

基于特征融合动态图网络的多标签文本分类算法^{*}

黄 靖¹, 陶竹林², 杜晓宇¹, 项欣光¹

¹(南京理工大学 计算机科学与工程学院, 江苏 南京 210094)

²(中国传媒大学 信息与通信工程学院, 北京 100024)

通信作者: 杜晓宇, E-mail: duxy@njut.edu.cn



摘 要: 多标签文本分类旨在为文本分配若干预定义的标签或类别. 为了充分发掘标签间的关联, 目前的方法通常使用标签关系图并结合图神经网络获取标签特征表示. 然而, 这类方法过度依赖初始建图策略, 忽视了当前文本中固有的标签相关性, 使得分类结果更依赖于数据集统计信息, 而容易忽视当前文本段中的标签相关信息. 因此, 提出一种基于特征融合动态图网络的多标签文本分类算法, 设计动态图来建模当前文本中的标签相关性, 并结合特征融合与图神经网络, 形成基于当前文本的标签表示, 并由此形成更为准确的多标签文本结果. 随后, 设计实验进行验证, 在 3 个数据集实验结果表明, 所提出的模型在多标签分类任务中取得优秀的性能, 验证其有效性和可行性.

关键词: 多标签文本分类; 图卷积网络; 注意力机制; 动态图

中图法分类号: TP18

中文引用格式: 黄靖, 陶竹林, 杜晓宇, 项欣光. 基于特征融合动态图网络的多标签文本分类算法. 软件学报, 2025, 36(7): 3239–3252. <http://www.jos.org.cn/1000-9825/7229.htm>

英文引用格式: Huang J, Tao ZL, Du XY, Xiang XG. Multi-label Text Classification Algorithm Based on Feature-fused Dynamic Graph Network. Ruan Jian Xue Bao/Journal of Software, 2025, 36(7): 3239–3252 (in Chinese). <http://www.jos.org.cn/1000-9825/7229.htm>

Multi-label Text Classification Algorithm Based on Feature-fused Dynamic Graph Network

HUANG Jing¹, TAO Zhu-Lin², DU Xiao-Yu¹, XIANG Xin-Guang¹

¹(School of Computer Science and Engineering, Nanjing University of Science and Technology, Nanjing 210094, China)

²(School of Information and Communication Engineering, Communication University of China, Beijing 100024, China)

Abstract: Multi-label text classification aims to assign several predefined labels or categories to text. To fully explore the correlations among labels, current methods typically utilize a label relation graph and integrate it with graph neural networks to obtain the representations of label features. However, such methods often overly rely on the initial graph construction, overlooking the inherent label correlations in the current text. Consequently, classification results heavily depend on the statistics of datasets and may overlook label-related information within the text. Therefore, this study proposes an algorithm for multi-label text classification based on feature-fused dynamic graph networks. It designs dynamic graphs to model label correlations within the current text and integrates feature fusion with graph neural networks to form label representations based on the current text, thus achieving more accurate multi-label text classifications. Experimental results on three datasets demonstrate the effectiveness and feasibility of the proposed model as it shows excellent performance in multi-label text classifications.

Key words: multi-label text classification; graph convolutional network (GCN); attention mechanism; dynamic graph

文本分类是自然语言处理中的一个重要任务, 旨在为句子、文档等文本分配预定义类别或标签^[1]. 随着互联网的迅猛发展和信息爆炸式增长, 文本数据的数量和复杂性不断增加, 在此背景下, 文本分类在诸多领域得到了广泛应用, 如专利分类^[2]、网页标记^[3]、推荐系统^[4]和情感分析^[5]等. 在实际应用中, 许多文本数据具有复杂的语义

* 基金项目: 国家自然科学基金 (62172226, 62272230); 江苏省“双创博士”人才计划 (JSSCBS20210200)

收稿时间: 2023-11-09; 修改时间: 2024-01-04, 2024-02-25; 采用时间: 2024-05-20; jos 在线出版时间: 2024-12-11

CNKI 网络首发时间: 2024-12-12

和主题, 可能涉及多个标签, 而传统的单标签文本分类方法只能为每个文本分配一个最相关的标签. 因此, 在处理涉及多主题文本数据时, 多标签文本分类是一种更有效的方法.

多标签文本分类任务是为每个文本分配多个标签, 与传统的单标签文本分类任务不同, 多标签文本分类面临着更大的挑战和复杂性, 如标签依赖^[6]、标签不平衡^[7]、标签噪声^[8]等. 早期的研究方法将多标签分类问题转化为多个独立的单标签分类问题, 针对每个标签独立训练一个分类器来预测该标签, 然后将所有分类器的输出组合成最终的多标签预测结果^[9]. 这类方法简单易实现, 但忽略了标签之间的关系. 随着深度学习的不断发展, 基于深度神经网络的方法被陆续提出, 这类方法主要通过提取增强文本表示和建模标签关系来改善分类效果. 在提取增强文本表示上, 深度学习模型通过学习文本的语义表示和上下文信息, 能够更好地捕捉文本中的关键信息, 从而提高分类的性能^[10,11]. 在建模标签关系上, 利用标签的先验知识 (如共现等) 构造标签图, 采用不同的方法在图上学习标签表示, 通过捕获标签之间的依赖关系提升分类性能^[12,13]. 在这些方法当中, 标签图通常基于数据集统计信息构建, 也即是静态图. 然而, 从训练集获取到的标签共现数据依赖全局数据分布, 不能代替文档内容特有的标签关系. 在实际情况中, 不同文本依赖不同的标签关系, 需要一种动态的方式描述标签关联.

图 1 展示了动态建图与静态建图的差异. 为文本中的词汇标注标签, 并根据标签共现关系动态调整标签关系图.

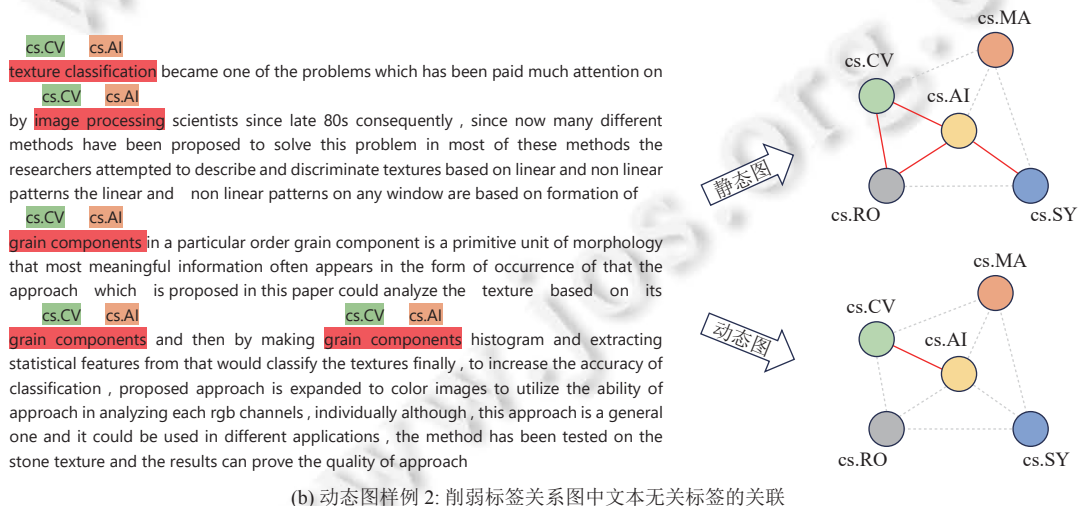
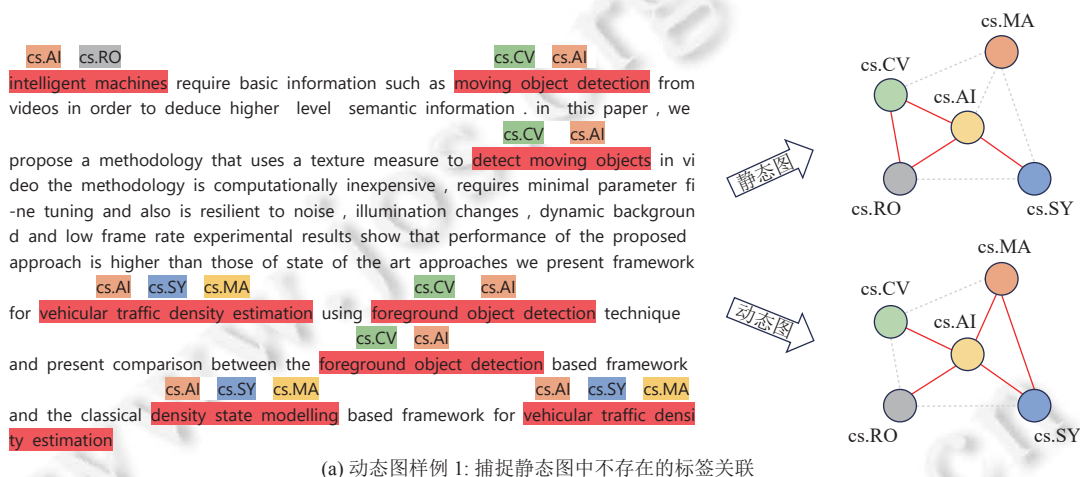


