Conference on Research and Development in Information Retrieval*, pages 1456–1465.
- Covington, P., Adams, J., and Sargin, E. (2016). Deep neural networks for youtube recommendations. In *Proceedings of the 10th ACM conference on recommender systems*, pages 191–198.
- Dacrema, M. F., Cremonesi, P., and Jannach, D. (2019). Are we really making much progress? a worrying analysis of recent neural recommendation approaches. In *Proceedings of the 13th ACM Conference on Recommender Systems*, pages 101–109.
- Fan, W., Ma, Y., Li, Q., He, Y., Zhao, E., Tang, J., and Yin, D. (2019). Graph neural networks for social recommendation. In *The world wide web conference*, pages 417–426.
- Feng, F., He, X., Wang, X., Luo, C., Liu, Y., and Chua, T.-S. (2019). Temporal relational ranking for stock prediction. *ACM Transactions on Information Systems*, 37(2):1–30.
- Ferrari Dacrema, M., Boglio, S., Cremonesi, P., and Jannach, D. (2021). A troubling analysis of reproducibility and progress in recommender systems research. *ACM Transactions on Information Systems*, 39(2):1–49.
- Gao, C., Zheng, Y., Li, N., Li, Y., Qin, Y., Piao, J., Quan, Y., Chang, J., Jin, D., He, X., and Li, Y. (2023). A survey of graph neural networks for recommender systems: Challenges, methods, and directions. *ACM Transactions on Recommender Systems*, 1(1):1–50.

- Goldberg, K., Roeder, T., Gupta, D., and Perkins, C. (2001). Eigentaste: A constant time collaborative filtering algorithm. *Information retrieval*, 4(2):133–151.
- Gomez-Uribe, C. A. and Hunt, N. (2015). The netflix recommender system: Algorithms, business value, and innovation. *ACM Transactions on Management Information Systems*, 6(4):1–19.
- Guo, Q., Zhuang, F., Qin, C., Zhu, H., Xie, X., Xiong, H., and He, Q. (2020). A survey on knowledge graph-based recommender systems. *IEEE Transactions on Knowledge and Data Engineering*, 34(8):3549–3568.
- Hamilton, W. L., Ying, R., and Leskovec, J. (2017). Representation learning on graphs: Methods and applications. *IEEE Data Engineering Bulletin*, 40(3):52–74.
- He, X., Deng, K., Wang, X., Li, Y., Zhang, Y., and Wang, M. (2020). Lightgcn: Simplifying and powering graph convolution network for recommendation. In *Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval*, pages 639–648.
- He, X., Liao, L., Zhang, H., Nie, L., Hu, X., and Chua, T.-S. (2017). Neural collaborative filtering. In *Proceedings of the 26th international conference on world wide web*, pages 173–182.
- Hidasi, B., Karatzoglou, A., Baltrunas, L., and Tikk, D. (2015). Session-based recommendations with recurrent neural networks. In *International Conference on Learning Representations*.
- Kang, W.-C. and McAuley, J. (2018). Self-attentive sequential recommendation. In *IEEE International Conference on Data Mining (ICDM)*, pages 197–206.
- Koren, Y. (2008). Factorization meets the neighborhood: a multifaceted collaborative filtering model. In *Proceedings of the 14th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 426–434.
- Koren, Y. (2010). Collaborative filtering with temporal dynamics. *Communications of the ACM*, 53(4):89–97.
- Koren, Y., Bell, R., and Volinsky, C. (2009). Matrix factorization techniques for recommender systems. *Computer*, 42(8):30–37.
- Lee, D. D. and Seung, H. S. (1999). Learning the parts of objects by non-negative matrix factorization. *Nature*, 401(6755):788–791.
- Li, J., Wang, Y., and McAuley, J. (2020). Time interval aware self-attention for sequential recommendation. In *Proceedings of the 13th International Conference on Web Search and Data Mining*, pages 322–330.
- Li, Y., Chen, H., Xu, S., Ge, Y., and Zhang, Y. (2023). Information geometric analysis of neural recommendation models. In *International Conference on Machine Learning*, pages 19234–19251.
