

基于多粒度标签扰动的文本分类研究

姚汝婧^{1,2} 王芳^{1,2*}

(1. 南开大学商学院信息资源管理系, 天津 300071; 2. 南开大学网络社会治理研究中心, 天津 300071)

摘要: [目的/意义] 基于深度学习的有监督学习算法是当前文本分类主要的研究方法。然而, 有监督的深度学习算法的训练严重依赖于样本标签的准确性, 由于标注者的标注经验和主观性等原因, 样本标签不可避免地会存在噪声。标签扰动是一种有效应对噪声标签的思路, 但当前基于标签扰动的噪声标签学习算法缺乏对多种粒度信息的有效利用, 从而限制了算法的性能。[方法/过程] 为解决这一问题, 本文提出了一种多粒度标签扰动算法(Multi-granularity Label Perturbation, MGLP), 融合了样本级粒度和类别级粒度的扰动方式, 并利用元学习的思想对不同粒度扰动方式的融合权重进行学习, 使该算法能够根据不同的数据特点自适应地对融合权重进行调整。[结果/结论] 本文在推文情感分类、电影评论情感分类、引文意图分类3个文本分类数据集上展开实验, 结果表明MGLP算法有效地提升了深度学习模型在文本分类任务上的性能, 在信息组织和信息分析中具有广泛的应用前景。

关键词: 文本分类; 深度学习; 标签扰动; 元学习; 多粒度

DOI: 10.3969/j.issn.1008-0821.2024.01.003

[中图分类号] TP391 [文献标识码] A [文章编号] 1008-0821 (2024) 01-0025-12

Text Classification Based on Multi-granularity Label Perturbation

Yao Rujing^{1,2} Wang Fang^{1,2*}

(1. Department of Information Resources Management, Business School, Nankai University, Tianjin 300071, China;

2. Center for Network Society Governance, Nankai University, Tianjin 300071, China)

Abstract: [Purpose/Significance] Supervised learning algorithms based on deep learning are currently the main research methods for text classification. However, the training of supervised deep learning algorithms heavily relies on the accuracy of the sample labels. Due to the annotator's experience and subjectivity, sample labels inevitably contain noise. Label perturbation is an effective way to deal with noisy labels. However, noisy label learning algorithms based on label perturbation lack effective utilization of multiple granularity information at present, which limits the performance of the algorithms. [Method/Process] In order to address the problem, the paper proposed a multi-granularity label perturbation algorithm (MGLP), which combined sample-level granularity and category-level granularity perturbation methods. The MGLP algorithm used the idea of meta-learning to learn the fusion weights of different granularity perturbation methods, which could adaptively adjust the fusion weights according to different data characteristics. [Result/Conclusion] The paper conducts experiment on three text classification datasets, including tweet sentiment classification, movie review sentiment classification, and citation intent classification. The results show that the proposed MGLP algorithm effectively improves the performance of deep learning models in text classification tasks and has broad application prospects in information organization and information analysis.

Key words: text classification; deep learning; label perturbation; meta-learning; multi-granularity

收稿日期: 2023-07-18

基金项目: 国家社会科学基金重大项目“基于数据共享与知识复用的数字政府智能化治理研究”(项目编号: 20ZDA039)。

作者简介: 姚汝婧(1995-), 女, 博士研究生, 研究方向: 数据挖掘与深度学习。

通讯作者: 王芳(1970-), 女, 教授, 博士生导师, 研究方向: 知识发现与情感挖掘、情报学理论与方法、政府信息资源管理、网络社会治理。

文本分类是信息组织和信息分析中的重要内容,其涉及的范围十分广泛,不仅包括学术文献的分类任务,如文献结构分类、引文情感分类、引文意图分类等,也包含社交媒体信息分类、突发事件的识别与分类、政策文本分类等。随着文献数量的飞速增长,对于学术文献相关内容的分类能够帮助学者在面对浩如烟海的文献资料时,快速地识别出所需信息,了解目标文献的研究内容与研究价值^[1-2]。政策文本分类能够帮助政府、高校、企业等高效地获取自身所需的科技政策^[3]。社交媒体信息分类能够帮助决策者快速了解公众的情感和观点,并利用这些有价值的分类信息优化和改进现有的解决方案^[4]。突发事件的自动识别与分类能够从海量信息中快速获取有效信息,并为应急管理机构提供帮助^[5]。

深度学习算法因其良好的特征表示能力,逐渐成为了解决上述问题的首选方案。然而,有监督的深度学习算法的训练依赖于训练数据,一个大型高质量的训练样本集对能否有效训练深度学习算法几乎起着决定性作用。但是,由于标注经验、主观性以及责任心等因素的影响,人工标注的数据集不可避免地存在着噪声。如有研究表明图像分类中最为知名的大型数据集 ImageNet 约包含了 6% 的噪声标签^[6],而 NER 领域应用最为广泛的数据集 CoNLL-2003 也被发现约 5% 的标签错误存在于测试句中^[7]。可想而知,对于深度学习任务所采用的非基准数据集,其噪声标签比例极可能更高,而这些噪声标签会对深度学习的模型训练造成负面影响。因此,在文本分类研究中,寻求有效的噪声标签学习方法也已经成为一个热点问题。

迄今为止,研究者们已经提出了各种各样的噪声标签学习算法^[8],通过降低噪声标签对模型训练的负面影响进而提升模型的性能,基于标签扰动的算法就是其中一类典型的学习策略。顾名思义,标签扰动的目标是训练样本的标签,即通过对一部分训练样本的标签添加扰动来减少噪声标签对模型造成的负面影响,增强模型的鲁棒性。按照扰动粒度的不同,标签扰动算法可以分为样本级粒度的扰动、类别级粒度的扰动、数据集级粒度的扰动。目前,有 3 种具有代表性的单一粒度的标签扰动算法: Label Smoothing(标签平滑)算法^[9]、Bootstrapping 算

法^[10]、Online Label Smoothing(在线标签平滑)算法^[11]。Label Smoothing 和 Online Label Smoothing 为类别级粒度的标签扰动算法, Bootstrapping 为样本级粒度的标签扰动算法。

多项研究表明,不同粒度的标签扰动算法都能有效地提升模型性能,然而,现有的算法大多是从同一粒度下的深入探索,缺乏对不同粒度信息的有效利用,而不同级别的粒度信息能够进行互补从而提高模型的性能。基于此,本文首先分析了 Label Smoothing、Bootstrapping、Online Label Smoothing 3 种单粒度的标签扰动算法如何具体地对学习过程起着调节作用。然后,提出了一种多粒度的标签扰动算法(Multi-granularity Label Perturbation, MGLP),该算法通过加权的方式将前述的 3 种单粒度的标签扰动算法融合在一起。随着融合权重取值的不同,本文提出的 MGLP 算法可以简化为 3 种扰动方式中的任意一种或者两种的组合。对于融合权重,本文采用元学习的思想对其进行学习,使之能够根据不同的数据特点自适应地进行调整,减轻了人工调参的负担,并减少了主观性偏差对结果造成的负面影响。最后,本文将提出的 MGLP 算法应用在推文情感分类、电影评论情感分类、引文意图分类 3 个文本分类数据集上,通过施加不同类型噪声的方式验证算法的性能,实验结果表明本文提出的 MGLP 算法有效地提升了深度学习算法在文本分类任务上的准确性,对于深度学习算法更准确地在信息组织和信息分析领域的应用具有十分重要的价值和意义。

