计算机科学 COMPUTER SCIENCE
http://www.jsjkx.com
DOI:10.11896/jsjkx.190300137

基于深度学习的多标签生成研究进展

刘晓玲　刘柏嵩　王洋洋　唐　浩
宁波大学信息科学与工程学院　浙江 宁波 315211
(liu_xiaol@163.com)

摘　要　大数据时代,数据呈现维度高、数据量大和增长快等特点。如何有效利用其中蕴含的有价值信息,以实现数据的智能化处理,已成为当前理论和应用的研究热点。针对现实普遍存在的多义性对象,数据多标签被提出并被广泛应用于数据智能化组织。近年来,深度学习在数据特征提取方面呈现出高速、高精度等优异性,使基于深度学习的多标签生成得到广泛关注。文中分五大类别总结了最新研究成果,并进一步从数据、关系类型、应用场景、适应性及实验性能方面对其进行对比和分析,最后探讨了多标签生成面临的挑战和未来的研究方向。

关键词:深度学习;多标签学习;标签相关性;神经网络
中图法分类号　TP391

Research and Development of Multi-label Generation Based on Deep Learning

LIU Xiao-ling,LIU Bai-song,WANG Yang-yang and TANG Hao
Faculty of Information Science and Technology,Ningbo University,Ningbo,Zhejiang 315211,China

Abstract　In the era of big data,data show the characteristics of high dimension,large amount and rapid growth. Efficiently discovering knowledge from these data is a research focus. Multi-label has been proposed for ambiguous objects in reality,and is widely used in data intelligent processing. In recent years,Multi-label generation receives widespread attention due to the excellent performance of deep learning. The latest research results were summarized from five categories and were further compared and analyzed from the aspects of data,relationship types,application scene,adaptability and experimental performance. Finally,the challenges of multi-label generation were discussed,followed with the prospects for future work.

Keywords　Deep learning,Multi-label learning,Label correlation,Neural networks

1　引言

随着在线信息数量、复杂性和动态性的不断增长,多标签学习已成为组织信息过载、数据挖掘的有效关键方法。通过标签学习可以对未知实例进行正确预测或检测,以达到标签生成的目的[1]。若实例具有唯一语义,则其具有唯一标签,若互斥标签总数大于2,即为多分类问题。多标签学习与之不同,允许实例同时与多个标签关联,且标签之间并不完全互斥。因此,多标签学习的目标是从一组实例中学习,每个实例同时属于一个或多个标签[2]。实际应用中存在许多多标签对象,例如,一篇政府改革的新闻可以涵盖多个主题,如政治、经济和社会[3];自然图像可以描绘多类型景物,如天空、沙滩、海洋和游艇[4]。越来越多的现代化应用需要多标签学习的成果,如信息检索[5]、生物信息学[6]、多媒体内容注释[7-8]和语义挖掘[9]。

多标签学习的形式化定义即 $X=R^d$ 表示 d 维的实例特征空间,$Y=\{y_1,y_2,\cdots,y_q\}$ 为由 q 个标签组成的标签空间,多标签训练的实例集合为 $D=\{x_i,Y_i\mid 1\leqslant i\leqslant m\}$,通过学习一个从实例空间映射到标签空间的模式 $h:X\rightarrow 2^Y$ 实现多标签生成。在每个实例 (x_i,Y_i) 中,$x_i\in X$ 是 d 维的特征向量,$Y_i\subseteq Y$ 是实例 x_i 的标签集合,学习到的模式 $h(\cdot)$ 能为未知实例生成标签[10]。

多标签学习是一项具有挑战性的研究[11]。传统方法主要有问题转换法和适应法。问题转换法是将多标签转换为单标签分类和回归问题,容易忽略或不够重视标签之间的联系。适应法改进了单标签算法,使其适应多标签数据,包括集成或链式的方式[12],易受先验知识影响,效果不稳定。传统方式独立的处理方法忽略了标签关系。但实际中,若存在标签"游乐园"和"米老鼠",很可能出现标签"迪士尼";类似地,图像出现"蓝天"和"白云",则出现"雾"的概率会降低。因此,标签数量增加时,标签组合呈指数级增长,若标签空间的标签数量为20,将导致超过 2^{20} 个可能的标签集。此外,传统方法还面临数据高维、不平衡以及标签噪音等挑战[13-14]。

许多学者致力于利用标签的相关性完成标签学习。基于考查方式的不同,我们将关系类型分为3种,如表1所列。有学者将具有标签关系的领域知识作为多标签学习的输入来利

到稿日期:2019-03-26　返修日期:2019-08-01　本文已加入开放科学计划(OSID),请扫描上方二维码获取补充信息。
基金项目:国家社会科学基金(15FTQ002)
This work was supported by the National Social Science Foundation of China(15FTQ002).
通信作者:刘柏嵩(lbs@nbu.edu.cn)

用领域知识的标签层次结构，也有学者试图在多标签学习过程中发现并利用标签关系[15]。

表1 标签关系

Table 1 Label correlation

关系	特点
一阶	假设各标签间相互独立，忽略标签间的关系，简单高效； 完全忽略标签之间的相关性，泛化能力较低；
二阶	考虑两两标签间的关系，如标签排序关系、交互关系 标签关系超越二阶时，算法性能受到很大影响
高阶	处理标签对其他标签或一组标签的标签相关性

近年来，深度学习技术取得突破性进展，利用深度学习技术将数据编码成更高维的数据进行标签学习能表现出较高的性能[16]，这种学习方式从丰富的信息中捕获数据内在的错综复杂关系，有利于标签特征和关系的建模。基于深度学习的多标签生成方法因克服了传统模型的缺点而获得高度关注。本文对这些研究成果进行了梳理、组织和探讨，整体架构如图1所示。

