

面向少样本的知识与数据的跨模态特征融合模型

柴传江¹⁾, 王成猛²⁾, 邵立珍^{1,3)✉}, 付冬梅^{1,4)✉}, 周珂¹⁾, 赵志毅²⁾

1) 北京科技大学自动化学院, 北京 100083 2) 北京科技大学材料科学与工程学院, 北京 100083 3) 北京市工业波谱成像工程技术研究中心, 北京 100083 4) 北京科技大学自动化学院工业过程知识自动化教育部重点实验室, 北京 100083

✉通信作者, 邵立珍, E-mail: lshao@ustb.edu.cn; 付冬梅, E-mail: fdm_ustb@ustb.edu.cn

摘要 在少样本学习中, 纯数据驱动式学习容易出现过拟合和泛化能力下降等问题, 而知识与数据的融合不足也会导致模型性能受限. 因此, 本文针对少样本学习问题, 提出了一种深度跨模态特征融合模型(KDFM)以融合领域知识和结构化数据特征, 进而提升下游任务的性能. KDFM 采用多特征交互框架: 首先, 基于知识图谱建模领域中用语义表达的知识的模态, 利用 TransE 算法提取知识节点的嵌入表示; 其次, 将结构化的数据模态映射为图网络, 通过多通道图卷积网络捕捉特征间的高阶关联; 最后, 设计注意力机制动态对齐知识嵌入与数据特征, 实现跨模态信息的自适应融合. 本文将所提模型分别在材料回归和医学分类两个少样本数据集上进行了验证. 相比其他纯数据驱动模型, 所提模型在各项回归和分类任务上均取得了较好的结果. 消融实验结果表明了所提模型的知识建模部分和跨模态融合部分的有效性. 这也说明 KDFM 通过多特征协同建模与高效融合策略, 在一定程度上解决了少样本下模型泛化能力弱, 知识与数据模态融合困难的问题.

关键词 少样本; 混合驱动; 知识图谱; 图卷积网络; 特征融合

分类号 TP182

Integrating knowledge and data: a cross-modal feature fusion model for few-shot problems

QI Chuanjiang¹⁾, WANG Chengmeng²⁾, SHAO Lizhen^{1,3)✉}, FU Dongmei^{1,4)✉}, ZHOU Ke¹⁾, ZHAO Zhiyi²⁾

1) School of Automation and Electrical Engineering, University of Science and Technology Beijing, Beijing 100083, China

2) School of Materials Science and Engineering, University of Science and Technology Beijing, Beijing 100083, China

3) Beijing Engineering Research Center of Industrial Spectrum Imaging, Beijing 100083, China

4) Key Laboratory of Knowledge Automation for Industrial Processes of Ministry of Education, University of Science and Technology Beijing, Beijing 100083, China

✉Corresponding author, SHAO Lizhen, E-mail: lshao@ustb.edu.cn; FU Dongmei, E-mail: fdm_ustb@ustb.edu.cn

ABSTRACT The few-shot problem is a common phenomenon in machine learning, particularly in experimental science and medical research. Pure data-driven learning relies heavily on the quality and quantity of data. When data is scarce, the model is prone to overfitting and its generalization ability will decrease. However, most fields have accumulated extensive experience and knowledge. A hybrid approach that combines domain knowledge with data can effectively improve model performance. However, in the context of few-shot problems, achieving effective cross-modal feature fusion of knowledge and data is challenging. This study proposes a knowledge and data cross-modal fusion model (KDFM) to address the few-shot problem. First, numerical modal features are categorized into different feature types and modeled using graphs. For each feature type, edges within the graphs are constructed based on K-means clustering. Then, the different types of numerical features are processed through multichannel graph convolution. These graphs convert

收稿日期: 2024–10–30

基金项目: 国家重点研发计划资助项目(2021YFB3702402); 中央高校基本科研业务费(FRF-DF-23-003)

numerical modal features into graph-level features, enhancing their expressiveness. Subsequently, domain knowledge features from semantic modalities are represented by a knowledge graph. Key entities and relationships are extracted from professional books and expert experiences. The knowledge graph consists of triples formed by combinations of entities and relationships, enabling the transformation of unstructured text features into graph-level features. Textual domain knowledge and experience are organized and converted into the neural network model. A graph convolutional neural network and attention mechanisms are employed for cross-modal feature fusion between knowledge and data. The input of the graph convolutional network includes different graphs constructed from numerical data, feature vectors obtained from the knowledge graph, and numerical vectors from the data. Based on the number of feature types, multichannel graph convolution is applied to achieve deep feature fusion of knowledge and data. The output is a fused multichannel feature vector, computed using the attention mechanism, which serves as the input feature vector for downstream tasks. The proposed model was validated using two small sample datasets: one for a regression task in the materials field and the other for a classification task in the medical field. Simulation results show that, compared with other data-driven models, the proposed KDFM model exhibits excellent performance across various regression and classification tasks. In the regression task, the model achieved the best results in terms of mean squared error, mean absolute error, and R^2 , with R^2 exceeding the suboptimal multilayer perceptron model by over 7%. In the classification task, the model was optimal in five out of seven indicators, with the remaining two indicators being suboptimal. Additionally, multiple ablation experiments were conducted to verify the effectiveness of the proposed model. By removing the modules of the knowledge graph and graph convolutional network from the full model, the study confirmed the effectiveness of both the knowledge modeling and cross-modal fusion mechanism. The proposed model addresses, to some extent, the challenges of weak generalization ability and the integration of knowledge and data modalities in few-shot problems.

