

# 一种基于电子健康记录的多尺度图表示学习模型

樊捷杰<sup>1</sup>, 班晓娟<sup>1</sup>, 张志研<sup>2</sup>

(1. 北京科技大学 人工智能学院, 北京 100083; 2. 北京科技大学 钢铁技术协同创新中心, 北京 100083)

**摘要:** 现有的电子健康记录(electronic health records, EHR)的图表示学习方法多依赖单个患者的局部信息, 忽视了群体患者在疾病演化和诊疗路径上的潜在关联, 从而限制了模型的泛化性与鲁棒性. 针对这一问题, 本文提出一种混合多层级图神经网络(hybrid multi-level graph neural network, H-MGNN)模型, 并将其应用于重症监护室(intensive care unit, ICU)患者的死亡预测. 该模型通过构建宏观层面的患者关系图(patient-patient graph, P-P)、微观层面的分类-笔记-词汇超图(taxonomy-note-word hypergraph, T-N-W), 结合超图的时序依赖关系, 实现多尺度上的患者特征融合. 同时, 本文设计了融合算法(hybrid embedding, Hybrid-E), 用于提取和整合患者嵌入的潜在特征, 以提升预测准确性. 实验结果表明, H-MGNN在MIMIC-III(medical information mart for intensive care III)数据集上的住院死亡率预测等任务中显著优于现有方法, 验证了其在复杂EHR数据挖掘中的有效性和先进性.

**关键词:** 电子健康记录; 多尺度; 超图; 图神经网络

中图分类号: TP 391.4 文献标志码: A 文章编号: 1005-3026(2026)01-0031-11

## A Multi-scale Graph Representation Learning Model Based on Electronic Health Records

FAN Jie-jie<sup>1</sup>, BAN Xiao-juan<sup>1</sup>, ZHANG Zhi-yan<sup>2</sup>

(1. School of Intelligence Science and Technology, University of Science and Technology Beijing, Beijing 100083, China; 2. Collaborative Innovation Center of Steel Technology, University of Science and Technology Beijing, Beijing 100083, China. Corresponding author: BAN Xiao-juan, E-mail: banxj@ustb.edu.cn)

**Abstract:** Existing graph representation learning methods for electronic health records (EHR) primarily rely on local information of a single patient, overlooking potential associations among patients in disease progression and treatment pathways. This limits the models' generalizability and robustness. To address this issue, a hybrid multi-level graph neural network (H-MGNN) model was proposed, and it was applied to mortality prediction for intensive care unit (ICU) patients. The model constructed a patient-patient graph (P-P) at the macroscopic level and a taxonomy-note-word hypergraph (T-N-W) at the microscopic level, while incorporating temporal dependencies within the hypergraph to achieve multi-scale fusion of patient features. Meanwhile, a hybrid embedding (Hybrid-E) algorithm was designed to extract and integrate latent patient features and improve the prediction accuracy. Experimental results demonstrate that H-MGNN on the medical information mart for intensive care III (MIMIC-III) dataset significantly outperforms existing methods in terms of in-hospital mortality prediction and other tasks, validating its effectiveness and superiority in complex EHR data mining.

**Key words:** electronic health record; multi-scale; hypergraph; graph neural network

在医疗领域,重症监护室(ICU)的治疗安排对于危重症患者的救治至关重要.引入电子健康

记录(EHR)使医院能够持续监测患者的状况,辅助医疗决策并优化治疗方案.作为患者状态的系

收稿日期: 2025-09-14

基金项目: 国家自然科学基金资助项目(62106020,62332017).

作者简介: 樊捷杰(1985—),男,江西余干人,北京科技大学博士研究生.

通信作者: 班晓娟, E-mail: banxj@ustb.edu.cn.

统性总结, EHR 涵盖了仪器读数、医疗程序、临床记录及医护人员笔记等信息. 通过对 EHR 进行深入分析, 可以预测患者在住院期间的健康状况及未来风险, 从而更加合理地分配医疗资源. EHR 属于典型的混合型数据, 包含结构化与非结构化两部分. 结构化数据包括生命体征、药物信息及治疗方案等; 非结构化数据主要指临床笔记, 记录医护人员对患者状态的观察、分析与判断. 尽管笔记内容分散且格式不一, 但其蕴含的关键医学信息对于临床预测具有重要价值. 目前由 EHR 驱动的预测任务<sup>[1]</sup>包括住院死亡率预测、滞留时间预测、呼吸衰竭预测及表型分类等.

早期的研究多利用时序数据预测患者健康状态<sup>[2-5]</sup>. 随后, 学者将非结构化的临床笔记引入预测任务中<sup>[6-9]</sup>. 然而临床笔记内容难以理解且结构复杂, 因此信息挖掘难度大. 随着预训练语言模型 (pre-trained language models, PLMs) 的发展逐步完善<sup>[10-11]</sup>, 研究者通过在大规模语料库上的预训练, 并结合针对临床笔记的微调 (fine-tuning) 或迁移学习, 显著提升了模型的文本理解能力和预测精度. 进一步地, 大语言模型 (large language models, LLMs) 的发展以及其检索与推理能力增强方法的成熟, 如检索增强生成 (retrieval-augmented generation, RAG), 为临床笔记的信息挖掘提供了新的研究范式<sup>[12]</sup>.

在 EHR 研究中, 时序特性同样是影响预测性能的关键因素. 例如, 文献[13-15]将时序特征与预训练模型相结合, 构建了一系列基于 Transformer 的模型, 在多项临床任务中取得了良好表现. 针对临床笔记的时序建模, 文献[16-17]使用插值等方法处理时间间隔不规则问题, 并结合 BERT (bidirectional encoder representations from Transformers) 及 Transformer 等预训练模型, 进一步提升了对非结构化临床数据的建模能力. 尽管这些模型性能优越, 但通常计算开销大, 且侧重于时序依赖建模, 对 EHR 中复杂的结构关系关注不足.

近年来, 图神经网络 (graph neural network, GNN)<sup>[18-19]</sup> 的兴起为复杂非欧几何结构的表示提供了有效的工具, 并在自然语言处理领域<sup>[20-21]</sup> 取得显著成效. 归纳式图神经网络, 如 TextING (text-based interaction graph neural network)<sup>[22]</sup>、InducT-GCN (inductive text graph convolutional network)<sup>[23]</sup> 和 SSL-GNN (sparse structure learning via graph neural networks)<sup>[24]</sup>, 有效缓解

了全局图初始化与泛化能力弱的问题; 在长文本建图过程中, 为应对边密集问题, 研究者提出图稀疏化策略与超图建模方法, 如 HyperGAT (hypergraph attention network)<sup>[25]</sup>、HEGEL (hypergraph embedding with graph-enhanced learning)<sup>[26]</sup>等; 在复杂关系构建中, 超图通过超边建立多对多联系, 天然适配层次化结构与高阶关系建模. 然而现有的 GNN 方法多聚焦于单个患者 EHR 的内部结构, 忽略了患者之间存在的关联关系, 如语义、病理关联等. 如何在大规模 EHR 数据中建立跨患者间的关系图, 并通过图学习增强个体嵌入的表达能力, 仍是一项具有临床意义的任务. 尤其是当关系图包含数万节点和数百万条边时, EHR 的图表示问题更具复杂性与挑战性.

针对上述问题, 本文提出了一种混合多层次图神经网络 (H-MGNN), 用于解决复杂异构 EHR 的图表示问题, 并将其应用于 ICU 患者的死亡预测任务. 该方法将 ICU 死亡预测问题建模为单个患者信息与群体患者结构相融合的问题, 构建了群体患者关系图 (P-P), 设计了高效的 Hybrid-E 算法以实现大规模异构图的无监督嵌入学习; 提出了单个患者 EHR 的混合多层次超图 T-N-W, 通过融合患者 EHR 内部多层次结构信息及其时序特性, 全面建模患者内部的复杂关系. 在 MIMIC-III 数据集上的实验结果表明, H-MGNN 能够显著提升预测性能, 在死亡预测任务上超过现有方法, 验证了其有效性与先进性.