图 1 动态图的生成

图 1(a) 中描述的文档中包含 detect moving objects、moving object detection、vehicular traffic density estimation、foreground object detection 等关键词, 这些关键词不仅与人工智能 (cs.AI) 和计算机视觉 (cs.CV) 相关, 也与系统控

制 (cs.SY) 和机器自动化 (cs.RO) 有关, 这是静态图难以描述的, 需要根据文档动态扩充. 但在图 1(b) 中描述文档中正好相反, 它仅包含 texture classification、image processing、color images 等与人工智能 (cs.AI) 和计算机视觉 (cs.CV) 相关的关键词, 对其余标签并不相关, 需要削弱无关关系的关联.

针对上述问题, 本文提出了一种基于特征融合动态图网络的多标签文本分类算法 (multi-label text classification via feature-fused dynamic graph network, FDGN), 旨在通过分析当前待分类文本形成更优的个性化标签特征. 详细来说, FDGN 基于当前文本中存在的标签相关性构造动态图, 以期能够降低数据集统计偏差的影响, 准确反映标签之间的关联关系. FDGN 在动态图的基础上, 设计基于当前文本的特征融合模块, 运用图神经网络生成标签的增强语义表示, 综合文本上下文信息, 提升标签预测的效果.

本文的主要贡献如下.

(1) 设计动态图模块, 根据当前文本上下文动态调整标签之间的关联性. 与传统的静态图相比, 动态图能够更准确地捕捉标签的语义信息, 使得所提出的模型在处理复杂多样的文本数据时更具优势.

(2) 设计全局与局部特征融合模块, 综合不同层次和角度的特征信息, 提高对标签的预测能力. 有效融合多个信息源, 能够更全面地理解文本的语义内容, 从而提高多标签分类的准确性.

(3) 设计实验评估所提算法, 在 3 个基准数据集上进行实验, 验证了提出的方法的有效性. 实验结果表明, 本文的方法在很大程度上优于基线算法, 证明了其在多标签文本分类任务中的优越性和可行性.

1 相关工作

1.1 多标签文本分类

在过去的几十年里, 已经提出了许多用于多标签文本分类的方法. 总的来说, 这些方法聚焦于提取增强文本表示和建模标签关系这两方面上.

在提取增强文本表示上, 早期的方法主要依赖于手工提取特征 (如词袋模型、TF-IDF 等), 然后使用机器学习算法进行分类^[14-16]. 虽然这些传统方法在文本分类任务中取得了一定成就, 但它们对于复杂语义和长文本的处理效果有限. 随着深度学习技术的发展, 基于神经网络的方法成为研究的主方向^[17-22], 包含卷积神经网络 CNN、循环神经网络 RNN、注意力模型 BERT 等. 目前已经提出了许多用于多标签文本分类的深度学习模型, 如 TextCNN^[23]、TopicRNN^[24]、HAN^[25]等. 这些深度学习方法能够学习丰富的文本语义表示, 显著提高多标签文本分类的性能. 然而, 这些方法侧重于提取文本表示, 没有考虑到标签和文本之间的交互关系. 肖琳等人^[26]注意到这个问题, 利用标签语义注意力机制捕获每个标签关注的单词, 为每个标签学习一个标签相关的文本表示. 该方法虽然通过在文本中提取标签相关的文本信息进而考虑了标签与文本之间的关系, 但忽视了标签之间的关系.

在建模标签关系上, 早期的方法主要基于链式规则和序列生成算法. Read 等人^[27]使用链式规则来进行分类, 将前一个分类器的输出作为后一个分类器的输入, 每个分类器都可以考虑前面的标签作为上下文, 从而能够捕捉标签之间的关联. Nam 等人^[28]将标签组合视为序列, 将多标签任务转变为一个标签序列生成任务, 通过递归神经网络来进行分类. 但是这类方法对于标签顺序较敏感, 不适合关系复杂的场景. 为了解决这个问题, 基于图神经网络的方法被提出. 这类方法通过构建标签之间的关系图 (如标签共现图、标签依赖图), 然后将标签表示通过图上的边进行传递和聚合, 通过学习到信息更为丰富的标签表示来提升长尾标签的预测效果. Ozmen 等人^[29]基于推拉关系设计了一个多关系图, 通过多关系神经网络 (CompGCN^[30]) 来建模标签之间的推拉关系.

1.2 注意力机制

在自然语言处理任务中, 注意力机制能够帮助模型关注输入文本中与当前预测有关的词语或片段, 从而提取更重要的语义信息. 在多标签文本分类任务当中, 注意力机制允许模型集中关注输入文本中与特定标签相关的关键信息, 从而提高分类的准确性和效率^[31]. 近年来, 一些研究基于注意力机制来探索标签和文档之间的语义联系, 它们将深度学习模型与注意力机制相结合, 以提高模型的效率. 肖琳等人^[26]利用交叉注意力机制从文档内容信息中提取标签特定的文档表示, 可以更好地学习标签的表示. Guo 等人^[32]将注意力机制与 CNN-RNN 模型相结合,

取得了良好的效果. Zhou 等人^[33]将长短期记忆网络和注意力机制相结合提取文本中的语义信息. 这类方法聚焦于建模标签和文本之间的关系, 能够充分利用标签信息, 但忽略了标签之间存在的关系. 基于此, 本文首先通过注意力机制建模标签与文本之间的关系, 为每个标签生成一个文本表示. 同时, 为了建模标签之间的关系, 为每个文本生成了一个标签动态图, 通过图神经网络优化标签表示.

1.3 图神经网络

随着文本数据的快速增长和多标签分类任务的复杂性不断增加, 图神经网络作为一种强大的图结构建模工具, 为解决多标签文本分类问题提供了新的方向. 图神经网络能够利用标签之间的关联关系构建标签图, 并在图上进行特征提取和表示学习. 常见的图神经网络主要有以下几种: 一是图卷积网络 (GCN)^[34], 通过聚合每个节点的邻居的信息来更新节点的表示. 二是图注意力网络 (GAT)^[35], 通过引入注意力机制来改进节点更新过程, 使得节点更专注于与其相关的邻居节点. 三是多关系神经网络 (CompGCN)^[30], 这类方法将节点之间多种关系与节点联合学习, 能够捕获节点之间的多种关系. 在多标签文本分类领域, 图神经网络的应用十分广泛. Liu 等人^[36]提出了一种结合了图卷积网络 (GCN)、BERT 和记忆网络的混合模型. Pal 等人^[37]提出了一种基于 GAT 模型 MAGNET, 通过计算标签之间的共现频率来建立标签之间的边, 然后利用图神经网络来学习标签的表示. Ozmen 等人^[29]将自注意力机制和多关系图卷积相结合, 获得了基于推拉关系的标签表示, 并且提出了基于卡方检验的标签图构造方法. 这类方法基于数据集当中的先验信息为所有样本构造同一个标签图, 忽视了文本段中固有的标签相关性, 使得图中信息的传播受到数据集统计信息的影响, 限制了初始信息较少的长尾标签的应用, 因此具有一定的局限性. 因此, 本文引入了动态图, 基于文本内容、标签信息来为每个样本生成一个标签图, 从而能够挖掘文本当中蕴含的标签关系.

