



计算机科学

COMPUTER SCIENCE

监督和半监督学习下的多标签分类综述

武红鑫, 韩萌, 陈志强, 张喜龙, 李慕航

引用本文

武红鑫, 韩萌, 陈志强, 张喜龙, 李慕航. [监督和半监督学习下的多标签分类综述](#)[J]. 计算机科学, 2022, 49(8): 12-25.

WU Hong-xin, HAN Meng, CHEN Zhi-qiang, ZHANG Xi-long, LI Mu-hang. [Survey of Multi-label Classification Based on Supervised and Semi-supervised Learning](#)[J]. Computer Science, 2022, 49(8): 12-25.

相似文章推荐 (请使用火狐或 IE 浏览器查看文章)

Similar articles recommended (Please use Firefox or IE to view the article)

[基于图卷积神经网络的文本分类方法研究综述](#)

Review of Text Classification Methods Based on Graph Convolutional Network

计算机科学, 2022, 49(8): 205-216. <https://doi.org/10.11896/jsjcx.210800064>

[基于双图神经网络信息融合的文本分类方法](#)

Text Classification Method Based on Information Fusion of Dual-graph Neural Network

计算机科学, 2022, 49(8): 230-236. <https://doi.org/10.11896/jsjcx.210600042>

[面向文本分类的类别区分式通用对抗攻击方法](#)

Class Discriminative Universal Adversarial Attack for Text Classification

计算机科学, 2022, 49(8): 323-329. <https://doi.org/10.11896/jsjcx.220200077>

[一种面向电商网络的异常用户检测方法](#)

Method for Abnormal Users Detection Oriented to E-commerce Network

计算机科学, 2022, 49(7): 170-178. <https://doi.org/10.11896/jsjcx.210600092>

[基于注意力机制和多任务学习的阿尔茨海默症分类](#)

Alzheimer's Disease Classification Method Based on Attention Mechanism and Multi-task Learning

计算机科学, 2022, 49(6A): 60-65. <https://doi.org/10.11896/jsjcx.201200072>

监督和半监督学习下的多标签分类综述

武红鑫 韩萌 陈志强 张喜龙 李慕航

北方民族大学计算机科学与工程学院 银川 750021

(893319518@qq.com)

摘要 传统的多标签分类算法大多数采用监督学习的方式,但现实生活中有许多数据没有被标记。通过人工的方式对需要的全部数据进行标记耗费的成本较高。半监督学习算法可以使用大量未标记数据和标记数据来进行工作,因此受到了人们的重视。文中首次从监督和半监督学习的角度对多标签分类算法进行阐述,同时全面地对多标签分类算法的应用领域进行了总结。从决策树、贝叶斯、支持向量机、神经网络和集成等多个方向对标签非相关性和标签相关性的监督学习算法进行概述,从批处理和在线的方向对半监督学习算法进行综述,从图像分类、文本分类和其他等角度对多标签的实际应用领域进行介绍。文中还简要分析了多标签的评估指标,最后给出了关于半监督学习下的复杂概念漂移处理、特征选择处理、标签复杂相关性处理和类不平衡处理的研究方向。

关键词 监督学习;半监督学习;多标签分类;图像分类;文本分类

中图分类号 TP391

Survey of Multi-label Classification Based on Supervised and Semi-supervised Learning

WU Hong-xin, HAN Meng, CHEN Zhi-qiang, ZHANG Xi-long and LI Mu-hang

School of Computer Science and Engineering, North Minzu University, Yinchuan 750021, China

Abstract Most of the traditional multi-label classification algorithms use supervised learning, but in real life, there are many unlabeled data. Manual tagging of all required data is costly. Semi-supervised learning algorithms can work with a large amount of unlabeled data and labeled data, so they have received more attention from people. For the first time, multi-label classification algorithms are explained from the perspective of supervised learning and semi-supervised learning, and application fields of multi-label classification algorithms are comprehensively summarized. Among them, supervised learning algorithms of label non-correlation and label correlation are described in terms of decision trees, Bayesian, support vector machines, neural networks, and ensemble, semi-supervised learning algorithms are summarized from the perspectives of batch and online learning. The real-world application areas are introduced from the perspectives of image classification, text classification and other fields. Secondly, this paper briefly introduces evaluation metrics of multi-label. Finally, research directions of complex concept drift under semi-supervised learning, feature selection, complex correlation of labels and class imbalance are given.

Keywords Supervised learning, Semi-supervised learning, Multi-label classification, Image classification, Text classification

随着大数据技术的快速发展,生活中产生了大量的数据,这些数据中包含着与人类生活密切相关的许多信息,为了从中获得需要的数据,学者开展了许多与数据挖掘有关的研究^[1]。传统的分类方法侧重于对单标签进行分类,然而,许多现实问题却需要使用多标签分类(Multi-label Classification, MLC)^[2]来解决。

给定一个 d 维输入空间 $\mathbf{X} = X_1 \times \cdots \times X_d$ 和一个输出标签 $Y = \{\lambda_1, \lambda_2, \cdots, \lambda_q\}$, $q > 1$ 。多标签实例可以定义为一对 (x, y) , 其中 $x = (x_1, \cdots, x_d) \in \mathbf{X}$ 并且 $y \subseteq Y$, 其中 y 被称为标签集。当标签 j 和实例 x 相关时, λ_j 等于 1, 否则等于 0。MLC

的目标是构建一个预测模型 $h: \mathbf{X} \rightarrow 2^Y$, 为未知实例提供一组相关标签。每个实例可能有几个与其相关的标签,这些标签来自先前定义的标签集。因此,对于每个 $x \in \mathbf{X}$, 有标签空间 Y 的二分集合 (y, \bar{y}) , 其中 y 是相关标签的集合, \bar{y} 是不相关标签的集合。

MLC 主要可以应用于文本分类^[3]、医学诊断分类^[4]、蛋白质分类^[5]、音乐^[6]或视频分类^[7]等。例如,在文本分类中,关于神州十二号发射的报道可以归类为社会类别,也可以归类为科技类别。类似地,在医学诊断分类中,一位病人可以同时患有糖尿病和高血压疾病。

到稿日期:2021-07-11 返修日期:2021-12-10

基金项目:国家自然科学基金(62062004);宁夏自然科学基金(2020AAC03216)

This work was supported by the National Natural Science Foundation of China(62062004) and Ningxia Natural Science Foundation Project(2020AAC03216).

通信作者:韩萌(2003051@nmu.edu.cn)

经典的 MLC 方法主要分为问题转化(Problem Transformation, PT)和算法自适应(Algorithm Adaptation, AA)。PT 是将多标签问题转化成多个简单的单标签分类问题,其中最常用的是二元相关(Binary Relevance, BR)方法。BR 方法将多标签分类问题转化为 $|L|$ 个不同的二元单标签分类问题,转化后,选择任何现有的二元单标签分类算法作为基分类器。BR 方法未考虑标签之间的相关性,为了解决这一问题,研究人员提出了分类器链方法(Classifier Chains, CC)^[8],它是在 BR 算法的基础上,通过链的方法连接由 BR 获得的二元分类器。标签幂集(Label Power-Set, LP)方法也是一种 PT 方法,LP 是将多标签问题转化为具有 $2^{|L|}$ 个类标签的多类问题。随机子标签集成算法(Random K-Labelsets, RAKEL)^[9] 是对 LP 方法的一种集成使用,其中每个 LP 基分类器都是由随机生成的且互不相同的小规模标签子集训练而成。AA 是修改现有的算法以适应需要解决的新问题,具体表现为将现有的单标签分类问题调整为 MLC 问题。构建 AA 的流行模型包括

k 近邻^[10]、决策树^[11]、支持向量机^[12]、神经网络^[13]等。

近年来,已有一些关于 MLC 的综述。Zhang 等^[14]介绍了多标签学习的基础知识,对 8 种经典的算法进行了分析和讨论,总结了多标签学习的在线资源等。Tsoumakas 等^[15]从 PT 和 AA 的角度对多标签进行分类的方法进行了详细的介绍,简单介绍了一些评估指标,最后比较了 MLC 方法的实验结果。Moyano 等^[16]对 20 个数据集上的多标签集成分类算法进行了比较,根据数据不平衡、标签间相关性的特征来评价它们的性能。Zhou 等^[17]针对评估措施和标签相关性问题对多标签学习进行了讨论,同时讲述了 4 种具有代表性的多标签算法的基本思想和技术细节。Zheng 等^[18]从多标签数据流分类的角度介绍了传统的 MLC 方法并讨论了它们的优缺点,确定了多标签数据流分类的挖掘约束。至今为止,还没有综述从监督学习和半监督学习的角度对 MLC 进行介绍,同时也没有综述对多标签的实际应用进行全面的介绍。本文的总体框架如图 1 所示。

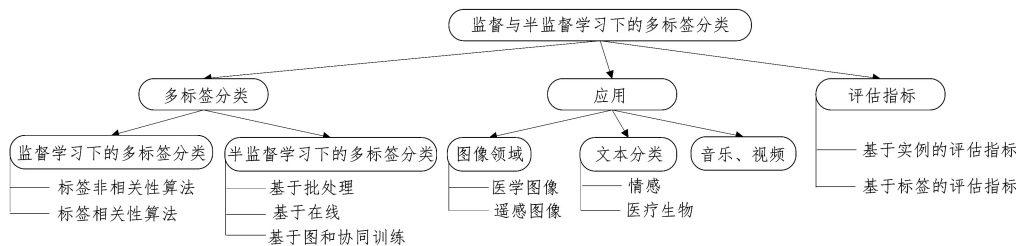


Fig. 1 Overall framework diagram of multi-label classification

本文的主要贡献有:

(1)首次从监督学习和半监督学习两个方面对 MLC 算法进行了综述,角度更加新颖、全面。

(2)首次从实际应用领域对 MLC 算法进行了全面的综述,总结了近年来应用领域的先进算法。

(3)深入分析了 MLC 算法中存在的问题,并提出了未来的研究方向。

1 多标签分类

监督学习算法和半监督学习算法在 MLC 中都得到了广泛的应用,本节将从监督学习和半监督学习的角度对其进行综述。

1.1 监督学习

监督学习是从已经被标记的训练集中推断一个功能的机器学习任务^[19]。下文将从标签非相关下的监督学习和标签相关下的监督学习两个方面对 MLC 算法进行综述。

1.1.1 标签非相关性算法

在 MLC 中,标签之间可能存在的联系使多标签的相关性变得非常复杂^[20],若不考虑该问题,可以降低算法的难度。基于标签非相关性,本节从决策树、贝叶斯、支持向量机、神经网络和集成等多个方向对 MLC 的算法进行综述。

