

基于图神经网络预测药物 - 靶标相互作用的方法综述

左乐, 张琪

延安大学数学与计算机科学学院, 陕西 延安

收稿日期: 2024年8月23日; 录用日期: 2024年9月30日; 发布日期: 2024年10月12日

摘要

药物 - 靶标相互作用(DTI)在药物发现中至关重要, 利用人工智能模型对DTI进行预测不仅减少了时间和成本, 还能够提高预测的准确性和精度。然而, DTI数据异质性的特点以及预测模型的可解释性为研究人员带来挑战。在目前的研究中, 图神经网络(GNN)在分析和处理多源数据挖掘其中的复杂关系及可解释性方面具有显著的优势, 是DTI预测研究的强大工具。文章从三个方面对GNN在DTI预测中的方法进行综述。首先, 介绍了DTI预测的研究现状。其次, 分别探讨了GNN在药物(靶标)特征提取和DTI预测的最新研究进展, 整理了常用的公开数据集。最后, 总结了基于GNN预测DTI研究中存在的挑战并为未来相关研究提供展望。

关键词

图神经网络, 药物与靶标相互作用, 异构网络

A Review of Methods for Predicting Drug-Target Interactions Based on Graph Neural Networks

Le Zuo, Qi Zhang

College of Mathematics and Computer Science, Yan'an University, Yan'an Shaanxi

Received: Aug. 23rd, 2024; accepted: Sep. 30th, 2024; published: Oct. 12th, 2024

Abstract

Drug-target interaction (DTI) is crucial in drug discovery, and utilizing artificial intelligence models

to predict DTI not only reduces time and costs but also enhances prediction accuracy and precision. However, the heterogeneity of DTI data and the interpretability of prediction models pose challenges for researchers. In current research, graph neural networks (GNNs) exhibit significant advantages in analyzing and processing complex relationships among multiple data sources and enhancing interpretability, making them powerful tools in DTI prediction studies. This article provides a comprehensive review of GNN methods in DTI prediction from three aspects. Firstly, it introduces the current status of DTI prediction research. Secondly, it discusses the latest research progress of GNNs in drug (target) feature extraction and DTI prediction, and compiles commonly used public datasets. Finally, it summarizes the challenges in DTI prediction research based on GNNs and provides prospects for future related studies.

Keywords

Graph Neural Networks, Drug-Target Interactions, Heterogeneous Networks

Copyright © 2024 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

药物分子通过作为配体与靶标蛋白相互作用并激发或抑制靶标的生物过程来发挥其治疗作用。药物-靶标相互作用(Drug-Target Interaction, DTI)预测的研究在药物发现中是至关重要的,有助于促进对药物作用机制、药物耐药性[1]、副作用预测[2]等研究的理解。虽然传统的体内或体外实验可用于预测 DTI,但十分耗时且昂贵,花费取决于实验规模和复杂性[3]。因此,高效准确的计算方法在 DTI 预测方面受到了广泛的关注[4]。

现有的主流 DTI 预测方法都是基于机器学习的方法,可以进一步分为基于特征的方法[5]和基于图的方法[6]。在基于特征的方法中,利用不同的特征提取策略来提取药物和靶标的生物特征[7]。基于特征的方法将 DTI 预测转化为二分类问题并使用机器学习方法,如支持向量机(Support Vector Machine, SVM)和随机森林作为分类器[8]。例如:基于 SVM 构建了双分辨率局部模型,该模型利用化学和基因组数据计算的药物-药物相似性、靶标-靶标相似性[9]。基于图的方法描述了不同类型生物实体之间复杂的相互作用,Ding 等提出一种多视图图形正则化链接传播模型 MvGRLP,通过融合不同视图之间的互补信息来预测 DTI [10],在此方法中药物和靶标之间的相似性是基于局部或全局拓扑信息计算的[11]。Ezzat 等构建了药物-蛋白质网络,并基于正则化矩阵分解建立了预测模型[12]。然而,这些 DTI 预测方法是浅层学习方法,不能充分提取药物和蛋白质之间的深层复杂关联。

近年来,异构网络与深度学习算法相结合,通过整合来自不同数据源的药物和靶标相关信息,充分学习生物实体间的复杂交互,已被应用到 DTI 预测中。异构网络涵盖多种类型的生物实体以及不同类型实体之间的复杂关联。例如:Zheng 等通过结合长短期记忆(Long Short-Term Memory, LSTM)网络和卷积神经网络(Convolutional Neural Network, CNN)开发了一种用于 DTI 预测的方法[13]。与 CNN 相比,图神经网络(Graph Neural Networks, GNN) [14]可以有效地提取图特征。依赖于 GNN 在图数据挖掘中的突出性能,基于 GNN 预测 DTI 的模型能够有效地学习和表示异构网络中的关联关系,同时融合异构网络的结构和语义特征。目前使用 GNN 预测 DTI 的很多方法中利用 GNN 的不同变体同其他深度学习方法相结合以提高预测性能。故以下总结并分析了一些在 DTI 预测研究中的 GNN 方法以及常用数据集,旨在为后