## 1 相关工作

### 1.1 电子健康记录学习

EHR 数据来源广泛, 涵盖临床笔记、实验室检测、诊断编码、手术操作、药物信息及影像数据等多种模态. 随着多模态学习的发展, 研究者尝试将这些异构信息整合以提升建模效果. 例如, 文献[27]将结构化数据构建为图, 并提出图-文本融合的多模态框架以支持复杂医疗任务; M3Care<sup>[28]</sup>引入模态插补机制, 有效应对模态缺失问题; Xu 等<sup>[29]</sup>从就诊序列与临床笔记联合建模出发, 基于 Gromov-Wasserstein 距离构建对比学习框架, 实现患者间相似性增强; Chen 等<sup>[30]</sup>提出统一的模态学习框架, 利用模态感知注意力机制避免烦琐的插补流程.

在医学知识图谱 (knowledge graph, KG) 融合方面, 文献[31]基于医学本体的层次结构, 引入

图注意力机制细化诊断表示; Qiu 等<sup>[32]</sup>将本体信息嵌入预测流程中,提升模型的上下文感知能力; Ma 等<sup>[33]</sup>采用 GNN 将 KG 中高阶结构整合至患者嵌入表示中. 此外,文献[34]通过结合结构化 EHR 嵌入与疾病相关文档特征,提升健康风险预测的准确性与可解释性;文献[35]探索了患者-疾病图中的协同建模机制; Gaupp 等<sup>[36]</sup>针对多次就诊过程中的知识遗忘问题提出有效缓解策略.

近年来,LLMs 被视为医学综合知识载体的新范式. 例如,文献[37]从结构化 EHR 中构建知识图以供图神经网络学习,尽管仍面临幻觉(hallucination)等挑战,但展现出强大潜力. 此外,部分研究将注意力转向非结构化的临床笔记,并尝试与结构化数据联合建模. 如 Zhang 等<sup>[38]</sup>、Gayathri 等<sup>[39]</sup>提出将时间序列与文本融合; Huang 等<sup>[40]</sup>、Hou 等<sup>[41]</sup>、Hastuti 等<sup>[42]</sup>、Mulyar 等<sup>[43]</sup>基于 BERT 与 XLNet 的预训练模型进行扩展,在临床笔记建模任务中取得了显著提升. Kruskal 等<sup>[44]</sup>进一步提出从文本中提取上下文的表型特征,为表型分类任务提供有力支撑.

## 1.2 图嵌入学习

图嵌入学习旨在将图的结构信息映射到低维空间,同时保留节点之间的语义关系,广泛应用于节点分类、链接预测等下游任务. 其理论基础来源于自然语言处理中“相似语境的词具有相似意义”的分布假设,即“You shall know a word by the company it keeps”.

早期经典方法如多维缩放(multidimensional scaling, MDS)<sup>[45]</sup>、IsoMap<sup>[46]</sup>、局部线性嵌入(locally linear embedding, LLE)<sup>[47]</sup>以及拉普拉斯特征映射(Laplacian eigenmaps)<sup>[48]</sup>等,通常先基于数据特征向量构建邻接图(如 K 近邻图),再进行维度压缩. 这些方法大多依赖于对邻接矩阵或拉普拉斯矩阵的特征分解,计算复杂度至少为节点数的平方,因此难以扩展至大规模图. 为缓解大图建模的效率瓶颈,研究者提出了多种基于采样与矩阵分解的图嵌入方法. 例如,图分解类方法 GraRep<sup>[49]</sup> 和 HOPE (high-order proximity preserved embedding)<sup>[50]</sup>等,均可通过矩阵分解捕捉图的高阶结构,但可能损失全局语义信息. DeepWalk<sup>[51]</sup>首次引入截断随机游走与 Skip-Gram 机制,将图嵌入问题转化为类语言建模问题,显著提升了可扩展性. LINE (large-scale information network embedding)<sup>[52]</sup>在此基础上进

一步建模一阶与二阶邻近关系,有效保留图的局部和整体结构;后续提出的 PTE (predictive text embedding)<sup>[53]</sup>则扩展至异构图,支持包含类别信息的文本节点建模. 需要指出的是,传统图嵌入方法多采用无监督学习策略,即在学习过程中未利用节点的标签信息,而仅在后续分类器中引入监督信号. 尽管如此,这类方法在训练过程中能够整合节点上下文、局部邻近关系与全局拓扑特征,其学习到的表示具有良好的通用性,可广泛用于多种图相关任务. 随着 GNN 的发展,嵌入学习逐渐与半监督学习结合,为下游任务提供更丰富且结构感知能力更强的表示. 近年来,面向高阶结构关系的图建模中引入了超图表示学习,进一步拓展了图嵌入学习的研究边界.

## 2 模型方法

### 2.1 问题定义

给定图  $G=(V, E)$ , 其中  $V$  表示节点集合, 对应数据对象,  $E$  表示边集合, 表示 2 个数据对象之间的关系. 每条边  $e \in E$  关联一个权重  $w_{uv} > 0$ , 表示节点  $u$  与  $v$  之间关系的强度. 在患者院内死亡率预测任务中, 设患者集合为  $P = \{p^1, p^2, \dots, p^j, \dots\}$ . 对于任意患者  $p \in P$ ,  $p$  的 EHR 包含一系列临床笔记  $N_p = \{n_{t_1}^1, n_{t_2}^1, \dots, n_{t_k}^1, \dots\}$ , 每篇临床笔记与分类学信息  $t \in T_p$  相关联, 包含单词序列  $W_{n_t} = \{w_{n_t}^1, w_{n_t}^2, \dots, w_{n_t}^i, \dots\}$ , 其中  $j, k, i$  分别表示患者  $p$  的临床笔记  $n$ 、分类学  $t$  与单词  $w$  的索引. 分类学集合为  $T_p = \{t_1, t_2, \dots, t_k\}$ . 在建模过程中, 本文旨在为每位患者  $p$  的 EHR 构建个体超图  $G_h$  (表示为 T-N-W 图), 并通过多层消息传递机制学习患者级表示. 同时, 在患者之间构建关系图  $G_p$  (表示为 P-P 图), 利用患者间的关联进一步增强表示学习能力.

### 2.2 模型总览

为了充分建模患者之间的复杂联系, 以及 EHR 内部的数据结构与时序特征, 本文构建了 2 个关系图: 患者关系图  $G_p$  与单个患者 EHR 的混合多层级超图  $G_h$ . 在  $G_p$  中, 边权重基于互信息的词共现关系进行初始化, 并设计了 Hybrid-E 算法用于学习和更新患者嵌入  $p_r$ . 在  $G_h$  中, 采用两层超图结构: T-W 表示分类学 (taxonomy) 与单词 (word) 之间的超图; N-W 表示临床笔记 (note) 与单词 (word) 之间的超图. 通过图消息扩散机制, 信息在超边与节点之间交替传播并更新嵌入, 最终融合生成患者嵌入  $p_h$ . 鉴于 EHR 内部文本 (如

临床笔记)的字符顺序通常意义有限,而EHR自身时序特性至关重要(例如临床笔记与分类学内容会随时间动态演化),本文进一步单独建模这种时序依赖.具体而言,采用LSTM(long short-

term memory)提取患者的分类学内容与临床笔记的时序特征,生成时序嵌入 $p_t$ .最后,将三部分嵌入 $p_t$ 、 $p_h$ 和 $p_i$ 输入至融合模块,得到最终患者表示,用于下游预测任务,如图1所示.

