

基于图卷积和图嵌入的药物-靶标反应识别方法

目录

1 引言	1
1.1 研究背景及其意义.....	1
1.2 国内外研究现状.....	2
1.3 主要工作	4
1.4 论文组织结构	4
2 相关理论概述	5
2.1 构建相互作用网络图.....	5
2.2 相似度度量	5
2.3 图卷积神经网络.....	6
2.4 激活函数	7
2.5 图注意力机制	7
2.6 关系发现	8
3 DPPM 模型介绍	9
3.1 模型概念与结构.....	9
4 实验结果分析	11
4.1 实验设置	11
4.2 实验指标	11
4.2.1 AUROC.....	11
4.2.2 AUPR	12
4.3 实验结果	13
4.4 消融实验	14
4.5 关系发现	15
5 总结与展望	16
5.1 研究结论	16
参考文献	18

摘要: 本文主要聚焦于药物-靶标反应的预测。传统方法受限于靶蛋白 3D 结构未知或已知结合配体数量有限。基于机器学习的方法考虑药物的化学结构和靶标序列，将 DTI 视为二元分类问题。DTI 所面临的挑战有大型数据集中无法学习到良好的节点表征，同时多路径融合方面面临许多问题，这些问题限制了 DTI 预测的性能。本文提出了基于图卷积和图嵌入的药物-靶标反应预测模型 DPPM，通过学习异构网络中节点的低维表示，并引入图注意力机制进行特征融合，实现了在统一空间中对药物-靶标反应的预测。实验结果表明，DPPM 在预测准确率方面优于传统方法，跟之前的方法相比，在 AUROC 上提升了 13%，在 AUPR 上提升了 15%。这一研究为药物发现和再定位提供了新的工具，有助于理解药物的作用机制。

关键词: 图卷积神经网络；图注意力机制；DTI；异构图

1 引言

1.1 研究背景及其意义

药物-靶标相互作用 (DTI) 的计算预测通过识别新药或现有药物的潜在新药靶点，在药物发现和重新定位中起着至关重要的作用。DTI 鉴定最初是用生物的方法进行的，但是此类方法不仅耗时，而且十分高昂。但是，有关的计算机科学技术能够预测出潜在的 DTI，从而简化了药物发现和重新定位过程。传统的计算方法主要利用基于分子对接和基于配体的策略。但分子对接的功效很容易受到限制，是因为靶蛋白的 3D 结构未知，而当靶蛋白已知配体有限时，基于配体的方法也无法进行准确的预测。

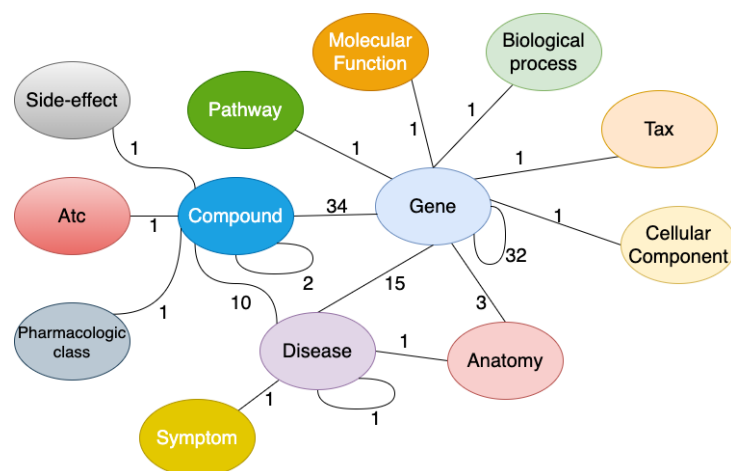


图 1 医疗实体关系图

近年来，人们非常关注旨在预测药物-靶点相互作用 (DTI) 的基于机器学习的计算技术的发展。这些方法基于“guilt-by-association”^[1]的概念，该概念认为具有相似特征的药物可能具有共同的靶标，反之亦然。例如，Bleakley 和 Yamanishi^[2]采用支持向量机框架来使用两部分局部模型 (BLM) 来预测

DTI。在此基础上，梅等人^[3]通过将 BLM 与邻近相互作用概况推断 (NII) 技术（称为 BLMNII）集成来增强模型，该技术能够从邻近实体中提取 DTI 特征并预测涉及新药或候选靶点的相互作用。Liat Perlman 等^[4]使用药物与药物和基因与基因的相似度度量构建 logistics 回归预测模型。A Kastrin 等人^[5]采用药物边际效用作为 logistics 回归模型的特征预测药物与蛋白质之间的反应。然而这些机器学习方法受限于手工提取特征使得模型的性能依赖于特征提取的质量并且在含有大量实体的数据集中面临挑战。得益于深度学习的发展，大规模数据集的 DTI 预测工作开始涌现。应用于知识图谱的图嵌入技术通过学习用于基于 DL 的方法中使用的药物或靶点的低维特征表示，提高了 DTI 预测性能^[6]。RS Olayan 等人^[7] 基于网络的方法将 DTI 的预测表述为异构图中的链路预测问题，这些方法利用了异构数据源提供的多样化的信息和多视角。G Fu 等人^[8]还提出了一种基于元路径的方法，从异构网络中提取 DTI 的语义特征。W Wang 等人^[9]提出了 HNM 以随机游走的方法学习异构图的节点嵌入和平均池化的方法融合元路径的特征。

上述的方法在大型数据集中无法学习到良好的节点表征。同时，在多路径融合方面面临挑战。而本文提出了一种新方法 DPPM (Drug Protein Prediction Model)，该方法提出了

- 融合多路径特征的框架，该框架综合了多种数据源，能提高预测准确性，并增强模型鲁棒性，减少过拟合的风险，而且可以发现新关联，适应复杂的生物系统。
- 图卷积嵌入学习，该方法在药物-靶标相互作用预测中的应用，其核心思想是将药物和蛋白质表示为图形结构，使用图卷积神经网络提取节点特征，嵌入到一个低维、连续、稠密的向量当中，然后通过结合这些嵌入向量来预测它们之间的相互作用。
- 图注意力特征融合，它允许模型在不同的节点之间分配注意力，以捕捉它们之间的重要关系。它能够帮助模型更好地学习节点特征，捕捉复杂关系，提高模型的可解释性和泛化能力。

1.2 国内外研究现状

以上内容仅为本文档的试下载部分，为可阅读页数的一半内容。

如要下载或阅读全文，请访问：

<https://d.book118.com/145020030013012034>