续研究提供参考。

2. 基于图神经网络的 DTI 预测的方法

随着 GNN 在图分析中的广泛应用和突出性能, 研究人员自然地使用 GNN 对多类型实体之间复杂的关联关系进行建模。一些 GNN 的变体具有多尺度建模能力, 能够灵活地融入不同类型的数据信息, 包括药物的结构特征、蛋白质的序列特征以及药物和蛋白质相互作用网络。这种跨模态信息融合的能力可以提高模型对复杂相互作用的建模能力, 从而提高预测的准确性。如图卷积网络(Graph Convolutional Network, GCN) [15]、图注意力网络(Graph Attention Network, GAT) [16]等, 在 DTI 预测任务中广泛使用并展现出了优秀的性能。表 1 展示了近年来具有代表性的基于 GNN 的 DTI 预测方法总结:

Table 1. Summary of DTI prediction methods based on GNN

表 1. 基于 GNN 的 DTI 预测方法总结

方法	模型	特点	优势	缺陷
基于 GCN	GCN-DTI	应用 GCN 识别每一个 DPP 的特征	考虑 DPP 之间的关联	忽略了 DPP 网络的语义特征
	SGCL-DTI	对 DPP 网络的拓扑结构和语义特征进行对比	多视图的对比学习并结合监督信息	模型复杂
	EEG-DTI	提出一种新的异构图卷积网络	考虑多种生物实体之间的联系	未能考虑不同相邻节点的贡献
	AMGDTI	自动从异构网络中搜索自适应图	不依赖于领域知识	复杂度高, 计算成本大
	SHGCL-DTI	监督 DTI 预测任务与辅助图对比学习	充分利用语义信息与异构数据的交互	未利用元路径的上下文信息
基于 GAT	DTI-GAT	自动提取药物与靶标之间重要的高阶关系	通过添加相似性网络扩展异构图	容易产生过拟合问题
	MHGNN	设计为双通道架构分别学习药物和靶标表示	将高阶关系考虑在内	元路径的获取依赖领域知识
	DTIHNC	利用 GAT 集成药物和靶标的多模态相似性	充分探索药物和靶标的相似性	泛化性和可解释性较差
	IMCHGAN	采用两级注意力机制获取特征表示	GAT 采用动态权重聚合特征	学习过程中未考虑药物和靶标的相互作用
基于组合	DTI-MGNN	将 DPP 视为节点构成特征图	探索 DPP 表示中的拓扑结构和语义信息	未利用高阶关系
	DTI-HETA	结合 GCN 和 GAT 充分利用源数据信息	突出相邻节点贡献	异构网络未提供丰富的语义信息

缩写: GNN——图神经网络; DTI——药物-靶相互作用; GAT——图注意力网络; GCN——图卷积网络; DNN——深度神经网络; DPP——药物-靶标对。

基于 GNN 预测 DTI 的方法一般包含三个步骤, 首先构建生物异构网络, 然后使用基于 GNN 的方法学习药物和靶标的节点嵌入, 最后进行 DTI 预测。图 1 描绘了基于 GNN 预测 DTI 的整体流程。

2.1. 生物异构网络的构建

生物实体(如: 药物、靶标、疾病和副作用)之间的相互关系包含丰富的语义信息, 这些复杂的相互

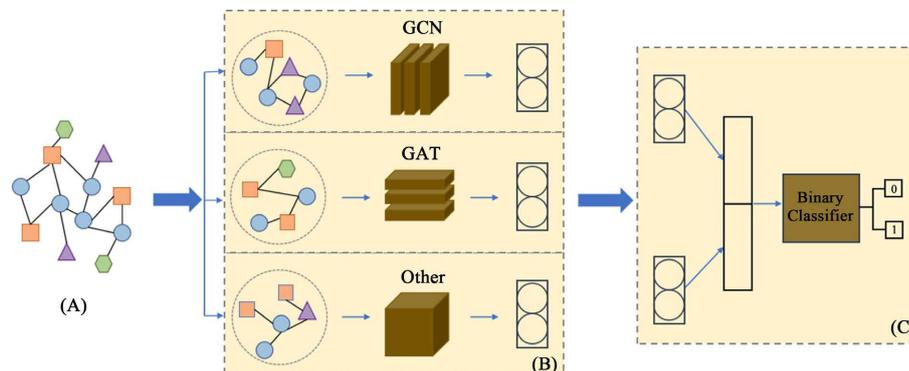


Figure 1. Flowchart of the GNN-based DTI prediction method. (A) Construction of biological heterogeneous network; (B) Learning of drug (target) features; (C) DTI prediction
图 1. 基于 GNN 预测 DTI 方法流程图。(A) 生物异构网络的构建; (B) 药物(靶标)特征学习; (C) DTI 预测