预测概率的多标签决策树^[21]在监督学习的情况下利用归一化方法将多标签数据转换成单个标签的实例,使用传统、

单标签的决策树算法建立一棵树。该算法基于树的复杂性和预测准确性引入了新的度量方法。基于 NPI-M 的 ML-DT^[22]是一种基于多项数据的非参数预测推理模型,对噪声的鲁棒性很高。随着数据中引入更多噪声,该算法的分类优势更为明显。监督学习下的 LdSM^[23]可以用于构建和训练多标签决策树,在树的每个节点中优化一个新的目标函数,该函数有利于平衡分裂,保持子节点的高类纯度,并允许向多个方向发送实例,同时有一个防止树过度增长的惩罚。一旦前一个节点完成,就对树的每个节点进行训练。

Yang 等^[24]以贝叶斯网络为基础模型,提出了一种特征加权方法来提高决策函数的分类精度。该方法通过计算训练数据中特征的深度频率来估计正类的条件概率,对于概率与决策函数无关的变量,可以通过消除多余的变量来简化决策函数。

AEML-LLSVM^[25]是一种基于多核低秩线性化 SVM 的快速分类算法,该算法首先利用 BR 转换策略将多标签数据集分解为多个二元数据集,然后使用近似极值法从每个二元数据集中获取代表集,然后在每个代表集上进行训练。AEDC-MLSVM^[26]是一种监督学习下的将近似极值法和分治策略相结合的一种算法,是上一种算法的改进方法。它采用改进的分治策略将代表集划分为子集,这样可以确保每个代表子集都包含一定数量的正负实例,避免奇异问题并解决负载不平衡问题。Bellmann 等^[27]使用分类器链的集成处理

模糊的 MLC,该方法使用分类器链对 8 种基本图书类型进行分类,并将嵌入的单词和作者信息作为特征。对于每个分类器,使用具有径向核函数的支持向量机以及与作者信息相结合的全局矢量表示。

在监督学习下, BP-AEPML^[28]采用近似极限点提取具有代表性的数据集来缩小原始数据集的大小,首先使用 BR 转换策略将原始的多标签数据集转换为多个单独的二元子数据集,同时使用近似极限值提取每个二元子数据集的代表数据集,最后使用 BP 神经网络训练该代表数据集。基于双向关联池化层的深度神经网络结构的算法^[29]采用关联函数来检测聚集到合并神经元中的不同神经元对,同时提出了一种迭代过程,该过程可以估计较深层中合并的神经元之间的关联度,而无须重新计算相关矩阵。关于遥感的一种新型的多注意驱动系统^[30]使用一种新的 k 分支 CNN 架构来提取与不同空间分辨率相关的遥感图像波段的初步局部描述符。在 RNN 中,它所有的输出都被用来预测遥感图像的多标签,而不是通过考虑单一的特定类别节点来确定每个标签。

基于监督学习的 ELIFT^[31]算法使用集成方法来减轻对高分类精度的限制,使用装袋策略生成的多个训练集构造多个 LIFT 分类器,根据每个分类器的损失,自动对不同的分类器进行加权。对于每个新实例,通过学习的加权集成分类器来获得预测标签。

1.1.2 标签相关性算法

随着数据集中类标签数量的增加,需要预测的标签集数量会呈指数增长,考虑标签相关性可以促进学习过程^[17]。本节从决策树、贝叶斯、支持向量机、神经网络和集成等多个方面对多标签的标签相关性算法进行介绍。

为了解决标签之间的相关性,在集成分类器链(Ensemble Algorithm Based on Classifier Chains, ECC)的基础上对其进行扩展,将 Credal C4.5^[32]作为基分类器的集成模型在嘈杂的数据中有更好的性能,同时也可以很好地处理标签之间的相关性。Credal C4.5 使用的分裂准则是不精确信息增益比。为了对新实例进行分类,对于每个标签,对每个链预测的与该实例相关的后验概率进行平均,根据 ECC 给出标签与实例相关性的后验概率。最后的预测标签集由集成预测的后验概率大于 0.5 的实例组成。学习标签相关性可能会产生循环依赖性,为了解决这个问题,提出了 3RC 算法。3RC^[33]是一种监督学习下的 MLC 方法,这种新方法通过使用多个决策树作为二元分类器来遵循 BR 方法。这种新颖的方法旨在学习标签的依赖关系,并给出仅考虑相关依赖关系的模型的结果,以便更好地进行预测并减少由于无关和弱依赖关系而导致的错误传播。

监督学习下的 MLNB-LD^[34]提出了具有强独立假设的贝叶斯定理,是针对多标签问题的一种新的后验概率估计方法。所提方法利用标签对之间的相关性来确定给定未见实例的最可能标签集。监督学习下的 TreeStacking^[35]提出了一种针对树型依赖表示模型的贝叶斯网络算法。该算法确定了每个类标签可能依赖的其他若干标签以及确定每个类标签的

预测顺序;在确定依赖关系的过程中,减小了无关属性对当前标签分类效果及性能的影响。BCC^[36]考虑到特征选择在分类任务中的重要性,通过改进每个内部分类器来提升分类的性能,提出了两种算法来测试 BCC,即 BF-FS-BCC 和 GS-FS-BCC。鉴于 BCC 的结构和链序,对于每个标签,选择一个子集然后建立分类器。

在监督学习下, RBRL^[37]是一种将排序支持向量机、BR 与鲁棒低秩学习相结合的算法,该算法捕捉输入和输出之间的非线性关系,同时采用两种加速近端梯度方法来有效地解决快速收敛的优化问题。基于监督学习的 SSSVM^[38]是一种超高分辨率遥感图像的新型 MLC 方法,其基本思想是通过结构化 SVM、标签间的关系以及在成本函数中添加一个鼓励空间平滑的术语,来将空间上下文信息纳入结构 SVM 优化过程中。

LCL-Net 算法^[39]为了对 ChestX-ray14 数据集中病理之间的相关性进行建模,将多层感知器引入 LCL 模块中。多层感知器是一种通用函数逼近器,可以在训练阶段自适应地重新校准多标签输出,并提高 LCL-Net 的性能。同时, LCL 模块可以轻松插入任何基于 CNN 的模型的末尾。SDLML^[40]模型在图像特征学习部分使用 VGG 的 CNN 学习全脑 CT 图像。在监督学习下,可变长度切片之间的切片相关性以及多种疾病之间的因果关系可从 RNN 获得。基于监督学习的 DCNet^[41]模型主要由 3 个模块组成,一个用于特征提取的主干 CNN,一个用于特征相关的空间相关模块,以及一个用于生成分类分数的分类器。其中主干 CNN 分别从左侧和右侧的彩色眼底照片中提取两组特征。

基于监督学习的 ML-CIB^[42]是一种集成 MLC 方法,通过学习一个新的标签矩阵并捕获新的标签相关性,提出了一种标签正则化方法来处理新标签中不平衡的多标签问题,并在目标函数中引入正则化范数来选择相关的稀疏特征。ACkEL^[43]是一个随机 K 标签集集成,它借用主动学习的思想,提出了标签选择标准,以评估从标签子集转换的类的可分离性和平衡水平。基于互信息和联合熵的 K -标签的集成方法^[44]评估每个 K 标签集的冗余度和不平衡度。该算法迭代执行不相交采样,将互信息量较少的多个 K 标签集保留为候选项,并选择具有最高联合熵的 K 标签集。

1.1.3 小结

本节从标签非相关性下的算法和标签相关性下的算法对监督学习的 MLC 算法进行了总结,主要从时间复杂度和数据集的角度对其进行了综述。

(1) 时间复杂度

在多标签算法中, BR 方法由于实现简单、运行速度快而在许多领域被广泛应用,但它会忽略标签之间的关系,将每个标签单独处理,丢失了许多的信息。标签相关下的算法比标签非相关下的算法可以获得更好的准确度。对准确性有较高要求的应用可以通过解决相关性问题来达成目的。除了相关性问题会影响分类问题,标签类不平衡也会对结果产生影响。

总体来说,相比决策树、贝叶斯和集成算法,支持向量机

由于时间复杂度问题限制了其在大规模数据集中的应用。但本节的 RBRL^[37]和 AEDC-MLSVM^[26]算法可以有效解决这个问题并处理类不平衡问题。神经网络也存在训练时间太长的问題,原因是 BP 神经网络算法的学习率是固定的,为了使输出向量尽可能接近期望值,在数据训练过程中需要反复调整网络的权重和偏差,在这个过程中,训练数据集的规模越大,需要的调整时间就越长,尤其是在隐藏层较多的情况下。采用基于近似极限点提取具有代表性的数据集可以缩小原始数据集的大小,从而缩短在数据训练上花费的时间。

标准 SVM 分类算法的训练时间复杂度为 $O(N^3)$ 。获得代表性集合时,AEML-LLSVM^[25]分类算法的时间复杂度为 $O(mN)$,同时 AEML-LLSVM 分类算法采用 AELLSVM 处理每个二进制数据子集,训练时间复杂度为 $O(mMk^2)$ 。AEDC-MLSVM^[26]获得代表集的时间复杂度为 $O(mN)$,AEDC-MLSVM 分类算法的时间复杂度至少为 $O(mM^2/\omega)$ 。在神经网络中,BP-AEPML^[28]算法将代表性数据集的大小缩小到 M ,同时 $M < N$,它的时间复杂度是 $O(huM)$,其中 $O(huM) < O(huN)$, N 为训练数据集的大小, k 表示地标数据实例的数量, m 是标签的数量, M 是代表集中数据实例的数量, ω 代表聚类中心, h 代表隐藏层的数量, u 是每个隐藏层的单元个数。

(2)数据集

对于 Mediamill 数据集,实例数有 43 907 个,其中训练数据集有 30 993 个,测试数据集有 12 914 个,标签数为 101,AEDC-MLSVM 算法^[26]相比其他算法有很好的测量结果,但比 3RC 算法^[33]略差。3RC 算法在其上可以获得很好的平均精确度,达到 97.09%,是算法 AEDC-MLSVM 的 1.5 倍,汉明损失达到 2.08%,是算法 AEDC-MLSVM 的 2.5 倍。但 3RC 在其他数据集(如 yeast, scene, emotions)上的测量结果并不是很好。相比之下,它更适合分类实例数较多的数据集。

对于 TMC2007-500 数据集,实例数为 28 596 个,其中训练数据集有 21 519 个,测试数据集有 7 077 个,标签数为 22,AEDC-MLSVM 算法的训练时间为 896.1s,而当 BP-AEPML 算法的隐藏层为 5 时,训练时间为 387.2s,比 AEDC-MLSVM 算法节约了 508.9s。但 BP-AEPML 的平均精确度只有 79.56%,比 AEDC-MLSVM 算法降低了 18.87%。因此,如果追求时间效率且对准确性要求不是特别高时,可以选择 BP-AEPML 算法。

