AVALIANDO A COMBINAÇÃO DE SISTEMAS DE RECOMENDAÇÃO CIENTES DE SESSÃO E SENSÍVEIS AO CONTEXTO

Giovani Oliveira Santos (PIBIC/CNPq/FA/UEM), Matheus Henrique Cecilio Leme (Coorientador), Marcos Aurelio Domingues (Orientador). E-mail: ra133166@uem.br.

Universidade Estadual de Maringá, Centro de Tecnologia, Maringá, PR.

Área e subárea do conhecimento: Ciências Exatas e da Terra / Ciência da Computação.

Palavras-chave: Sistemas de Recomendação; Redes Neurais em Grafos; Atenção Sequencial.

RESUMO

A sobrecarga de informação e a complexidade das interações dificultam a personalização em Sistemas de Recomendação Baseados em Sessão. Esta pesquisa avalia 16 combinações entre quatro Redes Neurais em Grafos (GCN, GAT, GraphSAGE e GIN) e quatro métodos de aprendizado auto-supervisionado (DGI, GRACE, LE e VGAE) e sua hibridização com o modelo de atenção sequencial SASRec para a recomendação do próximo item. Foram utilizados dados dos domínios musical, comercial e audiovisual, provenientes dos conjuntos Music4All, RetailRocket e The Movies Dataset. Os resultados indicaram que o desempenho das GNNs isoladas variou conforme o conjunto de dados avaliado, destacando-se o GraphSAGE no domínio musical e a GAT no domínio comercial. A hibridização com o SASRec apresentou melhorias em todos os cenários avaliados, alcançando HR@10 de 0,337 no domínio musical, 0,511 no comércio eletrônico e 0,086 no audiovisual. O VGAE apresentou-se como a estratégia de treinamento mais consistente para a inicialização estrutural do modelo sequencial. Assim, a pesquisa mostrou que a combinação entre representações topológicas e modelagem temporal é mais eficaz do que o uso isolado de GNNs.

INTRODUÇÃO

A expansão digital intensificou a sobrecarga de informação, tornando os Sistemas de Recomendação (SRs) essenciais à personalização. SRs Baseados em Sessão utilizam interações recentes para inferir interesses e prever itens subsequentes. Uma abordagem adotada para representar essas interações consiste em modelar as sessões como grafos, nos quais os itens são nós e as transições entre eles são arestas, o que permite capturar relações topológicas que vão além da ordem linear das sequências (WU et al., 2019). Nesse contexto, Redes Neurais em Grafos (GNNs) podem produzir representações vetoriais (*embeddings*) capazes de incorporar características estruturais dos itens. Porém, essas representações não

capturam necessariamente a dinâmica temporal estrita das sessões. Para explorar a complementaridade entre as duas perspectivas, esta pesquisa utiliza os *embeddings* produzidos pelas GNNs como inicialização para um modelo sequencial baseado em autoatenção, adaptado da arquitetura SASRec (KANG; MCAULEY, 2018).
O objetivo dessa pesquisa é avaliar a qualidade das representações geradas por diferentes arquiteturas de GNN e métodos de aprendizado auto-supervisionado, verificando sua eficácia tanto isoladamente quanto após a hibridização com o SASRec. São comparados quatro modelos de GNN e quatro métodos de treinamento em três domínios de aplicação. Busca-se, ainda, identificar quais configurações apresentam maior capacidade de adaptação ao refinamento temporal realizado pelo modelo sequencial.

MATERIAIS E MÉTODOS

Nesta pesquisa, foram utilizados os conjuntos de dados Music4All (musical), Retail Rocket (comercial) e The Movies (audiovisual). Nos domínios de música e comercial, as interações foram segmentadas em sessões, com um intervalo de inatividade de 30 minutos. No domínio audiovisual, utilizou-se o *Time Span Partitioning* (TSP), com janelas de 30 dias, para representar comportamentos de consumo ao longo do tempo (LUO et al., 2026). Como forma de pré-processamento, foram removidas as sessões com menos de três interações e os usuários com apenas uma sessão; no conjunto audiovisual, foram mantidas apenas as avaliações iguais ou superiores a 3,0. A partir das interações, construiu-se um grafo direcionado, no qual os nós representam itens e as arestas representam as transições de um item a outro. As características iniciais dos nós foram obtidas com o *Skip-Gram*. Com base na modelagem dos dados em grafo, foram treinadas quatro arquiteturas (LEME et al., 2026): GCN, GAT, GraphSAGE e GIN. A GCN realiza agregação normalizada dos vizinhos, a GAT utiliza pesos de atenção, o GraphSAGE usa amostragem de vizinhos e a GIN busca discriminação estrutural. Cada arquitetura foi combinada com quatro métodos auto-supervisionados: DGI, GRACE, Laplacian Eigenmaps e VGAE (LEME et al., 2026). Assim, avaliaram-se 16 configurações de extração. Os *embeddings* resultantes foram utilizados como inicialização da camada de entrada do SASRec e permaneceram ajustáveis durante o treinamento, permitindo que o modelo sequencial refinasse as representações ao longo da ordem temporal das interações.
A avaliação utilizou separação cronológica entre treinamento e teste e o protocolo *leave-one-out* para a tarefa de predição do próximo item da sessão atual de cada usuário. Os resultados (Top-10) foram mensurados pelo HR@10 (taxa de acerto do item alvo), pelo NDCG@10 (qualidade do ranqueamento, com peso maior no topo) e pelo *Coverage* (proporção do catálogo recomendado, indicando o alcance de itens de nicho). As configurações foram comparadas tanto como extratoras GNN isoladas quanto após a hibridização com o SASRec.