关系可以表示为不同的药物和靶标相关网络(图 2(C)), 生物异构网络(图 2(D))由节点和边构建而成, 其中节点是生物实体(图 2(A)), 即: 药物(D)、靶标(T)、疾病(I)和副作用(S), 边是生物实体间的关联关系或相互作用, 例如: 药物-靶标相互作用。DTI 预测被看作是一个二分类问题(图 2(D)中的虚线), 即: 药物节点和靶标节点之间是否存在相互作用。

元路径描述两个节点之间的复合关系, 在异构网络中广泛使用。通过定义不同类型的元路径, 可以探索异构网络中不同节点之间的关系, 并发现潜在的语义信息[17]。如图 2(B), 其中 DTD 表示两种药物具有相同的靶标, DID 表示一种疾病可以用两种药物进行治疗。从公共数据库中收集异构生物数据, 对不同类型的生物实体及其之间的复杂关系进行建模, 利用异构网络来描述药物、蛋白质及其之间的异构关系(如图 2(D))。

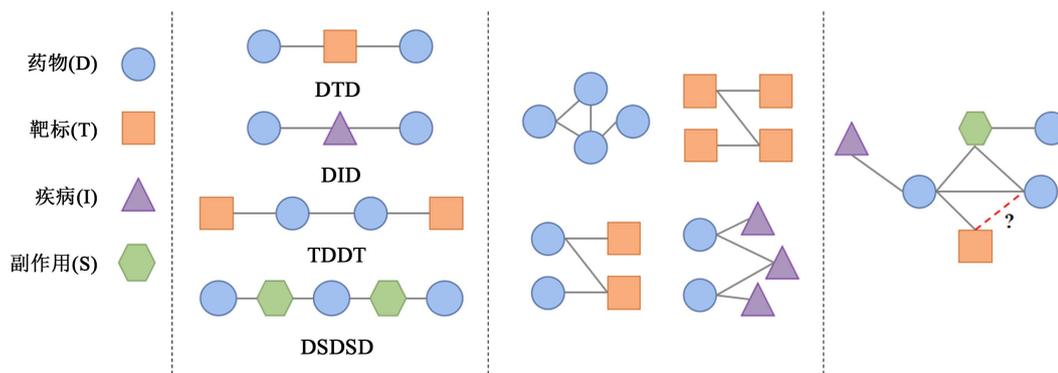


Figure 2. (A) Four node types represent four types of biological entities; (B) Examples of four different meta-paths; (C) Networks related to drugs and targets; (D) Biological heterogeneous networks (the red dashed line indicates the predicted interaction between nodes D and T)

图 2. (A) 四种节点类型代表四种类型的生物实体; (B) 举例四个不同的元路径; (C) 与药物和靶标相关的网络; (D) 生物异构网络(红色虚线表示节点 D 和 T 之间相互作用预测)

2.2. 药物(靶标)特征学习

将构建好的生物异构网络输入到 GNN 中, GNN 通过迭代更新每个节点的特征, 利用节点的邻居信息来更新节点的特征表示。这样, 节点的特征能够捕捉到节点自身的属性以及其在网络中的相互作用关系。经过 GNN 的多轮迭代后, 得到药物和靶标的特征向量, 这个向量包含了节点在整个生物网络中的信息。

2.2.1. 基于图卷积网络的方法

GCN 是 GNN 中最具代表性的模型, 能够有效地利用图结构信息, 从而学习到每个节点的表示, 捕捉异构图的局部结构和全局拓扑信息。

先前的研究通常是基于每种类型的相互作用分别提取特征, 然后基于每种类型的交互特征被连接在一起, 但没有对单个异构网络中的异构信息进行建模。因此, 这些方法不能很好地考虑多种类型的生物实体之间的关联。为解决该限制, Zhao 等[18]提出了名为“GCN-DTI”的学习框架, 将药物-蛋白质对(DPP)之间的关联考虑在内, 从一个非常大的 DPP 网络中区分真假 DPP。为了将 DPP 之间的关联纳入 DTI 建模, 该模型使用 GCN 来学习每个 DPP 的特征。类似的, Peng 等[19]提出一个基于异构图卷积网络的端到端学习框架 EEG-DTI 用于 DTI 预测, EEG-DTI 建立了一个基于 GCN 的模型以学习药物和靶标的低维特征表示。

Yang 等[20]首次尝试以监督方式对异构网络的拓扑结构和语义特征进行对比, 提出了一个端到端的监督图对比学习模型 SGCL-DTI。该模型通过 GCN 学习基于不同元路径的药物和靶标的表示, 然后构造并使用两个多层 GCN 学习药物-蛋白质对的拓扑网络和语义网络。Yao 等[21]提出了一种半监督异构图对比学习的 DTI 预测框架 SHGCL-DTI, 添加了与辅助图对比学习任务。分别通过异构卷积网络(HGCN)和具有语义注意机制的 GCN 从邻居视图和元路径视图中提取节点特征。在两个视图之间执行图形对比学习任务, 充分利用异构图中丰富的语义信息和异构数据之间复杂的交互。