KEY WORDS few-shot; hybrid drive; knowledge graph; graph convolutional network; feature fusion

在机器学习中,结构化数据是最常见的数据表现形式之一,而数据驱动式学习是处理结构化数据的一种流行范式,它利用大量数据来训练和优化模型,不依赖于先验的假设或规则。但在实际情境中,少样本的现实困境在于数据稀缺性不仅会导致模型过拟合风险骤增,更使得特征间潜在关联难以被可靠捕获。这一挑战在实验科学与医疗领域尤为突出^[1-3],使得上述的学习范式不再适用。因此,在少样本情况下如何训练出具有良好泛化能力的模型对解决实际生产下的数据稀缺问题有重要意义。

对于表格型的结构化数据,传统的数据驱动式模型能自动地从大量记录中学习特征和模式,从而适应复杂和多变的任务场景^[4-7]。而在少样本下,传统做法通常采用数据增强、迁移学习等方式来提升模型性能,其中数据增强是处理少样本最常见的做法。例如,在图像少样本问题下,通过模型生成、图像变换的数据扩充方式^[8-10],能够增强图像的分割或分类准确性。尽管这种方式在图像任务上效果显著,但结构化数据的表格式特征受严格语义约束,例如医疗指标的数值范围、材料参数的物理意义,盲目生成易破坏特征间逻辑一致性,甚至引入噪声。迁移学习通过将模型在其他任务的大数据集上学习的信息,迁移到自身任务中的小数据集中,进而提高模型在小样本数据上的

预测表现^[11-12]。但是,迁移学习的方法高度依赖源-目标领域的相关性,而垂直领域往往缺乏可用的大规模相似数据集,导致域间差异难以弥合。

尽管在垂直领域比如材料、医学领域结构化数据严重稀缺,但其中也积累了成熟的经验和知识^[13-14]。当纯数据驱动式学习无法解决少样本问题时,知识和数据的双驱动方式能有效提升模型的泛化性能^[15-19]。这引申出的另一个问题是,知识与数据应如何进行跨模态融合。目前在知识与数据的融合上主要有四种方式:数据层融合、结构层融合、训练层融合以及决策层融合。数据层面融合主要通过生成包含依赖关系的经验知识的伪训练数据方式,增强现有的训练数据,比如将变量间的条件独立关系作为先验知识加入模型中增强预测性能^[20]。结构层的融合方式是将变量间的语义关系表示,集成到学习模型中,其中知识图谱作为一种建模专家经验或领域知识等语义关系的有效工具常常被采用^[21],比如在使用卷积神经网络提取目标对象的多级组织特征中,将具有先验知识的知识图谱作为一层集成到神经网络的体系结构中^[22],或设计一个带有知识推理的结构块以增强下游任务性能^[23]。训练层的融合将经验知识看作指导模型训练的某种约束,可通过修改损失函数的形式干预模型的学习^[24-25]。决策层的融合方法则是通过将数据和先验知识的预测结合起来,最终在输出

层进行融合^[26]. 典型的工作有将分子的宏观化学特性知识, 三维信息知识等与分子式信息相结合, 使用多通道的分子聚合表示来进行性能预测^[27-28]. 在这四种融合方式中, 对于垂直领域数据的少样本问题, 结构层的融合是一种潜在的解决方案. 考虑到知识图谱出色的知识总结与推理能力, 以及图卷积网络对复杂特征的捕捉能力, 本文采用知识图谱结合图卷积网络进行跨模态特征融合.

本文提出了一个基于知识图谱和图卷积网络的深度跨模态特征融合模型, 其核心在于突破传统的纯数据模态式的驱动学习, 首次将领域知识图谱的用语义表达的知识的模态与结构化数据的数值模态统一于图网络框架中. 不同于纯数据驱动模型仅挖掘表格特征的表层统计关联, 所提模型通过知识图谱建模领域内隐逻辑, 将其编码为语义嵌入向量, 并与结构化数据的图网络表征通过多通道图卷积网络进行层次化交互. 进一步地, 模型通过自适应注意力融合机制, 在特征空间内校准语义先验与数据分布的一致性, 从而在少样本下实现更鲁棒的特征表示. 此设计不仅缓解了数据稀缺导致的过拟合风险, 更通过知识-数据约束提升了模型在多任务场景中的泛化能力. 在实验设计上, 本文选取了两个任务来验证模型普适性. 其一, 选取了典型的材料回归任务, 聚焦于成分-性能预测, 对非线性耦合关系的稀疏样本进行建模. 其二, 选取了具有代表性的医学分类任务, 对于复杂的多维特征进行风险判别. 二者分别代表了少样本学习中连续空间优化与离散决策边界强化的典型问题. 此外, 本文还进行了消融实验来验证模型中各个模块的有效性.