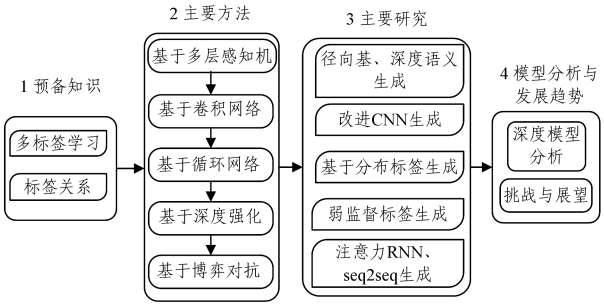


图1 整体框架

Fig. 1 Overall framework

2 基于深度学习的多标签生成

基于深度学习的多标签生成过程通常将各类实例数据和标签关系数据作为输入，利用深度模型学习隐表示，并在此基础上产生标签序列，其基本框架如图2所示。

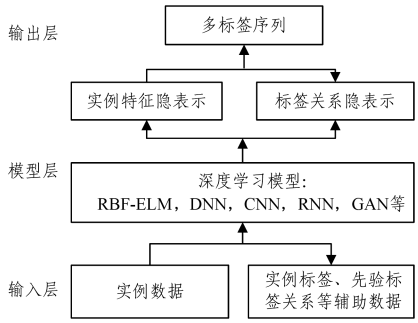


图2 基于深度学习的多标签生成的基本流程

Fig. 2 Flow of multi-label generation based on deep learning

根据多标签生成网络结构的不同，本文把研究方法分为5个主要类别：基于多层感知机、基于卷积神经网络(Convolutional Neural Network，CNN)、循环神经网络(Recurrent Neural Network，RNN)、基于深度强化和基于博弈对抗。2014年，基于径向基神经网络进行多标签学习的方法首次被提出，之后出现了不同的改进多标签方法。卷积神经网络的池化操作能降低网络复杂度，为准确建模数据特征和标签关系，基于卷积网络的多标签方法不断涌现。鉴于普通卷积网络建模序列数据时无能为力，而循环神经网络能够建模序列中不同时刻数据之间的依赖关系，研究者们提出了基于RNN建模标签序列关系的模型。受机器翻译等领域的启发，研究者又提出了一种序列标签生成方法。基于深度强化的方法在解决序列标签顺序问题方面占优势，解决了上一阶段面临的问题。近几年，生成对抗网络(Generative Adversarial Networks，GAN)发展迅速，研究者从数据多样性及鲁棒性出发，提出了博弈对抗的多标签研究方法。

2.1 基于多层感知机的方法

2.1.1 径向基模型

考虑到多个样本标签之间的关联性，研究者将AP-RBF(Affinity Propagation-Radial Basis Function)[17]与ML-RBF[18]融合并进行改进，将体现标签关联性的计数向量与优化后的基函数中心线性叠乘进行多标签学习，图3展示了其流程。在输入层输入样本特征，采用AP聚类确定隐藏层节点数目，然后将蕴含标签关联的计数向量加权到RBF网络基函数中心进行训练。学习到隐含层与输出层之间的权重后，输出层产生标签向量。但此模型良好的结果是以牺牲计算效率为代价的。之后，Zhang等[19]将具有快速学习能力的极限学习机(ELM)[20]融入到径向基网络，提出了ML-ELM-RBF，其在雅虎网页标签化上取得了显著效果。也有研究者对ELM进行改进，提出了RBF与正则化极限学习机(RELM)结合的多标签学习模型AP-RBF-RELM[21]，即在RBF的隐层嵌入RELM。与上述模型不同，该模型中的RELM平衡了训练过程中的结构风险和经验风险，确保了模型具有良好的全局能力和泛化能力。

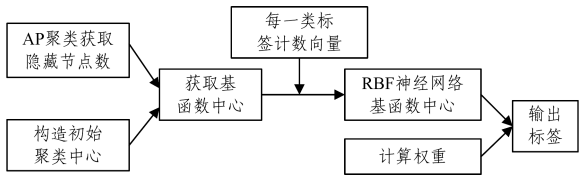


图3 AP-RBF模型流程图

Fig. 3 Flow chart of AP-RBF model

基于RBF的学习方式是根据权值将输入向量映射到输出空间，在输入空间中，该模型改进的重点是构造以多标签数据为基础的高斯函数形式的基函数中心。但此类模型在处理大规模样本集时欠缺语义考虑，深度语义模型应运而生。

2.1.2 深度语义模型

为更好地利用标签高阶依赖性和语义信息，C2AE(Canonical Correlated AutoEncoder)[22]在深度网络(DNN)[23]的基础上进行扩展，以实现特征感知和标签嵌入。其基本思想是：设置两类映射，将实例特征和标签信息映射到同一个隐空间，然后引入敏感损失函数进行反馈学习，进而输出关联的标签。该模型架构如图4所示。在训练阶段，输入实例 X 及其标签 Y，采用深度典型相关分析和自编码器得到特征映射 F_x 和标签编码 F_e 的联合嵌入隐空间；然后解码函数 F_d，学习标签关系 Y。C2AE的优化目标为：

$$\theta=\min_{F_x,F_e,F_d}\Phi(F_x,F_e)+\alpha\Gamma(F_e,F_d) \tag{1}$$

其中，$\Phi(F_x,F_e)$ 和 $\Gamma(F_e,F_d)$ 分别表示隐空间和输出的损失，参数 α 发挥平衡作用。学习流程为：输入测试实例 $\hat{x}$，首先采

用 F_x 将其变换到隐空间，再由 F_d 解码，最后输出标签 $\hat{y}$（即 $\hat{y}=F_d(F_x(\hat{x}))$）。C2AE 联合特征和标签进行学习，其采用的损失函数可适应标签缺失时的情境。DDPP（Deep Determinantal Point Process）[24] 考虑到计算效率问题，结合 DPP 和 DNN 进行端到端的多标签学习和深层表示学习，该模型中的低秩核学习算法提升了计算效率。

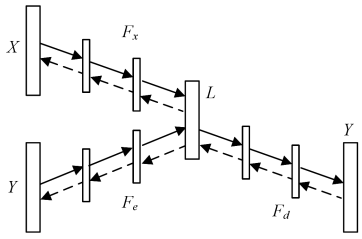


图 4 C2AE 模型架构

Fig. 4 Model architecture of C2AE

2.2 基于卷积神经网络的方法

CNN 已成为当前数据理解领域的研究热点。其由输入层、卷积层、池化层、全连接层和输出层组成，通过将数据直接输入网络避免了复杂的特征提取和数据重建过程。

2.2.1 基于改进 CNN 的标签生成

Kurata 等[25] 采用融入标签关系神经元的 CNN 生成多标签，其依据标签超集思想[26] 在隐藏层设置标签共现模式的专用神经元来模拟标签关系，并根据共现强度初始化神经元权重。模型训练完成时依据标签输出概率得到标签序列。该模型易操作，但忽略了标签频繁项和稀疏项的影响，对未出现的标签组合的预测能力有限。XML-CNN[27] 基于此网络结构进行改进，利用动态池化提取特征，同时又考虑其位置，并在池化层和输出层增加隐藏单元，建模复杂组合关系，进而实现 wiki 数据的标签化。CNN-RNN[28] 采用提取全局与局部语义信息的方式建模高阶关系，首先采用 CNN 提取全局固定长度的特征向量作为 RNN 的输入，再在 RNN 中学习标签隐关系，最后根据首尾特殊标注生成不同长度的标签序列。类似的模型有 Att-CNN-RNN[29]，其采用通道注意力[30] 关注数据最相关特征并生成了天气图像的多标签。此方法提取的特征准确，但受训练集大小的影响，数据量少时容易出现过拟合问题。

