

Explorando Graph Neural Networks (GNNs): Um Mapeamento Sistemático da Literatura para Análise de Domínios de Aplicação e Métricas de Avaliação

Ronilson W. S. Pereira¹

PPG-COMP/UERJ, Rio de Janeiro, RJ

Paulo A. I. Pontes²

Instituto Federal do Pará - IFPA, Belém Brasil

Marcos C. R. Seruffo³

PPGEE/UFPA, Belém, PA

Karla T. F. Leite⁴

PPG-COMP/UERJ, Rio de Janeiro, RJ

Resumo. Este artigo apresenta um mapeamento sistemático das Redes Neurais de Grafos (GNNs, do inglês *Graph Neural Networks*), destacando sua crescente relevância em diversos campos, como ciência de dados, aprendizado de máquina e redes complexas. Para modelar dados estruturados em forma de grafos, as GNNs são estruturas eficazes que permitem uma análise eficiente de informações interconectadas. Examinar os domínios de aplicações e investigar as principais métricas para validar / avaliar as GNNs é o principal objetivo deste estudo, que oferece uma visão abrangente das contribuições recentes nesse campo emergente. Para atingir este objetivo, foi realizado um Mapeamento Sistemático da Literatura (MSL), seguindo um protocolo bem definido, incluindo critérios específicos para seleção de estudos, busca abrangente em diversas fontes, análise dos resultados e síntese do conhecimento obtido. Os resultados deste mapeamento destacam uma ampla gama de aplicações de GNNs, como classificação grafo, bioinformática, medicina, ciência da computação, recomendação de itens em sistemas de recomendação, apresentando excelentes desempenhos em seus testes.

Palavras-chave. MSL, Aprendizado de Máquina, Grafos, *Graph Neural Networks*, GNNs.

1 Introdução

Nos últimos anos, as Redes Neurais de Grafos (GNNs, do inglês *Graph Neural Networks*) surgiram como uma ferramenta poderosa para análise de dados em domínios onde os grafos fornecem uma melhor representação e modelagem da estrutura de dados. Grafos, que consistem em vértices (ou nós) conectados por arestas (ou *links*), são uma estrutura matemática frequentemente usada para modelar diversos sistemas complexos, como recomendações, transportes, redes sociais, biologia molecular [7]. A capacidade das GNNs de aprender representações gráficas eficazes diretamente a partir de dados brutos despertou interesse crescente em várias comunidades de pesquisa, que podem aplicar em diferentes áreas, como aprendizado de máquina, ciência da computação, física e

¹ronilsonengenharia@gmail.com

²paulo.pontes@ifpa.edu.br

³seruffo@ufpa.br

⁴karlafigueiredo@ime.uerj.br

biologia [6]. O desenvolvimento de GNNs foi estimulado por avanços no aprendizado profundo e na teoria dos grafos [2].

Embora as tarefas, como classificação de grafos, predição de *links* e geração de grafos, possuem rápido progresso e sucesso [1], há desafios significativos relacionados à compreensão e ao aprimoramento das GNNs. Esses aspectos englobam, entre outros, as interpretabilidade dos modelos, a generalização para grafos maiores e mais complexos, entre outros. Neste contexto, o presente MSL tem como propósito fornecer uma visão ampla do estado atual da investigação sobre GNNs e fornecer um guia prático descrevendo os passos necessários para realizar este mapeamento. O objetivo principal é identificar e categorizar os principais estudos que investigam o uso das GNNs, fornecendo uma referência útil para pesquisadores interessados em investigar e contribuir para este campo dinâmico.

O restante do artigo é organizado da seguinte forma: O processo de mapeamento sistemático é detalhado na Seção 2, que abrange informações sobre os procedimentos empregados. A Seção 3 apresenta os resultados obtidos nas etapas descritas na Seção 2. Por fim, as principais conclusões alcançadas são resumidas e os trabalhos futuros explicados na Seção 4.

