

文章编号: 1003-0077(2020)04-0092-09

融合 CNN 和 EWC 算法的不平衡文本情绪分类方法

程 艳¹, 朱 海¹, 项国雄², 唐天伟³, 钟林辉¹, 王国玮¹

(1. 江西师范大学 计算机信息工程学院, 江西 南昌 330022;

2. 江西师范大学 新闻与传播学院, 江西 南昌 330022;

3. 江西师范大学 管理决策评价研究中心, 江西 南昌 330022)

摘 要: 文本情绪分类是自然语言处理领域的一个基本任务。然而, 基于不平衡数据的学习使得传统文本情绪分类方法的分类性能降低。针对这个问题, 该文提出了一种融合 CNN 和 EWC 算法的不平衡文本情绪分类方法。首先, 该方法使用随机欠采样方法得到多组平衡数据; 其次, 按顺序单独使用每一组平衡数据输入 CNN 训练, 同时在训练过程中引入 EWC 算法用以克服 CNN 中的灾难性遗忘; 最后, 把使用最后一组平衡数据输入 CNN 训练得到的模型作为最终分类模型。实验结果表明, 该方法在分类性能上明显优于基于欠采样和多分类算法的集成学习框架, 且该方法比基于多通道 LSTM 神经网络的不平衡情绪分类方法在 Accuracy 和 G-mean 上分别提高了 1.9% 和 2.1%。

关键词: 情绪分类; 不平衡分类; CNN; EWC 算法

中图分类号: TP391

文献标识码: A

Emotion Classification Based on CNN and EWC Algorithm for Unbalanced Texts

CHENG Yan¹, ZHU Hai¹, XIANG Guoxiong², TANG Tianwei³, ZHONG Linhui¹, WANG Guowei¹

(1. School of Computer Information Engineering, Jiangxi Normal

University, Nanchang, Jiangxi 330022, China;

2. School of Journalism and Communication, Jiangxi Normal University, Nanchang, Jiangxi 330022, China;

3. Management Decision Evaluation Research Center, Jiangxi Normal
University, Nanchang, Jiangxi 330022, China)

Abstract: Text emotion classification is a well-addressed task in the field of natural language processing. To deal with the unbalanced data which hurt the classification performance, this paper proposes an emotion classification method combining CNN and EWC algorithms. First, the method uses the random under-sampling method to obtain multiple sets of balanced data for training. Then it feeds each balanced dataset to CNN training in sequence, introducing EWC algorithm in the training process to overcome the catastrophic forgetting issue in CNN. Finally, the CNN model trained by the last data set is treated as the final classification model. The experimental results show that the proposed method is superior to the ensemble learning framework based on under-sampling and multi-classification algorithms, and outperforms the multi-channel LSTM neural network with 1.9% and 2.1% improvements in accuracy and G-mean, respectively.

Keywords: emotion classification; imbalanced classification; CNN; EWC algorithm

0 引言

随着微博等网上社交平台的迅猛发展, 数以

亿计的网民在互联网上通过文字和他人进行交流, 从而使得互联网上充斥着大量的带有发表者个人情绪的文本数据。对这些文本数据中的情绪进行分析可以帮助研究机构、信息咨询组织和政

收稿日期: 2019-08-22 定稿日期: 2019-11-28

基金项目: 国家自然科学基金(61967011); 江西省科技重点项目(20161BBE50086); 江西省教育厅科技重点项目(GJJ150299); 江西省教育厅人文社科重点(重大)项目(文)(JD19056); 江西省教育厅科学技术项目(GJJ170207)

府决策部门掌握社会情绪动态^[1]。因此,文本情绪分析受到了自然语言处理领域研究者的密切关注,现已成为自然语言处理领域的一个研究热点。

现有文本情绪分析工作主要集中在文本情绪分类的研究上。文本情绪分类任务主要指通过提取文本内容中的情绪要素,将文本划分到预定义的情绪类别中。目前,情绪分类方法主要包括:基于词典