实际场景中，数据标签的完备性不能得到有效保证，因此 Jing 等[31] 提出一种基于标签依赖关系型的算法 NN_AD_Omega。该算法通过在神经网络顶层加入关系矩阵来刻画标签之间的依赖关系，在实例部分的标签缺失时，根据学习到的标签依赖关系提升标签效果。但随着标签缺失率的增大，该算法的性能逐渐降低。

2.2.2 分布式多标签生成

上述基于改进 CNN 的多标签生成方法采用数据标签的逻辑表示；而分布式多标签[32] 涵盖一定数量的标签，代表每个标签描述实例的程度，其与逻辑表示的标签的区别在于标签增强过程，即将训练集中的逻辑标签转化为标签分布。标签增强形式化表示为：将原 $\varphi=\{0,1\}^p$ 转为 $D=[0,1]^p$ 即增强，其中 φ 为逻辑标签空间，D 为增强后的标签空间，p 是标签数。DLDL（Deep Label Distribution Learning）[33] 是一种分布式多标签学习模型，该模型首先用标签增强将逻辑表示转为分布表示，然后根据相似实例具有的强关联标签，在深度网络中最小化预测标签和真实标签间的 KL 散度，进而在特征学习和分类器学习中生成标签。将 N 个实例表示为 $D=\{(X^1,y^1),\cdots,(X^N,y^N)\}$，用 $\hat{y}_i$ 表示标签向量，则该模型的目标是从实例中学习条件概率质量函数，找到使 y 和 $\hat{y}$ 分布相似的 θ 并最小化 T：

$$T=-\sum_k y_k \ln \hat{y}_k \tag{2}$$

$$\theta=\arg\min_\theta \sum_k y_k \ln \frac{y_k}{\hat{y}_k} \tag{3}$$

$$\hat{y}=p(y|x,\theta) \tag{4}$$

其中，$x=\phi(X;\theta)$ 是卷积网络最后全连接层的激活函数。softmax 函数[34] 将激活函数转为概率分布，按照正、负和难以辨别的程度对标签归属度进行定义。该模型利用概率分布产生相似数据，实现了图片的多标签化。

DLDL 降低了收集大量精确训练数据的难度，防止网络在训练集较小时出现过拟合的问题。标签增强利用隐含在数据中的标签相关性有效地获取实例的监督信息，进而通过学习标签分布获得更好的效果。但是目前大多数实例数据仅包含简单逻辑标签而非标签分布。

2.3 基于循环神经网络的方法

2.3.1 弱监督标签学习

弱监督标签学习[35] 即在数据资源仅具有部分标签或有空标签组时的学习形式。传统方法[36] 通过增加标签平滑正则化来处理部分标签，但忽略了无标签数据信息，适应范围较小；还有方法将部分标签学习作为凸二次优化的多分类问题[37]，但需要在训练时呈现要预测的数据。基于图[38] 提出的方法虽在处理未标签数据方面表现突出，但基于图的一些标签传播算法[39-40] 只能对呈现的未标签数据进行预测，对未见实例效果甚微。另一类基于流形正则化的算法[41-42] 虽有归纳学习能力，但对图结构和相关距离测量非常敏感。DSGM（Deep Sequential Generative Model）[43] 考虑了上述因素，集成了一个独特的深度序列生成模型和分类的新型深度序列分类模型，引入顺序网络构建生成模型并观察数据后验分布，进而将其用于描述多标签数据以及后向推理，在序列生成模型的学习及有效的优化下得到变分下界。该方法首先利用缺少部分标签或完全未标签的训练数据信息进行可扩展的变分推理，然后通过分类模型来生成标签。实验证明，DSGM 模型比深度半监督模型 SS-MLC[44] 和部分标签的 NL-MLL[45] 更有效。

2.3.2 基于注意力 RNN 的生成

RNN 是针对序列问题建模而提出的。与 CNN 不同，RNN 能建模标签序列的依赖关系。注意力机制（Attention）可以更多地关注和提供有用信息的关键点，基于注意力的 Att-RNN 通过学习区分特征重要度产生特征隐表示，再由前面生成的标签对当前时刻的影响产生标签关系隐表示，进而生成多标签。这种方法的关键优势在于考虑的子集合比完整标签集小，利于解决复杂输出空间问题。其学习目标是最大化正标签联合概率 P：

$$\begin{aligned} P &= (y_{p1}, y_{p2}, \cdots, y_{pT} \mid x) \\ &= \prod_{i=1}^{T} P(y_{pi} \mid y<pi, x)(T<L) \end{aligned} \tag{5}$$

其中，x 是实例，T 是正标签，L 是标签总数，y 是 L 维标签向量。网络嵌入图层模型[46]实现了音乐风格的标签生成，通过设计两级注意力从音乐数据中收集风格的相关信息，在网络中添加标签图层，最终将两个信息流组合，输出标签。Zhu 等[47]对其进行改进：1）提出了空间正则化网络（SRN），为标签生成注意力图，模拟空间关系，并通过卷积捕获标签的潜在语义关系；2）对整个深度神经网络进行端到端的训练，不需要额外的标签处理工作。由于预定义标签序列对模型性能有明显影响[48]，Chen 等[49]提出了一种 Order-Free 的模型 Att-LSTM，该模型由视觉注意力和置信度排序的 LSTM 组成。注意力专注每个标签相关区域并可检测较小对象，LSTM 学习标签依赖性，因此模型性能通过对组合网络的联合学习得以确保。

预定义标签序列最直接的方法是按频率顺序排列正标签，但降序或升序都会破坏标签相关性；若标签附加信息如标签层次结构可用，则可用此信息确定预序列。提高此类模型效果的关键，是利用数据先验知识进行序列预定义及缓解标签顺序的严格要求。