1 数据来源

结构化数据通常可以表示为 $\mathbf{X} = \{\mathbf{x}^1, \mathbf{x}^2, \dots, \mathbf{x}^i, \dots, \mathbf{x}^N\}$, 其中, \mathbf{x}^i 表示第 i 条数据对象, N 表示数据集大小. 每个数据对象又可以表示为 $\mathbf{x}^i = [x_1^i, x_2^i, \dots, x_n^i]^T$, 其中 n 为特征个数. 对于少样本数据集, N 的取值通常为几十至几百, 特征个数则不定. 此外, 需要从相关图书和领域专家中搜集、整理与数据集特征关系密切的领域知识, 并构建为知识图谱, 完成本文所提方法的数据准备工作. 因此, 本文面向的数据对象主要是数值模态的结构化数据和语义模态的知识.

本文在材料和医学两个领域验证了所提方法的泛化性. 其中医学领域的加州大学尔湾分校

(The University of California, Irvine, UCI) 心力衰竭公开数据集^[29]属于分类任务范畴. 该数据集包含在随访期间收集的 299 例心力衰竭患者的病历, 其中每个患者档案具有 13 种临床特征, 输入特征中一些是二元的: 贫血、高血压、糖尿病、性别和吸烟, 一些特征是连续的: 年龄、肌酸磷酸激酶 (Creatine phosphokinase, CPK) 水平、射血分数、血小板水平、血清肌酐水平、血清钠水平和随访日期, 输出特征为一维的分类变量, 代表患者是否死亡. 本文借鉴文献^[30]的处理方式, 亦舍弃“随访时间”变量, 以开展患者死亡状态分类任务预测.

材料科学领域数据集来源于科学文献^[31], 仅有 140 条实验数据. 与医学领域不同, 材料数据集任务目标属于回归任务范畴. 为保证低合金钢各元素特征值的独立性, 同时便于钢铁领域知识的引入, 本文利用数据集中已有的 Mn 的质量分数与 Mn 和 S 的质量比, 反推得到 S 的质量分数, 并以此取代原始的 Mn/S 比值特征变量. 最终数据集包括 C、Si、Mn、P、S、Mo、Cr、Ni 八种元素含量和回火温度 (Tempering temperature, TT)、冷却速率 (Cooling rate, CR) 两种热处理工艺共十维输入, 输出为屈服强度 (Yield strength, YS)、抗拉强度 (Ultimate tensile strength, UTS) 和延伸率 (Elongation, EL) 三个常见的钢材性能.

2 方法

2.1 模型框架

本文提出的深度多模态融合模型 (Knowledge and data cross-modal fusion model, KDFM) 框架如图 1 所示, 其核心流程包含数据准备、数据特征工程与跨模态融合三部分. 在数据准备阶段, 基于领域文献与专家规则提取实体及关系, 以三元组形式构建结构化知识图谱, 并通过 TransE 嵌入算法将实体与关系映射为低维向量. 数据特征工程阶段, 根据特征类型差异化定义图结构, 通过一定策略形成多组特征子图以保留模态特异性. 模型采用多通道图卷积网络处理各特征子图, 通过邻域信息聚合提取高阶特征表示; 同时, 设计注意力机制动态对齐融合的知识嵌入向量与数据特征, 计算跨模态权重并整合多源信息. 最终, 融合特征经全连接层适配下游任务.

2.2 知识特征建模

本文通过知识图谱来建模领域知识. 给定一个实体集 \mathbf{U} 和边集 \mathbf{D} , 知识图谱可以被表示为 $\mathbf{T} = (\mathbf{U}, \mathbf{D})$, 知识图谱通常由多个三元组构成, 三元

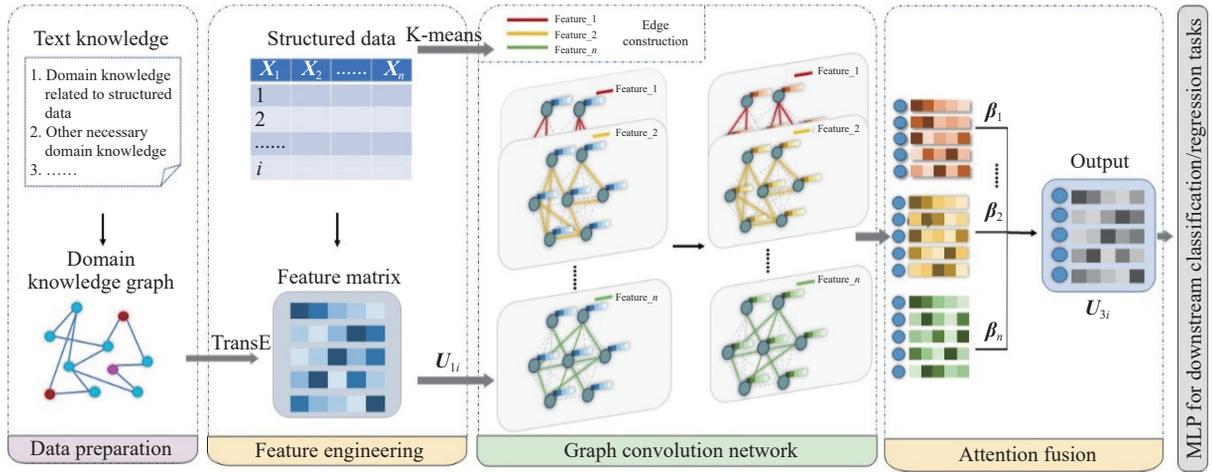


图1 KDFM 模型框架

Fig.1 KDFM framework

组可表示为 \$(u_1, d, u_2)\$, 其中 \$u_1, u_2 \in U, d \in D\$.