和规则的方法、基于机器学习的方法、复合方法以及其他方法^[1]。其中,基于机器学习的方法又分为有监督学习方法和无监督学习方法。有监督学习方法的主要思想是使用训练数据训练分类器得到最终分类模型,再使用测试数据评估最终分类模型的性能。训练数据和测试数据的情绪类别一般都是通过专家进行人为标注。表 1 给出了一些文本数据及其对应的情绪类别。

表 1 文本数据及其对应的情绪类别

文本数据	情绪类别
我的第六本书今天拿到清样了,是一本专门写投资的书,书的内容涵盖股票、基金、房地产和实业。哈哈,今天很开心。	高兴
今天在论坛上看到的韩国整容手术的广告,这太神奇了!眼睛、鼻子、嘴巴……就连发际线都整了!	惊讶
走着走着,就散了,回忆都淡了;看着看着,就累了,星光也暗了;听着听着,就醒了,开始埋怨了;回头发现,你不见了,突然我乱了。	悲伤

目前大多数传统有监督学习方法都是基于训练数据是平衡数据(各个情绪类别样本分布平衡)的假设。当这个假设成立时应用这些方法一般可以取得良好的分类效果。然而,当训练数据是不平衡数据(各个情绪类别样本的分布不平衡)时,应用这些方法得到的分类结果严重偏向样本数量较多的情绪类别,忽略样本数量较少的情绪类别,从而使得分类性能降低^[2]。而在实际收集到的文本数据中,各个情绪类别样本的分布往往会非常不平衡^[3]。因此,对文本情绪分类任务中的不平衡数据分类问题进行研究是至关重要的。

不平衡数据分类问题是机器学习领域的一个经典问题。针对不平衡数据分类问题,目前主流的数据预处理方法是欠采样方法。该方法的主要思想是通过减少样本数量较多的类别中的样本,使得不平衡数据在一定程度上达到平衡状态,从而消除不平衡数据分类问题^[4]。但是由于该方法使得样本数量较多的类别中的样本只有部分参与训练,从而丢失了很多可能对分类有帮助的样本^[3]。为了充分利用已有样本,本文提出了一种融合 CNN(convolutional neural networks)和 EWC(elastic weight consolidation)算法的不平衡文本情绪分类方法,用来解决文本情绪分类任务中的不平衡数据分类问题。CNN 可以提出句子中类似 n -gram 的关键信息,能够更好地捕捉局部相关性,从而有助于提高分类性能。具体而言,首先,该方法使用随机欠采样方法得到多组平衡数据;其次,按顺序单独把每一组平衡数据输入 CNN 训练,同

时在训练过程中引入 EWC 算法,用以克服 CNN 中的灾难性遗忘;最后,把使用最后一组平衡数据输入 CNN 训练得到的模型作为最终分类模型。实验结果表明,该方法在分类性能上明显优于基于欠采样和多分类算法的集成学习框架,且该方法比基于多通道 LSTM(long short-term memory)神经网络的不平衡情绪分类方法在 Accuracy 和 G-mean 上分别提高了 1.9% 和 2.1%。

1 相关工作

1.1 基于 CNN 的文本情绪分类

CNN 特有的卷积、池化结构能够提取图像中各种不同程度的纹理、结构。基于 CNN 的这个特点,将 CNN 应用到计算机视觉领域的图像分类任务中,在分类准确率上取得巨大突破。例如, Alex Krizhevsky 等^[5]提出了 Alexnet 模型,并将其应用到图像分类任务,以极大的优势(在分类准确率上远远超过第二名)赢得了 2012 年度 ILSVRC 挑战冠军。之后, Yoon Kim^[6]提出了 TextCNN 模型,并将其应用于自然语言处理领域中的文本分类任务,取得了良好的分类效果。在此基础上,自然语言处理领域研究者们提出了一些基于 CNN 的文本情绪分类方法。例如, Ouyang Xi 等^[7]提出了一种基于 CNN 的情绪分类架构,该架构使用 Word2Vec 方法得到文本的词向量表示,再将文本的词向量输入 CNN 训练,在 5 类情绪分类任务上均取得了满意的