1 相关研究

1.1 文本分类

文本分类在信息组织和信息分析中发挥着日益重要的作用。早期的文本分类利用信息增益^[12]、互信息^[13]或者主题模型^[14]等提取特征,然后利用浅层分类器进行分类。近年来,以词向量为基础的分布表示和以 LSTM^[15]、Transformer^[16]等为代表的深度学习算法逐步取代了早期文本分类的方法。如 BERT、ERNIE 等模型被用来对文献的学科进行分类^[17]。基于 SciBert 的模型被用于学术文献致谢的识别^[18]。融合多种特征的深度学习模型可以较好地实现对评论中的用户意见的分类^[19]。除了在上述文本分类任务上以外,深度学习算法也广泛地应

用在突发事件的识别与分类任务以及政策文本分类方面。吴雪华等^[5]提出了一个两阶段的突发事件应急行动支撑信息的识别与分类框架，且利用 SVM、LR、TextCNN 以及 BERT 等算法进行实验来验证其性能。一种 BERT 与多尺度 CNN 融合的算法被提出且用来捕获科技政策文本的特征信息，对政策文本的主题进行分类^[20]。深度学习算法具有较强的特征表示能力，被广泛应用于各种文本分类任务，且使得文本分类任务的准确性有了进一步的提高。

除了词汇的分布表示、更为有效的深度学习网络架构等研究之外，针对训练数据的不完美特点设计有效的学习策略，如噪声标签、类别不平衡等问题，也是文本分类领域的研究重点。针对类别不平衡问题，研究者们提出了多种解决方案。Zong D 等^[21]设计了一个双通道的学习策略来解决文本分类中的长尾分布问题。卢小宾等^[22]提出了综合数据、算法、评估 3 个层面的优化框架以解决新兴技术识别中的数据类别不平衡问题。为了更好地处理虚假评论识别任务，一种基于类别可分性计算的代价敏感学习方法被提出^[23]。基于类别先验 Mixup 数据增强策略被用来解决罪名分类任务中的不平衡问题^[24]。此外，元学习和小样本学习等方法也被引入来处理数据中存在问题。一种在不平衡少样本情况下基于元学习的文本分类模型被提出^[25]。小样本数据增强技术被用于对科技文档的不平衡分类问题进行解决^[26]。通过对上述研究的总结和分析发现，文本分类在信息组织和信息分析中发挥着越来越重要的作用。此外，针对数据的不完美特点探讨有效的学习策略逐步成为研究的热点。

1.2 噪声标签学习

在文本分类任务中，基于浅层机器学习和深度学习的算法逐渐成为主要方法，而在算法的训练过程中，训练数据中存在的噪声标签会对算法的训练造成负面影响，比如容易导致所学习到的模型产生过拟合等问题，因此，噪声标签学习逐渐成为一个重点的研究方向。在浅层机器学习时代，噪声标签学习就是一个极受关注的研究问题。如经典的支持向量机算法^[27]，所引入的松弛变量的一大动机就是抑制噪声标签的不利影响。相对于浅层学习时代相对规模较小的训练数据，深度学习的训练对人工

标注的数据集规模有着更高的要求，不可避免地会进一步带来噪声标签问题。

目前解决噪声标签问题的途径主要有两种，一种是对噪声标签样本进行离线检测，另一种是基于噪声标签样本进行在线检测。离线检测的方法主要利用损失^[28]、交叉验证错误率^[7]、几何边界距离^[29]等量化指标来区分正常标签样本与噪声标签样本。如置信学习被用来对样本集进行清洗，降低噪声数据对模型的负面影响，进而有效地提升了模型的性能^[30]。为了检测命名实体识别任务中的噪声样本，一种基于交叉验证的方式被提出来计算每个样本的预测正确率^[7]。该正确率越小那么该样本更可能是噪声样本，然后将正确率作为样本权重重新进行训练。此外，汪敏等^[31]提出了一种噪声识别与纠正算法，通过筛选可信样本对样本标签的置信度进行预测，然后识别噪声标签，对噪声标签进行纠正。Huang J 等^[32]提出了一个基于过拟合—欠拟合过程的策略来识别噪声样本。对噪声标签样本的离线检测方法的主要缺陷在于通常需要增加大量的训练时间，为此，研究者们开发出了基于噪声标签样本的在线检测途径。

基于噪声标签样本的在线检测方法隐式地降低噪声标签的不利影响，基于标签扰动的方法就是其中一类典型的学习策略。许多基于噪声标签学习的深度学习算法，甚至其一些研究分支，本质上都可以归结为对训练数据的扰动。如当前深度学习中的热点研究方向：对抗攻击^[33]以及基于对抗攻击的对抗训练^[34]。对抗攻击的根本性问题就是寻求一个满足特定目标的样本扰动并叠加到输入样本上。从数据对象上看，现有的方法可以分为特征扰动、逻辑向量扰动以及标签扰动等几大类别。对抗攻击可以看作是特征扰动。此外，近期一些代表性的研究从不同角度来提升算法的泛化性能，如 IS-DA^[35]、Logit Adjustment^[36]等在数学上都可以归结为逻辑向量扰动。Label Smoothing、Bootstrapping、Online Label Smoothing 可以归结为标签扰动。从扰动粒度上，现有的方法可以分为训练集级别、类别级别以及样本级别。Label Smoothing 和 Online Label Smoothing 是类别级粒度的扰动，Bootstrapping 是样本级粒度的扰动。然而，目前大多数研究通常在同

一粒度下进行探索和创新,极少有研究综合考虑利用不同粒度的有效信息,而多种粒度信息的有效利用能够帮助模型更好地学习特征表示,有利于提升模型的性能。基于此,本文针对单一粒度的不足,探讨多种粒度级别下的标签扰动,以期通过多粒度信息的有效利用提升模型的性能。

2 研究方法

本节首先对3种单粒度的经典标签扰动算法进行了分析,然后针对单粒度算法没有有效地利用不同粒度级别信息的缺陷,提出了一种多粒度标签扰动算法。该算法综合考虑了样本级粒度和类别级粒度信息,弥补了单粒度算法的不足。对于不同粒度信息的融合权重,本文采用了元学习的思想对其进行学习,使本文提出的方法能够根据不同的数据特点自适应地调整融合权重,减少了人工调参的负担,降低了主观因素对结果产生的不利影响。

2.1 经典标签扰动算法的分析

在噪声标签学习中,Label Smoothing、Bootstrapping、Online Label Smoothing是3种经典的算法,本文首先对上述3种噪声标签学习算法进行分析。给定一组训练集 $D = \{(x_i, y_i)\}, i \in \{1, \dots, N\}$,其中 N 为样本总数, y_i 为样本 x_i 的类别标签(one-hot形式)。在一轮深度神经网络训练中,假定当前模型对该样本的预测输出(softmax形式)为 p_i ,总类别数为 C ,扰动之后的标签为 y'_i (one-hot形式)。对于标准的交叉熵损失函数,其定义如式(1)所示。

$$L = \sum_{i=1}^N l(p_i, y_i) \quad (1)$$

Label Smoothing的原理是通过引入一个超参数 λ 来对标签 y_i 进行扰动,避免模型过度自信造成过拟合的情况,提升模型的泛化性能,其扰动公式为:

$$y'_i = y_i + \lambda \left(\frac{1}{C} - y_i \right) \quad (2)$$

其中, $\lambda \in [0, 1]$, $\frac{1}{C} = \left[\frac{1}{C}, \dots, \frac{1}{C} \right]^T \in R^C$, $\lambda \left(\frac{1}{C} - y_i \right)$ 为扰动项,扰动之后, y'_i 为一个软性的one-hot向量。