在元路径的构建中, 现有的元路径通常是经验设计的, 严重依赖领域知识[22]。虽然异构网络包括多种信息, 但并非所有的信息都对 DTI 预测有用, 构建不恰当的元路径可能影响到最后的预测结果。Su 等[23]提出了一种自动从异质网络中搜索自适应元图, 无需领域知识的 DTI 预测方法 AMGDIT。AMGDIT 使用 GCN 分别在异构网络中寻找药物和蛋白质靶点信息聚合的自适应元图。该方法考虑了在无需领域知识的情况下自动搜索有效的信息, 集成异构网络中复杂的多种语义关系和结构信息, 这是获得好的预测结果的关键。

2.2.2. 基于图注意力网络的方法

GAT 通过引入注意力机制, 使模型能够灵活地学习节点之间的关系权重, 具有更强的适应性和表征能力。GAT 在建模复杂图数据方面表现出巨大的潜力, 这在不同的研究领域中引起广泛的关注, 如: 图分类[24]和推荐系统[25]。图注意力层应用自注意力机制, 允许为邻居节点分配不同的注意力权重, 从而实现信息的聚合和传播。

异构网络中的不同邻居节点对 DTI 的重要性不同, 而 GCN 在卷积中平等对待所有邻居节点, 不能根据节点的重要性分配不同的权重[26]。Wang 等[27]提出一个端对端的框架 DTI-GAT 用于 DTI 预测。在构建的图上应用 GAT 来为每个蛋白质和药物生成嵌入向量。证明了注意力机制可以通过为每条边分配不同的权重, 从而自动提取蛋白质和药物之间的重要高阶关系。Li 等[28]提出了一种元路径聚合异构图神经网络 MHGNN, 该模型设置为双通道架构, 在每个通道中, 都使用 GAT 从元路径中学习具有丰富语义信息的药物(靶标)表示。

以往有些研究忽略了异构网络中的元路径信息或不能提供良好的 DTI 异构网络分析解释, 为了克服该局限性, Li 等[29]提出一种具有异构图注意力网络方法 IMCHGAN。首先, IMCHGAN 采用两级表示学习方法, 分别从异构网络中学习药物和目标潜在特征表示。在底层, 利用 GAT 来学习特定元路径的药物(或靶标)潜在特征表示; 在最上层, 进一步采用基于注意力的学习方法将不同的元路径潜在表示集成到最终的潜在特征中。IMCHGAN 充分考虑了 DTI 网络的异质性, 采用两级注意机制, 获得更全面的药物和靶点特征表示。

Jiang 等[30]提出一个从多实体异构网络中集成多模态相似性和邻居拓扑信息的 DTI 预测方法

DTIHNC。DTIHNC 从异构信息网络中学习邻居拓扑表示。通过 GAT 集成一阶邻居拓扑信息和异构网络中节点的低维特征。

2.2.3. 基于组合的方法

将 GAT 和 GCN 结合起来预测 DTI 可以充分利用各自的优势, 提高模型对图数据的表征能力和预测性能。既可以有效地聚合邻居节点的信息, 又可以探索生物异质图中的高阶依赖关系, 将 DTP 之间的高阶关联考虑在内。综合利用 GCN 对图结构信息的敏感性和参数共享特性以及 GAT 中注意力机制的灵活性, 从而提高模型对药物和蛋白质特征表示的学习能力, 在 DTI 预测任务中获得更高的预测性能。

为了解决在数据集成过程中容易造成部分信息丢失, 在聚合中心节点信息时也未考虑不同相邻节点的贡献, 导致预测性能不佳的问题。Shao 等[31]提出一种基于异构图具有注意力机制的端对端模型 DTI-HETA。使用 GCN 来获得药物和靶标的嵌入表示。同时, 采用 GAT 来突出不同邻域节点在聚合图卷积信息时对中心节点的贡献。

除此之外, 以往的研究大多集中在学习 DPP 网络的拓扑结构上, 而忽略了 DPP 表征学习过程中拓扑结构与语义特征的融合。Li 等[32]提出了一种多通道图神经网络 DTI-MGNN, 为了在学习在药物-蛋白质对特征向量中更好地融合拓扑结构和语义特征, 一方面, 使用两个独立的 GAT 来学习拓扑图和特征图的节点之间的不同相互作用。另一方面, 使用具有共享权重矩阵的 GCN 来学习两个图的公共信息。

2.3. DTI 预测

根据学习到的药物和靶标的特征向量, 选择合适的方法进行药物-靶标相互作用预测。构建一个预测模型, 可以采用机器学习方法或深度学习模型等, 将融合后的特征向量作为输入, 预测药物-靶标相互作用的结果。