为了方便地分析算法的性能和优缺点,表 1 从分类方法、实验数据集、测试领域、优缺点方面对所提出的监督学习算法进行了总结。

表 1 多标签分类监督学习算法

Table 1 Multi-label classification supervised learning algorithm

算法	分类方法	实验数据集	测试领域	优点	缺点
ML-DT ^[22]	决策树	Bibtex; birds; cal500; corel5k; emotions; enron; flags; genbase; mediamill; medical; scene; yeast	文本; 音频; 音乐; 图像; 生物; 医疗	基于非参数推理模型的多标签决策树方法对噪声具有很好的鲁棒性	没有将非参数预测模型应用于 MLC 的决策树集成中
LdSM ^[23]	决策树	Bibtex; Mediamill; Delicious; AmazonCat-13k; Wiki10-31k 等	文本; 视频	该算法有着高预测精度和低预测时间	没有在更多领域测试该算法的性能
3RC ^[33]	决策树	Scene; Yeast; emotions; Mediamill	图像; 生物; 音乐; 文本	该算法引入了依赖性权重来确定标签之间的相关性,然后仅保留标签之间重要的依赖关系	没有使用更多具有相关标签的数据集进行实验,没有与更具有竞争力的算法进行性能的比较
MLNB-LD ^[34]	贝叶斯	Arts; Education; Entertain; Health; Recreation; Reference; Science; Social 等	文本; 图像; 音乐	所提方法利用标签对之间的相关性来确定给定未见实例的最可能标签集	该算法的计算效率较低,没有去除不必要的噪声信息
TreeStacking ^[35]	贝叶斯	Birds; CAL500; emotions; flags; genbase; medical; yeast; enron; bibtex	音频; 音乐; 图像; 生物; 医疗; 文本	该算法只建立一棵树,有效地降低了算法的复杂度	只是简单地选取互信息值最小的节点,没有研究更好的确定树根节点的方法
BCC ^[36]	贝叶斯	Music; Scene; Slashdot; Yeast; 20NG; Enron; LangLog; Medical; Ohsumed; IMDB; Bibtex	音乐; 图像; 文本; 生物	该算法使用特征选择技术改进了单个内部分类器,使结果性能提高	需要考虑找到特征子集所需的时间,尽管这段时间只影响模型的构建阶段
AEML-LLSVM ^[25]	SVM	Mediamil; Tmc2007-500; Nuc-wide	音频; 文本	在保持性能优势的前提下大大缩短了训练和测试时间	未考虑标签之间的相关性
AEDC-MLSVM ^[26]	SVM	TMC2007-500; Mediamill (exp1); EukaryoteGO	文本; 音频; 生物	该算法的训练时间和测试时间大大缩短	未使用标签之间的相关信息来提升算法的分类性能
RBRL ^[37]	SVM	emotions; image; scene; yeast; enron; arts; education; recreation 等	音乐; 图像; 生物; 文本	提出的 RBRL 方法可以同时减少汉明损失和排序损失	未考虑在类不平衡的情况下算法的性能
SSSVM ^[38]	SVM	UAV 数据集; 机载图像数据集	图像	该方法可以对空间连续性以及输出结构进行建模	因为需要在每个步骤中解决所有样本的线性规划问题,计算成本较高
BP-AEPML ^[28]	神经网络	Mediamill	视频	该算法的训练时间大大缩短	未考虑在多个数据集上进行实验
LCL-Net ^[39]	神经网络	ChestX-ray14	医疗; 图像	CNN 模型可以自适应地更新校准多标签的输出,并提升 LCL-Net 的性能	未在更多的领域对算法进行测试

(续表)

算法	分类方法	实验数据集	测试领域	优点	缺点
SDLM ^[40]	神经网络	CQ500 数据集	医疗; 图像	缩短了 在 标签 图像 上 花费 的 时间	未 在 顺序 图像 上 进行 图像 增强
DCNet ^[41]	神经网络	ODIR2019	图像; 医疗	该 算法 具有 较低 的 计算 复杂性	无法 将 所 提 方法 的 结果 与 进行 图像 级别 分类 的 患者 进行 比较, 未 解决 类 不平衡 的 问题
ELIFT ^[31]	其他	Birds; Enron; Arts; Corel5k; Pascal	音频; 文本; 图像	该 算法 在 高维 数据 集 上 的 效果 更好	未 考虑 标签 之间 的 相关性
ML-CIB ^[42]	其他	Bibtex; Medical; Enron; Genebase; Image; MIMLtext; CAL500; Corel5k 等	音频; 文本; 图像; 音乐; 生物	该 方法 在 常规 规模 和 大规模 的 不平衡 多 标签 数据 上 可以 获得 很好 的 结果	该 算法 对 较 宽 范围 的 参数 不 敏感
ACkEL ^[43]	其他	Birds; CAL500; CHD49; Emotions; Flags; Foodtruck; Genbase; GpositiveGo; Medical; PlantGo 等	音频; 音乐; 文本; 生物; 医疗; 图像; 化学	该 算法 解决 了 转化 后 的 类 在 特征 空间 中 难以 分离 的 问题	未 找到 针对 不同 数据 集 的 最优 K , 后期 可以 尝试 其他 的 基 分类 器, 训练 复杂度 较高

1.2 半监督学习

随着大数据的快速发展,数据信息呈现爆炸式的增长。大多数的分类方法都假设训练数据已经完全被标记,但在现实应用中,获得完全标记的样本既昂贵又费时,并且使用大量未标记和少量标记的样本进行训练是一种实用的方法。半监督技术是处理不完全标记的一种流行的方法。接下来,半监督学习方法将从批处理学习和在线学习等角度进行综述。

1.2.1 批处理

批处理学习指在学习过程开始之前,进行整个训练数据集的收集,算法需要在任何时候都将这个训练集存储在内存中,以计算模型^[45]。

RML-CNN^[46]是一种将标签从标记图像传播到特征空间中未标记的相邻图像的基于卷积神经网络的批处理半监督学习分类算法。为了在开始时初始化未标记数据的分数和标签,该算法仅使用标记数据来训练 CNN 模型。为了与特征提取器联合训练,将提出的分类器集成在 CNN 模型中,并将其作为最终分类阶段的损失函数。ALSM^[47]是自适应低秩半监督学习,它通过低秩矩阵恢复为标记和未标记训练样本学习中间特征空间,并采用自适应半监督学习策略来训练多标签分类器。Liu 等^[48]提出了一种新颖的基于核规范的 MLC 半监督学习框架。该框架采用核规范正则化来实现类级别的平滑度,以在准则函数中自适应地构建类别图。OPF-SEMImst+knn^[49]是使用最佳路径森林框架的分类算法。由于被错误分类的样本通常出现在聚类之间的边界,该方法通过从概率密度函数的最大值重新传播标签来减少训练集中标签传播的误差,对接近其最大值的训练样本给予更高的优先权,并在分类期间将其标签分配给新样本。Abburi 等^[50]提出了一种基于自我训练的半监督技术,该技术利用未标记的样本来扩充标签集,相对于现有的标签集而言,高度文本多样性是候选的无标签实例的理想样例,还引入了缓解 MLC 的类不平衡的方法。

QdS² ML^[51]是一种量化的分布式批处理半监督多标签学习算法。为了利用已标记和未标记数据的信息,使用由核逻辑回归函数建模的信息论方法来设计损失函数。SLRM^[52]使用半监督奇异值分解来应对标签之间的相关性和弥补缺少标签数据甚至丢失标签数据的问题,同时提出了一种交替方向乘子算法来有效地处理大规模数据集。MixLab^[53]是一种基于一致性正则化技术的新颖方法,用于解决半监督 MLC

问题。该算法使用标签相关性来提高对无标记实例的预测,并将其加入到一致性正则化中,使用混合技术在已标记和未标记的实例之间进行一致性预测。

1.2.2 在线

由于批处理学习算法不能从数据流和动态数据中进行学习,研究者们开始提出在线学习的概念。在线学习中,模型会随着观察值的变化而迭代更新。由于一次迭代只需要在内存中进行一次观察,因此该类型的模型可以分析非常庞大的或者动态的数据集^[45]。

OnSeML^[54]是一种基于标签压缩和局部平滑回归的在线半监督多标签学习算法。该算法通过一个固定的正交标签编码器将标签集压缩到一个低维空间中。通过一个封闭形式的解,为每个引入的实例获得一个局部定义的回归函数。RACE^[55]是一种新颖的在线线性标签压缩的半监督方法,对于每一批新数据,该算法将标签空间压缩为较小的随机空间,然后在此压缩标签集上更新单标签在线分类器或回归器。之后使用该批数据更新现有的解压缩函数。

在半监督的情况下,基于 GNG 的在线多标签算法^[45]可以在有限的内存使用情况下动态地提供数据的注释,同时引入在预测一个标签时考虑网络拓扑结构的非局部标记函数,还改进了标签策略对网络拓扑的影响并使用标签来改进突触链接。在线多标签 LVQ 神经网络^[56]作为基分类器的半监督算法是一种自学习框架。对于每个输入数据,计算输入和每个神经元之间的距离。在更新过程中引入了实例的权重矩阵,使得给定的输入对 BMU 类标签的隶属度具有累积影响。

1.2.3 其他

在半监督学习中,除了基于批处理学习和在线学习的 MLC 方法外,基于协同训练的半监督学习方法和基于图的半监督学习方法也受到了广泛研究。

协同训练方法是一种半监督学习技术,需要两个视图的数据。它假设每个实例都使用两个不同的特性集来描述,这些特性集提供了实例的不同且互补的信息^[57]。Coins^[58]采用在归纳设置下自然起作用的协同训练策略来从标记和未标记的数据中学习。在每个协同训练回合中,通过最大限度地提高两个分类器之间的多样性来学习特征空间上的二分类。MLCT^[59]是基于协同训练的一种半监督集成分类算法。该算法使用标签的共现信息来调整预测的可能性并处理类不平衡问题。CMLDSE^[60]是一种基于协同训练的多标签数据流

的半监督集成分类算法。针对多标签数据流中标签丢失的问题,引入了一种新的标签出现检测机制来解决海量和快速数据流中标签丢失的问题。文献[61]提出了一种名为 MNB-VWCO 的算法,该算法采用协同训练方法在特征和标签集上学习两个独立的分类器,其中协同训练算法用于捕获类标签之间的相关关系。

SMILE^[62]是一种使用不完整标签信息的基于图的半监督 MLC 方法,该方法可以根据所采用的邻域图来进一步补充训练实例的缺失标签,同时可以直接预测完全未标记的新实例的标签。Chaudhuri 等^[63]提出了一种半监督图论的方法,将特征空间中标记和未标记数据的结构与训练图像中的区域标记在一起使用。

1.2.4 小结

当训练数据中有很少数据被标记时,因为基于监督学习的算法只考虑标记的数据,使用半监督学习算法可以比监督学习算法获得更好的结果。本节对批处理学习和在线学习的半监督 MLC 算法进行了总结,相比批处理算法,在线算法具有实时性,可以处理不断变化的标签分类。