由此,可以得到基于Label Smoothing的损失函数为:

$$L = \sum_{i=1}^N l \left(p_i, y_i + \lambda \left(\frac{1}{C} - y_i \right) \right) \quad (3)$$

Label Smoothing对相同类别中的样本进行了相同的扰动,而不同类别的样本的扰动是不同的,即Label Smoothing进行了类别级的扰动,从而减轻了噪声标签对模型训练产生的负面影响。

Bootstrapping方法与Label Smoothing具有明显的差异,Bootstrapping根据模型的当前状态对样本标签进行扰动,其扰动公式为:

$$y'_i = y_i + \lambda (p_i - y_i) \quad (4)$$

其中, λ 是超参数, $\lambda \in [0, 1]$, $\lambda (p_i - y_i)$ 为扰动项。

Bootstrapping方法有Soft Bootstrapping和Hard Bootstrapping两种类型,Soft Bootstrapping的损失函数为:

$$L_{soft} = \sum_{i=1}^N l(p_i, y_i + \lambda (p_i - y_i)) \quad (5)$$

假设 p_i 的one-hot形式为 y_{p_i} ,则Hard Bootstrapping的损失函数为:

$$L_{hard} = \sum_{i=1}^N l(p_i, y_i + \lambda (y_{p_i} - y_i)) \quad (6)$$

Bootstrapping方法可以看作对每个样本的标签都进行了个性化的扰动,即样本级的扰动。随着训练的迭代,模型的预测结果会更加可信。通过利用模型的当前状态对标签进行扰动,缓解噪声标签对模型训练的负面影响。

Online Label Smoothing算法是一种典型的类别级标签扰动方法。Online Label Smoothing算法利用每个类别中预测正确的样本的预测概率来对标签进行扰动,其扰动公式为:

$$y'_i = y_i + \lambda (p_{y_i} - y_i) \quad (7)$$

其中, $\lambda (p_{y_i} - y_i)$ 为扰动项。 p_{y_i} 由前一个epoch产生,定义如式(8)所示。

$$p_{y_i} = \frac{1}{N} \sum_{j=1}^N p_{j, y_i} \quad (8)$$

其中, N 为标签是 y_i 的样本中被模型预测正确的样本数量, p_{j, y_i} 为标签是 y_i 的样本中被模型预测正确的样本的预测概率。

该方法的损失函数如式 (9) 所示:

$$L = \sum_{i=1}^N l(p_i, y_i + \lambda(p_{y_i} - y_i)) \quad (9)$$

Online Label Smoothing 能够随着模型的训练动态地对扰动项的取值进行调整, 减少对噪声标签样本的过度拟合, 降低噪声标签对模型的负面影响。

通过对 Bootstrapping 和 Online Label Smoothing 的分析发现, Online Label Smoothing 本质上算是一种类别级的 Bootstrapping 的思路。分析 Bootstrapping 的扰动式 (4) 和 Online Label Smoothing 算法的扰动式 (7) 可知, 当超参数 λ 取值为 0 的时候, Bootstrapping 方法和 Online Label Smoothing 方法的思路中均有 $y'_i = y_i$ 。当 λ 取值为 1 的时候, Bootstrapping 的扰动结果为 $y'_i = p_i$, 对于不同的样本, p_i 的取值不同。Online Label Smoothing 的扰动结果为 $y'_i = p_{y_i}$, 对于不同类别下的样本, p_{y_i} 取值不同, 对于相同类别下的不同样本, p_{y_i} 的取值相同。分析 Label Smoothing 的扰动式 (2) 和 Online Label Smoothing 算法的扰动式 (7) 可知, 当超参数 λ 取值为 0 的时候, Label Smoothing 方法和 Online Label Smoothing 方法的思路中均有 $y'_i = y_i$ 。当 λ 取值为 1 的时候, Label Smoothing 的扰动结果为 $y'_i = \frac{I}{C}$, Online Label Smoothing 的扰动结果为 $y'_i = p_{y_i}$ 。Label Smoothing 算法和 Online Label Smoothing 虽然

都是类别级的标签扰动算法, 但是其扰动的取值方式不同, 将 Label Smoothing 和 Online Label Smoothing 二者结合可进行互补, 充分发挥不同角度的优势, 降低噪声标签对模型的影响。

2.2 多粒度标签扰动算法 (Multi-granularity Label Perturbation, MGLP)

2.2.1 MGLP 算法设计

大量的理论和实验已经证明, 类别级的标签扰动和样本级的标签扰动对于解决噪声标签问题非常有效和高效。然而, 当前大部分研究都是在同一粒度下对算法的深入探索, 鲜有研究综合考虑利用这些不同粒度的标签扰动策略提升模型的性能, 而不同粒度的扰动能够从不同的角度对算法性能进行补充。基于此, 本文提出了一种多粒度标签扰动算法 (Multi-granularity Label Perturbation, MGLP), 该算法对 Label Smoothing、Bootstrapping、Online Label Smoothing 3 种单粒度的标签扰动策略进行了集成, 综合考虑了样本级粒度和类别级粒度的扰动。同时, 利用元学习的思想对 3 种标签扰动策略的融合系数进行学习, 技术路线如图 1 所示。由于在文本分类任务中, BERT 和 BiLSTM 具有良好的特征表示能力^[37,17], 因此, 为了全面地验证本文所提出的算法的性能, 本研究分别利用 BERT-Base 和 BiLSTM-Attention 作为基线模型进行实验。