2 Metodologia

Uma avaliação abrangente foi requerida para o desenvolvimento desta pesquisa, visando a triagem dos artigos científicos. A metodologia empregada pretende avaliar e compreender estudos a **partir de uma Questão de Pesquisa**. O método de pesquisa utilizado no MSL foi definido com base nas orientações dadas por [4] e envolve três etapas principais: planejamento, condução e revisão, conforme se verifica na Figura 1.

Encontrar termos de pesquisa adequados é essencial para fazer uma busca adequada de estudos relevantes. Nesse sentido, os pontos de vista do PICOC (*population, intervention, comparison, outcomes, e context*), foram propostos por [4]. Para estabelecer os objetivos específicos, a metodologia PICOC foi empregado, permitindo uma estruturação mais fácil das questões de pesquisa, *string* de busca, palavras-chave, sinônimos e critérios de inclusão e exclusão [5]. Foi utilizado o *Parsifal*⁵, uma ferramenta online desenvolvida para apoiar pesquisadores na realização de revisões sistemáticas de literatura no contexto da Engenharia de *Software*.

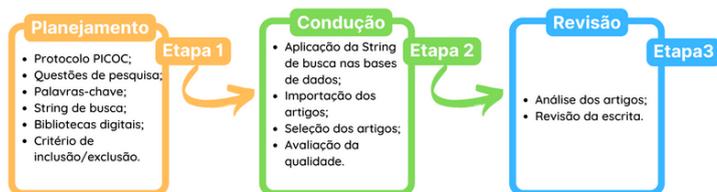


Figura 1: Etapas do desenvolvimento do Mapeamento Sistemático. Fonte: dos autores.

2.1 Planejamento

A definição das questões de pesquisa, o desenvolvimento de um protocolo de revisão e a validação desse protocolo são componentes da etapa de planejamento. Durante esta etapa, os fundamentos do MSL são definidos, onde a formulação de objetivos detalhados e a elaboração de uma questão de pesquisa a ser respondida são cruciais para qualquer projeto. Além disso, a pesquisa principal é definida pelos parâmetros de seleção que estruturam o estudo.

⁵<https://parsif.al/>

2.1.1 Protocolo PICOC

Os critérios PICOC ajudam na formulação de questões de pesquisa e dividem os objetivos do MSL em palavras-chave pesquisáveis [3]. Os cinco elementos fundamentais da metodologia PICOC são, população, intervenção, comparação, resultados e contexto. Esses elementos são utilizados para descrever todos os componentes relacionados ao problema identificado e estruturar as questões de pesquisa. A Tabela 1 apresenta a lista de elementos PICOC, suas descrições e seus termos correspondentes para GNNs.

Tabela 1: Definindo elementos da metodologia PICOC.

Elementos	Descrição	Exemplo(PICOC)
População	Uma função/paradigma que será observado.	Aplicações, Datasets, GNNs.
Intervenção	Método/procedimento que será investigado.	Otimização, Problemas, GNNs.
Comparação	Método/procedimento que será comparado.	Não se aplica.
Resultados	Uma função/paradigma que será observado.	Análises, aplicações, pesquisas em GNNs.
Contexto	Efeito que será investigado.	Bioinformática, Ciência de Dados, Análise de Redes Sociais.

2.1.2 Questões de Pesquisa

Os componentes cruciais que estabelecem a direção para a identificação do estudo e extração de dados são um tópico ou questões de pesquisa bem definidas. A partir deste ponto de vista, o objetivo do trabalho é responder às seguintes Questões de Pesquisa (QP), conforme ilustrado na Tabela 2.

Tabela 2: Questões de Pesquisa.

Questões de Pesquisa (QP)	
QP1	Em quais domínios específicos as GNNs têm sido mais aplicadas e bem-sucedidas?
QP2	Quais as métricas utilizadas para validar/avaliar as GNNs?

Definir as questões de pesquisa é uma etapa essencial do planejamento do estudo, já que estabelece claramente os objetivos a serem alcançados e oferece uma estrutura para a investigação. As questões principais foram formuladas para orientar a pesquisa em direção aos seus objetivos principais, abordando os aspectos fundamentais do problema em questão.