基于批处理算法 OPFSEMI_{st}+knn^[49]的结果会受到特征标签维度的影响。在大多数情况下,当使用较小的数据集时,使用 LP 的算法超过了相同条件下分析的大多数情况的最佳结果,而使用较大的数据集时,使用 HOMER 的算法可以得到更好的结果。批处理的算法 ALSM^[47]使用自适应学习策略可以隐式地考虑缺失标签问题。当改变丢失标签的百分比时,该方法的性能会有所下降。例如数据集 Yeast,当丢失标签的百分比为 10%时,Micro-F1 的值为 68%左右,而当丢失标签的百分比为 50%时,Micro-F1 的值只有 60%左右。该算法可以通过引入低秩来有效地处理噪音数据。

对于在线算法 RACE^[55],在数据集 CAL500 和 delicious 上可以达到很好的准确度和汉明损失,并且在所有的数据集上,由于其随机压缩的性质,运行时间大大缩短,特别地,在数据集 rev1v2 和 NUS-WIDE 上,该算法的运行时间分别为 18400 s 和 69700 s,而对批算法 PLST 甚至不能在合理的时间内完成。因此,该算法适合大部分的数据集,尤其可以完成一些别的算法无法处理的大规模数据集。在线算法 OnSeML^[54]具有自适应更新策略和调整策略,两种不同的调整策略在不同的数据集上会表现出不同的性能。使用调整策略,OnSeML 在 Enron 和 Corel5k 上获得了更好的预测性能。而在 Mirflickr 和 Mediamill 上,自适应更新策略会产生更好的性能,造成这种区别的原因是,Mirflickr 和 Mediamill 的标签集很小,并且来自不同时间片的实例可能具有相似的标签分布,而调整策略可能会过度拟合最近到达的数据,因此该算法适用于标签集很小的数据集。基于在线的 GNG 算法^[45]在具有不同数量的未标记数据的数据集上具有较低的汉明损失和较高的 F1 测量值。它对数据集中的复杂性的增加不是很敏感,特别地,在具有较高的 TCS 的数据集上有很好的表现,如数据集 Medical 和 TMC2007。

为了便于分析算法的性能和优缺点,从批处理和在线两个方向对上文提到的半监督下的 MLC 算法进行了详细的介绍,如图 2 所示。

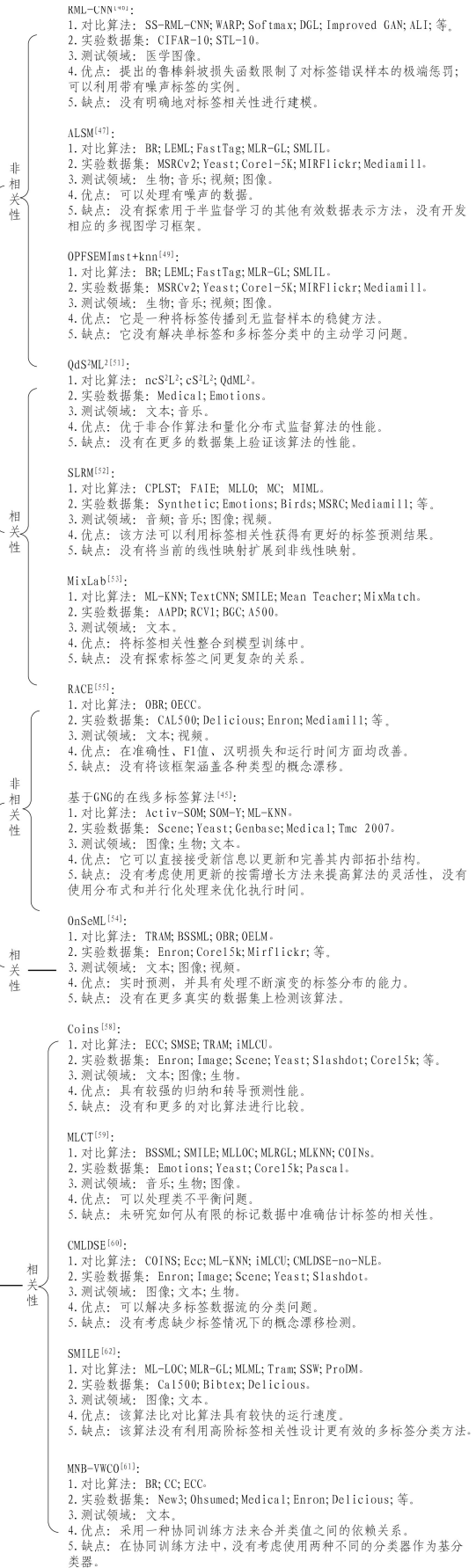


图 2 多标签分类半监督学习算法
Fig. 2 Semi-supervised learning algorithm of multi-label classification

2 应用

MLC 问题因被广泛应用而引起了越来越多研究者的关注^[15]。下文分别从图像分类、文本分类、其他等多个领域对 MLC 算法展开阐述。

2.1 图像领域

图像分类是一项艰巨的任务,近年来引起了学术界的极大关注,它更适合使用 MLC 算法进行分类。大多数图像可以与多个标签关联以描述语义内容,如对象、场景、动作和属性等^[43]。本文主要从基于医学领域和基于遥感图像对图像领域进行介绍。

2.1.1 医学领域

多标签图像分类在医学领域有着广泛的应用,如胸部 X 射线、心电图、脑部 CT、眼部疾病等。通过 MLC 算法,可以很好地弥补医生的不足并一定程度地减轻医生的工作量。

Chen 等^[39]提出了一种用于多标签胸部 X 射线图像分类的新颖标签共现学习模型,该模型利用标签共现和依赖信息来探索图像中潜在的共现标签。Guan 等^[64]为了解决胸部 X 射线图像上进行多标签胸腔疾病分类的问题,提出了 CRAL 模型。它可以预测在一个特定类别的注意力视图中的多种病变,通过赋予相应的特征表示较小的权重来抑制不相关类别的影响。

Cai 等^[65]提出了一种基于心电图数据集的心律失常多标签图像分类方法,可以同时检测 55 种心脏病症状,被称为 Multi-ECGNet。该模型提出端到端深度学习模型的整套心电图监测分析建模方法和研究思路,同时在指标方面优于普通的心脏病专家。Li 等^[40]提出了一种多标签切片依赖学习模型,称为 SDLM。它是一种序列到序列的模型,以端到端的方式有效地学习图像特征和切片依赖性。He 等^[41]提出了考虑双眼相关联的患者水平诊断和多标签疾病分类的模型,该模型包含 3 部分,一个主干 CNN 用于特征提取,一个空间相关模块用于特征关联,一个分类器用于分类评分生成。主干 CNN 分别从左右彩色眼底照片中提取两组特征,空间相关模块捕获两个特征集之间的像素级相关性,然后将处理后的特征进行融合,得到患者级表征,最后的疾病分类是在患者水平上进行的。

Mercan 等^[66]在弱监督的学习场景中利用病理学家的观察记录及其幻灯片级别的注释提出了新的模型。该模型根据不同的行为从病理学家的图像筛选日志中提取候选 ROI;使用由候选 ROI 代表的实例包和从病理形式中提取的一组类别标签对每个幻灯片进行建模。Xu 等^[67]探索了易于访问的标签,以帮助对病变类型进行分类,使用病变类型和患者 ID 标签构建基于 DML 的损失函数,利用五重输入建立利用转移学习的深度模型,最后提出标签选择训练样本的五重挖掘算法。

2.1.2 遥感图像

多标签图像分类在遥感图像的复杂内容方面有着举足轻重的作用,并且在过去几年中引发了一些相关研究^[68]。

Alshehri 等^[69]提出了一种带有通道和空间注意机制的编解码器神经网络架构模型,用于处理基于遥感的无人机

问题。该模型基于预训练 CNN 编码器模块,使用适当的特征组合将输入图像转换为的一组特征图。基于长期短期记忆网络的解码器模块是以顺序方式生成图像中存在的类。Hua 等^[68]提出了一种新颖的遥感技术 MLC 网络,即注意感知标签关系推理网络,该网络由 3 个基本模块组成。特征提取模块负责从原始图像中捕获细粒度的高级特征图;注意区域提取模块生成注意标签特定的功能;标签关系推断模块使用前一个模块的输出中得出的标签关系来预测标签的存在。一种新颖的多注意驱动系统于 2020 年被 Sumbul 等^[30]提出,主要用于处理遥感图像。该系统主要由 4 个模块组成。第一个模块提取与不同空间分辨率关联的遥感图像波段的初步本地描述符;第二个模块通过双向 RNN 架构来实现,其中长短期内存节点通过考虑局部区域的空间关系来丰富局部描述符;第三个模块通过基于补丁的多注意机制来实现,考虑了多个土地覆盖类别共同出现的情况;最后一个模块利用这些描述符进行多标签遥感图像分类。Chaudhuri 等^[63]在多标签遥感图像检索问题的框架内引入了一种半监督图论模型。所提模型基于 4 个主要步骤:第一步是对档案中的每个图像进行分割,并提取每个区域的特征;第二步构建图像邻域图,并使用相关的标签传播算法;第三步通过一种新颖的区域标记策略,将类标记与图像区域相关联;最后一步通过子图匹配策略检索与给定查询图像相似的图像。

Dai 等^[70]提出了一种新颖的 CBIR 模型。该模型考虑了基于尺度不变特征变换的视觉词袋方法,结合空间和频谱描述符,通过基于稀疏重构的新颖遥感图像检索方法实现了图像检索,考虑了一种新的标签似然度量,并将原始的稀疏分类器推广到单标签和多标签遥感图像检索问题,引入了一种策略来利用基于稀疏重构的方法对不同词典单词的敏感性。Koda 等^[38]提出了一种新颖的基于 SVM 的 MLC 方法,以实现遥感图像的准确土地覆盖分类,将增强整个图像的平滑度的模型称为空间结构化 SVM。

2.2 文本分类

文本分类的应用也可以采用 MLC 算法来解决,主要的应用领域有情感分类和医学生物等。

2.2.1 情感

多标签情感分类是文本情感分类的子任务,其目的是识别文本中表达的共存情感(如欢乐、愤怒和焦虑等),由于其广泛的潜力而备受研究者的关注^[3]。

He 等^[3]提出联合二元神经网络(Joint Binary Neural Network, JBNN)来处理情感分类问题。在 JBNN 中,文本的表示由一组 logistic 函数代替 softmax 函数,多个二元分类在一个神经网络框架中同步进行。此外,通过训练联合二进制交叉熵(Joint Binary Cross Entropy, JBCE)损失来获取标签之间的关系,同时将之前的标签关系纳入 JBCE 损失中。Yu 等^[71]提出了一种新的迁移学习架构。该模型使用共享 LSTM 层来提取情感分类任务的共享情绪特征,并使用目标特定的 LSTM 层来提取特定的仅对人们的情感分类任务敏感的情感特征。Fei 等^[72]提出了一个主题增强的 capsule 网络,它能够在没有外部知识的情况下学习潜在主题信息,从而促进多标签情感分类。