基于深度学习的方法在最后一步 DTI 预测能够很好地融合特征向量, 从而提高预测的准确性和鲁棒性。例如: 多层感知器(Multilayer Perceptron, MLP)、深度神经网络(Deep Neural Network, DNN) [33]、卷积神经网络(Convolutional Neural Network, CNN) [13]和 GCN 等。具体而言, 模型 DTI-MGNN 在最后训练了一个 MLP 模型来预测药物和蛋白质之间是否存在关系; 模型 GCN-DTI 使用 DNN 模型被用作监督学习模型来确认 DPP 的真实性; 模型 DTIHNC 设计了一种 CNN, 将药物和蛋白质的相似性信息与从异构网络中获得的嵌入特征深度融合; 模型 MHGNN 对于每个 DTP, 将药物表示和靶标表示拼接起来获得最终的表示, 输入到两个 GCN 层中预测结果。另外, 模型 SGCL-DTI 提出一种监督对比(contrastive learning)学习任务, 通过将属于同一类别的样本聚集在一起来最小化类内方差, 并将来自不同类别的样本推开来最大化类间方差。在模型 SGCL-DTI 中, 监督对比学习(SCL)策略类似于自监督图对比学习中使用的对比目标。不同之处在于, 对比目标被用来监督最终预测任务的学习, 而不是对比不同增强视图的样本。

除此之外, 模型 DTI-GAT 中所使用的逐元素乘法(Element-wise multiplication)和模型 SHGCL-DTI 中使用的距离乘积(Distance Multiplication, DistMult) [34]都是知识图谱中的关系预测模型。具体来说, 模型 DTI-GAT 采用由 GAT 层生成的药物-靶标嵌入对作为输入, 然后两个节点向量进行逐元素乘法, 最后通过一层神经网络和 sigmoid 激活函数产生一个概率分数。模型 SHGCL-DTI 应用 DistMult 方法来重建异构网络中八种类型的边, 其中重建的 DTI 网络是 DTI 预测的结果。归纳矩阵补全(Inductive Matrix Completion, IMC)是基于矩阵分解的模型。通过将实体和关系映射到低维连续向量空间中, 来表示实体之间的语义关联。IMCHGAN 学习到样本药物和靶标的嵌入后使用归纳矩阵补全(IMC) [35]的方法预测药物靶标关联得分。

3. 公开数据集

药物与靶标相互作用数据集对于模型的评估及其优化有着关键影响, 不同的 DTI 预测方法所需要的

数据集类型也不尽相同。本节将对 DTI 预测研究中常用到的三个异构数据集做出简要归纳, 数据集的基本信息如表 2 所示。

Luo 数据集[36]: 有四种类型的节点(药物、蛋白质、疾病和副作用)和八种类型的相互作用, 其中包含六种异质性相互作用类型(药物-药物相互作用、药物-蛋白质相互作用、药物-疾病相互作用、药物-副作用相互作用、蛋白质-蛋白质相互作用和蛋白质-疾病相互作用)和两个相似性网络(药物化学结构相似性网络和蛋白质序列相似性网络), 共有 12,015 个节点和 894,854 条边。

Zheng 数据集[37]: 数据集中 1094 种药物和 1556 种靶蛋白之间有 11819 种实验验证的相互作用。除药物和靶标外, 它还包含四种异构信息: 药物取代基、化学结构、药物副作用和基因本体。

Yamanishi 数据集[38]包含四个子数据集: 酶(E)、离子通道(IC)、G 蛋白偶联受体(GPCR)和核受体(NR)。每个子数据集包含三个网络: 药物-药物结构相似性网络、靶点-靶点相似性网络和 DTI 网络。

Table 2. Statistics of the dataset

表 2. 数据集的统计

(a) Luo 数据集				
节点类型	数量	边类型	数量	
药物	708	药物-药物相互作用	10,036	
蛋白质	1512	药物-靶标相互作用	1923	
疾病	5603	靶标-靶标相互作用	7363	
副作用	4192	药物-疾病关联关系	199,214	
		药物-副作用关联关系	80,164	
		靶标-疾病关联关系	159,6745	
		药物-药物相似性关系	501,264	
		靶标-靶标相似性关系	2,286,144	
(b) Zheng 数据集				
节点类型	数量	边类型	数量	
药物	1094	药物-药物	1,196,836	
蛋白质	1556	药物-蛋白质	11,819	
化学结构	881	药物-化学亚结构	133,880	
副作用	4063	药物-副作用	122,792	
取代基	738	药物-取代基	20,798	
基因本体论术语 (Gene Ontology term, GO term)	4098	蛋白质-GO term	35,980	
		蛋白质-蛋白质	2,421,136	
(c) Yamanishi 数据集				
	节点类型	数量	边类型	数量
酶	药物	445	药物-靶标相互作用	2925
	靶标	663		

续表

离子通道	药物	210	药物 - 靶标相互作用	1476
	靶标	204		
G 蛋白偶联受体	药物	223	药物 - 靶标相互作用	635
	靶标	95		
核受体	药物	54	药物 - 靶标相互作用	90
	靶标	26		