2.1.3 Palavras-chave, Sinônimos e String de Busca

Encontrar pesquisas relevantes envolveu a criação de termos de pesquisa baseados em questões de pesquisa usando o protocolo PICOC usado na plataforma *Parsifal*. A partir daí, é importante pensar em termos de palavras-chave e sinônimos que possam posteriormente ser utilizadas para construir consultas nas bibliotecas digitais escolhidas. As palavras-chave e sinônimos utilizados na pesquisa estão apresentadas na Tabela 3.

Tabela 3: Lista de palavras-chave e Sinônimos.

Palavras-chave	Sinônimos	Relacionado a
<i>Graph Neural Networks</i> (GNNs)	Redes Neurais de Grafos	População
<i>Graph Representation Learning</i>	Representação de Grafos.	População
<i>Heterogeneous Graphs</i>	Grafos Heterogêneos	Intervenção
Otimização para GNNs	Maximização de Desempenho	Intervenção
<i>Graph Embeddings</i>	Incorporação de Grafos	Resultado
<i>Graph-based Data Mining</i>	Mineração de dados em grafos	Resultado

Com base nas palavras-chave, realizaram-se algumas alterações para refinar a pesquisa, desta forma, uma *string* de busca foi definida conforme mostrado na Tabela 4. OR e AND são valores booleanos que separam as palavras logicamente em palavras compostas e parênteses.

Tabela 4: String de busca.

("Graph Neural Networks (GNNs)"OR "Redes Neurais de Grafos"OR "*Graph Representation Learning*"OR "Representação de Grafos") AND ("*Heterogeneous Graphs*"OR "Grafos Heterogêneos"OR "Otimização para GNNs"OR "Maximização de Desempenho") AND ("*Graph Embeddings*"OR "Incorporação de Grafos"OR "*Graph-based Data Mining*"OR "Mineração de Dados Baseada em Grafos")

2.1.4 Bibliotecas Digitais

A validade de um estudo dependerá da seleção adequada de uma base de dados, uma vez que esta deve cobrir adequadamente a área sob investigação [3]. As seguintes bases de dados foram selecionadas para a realização da revisão sistemática devido à sua reputação e reconhecimento do âmbito acadêmico e à relevância dos artigos científicos publicados. A Tabela 5 apresentada as bases de dados utilizadas.

Tabela 5: Bases de Dados Selecionadas.

Base de dados	URL	Área
<i>Springer Link</i>	http://link.springer.com	Interdisciplinar
<i>Science@Direct</i>	http://www.sciencedirect.com	Interdisciplinar
<i>ACM Digital Library</i>	http://portal.acm.org	Computação e T.I
<i>Google Scholar</i>	https://scholar.google.com.br	Interdisciplinar

2.1.5 Critérios de Inclusão/Exclusão

Neste estudo de mapeamento sistemático, critérios de inclusão e exclusão foram empregados para selecionar e desmarcar estudos das fontes de dados para responder às Questões de Pesquisa. Esses critérios da primeira etapa foram definidos da seguinte forma: Critérios Inclusivos (CI): a) Estudos que investigam GNNs como componente central de sua metodologia; b) estudos que fornecem detalhes suficientes sobre as arquiteturas, algoritmos, técnicas de treinamento e conjunto de dados utilizados em GNNs; e Critérios Exclusivos (CE): a) Estudos que não estão diretamente relacionados aos objetivos do mapeamento ou não contribuem significativamente para o entendimento das GNNs; b) Estudos que não se concentram em GNNs como principal área de investigação; c) Estudos duplicados de outros estudos já incluído no mapeamento.