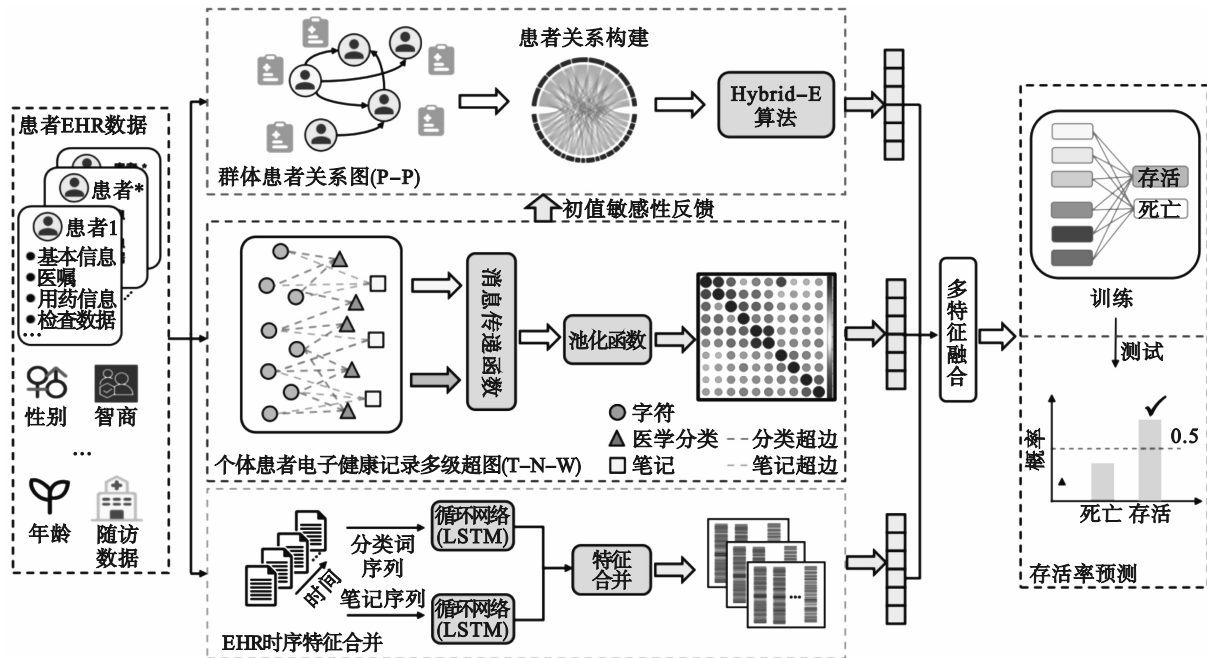


图 1 H-MGNN 框架图

Fig. 1 Architecture diagram of H-MGNN

### 2.3 群体患者关系图(P-P)

为了充分利用患者EHR之间的复杂关联,本文构建了患者关系图 $G_p=(V_p, E_p)$ .其中, $V_p$ 表示患者集合, $E_p$ 表示患者之间的关系集合,边 $e \in E_p$ 的权重 $w_{uv} > 0$ 表示患者 $u$ 与 $v$ 之间关系的强度.本文对 $G_p$ 进行了建模与可视化实验,选取5类疾病,每类样本按照患者的存活时间划分为3个区间:48 h内、48 h至1周、1周以上.每类每个区间各选取10例,共150例样本.如图2所示,在嵌入空间中,存活时间较短的患者(深色节点)聚集得更紧密,表明其特征具有高度相似性;同一疾病的患者则呈现基于色系的层次分布,说明在同一疾病类别内部,患者之间还具有邻近结构的相似性.

初始化策略:本文采用互信息与词共现相结合的方式初始化 $G_p$ 的边权重.首先,从临床语料中选取1 800个关键词,计算其互信息以确定语义相关性.随后,在设定的词共现窗口内统计词共现频率,并将互信息与词共现频率的乘积(归一化后)定义为边权重.为了保证图的稀疏性,引入超参数阈值 $\alpha$ ,屏蔽权重过小的边.

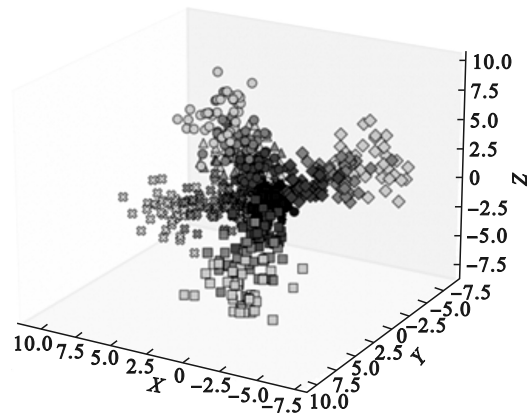


图 2 患者样本 3D 示例(不同形状代表不同疾病,颜色深浅代表疾病的严重程度)

Fig. 2 3D visualization of patient samples (different shapes represent different diseases, color intensity indicates disease severity)

建模目标:患者关系图旨在捕捉节点之间的一阶邻近性(first-order proximity)和高阶邻近性(multi-hop proximity).一阶邻近性:表示2个节点直接连接的强度.若边 $e_{uv} \in E_p$ 存在,其权重 $w_{uv}$ 即为一阶邻近强度;若不存在边,则邻近强度为0.一阶邻近性主要体现节点的局部邻近关系.高阶邻近性:扩展了一阶邻近性,用于刻画节点在更

大范围网络结构中的相似性. 定义  $\mathbf{p}_u = \{w_{u1}, \dots, w_{u|V|}\}$  为节点  $u$  与其他所有节点的一阶邻近性向量. 若  $u$  与  $v$  通过  $N$  个中间节点相连, 则为  $N+1$  阶邻近性. 以三阶邻近为例,  $u$  与  $v$  通过 2 个中间节点  $s$  与  $z$  连接, 其邻近值为

$$w_3(u, v) = \sum_{s \in N(u)} \sum_{z \in N(s) \cap N(v)} w_{us} \cdot w_{sz} \cdot w_{zv}. \quad (1)$$

Hybrid-E 算法: 算法分为一阶邻近性与高阶邻近性两部分. 一阶邻近性建模: 通过最小化节点嵌入的联合概率与边权重对应的经验概率之间的 KL 散度 (Kullback-Leibler divergence) 来优化. 对于边  $(i, j)$ , 其节点  $v_i$  和  $v_j$  的联合概率  $p_1$  为

$$p_1(v_i, v_j) = \frac{1}{1 + \exp(-\mathbf{u}_i^T \cdot \mathbf{u}_j)}. \quad (2)$$

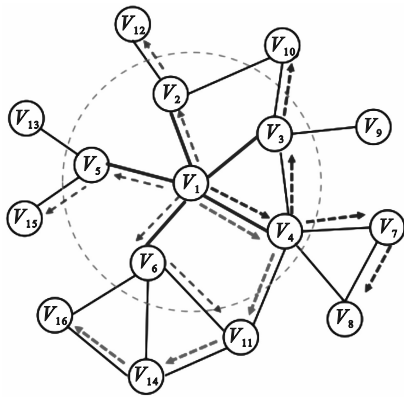


图 3 Hybrid-E 算法一阶邻近和高阶邻近路径生成图

Fig. 3 Graphs generated by Hybrid-E algorithm for first-order and high-order neighboring paths

高阶邻近性建模: 采用随机游走 (random walk) 策略来获取节点的“上下文”信息. 选择任意节点  $v_i$  为根节点, 以边的权重作为转移概率进行游走. 设  $v_k$  是游走中的一个节点, 下一个被选中的节点由以下概率  $p_2$  分布生成:

$$p_2(u_l | u_k) = \begin{cases} \pi_{kl}, & e \in (v_k, v_l), \\ 0, & e \notin (v_k, v_l). \end{cases} \quad (6)$$