2 方 法

2.1 问题描述

在多标签文本分类任务当中, 训练集 $D = \{(x_i, y_i)\}_{i=1}^N$ 由 N 篇文档 x_i 和相应的标签 $Y = \{y_i \in \{0, 1\}^m\}$ 组成, m 是标签的总个数. 每篇文档 x_i 都由一系列单词组成, $x_i = \{w_1, w_2, \dots, w_n\}$, 其中 n 是文档的长度, $w_i \in \mathbb{R}^d$ 是第 i 个词向量. 标签是文本, 将其表示为 $L = \{l_1, l_2, \dots, l_m\}$, 每个词向量 $l_i \in \mathbb{R}^d$. 多标签文本分类的目标是根据训练集训练一个预测模型, 能够将一个新的未标记的样本分类到 m 个语义标签中.

2.2 模型框架

本文提出了一种多标签文本分类模型 FDGN, 如图 2 所示. 该模型采用了三阶段来充分利用文档和标签之间的关联信息. 首先, 该模型在第 1 阶段分别利用自注意力网络和图神经网络来提取文本和标签表示. 在得到文本和标签表示后, 为了建模标签和文本之间的关系, 通过注意力机制在文本中为每个标签提取相关内容, 得到标签相关的文本表示. 然后, 考虑到文本之间标签关系是不同的, FDGN 在第 2 阶段设计了一个特征融合动态图模块. 该模块结合标签表示和文本表示为每个文本生成一个标签动态图, 在图上优化标签表示. 最后, FDGN 在第 3 阶段将标签相关的文本表示与标签语义表示进行匹配得到最后的分类结果. 下面对 FDGN 进行分步骤描述.

在第 1 阶段, 与之前的一些研究类似, 本文采用自注意力机制来提取文档特征, 这样可以有效捕捉文档内部的语义信息和重要单词. 给定文档 $x_i = \{w_1, w_2, \dots, w_n\}$, 受 Transformer^[38]的启发, 文档中的每个词向量将经过多头自注意力机制和前馈神经网络层的处理, 从而得到文档特征表示. 公式如下:

$$\text{MultiHead}(Q, K, V) = \text{Concat}(\text{head}_1, \text{head}_2, \dots, \text{head}_n)W^o \quad (1)$$

$$\text{head}_i = \text{Attention}(QW_i^Q, KW_i^K, VW_i^V) \quad (2)$$

$$\text{Attention}(Q, K, V) = \text{Softmax}\left(\frac{QK^T}{\sqrt{d}}\right)V \quad (3)$$

$$\text{FNN}(x) = \max(0, xW_1 + b_1)W_2 + b_2 \quad (4)$$

其中, W_i^Q 、 W_i^K 、 W_i^V 、 W^o 、 W_1 、 W_2 、 b_1 、 b_2 是可训练参数, d 是向量维度, n 是注意力头个数.

通过叠加多层的多头自注意力机制和前馈神经网络后, 最后得到了文档的语义表示 $Z \in \mathbb{R}^{n \times d}$. 在建模标签关系

方面, 本文定义了一个图 $G = (\mathcal{V}, \mathcal{E})$, 其中, 节点指的是标签, 边是由标签先验共现信息构建的, 如果两个标签之间存在共现关系, 则对应节点之间由一条边相连. 接下来利用 LightGCN^[39] 在标签共现图上学习标签表示, 最后得到标签表示 $E \in R^{m \times d}$. 公式如下:

$$\hat{A} = D^{-\frac{1}{2}} \tilde{A} D^{-\frac{1}{2}} \quad (5)$$

$$e_i^{(k+1)} = \sum_{z \in N_i} \frac{1}{\sqrt{|N_i|} \sqrt{|N_z|}} e_z^{(k)} \quad (6)$$

其中, D 为节点的度矩阵, 当节点度数为 0, 会存在分母为 0 的问题. 于是本文先对邻接矩阵进行归一化, 其中 $\tilde{A} = A + A^E$, A^E 为单位对角矩阵, A 为标签共现二值矩阵, 是基于训练集构建的. 若标签 i 和标签 j 之间共现次数大于 0, 则 A_{ij} 为 1. $e_i^{(k+1)}$ 为节点 i 在第 $k+1$ 层的表示, $|N_i|$ 是节点 i 的度. 可见节点的表示是通过对其邻居节点的表示进行加权平均来得到的.

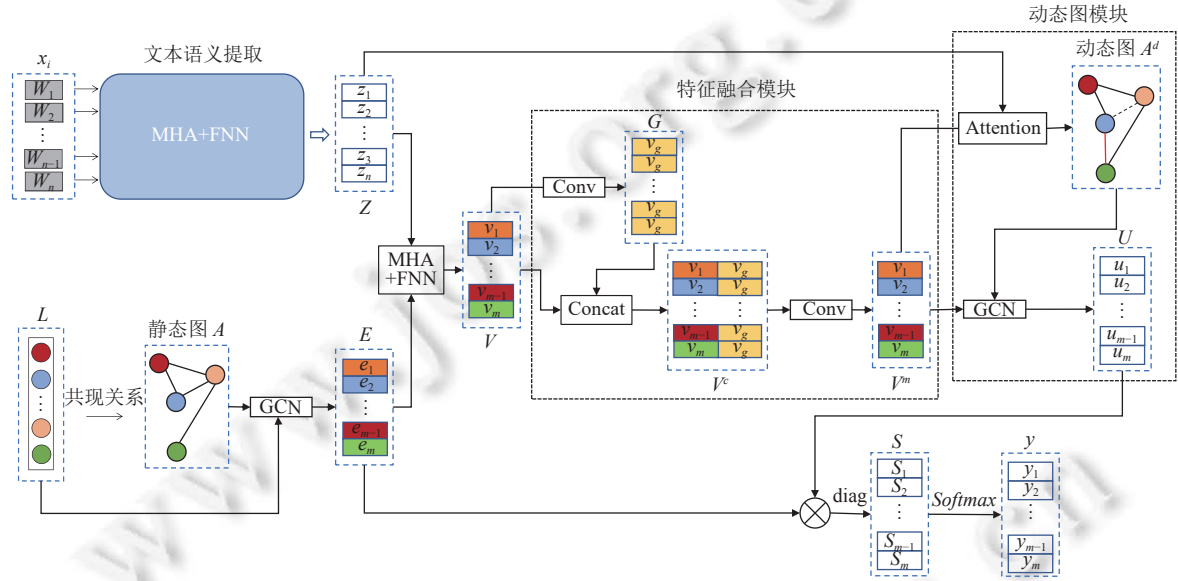


图2 本文提出的 FDGN 模型

在得到文本的表示 Z 和标签的表示 E 后, 由于文本通常较长且包含大量信息, 其中只有少部分与标签相关. 因此, FDGN 使用标签引导的注意力机制从文本中提取与标签相关的语义信息. 公式如下:

$$\alpha_{ij} = \frac{\exp(h_j e_i^T)}{\sum_j \exp(h_j e_i^T)} \quad (7)$$

$$v_i = \sum_j \alpha_{ij} h_j \quad (8)$$

其中, α_{ij} 为文本中第 j 个单词对于第 i 个标签的权重, h_j 是第 j 个单词的词向量表示, e_i 是第 i 个标签的语义表示. 因此单词和标签语义越接近, 那么权重更大, 最后得到标签相关的文本特征 V . 该操作能够充分利用标签语义信息, 有助于模块在文本中挖掘标签相关的内容.

在获得标签相关语义特征后, 本文在第 2 阶段引入了特征融合动态图网络来捕获标签之间的关系. 与第 1 阶段提到的基于共现关系的标签图不同, FDGN 的动态图是基于标签之间的关系构建的, 这意味着不同文本之间的标签关系是不同的. 与 Ma 等人^[40]不同的是, FDGN 认为标签之间的关系还与文本的全局特征有关. 例如, 船和海在与航海相关的文本中可能存在较强的关系, 而在其他文本中可能没有那么明显. 因此, 本文首先设计了一个特征融合模块, 以综合考虑标签相关的文本特征和文本全局特征. 在得到标签融合特征后, 本文基于注意力机制构建了

标签动态图, 通过在这个动态图上进行图卷积操作, 提取出该文本中蕴含的标签关系. 这一步骤有助于更准确地理解标签之间的关联性, 从而增强多标签文本分类的性能. 在上述过程之后, 能够获得标签的深层次表示. 第 3 阶段, 本文进一步将标签的深层次表示与原始标签进行匹配, 最终得到对标签的预测概率. 通过这种方法, 能够充分挖掘文本和标签之间的关联信息. 特别是, 动态图的引入能够使所提出模型能够适应不同文本和标签之间的多样性关系, 从而更好地解决实际应用中的多标签文本分类问题.