根据原始数值数据对象中蕴含的知识规律, 首先调研相关领域专业书籍和文献, 将所需的特定领域知识进行实体、关系抽取, 从大量非结构化数据中总结出关键的实体与关系. 其次, 将各类实体依据关系依赖构建为多个三元组. 在抽取的实体与关系中, 将原始数据对象 \$X\$ 的特征名抽取为实体 \$F = \{f_1, f_2, \dots, f_n\}\$, 从领域知识中抽取实体集 \$K = \{k_1, k_2, \dots, k_l\}\$ 和关系集 \$D = \{d_1, d_2, \dots, d_b\}\$. 则构建的垂直领域知识图谱实体集 \$U = F \cup K\$. 三元组可按照 \$(f_1, d_1, k_1), (k_1, d_2, f_3), (f_2, d_2, k_1)\$ 等形式组织, 最终构成的垂直领域知识图谱作为本文的知识特征建模对象.

在构建的垂直领域知识图谱基础上, 本文采用 TransE 算法进行知识特征的嵌入表示. TransE 是一种知识图谱嵌入表示的方法^[32], 其输入是知识图谱的三元组, 其输出是表征了实体和关系且能被下游任务接受的向量. 本文仅保留知识图谱实体集 \$U\$ 中的 \$F\$ 部分, 舍弃 \$K\$ 部分, 将实体集 \$F\$ 的特征表示作为下游任务的特征向量输入. 在保证实体集 \$F\$ 和数值特征 \$X\$ 内部元素顺序一致性的前提下, 将实体集 \$F\$ 的特征表示按序记为符号 \$s_1, s_2, \dots, s_n \in \mathbf{R}^d\$, 其中 \$d\$ 是 TransE 向量的特征维度.

2.3 数据特征工程

对于每一个数据对象, 有 \$\mathbf{x}^i = [x_1^i, x_2^i, \dots, x_n^i]^T\$, 特征共 \$n\$ 个. 此外, \$n\$ 个特征可划分 \$p\$ 个类型, 每个特征类型命名为 \$M_1, M_2, \dots, M_p (n > p)\$, 其中特征类型个数 \$p\$ 根据实际情况而定. 定义图 \$G = (V_m, E_n)\$, 其中数据对象 \$\mathbf{x}^i\$ 与节点 \$V_{mi}\$ 一一对应. 利用相应方法产生图 \$G\$ 的边连接, 其过程如下:

(1) 根据不同特征类型的数据特性, 采用差异

化划分策略: 对离散型特征, 可由特征值的唯一性直接确定; 对连续型特征, 优先基于领域知识分段, 若无先验知识则使用 K-means 聚类法产生标签值. 例如, 对于离散型特征如性别特征, 标签值由特征值唯一性直接确定; 对于连续型特征如温度值, 若存在领域划分标准如低温 \$< 100^\circ\text{C}\$, 中温 \$100 \sim 300^\circ\text{C}\$, 高温 \$> 300^\circ\text{C}\$, 则按规则分组, 否则采用 K-means 聚类.

(2) 利用标签值构造图节点的边连接. 在每个特征类型下, 在具有相同标签值的节点之间添加一条边. 例如在某个特征类型下, 聚类结果为 3 个标签值, 分别为 1, 2, 3, 那么在所有标签值为 1 的节点之间添加一条边, 2 与 3 同理.

然后对原始数值特征进行归一化处理, 得到数值特征 \$\hat{\mathbf{x}}^i = [\hat{x}_1^i, \hat{x}_2^i, \dots, \hat{x}_n^i]^T\$, 另有语义特征 \$\mathbf{s}^i = [s_1^i, s_2^i, \dots, s_n^i]^T\$, 融合获得一个高维向量, 记为 \$U_{li}\$, 它也是图节点 \$V_{mi}\$ 的特征向量. 融合过程如式 1 所示.

$$U_{li} = \hat{\mathbf{x}}^i \odot \mathbf{s}^i \quad (1)$$

其中, \$U_{li} \in \mathbf{R}^m\$, \$\odot\$ 表示元素级的乘法.

2.4 多通道图卷积融合模型

根据构造图可知, 相连的节点应有相似的特征. 因此, 采用图卷积网络来进行节点间的特征聚合. 模型的输入是具有不同边连接的多张图, 且图中每个节点 \$V_m\$ 具有特征向量 \$U_{li}\$. 那么根据边连接的不同可划分为多个通道, 分别进行图卷积网络的前向计算, 此外考虑到深层神经网络在训练过程中存在难以训练的问题, 本文在图卷积网络层的前向计算过程中添加了残差连接层. 前向计算过程用公式描述如下:

$$\text{初始化: } U_{2i(0)}^{M_1} = U_{2i(0)}^{M_2} = \dots = U_{2i(0)}^{M_p} = U_{li} \quad (2)$$

聚类算法对数据集中每个特征类型的数据分别进行聚类. 对于医学数据集, 肌酸磷酸激酶、血清肌酐、射血分数和血清钠被归为同一个特征类型, 因为它们都属于血液微观统计信息. 是否患有糖尿病、是否抽烟等离散型变量被划分为另一个特征类型, 年龄则被单独划分为一个特征类型. 因此, 医学数据集中共有三种特征类型, 对应多通道图卷积模型的三个通道. 对于血液类特征类型, 聚类算法将所有数据对象聚类为三个簇, 即三类. 同一类的数据对象对应的节点之间添加边, 以便于后续计算. 离散类变量和年龄变量则采用分段划分类别的方式处理.