2.3.3 基于 Seq2Seq 的标签生成

Seq2Seq（Sequence-to-Sequence）是循环神经网络的变种。受机器翻译[50]、摘要生成[51]、风格迁移[52-53]等领域中 Seq2Seq 模型的启发，学者们提出 SGM（Sequence Generation Model）[54]来对学术论文进行主题标注，其框架如图 5 所示。

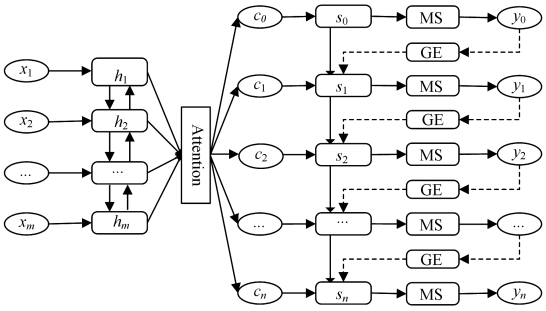


图 5 SGM 模型框架

Fig. 5 Model architecture of SGM

从序列生成角度看，要生成条件概率 $P(y|x)$ 最大化时的标签序列，即：

$$P(y|x)=\prod_{i=1}^{n}P(y_i|y_1,y_2,\cdots,y_{i-1},x) \tag{6}$$

学习流程为：论文 x 在网络编码器中编码为隐藏状态，利用注意力机制在时刻 t 将其聚合为上下文向量 c_t，解码器将上下文向量 c_t、最后的隐藏状态 S_{t-1} 和嵌入向量作为输入，在时刻 t 产生状态 S_t，继而将状态传递到下一时刻，通过遮掩 softmax 层（Masked Softmax，MS）在论文主题标签空间 L 上输出标签概率分布 y。SGM 在解码结构中采用全局标签嵌入（Global Embedding，GE）检查整体标签信息，目的是降低所生成错误标签后传的概率。与无全局嵌入设置相比，全局嵌入使标签生成的汉明损失降低了 7.41%，$F1$ score 提高了 1.04%。

MDC-HA（Multi-level Dilated Convolution-Hybrid Attention）[55]考虑了传统注意机制在文本主题标注应用上的不足，融入了复合注意力[56]。复合注意力是将高级别的段注意力和词注意力相结合来增强特征隐表示，进而提升标签生成的效果。之后，LD-seq2seq[57]针对标签预序列问题，采用新的损失函数代替原始交叉熵损失函数，用弱有序的多标签学习方法在维基网站中根据词语描述自动生成标签。

基于 Seq2Seq 的多标签学习方法能有效建模高阶标签关系，通过注意力机制在预测不同标签时自动选择最有信息量的单词来提升标签效果，但这类方法对标签顺序仍有很强的依赖。

2.4 基于深度强化的方法

深度强化学习探索智能体如何采取行动与环境进行交互以获取最大累积奖励，在自然语言处理领域被广泛应用，如使用策略梯度方法进行机器翻译[58]，用 actor-critic 训练方法生成序列[59]，用强化学习解决零代词解析问题[60]。

在多标签生成过程中，依赖严格标签顺序假设的模型易导致错误的损失计算，进而影响标签结果。当真实标签为 $\{A,B,C\}$ 时，Seq2Seq 仍会因生成标签序列 $\{C,A,B\}$ 而受到极大惩罚，尽管此时标签都正确。Seq2Set（Sequence-to-Set）[61]提供了解决极大惩罚问题的思路，流程如图 6 所示，其包含编码器 E、解码器 D_1 和 D_2。首先输入文本 x；为学习符合人类先验知识的初步标签顺序，利用 E 和 D_1 联合生成隐藏表示 $s=(s_0,s_1,\cdots,s_n)$。D_2 是强化智能体，将 E 和 D_1 的隐藏状态作为输入，其在时刻 t 的状态是当前生成的标签 $(y_0,y_1,\cdots,y_{t-1})$。D_2 的参数 θ 决定执行下一个标签选择的动作，一旦生成完整的标签序列 y，D_2 将观察反馈奖励 R，R 评价生成的标签效果。采用 F1 score 作为奖励函数对生成标签进行检验，其学习目标是最小化负反馈：

$$\begin{aligned}&\min L(\theta)=-E_{y\sim p\theta}[R(y)]\\&R(y)=F1(y,y^*)\end{aligned} \tag{7}$$

其中，y 和 y^* 分别表示生成标签序列和实际标签序列。使用自批判策略梯度算法[62]训练模型。Seq2Set 可融入先验知识，在有结构化和无结构化标签关系时均可使用，其采用反馈机制在一定程度上降低了过度惩罚误差。

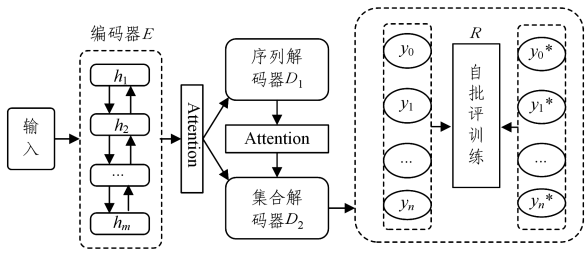


图 6 Seq2Set 流程图

Fig. 6 Flow chart of Seq2Set

RMIC（Reinforced Multi-label Image Classification）[63]从数据特征上考虑了上述误差形成的原因，即图像特征顺序并非固定，易受实体占比和语义联系程度的影响。考虑采用课程学习机制进行多标签生成更符合实际情景，智能体根据特征内容和标签关系学习生成策略，将特征和历史标签作为当前状态，下一个标签作为动作，利用系统向用户推荐标签并收集用户反馈作为奖励。RARL（Recurrent Attentional Rein-

forcement Learning)[64]同时考虑了图像局部区域易被忽略和过度简化的问题,采用循环注意力多标签学习方式,明确模拟局域间的标签关系并进行图像标记。与其余方法相比,RARL 探索了语义标签和注意区域之间的相互作用,同时明确了捕获区域间的依赖关系。

基于深度强化的多标签生成设计了反馈机制来反馈标记效果,减弱了标签顺序交换的敏感性,提高了准确度,但智能体与环境的实时交互将耗费内存和计算力。

2.5 基于博弈对抗的方法

GAN 是 Goodfellow 等于 2014 年提出的深度学习网络。该网络由生成器 G 和鉴别器 D 构成,G 试图产生更接近真实的数据,D 试图更完美地分辨真实数据与生成数据,两个网络在对抗中学习生成理想的输出。