- Liang, D., Krishnan, R. G., Hoffman, M. D., and Jebara, T. (2018). Variational autoencoders for collaborative filtering. In *Proceedings of the 2018 world wide web conference*, pages 689–698.

- Liu, Z., Chen, C., Yang, X., Zhou, J., Li, X., and Song, L. (2021). Heterogeneous graph neural networks for malicious account detection.
- McInnes, L., Healy, J., and Melville, J. (2018). Umap: Uniform manifold approximation and projection for dimension reduction.
- Mikolov, T., Sutskever, I., Chen, K., Corrado, G. S., and Dean, J. (2013). Distributed representations of words and phrases and their compositionality. In *Advances in neural information processing systems*, volume 26, pages 3111–3119.
- Mirylenka, K., Dallachiesa, M., and Palpanas, T. (2017a). Data series similarity using correlation-aware measures. In *SSDBM ’17: Proceedings of the 29th International Conference on Scientific and Statistical Database Management*, pages 1–12.
- Mirylenka, K., Giannakopoulos, G., Palpanas, T., and Bouros, P. (2017b). On the evaluation of local distance preservation in embedding techniques for dimensionality reduction. In *Proceedings of the 20th International Conference on Extending Database Technology*, pages 138–149.
- Petrov, A. and Macdonald, C. (2022). A systematic evaluation and comparison of three decades of ir evaluation. In *Proceedings of the 45th International ACM SIGIR Conference on Research and Development in Information Retrieval*, pages 503–515.
- Quadrana, M., Cremonesi, P., and Jannach, D. (2018). Sequence-aware recommender systems. *ACM computing surveys*, 51(4):1–36.
- Rendle, S., Freudenthaler, C., and Schmidt-Thieme, L. (2010). Factorizing personalized markov chains for next-basket recommendation. In *Proceedings of the 19th international conference on World wide web*, pages 811–820.
- Rendle, S., Krichene, W., Zhang, L., and Anderson, J. (2020). Neural collaborative filtering vs. matrix factorization revisited. In *Proceedings of the 14th ACM Conference on Recommender Systems*, pages 240–248.
- Resnick, P., Iacovou, N., Suchak, M., Bergstrom, P., and Riedl, J. (1994). Grouplens: an open architecture for collaborative filtering of netnews. In *Proceedings of the 1994 ACM conference on Computer supported cooperative work*, pages 175–186.
- Roweis, S. T. and Saul, L. K. (2000). Nonlinear dimensionality reduction by locally linear embedding. *Science*, 290(5500):2323–2326.
- Sarwar, B., Karypis, G., Konstan, J., and Riedl, J. (2001). Item-based collaborative filtering recommendation algorithms. In *Proceedings of the 10th international conference on World Wide Web*, pages 285–295.
- Su, X. and Khoshgoftaar, T. M. (2009). A survey of collaborative filtering techniques. *Advances in artificial intelligence*, 2009.
- Sun, F., Liu, J., Wu, J., Pei, C., Lin, X., Ou, W., and Jiang, P. (2019). Bert4rec: Sequential recommendation with bidirectional encoder representations from

- transformer. In *Proceedings of the 28th ACM International Conference on Information and Knowledge Management*, pages 1441–1450.
- Sun, Z., Yu, D., Fang, H., Yang, J., Qu, X., Zhang, J., and Geng, C. (2020). Are we evaluating rigorously? benchmarking recommendation for reproducible evaluation and fair comparison. In *Proceedings of the 14th ACM Conference on Recommender Systems*, pages 23–32.
- Tenenbaum, J. B., De Silva, V., and Langford, J. C. (2000). A global geometric framework for nonlinear dimensionality reduction. *Science*, 290(5500):2319–2323.
- van den Berg, R., Kipf, T. N., and Welling, M. (2017). Graph convolutional matrix completion.
- van der Maaten, L. and Hinton, G. (2008a). Visualizing data using t-sne. *Journal of machine learning research*, 9(11):2579–2605.
- van der Maaten, L. J. P. and Hinton, G. E. (2008b). Visualizing data using t-sne. *Journal of Machine Learning Research*, 9(1):2579–2605.
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., and Polosukhin, I. (2017). Attention is all you need. *Advances in neural information processing systems*, 30.