本文采用 MCC 值, 精确率 (Precision), 准确率 (Accuracy), 召回率 (Recall), F1 值. 计算方式如式 (8) ~ (12) 所示.

$$MCC = \frac{TP \times TN - FP \times FN}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}} \quad (8)$$

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (9)$$

$$Precision = \frac{TP}{TP + FP} \quad (10)$$

$$Recall = \frac{TP}{TP + FN} \quad (11)$$

$$F1 = \frac{2 \times Precision \times Recall}{Precision + Recall} \quad (12)$$

其中, TP 表示实际为正类的样本被正确地分类为正类的数量; TN 表示实际为负类的样本被正确地分类为负类的数量; FP 表示实际为负类的样本被错误地分类为正类的数量; FN 表示实际为正类的样本被错误地分类为负类的数量. 特别地, MCC 值为分类综合指标, 其值越大说明分类性能越好. 除了上述指标外, 还计算了 ROC (Receiver operating characteristic) 曲线下面积 (Area under curve, AUC) 和

PR (Precision-recall) 曲线下面积 (Average precision, AP), 其中 ROC 曲线描绘了不同阈值下的真正率和假正率之间的关系, PR 曲线反应在不同阈值下精确率和召回率之间的关系, 如果曲线下面积值越大, 说明模型分类性能越好.

医学领域分类任务的目标是根据患者特征判断患者是否死亡, 训练集和测试集由原始数据集进行 8 : 2 的随机划分方式产生. 本文所提方法与数据驱动式的方法进行了对比, 包括支持向量机 (SVC)、随机森林 (RF), 多层感知器 (MLP), K 近邻 (KNN), 决策树 (Decision tree)、线性回归 (LR)、极端梯度提升树 (XGBoost)、梯度提升决策树 (GDBT) 算法. 值得注意的是, 对于数据驱动方法, 统一采用原始数据集的归一化数值 \hat{x}^i 为输入, 使用 Python 下的机器学习库进行实现, 模型超参数均为多次试验下的最优值. 对比各种方法, 测试集上的预测结果如表 1 所示. 综合多个分类任务指标, 本文提出的方法相比其他模型拥有更高的分类性能, 其中 MCC 值相比次优模型决策树提升了近 10%, 仅在 ROC 和 PR 曲线下面积指标中落后于随机森林模型, 相关数值在表 1 中以*额外标注.

3.2 材料领域实验过程及结果

钢铁材料领域知识图谱同样由多个专家的知识 and 经验构建而成, 共包括了 30 个实体, 3 种关系和 105 个三元组. 钢材知识图谱的关系被简化为 (正影响、负影响、属性) 三种. 原始数据集的特征包括材料成分和工艺参数, 进行钢铁强塑性性能预测. 除此以外, 本文另外整理的材料知识型实体包括: 合金元素、碳化物形成元素、碳化物的聚集与长大、固溶强化、相变温度等. 与医学知识图谱相比, 抽取了的知识实体更多, 但两种知识图谱的组织结构是一致的. 具体构建的材料领域知识图谱实体、关系和三元组如图 3 所示.

表 1 医学领域分类数据集各模型预测结果

Table 1 Prediction results of various models in medical field classification datasets

Methods	MCC	Precision	Acc	Recall	F1	AUC	AP
SVC	0.1166	0.6071	0.7000	0.5317	0.4990	0.5317	0.3222
Decision Tree	0.4057	0.7182	0.7166	0.6885	0.6919	0.6885	0.5755
RF	0.3688	0.7017	0.7000	0.6685	0.6703	0.7794	0.6995
MLP	0.3646	0.7644	0.6833	0.6257	0.6029	0.6257	0.5450
LR	0.2152	0.6536	0.7166	0.5753	0.5705	0.5753	0.3603
XGBoost	0.2938	0.6607	0.6667	0.6342	0.6336	0.6342	0.5180
GDBT	0.3688	0.7017	0.7000	0.6685	0.6703	0.6685	0.5554
Ours	0.4929	0.7873	0.7500	0.7114	0.7151	0.7114*	0.6281*