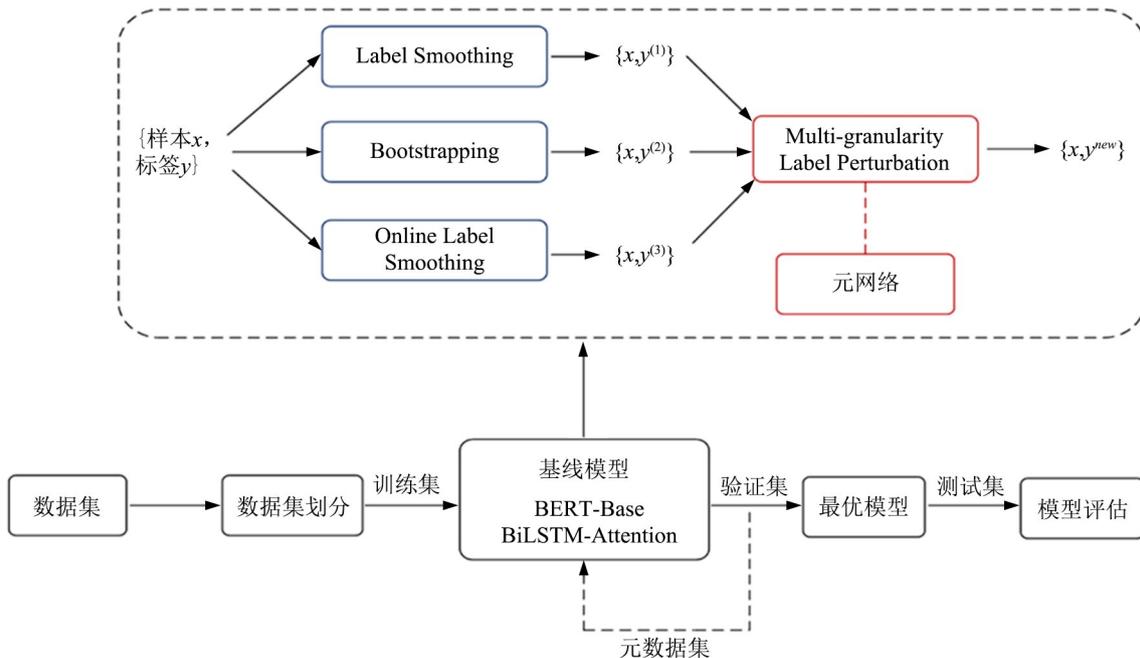


图 1 技术路线

Fig. 1 Technical Roadmap

具体来说, 在将数据集进行划分后, MGLP 算法作用于模型的训练过程, 以降低噪声标签对模型训练的负面影响, 提高模型的性能。MGLP 算法通过加权的方式将 Label Smoothing、Bootstrapping、Online Label Smoothing 3 种算法进行了融合, 在进行扰动时, 通过融合系数的取值来控制不同粒度的扰动比例, 融合系数的取值通过元学习的思想来求得。假定 α_i 、 β_i 、 γ_i 为样本 i 的 3 个控制扰动比例的融合系数, 则 MGLP 算法的扰动后的标签如式 (10) 所示:

$$y'_i = y_i + \alpha_i \left(\beta_i p_{y_i} + \gamma_i p_i + (1 - \beta_i - \gamma_i) \frac{I}{C} - y_i \right) \quad (10)$$

其中, $\alpha_i \left(\beta_i p_{y_i} + \gamma_i p_i + (1 - \beta_i - \gamma_i) \frac{I}{C} - y_i \right)$ 为扰动项。

MGLP 算法的损失函数如式 (11) 所示:

$$L = \sum_{i=1}^N l \left(p_i, y_i + \alpha_i \left(\beta_i p_{y_i} + \gamma_i p_i + (1 - \beta_i - \gamma_i) \frac{I}{C} - y_i \right) \right) \quad (11)$$

MGLP 算法融合了类别级标签扰动和样本级标签扰动算法的特点, α_i 、 β_i 、 γ_i 3 个融合系数可以根据不同的数据特点对不同粒度的扰动比例进行控制。当 $\alpha_i \neq 0$ 、 $\beta_i = 0$ 、 $\gamma_i \in (0, 1)$ 时, MGLP 算法的损失函数如式 (12) 所示, 此时 MGLP 算法集成了样本级粒度的 Bootstrapping 和类别级粒度 Label Smoothing。

$$L = \sum_{i=1}^N l \left(p_i, y_i + \alpha_i \left(\gamma_i p_i + (1 - \gamma_i) \frac{I}{C} - y_i \right) \right) \quad (12)$$

当 $\alpha_i \neq 0$ 、 $\beta_i \in (0, 1)$ 、 $\gamma_i = 0$ 时, MGLP 算法的损失函数如式 (13) 所示, 此时 MGLP 算法集成了类别级粒度的 Online Label Smoothing 和类别级粒度 Label Smoothing。

$$L = \sum_{i=1}^N l \left(p_i, y_i + \alpha_i \left(\beta_i p_{y_i} + (1 - \beta_i) \frac{I}{C} - y_i \right) \right) \quad (13)$$

当 $\alpha_i \neq 0$ 、 $\beta_i \in (0, 1)$ 、 $\gamma_i \in (0, 1)$ 、 $\beta_i + \gamma_i = 1$ 时, MGLP 算法的损失函数如公 (14) 所示, 此时 MGLP 算法集成了类别级粒度的 Online Label Smoothing 和样本级粒度的 Bootstrapping。

$$L = \sum_{i=1}^N l \left(p_i, y_i + \alpha_i (\beta_i p_{y_i} + \gamma_i p_i - y_i) \right) \quad (14)$$

MGLP 算法除了能够利用不同粒度的扰动信息

以外, 还可以简化为单粒度的扰动策略。具体来说, 当 $\alpha_i \neq 0$ 、 $\beta_i = 0$ 、 $\gamma_i = 1$ 时, MGLP 算法简化为 Soft Bootstrapping 算法, 即样本级的扰动算法。当 $\alpha_i \neq 0$ 、 $\beta_i = 1$ 、 $\gamma_i = 0$ 时, MGLP 算法简化为 Online Label Smoothing 算法, 即类别级的扰动算法。当 $\alpha_i \neq 0$ 、 $\beta_i = 0$ 、 $\gamma_i = 0$ 时, MGLP 算法简化为 Label Smoothing 算法, 即类别级的扰动算法。

2.2.2 基于元学习的训练

对于 α_i 、 β_i 、 γ_i 3 个融合系数, 本文利用元学习的思想进行学习^[38]。因此, 本小节首先对基于元学习的神经网络训练策略进行介绍。在深度学习领域, 元学习又被称为“学会学习 (Learning to Learn)^[39]”, 其有很多不同的学习范式, 一种典型的学习范式就是利用元学习来自动对超参数进行优化。与传统的利用网格化的超参数优化的策略相比, 基于元学习的超参数优化既有较高的效率, 也有较好的理论保障。

基于元学习的超参数优化策略首先需要专门构建一个元数据集, 记为 \hat{D} , 以及一个专门用来超参数预测的元网络, 其参数记为 θ 。在神经网络的一个训练 epoch 中, 该策略主要包含 3 个步骤: 首先, 根据上一轮元网络输出的超参数来计算当前 epoch 的训练损失, 并基于损失来得到主干网络临时参数, 记为 \hat{w} ; 其次, 基于临时参数的主干网络来计算元数据集的损失, 并基于该损失来得到元网络新的参数, 记为 θ^{t+1} ; 最后, 基于新的元网络参数 θ^{t+1} 计算出的新的超参数来重新计算当前 epoch 的训练损失, 利用该损失来得到该 epoch 最终的主干网络参数, 记为 w^{t+1} 。经过上述 3 个步骤后, 主干网络参数、元网络参数以及超参数都得到了优化。下面具体介绍本研究所涉及的元学习训练细节。

给定一组训练集元数据集 $\hat{D} = \{ (x_i^{meta}, y_i^{meta}) \}$, $i \in \{1, \dots, M\}$, 其中 M 为元数据集中的样本总数, y_i^{meta} 为样本 x_i^{meta} 的类别标签 (one-hot 形式)。对于深度学习模型的参数 w , 通过最小化式 (15) 来计算。

$$w^* (\theta) = \arg \min_w L^{train} (w; \theta) \triangleq \frac{1}{N} \sum_{i=1}^N l_i^{train} (w; \theta) \quad (15)$$

其中, θ 为元网络的参数。

对于元网络中的参数 θ , 通过最小化式 (16) 来计算。

$$\theta^* = \arg \min_{\theta} L^{meta}(w^*(\theta)) \triangleq \frac{1}{M} \sum_{i=1}^M l_i^{meta}(w^*(\theta)) \quad (16)$$

本文利用多层感知机作为元网络, 具体来说, 首先使用深度学习模型 (Bert-base/BiLSTM-Attention) 得到当前批次中每个样本对应的 p_i 和 p_{y_i} , 然后将 p_i 、 p_{y_i} 以及样本的标签 y_i 作为元网络的输入, 得到融合系数 α_i 、 β_i 、 γ_i 以及扰动后的标签 y'_i , 然后利用式 (11) 求得损失, 即添加扰动后每个训练样本的损失 $l_i^{train}(w)$ 如式 (17) 所示。

$$y'_i = V(p_i, p_{y_i}, y_i; \theta) \quad (17)$$

$$l_i^{train}(w) = l(p_i, y'_i)$$

其中, V 表示元网络。

然后, 利用式 (18) 对深度学习模型中的参数进行更新。

$$\hat{w}^t(\theta) = w^t - a \frac{1}{n} \sum_{i=1}^n \nabla_w l_i^{train}(w) \Big|_{w^t} \quad (18)$$