4. 讨论与总结

4.1. 挑战

目前, GNN 已成为 DTI 预测的有力工具, 尽管 GNN 在 DTI 预测中展现出了良好的性能, 但这项研究面临着一些挑战。首先是异构数据集的构建, 异构数据集包含了多种类型的生物信息, 如药物结构、蛋白质序列等。这些数据提供了不同层面和不同尺度的语义信息。在建立药物 - 靶标相互作用预测模型的异构数据集中, 需要考虑数据噪声、缺失数据和未知结构等不确定因素。这些不确定性可能会降低预测结果的准确性和模型的可解释性。另外, 目前大量方法使用的负样本是由未知的 DTI 构成, 已知的 DTI 数量较少, 未知 DTI 数量则相对较多, 导致数据集出现正负样本不平衡的情况。因此, 构建合适的异构数据集是一个未来研究中存在的挑战。

另一挑战是模型的设计, 尽管现有的 GNN 模型在 DTI 预测任务中表现出一定的效果, 但模型还是存在未充分利用异构网络中包含的大量丰富信息, 无法完全捕获复杂网络中的拓扑结构和语义特征的限制。在 DTI 预测过程中, 需要综合考虑数据的异构性, 特征选择与提取, 数据融合策略、信息传递与交互以及模型训练与优化多方面的问题, 以设计出有效的模型进行准确的 DTI 预测。

4.2. 未来展望

为了解决这些挑战, 首先, 通过整合更多的异构属性信息来扩展的异构数据集, 越来越多的药物和靶标相关的异构网络可以被纳入和进一步探索, 如代谢网络等。这些不同的异构网络将提供丰富的语义信息, 以增强模型学习药物和靶标之间关系的能力。

其次, 需要设计更加复杂和精细的 GNN 模型。这些模型应该能够灵活地利用不同类型的信息, 同时能够充分捕获复杂网络中的拓扑结构和语义特征。可能的解决方案包括引入多层次的注意力机制、设计更复杂的图卷积操作等。通过这些措施, 使其能够更好地适应不同类型的关系和网络结构。

未来, 研究人员通过加深对模型可解释性理解, 设计出更好、更具有可解释性的模型来预测 DTI。一方面可以增强对相关生物过程和复杂生物相互作用的理解, 另一方面可以通过预测现有药物的新靶标来加速药物进入临床试验的进程, 有利于研究人员通过缩小搜索空间来实验验证预测的 DTI。

4.3. 总结

随着近年来 GNN 的研究日益深入, GNN 已被证明在捕获和利用药物和蛋白质之间的关系上是有效的, GNN 可以对药物和蛋白质之间的结构关系、功能区域和多模态特征进行建模, 生成可解释的预测和可视化。这些过程为药物设计和开发提供更深入的研究和指导。在结构特征建模方面, GNN 可以将各生物实体表示为异构网络中不同类型的节点和边。通过学习节点之间的相互作用和信息传递, GNN 以此提取药物和蛋白质的结构特征, 将学习到的特征向量输入到指定的解码器中进行 DTI 预测, 可以帮助理解药物和靶标蛋白相互作用的机制, 为 DTI 应用中的生物可解释性提供了强有力的支持。

使用 GNN 的方法进行 DTI 预测的研究得到了广泛的关注, DTI 预测模型也得到了不断更新和性能的显著提升, 并逐步走向计算生物学与生物信息学领域的实际应用场景中, 成为新药物研发和药物重定位的关键技术。