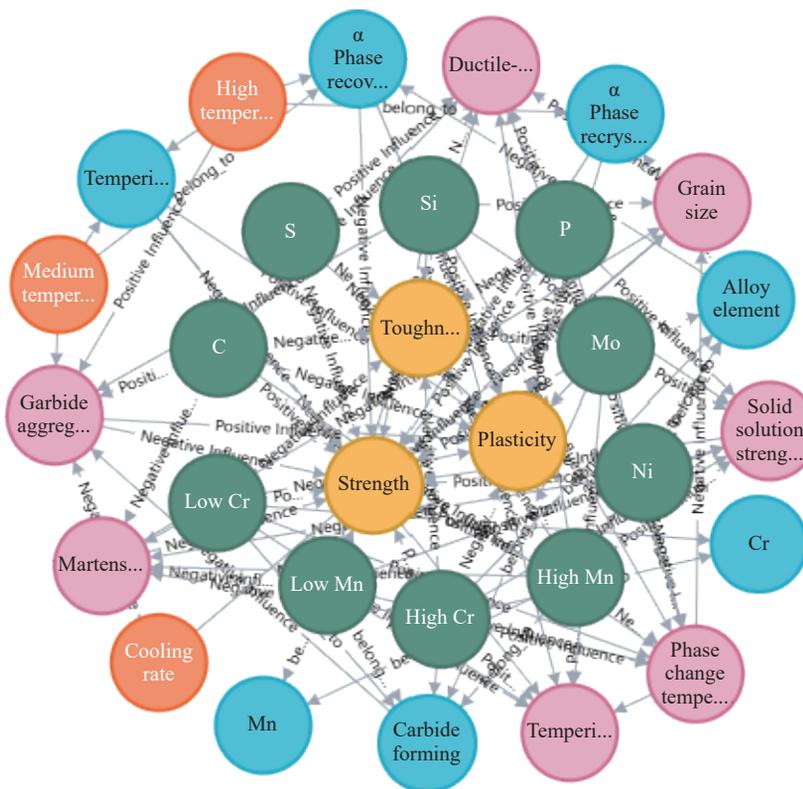


图 3 材料领域知识图谱概况

Fig.3 Overview of the knowledge graph in the field of materials

对于钢铁材料数据集, 首先根据特征的特点, 将所有特征划分为成分含量、回火温度和冷却速度三种特征类型. 因为回火温度与冷却速度均为一维变量, 采用分段式划分的方式即可. 对于成分变量, 则采用聚类的方式划分为多个簇并使用与医学数据集相同的方式添加数据对象间的边连接. 由于材料数据集划分了三个特征类型, 因此同样进行三通道的图卷积计算. 通过这种方法能够有效地处理不同特征类型的数据, 并在图卷积网络中进行特征聚合.

材料回归任务采用均方误差 (Mean squared error, MSE), 平均绝对误差 (Mean absolute error, MAE) 和决定系数 R^2 来评估性能预测的表现. 其计算方式如式 (13)~(15) 所示.

$$MSE = \frac{1}{m} \sum_{i=1}^m (Y_{pred_i} - Y_{real_i})^2 \quad (13)$$

$$MAE = \frac{1}{m} \sum_{i=1}^m |Y_{pred_i} - Y_{real_i}| \quad (14)$$

$$R^2 = 1 - \frac{\sum_{i=0}^m (Y_{pred_i} - Y_{real_i})^2}{\sum_{i=1}^m (Y_{real_i} - \bar{Y})^2} \quad (15)$$

其中, Y_{pred_i} 代表样本 i 模型的预测标签值, Y_{real_i} 代表样本 i 的实际标签值, \bar{Y} 表示测试集标签的平均值, m 表示测试集中样本总数. MSE 与 MAE 的值越小, R^2 的值越大, 说明模型的拟合精度越高.

材料领域回归任务训练集和测试集同样采用 8 : 2 的比例划分, 本文将所提方法与支持向量机 (SVM)、随机森林 (RF)、多层感知器 (MLP)、K 近邻 (KNN)、决策树 (Decision tree)、线性回归 (LR)、极端梯度提升树 (XGBoost) 和自适应提升算法 (AdaBoost) 进行了对比, 结果如表 2 和表 3 所示, 其中表 2 展示了测试集三个性能指标屈服强度 YS 、抗拉强度 UTS 和延伸率 EL 的预测结果, 表 3 展示了各种方法输出性能预测结果的平均值.

从表 2 中可以看出对比多种方法, 所提方法在各项指标中均表现出最优的结果. 表 3 结果反映了本文模型预测结果的平均 MSE、MAE 以及 R^2 均优于其他对比方法, 其中 R^2 相比其他模型的最优结果提升了约 7%.

综合两个领域下回归任务与分类任务的结果, 说明本文提出的 KDFM 方法在各项统计指标上相比其他数据驱动式的模型拥有更好的泛化性能. 在建模材料领域的稀缺数据时, 某些数据驱动式方法尽管能一定程度表达出研究对象输入与输出之间的关系, 但是由于数据量的约束, 使模型出

表2 材料领域回归数据集各模型预测结果

Table 2 Prediction results of various models in the regression dataset of the materials field

Methods	YS			UTS			EL		
	MSE	MAE	R^2	MSE	MAE	R^2	MSE	MAE	R^2
Decision tree	9604	75.39	0.5888	7508	65.07	0.5008	6.73	2.00	0.3643
RF	7406	72.76	0.6828	4796	57.11	0.6811	5.00	1.78	0.5281
XGBoost	8606	77.06	0.6314	6513	59.95	0.5669	6.18	1.77	0.4163
AdaBoost	5951	60.96	0.7452	5653	61.35	0.6241	5.56	1.77	0.4751
SVM	10462	69.01	0.3126	11625	87.47	0.2271	4.69	1.59	0.5567
LR	16304	110.97	0.3019	10087	87.76	0.3293	7.10	2.12	0.3295
MLP	5099	60.23	0.7816	5006	58.81	0.6671	4.80	1.61	0.5471
Ours	3399	50.20	0.8544	4188	50.12	0.7215	4.02	1.48	0.6204

表3 材料领域回归数据集各模型预测结果平均

Table 3 Average prediction results of various models in the regression dataset of the materials field

Methods	Average		
	MSE	MAE	R^2
Decision Tree	4296	47.48	0.4846
RF	4069	43.88	0.6307
XGBoost	5042	46.26	0.5383
AdaBoost	3870	41.36	0.6148
SVM	10462	69.01	0.3126
LR	8709	66.94	0.3203
MLP	3370	40.22	0.6653
Ours	2530	33.90	0.7322

现了泛化能力和精度大幅下降的问题. 而 KDFM 模型借助领域知识的引导, 采用知识与数据双驱

动的策略进行跨模态融合, 缓解了垂直领域数据不足的问题, 因而展现了优异的结果.