Wen 等^[73]深入研究读者情绪分类问题,针对真实的情绪文本复杂多样、具有多视角多标签的特性,提出了一种改进模型 MV-MSI。该模型融合了包括新闻文本和评论在内的多视角信息,对读者情绪进行了更加准确的判断。Alzubi 等^[74]提出了一种解决阿拉伯社交媒体的情感分析模型,使用了多个支持 MLC 的分类器,为了使带注释的数据集更加准确,该模型使用仲裁过程来检查和更新带注释的数据集。Bravo-Marquez 等^[75]提出了带有注释的情感词典模型,该模型将单词级功能和学习技术相组合以最有效地完成注释情感词典的任务,即可以使用未标记的推文从任何特定领域的集合中识别情感词。

Kim 等^[76]提出了一种基于注意力机制的分类器,它可以预测给定句子的多种情绪。该模型由注意力机制和多个独立的 CNN 组成,通过对表情符号的语义预处理和对额外的词汇进行使用,其性能得到进一步提高。Mulki 等^[77]开发了一种 MLC 模型 Tw-StAR,用于识别嵌入在阿拉伯文、英文和西班牙文推文中的情绪。该模型对推文进行了预处理技术的一种或多种组合,采用 BR 转换策略,同时使用 TF-IDF 方案生成推文功能。Alhuzali 等^[78]提出了一种新的模型 SpanEmo,该模型将多标签情感分类作为跨度预测,它可以帮助情感分析模型学习标签和句子中单词之间的关联,同时还引入了一个专注于对输入句子中的多种共存情绪进行建模的损失函数。Hyun 等^[79]提出了一种基于深度学习的模型,该模型将语言嵌入和情感嵌入相结合,用于 CL-AFF 共享任务的文本分类,同时将从嵌入模型中提取的句子特征作为 TextCNN 输入,从而提供给文本进行分类。Ying 等^[80]选择流行的 BERT 语言模型,以提供用于建模句子的一般语言知识,使用特定于 Twitter 的预处理器来解码与 Twitter 相关的表达式,训练令牌模式检测器以识别影响情感的令牌模式,引入了两步训练过程来整合常识和检测到的领域知识以进行情感分类。

2.2.2 医学生物

多标签文本分类在信息检索领域起着重要作用,并对医疗生物领域的信息检索产生了一定影响^[81]。

Du 等^[4]提出了 ML-Net 模型,这是一种新颖的端到端深度学习框架,用于生物医学文本的 MLC。该模型是一种高效且可扩展的方法,将标签预测网络与自动标签数量预测机制相结合,利用每个标签的预测置信度得分和目标文档中的深层上下文信息来实现。Glinka 等^[82]提出了一个改进多标签医学文本分类的特征选择方法的模型,研究了过滤器和包装器方法以及混合方法。

Hughes 等^[83]允许自动生成基于上下文的、丰富的健康相关信息表示形式。基于从医学文本语料库中提取的紧急语义,将基于 CNN 的方法在句子级别上对文本片段进行分类,同时将提出的方法与句子嵌入、平均词嵌入和 BOW 单词嵌入 3 种方法进行比较。Yogarajan 等^[84]承认患者的多发性,并利用医学文本分类的多标签变体来增强对并发医学代码的预测。在与健康相关的文本上展示了新的嵌入,并在处理不平衡的多标签医学文本分类问题时比较了嵌入模型的几种变体。

Wasim 等^[85]针对事实类和列表类问题的生物医学问答系统的流程,提出了多标签问题的分类模型。针对预处理的 MLBioMedLAT 语料库是一个多标签问题分类数据集,而且语料库非常稀疏且偏斜,因此在预测阶段,列表类型的问题使用 COPY 的 LAT 预测模型,事实类问题使用 BR 的 LAT 预测模型。Baumel 等^[81]提出了分层注意双向门控循环单元,该模型可以使用注意力权重来更好地了解对决策影响最大的句子以及该句子中对每个决策影响最大的单词,同时可以找到每个标签中得分最高的句子,并通过这个最重要的句子找到得分最高的单词。

2.3 其他

MLC 方法除了在图像和文本分类方面得到了很好的应用,在其他方面也得到了广泛的研究。本节将从音乐和视频两个方面对 MLC 的应用进行介绍。

Oramas 等^[6]提出了一种使用深度学习架构的多标签音乐流派分类模型,该模型将基于学习的特征嵌入与最新的深度学习方法相结合。对于每张专辑,它收集了封面图像、文字评论和音轨。Zhao 等^[86]提出了一种通过用户评论对多标签音乐风格进行分类的模型,着重于评论驱动的多标签音乐风格分类。该模型分为两个机制,基于标签图的神经网络机制负责根据评论和样式相关性对音乐样式进行分类,基于软训练机制引入了具有连续标签表示的新损失函数。Ma 等^[87]提出了一种新颖的知识关系集成框架(Knowledge Relations Integrated Framework, KRF),使用图卷积网络自动学习样式之间的深层关联。该方法专注于整合有关音乐风格和统计信息的外部知识,以共同推导不同风格之间正确而完整的依赖关系,缓解过拟合和欠拟合的问题。

Kim 等^[88]提出了一种基于神经网络的方法。为了忽略嘈杂和无意义的帧,该方法将注意机制用于空间维度和时间维度,还将标签的联合概率分解为条件项来考虑标签之间的相关性。Karagoz 等^[7]提出了用于减少视频数据集维度的自动编码器,并结合了多目标进化非支配排序遗传算法和自动编码器提取的特征来提高视频数据集的预测准确性。Araujo 等^[89]基于线性代数和深度学习交叉路口的研究,提出了基于最先进的网络架构的视频分类模型。该模型将经典形式的图层表示为“密集”,将循环和对角线表示的图层称为“紧凑”,所需要的尺寸使用级联和切片来表示。Jiang 等^[90]提出了一个新的系统来实现短视频的实时和 MLC,在输出的 S 型功能之前添加激活调节层,以增强 CNN 对每个标签的判别能力,使用类不平衡感知训练损失来减小大多数不相关标签带来的影响,标签相关的后处理涉及自适应地融合图像级别的结果,并利用标签关系进行多标签视频分类。

2.4 小结

本节从图像领域和文本分类领域以及其他的应用领域对 MLC 算法进行了总结。在医学图像领域,文献[39]在数据集 CHESTX-RAY14 上可以获得很好的结果,平均值可以达到 82.6%,比文献[64]快 1%;在情感分析领域,文献[72]在数据集 Ren-CECps 上的汉明损失和 Ranking Loss 比文献[3]的低,但准确率却比其高大约 3.69%。

3 评估指标

在多标签分类中,单一地使用某些评估指标是不合适的。针对多标签分类设计了许多评估指标,这些指标可以粗略地被分为两类:基于实例的指标和基于标签的指标。

为了方便理解公式,表2列出了公式所使用的符号和数学含义。

表2 多标签分类数学符号

Table 2 Mathematical symbols for multi-label classification

符号	含义
\mathcal{X}	d 维实例空间 \mathbb{R}^d
\mathcal{Y}	用 q 个可能的类别标签标记空间, $\{y_1, y_2, \dots, y_q\}$
\mathbf{x}	d 维特征向量 $(x_1, x_2, \dots, x_d)^T, (x \in \mathcal{X})$
y'	y' 表示相关标签, 其中 $y' \in Y$
y''	y'' 表示不相关标签, 其中 $y'' \notin Y$
p	实例个数, $1 \leq i \leq p$
Y	与 $\mathbf{x}(Y \subseteq \mathcal{Y})$ 关联的标签集
\bar{Y}	在 \mathcal{Y} 中 Y 的互补集
$f(\mathbf{x}, y)$	实值函数 $f: \mathcal{X} \times \mathcal{Y} \rightarrow \mathbb{R}$, 其中 $f(\mathbf{x}, y)$ 返回为 \mathbf{x} 正确标签的置信度
$h(\mathbf{x})$	MLC 器 $h: \mathcal{X} \rightarrow 2^{\mathcal{Y}}$, 其中 $h(\mathbf{x})$ 返回 \mathbf{x} 的正确标签集

3.1 基于实例的评估指标

基于实例的指标是评估每个测试实例上的性能,然后返回整个测试集的平均值。基于实例的评估指标可以分为分类和 Ranking, 其中分类中包含子集准确度、汉明损失、精确度、召回率和 F 值, Ranking 分为错误率、覆盖率、Ranking Loss 和平均精确度。

3.1.1 分类公式

(1)子集准确度。子集准确度可以评估标签的所有正确分类实例,如式(1)所示。但对于评估来说,它非常严格,会导致非常低的度量值。

$$SA(h) = \frac{1}{p} \sum_{i=1}^p [h(\mathbf{x}_i) = Y_i] \quad (1)$$

(2)汉明损失。汉明损失评估错误分类的实例标签对的比例,即相关标签被遗漏或无关标签被预测。如式(2)所示, Δ 代表两个集合之间的对称差。请注意,当数据集集中的每个实例仅与一个标签相关联时, $HL(h)$ 将是传统误分类率的 $2/q$ 倍。

$$HL(h) = \frac{1}{p} \sum_{i=1}^p |h(\mathbf{x}_i) \Delta Y_i| \quad (2)$$

(3)准确度。它测量正确分类标签的比例,如式(3)所示。

$$Acc(h) = \frac{1}{p} \sum_{i=1}^p \frac{|Y_i \cap h(\mathbf{x}_i)|}{|Y_i \cup h(\mathbf{x}_i)|} \quad (3)$$

(4)精确度。它测量预测样本中被正确分类的比例,如式(4)所示。

$$Pre(h) = \frac{1}{p} \sum_{i=1}^p \frac{|Y_i \cap h(\mathbf{x}_i)|}{|h(\mathbf{x}_i)|} \quad (4)$$

(5)召回率。它测量预测为相关实例的相关标签的平均比例,如式(5)所示。

$$Re(h) = \frac{1}{p} \sum_{i=1}^p \frac{|Y_i \cap h(\mathbf{x}_i)|}{|Y_i|} \quad (5)$$

(6)F 值。精确度和召回度可用于计算加权 F 值,该评估指标通常被认为是比精确度和召回度更好的性能评估指标。其中, β 为平衡因子,通常使 β 等于 1,如式(6)所示。

$$F^\beta(h) = \frac{(1+\beta^2) \cdot Pre(h) \cdot Re(h)}{\beta^2 \cdot Pre(h) + Re(h)} \quad (6)$$

3.1.2 Ranking 公式

(1)1-错误率。评估排名靠前的标签不在相关标签集中实例的比例,如式(7)所示。

$$OE(f) = \frac{1}{p} \sum_{i=1}^p [[\arg \max_{y \in Y} f(\mathbf{x}_i, y)] \notin Y_i] \quad (7)$$