其中:  $\pi_{kl}$  是节点  $v_k$  和  $v_l$  之间的转移概率. 设  $G_p$  的邻接矩阵为  $\mathbf{A}$ , 其非零项表示边的强度, 使用归一化的邻接矩阵 (随机游走拉普拉斯矩阵)  $\mathbf{L} = \mathbf{D}^{-1} \mathbf{A}$  作为转移矩阵, 其中  $\mathbf{D}$  是  $\mathbf{A}$  的度矩阵. 设定游走长度  $l$ , 可以获得 1 条以根节点  $v_i$  为锚点的游走路径, 反复采样多条游走路径, 即可得到 1 个  $l$  跳路径集合. 通过设置不同的  $l \in \{2, 3, \dots, L\}$  即可得到多尺度的路径集合, 用来共同刻画  $v_i$  的高阶邻近性. 如图 3 所示, 在  $l$  上的一组路径  $W_{v_i}^l$  可表示为

$$W_{v_i}^l = \{W_{v_i}^{l,k} = (v_{k0} = v_i, v_{k1}, \dots, v_{kl}) \mid k=1, \dots, K\}. \quad (7)$$

其中:  $\mathbf{u}_i, \mathbf{u}_j \in \mathbf{R}^d$  为节点  $v_i$  和节点  $v_j$  的嵌入. 经验概率  $\hat{p}_1$  的定义为

$$\hat{p}_1(i, j) = \frac{w_{ij}}{W}, W = \sum_{(i,j) \in E} w_{ij}. \quad (3)$$

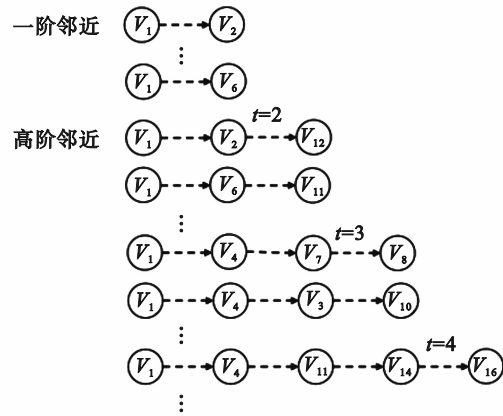
其中:  $w_{ij}$  为边  $e_{ij}$  的权重. 算法的优化目标函数  $O_1$  为

$$O_1 = d(\hat{p}_1(\cdot, \cdot), p_1(\cdot, \cdot)). \quad (4)$$

其中:  $d(\cdot, \cdot)$  是 2 个分布之间的距离, 可变换成 2 个概率分布的 KL 散度:

$$O_1 = - \sum_{(i,j) \in E} w_{ij} \log(p_1(v_i, v_j)). \quad (5)$$

结合负采样策略 (negative sampling), 可以学习得到包含一阶邻近信息的低维嵌入  $\{\mathbf{u}_i\}_{i=1}^{|V|}$ .



其中:  $W_{v_i}^{l,k}$  表示从根节点  $v_i$  开始的第  $k$  条长度为  $l$  的路径. 由于  $G_p$  是同构图, 可以直接将路径集合作为节点的上下文输入 Skip-gram 模型, 学习到包含高阶邻近信息的节点嵌入  $\{s_i\}_{i=1}^{|V|}$ .

通过融合一阶邻近和高阶邻近的嵌入可以得到  $G_p$  的嵌入:

$$p_r = \text{Concat}(\{\mathbf{u}_i\}_{i=1}^{|V|}, \{s_i\}_{i=1}^{|V|}). \quad (8)$$

其中: Concat 为融合函数, 本文默认采用向量拼接操作. 尽管 P-P 患者关系图规模庞大、训练收敛较慢, 但通过融合一阶与多阶邻近性嵌入, 能够获得兼具拓扑结构与语义相似性的患者嵌入  $p_r$ , 从而有效提升下游预测任务的性能.

#### 2.4 单个患者 EHR 混合多层次超图 (T-N-W)

混合多层次超图定义: 为了充分建模单个患者 EHR 的复杂层次结构, 引入分类学 (taxonomy)、临床笔记 (note) 和单词 (word) 这三类对象. 其中, 分类学用于对医疗数据进行系统化组织; 临床笔

记录患者在诊疗过程中的详细信息;单词则作为笔记和分类学内容的最小语义单元.三者的关系表现为:分类学与临床笔记之间的关系类比于“分类体系”与“笔记实例”,而单词则是临床笔记和分类学的具体内容的载体.基于此,构建超图  $G_h=(V_h, E_h)$ ,其中包含单词节点  $V_w$  以及两类超边  $\{E_N \cup E_T\} \in E_h$ ,临床笔记超边记为  $e_n \in E_N$ ,分类学超边记为  $e_t \in E_T$ ,其初始嵌入分别为  $h_{e_n}$  和  $h_{e_t}$ .超边也存在交互关系,假设 1 个单词节点  $v$  同时属于笔记  $n$  和分类学  $t$ ,则该节点的超边为  $E(v) = \{e_n, e_t | v \in n, n \in t\}$ .

层次化消息传递机制:为充分利用不同超边的语义特性,本文设计了分层超图卷积网络,使不同类型超边能够独立完成消息传递与表示更新.其核心消息传递函数定义如下:

$$\left. \begin{aligned} F_w(h, E, \theta) &= \sigma \left( W \sum_{u \in E(v)} \frac{1}{\sqrt{\hat{d}_v \hat{d}_u}} h_u \right), \\ F_\tau(h, E, \theta) &= \sigma \left( W \sum_{z \in V_\tau(e)} \frac{1}{\sqrt{\hat{d}_e \hat{d}_z}} h_z \right), \tau \in \{1, 2\}. \end{aligned} \right\} (9)$$

其中:  $F_w(\cdot)$  表示节点的消息传递函数,通过聚合连接的超边  $E(v)$  的嵌入来更新节点嵌入  $h_v$ ;  $F_\tau(\cdot)$  表示类型为  $\tau \in \{1, 2\}$  的超边的消息传递函数(1 和 2 分别表示分类学超边和临床笔记超边),通过聚合连接的节点  $V_\tau(e)$  的嵌入来更新超边嵌入  $h_e$ ;  $\sigma$  表示非线性激活函数(如 ReLU);  $W \in \mathbf{R}^{d \times d}$  是维度为  $d$  的权重矩阵,不同层次的超图可以定义和学习不同的权重矩阵,实现跨层次的语义捕捉.在分层学习过程中,通过掩蔽不同的超边来实现层级学习.

学习过程:该过程包含临床笔记消息传递层和分类学消息传递层两部分.

临床笔记消息传递层:在该层中,仅保留单词节点  $v \in V_w$  与临床笔记超边  $e_n \in E_N$ ,屏蔽分类学超边.单词节点仅与临床笔记超边交互,以学习临床笔记语义信息.其学习过程定义如下:

$$\left. \begin{aligned} h_N(v) &= F_w(h_\tau(v), E_N(v), W_N), \\ h_N(e_n) &= F_\tau(h_\tau(e_n), V_\tau(e_n), W_N), \\ h_N(e_t) &= h_\tau(e_t), \tau = 1. \end{aligned} \right\} (10)$$

其中:  $h_N(v)$ ,  $h_N(e_n)$  分别为当前临床笔记层的节点和超边的嵌入;  $h_\tau(v)$ ,  $h_\tau(e_n)$  分别为分类学层的节点和超边的嵌入.

