Geometria dos Embeddings em Recomendação: GNNs e Transformers sob Manifold Learning

Arthur Andreazza Marinho[1], **Daniel Carlos Guimarães Pedronette**[1]

[1]Universidade Estadual Paulista (UNESP) – Rio Claro, SP, Brasil

arthur.andreazza-marinho@unesp.br, daniel.pedronette@unesp.br

Abstract. *We compare three recommender paradigms - LightGCN, KGAT, and SASRec - from a manifold learning perspective, characterizing how each architecture shapes the geometry of learned embeddings. Two evaluation protocols are used: chronological next-item prediction and general top-K recommendation, both under full-corpus ranking. Neighborhood Preservation (NP@10) and Local Distance Distortion (LDD@10) are defined with the item interaction graph (Jaccard) as the reference space. Across MovieLens-1M, Yelp2018, and Amazon-Book, SASRec dominates next-item prediction, while LightGCN matches or surpasses KGAT in general recommendation, contradicting the assumption that knowledge graphs systematically improve performance. Geometric analysis sheds light on this finding.*

Resumo. *Foram comparados três paradigmas de recomendação - LightGCN, KGAT e SASRec - sob a perspectiva de manifold learning, com o intuito de verificar como cada arquitetura molda a geometria dos embeddings aprendidos. Foram adotados dois protocolos de avaliação: previsão cronológica do próximo item e recomendação geral top-K, ambos com ranqueamento full-corpus. As métricas de Preservação de Vizinhança (NP@10) e Distorção de Distância Local (LDD@10) são definidas com o grafo de interação de itens (Jaccard) como espaço de referência. Em MovieLens-1M, Yelp2018 e Amazon-Book, o SASRec domina a previsão sequencial, enquanto o LightGCN iguala ou supera o KGAT em recomendação geral, contrariando a hipótese de que grafos de conhecimento aprimoram sistematicamente o desempenho. A análise geométrica esclarece esse resultado.*

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 [Wu et al. 2022]. O desafio central desses sistemas reside na modelagem eficaz das complexas interações usuário-item, que frequentemente exibem propriedades não-euclidianas e estruturas de *manifolds* intrincadas [Belkin and Niyogi 2003, Wu et al. 2022].

As *Graph Neural Networks* (GNNs) reconhecem que a maioria dos dados em recomendação possui estrutura de grafo essencial [Wu et al. 2022, Gao et al. 2023]. As interações usuário-item podem ser representadas como um grafo bipartido, onde as conexões representam preferências implícitas ou explícitas [Wang et al. 2019b]. O Trans-

former [Vaswani et al. 2017] tornou-se dominante em modelagem sequencial, como em SASRec [Kang and McAuley 2018].

Apesar dos avanços, há uma lacuna na compreensão de como diferentes arquiteturas influenciam a geometria dos *embeddings* aprendidos. A perspectiva de *manifold learning* [Roweis and Saul 2000, Tenenbaum et al. 2000] oferece um *framework* teórico para analisar essa estrutura intrínseca.

Este trabalho preenche essa lacuna através de uma análise comparativa de três paradigmas arquiteturais: filtragem colaborativa baseada em grafos (LightGCN), redes de grafos com conhecimento (KGAT) e Transformers sequenciais (SASRec). Para avaliar sistematicamente como os priors arquiteturais de cada modelo influenciam a estrutura geométrica dos espaços de *embedding*, são empregadas métricas quantitativas de Preservação de Vizinhança (NP) e Distorção de Distância Local (LDD), complementadas por visualizações qualitativas via t-SNE (*t-distributed stochastic neighbor embedding*).

2. Trabalhos Relacionados

A evolução dos sistemas de recomendação, de métodos de vizinhança e fatoração de matrizes a *deep learning*, reflete a busca por representações mais expressivas, culminando no reconhecimento de que dados de recomendação possuem estrutura de grafo inerente [Wu et al. 2022, Gao et al. 2023].

O *Neural Graph Collaborative Filtering* (NGCF) [Wang et al. 2019b] introduziu sinal colaborativo explícito via propagação de *embeddings* no grafo de interações. O LightGCN [He et al. 2020] simplificou esse *design*, mostrando que apenas a agregação de vizinhança é essencial, com ganhos de aproximadamente 16% sobre o NGCF – evidência de que a topologia do grafo basta para modelar preferências.