2.2 Condução e Relatório da Revisão

Após estabelecer o protocolo, foi necessário adotar cada uma das etapas previamente detalhadas para realizar o mapeamento sistemático. Esta subseção explica como a condução do mapeamento ocorreu. Utilizou-se o Parsifal para definir metas e objetivos, importar artigos usando arquivos BibTeX, eliminar duplicatas, definir critérios de seleção e gerar relatórios. O objetivo consiste em localizar todos os estudos em que as questões de pesquisa foram obtidas por meio da *string* de busca. A Figura 2 ilustra o procedimento de mapeamento sistemático realizado na plataforma *Parsifal*.

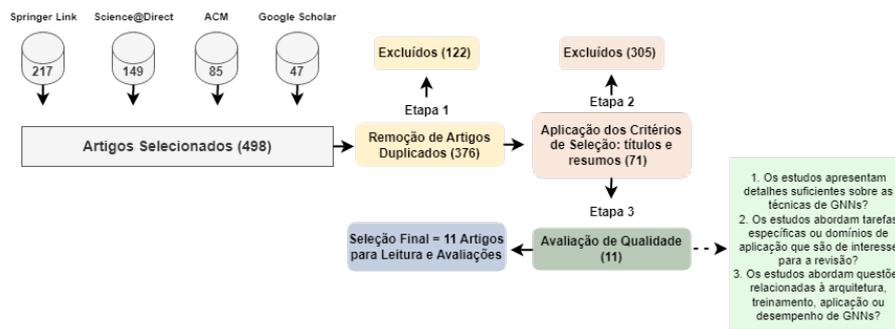


Figura 2: Ilustração esquemática da abstração dos artigos seguindo a metodologia. Fonte: dos autores.

A utilização da *string* de busca nas bibliotecas digitais escolhidas foi o primeiro passo do processo de mapeamento, que obteve um total de 498 artigos, no qual, a busca foi realizada entre os anos de 2022 e 2024 por meio da função de busca avançada disponível nas bibliotecas digitais. Conforme as com as bibliotecas pesquisadas, foram escolhidos diferentes números de artigos. Observou-se que a *Springer Link* apresentou o maior número de artigos identificados, com 217 (43,6%), enquanto o *Google Scholar* teve menos artigos, com apenas 47 (9,4%) trabalhos. Por outro lado, a *Science@Direct* e *ACM Digital Library* retornaram exatamente 149 (29,9%) e 85 (17,1%) artigos, respectivamente, totalizando cerca de 498 artigos.

Uma vez concluída a busca pelos artigos nas bibliotecas digitais e ordenadas as informações sobre os artigos retornados, foi necessário remover os artigos duplicados, excluindo um total de 122 artigos. Na etapa 2, conforme ilustrado na Figura 2, os artigos são incluídos ou retirados com base nos critérios de seleção, principalmente por meio da leitura dos títulos e resumos. Estes critérios foram rigorosamente implementados para assegurar a seleção dos estudos mais pertinentes e de ótima qualidade para a revisão, conforme mencionado na Subseção 2.1.5 deste estudo. Após a aplicação dos critérios, foram eliminados 305 artigos, resultando no final com 71 artigos considerados relevantes para o MSL.

Por fim, na etapa 3 (Figura 2), uma avaliação da qualidade desses estudos foi realizada utilizando critérios após a escolha dos 71 artigos pertinentes para o MSL sobre GNNs. Esta avaliação foi conduzida conforme as com as questões de pesquisa fornecidas na Subseção 2.1.2 deste estudo. Uma escala simples foi utilizada para a avaliação do controle de qualidade dos artigos no *Parsifal*, o procedimento de pontuação é o seguinte: SIM = 1, PARCIALMENTE = 0,5 e NÃO = 0. Para filtrar os artigos que não passam do controle de qualidade, uma pontuação de corte = 0,5 foi estabelecida, e obteve-se uma pontuação máxima de 7,1. Após a avaliação da qualidade dos 71 artigos, 11 artigos foram selecionados, então resumidos com seus pontos máximos, indicando os estudos mais destacados e significativos.