(2)覆盖率。它用于评估平均需要多少步才能沿着已排序的标签列表移动以获得实例的所有相关标签,如式(8)所示。

$$Co(f) = \frac{1}{p} \sum_{i=1}^p \max_{y \in Y_i} rank_f(\mathbf{x}_i, y) - 1 \quad (8)$$

(3)Ranking Loss。计算错误排序的标签对的分数,如式(9)所示。

$$RL(f) = \frac{1}{p} \sum_{i=1}^p \frac{1}{|Y_i| \|\bar{Y}_i\|} |\{(y', y'') \mid f(\mathbf{x}_i, y') \leq f(\mathbf{x}_i, y''), (y', y'') \in Y_i \times \bar{Y}_i\}| \quad (9)$$

(4)平均精确度。它用于评估排名高于特定标签的相关标签的平均分数,如式(10)所示。

$$AP(f) = \frac{1}{p} \sum_{i=1}^p \frac{1}{|Y_i|} \sum_{y \in Y_i} \frac{|\{y' \mid rank_f(\mathbf{x}, y') \leq rank_f(\mathbf{x}_i, y), y' \in Y_i\}|}{rank_f(\mathbf{x}_i, y)} \quad (10)$$

3.2 基于标签的评估指标

基于标签的指标评估每个类标签上的性能,然后返回所有类标签的宏观/微观平均值。基于标签的评估指标可以分为分类和 Ranking, 其中分类中包含宏观平均值和微观平均值; Ranking 分为宏观 AUC 和微观 AUC。

其中,曲线下面积(Area Under Curve, AUC)是一种统计度量,对应于特征曲线下的总面积。该曲线表示真实阳性占总实际阳性(即召回率)的比例与假阳性占总实际阴性的比例,将其引入多标签中是为了衡量分类器避免错误分类的能力^[91]。由于 AUC 是对单标签分类域的度量,因此在多标签中,它也可以对其结果进行宏观 AUC(macro AUC)或微观 AUC(micro AUC)。因它在多标签分类中使用得较少,所以这里重点介绍分类公式。

任何用于单标签分类的度量都适用于多标签分类的基于标签的度量,可以使用两个平均操作来计算所有标签的这些度量,本文称其为宏平均和微平均。宏观平均是在单个类标签上计算的,然后在所有类上平均,微观平均是在所有实例/类标签上全局计算这些指标。TP, TN, FP 和 FN 分别表示由单标签分类器评估的真阳性、真阴性、假阳性和假阴性的数量。

宏观平均如式(11)–式(13)所示。

$$Macro_Precision = \frac{1}{q} \sum_{i=1}^q \frac{TP_i}{TP_i + FP_i} \quad (11)$$

$$Macro_recall = \frac{1}{q} \sum_{i=1}^q \frac{TP_i}{TP_i + FN_i} \quad (12)$$

$$Macro_F1 = \frac{1}{q} \sum_{i=1}^q \frac{2 \times R_i \times P_i}{R_i + P_i} \quad (13)$$

其中, R_i 和 P_i 是第 i 个标签的召回率与精确度。

微观平均如式(14)–式(16)所示。

$$Micro_Precision = \frac{\sum_{i=1}^q TP_i}{\sum_{i=1}^q TP_i + \sum_{i=1}^q FP_i} \quad (14)$$

$$Micro_recall = \frac{\sum_{i=1}^q TP_i}{\sum_{i=1}^q TP_i + \sum_{i=1}^q FN_i} \quad (15)$$

$$Micro_F1 = \frac{2 \times Micro_Precision \times Micro_Recall}{Micro_Precision + Micro_Recall} \quad (16)$$

3.3 小结

一个特定的评估指标在单独考虑时可能不如另一个,但这并不意味着该评估指标在评估中完全没有作用,可以选择与更敏感的评估指标结合使用。一般而言,在衡量算法性能

时应提供各种评估指标,而不是将性能由单个评估指标决定。

对于基于实例的评估指标,在研究分类器性能时,可以采用基于分类的评估指标来进行评估。但是大多数的多标签学习算法需要返回实值函数,因此将从 Ranking 指标的角度进行评估。

为了更好地了解评估指标在多标签中的作用,将通过图 3 对上述监督学习算法的评估指标进行描述。图 3 中,SA 为子集精确度,Acc 为准确度,Pre 为精确度,Re 为召回度,Micro Pre 为微预测,Macro Pre 为宏预测,Micro Re 为微召回度,Macro Re 为宏召回度,Macro Av 为宏平均,Co 为覆盖度,RL 为 Ranking Loss,AP 为平均精度,OE 为 1-错误率。

表 3 多标签分类监督学习算法的评估指标

Table 3 Evaluation indexes of multi-label classification supervised learning algorithm

算法	评估指标																
	HL	SA	Acc	Pre	Re	F1	Micro Pre	Macro Pre	Micro Re	Macro Re	Micro F1	Macro F1	Macro Av	Co	RL	AP	OE
ML-DT ^[22]	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓		✓	✓	✓	✓
LdSM ^[23]				✓													
AEML-LLSVM ^[25]	✓													✓	✓	✓	✓
AEDC-MLSVM ^[26]	✓													✓	✓	✓	
BP-AEPMML ^[28]	✓													✓	✓	✓	✓
NNMLInf ^[57]	✓													✓	✓	✓	✓
ELIFT ^[31]		✓	✓														
3RC ^[33]	✓		✓		✓	✓											✓
MLNB-LD ^[34]			✓								✓	✓					
TreeStacking ^[35]	✓	✓	✓	✓	✓	✓											✓
BCC ^[36]	✓		✓														
RBRL ^[37]	✓	✓				✓								✓	✓	✓	
SSSVM ^[38]	✓																
LCL-Net ^[39]				✓	✓	✓											
SDLM ^[40]						✓											
DCNet ^[41]	✓													✓	✓	✓	✓
ACKEL ^[43]	✓	✓									✓			✓	✓	✓	✓

4 未来工作展望

现有的基于监督学习和半监督学习的 MLC 文献已经能很好地解决目前的分类问题,但是对于研究人员来说,仍然有一些严峻的问题需要解决,例如半监督学习下的复杂概念漂移处理、特征选择处理、标签复杂相关性处理和类不平衡处理。下文将对这 4 类问题进行分析并将其作为未来的研究方向。

(1) 半监督学习下的复杂概念漂移处理

随着数据流变得越来越普遍,挖掘瞬态和动态数据的在线算法的必要性变得越来越明确^[92]。随着数据流实例的引入,数据的分布与输入变量和输出变量之间的关系可能会随时间发生变化,这称为概念漂移^[93]。目前用于解决概念漂移的多标签算法很少,同时将全部标签进行标记需要消耗大量的精力,这就使得半监督下的概念漂移成为 MLC 问题中一个值得研究的方向。由于概念漂移的类型多种多样,如渐变漂移、突变漂移、重复漂移等,如何有效地检测概念漂移成为目前需要解决的难点。基于块和增量更新策略在单标签算法中被广泛应用并取得了很好的效果。本文使用 BR 方法将其转换成二元分类器,采用基于块的和增量更新策略的算法来检测概念漂移,同时将现有的半监督算法通过凸优化等策略

进行优化,降低了时间复杂度并提高了性能。

(2) 特征选择处理

特征选择是数据预处理的过程,通过特征选择可以降低问题的复杂度,提高学习算法的预测精度和鲁棒性。一般而言,冗余或无关功能的存在可能会导致分类性能较差等其他问题,并且可能对计算和内存存储有更高的要求^[94]。传统的多标签方法认为标签的重要性相等,并在整个标签集上进行选择,但这造成了高维的数据灾难并产生了昂贵的经济代价。在监督学习、半监督学习下的多标签算法中,选择包含高度相关且非冗余特征的特征子集可以很大程度地滤除冗余特征,目前最基本的特征选择方法可以大致分为包装方法、嵌入方法和过滤方法,但是它们必须在特征选择开始执行之前收集完整的特征集,这具有一定的局限性,因为现实中很多功能是在动态变化的。针对这一特点,可以采用在线特征选择的方法对多标签问题进行处理,以解决特征选择等问题。

(3) 标签复杂相关性处理

现有的分类方法只简单考虑标签之间的相关性,但有些标签之间的相关性却存在非常复杂的关系。例如,有些数据集内的标签存在双向关系和多个周期性依存关系。例如,对“海滩”类别的预测取决于“城市”值,而对“城市”类别的预测取决于“海滩”值^[33]。这使得标签之间的相关性在监督学习

与半监督学习下的 MLC 问题中有重要的研究意义。目前所提出的算法只有一部分考虑了复杂标签相关的问题,有效考虑标签相关性可以提高算法分类性能,可以考虑引入依赖权重来确定相关性,然后仅保留重要的依赖。

(4)类不平衡处理

大多数多标签数据集都存在严重的类不平衡问题,这将严重影响分类性能^[26]。类不平衡数据包括两个方面:1)对于特定的类标签,肯定实例的数量大大少于否定实例的数量;2)对于特定实例,相关标签的数量通常少于无关标签的数量^[37]。传统的分类器更适合对平衡数据进行分类,因为当多标签数据中的类大小不平衡时,分类性能会急剧下降。因此,将监督学习、半监督学习对类不平衡的处理作为下一步的研究有着重要的意义,将主动学习应用于训练集的选择,可以有效提高分类器的精度。

结束语 本文对现有的基于监督学习以及半监督学习的 MLC 算法进行了综述。在监督学习中,详细描述了标签非相关性和标签相关性算法,从时间复杂度和数据集角度对算法的性能进行了比较和分析,并绘制表格对其优缺点进行了描述。在半监督学习中,具体介绍了批处理和在线的相关算法,通过数据集的对比来介绍算法的性能,以图的方式概述了其优缺点。同时,对多标签的图像和文本分类等实际应用领域的算法进行了综述,简略地总结了多标签领域的评估指标,并通过表格的方式对监督学习算法的评估指标进行了介绍。最后,基于目前 MLC 所面临的挑战,例如半监督学习下的复杂概念漂移处理、特征选择处理、标签复杂相关性处理和类不平衡处理,提出了一定的解决方法。