O KGAT [Wang et al. 2019a] integra grafos de conhecimento à filtragem colaborativa via relações semânticas entre entidades [Guo et al. 2020].

A recomendação sequencial, por sua vez, reconhece que preferências evoluem temporalmente: o SASRec [Kang and McAuley 2018] adapta a arquitetura Transformer para capturar dependências de longo alcance em sequências de comportamento, complementando as abordagens baseadas em grafos.

A perspectiva de *manifold learning* [Roweis and Saul 2000, Tenenbaum et al. 2000] fornece ferramentas para compreender como esses paradigmas organizam suas representações: t-SNE [van der Maaten and Hinton 2008] permite visualizá-las, enquanto *Neighborhood Preservation* [Venna and Kaski 2006] quantifica sua fidelidade estrutural. Trabalhos anteriores mostram que preferências de usuários residem em *manifolds* de baixa dimensão [Belkin and Niyogi 2003], motivando a investigação de como diferentes arquiteturas preservam essa estrutura.

3. Métodos

LightGCN [He et al. 2020] representa a filtragem colaborativa baseada em grafo puro, simplificando GCNs tradicionais ao remover transformações de características e ativações não-lineares. Sua operação central propaga linearmente *embeddings* no grafo de

interação, dada por

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

onde, $\mathbf{e}_u^{(k)}, \mathbf{e}_i^{(k)} \in \mathbb{R}^d$ são os *embeddings* de dimensão d do usuário u e do item i após k passos de propagação no grafo bipartido de interações, $\alpha_k \in \mathbb{R}$ é o peso escalar atribuído à k-ésima camada e $K \in \mathbb{N}$ é o número total de camadas. Seguindo o artigo original, foi adotado $\alpha_k = 1/(K+1)$ uniforme.

KGAT [Wang et al. 2019a] integra o grafo de interação com um grafo de conhecimento externo, empregando atenção durante a passagem de mensagens, representada por

$$\pi(h, r, t) = (\mathbf{W}_r \mathbf{e}_t)^\top \tanh(\mathbf{W}_r \mathbf{e}_h + \mathbf{e}_r), \quad \alpha(h, r, t) = \mathrm{softmax}_{\mathcal{N}_h} \pi(h, r, t) \tag{2}$$

onde, h, r e t denotam, respectivamente, a *head* (entidade origem), a relação e a *tail* (entidade destino) de uma tripla do grafo de conhecimento; $\mathbf{e}_h, \mathbf{e}_t \in \mathbb{R}^d$ são *embeddings* de entidades e $\mathbf{e}_r \in \mathbb{R}^{d_r}$ é o *embedding* da relação r; $\mathbf{W}_r \in \mathbb{R}^{d_r \times d}$ é a matriz de projeção TransR específica da relação r e $\mathcal{N}_h$ denota a vizinhança de h no grafo. A função $\tanh$ é a tangente hiperbólica. Os pesos $\alpha(h, r, t)$ são recomputados a cada época de treinamento e usados como pesos de agregação na camada *Bi-Interaction*.

SASRec [Kang and McAuley 2018] modela o histórico de interações como uma sequência ordenada, usando auto-atenção para prever o próximo item, dado por

$$\mathrm{Attention}(\mathbf{Q}, \mathbf{K}, \mathbf{V}) = \mathrm{softmax}\left(\frac{\mathbf{Q}\mathbf{K}^\top}{\sqrt{d_k}}\right)\mathbf{V} \tag{3}$$

onde, $\mathbf{Q}, \mathbf{K}, \mathbf{V} \in \mathbb{R}^{n \times d_k}$ são, respectivamente, as matrizes de *query*, *key* e *value*, obtidas via transformações lineares dos *embeddings* de entrada da sequência, n é o comprimento da sequência, d_k é a dimensão das chaves, com o fator $1/\sqrt{d_k}$ estabilizando os gradientes e $\mathrm{softmax}(\cdot)$ é aplicado linha a linha.