第 2.3 节将介绍特征融合动态图网络, 该节主要涉及特征融合模块和动态图模块. 第 2.4 节将介绍多标签概率预测, 该节主要涉及最后的标签概率计算.

2.3 特征融合动态图网络

FDGN 首先通过特征融合模块提取标签的混合特征. 与 Ye 等人^[41]类似, FDGN 利用卷积网络提取出文本的全局特征 G , 然后将 G 与标签相关的文本特征拼接, 最后通过卷积操作得到标签融合特征 V^m . 公式如下:

$$G = (\text{Conv1}(V))^T \delta(\text{Conv2}(V)) \quad (9)$$

$$V^c = [(v_1; v_g), (v_2; v_g), \dots, (v_{|C|}; v_g)] \quad (10)$$

$$V^m = W^c V^c \quad (11)$$

其中, Conv1 、 Conv2 为两个 1×1 的卷积, δ 为 ReLU 函数, W^c 是参数矩阵.

在构造动态图上, FDGN 先采用注意力机制得到标签与文本中每个词的注意力得分 S^c , 注意力得分能在词级别上体现标签与文本上下文的关系. 接下来 FDGN 设计了一种注意力机制来捕捉标签之间的关系. 该机制首先结合单词和文本得到每个单词在文本中的权重 W^t , 然后在每个单词上计算标签之间的相似度, 最后按权重累加得到标签之间的相关性分数, 最后得到动态图 A^d . 这个图充分体现了标签之间的动态关系, 能够结合文本更加真实地反应标签之间的关系, 公式如下:

$$A_{ij}^d = \text{Sigmoid} \left(\sum_{k=1}^n s_{i,k}^c \times s_{j,k}^c \times W_k^t \right) \quad (12)$$

$$S^c = \frac{V^m W^q Z^T W^k}{\sqrt{d}} \quad (13)$$

$$W^t = \text{Softmax}(S^s) \quad (14)$$

$$S_k^s = \sum_{j=1}^n S_{kj}^t \quad (15)$$

$$S^t = \frac{Z W^Z Z^T W^L}{\sqrt{d}} \quad (16)$$

其中, W^q 、 W^k 、 W^Z 、 W^L 为参数矩阵, d 为特征维度. S^t 为文本单词之间的注意力得分, S^c 为标签和文本单词之间的注意力得分.

在得到特定于文本的动态图后, 接下来通过公式 (5) 对邻接矩阵 A^d 进行归一化, 然后对标签融合特征 V^m 进行图卷积操作, 最后得到标签动态语义表示 $U \in \mathbb{R}^{m \times d}$.

2.4 模型训练

在上述过程之后, 获得了标签的动态语义表示 U , 然后将其与标签表示 E 进行匹配, 这里本文使用向量点积得到标签的预测得分 S . 公式如下:

$$S = \text{diag}(U^T E) \quad (17)$$

然后为了获得标签的预测概率, 本文使用 Sigmoid 函数对得分进行归一化, 公式如下:

$$y = \text{Sigmoid}(S) \quad (18)$$

本文使用 $y \in \mathbb{R}^m$ 来代表这篇文档预测的标签, 其中 y^i 为 1 表示该标签属于该文档, y^i 为 0 表示该标签不属于该文档, 损失函数为交叉熵损失函数. 本文在验证集上确定预测最佳阈值, 以此阈值为标准来得到最后的预测结果.

$$\mathcal{L} = - \sum_{i=1} y^i \log(\hat{y}^i) + (1 - y^i) \log(1 - \hat{y}^i) \quad (19)$$

3 实 验

当前大部分方法为所有样本生成一个动态图, 因而难以适应不同样本的特殊情况. 基于此, FDGN 引入了特征融合动态图网络来挖掘文本当中蕴含的标签关系. 为了评估基于特征融合动态图网络的多标签文本分类模型 FDGN 的有效性, 本文选取了精确率、汉明损失、微平均 $F1$ 分数为评价指标, 并尝试验证了以下 3 个问题.

- 问题 1: FDGN 是否比其他多标签文本分类的方法更好?
- 问题 2: FDGN 的各模块如何改善分类结果?
- 问题 3: FDGN 的实践效果如何?

3.1 数据集

为了验证本文提出的模型是有效的, 本文在学术论文引用数据集 (BIBTEX)^[42]、路透社语料库 (RCV1-V2)^[43] 和 arXiv 学术论文数据集 (AAPD)^[44] 上进行了实验. 3 个数据集的对比在表 1 中.

• BIBTEX 数据集是由 Katakis 等人^[42]创建的, 包含了一篇学术论文的元数据信息, 如标题、作者、摘要、会议或期刊信息等, 该数据集共有 159 个标签, 每个文档平均有 2.38 个标签.

• RCV1-V2 是路透社新闻文本和相应的新闻类别数据. 它包括 804414 条新闻报道, 每个新闻项目都是根据标签类别手动分类的, 每个新闻项目属于多个标签. 该数据集共有 103 个标签, 每个文档平均有 3.24 个标签.

• AAPD 是由 Yang 等人^[44]创建的多标签分类领域的一个新数据集, 包括计算机科学领域 55840 篇学术论文的摘要和主题. 该数据集共有 54 个标签, 每个文档平均有 2.41 个标签.

表 1 实验中使用的数据集

Dataset	Train	Valid	Test	Labels	Words/S	L/S
BIBTEX	4377	777	3019	159	46.47	2.38
RCV1-V2	644410	80000	80000	103	123.94	3.24
AAPD	45850	5000	5000	54	163.42	2.41

注: Words/S 和 L/S 为平均每篇文档的单词数和标签数

3.2 评估指标

多标签文本分类可以用一组指标来评估, 这些指标从不同角度来评估模型的性能^[45]. 参照 Yang 等人^[46]的工作, 本文采用微精确率 (*Micro-Precision*)、汉明损失 (*HL*)、微平均 $F1$ 值 (*Micro-F1*) 作为主要评估指标. 微精确率 (*Micro-Precision*) 计算的是预测为正例的样本数中真正的正样本所占比例, 值越高表明模型在预测样本时的准确性较高. 汉明损失 (*HL*) 衡量的是分类器在所有标签上的预测结果与实际标签之间的不匹配程度, 汉明损失越低表明模型分类性能越好. 微平均 $F1$ 值 (*Micro-F1*) 是通过分别计算每个标签的精确率和召回率, 然后对其进行加权平均得到的, 值越高表明模型的整体分类性能越好. 具体公式如下:

$$\text{Micro-Precision} = \frac{\sum_{j=1}^Q tp_j}{\sum_{j=1}^Q tp_j + fp_j} \quad (20)$$

$$HL = \frac{1}{N} \sum_{i=1}^N \frac{XOR(y_i, \hat{y}_i)}{Q} \quad (21)$$

$$\text{Micro-F1} = \frac{\sum_{j=1}^Q 2tp_j}{\sum_{j=1}^Q 2tp_j + fp_j + fn_j} \quad (22)$$

其中, Q 是标签数量, N 是样本数量, XOR 是异或运算, y_i 是第 i 个样本的预测值, $\hat{y}_i$ 是第 i 个样本的真实值, tp_j 、 fp_j 、 fn_j 分别表示第 j 个标签下真阳性、假阳性、假阴性的样本数量.

3.3 基线模型

为了进行全方面对比, 本文选取了多种模型作为基线. 其中 Binary relevance (BR)、CNN-RNN、LSAN 未考虑标签之间的相关性, Classifier chain (CC)、MRMP、LANRTN、LACO、ML-R、SGM 考虑了标签之间的相关性.