效果。李德玉等^[8]提出了一种基于标签特征的 CNN 多标签情绪分类方法,该方法通过融合 CNN 和标签特征,在文本情绪多标签分类任务上取得了良好的分类效果。

然而,基于 CNN 的文本情绪分类方法有一个很明显的缺点:基于不平衡数据的学习使得分类性能降低。

1.2 不平衡文本情绪分类

目前大部分已有的文本情绪分类方法研究都是基于训练数据是平衡数据的假设^[9-13]。由于这些方法没有考虑到训练数据是不平衡数据的情况,并且认为各个情绪类别之间的误分类代价相同,从而导致不平衡文本情绪分类的效果不理想^[4]。基于不平衡数据的文本情绪分类方法研究还很缺乏^[3]。林怀逸等^[14]提出了一种基于词向量预训练的不平衡文本情绪分类方法,该方法将词向量预训练和均衡过采样方法相结合,在大部分无严重过拟合情况下取得了比过采样方法更好的平衡效果。虽然过采样方法可以消除不平衡数据分类问题,但是过采样方法使得数据出现冗余,模型训练复杂度增大,而且容易造成模型过拟合问题^[4]。王中卿等^[2]提出了一种基于欠采样和多分类算法的集成学习框架,该方法采用欠采样方法消除不平衡数据分类问题,避免了过采样方法带来的问题。同时该方法采用集成学习方法充分利用已有样本,避免了欠采样方法可能丢失对分类有帮助的样本的问题。虽然该方法取得了优于传统不平衡分类方法的分类性能,但是该方法使用朴素贝叶斯、最大熵和支持向量机等三种分类算法构建基分类器,而已有研究表明,在情绪分类任务上神经网络方法的分类性能要优于传统机器学习方法^[15]。殷浩等^[3]提出一种基于多通道 LSTM 神经网络的不平衡情绪分类方法,该方法采用随机欠采样方法消除不平衡数据分类问题,同时采用多通道神经网络方法充分利用已有样本,取得了优于前者的分类性能。但是融合多个神经网络模型使得模型变得更加复杂,大量参数的学习受限于训练数据的规模,当训练数据不足时,常常出现过拟合现象^[16]。为了降低神经网络模型的复杂度,又能充分利用已有样本,一种可行的方案是使用单通道神经网络连续学习多组平衡数据的知识。但这样做使得单通道神经网络遭受灾难性遗忘。

1.3 灾难性遗忘问题

M McCloskey 等^[17]描述了神经网络中的灾难

性遗忘问题,即使用新数据输入神经网络训练会覆盖神经网络在旧数据上所学的知识。之后,灾难性遗忘问题开始进入人们的视野,并被普遍认为是连接主义模型不可避免的特征。早期用于解决灾难性遗忘问题的两个主要策略分别是冻结/增长(通过冻结部分网络权重,同时增长网络以保持学习能力)策略^[18]和预演(不仅用最新数据,而且用早期数据不断刺激网络)策略^[19]。虽然这两个策略可以在一定程度上减轻神经网络中的灾难性遗忘,但是当使用新数据输入神经网络训练时遵循这两个策略的方法在旧数据上的分类性能远非理想。为了进一步减轻神经网络中的灾难性遗忘,James Kirkpatrick 等^[20]受神经科学中突触巩固的启发提出了 EWC 算法,实验结果表明,该算法允许深层神经网络按顺序学习一组复杂的分类任务,而不会出现灾难性遗忘问题。

本文主要的贡献是在前人的工作基础上提出了一种融合 CNN 和 EWC 算法的分类方法,用来解决文本情绪分类任务中的不平衡数据分类问题。具体而言,在 Ouyang Xi 等的工作基础上^[7],本文选择采用 CNN 作为分类器。在殷昊等的工作基础上^[3],本文遵循采用随机欠采样方法消除不平衡数据分类问题的同时充分利用已有样本的策略,不同的是本文使用随机欠采样方法得到多组平衡数据,再按顺序单独把每一组平衡数据输入单通道 CNN 训练。但是这样使得单通道 CNN 遭受灾难性遗忘。在 James Kirkpatrick 等的工作基础上^[20],本文在训练过程中引入 EWC 算法,用以克服单通道 CNN 中的灾难性遗忘。