其中, n 为当前批次中的训练样本的数量, a 为步长。

更新得到 $\hat{w}^t(\theta)$ 后, 利用式 (19), 对元网络中的参数 θ 进行更新。

$$\theta^{t+1} = \theta^t - b \frac{1}{m} \sum_{i=1}^m \nabla_{\theta} l_i^{meta}(\hat{w}^t(\theta)) \Big|_{\theta^t} \quad (19)$$

其中, m 为当前批次中元样本的数量, b 为步长。

然后利用更新得到的 θ^{t+1} 来优化深度学习模型中的参数 w^{t+1} , 如式 (20) 所示。

$$w^{t+1} = w^t - a \frac{1}{n} \sum_{i=1}^n \nabla_w l_i^{train}(w) \Big|_{w^t} \quad (20)$$

3 实验

3.1 数据

为了充分地验证本文所提出的 MGLP 算法的性能, 选择 3 个知名的国际公开的英文文本分类数据集进行实验: SemEval-2016 Task 4 Subtask A^[40]、Movie Reviews (MR)^[41] 和 SciCite^[42]。第一个数据集来自于 SemEval-2016 的任务 4 的子任务 A, 是

推文情感分类数据集, 包含正面情感、中性情感、负面情感 3 种类型, 本文采用官方给定的训练集、验证集和测试集的划分。第二个数据集是 MR 数据集, 其为一个电影评论情感分类数据集, 包含正面情感和负面情感两种类型的标签, 由于官方数据未划分训练集、验证集和测试集, 因此, 本文按照 7:1:2 的比例将 MR 数据集划分为训练集、验证集、测试集。SciCite 是一个引文意图分类数据集, 包含背景、方法和结果 3 种类型, 本文采用该数据集官方给定的训练集、验证集和测试集的划分。

3.2 对比算法及参数设置

为了有效地验证本文提出的 MGLP 算法的性能, 采用以下几种经典的以及最先进的噪声标签处理算法进行对比: Label Smoothing^[9]、Soft/Hard Bootstrapping^[10]、Online Label Smoothing^[11]、Self-Distillation from Last Mini-Batch (DLB)^[43]、Margin-based Label Smoothing (MbLS)^[44]。本文在原始数据集的基础上施加了两种类型的噪声, 一种是对称噪声, 一种是非对称噪声^[8]。对称噪声指的是样本的真实标签以相同的概率随机翻转成其他类别标签, 非对称噪声指的是样本的真实标签被翻转成某类特定的标签, 本文设置的样本的真实标签翻转比例 (即噪声比例) 为 10%、20%、30%。

对于本文采用的 BERT-Base, 其有 12 层 Transformer 结构, 隐藏层维度为 768, 学习率设为 $2e-5$, epoch 设为 10。对于 BiLSTM-Attention, 本文采用 300 维的 Glove 词向量, 隐藏层维度设为 300, epoch 设为 50。对于 Label Smoothing、Soft Bootstrapping、Hard Bootstrapping、Online Label Smoothing、DLB、MbLS 等对比算法, 按照其对应的原论文的设置进行实验。在 MGLP 算法中, 本文随机选取验证集中的每类样本 100 个作为元数据集。

3.3 实验结果

表 1 为基线模型为 BERT-Base 时不同噪声标签处理算法在 3 种不同数据集上的实验结果, 表 2 为基线模型为 BiLSTM-Attention 时不同噪声标签处理算法在 3 种不同数据集上的实验结果。采用分类任务中常用的准确率作为评价指标, 每个实验进行 3 次取其平均值作为最终结果。

表 1 BERT-Base 为基线模型时不同数据集的分类准确率(%)
Tab. 1 Classification Accuracy of Different Datasets with BERT-Base as the Baseline Model (%)

数据集	模 型	原始数据集	对称噪声			非对称噪声		
			10%	20%	30%	10%	20%	30%
SemEval-2016	BERT-Base	61.03	60.05	58.76	53.20	60.07	59.10	56.78
	Label Smoothing	61.70	60.69	59.35	55.37	60.19	59.21	57.04
	Soft Bootstrapping	61.08	60.77	59.74	55.32	60.97	59.28	57.65
	Hard Bootstrapping	62.13	61.31	58.93	54.56	61.22	60.02	57.27
	Online Label Smoothing	61.65	60.42	58.86	55.92	60.53	60.37	57.67
	DLB	62.01	60.31	59.12	55.84	61.25	59.39	57.89
	MbLS	61.94	61.02	59.18	57.49	61.34	60.56	57.81
	MGLP	63.05	62.16	61.53	59.57	62.23	61.68	59.62
MR	BERT-Base	85.67	84.20	81.77	80.02	84.15	81.71	71.39
	Label Smoothing	85.76	84.29	83.08	80.60	84.49	82.16	71.45
	Soft Bootstrapping	85.92	84.43	82.38	80.10	84.17	81.75	71.59
	Hard Bootstrapping	85.69	84.37	82.94	80.53	84.20	82.33	71.41
	Online Label Smoothing	85.71	84.34	83.06	80.71	84.46	81.93	71.52
	DLB	85.74	84.31	82.55	80.54	84.19	82.24	71.66
	MbLS	85.85	84.40	83.17	80.76	84.53	82.21	71.73
	MGLP	86.72	85.37	83.91	82.42	85.23	82.97	72.84
SciCite	BERT-Base	85.37	84.85	84.19	83.76	84.86	84.30	83.79
	Label Smoothing	85.65	85.09	84.30	84.02	85.38	84.88	84.42
	Soft Bootstrapping	85.41	85.06	84.38	84.11	85.25	84.79	84.31
	Hard Bootstrapping	85.70	84.89	84.67	84.16	85.14	84.76	84.28
	Online Label Smoothing	85.73	84.92	84.41	84.29	85.01	84.75	84.36
	DLB	85.44	85.02	84.59	84.31	85.27	84.81	84.35
	MbLS	85.71	85.10	84.62	84.47	85.43	84.92	84.49
	MGLP	86.43	85.97	85.46	85.05	85.94	85.42	84.97

通过表 1 的实验结果,可以看出本文提出的 MGLP 算法以 BERT-Base 为基线模型时,在 3 种数据集上都取得了最佳结果。

具体来说,在 SemEval-2016 数据集上,不添加噪声的情况下,MGLP 算法相比于基线模型 BERT-Base 提高了 2.02%。相比于单粒度的 Label Smoothing、Soft Bootstrapping、Hard Bootstrapping、Online Label Smoothing 算法,MGLP 算法分别提升了 1.35%、1.97%、0.92%、1.40%。说明 MGLP 算法充分利用了样本级粒度和类别级粒度的信息,结合了 Label Smoothing、Bootstrapping、Online Label Smoothing 3 种经典的单粒度噪声标签算法的优势。在施加噪声的情况下,相比于近年来的 MbLS 算法,本文提出的 MGLP 算法平均提升了 1.57%,并且在施加 20%对称噪声的情况下,准确率高于 MbLS 算