- BR^[47]为每个标签独立训练一个二元分类器 (线性 SVM).
- CC^[27]将 MLTC 问题转化为一系列二进制分类问题.
- CNN-RNN^[48]使用 CNN 和 RNN 提取全局和局部语义信息, 并对高阶标签相关性进行建模.
- SGM^[44]提出了一种新的序列生成模型, 用于解决多标签分类问题.
- LSAN^[26]借助注意力机制学习标签特定的文本表示.
- ML-R^[49]将推理机制引入多标签文本分类问题.
- LACO^[50]通过多任务学习来增强多标签文本分类中的标签相关性学习.
- MRMP^[29]在标签之间建立了推拉关系, 以此学习到标签之间更为丰富的信息.
- LANRTN^[51]使用基于注意力机制的序列建模方法用于学习文本序列的表示.

3.4 实验设置

本文基于 PyTorch 框架构建模型, PyTorch 版本为 2.1.1, Python 版本为 3.11.5. 实验程序运行在 Ubuntu 20.04.1 LTS 系统上, CPU 型号为 Xeon Platinum 8168, GPU 型号为 NVIDIA GeForce RTX 3090, CUDA 版本为 12.0. 在超参数设置方面, 本文将词向量维度和标签向量维度设为 512, 文本语义提取网络的层数设为 2, 注意力头的数量被设为 4, 图卷积层数设为 2, 标签相关语义提取网络的层数设为 2, 整个模型使用 Adam^[52]进行训练, 初始学习率为 0.0002, 步长为 10, 衰减率为 10%.

3.5 实验结果

- 问题 1: FDGN 是否比其他多标签文本分类的方法更好?

为了验证方法 FDGN 的有效性, 本文选取了经典的基线模型并比较了它们在 BIBTEX、RCV1-V2 和 AAPD 数据集上的结果, 如表 2 所示. 结果表明, 本文所提出的模型在这 3 个数据集上都具有最好的性能. 进一步的, 本文观察到如下现象.

表 2 实验结果

算法	BIBTEX			RCV1-V2			AAPD		
	<i>Micro-Precision</i>	<i>HL</i> ×10 ⁻²	<i>Micro-F1</i>	<i>Micro-Precision</i>	<i>HL</i> ×10 ⁻²	<i>Micro-F1</i>	<i>Micro-Precision</i>	<i>HL</i> ×10 ⁻²	<i>Micro-F1</i>
BR	0.531	1.66	0.388	0.904	0.85	0.857	0.664	3.16	0.646
CNN-RNN	0.537	1.48	0.411	0.889	0.87	0.856	0.718	2.80	0.664
LSAN	0.559	1.36	0.478	0.910	0.79	0.882	0.776	2.46	0.702
CC	0.535	1.62	0.407	0.887	0.87	0.857	0.657	3.06	0.654
LANRTN	0.564	1.33	0.472	0.906	0.73	0.892	0.762	2.40	0.718
LACO	0.558	<u>1.30</u>	<u>0.483</u>	0.907	<u>0.72</u>	0.883	<u>0.803</u>	2.13	0.748
SGM	0.534	1.42	0.475	0.887	0.81	0.869	0.746	2.50	0.699
ML-R	0.565	1.39	0.473	0.890	0.79	0.871	0.726	2.48	0.722
MRMP	<u>0.566</u>	1.32	0.479	<u>0.912</u>	0.73	<u>0.896</u>	0.799	<u>2.12</u>	<u>0.749</u>
FDGN	0.575*	1.27*	0.493*	0.915*	0.70*	0.912*	0.804*	2.10*	0.753*

注: 粗体格式的为最优结果, 下划线格式的为次优结果, *代表 p -value值小于0.05, 具有显著性

1) 与 BR 和 CC 等传统机器学习方法相比, 一些基于深度学习的经典模型 (如 CNN-RNN) 在多标签文本分类任务中具有明显优势, 这是因为它可以从文本中提取更深层的语义特征. 与 CNN-RNN 模型相比, LSAN 利用注意力机制为每个标签学习一个特定的文档表示, 其中在 AAPD 数据集上, *Micro-Precision* 提升了 8.08%, *HL* 减少

了 12.14%, *Micro-F1* 提升了 5.72%. 此外, 与 CNN-RNN、LSAN 等不考虑标签之间关系的方法相比, 一些基于标签之间相关性的经典模型 (如 MRMP) 在多标签文本分类任务中具有明显优势, 因而建模标签相关性对多标签分类问题非常重要. 此外还可以看出, 基于注意力机制和序列生成的 LANRTN 性能要高于 SGM 等基于序列生成的模型. 尽管 SGM 使用有序序列来对标签之间的关系进行建模, 但它忽略了标签和文档内容之间的交互. LANRTN 使用多标签注意力, 可以关注内容中最相关的部分, 并为每个标签提取不同的语义信息.

2) 本文所提出的模型在 3 个数据集上的性能都优于其他模型, 这是由于 FDGN 通过动态图为每个文本都生成了一个标签相关图, 因此可以改善基于共现的静态图的不足. 其中在 BIBTEX 数据集上, FDGN 的汉明损失为 0.0127, 是所有基线方法中最好的, 与基于注意力机制的 LSAN 相比, 本文模型在汉明损失指标上减少了 6.62%, 在 *Micro-F1* 指标上提升了 3.14%, 在 *Micro-Precision* 指标上提升了 2.86%. 与基于图卷积神经网络和注意力机制的 MRMP 相比, 本文模型在汉明损失指标上减少了 3.79%, 在 *Micro-F1* 指标上提升了 2.92%, 在 *Micro-Precision* 指标上提升了 1.59%. 在 RCV1-V2 数据集上, FDGN 的汉明损失为 0.007, 是所有基线方法中最好的, 与基于注意力机制的 LSAN 相比, 本文模型在汉明损失指标上减少了 11.40%, 在 *Micro-F1* 指标上提升了 3.40%, 在 *Micro-Precision* 指标上提升了 0.55%. 与基于图卷积神经网络和注意力机制的 MRMP 相比, 本文模型在汉明损失指标上减少了 4.11%, 在 *Micro-F1* 指标上提升了 3.40%, 在 *Micro-Precision* 指标上提升了 0.33%. 最后, 在 AAPD 数据集上, FDGN 的汉明损失为 0.021, 是所有基线方法中最好的, 与基于注意力机制的 LSAN 相比, 本文模型在汉明损失指标上减少了 14.63%, 在 *Micro-F1* 指标上提升了 7.26%, 在 *Micro-Precision* 指标上提升了 3.61%. 与 MRMP 相比, 本文模型在汉明损失指标上减少了 0.62%, 在 *Micro-F1* 指标上提升了 0.53%, 在 *Micro-Precision* 指标上提升了 0.62%. 因此, 本文提出的特征融合动态图网络是有效和有益的.

问题 2: FDGN 的各模块如何改善分类结果?

为了评估 FDGN 中所提改进对分类结果的影响, 本文通过选择不同的模型改进结构, 进行了一系列消融实验, 并将结果记录在了表 3 中. 实验表明: 1) FDGN 的各模块均能带来一定的提升, 其中的核心模块动态图对结果有着重要贡献; 2) 模型在 BIBTEX 和 RCV1-V2 数据集上的性能明显优于 AAPD, 这是由于 BIBTEX 和 RCV1-V2 数据集中的低频标签较多, FDGN 通过结合文本特征、重构标签关系, 改善了标签统计偏差对结果的影响; 3) 融合特征对实验结果有着积极意义, 这表明整合当前文档的标签特征确实能够更好地描述标签特性.

表 3 消融实验结果

模型改进结构		BIBTEX			RCV1-V2			AAPD		
融合特征	动态图	<i>Micro-Precision</i>	$HL \times 10^{-2}$	<i>Micro-F1</i>	<i>Micro-Precision</i>	$HL \times 10^{-2}$	<i>Micro-F1</i>	<i>Micro-Precision</i>	$HL \times 10^{-2}$	<i>Micro-F1</i>
√	√	0.575	1.27	0.493	0.915	0.70	0.912	0.801	2.10	0.753
×	√	0.573	1.29	0.489	0.914	0.71	0.907	0.800	2.11	0.752
×	×	0.566	1.32	0.480	0.912	0.73	0.896	0.798	2.13	0.749

问题 3: FDGN 的实践效果如何?