2 融合 CNN 和 EWC 算法的不平衡文本情绪分类方法

2.1 基于 CNN 的文本分类模型

Yoon Kim 提出了 TextCNN^[6],使得 CNN 被广泛应用于自然语言处理领域。TextCNN 模型如图 1 所示,该模型由一个输入层、一个卷积层、一个池化层和一个全连接层组成。具体而言,输入层为 $n \times k$ 的词向量矩阵,其中, n 表示输入向量长度, k 表示词向量的维度。输入层是双通道的,一个通道是静态词向量,另一个通道是动态词向量;卷积层使用若干个窗口大小分别为 $2 \times k$ 、 $3 \times k$ 和 $4 \times k$ 的卷

积核对输入层的词向量矩阵进行卷积操作,再使用激活函数激活得到与卷积核相同个数的 feature maps;池化层采用 Max-over-time pooling 方法提取

出卷积层的 feature maps 的最大值;全连接层对池化层的输出进行加权求和后,再使用激活函数激活得到模型的输出。

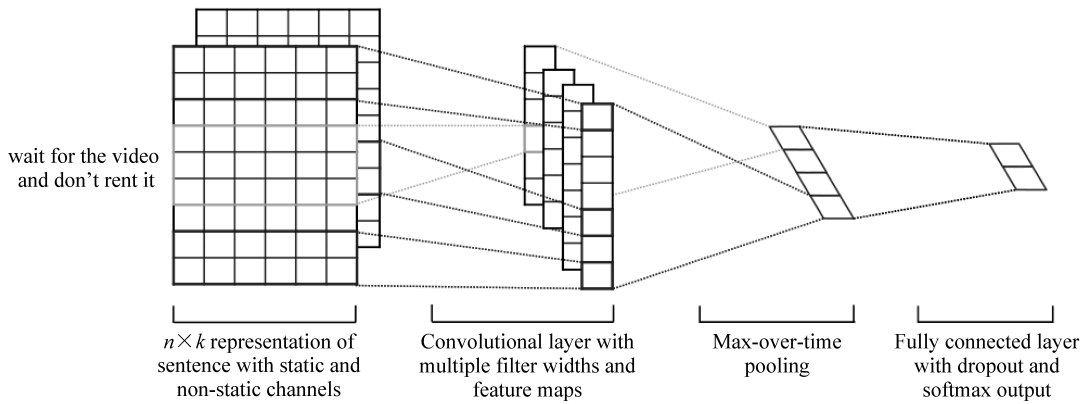


图 1 TextCNN 模型图

2.2 基于单通道 CNN 的不平衡文本情绪分类方法

基于单通道 CNN 的不平衡文本情绪分类方法框架如图 2 所示,第一个虚线框部分为不平衡数据的数据预处理过程,第二个虚线框部分为 CNN 模型内部结构。首先,该方法使用随机欠采样方法从不平衡数据中得到一组平衡数据。其次,使用该组平衡数据输入 CNN 训练。在 CNN 模型中,输入层为该组平衡数据对应的词向量表示。卷积层使用若

干个不同大小的卷积核对输入层的词向量表示进行卷积操作,卷积操作可以提取出句子中类似于 n -gram 的关键信息,从而能够更好地捕捉局部相关性,再使用激活函数激活得到与卷积核数量相同的 feature maps。池化层采用 Max-over-time pooling 方法提取出卷积层的 feature maps 的最大值,池化层得到的高维向量可以更好地描述文本数据。全连接层对池化层的高维向量进行加权求和,再经过激活函数激活得到 CNN 模型的输出。