基金项目

延安大学校级项目(YCX2024050)。

参考文献

- [1] Xue, H., Li, J., Xie, H. and Wang, Y. (2018) Review of Drug Repositioning Approaches and Resources. *International Journal of Biological Sciences*, **14**, 1232-1244. <https://doi.org/10.7150/ijbs.24612>
- [2] Mongia, A. and Majumdar, A. (2020) Drug-Target Interaction Prediction Using Multi Graph Regularized Nuclear Norm Minimization. *PLOS ONE*, **15**, e0226484. <https://doi.org/10.1371/journal.pone.0226484>
- [3] Yao, L., Evans, J.A. and Rzhetsky, A. (2010) Novel Opportunities for Computational Biology and Sociology in Drug Discovery. *Trends in Biotechnology*, **28**, 161-170. <https://doi.org/10.1016/j.tibtech.2010.01.004>
- [4] 张然, 王学志, 汪嘉蓓, 等. 药物-靶点相互作用预测的计算方法综述[J]. 计算机工程与应用, 2023, 59(12): 1-13.
- [5] Mei, J., Kwoh, C., Yang, P., Li, X. and Zheng, J. (2012) Drug-Target Interaction Prediction by Learning from Local Information and Neighbors. *Bioinformatics*, **29**, 238-245. <https://doi.org/10.1093/bioinformatics/bts670>
- [6] Lee, I. and Nam, H. (2018) Identification of Drug-Target Interaction by a Random Walk with Restart Method on an Interactome Network. *BMC Bioinformatics*, **19**, Article No. 208. <https://doi.org/10.1186/s12859-018-2199-x>
- [7] Shamima, M.K., Mehedi, M.H. and Hiroyuki, K. (2019) PreAIP: Computational Prediction of Anti-Inflammatory Peptides by Integrating Multiple Complementary Features. *Frontiers in Genetics*, **10**, Article 129.
- [8] Fu, G., Ding, Y., Seal, A., Chen, B., Sun, Y. and Bolton, E. (2016) Predicting Drug Target Interactions Using Meta-Path-Based Semantic Network Analysis. *BMC Bioinformatics*, **17**, Article No. 160. <https://doi.org/10.1186/s12859-016-1005-x>
- [9] Bleakley, K. and Yamanishi, Y. (2009) Supervised Prediction of Drug-Target Interactions Using Bipartite Local Models. *Bioinformatics*, **25**, 2397-2403. <https://doi.org/10.1093/bioinformatics/btp433>
- [10] Ding, Y.J., Tang, J.J. and Guo, F. (2021) Identification of Drug-Target Interactions via Multi-View Graph Regularized Link Propagation Model. *Neurocomputing*, **461**, 618-631.
- [11] Zhang, Z., Zhang, X., Wu, M., Ou-Yang, L., Zhao, X. and Li, X. (2020) A Graph Regularized Generalized Matrix Factorization Model for Predicting Links in Biomedical Bipartite Networks. *Bioinformatics*, **36**, 3474-3481. <https://doi.org/10.1093/bioinformatics/btaa157>
- [12] Ezzat, A., Zhao, P., Wu, M., Li, X. and Kwoh, C. (2017) Drug-Target Interaction Prediction with Graph Regularized Matrix Factorization. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, **14**, 646-656. <https://doi.org/10.1109/tcbb.2016.2530062>
- [13] Zheng, X., He, S., Song, X., Zhang, Z. and Bo, X. (2018) DTI-RCNN: New Efficient Hybrid Neural Network Model to Predict Drug-Target Interactions. In: Kůrková, V., Manolopoulos, Y., Hammer, B., Iliadis, L. and Maglogiannis, I., Eds, *Artificial Neural Networks and Machine Learning—ICANN 2018*, Springer, 104-114. https://doi.org/10.1007/978-3-030-01418-6_11
- [14] Scarselli, F., Gori, M., Chung Tsoi, A., Hagenbuchner, M. and Monfardini, G. (2009) The Graph Neural Network Model. *IEEE Transactions on Neural Networks*, **20**, 61-80. <https://doi.org/10.1109/tnn.2008.2005605>
- [15] Kipf, N.T. and Welling, M. (2016) Semi-Supervised Classification with Graph Convolutional Networks. arXiv: 1609.02907.
- [16] Velickovic, P., Cucurull, G., Casanova, A., et al. (2017) Graph Attention Networks. arXiv: 1710.10903.
- [17] Dong, Y., Chawla, N.V. and Swami, A. (2017) Metapath2vec: Scalable Representation Learning for Heterogeneous Networks. *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, Halifax, 13-17 August 2017, 135-144. <https://doi.org/10.1145/3097983.3098036>
- [18] Zhao, T., Hu, Y., Valsdottir, L.R., Zang, T. and Peng, J. (2020) Identifying Drug-Target Interactions Based on Graph Convolutional Network and Deep Neural Network. *Briefings in Bioinformatics*, **22**, 2141-2150. <https://doi.org/10.1093/bib/bbaa044>
- [19] Peng, J., Wang, Y., Guan, J., Li, J., Han, R., Hao, J., et al. (2021) An End-To-End Heterogeneous Graph Representation Learning-Based Framework for Drug-Target Interaction Prediction. *Briefings in Bioinformatics*, **22**, bbaa430. <https://doi.org/10.1093/bib/bbaa430>