分类学消息传递层:与临床笔记消息传递层

的学习机制相似,该层仅保留单词节点  $v \in V_w$  与分类学超边  $e_t \in E_T$ ,屏蔽临床笔记超边.两层的消息并非完全孤立,单词节点在前一层已融合临床笔记信息,因此该层能在分类学关系中进一步强化节点与超边的联合语义:

$$\left. \begin{aligned} h_T(v) &= F_w(h_N(v), E_T(v), W_T), h_T(e_n) = h_N(e_n), \\ h_T(e_t) &= F_\tau(h_N(e_t), V_\tau(e_t), W_T), \tau = 2. \end{aligned} \right\} (11)$$

通过上述层次化消息传递过程,最终获得融合多级语义的节点与超边嵌入  $h_\tau(v)$ ,  $h_\tau(e_n)$  和  $h_\tau(e_t)$ .随后,通过池化操作聚合这些嵌入,生成单个患者 EHR 的多级超图嵌入  $p_h$ .

## 2.5 EHR 时序特征建模

时序是 EHR 的关键特征之一.在分析相关研究与实验结果后发现:EHR 内部单词的字符顺序对患者死亡预测的影响有限;若直接在 T-N-W 超图中引入时序关系,无法充分体现时序特征的重要性.为此,将 EHR 视作时间序列数据  $X_{ET}$ (参考和遵循 Harutyunyan 等<sup>[54]</sup>的预处理方法),输入至 LSTM 以捕捉动态特征.具体而言,由于 EHR 中的临床笔记和分类学内容均随时间动态变化,为增强时序建模能力,将 T-N-W 超图中的临床笔记与分类学部分分别构建时间序列,并独立送入 LSTM 网络进行处理,最后融合两部分的隐藏状态,生成患者的时序嵌入  $p_t$ .

如图 1 中的时序部分,在临床笔记部分,对于每个时间步  $t$ ,以临床笔记的单词嵌入序列  $v = \{v_1, v_2, \dots, v_n\}$  作为输入,先通过均值池化(mean-pooling)得到一个  $d$  维特征向量  $x_t^{\text{note}}$ ,再将其输入 LSTM 更新隐藏状态:

$$\left. \begin{aligned} x_t^{\text{note}} &= \text{Pooling}(v_1, v_2, \dots, v_n), \\ h_t^{\text{note}} &= \text{LSTM}(x_t^{\text{note}}, h_{t-1}; \theta), t = 1, \dots, T. \end{aligned} \right\} (12)$$

其中:  $\theta$  是 LSTM 的可训练参数;  $h_t^{\text{note}}$  为临床笔记的时序嵌入.对于分类学部分,其操作和上面类似,但是分类学内容通常较为稀疏,直接使用会表达能力不足.为此,本文将同一时间步下的分类学节点嵌入  $\{v_1, v_2, \dots, v_k\}$  进行拼接与投影(Proj 为投影函数),得到低维嵌入  $x_t^{\text{tax}}$  再输入到 LSTM,得到分类学部分的时序嵌入  $h_t^{\text{tax}}$ :

$$\left. \begin{aligned} x_t^{\text{tax}} &= \text{Proj}(v_1 \otimes v_2 \otimes \dots \otimes v_k), \\ h_t^{\text{tax}} &= \text{LSTM}(x_t^{\text{tax}}, h_{t-1}; \theta), t = 1, \dots, T. \end{aligned} \right\} (13)$$

最终,将  $h_t^{\text{note}}$  和  $h_t^{\text{tax}}$  通过拼接或加权融合的方式组合,得到患者的时序嵌入  $p_t$ .

## 2.6 模型训练

在 H-MGNN 模型中, P-P 患者关系图规模庞

大,其无监督训练方式与T-N-W超图的监督训练存在差异,同时对初始嵌入较为敏感.为此,模型采用分阶段训练策略以保证稳定性与效率.具体流程分为四阶段:首先,对T-N-W超图训练至收敛,获得患者的EHR的嵌入 $p_h$ 以及对应的节点嵌入 $V_h$ ;其次,使用 $V_h$ 和 $p_h$ 初始化P-P图,并独立训练至收敛,得到包含患者关系信息的嵌入 $p_r$ ;再次,以 $V_h$ 及T-N-W超图的超边划分结果作为输入,初始化时序建模部分.最后,将时序部分、T-N-W超图 $p_h$ 及P-P图的嵌入 $p_r$ 融合训练.在该阶段,P-P图部分的参数冻结不再更新,仅优化其余模块,以生成最终的患者EHR嵌入.

### 3 实验分析

#### 3.1 数据集及预处理

MIMIC-III<sup>[1]</sup>是一个公开可用的临床数据库,收录了约46 520名于2001年至2012年期间在Beth Israel Deaconess Medical Center(贝斯以色列女执事医疗中心)接受ICU治疗的患者的医疗记录.该数据集涵盖人口统计信息、生理测量、实验室检查、医疗干预、药物使用、文本记录,以及诊断和手术编码等多维度信息,存储在26个关联表中.

为确保实验的公平性与可复现性,本文筛选出包含患者进入重症监护室(ICU)后48 h内临床笔记的ICU住院记录作为研究对象.数据预处理及训练/测试集划分遵循文献[55]提出的基准设置,并从训练集中随机划分20%作为验证集.在超图T-N-W部分,延续文献[55]提出的基准方法,每位患者的临床笔记数量上限设为30篇,保留出现频率最高的前6类分类学标签参与建模.在时序特征建模部分,时序数据被限制在患者住院后的前48 h内,采样间隔设为1 h.若某1 h内某变量存在多个记录,则保留该时间段内最晚的1条记录.缺失值通过前向填充法进行处理;若无可用历史值,则使用文献[54]提供的默认特征值替代.

#### 3.2 基准方法

本文选取了基于词、时序、图和超图的四类基准方法进行对比,并对数据集进行裁剪以适配不同的基线模型.词嵌入方法采用FastText<sup>[55]</sup>,该方法基于词的上下文信息生成词向量,能够有效缓解低频词汇带来的数据稀疏问题;时序建模方法包括双向长短期记忆网络(bi-directional long short-term memory, Bi-LSTM)<sup>[56]</sup>,通过捕捉文本

中的双向时序信息来提高文本分类性能,带注意力机制的双向LSTM(bi-LSTM with attention, Bi-LSTM-Att)<sup>[57]</sup>,通过注意力机制自动聚焦输入序列中的重要部分;图方法包括文本图神经网络(inductive text classification via graph neural network, TextING)<sup>[23]</sup>,通过构建文本词共现图,并利用GNN提取词之间的依赖关系;归纳图卷积网络(inductive text graph convolutional network, InducT-GCN),通过GNN对文本进行图结构建模,适用于未知类别的文本分类任务.超图方法包括超图注意力网络(hypergraph attention network, HyperGAT)<sup>[25]</sup>,结合超图结构和注意力机制来提升模型的表示能力;多层分类学超图神经网络(taxonomy-aware multi-layer hypergraph neural network, TM-HGNN)<sup>[58]</sup>,通过分类学超图结构建模复杂文本结构,提升分类效果.

#### 3.3 实验设置

所提方法基于TensorFlow实现.实验在配备8个Intel(R)Xeon(R)E5 CPU以及1块NVIDIA A800 GPU的阿里云服务器上进行.T-N-W超图的训练采用交叉熵损失函数和Adam优化器,学习率为0.05,批量大小设置为64.在P-P图的训练中,批量大小为8,负采样的比例为3,词嵌入初始化采用GLOVE 300<sup>[16]</sup>,随机游走长度 $l$ 设为3,顶点游走路径数为 $20 \times l$ .时序部分的2个LSTM的时序长度均设为48,以对应采样间隔.所有学习到的嵌入向量最后统一进行 $L_2$ 归一化.