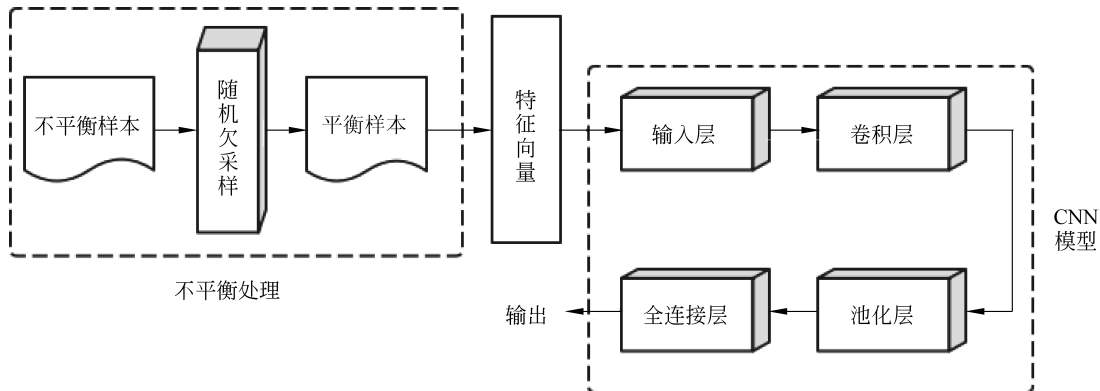


图 2 基于单通道 CNN 的不平衡文本情绪分类方法框架图

卷积层使用 Relu 函数作为激活函数,Relu 函数可以使得神经网络变得稀疏,从而缓解神经网络模型过拟合问题,Relu 函数如式(1)所示。

$$g(\mathbf{x}) = \max(0, \mathbf{x}) \quad (1)$$

其中, \mathbf{x} 为输出向量。

对池化层的输出使用 Dropout 方法,Dropout 方法可以使得神经网络学习那些在神经元的不同随机子集中更加健壮的特征,从而减轻神经网络模型

过拟合,Dropout 方法如式(2)所示。

$$g = h * \cdot D(p) \quad (2)$$

其中, D 表示 Dropout 操作符, p 为一个可调的超参数(保留隐层单元的比率)。

全连接层使用 Softmax 函数作为激活函数,Softmax 函数可以把神经网络模型的输出解释为概率,Softmax 函数如式(3)所示。

$$\text{label}_{\text{pred}} = \arg \max_i P(Y = i | \mathbf{x}, \mathbf{W}, \mathbf{b}) \quad (3)$$

其中, i 为标签预测值, x 为上一层输出向量, W, b 为 CNN 模型中的系数矩阵, $label_{pred}$ 为后验概率最大的预测标签。

2.3 EWC 算法

James Kirkpatrick 等^[20]受神经科学中突触巩固的启发提出了 EWC 算法, 该算法可以使得深层神经网络在按顺序学习一组分类任务后旧任务的准确率只是略有下降。由于深层神经网络存在过度参数化现象(参数的许多不同配置将导致相同的分类性能), 从而使得可能有一种参数的配置能够同时在不同的任务上都取得良好的分类效果。EWC 算法通过减缓在神经网络中那些对旧任务很重要的权重的学习, 试图找到这种能够同时在不同的任务上取得良好的分类效果的参数配置。EWC 算法如式(4)所示。

$$C = C_B + \sum_i \frac{\lambda}{2} F_i (\theta_i - \theta_{A,i})^2 \quad (4)$$

其中, C_B 为任务 B (新任务) 的损失函数, λ 为一个可调的超参数, 表示任务 A (旧任务) 相较于任务 B 的重要程度, F_i 为费希尔信息矩阵的第 i 个对角线元素, θ_i 为训练任务 B 时神经网络模型中的第 i 个参数, $\theta_{A,i}$ 为训练任务 A 时得到的神经网络模型中的第 i 个参数。