- Venna, J. and Kaski, S. (2006). Neighborhood preservation in nonlinear projection methods: An experimental study. In *International Conference on Artificial Neural Networks*, pages 485–491.
- Wang, H., Zhang, F., Wang, J., Zhao, M., Li, W., Xie, X., and Guo, M. (2018). Ripplenet: Propagating user preferences on the knowledge graph for recommender systems. In *Proceedings of the 27th ACM international conference on information and knowledge management*, pages 417–426.
- Wang, H., Zhang, F., Zhang, M., Leskovec, J., Zhao, M., Li, W., and Wang, Z. (2019a). Knowledge-aware graph neural networks with label smoothness regularization for recommender systems. In *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, pages 968–977.
- Wang, H., Zhao, M., Xie, X., Li, W., and Guo, M. (2019b). Knowledge graph convolutional networks for recommender systems. In *The world wide web conference*, pages 3307–3313.
- Wang, J., Yu, L., Zhang, W., Gong, Y., Xu, Y., Wang, B., Zhang, P., and Chen, D. (2017). Irgan: A minimax game for unifying generative and discriminative information retrieval models. In *Proceedings of the 40th International ACM SIGIR conference on Research and Development in Information Retrieval*, pages 515–524.
- Wang, S., Cao, L., Wang, Y., Sheng, Q. Z., Orgun, M. A., and Lian, D. (2023). A survey on session-based recommender systems. *ACM Computing Surveys*, 54(7):1–38.

- Wang, X., He, X., Cao, Y., Liu, M., and Chua, T.-S. (2019c). Kgat: Knowledge graph attention network for recommendation. In *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, pages 950–958.
- Wang, X., He, X., Wang, M., Feng, F., and Chua, T.-S. (2019d). Neural graph collaborative filtering. In *Proceedings of the 42nd international ACM SIGIR conference on Research and development in Information Retrieval*, pages 165–174.
- Wang, X., Huang, T., Wang, D., Yuan, Y., Liu, Z., He, X., and Chua, T.-S. (2021). Learning intents behind interactions with knowledge graph for recommendation. In *Proceedings of the Web Conference 2021*, pages 878–887.
- Wang, X., Jin, H., Zhang, A., He, X., Xu, T., and Chua, T.-S. (2020). Disentangled graph collaborative filtering. In *Proceedings of the 43rd International ACM SIGIR conference on research and development in Information Retrieval*, pages 1001–1010.
- Wu, J., Wang, X., Feng, F., He, X., Chen, L., Lian, J., and Xie, X. (2021). Self-supervised graph learning for recommendation. In *Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval*, pages 726–735.
- Wu, L., Li, S., Hsieh, C.-J., and Sharpnack, J. (2020). Sse-pt: Sequential recommendation via personalized transformer. In *Proceedings of the 14th ACM Conference on Recommender Systems*, pages 328–337.
- Wu, S., Sun, F., Zhang, W., and Cui, B. (2022). A survey of graph neural networks for recommender systems: Challenges, methods, and directions. *ACM Transactions on Recommender Systems*.
- Yao, T., Yi, X., Cheng, D. Z., Yu, F., Chen, T., Menon, A., Hong, L., Chi, E. H., Tjoa, S., Kang, J., et al. (2021). Self-supervised learning for large-scale item recommendations. In *Proceedings of the 30th ACM International Conference on Information & Knowledge Management*, pages 4321–4330.
- Zhang, M., Wu, S., Gao, M., Jiang, X., Xu, K., and Wang, L. (2022). Personalized graph neural networks with attention mechanism for session-based recommendation. *IEEE Transactions on Knowledge and Data Engineering*, 34(8):3946–3957.
- Zhang, S., Yao, L., Sun, A., and Tay, Y. (2019). Deep learning based recommender system: A survey and new perspectives. *ACM computing surveys*, 52(1):1–38.
- Zhou, K., Wang, H., Zhao, W. X., Zhu, Y., Wang, S., Zhang, F., Wang, Z., and Wen, J.-R. (2020). S³-rec: Self-supervised learning for sequential recommendation with mutual information maximization. In *Proceedings of the 29th ACM International Conference on Information & Knowledge Management*, pages 1893–1902.