法 2.35%。在施加 30%对称噪声时,对比算法中准确率最高的为 MbLS 算法,而本文提出的 MGLP 算法的准确率相比于 MbLS 算法提升了 2.08%,且相比于基线模型 BERT-Base 算法提升了 6.37%,具有明显的性能优势。在 MR 数据集上,在不添加噪声的情况下,MGLP 相对于基线模型 BERT-Base 提升了 1.05%。在施加噪声的情况下,相比于经典的 Label Smoothing、Soft Bootstrapping、Hard Bootstrapping、Online Label Smoothing 算法,MGLP 算法分别平均提升了 1.11%、1.39%、1.16%、1.12%。施加 30%对称噪声和 30%非对称噪声时,MGLP 算法相比于对比算法中准确率最高的 MbLS 算法分别提升了 1.66%和 1.11%,取得了最佳结果。在 SciCite 数据集上,不添加噪声的情况下,MGLP 相对于基线模型 BERT-Base 提升了 1.06%,相比于对

比算法中准确率最高的 Online Label Smoothing 提升了 0.70%。在施加噪声 20% 对称噪声的情况下，相比于对比算法中准确率最高的 MbLS 算法提升了 0.86%，实现了最优性能。以上结果表明 MGLP 算

法能够充分利用不同粒度的数据信息，减少噪声标签对模型性能产生的负面影响，使模型的性能具有明显的提升。

表 2 BiLSTM-Attention 为基线模型时不同数据集的分类准确率 (%)

Tab. 2 Classification Accuracy of Different Datasets with BiLSTM-Attention as the Baseline Model (%)

数据集	模 型	原始数据集	对称噪声			非对称噪声		
			10%	20%	30%	10%	20%	30%
SemEval-2016	BiLSTM-Attention	55.35	52.37	51.87	50.98	53.98	52.23	51.79
	Label Smoothing	55.42	53.25	51.92	51.10	54.32	53.09	52.78
	Soft Bootstrapping	55.61	53.31	52.08	51.55	54.40	53.05	52.13
	Hard Bootstrapping	55.40	52.51	51.97	51.57	54.01	52.57	51.95
	Online Label Smoothing	55.38	52.62	52.01	51.60	54.05	53.39	52.90
	DLB	55.53	52.98	52.04	51.56	54.41	53.22	52.96
	MbLS	55.72	53.34	52.13	51.63	54.43	53.35	53.09
	MGLP	56.45	54.01	53.59	52.87	55.76	54.39	53.97
MR	BiLSTM-Attention	78.42	77.14	74.11	71.48	76.43	73.90	64.57
	Label Smoothing	78.82	77.46	74.73	72.04	76.46	74.05	65.83
	Soft Bootstrapping	78.87	77.57	74.39	72.77	76.92	74.10	67.79
	Hard Bootstrapping	78.91	77.20	74.74	71.62	76.45	74.33	67.92
	Online Label Smoothing	78.93	77.54	75.01	72.19	76.78	74.21	65.25
	DLB	78.95	77.28	74.57	71.75	76.74	73.95	67.28
	MbLS	79.01	77.51	74.78	72.22	76.81	74.24	68.56
	MGLP	79.65	78.59	75.62	73.51	77.73	75.38	69.05
SciCite	BiLSTM-Attention	80.24	78.94	78.47	77.57	79.73	78.58	76.01
	Label Smoothing	80.63	79.96	78.58	77.72	80.03	78.80	76.19
	Soft Bootstrapping	80.71	79.54	78.66	77.63	80.01	78.69	76.10
	Hard Bootstrapping	80.39	79.76	78.73	77.87	80.12	78.75	76.49
	Online Label Smoothing	80.58	79.81	78.67	77.91	80.18	78.72	76.73
	DLB	80.79	79.86	78.74	77.85	80.24	78.83	76.58
	MbLS	80.84	79.93	78.89	77.98	80.16	78.91	76.68
	MGLP	81.45	80.40	79.95	79.19	80.96	79.98	77.43

通过表 2 可知，以 BiLSTM-Attention 为基线模型时，MGLP 算法在 3 个数据集上都实现了最优的性能。

在 SemEval-2016 数据集上，MGLP 算法在不施加噪声的情况下相比于基线模型提升了 1.10%，相比于准确率最高的 MbLS 算法提升了 0.73%。在 SemEval-2016 数据集施加对称噪声的情况下，MGLP 算法相比于基线模型平均提升了 1.75%；在施加非对称噪声的情况下，MGLP 算法相比于基线模型平均提升了 2.04%，相比于近年来的 DLB 算法

和 MbLS 算法，MGLP 算法分别平均提升了 1.18%、1.08%。在施加 20% 对称噪声时，MGLP 算法相比于对比算法中准确率最高的 MbLS 算法提升了 1.46%。在 MR 数据集上，相比于经典的 Label Smoothing、Soft Bootstrapping、Hard Bootstrapping、Online Label Smoothing 算法，MGLP 算法平均提升了 1.45%、1.02%、1.19%、1.37%。在施加 10% 对称噪声的情况下，MGLP 算法相比于对比算法中准确率最高的 Soft Bootstrapping 算法提升了 1.02%，实现了最佳性能。此外，在 MR 数据集施加 30% 非对

称噪声的情况下，MGLP 算法相比于基线模型提升了 4.48%。在 SciCite 数据集上，MGLP 算法准确率也是最高的，相比于 BiLSTM-Attention 基线模型平均提升了 1.40%。在 SciCite 上施加 30% 对称噪声的情况下，MGLP 算法相比于 Label Smoothing、Soft Bootstrapping、Hard Bootstrapping、Online Label Smoothing、DLB、MbLS 算法分别提升了 1.47%、1.56%、1.32%、1.28%、1.34%、1.21%。此外，在施加 20% 非对称噪声时，对比算法中准确率最高的为 MbLS 算法，本文提出的 MGLP 算法相比于 MbLS 算法提升了 1.07%，取得了最优结果。以上实验结果均表明，MGLP 算法充分利用了 3 种经典的单粒度的噪声标签算法的优势，融合了样本级粒度和类别级粒度的数据信息，提升了深度学习模型的性能。在模型进行学习的过程中，MGLP 算法能够根据不同数据的特点，自适应地选取不同的比例对样本级粒度和类别级粒度的数据信息进行融合，从而减少噪声样本对模型训练产生的负面影响。实验结果表明，不管是在原始数据集上还是在施加噪声的情况下，MGLP 算法均有效且明显地提升了深度学习模型的性能。

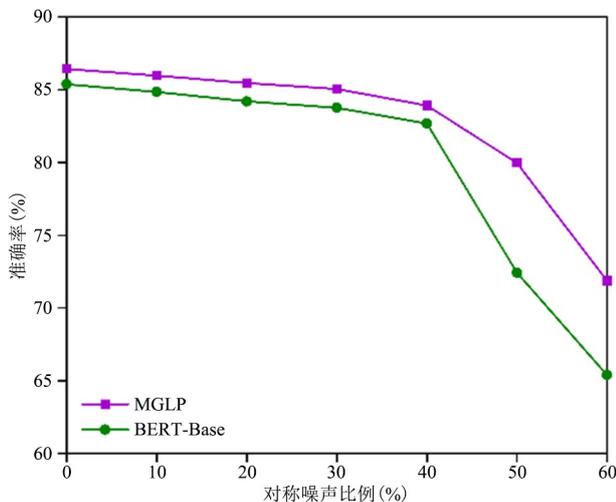


图 2 施加对称噪声时模型准确率 (BERT-Base 为基线模型)