3.3 消融实验

为了进一步验证模型中每一个部分的有效性, 本文进行了多项消融实验. 首先, 为了验证模型中知识型语义特征建模的作用, 本文删除了模型中知识特征建模部分, 将原有领域知识图谱中提取的特征替换为随机向量, 仅保留数值模态信息, 重新完成医学分类与材料回归任务. 此外, 为了验证模型中图卷积特征融合部分的有效性, 本文删除了进行知识与数据特征融合的多通道图卷积网络部分和注意力机制特征交互部分, 再次完成医学分类与材料回归任务. 本文在两个任务上均对比了消融实验的结果与完整模型结果, 如表4和表5所示, 其中表4是医学分类任务两项消融实验结果, 表5是材料回归任务两项消融实验结果.

表4 医学分类任务消融实验结果

Table 4 Ablation experiment results of the medical classification task

Methods	MCC	Precision	Accuracy	Recall	F1	AUC	AP
Without knowledge	0.3761	0.7250	0.7000	0.6571	0.6534	0.6571	0.5577
Without graph convolutional neural networks	0.3076	0.7024	0.6667	0.6114	0.5896	0.6114	0.5177
Total model	0.4929	0.7873	0.7500	0.7114	0.7151	0.7114	0.6281

表5 材料回归任务消融实验结果

Table 5 Ablation experiment results of the material regression task

Methods	MSE	MAE	R^2
Without knowledge	4238	41.25	0.6694
Without graph convolutional neural networks	3892	42.14	0.6440
Total model	2530	33.90	0.7322

不难发现, 当仅有数值模态时, 所提模型的精度与纯数据驱动式模型结果相差无几, 说明领域

知识的确能缓解数据稀缺问题, 提高模型的泛化性能. 然而当没有图卷积网络和注意力机制的特征融合模块时, 模型难以充分学习到高维复杂的特征, 导致精度急剧下降. 因此, 消融实验验证了本文所提模型的每一个模块都能充分发挥作用, 实现知识与数据的跨模态融合.

4 结论

本文提出了一个知识与数据的跨模态融合模

型, 并通过两个不同领域的下游任务进行了有效性验证, 所提模型在两项任务中均表现出最优的性能. 进一步地, 通过消融实验说明所提模型在加入领域知识和跨模态融合上均有可观的效果. 本文的主要结论如下:

(1) 本文通过构建领域知识图谱的方式, 将知识型特征建模为知识图谱并进行特征嵌入, 最后通过消融实验证明了引入知识的有效性.

(2) 本文提出了一种新的特征融合模型, 采用多通道图卷积神经网络和注意力机制融合知识和数据, 解决了实际生产下的数据稀缺和模态融合困难的问题.

(3) 本文提出的模型在医学诊断领域的分类任务和材料科学领域的回归任务中相比其他数据驱动模型表现出了更优异的结果. 因此, 对同类型的结构化数据预测任务, 本文的模型能够得到更好的预测精度.