4. Experimentos

4.1. Conjuntos de Dados

Os experimentos foram conduzidos em três *datasets* públicos (Tabela 1), selecionados por diferentes regimes de densidade e por permitirem grafos de conhecimento de qualidade distinta. Yelp2018 e Amazon-Book apresentam interações esparsas (0,057% e 0,048%, respectivamente) e dispõem de grafos de conhecimento (KG) ricos derivados do *Freebase* pelos autores do KGAT [Wang et al. 2019a], com 42 e 39 relações, respectivamente. O MovieLens-1M é denso (4,84%) mas não possui grafo de conhecimento externo associado para o KGAT. Foi construído um grafo simples a partir dos metadados de gênero (1 relação *has_genre*, 18 gêneros, 12.384 triplas), reconhecendo a baixa expressividade desse KG como uma limitação. O MovieLens-1M é empregado em ambos os protocolos: divisão cronológica *leave-one-out* no Protocolo A e divisão aleatória 80/20 por usuário no Protocolo B. Yelp2018 e Amazon-Book entram apenas no Protocolo B, com as divisões aleatórias disponibilizadas pelos autores do KGAT [Wang et al. 2019a].

Tabela 1. Conjuntos de Dados Utilizados

Dataset	# Usuários	# Itens	# Interações	Densidade
Yelp2018	45.919	45.538	1.183.610	0,057%
Amazon-Book	70.679	24.915	846.434	0,048%
MovieLens-1M	6.040	3.416	999.611	4,84%

4.2. Configuração Experimental

Foram adotados dois protocolos complementares para evitar a assimetria inerente à comparação direta entre modelos de recomendação geral (LightGCN, KGAT) e sequenciais (SASRec). **Protocolo A — previsão do próximo item**: divisão cronológica *leave-one-out* no MovieLens-1M (único *dataset* com *timestamps* válidos), avaliando os três modelos sob ranqueamento contra todo o catálogo de itens. **Protocolo B — recomendação geral**: divisão aleatória 80/20 nos três *datasets*, avaliando LightGCN e KGAT, também com ranqueamento *full-corpus*. Ambos os protocolos substituem o ranqueamento amostrado (n = 100) tipicamente usado em [Kang and McAuley 2018], evitando inconsistências documentadas em avaliações amostradas [Krichene and Rendle 2020].

Hiperparâmetros e robustez: todos os modelos seguem as configurações dos artigos originais (LightGCN: `embed_dim` = 64, 3 camadas, α_k uniforme e $\lambda = 10^{-4}$; KGAT: `embed_dim` = 64, camadas (64,32,16), agregador *bi-interaction* e atenção atualizada a cada época; SASRec: `hidden_units` = 50, 2 blocos e `maxlen` = 200). Treinamento até 1000 épocas (LightGCN, KGAT) ou 200 (SASRec) com *early stopping* por estagnação da métrica de validação principal (paciência de 100 épocas para LightGCN, 50 épocas para KGAT e 100 épocas para SASRec). Cada experimento foi repetido com três sementes distintas, exceto Amazon-Book, executado com semente única devido às restrições computacionais. Resultados estendidos serão disponibilizados publicamente no repositório do projeto, com média ± desvio padrão. Visualizações t-SNE usam `perplexity` = 30, `n_iter` = 1000, `init=pca` e `random_state` = 42.

Métricas geométricas: NP@10 e LDD@10 são computadas com o grafo de interação como espaço de referência. Dado um item i, seus k vizinhos de referência são os itens com maior similaridade de Jaccard $|U(i) \cap U(j)|/|U(i) \cup U(j)|$, onde $U(\cdot)$ é o conjunto de usuários que interagiram com cada item. No espaço de *embedding*, a similaridade é o cosseno. NP@k mede a fração de vizinhos preservados e LDD@k mede a distorção relativa das distâncias locais. Foram amostrados 1000 itens por *dataset*, mantendo a mesma amostra para t-SNE e métricas.

5. Resultados e Análise

Para a previsão do próximo item (Tabela 2), o SASRec domina com Recall@20 superior aos modelos baseados em grafo. O *gap* reflete a operação do LightGCN e KGAT fora de seu domínio de projeto: nenhum modela ordem temporal. A pequena diferença entre eles (HR@10 de 0,084 vs 0,079) sugere que o KG *has_genre* construído para o MovieLens-1M é insuficiente para enriquecer as representações de forma significativa.