Fig. 2 Model Accuracy with Imposed Symmetric Noise (BERT-Base as the Baseline Model)

为了探究 MGLP 算法的鲁棒性，本文以 SciCite 数据集为例，对更高比例噪声标签下算法的性能进行了评估，结果如图 2~图 5 所示。在施加 40%、50%、60% 比例的对称噪声情况下，本文提出的 MGLP 算法相比于 BERT-Base 和 BiLSTM-Attention

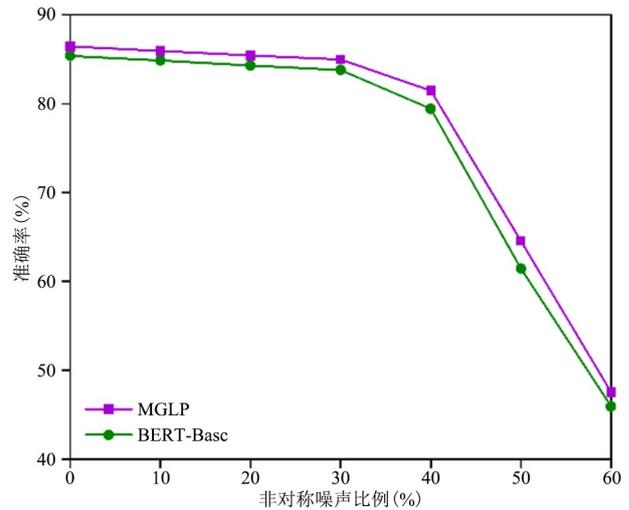


图 3 施加非对称噪声时模型准确率 (BERT-Base 为基线模型)

Fig. 3 Model Accuracy with Imposed Asymmetric Noise (BERT-Base as the Baseline Model)

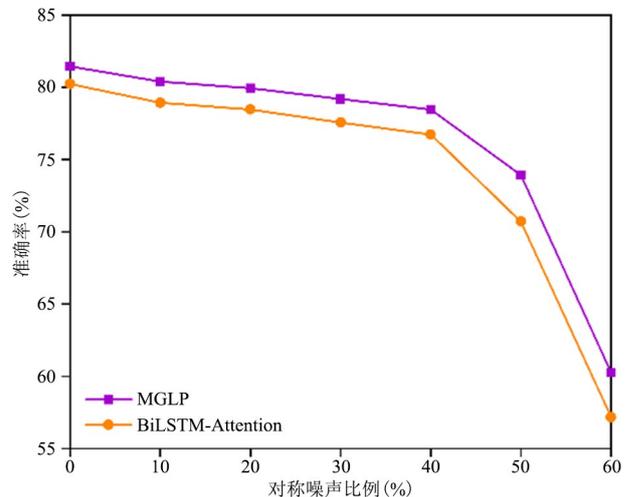


图 4 施加对称噪声时模型准确率 (BiLSTM-Attention 为基线模型)

Fig. 4 Model Accuracy with Imposed Symmetric Noise (BiLSTM-Attention as the Baseline Model)

基线模型分别平均提升了 5.09%、2.68%。在施加 40%、50%、60% 比例的非对称噪声情况下，MGLP 算法相比于 BERT-Base 和 BiLSTM-Attention 基线模型分别平均提升了 2.24%、2.22%。实验结果表明，即使在施加更高比例噪声的情况下，本文提出的 MGLP 算法仍保持较高的准确率，具有良好的鲁棒性。

4 结语

本文针对现有的标签扰动算法大都只在单一粒度层级下进行深入探索，而未有效利用多种粒度信息，从而限制了算法的性能这一不足之处，首先分

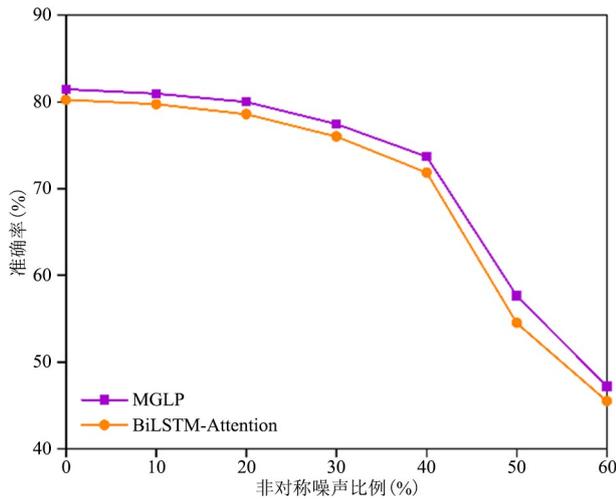


图5 施加非对称噪声时模型准确率
(BiLSTM-Attention 为基线模型)

Fig. 5 Model Accuracy with Imposed Asymmetric Noise
(BiLSTM-Attention as the Baseline Model)

析了 Label Smoothing、Bootstrapping 和 Online Label Smoothing 3 种经典的单一粒度的标签扰动算法的原理，然后提出了一种融合了样本级粒度和类别级粒度的多粒度标签扰动算法(MGLP)。该算法通过加权的方式将类别级的 Label Smoothing 和 Online Label Smoothing、样本级的 Bootstrapping 3 种单粒度的标签扰动算法融合在一起，集成了类别级粒度和样本级粒度的标签扰动算法的特点，通过融合系数来控制不同粒度扰动的比例，并利用元学习的思想对融合系数进行学习，使其能够根据不同的数据特点自适应地对融合系数进行调整，避免了人工调参所造成的主观性误差，提高了模型的性能。本文在推文情感分类数据集、电影评论情感分类数据集、引文意图分类数据集 3 个公开的文本数据集上进行了实验，结果表明本文提出的 MGLP 算法与其他算法相比性能有明显的提升，能够有效地减轻噪声标签对深度学习模型训练的负面影响，对于深度学习模型在信息组织和信息分析领域更准确地应用具有十分重要的价值和前景。

本文也存在一些局限性。首先，只考虑了利用样本级粒度和类别级粒度的数据信息，还未研究与数据集级粒度信息的结合；其次，本文只在英文文本分类数据集上对算法的性能进行了验证。在未来研究中，将探究如何将样本级、类别级以及数据集级粒度的信息进行融合，以期进一步提高对多粒度信息的有效利用，进一步提升深度学习模型的性能，并将在中文数据集上对多粒度标签扰动算法的

性能进行探究。此外，还将探究利用更多的信息对融合系数进行求解。在算法未来的应用层面，除了将本文提出的算法应用于文本分类领域之外，还将探究其在更多领域中的应用，如计算机视觉领域，以期在多个领域中发挥该算法的应用价值。