参 考 文 献

- [1] Xu P C, Ji X B, Li M J, et al. Small data machine learning in materials science. *NPJ Comput Mater*, 2023, 9: 42
- [2] Yu Z, Ye S, Sun Y L, et al. Deep learning method for predicting the mechanical properties of aluminum alloys with small data sets. *Mater Today Commun*, 2021, 28: 102570
- [3] Ge Y, Guo Y T, Das S, et al. Few-shot learning for medical text: A review of advances, trends, and opportunities. *J Biomed Inform*, 2023, 144: 104458
- [4] Guo S, Yu J X, Liu X J, et al. A predicting model for properties of steel using the industrial big data based on machine learning. *Comput Mater Sci*, 2019, 160: 95
- [5] Karamanou A, Kalampokis E, Tarabanis K. Linked open government data to predict and explain house prices: The case of Scottish statistics portal. *Big Data Res*, 2022, 30: 100355
- [6] Wang Z L, Adachi Y. Property prediction and properties-to-microstructure inverse analysis of steels by a machine-learning approach. *Mater Sci Eng A*, 2019, 744: 661
- [7] Li F F, Kuang J L, Ji J H, et al. Application of machine learning for predicting the service performance of metallic materials. *Chin J Eng*, 2024, 46(1): 120
(李丰范, 匡健隆, 季佳浩, 等. 机器学习在金属材料服役性能预测中的应用. 工程科学学报, 2024, 46(1): 120)
- [8] Xu Z H, Tang J Q, Qi C, et al. Cross-domain attention-guided generative data augmentation for medical image analysis with limited data. *Comput Biol Med*, 2024, 168: 107744
- [9] Jia X B, Guo X, Wang L, et al. A few-shot medical image segmentation network with iterative boundary refinement. *Acta Autom Sin*, 2024, 50(10): 1988
(贾熹滨, 郭雄, 王璐, 等. 一种迭代边界优化的医学图像小样本分割网络. 自动化学报, 2024, 50(10): 1988)
- [10] Sun D, Dornaika F, Charafeddine J. LCAMix: Local-and-contour aware grid mixing based data augmentation for medical image segmentation. *Inf Fusion*, 2024, 110: 102484
- [11] Zhou C L, Yuan R H, Su B L, et al. Creep rupture life prediction of high-temperature titanium alloy using cross-material transfer learning. *J Mater Sci Technol*, 2024, 178: 39
- [12] Gao Y, Li H J, Fu W Q. Few-shot learning for image-based bridge damage detection. *Eng Appl Artif Intell*, 2023, 126: 107078
- [13] Mullick A, Ghosh A, Chaitanya G S, et al. MatSciRE: Leveraging pointer networks to automate entity and relation extraction for material science knowledge-base construction. *Comput Mater Sci*, 2024, 233: 112659
- [14] Xie J, Jiang J C, Wang Y H, et al. Learning an expandable EMR-based medical knowledge network to enhance clinical diagnosis. *Artif Intell Med*, 2020, 107: 101927
- [15] Cui Z J, Gao T, Talamadupula K, et al. Knowledge-augmented deep learning and its applications: A survey. *IEEE Trans Neural Netw Learn Syst*, 2025, 36(2): 2133
- [16] Deng J D, Ni H P, Ji S, et al. Hybrid friction modeling method for robot joints integrating mechanistic model and ensemble learning. *Chin J Eng*, 2024, 46(6): 1140
(邓金栋, 倪鹤鹏, 姬帅, 等. 机理模型与集成学习混合驱动的机器人关节摩擦建模方法. 工程科学学报, 2024, 46(6): 1140)
- [17] Fang Y, Zhang Q, Yang H H, et al. Molecular contrastive learning with chemical element knowledge graph. *Proc AAAI Conf Artif Intell*, 2022, 36(4): 3968
- [18] Fang Y, Zhang Q, Zhang N Y, et al. Knowledge graph-enhanced molecular contrastive learning with functional prompt. *Nat Mach Intell*, 2023, 5(5): 542
- [19] Bo L, Pang L, Zhang C L, et al. ACRank: Injecting IR axiomatic knowledge to neural ranking models. *Chin J Comput*, 2023, 46(10): 2117
(薄琳, 庞亮, 张朝亮, 等. ACRank: 在神经排序模型中引入检索公理知识. 计算机学报, 2023, 46(10): 2117)
- [20] Takeshi T, Masashi S. Incorporating Causal Graphical Prior Knowledge into Predictive Modeling via Simple Data Augmentation [J/OL]. *arXiv preprint* (2021-08-17) [2024-10-31]. <https://arxiv.org/abs/2103.00136>
- [21] Singhal A. Official Google Blog: Introducing the knowledge graph: things, not strings [J/OL]. *Google Inc* (2012-05-02) [2024-10-31]. <http://googleblog.blogspot.pt/2012/05/introducing-know-ledge-graph-things-not.html>
- [22] Dönmez E. Enhancing classification capacity of CNN models with deep feature selection and fusion: A case study on maize seed classification. *Data Knowl Eng*, 2022, 141: 102075
- [23] Song L, Wang D L, Liu X W, et al. Prediction of mechanical properties of composite materials using multimodal fusion learning. *Sens Actuat A Phys*, 2023, 358: 114433
- [24] Xie Y, Xu Z, Kankanhalli M S, et al. Embedding symbolic knowledge into deep networks // *Proceedings of the 33rd*

- International Conference on Neural Information Processing Systems*. Vancouver, 2019: 4233
- [25] You S, Ding D, Canini K, et al. Deep lattice networks and partial monotonic functions // *Proceedings of the 31st International Conference on Neural Information Processing Systems*. California, 2017: 2985.
- [26] Marino K, Chen X L, Parikh D, et al. KRISP: Integrating implicit and symbolic knowledge for open-domain knowledge-based VQA // *2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. Nashville, 2021: 14106
- [27] Fang X M, Liu L H, Lei J Q, et al. Geometry-enhanced molecular representation learning for property prediction. *Nat Mach Intell*, 2022, 4(2): 127
- [28] Ma J B, Dai J X, Guo X, et al. Data-driven corrosion inhibition efficiency prediction model incorporating 2D–3D molecular graphs and inhibitor concentration. *Corros Sci*, 2023, 222: 111420
- [29] Ahmad T, Munir A, Bhatti S H, et al. Survival analysis of heart failure patients: A case study. *PLoS One*, 2017, 12(7): e0181001
- [30] Chicco D, Jurman G. Machine learning can predict survival of patients with heart failure from serum creatinine and ejection fraction alone. *BMC Med Inform Decis Mak*, 2020, 20(1): 16
- [31] Reddy N S, Krishnaiah J, Hong S G, et al. Modeling medium carbon steels by using artificial neural networks. *Mater Sci Eng A*, 2009, 508(1-2): 93
- [32] Antoine Bordes, Nicolas Usunier, Alberto, et al. Translating embeddings for modeling multi-relational data // *Proceedings of the 26th International Conference on Neural Information Processing Systems*. Nevada, 2013: 2787