Sob o protocolo de recomendação geral (Tabela 3), o LightGCN iguala ou supera o KGAT em todos os três *datasets*, contrariando a hipótese de que grafos de conheci-

Tabela 2. Protocolo A: Previsão do Próximo Item no MovieLens-1M (*leave-one-out* cronológico), média ± desvio padrão sobre 3 sementes. Avaliação *full-corpus*.

Modelo	HR@10	NDCG@10	Recall@20	NDCG@20
LightGCN	0,0836 ± 0,0021	0,0414 ± 0,0005	0,1494 ± 0,0009	0,0579 ± 0,0001
KGAT	0,0791 ± 0,0030	0,0390 ± 0,0013	0,1357 ± 0,0070	0,0532 ± 0,0022
SASRec	**0,2493 ± 0,0016**	**0,1340 ± 0,0008**	**0,3639 ± 0,0013**	**0,1628 ± 0,0006**

mento aprimoram sistematicamente a recomendação. No MovieLens-1M os modelos não diferem, dentro do desvio padrão, e nos *datasets* com KG rico do Freebase (Yelp2018 e Amazon-Book), o LightGCN supera o KGAT em Recall@20 por margens próximas de 15%.

Tabela 3. Protocolo B: Recomendação Geral, divisões aleatórias 80/20. Para MovieLens-1M e Yelp2018, média ± desvio padrão sobre 3 sementes; Amazon-Book usa semente única (2020) devido a restrições computacionais. Avaliação *full-corpus*.

Dataset	Modelo	HR@10	Recall@20	NDCG@10
ML-1M	LightGCN	**0,9151 ± 0,0011**	**0,2470 ± 0,0005**	**0,3908 ± 0,0001**
	KGAT	0,9139 ± 0,0012	0,2427 ± 0,0007	0,3889 ± 0,0018
Yelp2018	LightGCN	**0,1816 ± 0,0009**	**0,0730 ± 0,0004**	**0,0373 ± 0,0003**
	KGAT	0,1605 ± 0,0001	0,0637 ± 0,0001	0,0320 ± 0,0001
Amazon-Book	LightGCN	**0,1725**	**0,1421**	**0,0614**
	KGAT	0,1449	0,1237	0,0493

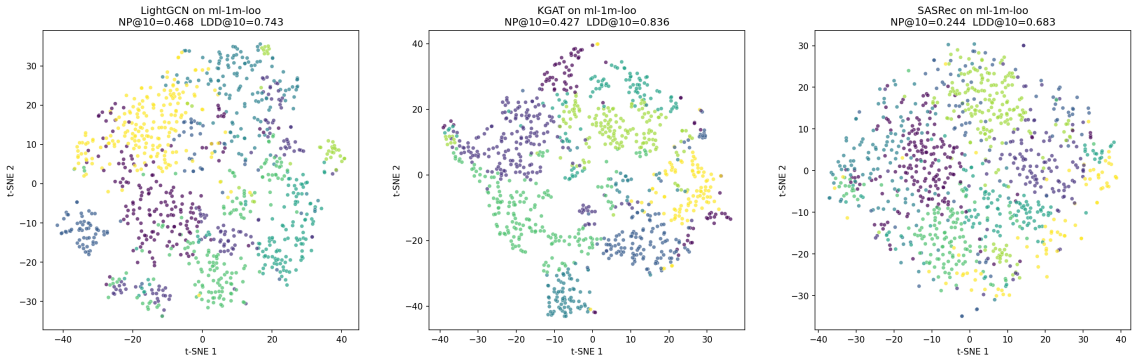
A análise geométrica (Tabela 4) ajuda a explicar essa contradição. O LightGCN mantém NP@10 mais alto, preservando vizinhança colaborativa - precisamente o sinal predominante em ranking top-K. O KGAT exibe NP@10 substancialmente menor (em Yelp2018, aproximadamente 3,5× menor que o LightGCN), indicando reorganização do espaço em torno da estrutura semântica do KG ao custo da preservação colaborativa. O SASRec apresenta LDD@10 mais baixo, compatível com sua modelagem sequencial. A correlação entre baixo NP e menor desempenho de ranking sugere uma explicação geométrica para os resultados da Tabela 3: o *trade-off* imposto pelo KGAT não se traduz em ganhos quando o sinal colaborativo é central à tarefa. As projeções t-SNE (Figura 1) sugerem organizações latentes consistentes com essas medidas, devendo ser interpretadas como indicativas.