Finalmente, a fase de revisão do relatório analisa os dados sintetizados e extrai informações significativas dos artigos escolhidos. Portanto, uma análise dos trabalhos selecionados na última

etapa incluiu uma extração e o mapeamento dos dados relacionados às questões de pesquisa. A seção 3 contém a descrição desta etapa metodológica.

3 Resultados e Discussões

Os seguintes resultados foram obtidos a partir da análise dos artigos obtidos através do MSL deste estudo. As seguintes categorias foram abordadas pelos resultados com base na Questão de Pesquisa 1 (QP1) "Em quais domínios específicos as GNNs têm sido mais aplicadas e bem-sucedidas?" e Questão de Pesquisa 2 (QP2) "Quais as métricas utilizadas para validar/avaliar as GNNs?". Além disso, os 11 artigos selecionados com base nas duas questões de pesquisa estão disponíveis para acesso no seguinte repositório: <https://github.com/ronilsonpereira/Base-Dados-Quest-o-de-Pesquisa.git>.

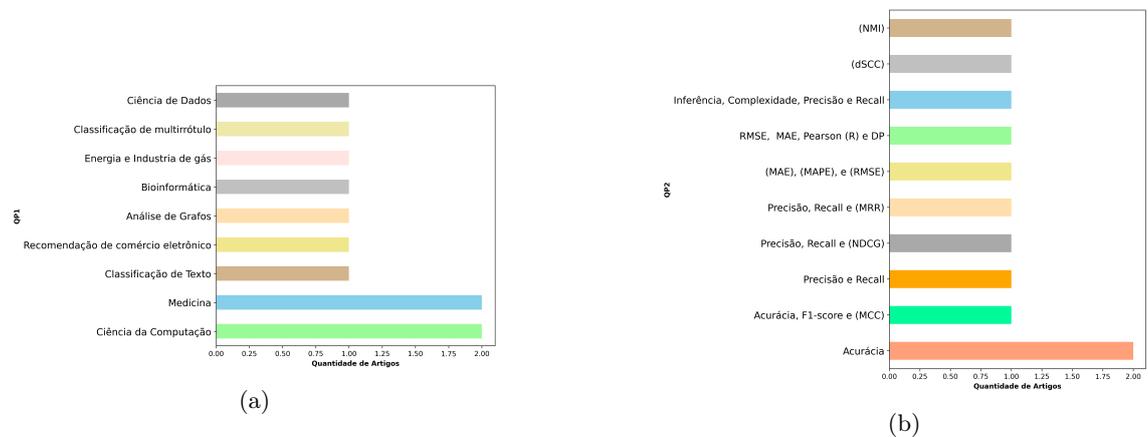


Figura 3: Distribuição dos artigos selecionados com base nas QP1(a) e QP2(b). Fonte: dos autores.

Conforme a QP1, a análise dos artigos obtidos por meio do Mapeamento Sistemático da Literatura (MSL) revelou importantes *insights* sobre os domínios em que as (GNNs) são aplicadas. Conforme ilustrado na Figura 3 (a), identificamos nove áreas específicas onde as GNNs têm sido amplamente utilizadas e bem sucedidas. Cada artigo examinado mostrou uma área específica de aplicação, mostrando a variedade de domínios em que as GNNs têm sido empregadas com sucesso. Notavelmente, dois dos estudos centraram-se em aplicações médicas, destacando o potencial das GNNs na análise e diagnóstico de condições de saúde. Além disso, dois artigos destacaram a importância das GNNs na ciência da computação, demonstrando a sua eficácia em tarefas como geração de texto e processamento de linguagem natural. O principal objetivo desta análise foi destacar os vários domínios de aplicações nos quais as GNNs foram utilizadas e implementadas com sucesso.