WGAN-gp[65]最早利用对抗进行图像多标签生成。G 是带激活函数的分类器,将图像 x 作为输入并得到标签的生成概率分布 $\hat{y}$。D 包含特征抽取器,将标签集 y(或 $\hat{y}$)和 x 作为输入,产生分数值 $D(y,x)$(或 $D(\hat{y},x)$)来衡量标签和图像的匹配程度,即标签的真实度。G 在最小化标签损失的同时最大化 D 的分数值,D 需要尽力区分真标签和假标签。G 和 D 的学习函数为:

$$\begin{aligned} G: L_G = & -E_{x\sim data,\hat{y}\sim G(x)}[D(\hat{y},x)]+\alpha[E_{(x,y)\sim data} \\ & [\sum_{i=1}^{|S|} y_i \log \hat{y}_i + (1-y_i)\log(1-\hat{y}_i)]] \\ D: L_D = & -E_{(x,y)\sim data}[D(y,x)]+1/2E_{x\sim data,\hat{y}\sim G(x)}[D(\hat{y},x)]+ \\ & \frac{1}{2}E_{x\sim data,y'\sim data}[D(y',x)]+\lambda L_{gp} \end{aligned} \tag{8}$$

其中,L_G 的第一项为 G 的生成标签获得的匹配分数,第二项表示损失函数;L_D 的第一项为 D 要最大化的图像和真实标签集的匹配分数,第二、三项表示 D 要最小化的生成标签及错误匹配的标签分数,第四项是保持训练稳定的梯度惩罚函数。

基于博弈对抗思想,对抗深度网络 I-FT-ASDN[66]在上述基础上侧重考虑多标签实例数据的多样性及鲁棒性。其采用 SPP[67]对不同尺寸的图像进行特征提取,并将单标签训练参数迁移到网络中,同时使用对抗空间损失网络(ASDN)生成遮挡图,这一方面解决了训练数据不足的问题,另一方面增强了对未见实例的标注能力。此模型在生成复杂图的多标签上具有一定优势。Behpour 等[68]考虑到模型存在铰链损失宽松的问题,从博弈论的视角为损失函数提供界限,提出了对抗性鲁棒切割模型(ARC),将多标签生成任务视为基于最小成本图切割的 G 和 D 之间的零和博弈。该模型在半监督学习下取得了较好的成效。

基于博弈对抗进行多标签学习能同时利用标记和无标记数据,同时通过对抗学习增强生成器的能力,但该类模型的稳定性有待进一步提升。

3 深度模型分析

基于深度学习生成多标签的模型采用端到端的学习方式,但不同模型各有侧重。深度学习技术的进一步发展使得多标签学习模型具有下述主要优势:1)避免设计复杂算法来提取特征,尤其是面对图像这类非结构化数据,模型(如 Att-RNN,RARL)能够有效地从非结构化数据中获得特征信息;2)学习非线性多层次的抽象特征,通常是稠密和低维度的(如 SGM,MDC-HA 等)特征,这是传统浅层模型所不具备的;3)部分模型可以对数据原标签质量进行检验,避免了噪音标签的影响(如 DLDL)并缓解了严格的标签序列要求(Seq2Set,Att-LSTM)。

基于不同深度学习技术的多标签模型在数据类型、学习方式、学习效果等方面存在的差异,表 2 总结了多标签学习的 4 种设置方式;表 3 归纳了主流模型的各项特点,并列举了相关应用。为增强不同方法的可比性,将各方法实验结果按数据集进行划分,详情如表 4 所列。由表 4 可见,同一模型在不同领域表现出差异性,大多数模型的效果与领域相关,其移植与扩展性须进一步改善。同时,对比结果说明标签空间较大时,模型学习效果降低;原始数据语义较丰富时,查全率与准确度有所提高。结合表 3 可知,监督学习下模型的准确度较高,但需要大量标记语料,比半监督学习方式更费时,且无法捕捉无标注数据的信息。弱监督学习方式较符合实际情形,适应性较强,但适应复杂情况时的准确度有待提升。

表 2 不同设置的多标签学习方式

Table 2 Multi-label learning with different settings

设置	全标记数据	部分标数据	无标记数据
①监督型	√	×	×
②半监督	√	×	√
③标签缺失	√	√	×
④弱监督	√	√	√

表 3 不同标签学习模型的特点

Table 3 Characteristics of different label learning models

基于深度模型	模型	数据类型	应用	关系	特点
多层感知机	AP-RBF ML-ELM-RBF AP-RBF-RELM	文本	基因分析 网页/图书标注 音乐风格识别	二阶	模型易构建,计算效率低,适应①②设置,但不足以处理大规模数据
	C2AE	文本/图像	主题标注	高阶	此模型关联标签跨域数据,适应设置③
卷积神经网络	CNN-RNN NN-AD-Omega	文本/图像	新闻/图片标注	二阶	有效挖掘数据特征,分别适应①③,③设置下的准确度有待提高
	DLDL	标签分布式图像	年龄、头部姿态估计及语义分割	高阶	适应领域广,缓解数据量约束,需要分布式表示数据
循环神经网络	Att-RNN SRN Att-LSTM	图像	天气标注 图像分类	高阶	依赖预定义标签顺序,适应设置①
	DSGM	图像	图像标注	高阶	适应设置④复杂情况,但评价方式单一
	SGM MDC-HA LDseq2seq	文本	主题标注 义原提取	高阶	自动选择信息量特征考虑高层语义,但需先验知识确定最优标签顺序
深度强化	Seq2Set RMIC RARL	文本 图像 图像	主题标注 图像识别	高阶	设计反馈机制降低惩罚误差,增强了标签效果,但耗内存与计算力
博弈对抗	WGAN-gp I-FT-ASDN ARC	图像 文本 图像	半监督分类	高阶	适应设置②,对未见实例学习标签能力强,但稳定性有待提高