Os padrões observados têm interpretação direta no contexto de redes sociais. O grafo bipartido usuário-item é estruturalmente análogo a redes sociais bipartidas (usuário-conteúdo, usuário-grupo). A preservação de vizinhança pelo LightGCN sugere que GNNs simples podem ser suficientes para capturar estrutura de comunidade, enquanto a reorganização agressiva do KGAT é análoga ao efeito que atributos de perfil rico podem ter sobre representações sociais aprendidas.

Tabela 4. Métricas Geométricas (NP@10 / LDD@10) com grafo de interação (Jaccard) como referência.

Modelo	MovieLens-1M	Yelp2018	Amazon-Book
LightGCN	0,458 / 0,728	0,234† / 0,740†	0,292 / 0,719
KGAT	0,418 / 0,832	0,068 / 0,768	0,186 / 0,709
SASRec	0,243 / 0,686	—	—

† 2 sementes; análise geométrica da terceira semente não executada.



(a) LightGCN (b) KGAT (c) SASRec

Figura 1. Projeção t-SNE dos embeddings de itens no MovieLens-1M, coloridos por agrupamento KMeans no espaço de *embedding* original (`perplexity=30`, `init=pca`, `random_state=42`). Os padrões qualitativos sugerem organizações latentes distintas, consistentes com os diferentes objetivos de aprendizado.

6. Conclusão

Este trabalho comparou três paradigmas de recomendação: LightGCN, KGAT e SASRec, sob a perspectiva de *manifold learning*, integrando avaliação preditiva e análise geométrica. SASRec domina a previsão do próximo item, refletindo sua adequação à modelagem sequencial. Em recomendação geral, contrariando uma hipótese comum, o LightGCN iguala ou supera o KGAT consistentemente nos três *datasets*, mesmo quando há grafos de conhecimento ricos disponíveis. A análise geométrica esclarece esse achado: o KGAT troca preservação de vizinhança colaborativa por estrutura semântica derivada do KG, mas esse *trade-off* não se converte sistematicamente em ganhos preditivos.

A análise geométrica focou em *embeddings* de itens; uma extensão natural é caracterizar os de usuários e sua relação com bolhas de filtro. O resultado do Amazon-Book usa semente única por restrições computacionais, mas resultados estendidos poderão ser obtidos do código-fonte, disponível em `https://github.com/Concius/brasnam2026-manifold-recsys`.

Agradecimentos

Apoio: FAPESP (procs. 2025/07727-0 e 2024/04890-5), CNPq (proc. 313193/2023-1). Uso de IA generativa em revisão de redação e suporte na implementação experimental.

Referências

Belkin, M. and Niyogi, P. (2003). Laplacian eigenmaps for dimensionality reduction and data representation. *Neural computation*, 15(6):1373–1396.

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):1–51.

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.

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.

Kang, W.-C. and McAuley, J. (2018). Self-attentive sequential recommendation. In *IEEE International Conference on Data Mining (ICDM)*, pages 197–206.

Krichene, W. and Rendle, S. (2020). On sampled metrics for item recommendation. In *Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, pages 1748–1757.

Roweis, S. T. and Saul, L. K. (2000). Nonlinear dimensionality reduction by locally linear embedding. *Science*, 290(5500):2323–2326.

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 der Maaten, L. J. P. and Hinton, G. E. (2008). 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, X., He, X., Cao, Y., Liu, M., and Chua, T.-S. (2019a). 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. (2019b). Neural graph collaborative filtering. In *Proceedings of the 42nd international ACM SIGIR conference on Research and development in Information Retrieval*, pages 165–174.

Wu, S., Sun, F., Zhang, W., Xie, X., and Cui, B. (2022). Graph Neural Networks in Recommender Systems: A Survey. *ACM Computing Surveys*, 55(5):1–37.