- [20] Li, Y., Qiao, G., Gao, X. and Wang, G. (2022) Supervised Graph Co-Contrastive Learning for Drug-Target Interaction Prediction. *Bioinformatics*, **38**, 2847-2854. <https://doi.org/10.1093/bioinformatics/btac164>
- [21] Yao, K., Wang, X., Li, W., Zhu, H., Jiang, Y., Li, Y., *et al.* (2023) Semi-Supervised Heterogeneous Graph Contrastive Learning for Drug-Target Interaction Prediction. *Computers in Biology and Medicine*, **163**, Article ID: 107199. <https://doi.org/10.1016/j.compbiomed.2023.107199>
- [22] Yang, C., Liu, M., He, F., Zhang, X., Peng, J. and Han, J. (2019) Similarity Modeling on Heterogeneous Networks via Automatic Path Discovery. In: Berlingerio, M., Bonchi, F., Gärtner, T., Hurley, N. and Ifrim, G., Eds., *Machine Learning and Knowledge Discovery in Databases*, Springer, 37-54. https://doi.org/10.1007/978-3-030-10928-8_3
- [23] Su, Y., Hu, Z., Wang, F., Bin, Y., Zheng, C., Li, H., *et al.* (2023) AMGDTI: Drug-Target Interaction Prediction Based on Adaptive Meta-Graph Learning in Heterogeneous Network. *Briefings in Bioinformatics*, **25**, bbad474. <https://doi.org/10.1093/bib/bbad474>
- [24] Gao, J., Gao, J., Ying, X., Lu, M. and Wang, J. (2021) Higher-Order Interaction Goes Neural: A Substructure Assembling Graph Attention Network for Graph Classification. *IEEE Transactions on Knowledge and Data Engineering*, **35**, 1594-1608. <https://doi.org/10.1109/tkde.2021.3105544>
- [25] Wu, Q., Zhang, H., Gao, X., He, P., Weng, P., Gao, H., *et al.* (2019) Dual Graph Attention Networks for Deep Latent Representation of Multifaceted Social Effects in Recommender Systems. *The World Wide Web Conference*, San Francisco, 13-17 May 2019, 2091-2102. <https://doi.org/10.1145/3308558.3313442>
- [26] Gao, H., Wang, Z. and Ji, S. (2018) Large-scale Learnable Graph Convolutional Networks. *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, London, 19-23 August 2018, 1416-1424. <https://doi.org/10.1145/3219819.3219947>
- [27] Wang, H., Zhou, G., Liu, S., *et al.* (2021) Drug-Target Interaction Prediction with Graph Attention Networks. arXiv: 2107.06099.
- [28] Li, M., Cai, X., Xu, S. and Ji, H. (2023) Metapath-Aggregated Heterogeneous Graph Neural Network for Drug-Target Interaction Prediction. *Briefings in Bioinformatics*, **24**, bbac578. <https://doi.org/10.1093/bib/bbac578>
- [29] Li, J., Wang, J., Lv, H., Zhang, Z. and Wang, Z. (2022) IMCHGAN: Inductive Matrix Completion with Heterogeneous Graph Attention Networks for Drug-Target Interactions Prediction. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, **19**, 655-665. <https://doi.org/10.1109/tcbb.2021.3088614>
- [30] Jiang, L., Sun, J., Wang, Y., Ning, Q., Luo, N. and Yin, M. (2022) Identifying Drug-Target Interactions via Heterogeneous Graph Attention Networks Combined with Cross-Modal Similarities. *Briefings in Bioinformatics*, **23**, bbac016. <https://doi.org/10.1093/bib/bbac016>
- [31] Shao, K., Zhang, Y., Wen, Y., Zhang, Z., He, S. and Bo, X. (2022) DTI-HETA: Prediction of Drug-Target Interactions Based on GCN and GAT on Heterogeneous Graph. *Briefings in Bioinformatics*, **23**, bbac109. <https://doi.org/10.1093/bib/bbac109>
- [32] Li, Y., Qiao, G., Wang, K. and Wang, G. (2021) Drug-Target Interaction Prediction via Multi-Channel Graph Neural Networks. *Briefings in Bioinformatics*, **23**, bbab346. <https://doi.org/10.1093/bib/bbab346>
- [33] Hinton, G., Deng, L., Yu, D., Dahl, G., Mohamed, A., Jaitly, N., *et al.* (2012) Deep Neural Networks for Acoustic Modeling in Speech Recognition: The Shared Views of Four Research Groups. *IEEE Signal Processing Magazine*, **29**, 82-97. <https://doi.org/10.1109/msp.2012.2205597>
- [34] Yang, B., Yih, W., He, X., *et al.* (2014) Embedding Entities and Relations for Learning and Inference in Knowledge Bases. arXiv: 1412.6575.
- [35] Zhang, M. and Chen, Y. (2019) Inductive Matrix Completion Based on Graph Neural Networks. arXiv: 1904.12058.
- [36] Luo, Y., Zhao, X., Zhou, J., Yang, J., Zhang, Y., Kuang, W., *et al.* (2017) A Network Integration Approach for Drug-Target Interaction Prediction and Computational Drug Repositioning from Heterogeneous Information. *Nature Communications*, **8**, Article No. 573. <https://doi.org/10.1038/s41467-017-00680-8>
- [37] Zheng, Y., Peng, H., Zhang, X., Gao, X. and Li, J. (2018) Predicting Drug Targets from Heterogeneous Spaces Using Anchor Graph Hashing and Ensemble Learning. 2018 *International Joint Conference on Neural Networks (IJCNN)*, Rio de Janeiro, 8-13 July 2018, 1-7. <https://doi.org/10.1109/ijcnn.2018.8489028>
- [38] Yamanishi, Y., Araki, M., Gutteridge, A., Honda, W. and Kanehisa, M. (2008) Prediction of Drug-Target Interaction Networks from the Integration of Chemical and Genomic Spaces. *Bioinformatics*, **24**, i232-i240. <https://doi.org/10.1093/bioinformatics/btn162>