为了进一步的验证 FDGN 在实践中的作用, 本文可视化各样本以展示 FDGN 在建模标签相关性的优势. 图 3 和图 4 展示了两组样本中不同标签对文本的注意力权重, 其中颜色越深表明其相关性越大. 从中可以清晰地观察到, 每个标签在文本中集中关注不同的部分, 而且每个标签都有自己关心的单词. 标签“cs.CV”, 在文本中比较相关的是“image”“images”“features”“extracted”和“region” (图 3), 而标签“cs.LG”, 在文本中比较相关的是“classification”“algorithm”“clustering”“optimization”和“classified” (图 4). 这充分证明了基于当前文本构造标签动态关系, 能够更准确地描述文本与标签之间的关联, 从而在多标签文本分类任务中取得更好的效果.

为了证明 FDGN 能够动态捕捉文本当中潜在的标签关联, 挑选了一篇与数据库工程相关的文本, 对其动态图邻接矩阵进行可视化. 如图 5 所示, 两个标签的相关性越大表格颜色越深, 可以看到在动态图中标签“cs.DB”与标签“cs.SE”之间的相关性比较强. 这明显表明了 FDGN 动态图模块的优势, 它能够结合文本内容挖掘标签之间的动

态关系, 针对不同文本和标签之间的多样性关系进行更准确的建模, 从而有效地减少长尾标签的影响. 通过自适应地调整标签之间的关联权重, FDGN 能够更好地适应复杂的多标签文本分类场景, 从而提高分类性能.

biogeography is the study of the geographical distribution of biological organisms the mindset of the engineer is that we can learn from nature biogeography based optimization is a burgeoning nature inspired technique to find the optimal solution of the problem satellite image classification is an important task because it is the only way we can know about the land cover map of inaccessible areas though satellite images have been classified in past by using various techniques, the researchers are always finding alternative strategies for satellite image classification so that they may be prepared to select the most appropriate technique for the feature extraction task in hand this paper is focused on classification of the satellite image of a particular land cover using the theory of biogeography based optimization the original bbo algorithm does not have the inbuilt property of clustering which is required during image classification hence modifications have been proposed to the original algorithm and the modified algorithm is used to classify the satellite image of a given region the results indicate that highly accurate land cover features can be extracted effectively when the proposed algorithm is used



图 3 标签 cs.CV 注意力权重可视化, 颜色深浅代表相关性程度

biogeography is the study of the geographical distribution of biological organisms the mindset of the engineer is that we can learn from nature biogeography based optimization is a burgeoning nature inspired technique to find the optimal solution of the problem satellite image classification is an important task because it is the only way we can know about the land cover map of inaccessible areas though satellite images have been classified in past by using various techniques, the researchers are always finding alternative strategies for satellite image classification so that they may be prepared to select the most appropriate technique for the feature extraction task in hand this paper is focused on classification of the satellite image of a particular land cover using the theory of biogeography based optimization the original bbo algorithm does not have the inbuilt property of clustering which is required during image classification hence modifications have been proposed to the original algorithm and the modified algorithm is used to classify the satellite image of a given region the results indicate that highly accurate land cover features can be extracted effectively when the proposed algorithm is used



图 4 标签 cs.LG 注意力权重可视化, 颜色深浅代表相关性程度

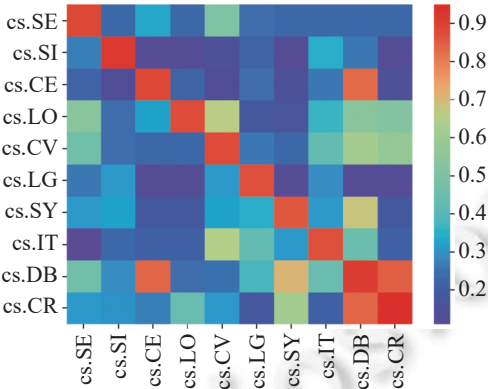


图 5 数据库工程相关文本的动态图邻接矩阵

4 总 结

本文提出了一种基于特征融合动态图网络的多标签文本分类算法, 旨在克服现有方法对于标签关系图构建策略的过度依赖以及对长尾标签的性能不足的问题. 算法通过动态图自适应地调整标签关系, 强调了重要关联并减弱了无关信息, 同时结合了文本的全局和局部特征, 以更准确地预测标签关联. 通过基于动态图的图卷积网络生成标签的最终表示, 为文本的多标签分类提供了有力支持. 实验结果验证了算法的有效性和可行性, 在多个数据集上展现出了卓越的性能. 本研究为多标签文本分类领域的发展和实际应用提供了有益的参考和启发. 在未来的研究

中将进一步探索和优化动态图网络在文本分类中的应用, 提高动态图网络的自适应性, 优化模型在不平衡数据集和带噪声标签数据中的表现, 构建适应性更强的多标签文本分类算法. 同时, 本文的研究也存在一定的局限性, 例如当数据集当中标签的数量过多时 (例如 AmazonCat-3M 数据集包含数百万个标签), 每个样本可能只与其中的小部分标签关联. 随着标签数量的增多, 计算复杂性也会增加, 这给模型的训练带来了巨大的困难.

References:

- [1] Minaee S, Kalchbrenner N, Cambria E, Nikzad N, Chenaghlou M, Gao JF. Deep learning-based text classification: A comprehensive review. *ACM Computing Surveys (CSUR)*, 2021, 54(3): 62. [doi: [10.1145/3439726](https://doi.org/10.1145/3439726)]
- [2] Tang PJ, Jiang M, Xia B, Pitera JW, Welser J, Chawla NV. Multi-label patent categorization with non-local attention-based graph convolutional network. In: *Proc. of the 34th AAAI Conf. on Artificial Intelligence*. New York: AAAI, 2020. 9024–9031. [doi: [10.1609/aaai.v34i05.6435](https://doi.org/10.1609/aaai.v34i05.6435)]
- [3] Jain H, Prabhu Y, Varma M. Extreme multi-label loss functions for recommendation, tagging, ranking & other missing label applications. In: *Proc. of the 22nd ACM SIGKDD Int'l Conf. on Knowledge Discovery and Data Mining*. San Francisco: ACM, 2016. 935–944. [doi: [10.1145/2939672.2939756](https://doi.org/10.1145/2939672.2939756)]
- [4] Fiallos A, Jimenes K. Using reddit data for multi-label text classification of Twitter users interests. In: *Proc. of the 6th Int'l Conf. on eDemocracy & eGovernment (ICEDEG)*. Quito: IEEE, 2019. 324–327. [doi: [10.1109/ICEDEG.2019.8734365](https://doi.org/10.1109/ICEDEG.2019.8734365)]
- [5] Wang YQ, Feng S, Wang DL, Yu G, Zhang YF. Multi-label Chinese microblog emotion classification via convolutional neural network. In: *Proc. of the 18th Asia-Pacific Web Conf.—Web Technologies and Applications*. Suzhou: Springer, 2016. 567–580. [doi: [10.1007/978-3-319-45814-4_46](https://doi.org/10.1007/978-3-319-45814-4_46)]
- [6] Vu HT, Nguyen MT, Nguyen VC, Tien MT, Nguyen, VH. Label correlation based graph convolutional network for multi-label text classification. In: *Proc. of the 2022 Int'l Joint Conf. on Neural Networks (IJCNN)*. Padua: IEEE, 2022. 1–8. [doi: [10.1109/IJCNN55064.2022.9892542](https://doi.org/10.1109/IJCNN55064.2022.9892542)]
- [7] Liu J, Chang WC, Wu Y, Yang Y. Deep learning for extreme multi-label text classification. In: *Proc. of the 40th Int'l ACM SIGIR Conf. on Research and Development in Information Retrieval*. Shinjuku: ACM, 2017. 115–124. [doi: [10.1145/3077136.308083](https://doi.org/10.1145/3077136.308083)]
- [8] Wei X, Huang JB, Zhao R, Yu H, Xu Z. Multi-label text classification model based on multi-level constraint augmentation and label association attention. *ACM Trans. on Asian and Low-resource Language Information Processing*, 2024, 23(1): 14. [doi: [10.1145/3586008](https://doi.org/10.1145/3586008)]
- [9] Luaces O, Díez J, Barranquero J, del Coz JJ, Bahamonde A. Binary relevance efficacy for multilabel classification. *Progress in Artificial Intelligence*, 2012, 1(4): 303–313. [doi: [10.1007/s13748-012-0030-x](https://doi.org/10.1007/s13748-012-0030-x)]
- [10] Zhang ML, Zhou ZH. Multilabel neural networks with applications to functional genomics and text categorization. *IEEE Trans. on Knowledge and Data Engineering*, 2006, 18(10): 1338–1351. [doi: [10.1109/TKDE.2006.162](https://doi.org/10.1109/TKDE.2006.162)]
- [11] Wu YJ, Li J, Wu J, Chang J. Siamese capsule networks with global and local features for text classification. *Neurocomputing*, 2020, 390: 88–98. [doi: [10.1016/j.neucom.2020.01.064](https://doi.org/10.1016/j.neucom.2020.01.064)]
- [12] Hu LM, Yang TC, Shi C, Ji HY, Li XL. Heterogeneous graph attention networks for semi-supervised short text classification. In: *Proc. of the 2019 Conf. on Empirical Methods in Natural Language Processing and the 9th Int'l Joint Conf. on Natural Language Processing (EMNLP-IJCNLP)*. Hong Kong: ACL, 2019. 4821–4830. [doi: [10.18653/v1/D19-1488](https://doi.org/10.18653/v1/D19-1488)]
- [13] Du XY, Chen Z, Xiang XG. Explainable tag-aware recommendation based on disentangled graph neural network. *Ruan Jian Xue Bao/Journal of Software*, 2023, 34(12): 5670–5685 (in Chinese with English abstract). <http://www.jos.org.cn/1000-9825/6754.htm> [doi: [10.13328/j.cnki.jos.006754](https://doi.org/10.13328/j.cnki.jos.006754)]
- [14] Clare A, King RD. Knowledge discovery in multi-label phenotype data. In: *Proc. of the 5th European Conf. on Principles of Data Mining and Knowledge Discovery*. Freiburg: Springer, 2001. 42–53. [doi: [10.1007/3-540-44794-6_4](https://doi.org/10.1007/3-540-44794-6_4)]
- [15] Elisseeff A, Weston J. A kernel method for multi-labelled classification. In: *Proc. of the 14th Int'l Conf. on Neural Information Processing Systems*. Vancouver: MIT Press, 2001. 681–687.
- [16] Zhang ML, Zhou ZH. ML-KNN: A lazy learning approach to multi-label learning. *Pattern Recognition*, 2007, 40(7): 2038–2048. [doi: [10.1016/j.patcog.2006.12.019](https://doi.org/10.1016/j.patcog.2006.12.019)]
- [17] Albawi S, Mohammed TA, Al-Zawi S. Understanding of a convolutional neural network. In: *Proc. of the 2017 Int'l Conf. on Engineering*