2.4 融合单通道 CNN 和 EWC 算法的不平衡文本情绪分类方法

基于单通道 CNN 的不平衡文本情绪分类方法有一个明显的缺点: 丢失了很多可能对分类有帮助的样本。为了充分利用已有样本, 本文提出了融合单通道 CNN 和 EWC 算法的不平衡文本情绪分类方法。该方法框架如图 3 所示, 下部虚线框部分为不平衡数据的数据预处理过程, 上部虚线框部分为 CNN 模型内部结构。首先, 该方法使用随机欠采样方法从不平衡数据中得到多组平衡数据; 其次, 按顺序单独把每一组平衡数据输入 CNN 训练, 同时在训练过程中引入 EWC 算法, 用以克服 CNN 中的灾难性遗忘。在 CNN 模型中, 输入层为每一组平衡数据对应的词向量表示。卷积层使用若干个不同大小的卷积核对输入层的词向量表示进行卷积操作, 再使用 Relu 函数激活得到与卷积核数量相同的 feature maps。池化层采用 Max-over-time pooling 方法提取出卷积层的 feature maps 的最大值, 再对最大值使用 Dropout 方法。全连接层对池化层的输

出进行加权求和, 再经过 Softmax 函数激活得到 CNN 模型的输出。

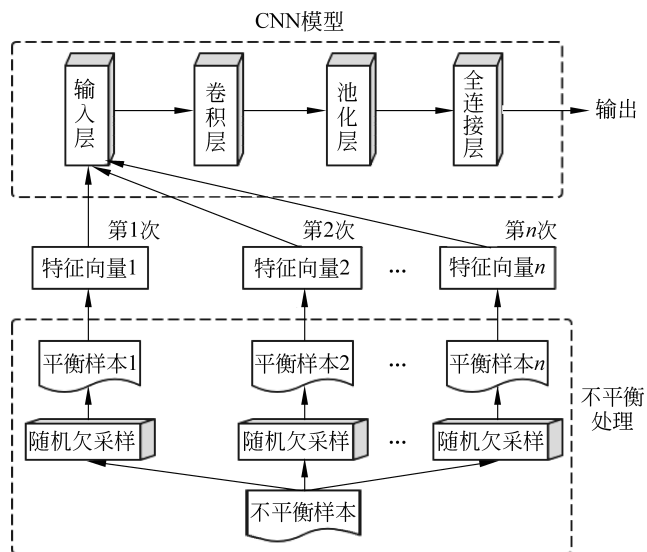


图3 融合单通道 CNN 和 EWC 算法的不平衡数据情绪分类方法框架图

在训练过程中采用交叉熵损失函数作为损失函数, 交叉熵损失函数如式(5)所示。

$$C = -\frac{1}{n} \sum_x (y \ln a + (1 - y) \ln(1 - a)) \quad (5)$$

其中, n 为训练集的样本大小, y 为训练数据的标签向量, a 为神经网络模型的输出向量。

在损失函数后面增加 L2 正则化项, L2 正则化方法可以使得神经网络的参数值变得更小, 从而减轻神经网络模型过拟合, L2 正则化项如式(6)所示。

$$L2 = \frac{\lambda}{2n} \sum_w w^2 \quad (6)$$

其中, λ 为规范化系数, n 为训练集的样本大小, w 为神经网络模型中的系数。

在第 1 次训练时, 损失函数为交叉熵损失函数和 L2 正则化项的和, 损失函数如式(7)所示。

$$C = C_0 + \frac{\lambda}{2n} \sum_w w^2 \quad (7)$$

其中, 第 1 项为交叉熵损失函数, 第 2 项为 L2 正则化项。

从第 2 次训练开始直到第 n 次训练结束, 损失函数为交叉熵损失函数、L2 正则化项和 EWC 算法项的和, 损失函数如式(8)所示。

$$C = C_0 + \frac{\lambda}{2n} \sum_w w^2 + \sum_i \frac{\lambda}{2} F_i (\theta_i - \theta_{A,i})^2 \quad (8)$$