RESULTADOS E DISCUSSÃO

A eficácia das GNNs variou conforme as características de cada conjunto de dados. No Music4All, o GraphSAGE apresentou o melhor desempenho entre as representações estruturais isoladas. Sua amostragem de vizinhos favorece a representação de itens menos frequentes. Na configuração híbrida, a combinação com VGAE e SASRec elevou o HR@10 para 0,337.

No RetailRocket, a GAT apresentou o melhor desempenho, beneficiada pelo mecanismo de atenção, que atribui diferentes pesos às conexões. A combinação GAT com VGAE e SASRec obteve o melhor resultado geral (HR@10 = 0,511 e NDCG@10 = 0,384). Embora GIN com GRACE tenha apresentado maior cobertura de catálogo (0,708), seu HR@10 foi de 0,152. A configuração híbrida GAT com VGAE manteve maior precisão, com cobertura de 0,664, indicando melhor equilíbrio entre assertividade e diversidade. Essa comparação evidencia por que HR@10, NDCG@10 e *Coverage* foram analisados em conjunto: HR@10 e NDCG@10 avaliam a assertividade e a qualidade do ranqueamento, enquanto *Coverage* verifica se esse desempenho é alcançado sem comprometer a diversidade do catálogo recomendado.

No The Movies Dataset, as GNNs apresentaram maior dificuldade, possivelmente associada ao viés de popularidade e à subjetividade das preferências. A configuração GraphSAGE com VGAE obteve HR@10 de 0,063, enquanto a hibridização com SASRec elevou o resultado para 0,086. De modo geral, as configurações híbridas superaram as isoladas, indicando que as informações topológicas do grafo são complementares à modelagem temporal do SASRec. Entre os métodos de treinamento, o VGAE apresentou maior consistência, favorecendo a integração entre a estrutura global do grafo e a dinâmica temporal das sessões.

CONCLUSÕES

Os experimentos demonstraram que, nos cenários avaliados, o desempenho das GNNs variou conforme o conjunto de dados, destacando-se o GraphSAGE no domínio musical e a GAT no domínio comercial. Entretanto, o principal resultado da pesquisa foi a superioridade da hibridização das representações estruturais com o SASRec em relação ao uso isolado das GNNs.

O VGAE apresentou-se como a estratégia de treinamento mais consistente para fornecer a inicialização dos *embeddings*, enquanto o SASRec mostrou capacidade de refinar essas representações conforme a dinâmica temporal das sessões. Assim, a combinação de informações estruturais globais e de dependências sequenciais mostrou-se uma alternativa eficaz para a recomendação do próximo item em diferentes domínios.

AGRADECIMENTOS

O presente trabalho foi realizado com o apoio indispensável do Programa Institucional de Bolsas de Iniciação Científica (PIBIC/CNPq-Fundação Araucária-UEM), ao qual agradecemos pelo fomento e financiamento que viabilizaram o desenvolvimento desta pesquisa.

REFERÊNCIAS

HAMILTON, W.; YING, Z.; LESKOVEC, J. **Inductive representation learning on large graphs**. In: *Advances in Neural Information Processing Systems*, v. 30, 2017.

KANG, W.-C.; MCAULEY, J. **Self-attentive sequential recommendation**. In: IEEE International Conference on Data Mining (ICDM'18), 2018. p. 197–206.

LEME, M. H. C.; SANTOS, G. O.; SIMM, V. S.; TANNO, D. R.; DOMINGUES, M. A. **A comparative study of GNN architectures and self-supervised objectives for embedding extraction in recommender systems**. In: 53º Seminário Integrado de Software e Hardware (SEMISH'26), 2026. p. 13–24.

LUO, J.; ZHANG, W.; ZHANG, X.; FANG, Y. **Time-Aware Adaptive Side Information Fusion for Sequential Recommendation.** In: ACM International Conference on Web Search and Data Mining (WSDM'26), 2026. p. 469–478.

VELIČKOVIĆ, P.; CUCURULL, G.; CASANOVA, A.; ROMERO, A.; LIO, P.; BENGIO, Y. **Graph attention networks**. 2017.

WU, S.; TANG, Y.; ZHU, Y.; WANG, L.; XIE, X.; TAN, T. **Session-based recommendation with graph neural networks**. In: AAAI Conference on Artificial Intelligence (AAAI'19), 2019. p. 346–353.