- and Technology (ICET). Antalya: IEEE, 2017. 1–6. [doi: [10.1109/ICEngTechnol.2017.8308186](https://doi.org/10.1109/ICEngTechnol.2017.8308186)]
- [18] Pouyanfar S, Sadiq S, Yan YL, Tian HM, Tao YD, Reyes MP, Shyu ML, Chen SC, Iyengar SS. A survey on deep learning: Algorithms, techniques, and applications. *ACM Computing Surveys (CSUR)*, 2018, 51(5): 92. [doi: [10.1145/3234150](https://doi.org/10.1145/3234150)]
- [19] Li FF, Su PZ, Duan JW, Zhang SC, Mao XL. Multi-label text classification with enhancing multi-granularity information relations. *Ruan Jian Xue Bao/Journal of Software*, 2023, 34(12): 5686–5703 (in Chinese with English abstract). <http://www.jos.org.cn/1000-9825/6802.htm> [doi: [10.13328/j.cnki.jos.006802](https://doi.org/10.13328/j.cnki.jos.006802)]
- [20] He JD, Xu BW, Yang Z, Han D, Yang CR, Lo D. PTM4Tag: Sharpening tag recommendation of stack overflow posts with pre-trained models. In: *Proc. of the 30th IEEE/ACM Int'l Conf. on Program Comprehension*. ACM, 2022. 1–11. [doi: [10.1145/3524610.3527897](https://doi.org/10.1145/3524610.3527897)]
- [21] Du XY, Wu ZK, Feng FL, He XN, Tang JH. Invariant representation learning for multimedia recommendation. In: *Proc. of the 30th ACM Int'l Conf. on Multimedia*. Lisboa: ACM, 2022. 619–628. [doi: [10.1145/3503161.3548405](https://doi.org/10.1145/3503161.3548405)]
- [22] Du XY, Wang X, He XN, Li ZC, Tang JH, Chua TS. How to learn item representation for cold-start multimedia recommendation? In: *Proc. of the 28th ACM Int'l Conf. on Multimedia*. Seattle: ACM, 2020. 3469–3477. [doi: [10.1145/3394171.3413628](https://doi.org/10.1145/3394171.3413628)]
- [23] Kim Y. Convolutional neural networks for sentence classification. In: *Proc. of the 2014 Conf. on Empirical Methods in Natural Language Processing (EMNLP)*. Doha: ACL, 2014. 1746–1751. [doi: [10.3115/v1/D14-1181](https://doi.org/10.3115/v1/D14-1181)]
- [24] Dieng AB, Wang C, Gao JF, Paisley J. TopicRNN: A recurrent neural network with long-range semantic dependency. In: *Proc. of the 5th Int'l Conf. on Learning Representations*. Toulon: OpenReview.net, 2017.
- [25] Yang ZC, Yang DY, Dyer C, He XD, Smola A, Hovy E. Hierarchical attention networks for document classification. In: *Proc. of the 2016 Conf. of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*. San Diego: ACL, 2016. 1480–1489. [doi: [10.18653/v1/N16-1174](https://doi.org/10.18653/v1/N16-1174)]
- [26] Xiao L, Chen BL, Huang X, Liu HF, Jing LP, Yu J. Multi-label text classification method based on label semantic information. *Ruan Jian Xue Bao/Journal of Software*, 2020, 31(4): 1079–1089 (in Chinese with English abstract). <http://www.jos.org.cn/1000-9825/5923.htm> [doi: [10.13328/j.cnki.jos.005923](https://doi.org/10.13328/j.cnki.jos.005923)]
- [27] Read J, Pfahringer B, Holmes G, Frank E. Classifier chains for multi-label classification. In: *Proc. of the 2009 European Conf. on Machine Learning and Knowledge Discovery in Databases*. Bled: Springer, 2009. 254–269. [doi: [10.1007/978-3-642-04174-7_17](https://doi.org/10.1007/978-3-642-04174-7_17)]
- [28] Nam J, Loza Mencía E, Kim HJ, Fürnkranz J. Maximizing subset accuracy with recurrent neural networks in multi-label classification. In: *Proc. of the 30th Conf. on Neural Information Processing Systems*. Long Beach, 2017. 30
- [29] Ozmen M, Zhang H, Wang PY, Coates M. Multi-relation message passing for multi-label text classification. In: *Proc. of the 2022 IEEE Int'l Conf. on Acoustics, Speech and Signal Processing (ICASSP)*. Singapore: IEEE, 2022. 3583–3587. [doi: [10.1109/ICASSP43922.2022.9747225](https://doi.org/10.1109/ICASSP43922.2022.9747225)]
- [30] Vashishth S, Sanyal S, Nitin V, Talukdar P. Composition-based multi-relational graph convolutional networks. In: *Proc. of the 8th Int'l Conf. on Learning Representations*. Addis Ababa: OpenReview.net, 2020.
- [31] Li BH, Xiang YX, Feng D, He ZC, Wu JJ, Dai TL, Li J. Short text classification model combining knowledge aware and dual attention. *Ruan Jian Xue Bao/Journal of Software*, 2022, 33(10): 3565–3581 (in Chinese with English abstract). <http://www.jos.org.cn/1000-9825/6630.htm> [doi: [10.13328/j.cnki.jos.006630](https://doi.org/10.13328/j.cnki.jos.006630)]
- [32] Guo L, Zhang DX, Wang L, Wang H, Cui B. CRAN: A hybrid CNN-RNN attention-based model for text classification. In: *Proc. of the 37th Int'l Conf. on Conceptual Modeling*. Xi'an: Springer, 2018. 571–585. [doi: [10.1007/978-3-030-00847-5_42](https://doi.org/10.1007/978-3-030-00847-5_42)]
- [33] Zhou P, Shi W, Tian J, Qi ZY, Li BC, Hao HW, Xu B. Attention-based bidirectional long short-term memory networks for relation classification. In: *Proc. of the 54th Annual Meeting of the Association for Computational Linguistics, Vol. 2 (Short Papers)*. Berlin: ACL, 2016. 207–212. [doi: [10.18653/v1/P16-2034](https://doi.org/10.18653/v1/P16-2034)]
- [34] Kipf TN, Welling M. Semi-supervised classification with graph convolutional networks. In: *Proc. of the 5th Int'l Conf. on Learning Representations*. Toulon: OpenReview.net, 2017.
- [35] Veličković P, Cucurull G, Casanova A, Romero A, Liò P, Bengio Y. Graph attention networks. In: *Proc. of the 6th Int'l Conf. on Learning Representations*. Vancouver: OpenReview.net, 2018.
- [36] Liu B. GCN-BERT and memory network based multi-label classification for event text of the Chinese government hotline. *IEEE Access*, 2022, 10: 109267–109276. [doi: [10.1109/ACCESS.2022.3213978](https://doi.org/10.1109/ACCESS.2022.3213978)]
- [37] Pal A, Selvakumar M, Sankarasubbu M. Multi-label text classification using attention-based graph neural network. *arXiv:2003.11644*, 2020.