#### 3.4 实验结果

本文采用文献[54]定义的院内死亡率预测作为基准任务,使用ICU住院前48 h的数据预测患者是否在出院前死亡.考虑到数据集样本不平衡(正负样本比例约为1:9),采用Precision-Recall曲线下面积(AUPRC)和ROC(receiver operating characteristic)曲线下面积(AUROC)两项指标评估模型性能.表1展示了H-MGNN与各类基线模型的实验结果,其中部分基线结果直接引用自文献[55](带\*表示对算法进行调整以匹配当前数据集).可以看出,相较于其他基线模型,H-MGNN在多项指标上均取得了显著提升.

综合对比来看,基于时序的方法普遍优于基于词的方法,表明时序信息是EHR中的关键特征;基于图的方法又优于纯时序的方法,说明图结构能够更好地表示EHR中的复杂关系.此外,分层超图模型在性能上超过了以往的图方法,说

明其所建模的高阶关系具有更强的表达能力. 本文提出的 H-MGNN 模型在超图方法的基础上进一步提升性能, 主要得益于两个方面: 其一, 引入了患者间的外部关系, 这些关系提供了更丰富的上下文信息; 其二, 将临床笔记和分类学内容中的时序结构从文本中抽离出来, 强调了结构层面的时序特征, 而非单纯的词序位置, 该改进被证明对提高预测效果尤为重要. 此外, 实验观察到

H-MGNN 在 AUPRC 上的提升远高于 AUROC, 证明 H-MGNN 在小样本和不平衡数据场景下依然保持了较高的鲁棒性, 展现了良好的泛化能力. 值得注意的是, H-MGNN 的多结构融合还增强了模型在临床层面的可解释性, 例如可以通过模型中的患者关系图来揭示患者群体间潜在的相似性, 这对于辅助医生制定个性化治疗方案具有重要参考价值.

表 1 患者院内死亡率预测结果  
Table 1 Prediction results of in-hospital mortality for patients

类别	模型	整体		高血压		糖尿病	
		AUPRC	AUROC	AUPRC	AUROC	AUPRC	AUROC
字符	FastText	17.06±0.08	62.37±0.11	25.56±0.28	62.39±0.18	31.33±0.33	67.89±0.20
时序	Bi-LSTM	17.67±4.19	58.75±5.78	21.75±5.25	57.39±6.11	27.52±7.57	61.86±8.38
	Bi-LSTM w/Att	17.96±0.61	62.63±1.31	26.05±1.80	63.24±1.57	33.01±3.53	68.89±1.58
图	TextING	34.50±7.79	78.20±4.27	36.63±8.30	80.12±4.05	36.13±8.66	80.28±3.84
	InducT-GCN	43.03±1.96	82.23±0.72	41.06±2.95	85.56±1.24	40.59±3.07	84.42±1.45
超图	HyperGAT	44.42±1.96	84.00±0.84	42.32±1.78	86.41±1.01	40.08±2.45	85.03±1.20
	TM-HGNN*	46.15±1.43	84.45±0.62	44.50±1.66	87.15±0.34	41.24±1.72	85.68±1.05
本文算法	H-MGNN	<b>49.80±0.54</b>	<b>86.35±0.30</b>	<b>48.67±0.72</b>	<b>87.95±0.45</b>	<b>44.62±0.80</b>	<b>87.85±1.16</b>

注: AUPRC 和 AUROC 指数结果取 5 次实验结果平均值, ± 表示标准差. 黑体为最佳结果. 下同.

### 3.5 消融实验

模块有效性分析: 为验证 H-MGNN 各模块的有效性, 本文依次移除部分组件进行对比实验. 首先, 屏蔽 P-P 患者关系图, 得到模型 T-MGNN. 随后, 进一步去除 T-MGNN 模型的时序

建模部分, 得到仅包含 T-N-W 图部分的简化模型, 简称 T-N-W. T-N-W 与超边模型 TM-HGNN\* 结构相似, 但不包含单词在超图中的位置信息. 实验结果如表 2 所示.

表 2 模块有效性消融实验结果  
Table 2 Ablation experiment results of module effectiveness

消融操作	模型	整体		高血压	
		AUPRC	AUROC	AUPRC	AUROC
去除 P-P 模块	T-MGNN	46.27±1.05	84.90±0.84	44.65±0.60	87.45±0.67
去除 P-P 及时序	T-N-W	45.63±1.26	84.23±0.55	43.36±0.85	86.68±0.78
增加内部位置	TM-HGNN*	46.15±1.43	84.45±0.62	44.50±1.66	87.15±0.34
本文算法	H-MGNN	<b>49.80±0.54</b>	<b>86.35±0.30</b>	<b>48.67±0.72</b>	<b>87.95±0.45</b>

从表 2 可以看出, T-MGNN 相较于 H-MGNN 的性能有显著下降, 表明通过 P-P 患者关系图生成的患者嵌入  $p_i$  在提升模型分类能力方面发挥了关键作用. T-MGNN 的表现优于 TM-HGNN\*, 而 T-N-W 的性能略低于 TM-HGNN\*, 说明临床笔记和分类学内容中的时序特性对模型整体性能具有较大影响; 相比之下, TM-HGNN\* 所保留的内部单词位置信息和顺序关系对模型贡献有限.

### 3.6 模型敏感性

为评估模型对初始化的敏感性, 本文比较了

2 种方式初始化患者嵌入  $p_i$  在 P-P 图无监督学习后的聚合效果: 一是直接对患者 EHR 中所有单词特征进行均值池化作为  $p_i$  的初始值; 二是使用 T-N-W 超图预训练得到的  $p_h$  来初始化  $p_i$ . 并设计了两类任务对模型聚合效果进行评估, 在实验中, 本文采用 t-SNE (t-distributed stochastic neighbor embedding) 对患者嵌入进行可视化, 以直观展示聚类效果.

疾病类型聚类: 根据国际疾病分类第九版 (ICD-9), 选取 5 种出现频率较高的疾病样本, 各取 95 例, 包括房颤、心脏病、高血压、糖尿病和急

性肾衰竭(部分样本存在疾病标签重叠).如图 4 所示,训练后的 P-P 关系图展现出较强的聚类能力.其中,采用 T-N-W 图进行预训练所获得的患

者表示(图 4a)在聚类分布上表现得更加紧凑和清晰,优于直接均值池化方式的效果(图 4b).

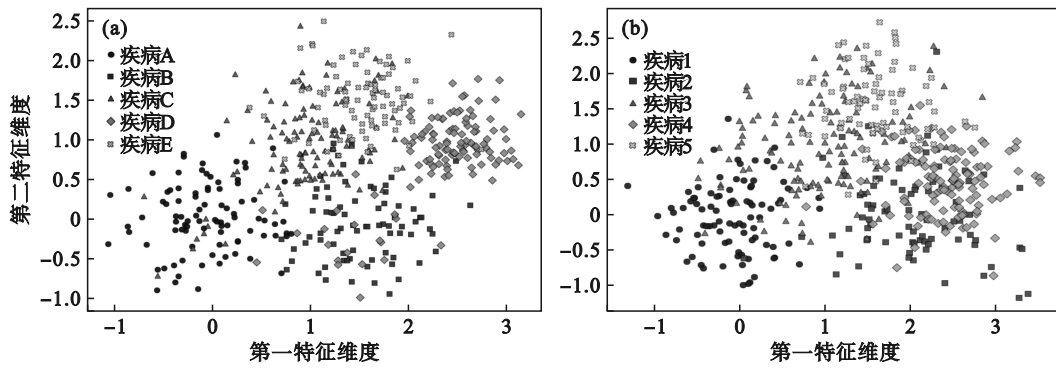


图 4 患者疾病种类聚合图

Fig. 4 Aggregation graph of patient disease categories

(a)—使用 T-N-W 预训练; (b)—不使用 T-N-W 预训练.