表 4 多标签模型的实验结果对比

Table 4 Experimental results comparison of multi-label learning methods

数据集	模型	实验结果				
		HL(−)	*O*(−)	*R*(+)	*F*1(+)	*P*(+)
Yeast 基因数据[17]	ML-ELM-RBF	0.1978	0.2372	—	—	0.75
	AP-RBF-RELM	0.1888	0.2280	—	—	0.76
	ARC	0.1860	—	—	—	—
Emotion 数据[70]	AP-RBF	0.5389	0.2833	—	—	0.79
	ARC	0.1860	—	—	—	—
Corel5K[68]	I-FT-ASDN	—	—	—	0.415	0.45
Reuters (RCV1-V2)[28]	CNN-RNN	0.0085	0.0830	0.825	0.855	0.88
	SGM	0.0075	—	0.860	0.878	0.89
	MD-CHA	0.0072	—	0.873	0.882	0.89
	Seq2Set	0.0070	—	0.884	0.887	0.89
CAL500[31]	NN_AD_Omega	—	0.0960	—	0.469	0.14
	ARC	0.102	—	—	—	—
天气数据[49]	Att-LSTM	—	—	0.870	0.870	0.87
Arxiv 论文[54]	SGM	0.0245	—	0.675	0.710	0.74
	CNN-RNN	0.0278	—	0.618	0.664	0.71
	Seq2Set	0.0249	—	0.659	0.698	0.74
MS-COCO 图像[65]	SRN	—	—	0.715	0.734	0.75
	CNN-RNN	—	—	0.664	0.692	0.69
	Att-LSTM	—	—	0.622	0.677	0.74
	RARL	—	—	0.572	0.662	0.78
	WGAN-gp	—	—	0.642	0.686	0.74
	DSGM	—	—	—	0.55	—
Ren-CECps 博客[55]	CNN-RNN	0.1876	—	0.538	0.556	0.57
	Att-LSTM	0.1793	—	0.573	0.581	0.59
	MD-CHA	0.1782	—	0.585	0.590	0.62
wiki[58]	LDseq2seq	—	—	0.419	0.448	0.48
PASCAL 图像[22]	C2AE	—	—	0.691	0.676	0.66
	RMIC	—	—	—	—	0.84
	RARL	—	—	—	—	0.85
	I-FT-ASDN	—	—	—	—	0.83

注：HL 表示 hamming loss，O 表示 one-error，R 表示 recall，F_1 表示 micro-F_1，P 表示 precision；“+”表示越大越好，“−”表示越小越好

4 面临的挑战与展望

基于深度学习的多标签生成已取得一系列研究成果，但仍存在以下挑战性问题有待进一步研究。

(1)大规模多标签学习问题。当数据特征空间和标签空间非常大时，标签子集呈指数级增长，计算复杂度和内存消耗较大[69]。如何克服数据采样、标签关系建模、子集选择等问题是难点。

(2)多标签噪音问题。实际情况中，贴标者忽略了部分不感兴趣的标签，因此难以获取较准确和完整的标签。

(3)标签分布问题。多标签学习中，现有方法一般采用均匀标签分布假设，即假设各个实例的相关标签的重要度相等[70]。但一个实例与多个标签相关时，其重要性一般不同。利用标签增强描述多标签程度是亟需解决的问题，目前面临的挑战有：建立适用标签增强实验的标准数据集，提出性能评价体系和新的标签增强算法。

(4)流式多标签学习问题。动态环境中，学习系统必须能够重用学习过的模型并使其适应不断变化的环境。已有的数据流处理方法仅关注数据流本身的动态变化，缺少新标签检测流程。为了在这种动态环境中保持良好的学习性能，学习模型有必要同时具有检测和分类新标签实例的功能。

针对上述挑战，未来可从以下 4 个方面展开多标签学习生成的研究。

(1)分布式深度学习的多标签方法

将集中式计算模式转变为分布式深度学习模式，实现计算并行化。基于以 MapReduce 为代表的大数据分布式计算平台，结合现有的分布式计算框架，设计以高效能为目标的分布式深度学习的多标签系统架构、计算框架和处理方法，值得探索。

(2)部分多标签学习方法

部分多标签学习需要在实例仅有部分标签正确的严苛条件下进行。设计同时利用全局和局部标签的相关性机制是处理全标签和丢失标签的一种策略。设计能充分利用多视图数据进行语义感知，以及用深度模型扩展多视角多标签模型，将是研究的趋势。

(3)标签增强的多标签方法

标签增强进一步将标签逻辑表示转成分布表示。目前，多标签数据集多是逻辑表示，少数实验借助人造数据进行初步实验，因此标签增强实验数据集应包括真实标签分布以及对应的逻辑标签；同时，要建立能反映从逻辑标签恢复到标签分布能力的评价体系；提出能充分利用实例空间相关性的标签增强算法。标签相关性既能从标签空间获取，也能从实例空间迁移得到，在理论和应用层面都有很大的研究空间。

(4)动态多标签学习方法

多标签学习环境中，当观察到新的实例时，系统必须能够修改预先训练的模型，并针对所有新出现的标签建立识别策略。增量学习中，历史模型可通过新增实例更新得到新模型来关注数据本身的动态。在此基础上，设计基于输入特性和标签属性的特殊检测器及模型更新机制，将检测的新标签合并，并设计具有容忍检测错误能力的学习模型来适应动态场景，值得进一步研究和关注。

结束语 随着数据规模、复杂性的不断增加，高效挖掘有价值的信息以实现数据智能化处理面临挑战，基于深度学习的多标签生成给数据知识智能化组织带来了新机遇。

参 考 文 献

[1] ALAZAIDAH R, THABTAH F, AL-RADAIDEH Q. A Multi-label Classification Approach Based on Correlations Among Labels[J]. International Journal of Advanced Computer Science and Applications, 2015, 6(2): 52-59.

[2] CORANI G, SCANAGATTA M. Air Pollution Prediction via Multi-label Classification[J]. Environmental Modelling & Software, 2016, 80(6): 259-264.

[3] ZHANG M L, WU L. Lift: Multi-label learning with label-specific features[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2015, 37(1): 107-120.

[4] SONG Z, QIU Q. Learn to Classify and Count: A Unified Framework for Object Classification and Counting[C]// Proceedings of the 2018 International Conference on Image and Graphics Processing. ACM, 2018: 110-114.

[5] HUANG C Q, YANG S M, PAN Y, et al. Object-location-aware hashing for multi-label image retrieval via automatic mask learning[J]. IEEE Transactions on Image Processing, 2018, 27(9): 4490-4502.

[6] LI Z, LIAO B, LI Y, et al. Gene Function Prediction Based on Combining Gene Ontology Hierarchy with Multi-instance Multi-label Learning[J]. Rsc Advances, 2018, 8(50): 28503-28509.