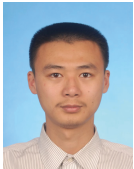
- [38] Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomez AN, Kaiser L, Polosukhin I. Attention is all you need. In: Proc. of the 31st Int'l Conf. on Neural Information Processing Systems. Long Beach: Curran Associates Inc., 2017. 6000–6010.
- [39] He XN, Deng K, Wang X, Li Y, Zhang YD, Wang M. LightGCN: Simplifying and powering graph convolution network for recommendation. In: Proc. of the 43rd Int'l ACM SIGIR Conf. on Research and Development in Information Retrieval. ACM, 2020. 639–648. [doi: [10.1145/3397271.34010](https://doi.org/10.1145/3397271.34010)]
- [40] Ma QW, Yuan CY, Zhou W, Hu SL. Label-specific dual graph neural network for multi-label text classification. In: Proc. of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th Int'l Joint Conf. on Natural Language Processing. ACL, 2021. 3855–3864. [doi: [10.18653/v1/2021.acl-long.298](https://doi.org/10.18653/v1/2021.acl-long.298)]
- [41] Ye J, He JJ, Peng XJ, Wu WH, Qiao Y. Attention-driven dynamic graph convolutional network for multi-label image recognition. In: Proc. of the 16th European Conf. on Computer Vision. Glasgow: Springer, 2020. 649–665. [doi: [10.1007/978-3-030-58589-1_39](https://doi.org/10.1007/978-3-030-58589-1_39)]
- [42] Katakis I, Tsoumakos G, Vlahavas I. Multilabel text classification for automated tag suggestion. In: Proc. of the 2008 ECML/PKDD Discovery Challenge. Antwerp: ECML, 2008. 1–9.
- [43] Lewis DD, Yang YM, Rose T G, Li F. RCV1: A new benchmark collection for text categorization research. The Journal of Machine Learning Research, 2004, 5: 361–397.
- [44] Yang PC, Sun X, Li W, Ma SM, Wu W, Wang HF. SGM: Sequence generation model for multi-label classification. In: Proc. of the 27th Int'l Conf. on Computational Linguistics. Santa Fe: ACL, 2018. 3915–3926.
- [45] Zhang ML, Zhou ZH. A review on multi-label learning algorithms. IEEE Trans. on Knowledge and Data Engineering, 2014, 26(8): 1819–1837. [doi: [10.1109/TKDE.2013.39](https://doi.org/10.1109/TKDE.2013.39)]
- [46] Yang PC, Luo FL, Ma SM, Lin JY, Sun X. A deep reinforced sequence-to-set model for multi-label classification. In: Proc. of the 57th Annual Meeting of the Association for Computational Linguistics. Florence: ACL, 2019. 5252–5258. [doi: [10.18653/v1/P19-1518](https://doi.org/10.18653/v1/P19-1518)]
- [47] Boutell MR, Luo JB, Shen XP, Brown CM. Learning multi-label scene classification. Pattern Recognition, 2004, 37(9): 1757–1771. [doi: [10.1016/j.patcog.2004.03.009](https://doi.org/10.1016/j.patcog.2004.03.009)]
- [48] Chen GB, Ye DH, Xing ZC, Chen JS, Cambria E. Ensemble application of convolutional and recurrent neural networks for multi-label text categorization. In: Proc. of the 2017 Int'l Joint Conf. on Neural Networks (IJCNN). Anchorage: IEEE, 2017. 2377–2383. [doi: [10.1109/IJCNN.2017.7966144](https://doi.org/10.1109/IJCNN.2017.7966144)]
- [49] Wang R, Ridley R, Su XA, Qu WG, Dai XY. A novel reasoning mechanism for multi-label text classification. Information Processing & Management, 2021, 58(2): 102441. [doi: [10.1016/j.ipm.2020.102441](https://doi.org/10.1016/j.ipm.2020.102441)]
- [50] Zhang XM, Zhang QW, Yan Z, Liu RF, Cao YB. Enhancing label correlation feedback in multi-label text classification via multi-task learning. In: Proc. of the 2021 Findings of the Association for Computational Linguistics (ACL-IJCNLP 2021). ACL, 2021. 1190–1200. [doi: [10.18653/v1/2021.findings-acl.101](https://doi.org/10.18653/v1/2021.findings-acl.101)]
- [51] Yan YY, Liu F, Zhuang XQ, Ju J. An R-Transformer_BiLSTM model based on attention for multi-label text classification. Neural Processing Letters, 2023, 55(2): 1293–1316. [doi: [10.1007/s11063-022-10938-y](https://doi.org/10.1007/s11063-022-10938-y)]
- [52] Kingma DP, Ba J. Adam: A method for stochastic optimization. In: Proc. of the 3rd Int'l Conf. on Learning Representations. San Diego, 2015. 6.

附中文参考文献:

- [13] 杜晓宇, 陈正, 项欣光. 基于解耦图神经网络的可解释标签感知推荐算法. 软件学报, 2023, 34(12): 5670–5685. <http://www.jos.org.cn/1000-9825/6754.htm> [doi: [10.13328/j.cnki.jos.006754](https://doi.org/10.13328/j.cnki.jos.006754)]
- [19] 李芳芳, 苏朴真, 段俊文, 张师超, 毛星亮. 多粒度信息关系增强的多标签文本分类. 软件学报, 2023, 34(12): 5686–5703. <http://www.jos.org.cn/1000-9825/6802.htm> [doi: [10.13328/j.cnki.jos.006802](https://doi.org/10.13328/j.cnki.jos.006802)]
- [26] 肖琳, 陈博理, 黄鑫, 刘华锋, 景丽萍, 于剑. 基于标签语义注意力的多标签文本分类. 软件学报, 2020, 31(4): 1079–1089. <http://www.jos.org.cn/1000-9825/5923.htm> [doi: [10.13328/j.cnki.jos.005923](https://doi.org/10.13328/j.cnki.jos.005923)]
- [31] 李博涵, 向宇轩, 封顶, 何志超, 吴佳骏, 戴天伦, 李静. 融合知识感知与双重注意力的短文本分类模型. 软件学报, 2022, 33(10): 3565–3581. <http://www.jos.org.cn/1000-9825/6630.htm> [doi: [10.13328/j.cnki.jos.006630](https://doi.org/10.13328/j.cnki.jos.006630)]



黄靖(1999—), 男, 硕士生, 主要研究领域为推荐系统.



杜晓宇(1986—), 男, 博士, 副教授, CCF 专业会员, 主要研究领域为推荐系统, 多媒体计算, 大模型应用.



陶竹林(1985—), 男, 博士, 副教授, 主要研究领域为认知传播, 推荐系统, 影像增强.



项欣光(1982—), 男, 博士, 教授, CCF 专业会员, 主要研究领域为计算机视觉, 媒体内容分析与理解.

www.jos.org.cn

www.jos.org.cn