存活时间聚类:按患者的生存时间,将共选取的 280 个样本划分为 4 类:小于 48 h、48 h 至 7 d、7 d 至 1 月、大于 1 月(若无死亡记录,视为存活超过 1 月).如图 5 所示,使用 T-N-W 超图预训练得

到的嵌入  $p_h$  进行初始化表现出明显更优的聚类效果,这表明利用超图结构进行预训练能够有效提升患者表示的判别性和聚类质量.

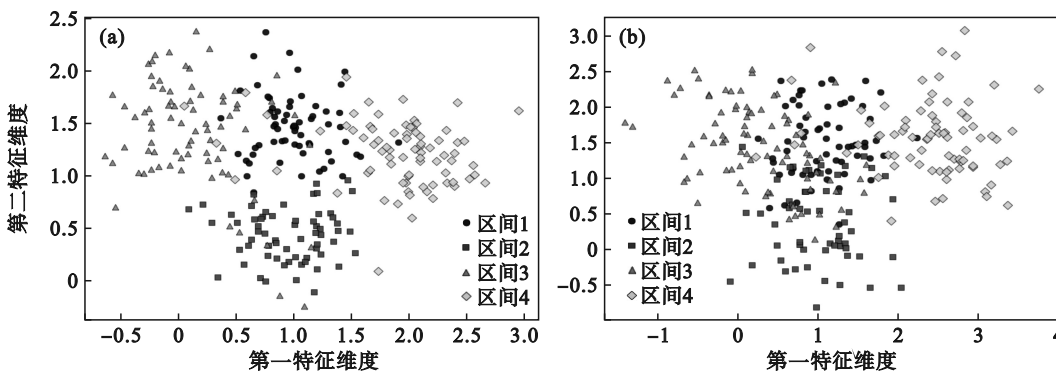


图 5 患者存活时间聚合图

Fig. 5 Aggregation graph of patient survival time

(a)—使用 T-N-W 预训练; (b)—不使用 T-N-W 预训练.

### 4 结 语

ICU 患者的 EHR 涵盖了其住院期间的生命体征、用药信息及临床笔记等多模态数据,蕴含丰富的临床信息.然而,EHR 数据具有规模庞大、结构异构、时序特性强等特点,给建模与分析带来了较大挑战.为此,本文提出了一种混合多尺度结构图表示学习模型(H-MGNN),从多个视角对 EHR 进行结构建模:包括反映患者间潜在病理或语义关联的患者关系图(P-P),捕捉 EHR 内部语义结构与高阶关系的超图(T-N-W),及用于表示临床笔记时间依赖性的时序模块.通过对这些不同来源的表示进行融合,H-MGNN 能够生成

全面且语义丰富的患者嵌入表示.在 MIMIC-III 数据集上的死亡预测任务中,H-MGNN 相较现有方法取得了显著性能提升.未来的研究将进一步融入医学统计学、临床规则建模与特征工程等领域的专业知识,以深入挖掘 EHR 中更具解释力的生命信息特征,推动模型在实际医疗决策中的落地与应用.

### 参考文献:

[ 1 ] Johnson A E W, Pollard T J, Shen L, et al. MIMIC-III, a freely accessible critical care database [J]. *Scientific Data*, 2016, 3: 160035.

[ 2 ] Lipton Z C, Kale D C, Elkan C, et al. Learning to diagnose with LSTM recurrent neural networks [C]// International Conference on Learning Representations (ICLR). San Juan, 2016:1-8.

- [ 3 ] Che Z P, Purushotham S, Cho K, et al. Recurrent neural networks for multivariate time series with missing values [J]. *Scientific Reports*, 2018, 8: 6085.
- [ 4 ] Malone B, Garcia-Duran A, Niepert M. Learning representations of missing data for predicting patient outcomes [EB/OL]. (2018-12-12) [2025-02-18]. <https://arxiv.org/abs/1811.04752>.
- [ 5 ] Xu Y B, Biswal S, Deshpande S R, et al. RAIM: recurrent attentive and intensive model of multimodal patient monitoring data [C]//Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD '18). London, 2018: 2565-2573.
- [ 6 ] Ngo Q H, Kechadi T, Le-Khac N A. Domain specific entity recognition with semantic-based deep learning approach[J]. *IEEE Access*, 2021, 9: 152892-152902.
- [ 7 ] Rasmy L, Xiang Y, Xie Z Q, et al. Med-BERT: pretrained contextualized embeddings on large-scale structured electronic health records for disease prediction [J]. *NPJ Digital Medicine*, 2021, 4: 86.
- [ 8 ] Lee J, Yoon W, Kim S, et al. BioBERT: a pre-trained biomedical language representation model for biomedical text mining[J]. *Bioinformatics*, 2020, 36(4): 1234-1240.
- [ 9 ] Alsentzer E, Murphy J R, Boag W, et al. Publicly available clinical BERT embeddings [EB/OL]. (2019-04-06) [2025-01-10]. <https://arxiv.org/abs/1904.03323>.
- [ 10 ] Vaswani A, Shazeer N, Parmar N, et al. Attention is all you need [J]. *Advances in Neural Information Processing Systems*, 2017, 30: 6000-6010.
- [ 11 ] Huang K X, Altsaar J, Ranganath R. Clinical BERT: modeling clinical notes and predicting hospital readmission [EB/OL]. (2019-04-10) [2021-04-15]. <https://arxiv.org/abs/1904.05342>.
- [ 12 ] Lewis P, Perez E, Piktus A, et al. Retrieval-augmented generation for knowledge-intensive NLP tasks[J]. *Advances in Neural Information Processing Systems*, 2020, 33: 9459-9474.
- [ 13 ] Song H, Rajan D, Thiagarajan J, et al. Attend and diagnose: clinical time series analysis using attention models [C]//AAAI Conference on Artificial Intelligence. New Orleans: AAAI Press, 2018: 4091-4098.
- [ 14 ] Hirszowicz O, Aran D. ICU bloodstream infection prediction: a transformer-based approach for EHR analysis [C]//Artificial Intelligence in Medicine. Cham: Springer, 2024: 279-292.
- [ 15 ] Li Y K, Rao S, Solares J R A, et al. BEHRT: transformer for electronic health records [J]. *Scientific Reports*, 2020, 10: 7155.
- [ 16 ] Pennington J, Socher R, Manning C. GLOVE: global vectors for word representation [C]// Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP). Doha: ACL, 2014: 1532-1543.
- [ 17 ] Tipirneni S, Reddy C K. Self-supervised transformer for sparse and irregularly sampled multivariate clinical time-series [J]. *ACM Transactions on Knowledge Discovery from Data*, 2022, 16(6): 105. 1-105.17.
- [ 18 ] Kipf T N, Welling M. Semi-supervised classification with graph convolutional networks [EB/OL]. (2016-09-09) [2020-01-05]. <https://arxiv.org/abs/1609.02907>.
- [ 19 ] Brody S, Yahav E, Levy O. Attentive neural processes [EB/OL]. (2021-01-17) [2022-01-05]. <https://doi.org/10.48550/arXiv.1901.05761>.
- [ 20 ] Liu X E, You X X, Zhang X, et al. Tensor graph convolutional networks for text classification [C]// Proceedings of the AAAI Conference on Artificial Intelligence. Philadelphia: AAAI Press, 2020: 8409-8416.
- [ 21 ] Yao L, Mao C S, Luo Y. Graph convolutional networks for text classification [C]// Proceedings of the AAAI Conference on Artificial Intelligence. Los Angeles: AAAI Press, 2019: 7370-7377.
- [ 22 ] Zhang Y F, Yu X L, Cui Z Y, et al. Every document owns its structure: inductive text classification via graph neural networks [EB/OL]. (2020-04-22) [2021-05-10]. <https://doi.org/10.48550/arXiv.2004.13826>.
- [ 23 ] Wang K Z, Han S C, Poon J. InducT-GCN: inductive graph convolutional networks for text classification [C]// 2022 26th International Conference on Pattern Recognition (ICPR). Montreal: IEEE, 2022: 1243-1249.
- [ 24 ] Piao Y h, Lee S S, Lee D, et al. Sparse structure learning via graph neural networks for inductive document classification [C]//Processing of the AAAI Conference on Artificial Intelligence. Vancouver, 2022: 11165-11173.
- [ 25 ] Ding K Z, Wang J L, Li J D, et al. Be more with less: hypergraph attention networks for inductive text classification [EB/OL]. (2020-11-01) [2023-05-10]. <https://arxiv.org/abs/2011.00387>.
- [ 26 ] Zhang H P, Liu X, Zhang J W. HEGEL: hypergraph transformer for long document summarization [EB/OL]. (2022-08-09) [2023-05-10]. <https://arxiv.org/abs/2210.04126>.
- [ 27 ] Park S, Bae S, Kim J, et al. Graph-text multi-modal pre-training for medical representation learning [C]// ACM Conference on Health, Inference, and Learning. Online, 2022: 261-281.
- [ 28 ] Zhang C H, Chu X, Ma L T, et al. M3Care: learning with missing modalities in multimodal healthcare data [C]// Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining. Washington DC, 2022: 2418-2428.
- [ 29 ] Xu Y X, Yang K, Zhang C H, et al. VeccoCare: visit sequences-clinical notes joint learning for diagnosis prediction in healthcare data [C]// Proceedings of the Thirty-Second International Joint Conference on Artificial Intelligence. Macau, 2023: 4921-4929.
- [ 30 ] Chen D X, O'Bray L, Borgwardt K M. Structure-aware transformer for graph representation learning [C]// International Conference on Machine Learning. Online, 2022: 3469-3489.
- [ 31 ] Choi E, Bahadori M T, Song L, et al. GRAM: graph-based attention model for healthcare representation learning [C]// Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. Halifax NS: ACM, 2017: 787-795.
- [ 32 ] Qiu L, Gorantla S, Rajan V, et al. Multi-disease predictive analytics: a clinical knowledge-aware approach [J]. *ACM Transactions on Management Information Systems*, 2021, 12(3): 1-34.
- [ 33 ] Ma J T, Liu B, Li K L, et al. A review of graph neural networks and pretrained language models for knowledge graph reasoning [J]. *Neurocomputing*, 2024, 609: 128490.
- [ 34 ] Mo X, Ding G H, Tang R, et al. Bipartite graphs contrastive learning with knowledge-aware diffusion-enhanced [J]. *IEEE Transaction Network Science and Engineering*, 2025, 12(5): 4182-4195.
- [ 35 ] Mishra R, Shridevi S. Knowledge graph driven medicine recommendation system using graph neural networks on longitudinal medical records [J]. *Scientific Reports*, 2024,