Os resultados relativos à Questão de Pesquisa 2 apresentaram as principais métricas utilizadas na validação e análise de GNNs nos estudos examinados, conforme mostrado na Figura 3 (b). 16 métricas distintas foram identificadas em nossa análise, sendo elas: "Acurácia, F1-score e *Matthew's Correlation Coefficient* (MCC), Precisão, Recall, *Normalized Discounted Cumulative Gain* (NDCG), *Mean Reciprocal Rank* (MRR), *Mean Absolute Error* (MAE), *Mean Absolute Percentage Error* (MAPE), *Root Mean Squared Error* (RMSE), coeficiente de correlação de Pearson (R), Desvio Padrão (DP), Inferência de Probabilidade, Complexidade do modelo, *distance Spearman Correlation Coefficient* (dSCC) e *Normalized Mutual Information* (NMI)", com a maioria dos estudos utilizando mais de uma delas. Tornou-se evidente que a seleção das métricas a serem utilizadas varia dependendo do domínio em que as GNNs são aplicadas. As métricas mais

comumente usadas para avaliar e validar GNNs é acurácia, precisão e recall. Nota-se que quatro estudos empregaram precisão e recall em conjunto com outras métricas, enquanto outra combinou a métrica de acurácia e outras. Dois estudos optaram para utilizar somente a métrica de acurácia. Os resultados mostram a relevância da seleção cuidadosa das medidas de avaliação, ajustando-as às peculiaridades de cada área onde as GNNs estão aplicadas. A avaliação do desempenho e eficácia das GNNs em cada contexto foi essencial, já que a variedade de métricas empregadas reflete a complexidade e diversidade de cenários nos quais elas são aplicadas.

4 Considerações Finais

Este artigo forneceu um guia prático descrevendo as etapas necessárias para realizar um Mapeamento Sistemático da Literatura no contexto das GNNs. O MSL apresentou uma visão abrangente e ampla sobre o estado atual da pesquisa neste campo.

De modo geral, os estudos analisados abordaram uma ampla gama de aspectos das GNNs. Observa-se um crescimento significativo na atividade de pesquisa sobre GNNs ao longo dos anos, refletido no aumento do número de publicações. Além disso, foram identificados diversos domínios de aplicação nesses estudos, incluindo ciência da computação, medicina, previsão de links, análise de redes sociais e recomendação de itens. Também foram investigadas as principais métricas utilizadas para validar e avaliar as GNNs, demonstrando que métricas específicas são aplicadas de acordo com o domínio de aplicação.

Em resumo, esta revisão sistemática ajudou a fortalecer o conhecimento atual sobre as *Graph Neural Networks*, fornecendo informações e orientações perspicazes para pesquisas futuras neste campo. Para trabalhos futuros, novas palavras podem ser adicionadas à *string* de busca e novos bancos de dados serão adicionados. Finalmente, criar um conjunto de dados específicos das GNNs, criar um *dashboard* e deixá-lo disponível para a comunidade científica.

Referências

- [1] C. Agarwal, O. Queen, H. Lakkaraju e M. Zitnik. “Evaluating explainability for graph neural networks”. Em: **Scientific Data** 10.1 (2023), p. 144.
- [2] S. Biswas. “Graph Neural Network and Its Applications”. Em: **Concepts and Techniques of Graph Neural Networks**. IGI Global, 2023, pp. 19–32.
- [3] A. Carrera-Rivera, W. Ochoa, F. Larrinaga e G. Lasa. “How-to conduct a systematic literature review: A quick guide for computer science research”. Em: **MethodsX** 9 (2022), p. 101895.
- [4] S. Keele et al. **Guidelines for performing systematic literature reviews in software engineering**. 2007.
- [5] M. Petticrew e H. Roberts. **Systematic reviews in the social sciences: A practical guide**. John Wiley & Sons, 2008.
- [6] L. Waikhom e R. Patgiri. “A survey of graph neural networks in various learning paradigms: methods, applications, and challenges”. Em: **Artificial Intelligence Review** 56.7 (2023), pp. 6295–6364.
- [7] Z. Wu, S. Pan, F. Chen, G. Long, C. Zhang e S. Y. Philip. “A comprehensive survey on graph neural networks”. Em: **IEEE transactions on neural networks and learning systems** 32.1 (2020), pp. 4–24.