参 考 文 献

- [1] DU S Y, HAN M, SHEN M Y, et al. Survey of ensemble classification algorithms for data streams with concept drift [J]. *Computer Engineering*, 2020, 46(1): 15-24, 30.
- [2] MA J, ZHANG H, CHOW T W S. Multilabel Classification With Label-Specific Features and Classifiers: A Coarse and Fine Tuned Framework [J]. *IEEE Transactions on Cybernetics*, 2021, 51(2): 1028-1042.
- [3] HE H, XIA R. Joint Binary Neural Network for Multi-label Learning with Applications to Emotion Classification [C] // *Proceedings of the Natural Language Processing and Chinese Computing*. 2018: 250-259.
- [4] DU J, CHEN Q, PENG Y, et al. ML-Net: multi-label classification of biomedical texts with deep neural networks [J]. *Journal of the American Medical Informatics Association*, 2019, 26(11): 1279-1285.
- [5] RANA P, MEIJERING E, SOWMYA A, et al. Multi-Label Classification Based On Subcellular Region-Guided Feature Description For Protein Localisation [C] // *Proceedings of the 18th International Symposium on Biomedical Imaging*. France: IEEE, 2021: 1929-1933.
- [6] ORAMAS S, NIETO O, BARBIERI F, et al. Multi-label Music Genre Classification from Audio, Text, and Images Using Deep Features [C] // *Proceedings of the 18th International Society for Music Information Retrieval Conference*. 2017: 23-30.
- [7] KARAGOZ G N, YAZICI A, DOKEROGLU T, et al. Analysis of Multiobjective Algorithms for the Classification of Multi-Label Video Datasets [J]. *IEEE Access*, 2020, 8: 163937-163952.
- [8] READ J, PFAHRINGER B, HOLMES G, et al. Classifier chains for multi-label classification [J]. *Machine Learning*, 2011, 85(3): 333-359.
- [9] TSOU MAKAS G, VLAHAVAS I. Randomk-Labelsets: An Ensemble Method for Multilabel Classification [C] // *Proceedings of the 18th European Conference on Machine Learning*. Berlin: Springer, Warsaw; *Lecture Notes in Computer Science*, 2007: 406-417.
- [10] OLIVEIRA E, CIARELLI P M, BADUE C, et al. A Comparison between a KNN Based Approach and a PNN Algorithm for a Multi-label Classification Problem [C] // *Proceedings of the Eighth International Conference on Intelligent Systems Design Applications*. Kaohsiung: IEEE, 2008: 628-633.
- [11] CIARELLI P M, KING R D. Knowledge Discovery in Multi-label Phenotype Data [C] // *Proceedings of the European Conference on Principles of Data Mining and Knowledge Discovery*. Freiburg: *Lecture Notes in Computer Science*, 2001: 42-53.
- [12] LI X, WANG L, SUNG E. Multi-label SVM active learning for image classification [C] // *Proceedings of the Image Processing*. Singapore: IEEE, 2004: 2207-2210.
- [13] SAPOZHNIKOVA E P. Multi-label Classification with ART Neural Networks [C] // *Proceedings of the 2009 Second International Workshop on Knowledge Discovery and Data Mining*. Moscow: IEEE, 2009: 144-147.
- [14] ZHANG M, ZHOU Z. A Review on Multi-Label Learning Algorithms [J]. *IEEE Transactions on Knowledge & Data Engineering*, 2014, 26(8): 1819-1837.
- [15] TSOU MAKAS G, KATAKIS I. Multi-Label Classification: An Overview [J]. *International Journal of Data Warehousing and Mining*, 2009, 3(3): 1-13.
- [16] MOYANO J M, GIBAJA E L, CIOS K J, et al. Review of Ensembles of Multi-Label Classifiers: Models, Experimental Study and Prospects [J]. *Information Fusion*, 2018, 44: 33-45.
- [17] ZHOU Z H, ZHANG M L. Multi-label Learning [J/OL]. *Encyclopedia of Machine Learning and Data Mining*, 2017: 875-881. https://link.springer.com/referenceworkentry/10.1007/978-1-4899-7687-1_910.
- [18] ZHENG X, LI P, CHU Z, et al. A Survey on Multi-Label Data Stream Classification [J]. *IEEE Access*, 2020, 8: 1249-1275.
- [19] LI P, WANG H, BHM C, et al. Online Semi-supervised Multi-label Classification with Label Compression and Local Smooth Regression [C] // *Proceedings of the Twenty-Ninth International Joint Conference on Artificial Intelligence*. Yokohama, 2020: 1359-1365.
- [20] WANG Z, WANG T, WAN B, et al. Partial Classifier Chains with Feature Selection by Exploiting Label Correlation in Multi-Label Classification [J]. *Entropy*, 2020, 22(10): 1-22.
- [21] BEZEMBINDER E M, WISMANS L J J, BERKUM E C V. Constructing multi-labelled decision trees for junction design using the predicted probabilities [C] // *Proceedings of the 20th IEEE International Conference on Intelligent Transportation*

- Systems. Yokohama: IEEE, 2017: 1-7.
- [22] MORAL-GARCÍA S, MANTAS C J, CASTELLANO J G, et al. Non-parametric predictive inference for solving multi-label classification[J/OL]. *Applied Soft Computing*, 2020, 88: 106011. <https://www.sciencedirect.com/science/article/pii/S1568494619307938?via%3Dihub>.
- [23] MAJZUBI M, CHOROMANSKA A. LdSM: Logarithm-depth Streaming Multi-label Decision Trees[C]// *Proceedings of the the 23rd International Conference on Artificial Intelligence and Statistics*. Palermo, 2019: 4247-4257.
- [24] YANG Y, DING M. Decision function with probability feature weighting based on Bayesian network for multi-label classification[J]. *Neural Computing and Applications*, 2019, 31(9): 4819-4828.
- [25] SUN Z, HU K, HU T, et al. Fast Multi-Label Low-Rank Linearized SVM Classification Algorithm Based on Approximate Extreme Points[J]. *IEEE Access*, 2018, 6: 42319-42326.
- [26] SUN Z, LIU X, HU K, et al. An Efficient Multi-Label SVM Classification Algorithm by Combining Approximate Extreme Points Method and Divide-and-Conquer Strategy[J]. *IEEE Access*, 2020, 8: 170967-170975.
- [27] BELLMANN F, BUNZEL L, DEMUS C, et al. Multi-Label Classification of Blurbs with SVM Classifier Chains[C]// *Proceedings of the 15th Conference on Natural Language Processing*. 2019.
- [28] WANG X, GUO Z, WANG X, et al. A Fast Neural Network Multi-label Classification Algorithm Based on Approximate Extreme Points[C]// *Proceedings of the 5th International Conference on Big Data Computing and Communications*. QingDao: IEEE, 2019: 17-21.
- [29] BELLO M, GONZALO N, RICARDO S, et al. Deep neural network to extract high-level features and labels in multi-label classification problems[J]. *Neurocomputing*, 2020, 413: 259-270.
- [30] SUMBUL G, BEGÜM D. A Deep Multi-Attention Driven Approach for Multi-Label Remote Sensing Image Classification [J]. *IEEE Access*, 2020, 8: 95934-95946.
- [31] WEI X, YU Z, ZHANG C, et al. Ensemble of Label Specific Features for Multi-Label Classification [C] // *Proceedings of the 2018 IEEE International Conference on Multimedia and Expo*. San Diego. IEEE, 2018: 1-6.
- [32] MORAL-GARCÍA S, MANTAS C J, CASTELLANO J G, et al. Ensemble of classifier chains and Credal C4. 5 for solving multi-label classification[J]. *Progress in Artificial Intelligence*, 2019, 8(2): 195-213.
- [33] LOTF H, RAMDANI M. Multi-Label Classification: A Novel approach using decision trees for learning Label-relations and preventing cyclical dependencies: Relations Recognition and Removing Cycles (3RC)[C]// *Proceedings of the 13th International Conference on Intelligent Systems: Theories and Applications*. Rabat, ACM, 2020: 1-6.
- [34] KIM H, PARK J, KIM D, et al. Multilabel naïve Bayes classification considering label dependence [J]. *Pattern Recognition Letters*, 2020, 136: 279-285.
- [35] ZHANG Z D, WANG Z H, LIU H Y. An Integrated Multi-label Classification Algorithm Based on Tree Bayesian Network[J]. *Computer Science*, 2018, 45: 195-201.
- [36] JIMÉNEZ R B, MORALES E F, ESCALANTE H J. Bayesian Chain Classifier with Feature Selection for Multi-label Classification[C] // *Proceedings of the Mexican International Conference on Artificial Intelligence*. 2018: 232-243.
- [37] WU G, ZHENG R, TIAN Y, et al. Joint Ranking SVM and Binary Relevance with robust Low-rank learning for multi-label classification[J]. *Neural Networks*, 2020, 122: 24-39.
- [38] KODA S, ZEGGADA A, MELGANI F, et al. Spatial and Structured SVM for Multilabel Image Classification[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2018, 56(10): 5948-5960.
- [39] CHEN B, LI J, LU G, et al. Label Co-occurrence Learning with Graph Convolutional Networks for Multi-label Chest X-ray Image Classification[J]. *IEEE Journal of Biomedical and Health Informatics*, 2019, 24(8): 2292-2302.
- [40] LI J, FU G, CHEN Y, et al. A multi-label classification model for full slice brain computerised tomography image [J]. *BMC Bioinformatics*, 2020, 21(6): 1-18.
- [41] HE J, LI C, YE J, et al. Multi-label ocular disease classification with a dense correlation deep neural network [J]. *Biomedical Signal Processing and Control*, 2021, 63: 102167.
- [42] BRAYTEE A, LIU W, ANAISSI A, et al. Correlated Multi-label Classification with Incomplete Label Space and Class Imbalance [J]. *ACM Transactions on Intelligent Systems and Technology*, 2019, 10(5): 1-26.
- [43] WANG R, KWONG S, JIA Y, et al. Active k-labelsets ensemble for multi-label classification[J/OL]. *Pattern Recognition*, 2021, 109. <https://www.sciencedirect.com/science/article/pii/S0031320320303861?via%3Dihub>.
- [44] WANG R, KWONG S, JIA Y, et al. Mutual Information Based K-Labelsets Ensemble for Multi-Label Classification[C] // *Proceedings of the 2018 IEEE International Conference on Fuzzy Systems*. Rio de Janeiro: IEEE, 2018: 1-7.
- [45] BOULBAZINE S, CABANES G, MATEI B, et al. Online Semi-supervised Growing Neural Gas for Multi-label Data Classification[C] // *Proceedings of the 2018 International Joint Conference on Neural Networks*. Rio de Janeiro: IEEE, 2018: 1-8.
- [46] CEVIKALP H, BENLIGIRAY B, GEREK O N. Semi-Supervised Robust Deep Neural Networks for Multi-Label Image Classification[J]. *Pattern Recognition*, 2019, 100: 107164.
- [47] LI S, FU Y. Robust multi-label semi-supervised classification [C]// *Proceedings of the 2017 IEEE International Conference on Big Data*. Boston: IEEE, 2017: 27-36.
- [48] LIU Y, NIE F, GAO Q X. Nuclear-norm Based Semi-supervised Multiple Labels Learning[J/OL]. *Neurocomputing*, 2017: 940-947. <https://linkinghub.elsevier.com/retrieve/pii/S0925231217315588>.
- [49] AMORIM W P, FALCAO A X, PAPA J P. Multi-label semi-supervised classification through optimum-path forest[J]. *Information Sciences*, 2018, 465: 86-104.
- [50] ABBURI H, PARIKH P, CHHAYA N, et al. Fine-grained Multi-label Sexism Classification Using Semi-supervised Lear-