- 14: 25449.
- [36] Gaupp R, Dinius J, Drazic I, et al. Long-term effects of an e-learning course on patient safety: a controlled longitudinal study with medical students[J]. *PLoS One*, 2019, 14(1): e0210947.
- [37] Gupta S, Sharma S, Sharma R, et al. Healing with hierarchy: hierarchical attention empowered graph neural networks for predictive analysis in medical data [J]. *Artificial Intelligence in Medicine*, 2025, 165: 103134.
- [38] Zhang D D, Yin C C, Zeng J C, et al. Combining structured and unstructured data for predictive models: a deep learning approach [J]. *BMC Medical Informatics and Decision Making*, 2020, 20: 280.
- [39] Gayathri R, Sangeetha S K B, Sangeetha R, et al. Dynamic AI-enhanced therapeutic framework for precision medicine using multi-modal data and patient-centric reinforcement learning[J]. *IEEE Access*, 2025, 13: 77709–77733.
- [40] Huang K X, Singh A, Chen S T, et al. Clinical XLNet: modeling sequential clinical notes and predicting prolonged mechanical ventilation [EB/OL]. (2019–12–27) [2020–10–10]. <https://arxiv.org/abs/1912.11975>.
- [41] Hou L X, Zhuang Y, Xie Y H, et al. Cross-modal generalizable visual-language models via inter-modal bidirectional supervision for enhanced pathology image recognition[J]. *Pattern Recognition*, 2026, 171: 112240.
- [42] Hastuti R P, Rajagede R A, Zheng M, et al. Clinic-prompt: few-shot discrete clinical prompt optimization[C]// Workshop on Large Language Models and Generative AI for Health at AAAI 2025. Philadelphia, 2025:2451490.
- [43] Mulyar A, Schumacher E, Rouhizadeh M, et al. Phenotyping of clinical notes with improved document classification models using contextualized neural language models [EB/OL]. (2019–10–30) [2021–01–02]. <https://arxiv.org/abs/1910.13664>.
- [44] Kruskal J B. Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis [J]. *Psychometrika*, 1964, 29(1): 1–27.
- [45] Haugen E, Firth J R. Papers in linguistics 1934–1951 [J]. *Language*, 1958, 34(4): 498–502.
- [46] Tenenbaum J B, de Silva V, Langford J C. A global geometric framework for nonlinear dimensionality reduction [J]. *Science*, 2000, 290(5500): 2319–2323.
- [47] Roweis S T, Saul L K. Nonlinear dimensionality reduction by locally linear embedding [J]. *Science*, 2000, 290(5500): 2323–2326.
- [48] Belkin M, Niyogi P. Laplacian eigenmaps and spectral techniques for embedding and clustering [C]// Advances in Neural Information Processing Systems 14. Cambridge, MA: MIT Press, 2002: 585–592.
- [49] Cao S S, Lu W, Xu Q K. GraRep: learning graph representations with global structural information [C]// Proceedings of the 24th ACM International on Conference on Information and Knowledge Management. Melbourne, 2015: 891–900.
- [50] Ou M D, Cui P, Pei J, et al. Asymmetric transitivity preserving graph embedding [C]// Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. San Francisco: ACM, 2016: 1105–1114.
- [51] Perozzi B, Al-Rfou R, Skiena S. DeepWalk: online learning of social representations [C]// Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. New York: ACM, 2014: 701–710.
- [52] Tang J, Qu M, Wang M Z, et al. LINE: large-scale information network embedding [EB/OL]. (2015–03–12) [2020–03–11]. <https://arxiv.org/abs/1503.03578>.
- [53] Tang J, Qu M, Mei Q Z. PTE: predictive text embedding through large-scale heterogeneous text networks [C]// Proceedings of the 21st ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. Sydney, NSW: ACM, 2015: 1165–1174.
- [54] Harutyunyan H, Khachatrian H, Kale D C, et al. Multitask learning and benchmarking with clinical time series data [J]. *Scientific Data*, 2019, 6: 96.
- [55] Kim N, Piao Y H, Kim S. Clinical note owns its hierarchy: multi-level hypergraph neural networks for patient-level representation learning [EB/OL]. (2023–05–16) [2025–02–20]. <https://arxiv.org/abs/2305.09756>.
- [56] Zhou P, Shi W, Tian J, et al. Attention-based bidirectional long short-term memory networks for relation classification [C]// Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics. Berlin: ACL, 2016: 207–212.
- [57] Wang Z H, Yang B. Attention-based bidirectional long short-term memory networks for relation classification using knowledge distillation from BERT [C]// 2020 IEEE International Conference on Dependable, Autonomic and Secure Computing, International Conference on Pervasive Intelligence and Computing, international conference on Cloud and Big Data Computing, international conference on Cyber Science and Technology Congress (DASC/PiCom/CBDCCom/CyberSciTech). Calgary: IEEE, 2020: 562–568.
- [58] Joulin A, Grave E, Bojanowski P, et al. Bag of tricks for efficient text classification [EB/OL]. (2016–07–06) [2023–05–10]. <https://arxiv.org/abs/1607.01759>.