[7] CAKIR E, HEITTOLA T, HUTTUNEN H, et al. Polyphonic sound event detection using multi label deep neural networks [C]// International Joint Conference on Neural Networks. IEEE, 2015: 1-7.

[8] ZHANG C. Multi-label Text Categorization of Zhi Hu Title Based on Deep Learning[D]. Beijing: Beijing Jiaotong University, 2018.

[9] SUN J J, PEI L, JIANG T. Research on Semantic Annotation in Academic Literature[J]. Journal of the China Society for Scientific and Technical Information, 2018, 37(11): 1077-1086.

[10] ZHANG M L, ZHOU Z H. A review on multi-label learning algorithms[J]. IEEE transactions on knowledge and data engineering, 2014, 26(8): 1819-1837.

[11] ZHANG J J. Research on Key Problems of Multi-label Learning [D]. Xi'an: Xidian University, 2016.

[12] SPOLAÔR N, MONARD M C, TSOUMAKAS G, et al. A systematic review of multi-label feature selection and a new method based on label construction [J]. Neurocomputing, 2016 180(C): 3-15.

[13] WANG S B, LI Y F. Classifier Circle Method for Multi-Label Learning[J]. Journal of Software, 2015, 26(11): 2811-2819.

[14] PENG Y. Research on multi-label learning problems based on topic model[D]. Nanjing: Nanjing University, 2018.

[15] FU D, ZHOU B, HU J. Improving SVM based multi-label classification by using label relationship[C]// 2015 International Joint Conference on Neural Networks (IJCNN). IEEE, 2015: 1-6.

[16] YAN Y. Text Representation and Classification with Deep Learning[D]. Beijing: University of Science & Technology Beijing, 2016.

[17] LI S L, LIU R, LIU H. Multi-label Learning for Improved RBF Neural Networks[J]. Computer Science, 2015, 42(4): 316-320.

[18] ABDECHIRI M, FAEZ K. Efficacy of utilizing a hybrid algorithmic method in enhancing the functionality of multi-instance multi-label radial basis function neural networks [J]. Applied Soft Computing, 2015, 34(C): 788-798.

[19] ZHANG N, DING S, ZHANG J. Multi-Layer Elmrbf for Multi-label Learning[J]. Applied Soft Computing, 2016, 43(C): 535-545.

[20] LI G, NIU P, DUAN X, et al. Fast learning network: a novel artificial neural network with a fast learning speed [J]. Neural Computing and Applications, 2014, 24(7/8): 1683-1695.

[21] DANG D, XU X Z. Multi-label learning Model Based on Radial Basis Function Neural Network and Regularized Extreme Learning Machine[J]. Pattern Recognition and Artificial Intelligence, 2017, 30(9): 67-74.

[22] Y YEH C K, WU W C, KO W J, et al. Learning deep latent space for multi-label classification[C]// Thirty-First AAAI Conference on Artificial Intelligence. AAAI, 2017: 2838-2844.

[23] NAM J, KIM J, MENCiA E L, et al. Large-scale multi-label text classification revisiting neural networks[C]// Machine Learning and Knowledge Discovery in Databases—European Conference. Springer, 2014: 437-452.

[24] XIE P, SALAKHUTDINOV R, MOU L, et al. Deep determinantal point process for large-scale multi-label classification[C]// Proceedings of the IEEE International Conference on Computer Vision. IEEE Computer Society, 2017: 473-482.

[25] KURATA G, XIANG B, ZHOU B. Improved Neural Network-based Multi-label Classification with Better Initialization Leveraging Label Co-occurrence [C]// Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics. ACL, 2016: 521-526.

[26] HÜLLERMEIER E, CHENG W. Superset Learning Based on Generalized Loss Minimization[C]// Joint European Conference on Machine Learning and Knowledge Discovery in Databases. Springer, 2015: 260-275.

[27] LIU J, CHANG W C, WU Y, et al. Deep learning for extreme multi-label text classification[C]// Proceedings of the 40th International ACM SIGIR Conference on Research and Development in Information Retrieval. ACM, 2017: 115-124.

[28] CHEN G, YE D, XING Z, et al. Ensemble application of convolutional and recurrent neural networks for multi-label text categorization[C]// 2017 International Joint Conference on Neural Networks. IEEE, 2017: 2377-2383.

[29] ZHAO B, LI X, LU X, et al. A CNN-RNN architecture for multi-label weather recognition[J]. Neurocomputing, 2018(12), 322: 47-57.

[30] YUAN Y, XUN G, MA F, et al. A novel channel-aware attention framework for multi-channel EEG seizure detection via multi-view deep learning[C]// 2018 IEEE EMBS International Conference on Biomedical & Health Informatics. IEEE, 2018: 206-209.

[31] SONG P, JING L P. Exploiting Label Relationships in Multi-Label Classification with Neural Networks[J]. Journal of Computer Research and Development, 2018, 55(8): 1751-1759.

[32] GENG X, XU N. Label distribution learning and label enhancement[J]. Scientia Sinica Informationis, 2018, 48(5): 521-530.

[33] GAO B B, XING C, XIE C W, et al. Deep label distribution learning with label ambiguity[J]. IEEE Transactions on Image Processing, 2017, 26(6): 2825-2838.

[34] LIU W, WEN Y, YU Z, et al. Large-Margin Softmax Loss for Convolutional Neural Networks [C]// Proceedings of the 33nd International Conference on Machine Learning. JMLR, 2016: 507-516.

[35] DURAND T, THOME N, CORD M. Weldon: Weakly supervised learning of deep convolutional neural networks[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. IEEE Computer Society, 2016: 4743-4752.

[36] WU B, LYU S, GHANEM B. Constrained submodular minimization for missing labels and class imbalance in multi-label learning[C]// Thirtieth AAAI Conference on Artificial Intelligence. AAAI, 2016: 2229-2236.

[37] WU B, JIA F, LIU W, et al. Multi-label learning with missing labels using mixed dependency graphs[J]. International Journal of Computer Vision, 2018, 126(8): 875-896.

[38] WANG M, FU W, HAO S, et al. Scalable semi-supervised learning by efficient anchor graph regu-larization[J]. IEEE Transactions on Knowledge and Data Engineering, 2016, 28(7): 1864-1877.

[39] JING L, YANG L, YU J, et al. Semi-supervised low-rank mapping learning for multi-label classification[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. IEEE Computer Society, 2015: 1483-1491.