- ning[C]//Proceedings of the 21st International Conference Amsterdam, Amsterdam: Lecture Notes in Computer Science, 2020: 531-547.
- [51] XU Z, ZHAI Y, LIU Y. Distributed Semi-supervised Multi-label Classification with Quantized Communication[C]//Proceedings of the 12th International Conference on Machine Learning and Computing, Shenzhen: ACM, 2020: 57-62.
- [52] JING L, SHEN C, YANG L, et al. Multi-Label Classification by Semi-Supervised Singular Value Decomposition[J]. IEEE Transactions on Image Processing, 2017, 26(10): 4612-4625.
- [53] QIU Y, GONG X, MA Z. MixLab: An Informative Semi-supervised Method for Multi-label Classification[C]//Proceedings of the 9th CCF International Conference. Zhengzhou: Lecture Notes in Computer Science, 2020: 506-518.
- [54] LI P, WANG H, BHM C, et al. Online Semi-supervised Multi-label Classification with Label Compression and Local Smooth Regression[C]//Proceedings of the Twenty-Ninth International Joint Conference on Artificial Intelligence. Yokohama, 2020: 1359-1365.
- [55] AHMADI Z, KRAMERS. A label compression method for online multi-label classification[J]. Pattern Recognition Letters, 2018, 11: 64-71.
- [56] CHEN N, RIBEIRO B, TANG C, et al. Multi-label learning vector quantization for semi-supervised classification[J]. Intelligent Data Analysis, 2019, 23(4): 839-853.
- [57] WANG X P, GUO Z W, WANG X, et al. NNMLInf: social influence prediction with neural network multi-label classification [C]//Proceedings of the ACM Turing Celebration Conference. Chengdu: ACM, 2019, 106: 1-5.
- [58] ZHAN W, ZHANG M L. Inductive Semi-supervised Multi-Label Learning with Co-Training[C]//Proceedings of the Acm Sigkdd International Conference on Knowledge Discovery & Data Mining. Halifax: ACM, 2017: 1305-1314.
- [59] XING Y, YU G, DOMENICONI C, et al. Multi-Label Co-Training[C] // Proceedings of the Twenty-Seventh International Joint Conference on Artificial Intelligence. Stockholm, 2018: 2882-2888.
- [60] CHU Z, LI P, HU X. Co-training Based on Semi-Supervised Ensemble Classification Approach for Multi-label DataStream [C]//Proceedings of the 2019 IEEE International Conference on Big Knowledge. Beijing: IEEE, 2019: 58-69.
- [61] LEE C H. Multi-label classification of documents using fine-grained weights and modified co-training [J]. Intelligent Data Analysis, 2018, 22(1): 103-115.
- [62] TAN Q, YU Y, YU G, et al. Semi-supervised multi-label classification using incomplete label information [J]. Neurocomputing, 2017, 260: 192-202.
- [63] CHAUDHURI B, DEMIR B, CHAUDHURI S, et al. Multilabel Remote Sensing Image Retrieval Using a Semisupervised Graph-Theoretic Method [J]. IEEE Transactions on Geoscience and Remote Sensing, 2018, 56(2): 1144-1158.
- [64] GUAN Q, HUANG Y. Multi-label Chest X-ray Image Classification via Category-wise Residual Attention Learning [J]. Pattern Recognition Letters, 2018, 130: 259-266.
- [65] CAI J, SUN W, GUAN J, et al. Multi-ECGNet for ECG Arrhythmia Multi-Label Classification [J]. IEEE Access, 2020, 8: 110848-110858.
- [66] MERCAN C, AKSOY S, MERCAN E, et al. Multi-Instance Multi-Label Learning for Multi-Class Classification of Whole Slide Breast Histopathology Images [J]. IEEE Trans. Med. Imaging, 2017, 37(1): 316-325.
- [67] XU S, YANG X, GUO J, et al. CXNet-m3: a deep quintuplet network for multi-lesion classification in Chest X-ray Images via multi-label supervision [J]. IEEE Access, 2020, 8: 98693-98704.
- [68] HUA Y, MOU L, ZHU X X. Relation Network for Multilabel Aerial Image Classification [J]. IEEE Transactions on Geoscience and Remote Sensing, 2020, 58(7): 4558-4572.
- [69] ALSHEHRI A, BAZI Y, AMMOUR N, et al. Deep Attention Neural Network for Multi-label Classification in Unmanned Aerial Vehicle Imagery [J]. IEEE Access, 2019, 7: 119873-119880.
- [70] DAI O E, DEMIR B, SANKUR B, et al. A Novel System for Content-Based Retrieval of Single and Multi-Label High-Dimensional Remote Sensing Images [J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2018, 11(7): 2473-2490.
- [71] YU J, LUÍS M, JIANG J, et al. Improving Multi-label Emotion Classification via Sentiment Classification with Dual Attention Transfer Network [C]//Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing, 2018: 1097-1102.
- [72] FEI H, JI D, ZHANG Y, et al. Topic-Enhanced Capsule Network for Multi-Label Emotion Classification [J]. IEEE/ACM Transactions on Audio, Speech, and Language Processing, 2020, 28: 1839-1848.
- [73] WEN W, CHEN Y, CAI R. Reader sentiment classification based on multi-perspective and multi-label learning [J]. Computer Science, 2018, 45(8): 198-204.
- [74] ALZU'BI S, BADARNEH O, HAWASHIN B, et al. Multi-Label Emotion Classification for Arabic Tweets [C]//Proceedings of the 2019 Sixth International Conference on Social Networks Analysis. Granada: IEEE, 2019: 499-504.
- [75] BRAVO-MARQUEZ F, FRANK E, MOHAMMAD S M, et al. Determining Word-Emotion Associations from Tweets by Multi-label Classification [C]//Proceedings of the IEEE/WIC/ACM International Conference on Web Intelligence. Omaha, NE: ACM, 2016: 536-539.
- [76] KIM Y, LEE H, JUNG K. AttnConvnet at SemEval-2018 Task 1: Attention-based Convolutional Neural Networks for Multi-label Emotion Classification [C]//Proceedings of the 12th International Workshop on Semantic Evaluation. New Orleans: Association for Computational Linguistics, 2018: 141-145.
- [77] MULKI H, ALI C B, HADDAD H, et al. Tw-StAR at SemEval-2018 Task 1: Preprocessing Impact on Multi-label Emotion Classification [C]//Proceedings of the SemEval-2018. New Orleans: Association for Computational Linguistics, 2018: 167-171.
- [78] ALHUZALI H, ANANIADOU S. SpanEmo: Casting Multi-la-

- bel Emotion Classification as Span-prediction[C]// Proceedings of the 16th Conference of the European Chapter of the Association for Computational Linguistics. 2021:1573-1584.
- [79] HYUN J,BAE B,CHEONG Y. [CL-AFF Shared Task] Multi-label Text Classification Using an Emotion Embedding Model [C]// Proceedings of the 3rd Workshop of Affective Content Analysis. New York:CEUR Workshop Proceedings, 2020: 169-178.
- [80] YING W,XIANG R,LU Q. Improving Multi-label Emotion Classification by Integrating both General and Domain-specific Knowledge[C]// Proceedings of the 5th Workshop on Noisy User-generated Text. Hong Kong: Association for Computational Linguistics, 2019: 316-321.
- [81] BAUMEL T, NASSOUR-KASSIS J, COHEN R, et al. Multi-Label Classification of Patient Notes; Case Study on ICD Code Assignment[C]// Proceedings of the Thirty-Second AAAI Conference on Artificial Intelligence. 2018: 409-416.
- [82] GLINKA K,WOZNIAK R,ZAKRZEWSKA D. Improving Multi-label Medical Text Classification by Feature Selection [C]// Proceedings of IEEE International Conference on Enabling Technologies: Infrastructure for Collaborative Enterprises. Poznan: IEEE, 2017: 176-181.
- [83] HUGHES M, LI I, KOTOULAS S, et al. Medical Text Classification using Convolutional Neural Networks[J/OL]. Computer Science, 2017. <https://arxiv.org/abs/1704.06841>.
- [84] YOGARAJAN V, MONTIEL J, SMITH T, et al. Seeing The Whole Patient: Using Multi-Label Medical Text Classification Techniques to Enhance Predictions of Medical Codes[J/OL]. Computer Science, 2020. <https://arxiv.org/abs/2004.00430>.
- [85] WASIM M, MAHMOOD W, ASIM M N, et al. Multi-Label Question Classification for Factoid and List Type Questions in Biomedical Question Answering[J]. IEEE Access, 2019, 7: 3882-3896.
- [86] ZHAO G, XU J, ZENG Q, et al. Review-Driven Multi-Label Music Style Classification by Exploiting Style Correlations [C]// Proceedings of the 2019 Conference of the North American Chapter. Minneapolis: Association for Computational Linguistics, 2019: 2884-2891.
- [87] MA Q, YUAN C, ZHOU W, et al. Beyond Statistical Relations: Integrating Knowledge Relations into Style Correlations for Multi-Label Music Style Classification [C]// Proceedings of the 13th International Conference on Web Search and Data Mining. Houston: ACM, 2020: 411-419.
- [88] KIM E, ON K, KIM J, et al. Temporal Attention Mechanism with Conditional Inference for Large-Scale Multi-label Video Classification [C]// Proceedings of the European Conference on Computer Vision. Munich: Lecture Notes in Computer Science, 2018: 306-316.
- [89] ARAUJO A, NEGREVERGNE B, CHEVALEYRE Y, et al. Training Compact Deep Learning Models for Video Classification Using Circulant Matrices [C]// Proceedings of European Conference on Computer Vision. Munich: Lecture Notes in Computer Science, 2018: 271-286.
- [90] JIANG B, ZHOU L, LIN L, et al. A Real-Time Multi-Label Classification System for Short Videos [C]// Proceedings of 2019 IEEE International Conference on Image Processing. Taipei: IEEE, 2019: 534-538.
- [91] PEREIRA R B, PLASTINO A, ZADROZNY B, et al. Correlation analysis of performance measures for multi-label classification[J]. Information Processing & Management, 2018, 54(3): 359-369.
- [92] BÜYÜKÇAKIR A, BONAB H, CAN F. A Novel Online Stacked Ensemble for Multi-Label Stream Classification [C]// Proceedings of the 27th ACM International Conference on Information and Knowledge Management. 2018: 1063-1072.
- [93] LU Y X, LIU B N, WANG C R, et al. Adaptive incremental ensemble classification algorithm for concept drift data stream[J]. Small Microcomputer System, 2019, 40(12): 2624-2630.
- [94] LIU J, LIN Y, LI Y, et al. Online multi-label streaming feature selection based on neighborhood rough set[J]. Pattern recognition, 2018, 84: 273-287.



WU Hong-xin, born in 1998, bachelor. Her main research interests include data mining and so on.



HAN Meng, born in 1982, Ph.D, professor, master's supervisor. Her main research interests include data mining and so on.

(责任编辑:喻藜)