参考文献

- [1] 周文远, 王名扬, 井钰. 基于 AttentionSBGMC 模型的引文情感和引文目的自动分类研究 [J]. 数据分析与知识发现, 2021, 5 (12): 48-59.
- [2] 罗卓然, 蔡乐, 钱佳佳, 等. 学术论文创新贡献句识别研究 [J]. 图书情报工作, 2021, 65 (12): 93-100.
- [3] 沈自强, 李晔, 丁青艳, 等. 基于 BERT 模型的科技政策文本分类研究 [J]. 数字图书馆论坛, 2022, (1): 10-16.
- [4] 谭炎, 张进, 夏立新. 社交媒体情境下的情感分析研究综述 [J]. 数据分析与知识发现, 2020, 4 (1): 1-11.
- [5] 吴雪华, 毛进, 陈思蓉, 等. 突发事件应急行动支撑信息的自动识别与分类研究 [J]. 情报学报, 2021, 40 (8): 817-830.
- [6] Northcutt C G, Athalye A, Mueller J. Pervasive Label Errors in Test Sets Destabilize Machine Learning Benchmarks [C] //Proceedings of the Neural Information Processing Systems (NIPS), 2021.
- [7] Wang Z, Shang J, Liu L, et al. Crossweigh: Training Named Entity Tagger from Imperfect Annotations [C] //Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP), 2019: 5153-5162.
- [8] Song H, Kim M, Park D, et al. Learning from Noisy Labels with Deep Neural Networks: A Survey [J]. IEEE Transactions on Neural Networks and Learning Systems, 2022.
- [9] Szegedy C, Vanhoucke V, Ioffe S, et al. Rethinking the Inception Architecture for Computer Vision [C] //Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016: 2818-2826.
- [10] Reed S, Lee H, Anguelov D, et al. Training Deep Neural Networks on Noisy Labels with Bootstrapping [C] //International Conference on Learning Representations, 2015.
- [11] Zhang C B, Jiang P T, Hou Q, et al. Delving Deep Into Label Smoothing [J]. IEEE Transactions on Image Processing, 2021, 30: 5984-5996.
- [12] 郭亚维, 刘晓霞. 文本分类中信息增益特征选择方法的研究 [J]. 计算机工程与应用, 2012, 48 (27): 119-122, 127.
- [13] 辛竹, 周亚建. 文本分类中互信息特征选择方法的研究与算法改进 [J]. 计算机应用, 2013, 33 (S2): 116-118, 152.
- [14] 杨萌萌, 黄浩, 程露红, 等. 基于 LDA 主题模型的短文本分类 [J]. 计算机工程与设计, 2016, 37 (12): 3371-3377.
- [15] Hochreiter S, Schmidhuber J. Long Short-term Memory [J]. Neural Computation, 1997, 9 (8): 1735-1780.
- [16] Vaswani A, Shazeer N, Parmar N, et al. Attention is All You

- Need [C] //Proceedings of the Neural Information Processing Systems (NIPS), 2017: 5998-6008.
- [17] 罗鹏程, 王一博, 王继民. 基于深度预训练语言模型的文献学科自动分类研究 [J]. 情报学报, 2020, 39 (10): 1046-1059.
- [18] 王佳敏, 吴乐艳, 李鹏程, 等. 学术文献致谢功能数据集构建与识别方法研究 [J]. 情报科学, 2021, 39 (11): 173-179.
- [19] Abdi A, Shamsuddin S M, Hasan S, et al. Deep Learning-based Sentiment Classification of Evaluative Text Based on Multi-feature Fusion [J]. Information Processing & Management, 2019, 56 (4): 1245-1259.
- [20] 马雨萌, 黄金霞, 王昉, 等. 融合BERT与多尺度CNN的科技政策内容多标签分类研究 [J/OL]. 情报杂志: 1-8 [2022-10-27].
- [21] Zong D, Sun S. GNN-XML: Graph Neural Networks for Extreme Multi-label Text Classification [J]. arXiv Preprint arXiv: 2012.05860, 2020.
- [22] 卢小宾, 张杨毅, 杨冠灿, 等. 新兴技术识别中的不平衡分类研究——基于代价敏感的随机森林算法 [J]. 情报学报, 2022, 41 (10): 1059-1070.
- [23] 刘美玲, 尚玥, 赵铁军, 等. 基于代价敏感学习的不平衡虚假评论处理模型 [J/OL]. 数据分析与知识发现: 1-13 [2023-03-25].
- [24] 线岩团, 陈文仲, 余正涛, 等. 融合类别先验 Mixup 数据增强的罪名预测方法 [J]. 自动化学报, 2022, 48 (8): 2097-2107.
- [25] 熊伟, 官禹. 基于元学习的不平衡少样本情况下的文本分类研究 [J]. 中文信息学报, 2022, 36 (1): 104-116.
- [26] 黄金凤, 高岩, 徐童, 等. 基于小样本数据增强的科技文档不平衡分类研究 [J]. 工程管理科技前沿, 2022, 41 (3): 23-30.
- [27] Vapnik V. Statistical Learning Theory [M]. New York: Wiley, 1998.
- [28] Algan G, Ulusoy I. Image Classification with Deep Learning in the Presence of Noisy Labels: A Survey [J]. Knowledge-Based Systems, 2021, 215: 106771.
- [29] Zhang J, Zhu J, Niu G, et al. Geometry-aware Instance-reweighted Adversarial Training [C] //International Conference on Learning Representations, 2021.
- [30] 李文娜, 张智雄. 基于置信学习的知识库错误检测方法研究 [J]. 数据分析与知识发现, 2021, 5 (9): 1-9.
- [31] 汪敏, 伍文静, 刘瀚阳, 等. 噪声标签识别与纠正的置信度预测方法 [J]. 西北大学学报 (自然科学版), 2022, 52 (5): 857-867.
- [32] Huang J, Qu L, Jia R, et al. O2u-net: A Simple Noisy Label Detection Approach for Deep Neural Networks [C] //Proceedings of the IEEE/CVF International Conference on Computer Vision, 2019: 3326-3334.
- [33] Zhou Y, Jiang J Y, Chang K W, et al. Learning to Discriminate Perturbations for Blocking Adversarial Attacks in Text Classification [C] //Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP), 2019: 4904-4913.
- [34] Liu P, Qiu X, Huang X J. Adversarial Multi-task Learning for Text Classification [C] //Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), 2017: 1-10.
- [35] Wang Y, Pan X, Song S, et al. Implicit Semantic Data Augmentation for Deep Networks [C] //Proceedings of the Neural Information Processing Systems (NIPS), 2019: 12614-12623.
- [36] Menon A K, Jayasumana S, Rawat A S, et al. Long-tail Learning Via Logit Adjustment [C] //International Conference on Learning Representations, 2021.
- [37] 张颖怡, 章成志. 基于学术论文全文的研究方法句自动抽取研究 [J]. 情报学报, 2020, 39 (6): 640-650.
- [38] Shu J, Xie Q, Yi L, et al. Meta-weight-net: Learning an Explicit Mapping for Sample Weighting [C] //Proceedings of the Neural Information Processing Systems (NIPS), 2019: 1917-1928.
- [39] Hospedales T, Antoniou A, Micaelli P, et al. Meta-learning in Neural Networks: A Survey [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2021, 44 (9): 5149-5169.
- [40] Nakov P, Ritter A, Rosenthal S, et al. SemEval-2016 Task 4: Sentiment Analysis in Twitter [J]. arXiv Preprint arXiv: 1912.01973, 2019.
- [41] Pang B, Lee L. Seeing Stars: Exploiting Class Relationships for Sentiment Categorization with Respect to Rating Scales [C] //Proceedings of the 43rd Annual Meeting of the Association for Computational Linguistics (ACL'05), 2005: 115-124.
- [42] Cohan A, Ammar W, Zuylen M, et al. Structural Scaffolds for Citation Intent Classification in Scientific Publications [C] //Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers), 2019: 3586-3596.
- [43] Shen Y, Xu L, Yang Y, et al. Self-distillation from the Last Mini-batch for Consistency Regularization [C] //Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2022: 11943-11952.
- [44] Liu B, Ben Ayed I, Galdran A, et al. The Devil is in the Margin: Margin-based Label Smoothing for Network Calibration [C] //Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2022: 80-88.

(责任编辑: 郭沫含)