[40] LIN G, LIAO K, SUN B, et al. Dynamic Graph Fusion Label Propagation for Semi-supervised Multi-modality Classification [J]. Pattern Recognition, 2017, 68(8): 14-23.

[41] WU F, WANG Z, ZHANG Z, et al. Weakly semi-supervised deep learning for multi-label image annotation[J]. IEEE Transa-

ctions on Big Data,2015,1(3):109-122.
[42] PAPANIKOLAOU Y,TSOUMAKAS G,KATAK-IS I. Hierarchical partitioning of the output space in multi-label data[J]. Data & Knowledge Engineering,2018,116(7):42-60.
[43] CHU H M,YEH C K,WANG Y C. Deep Generative Models for Weakly-Supervised Multi-Label Classification[C]// Proceedings of the European Conference on Computer Vision (ECCV). Springer,2018:400-415.
[44] WEI Y,LIANG X,CHEN Y,et al. Stc:A simple to complex framework for weakly-supervised semantic segmentation[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence,2017,39(11):2314-2320.
[45] LI Y,YANG J,SONG Y,et al. Learning from noisy labels with distillation[C]// Proceedings of the IEEE International Conference on Computer Vision. IEEE Computer Society,2017:1910-1918.
[46] ZHAO G,XU J,ZENG Q,et al. Driven Multi-label Music Style Classification by Exploiting Style Correlations[J]. arXiv:1808.07604,2018.
[47] ZHU F,LI H,OUYANG W,et al. Learning spatial regularization with image-level supervisions for multi-label image classification[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. IEEE Computer Society,2017:5513-5522.
[48] VINYALS O,BENGIO S,KUDLUR M. Order matters:Sequence to sequence for sets [J]. arXiv:1511.06391,2015.
[49] CHEN S F,CHEN Y C,YEH C K,et al. Order-free rnn with visual attention for multi-label classification[C]// Thirty-Second AAAI Conference on Artificial Intelligence. AAAI,2018:6714-6721.
[50] WEI J,PHAM K,O'CONNOR B,et al. Evaluating Grammaticality in Seq2seq Models with a Broad Coverage HPSG Grammar:A Case Study on Machine Translation[C]// Conference on Empirical Methods in Natural Language Processing. ACL,2018:298-305.
[51] HU X,LI G,XIA X,et al. Deep code comment generation[C]// Proceedings of the 26th Conference on Program Comprehension. ACM,2018:200-210.
[52] FU Z,TAN X,PENG N,et al. Style transfer in text:Exploration and evaluation[C]// Thirty-Second AAAI Conference on Artificial Intelligence. AAAI,2018:663-670.
[53] AKAMA R,INADA K,INOUE N,et al. Generating stylistically consistent dialog responses with transfer learning[C]// Proceedings of the Eighth International Joint Conference on Natural Language Processing. 2017:408-412.
[54] YANG P,SUN X,LI W. SGM:sequence generation model for multi-label classification[C]// Proceedings of the 27th International Conference on Computational Linguistics. 2018:3915-3926.
[55] LIN J,SU Q,YANG P,et al. Semantic-Unit-Based Dilated Convolution for Multi-Label Text Classification[C]// Proceedings of the 2018 Conference on Empirical Methods in Natural Language Procesing. ACL,2018:4554-4564.
[56] GEHRING J,AULI M,GRANGIER D,et al. Convolutional sequence to sequence learning[C]// Proceedings of the 34th International Conference on Machine Learning-Volume 70. JMLR,2017:1243-1252.
[57] LI W,REN X,DAI D,et al. Sememe prediction:Learning semantic knowledge from unstructured textual wiki descriptions [J]. arxiv:1808.05437,2018.
[58] HE D,XIA Y,QIN T,et al. Dual learning for machine translation[C]// Advances in Neural Information Processing Systems. MIT Press,2016:820-828.
[59] BAHDANAU D,BRAKEL P,XU K,et al. An actor-critic algorithm for sequence prediction[J]. arXiv:1607.07086,2016.
[60] YIN Q,ZHANG Y,ZHANG W,et al. Deep reinforcement learning for chinese zero pronoun resolution[C]// Annual Meeting of the Association for Computational Linguistics. ACL,2018:569-578.
[61] YANG P,MA S,ZHANG Y,et al. A Deep Reinforced Sequence-to-Set Model for Multi-Label Text Classification[J]. arXiv:1809.03118,2018.
[62] ZHAO L,ZHANG C,ZHANG X,et al. A Deep Reinforced Training Method for Location-Based Image Captioning[C]// Pacific Rim International Conference on Artificial Intelligence. Springer,2018:878-890.
[63] HE S,XU C,GUO T,et al. Reinforced Multi-Label Image Classification by Exploring Curriculum [C]// Thirty-Second AAAI Conference on Artificial Intelligence. AAAI,2018:3183-3190.
[64] CHEN T,WANG Z,LI G,et al. Recurrent attentional reinforcement learning for multi-label image recognition[C]// Thirty-Second AAAI Conference on Artificial Intelligence. AAAI,2018:6730-6737.
[65] TSAI C P,LEE H Y. Adversarial Learning of Label Dependency:A Novel Framework for Multi-class Classification[J]. arXiv:1811.04 689,2018.
[66] ZHOU T,LI Z,ZHANG C,et al. An Improved Convolutional Neural Network Model with Adversarial Net for Multi-label Image Classification[C]// Pacific Rim International Conference on Artificial Intelligence. Springer,2018:38-46.
[67] MEN K,BOIMEL P,JANOPAUL NAYLOR J,et al. Cascaded atrous convolution and spatial pyramid pooling for more accurate tumor target segmentation for rectal cancer radiotherapy[J]. Physics in Medicine & Biology,2018,63(18):185016.
[68] BEHPOUR S. Arc:Adversarial robust cuts for semi-supervised and multi-label classification[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. IEEE,2018:1905-1907.
[69] ZHOU Z H. Learn ware:on the future of machine learning [J]. Frontiers of Computer Science,2016,10(4):589-590.
[70] JIANG S,XU Y,WANG T. Multi-Label Metric Transfer Learning Jointly Considering Instance Space and Label Space Distribution Divergence[J]. IEEE Access,2019,7:10362-10373.



LIU Xiao-ling, born in 1994, postgraduate. Her main research interests include natural language processing and data mining.



LIU Bai-song, born in 1971, Ph.D, researcher, Ph.D supervisor, is member of China Computer Federation. His main research interests include natural language processing, big data and artificial intelligence.