其中, 第 1 项为交叉熵损失函数, 第 2 项为 L2 正则化项, 第 3 项为 EWC 算法项。

3 实验

3.1 实验设置

NLP&CC-2013 中文微博情绪分析评测任务提供的微博语料共有喜好、高兴、厌恶、悲伤、愤怒、惊讶和恐惧等 7 种情绪类别。在该数据中,各个情绪

类别的样本数量相差巨大。因此,本文选用该数据作为实验数据。由于该数据中恐惧情绪类别的样本数量太少,根据其样本数量生成的测试集得到的实验结果具有较大的偶然性,所以本文选取该数据中的惊讶情绪类别的样本数量的 20%(即 $362 \times 20\% \approx 72$)作为各个情绪类别的测试样本数。各个情绪类别的测试样本数和剩余样本数如表 2 所示。

表 2 各个情绪类别测试样本数和剩余样本数

情绪类别	喜好	高兴	厌恶	悲伤	愤怒	惊讶	恐惧
测试样本	72	72	72	72	72	72	72
剩余样本	2 131	1 388	1 320	1 101	597	290	76

在数据预处理部分,本文采用分词工具 FudanNLP 对文本数据进行分词。本文使用 Python 工具包 gensim 来生成文本数据的词向量模型。综合考虑实验性能和所需时间,本实验中词向量维度设为 100。CNN 使用深度学习开源框架 Tensorflow 搭建。CNN 模型的具体参数设置如表 3 所示。

表 3 CNN 模型的参数设置

参数表述	参数值
输入向量长度	120
卷积核大小	[2,3,4]
卷积核个数	128
Dropout 参数	0.5
迭代次数	20

实验采用正确率(Accuracy)、召回率(Recall)和几何平均数(G-mean)作为衡量分类效果的标准。正确率、召回率和几何平均数的算式分别如式(9)~式(11)所示。

$$\text{Accuracy} = \frac{\sum_{i=1}^n \text{TP}_i}{\sum_{i=1}^n \text{TP}_i + \sum_{i=1}^n \text{FP}_i} \quad (9)$$

其中, TP_i 为在类别 i 的全部测试集中模型预测正确的样本数量, FP_i 为在类别 i 的全部测试集中模型预测错误的样本数量, n 为类别个数。

$$\text{Recall}_i = \frac{\text{TP}_i}{\text{TP}_i + \text{FP}_i} \quad (10)$$

其中, TP_i 为在类别 i 的全部测试集中模型预测正确的样本数量, FP_i 为在类别 i 的全部测试集

中模型预测错误的样本数量。

$$\text{G-mean} = \sqrt[n]{\prod_{i=1}^n \text{Recall}_i} \quad (11)$$

其中, Recall_i 为类别 i 的召回率, n 为类别个数。

3.2 实验结果

为了进行充分比较,本文选取了以下几种基于不平衡数据的文本情绪分类方法。

(1) 完全训练+单通道 LSTM 神经网络,把各个情绪类别的剩余样本作为训练数据,使用单通道 LSTM 神经网络^[3]进行分类。超参数设置与文献[3]相同。

(2) 随机过采样+单通道 LSTM 神经网络,使用随机过采样方法从各个情绪类别的剩余样本中得到一组平衡数据(各个情绪类别的样本数量都为 2 131 个)作为训练数据,使用单通道 LSTM 神经网络^[3]进行分类。超参数设置与文献[3]相同。

(3) 随机欠采样+单通道 LSTM 神经网络,使用随机欠采样方法从各个情绪类别的剩余样本中得到一组平衡数据(各个情绪类别的样本数量都为 290 个)作为训练数据,使用单通道 LSTM 神经网络^[3]进行分类。超参数设置与文献[3]相同。

(4) 随机欠采样+集成学习,多次使用方法(3)中的随机欠采样方法从各个情绪类别的剩余样本中得到 5 组平衡数据作为训练数据,选择最大熵作为基分类器,使用集成学习^[2]方法进行分类。

(5) 随机欠采样+多通道 LSTM 神经网络,多次使用方法(3)中的随机欠采样方法从各个情绪类别的剩余样本中得到 5 组平衡数据作为训练数据,

使用多通道 LSTM 神经网络^[3]进行分类。超参数设置与文献[3]相同。

(6) 随机欠采样+单通道 CNN+多次训练+EWC 算法(本文方法),多次使用方法(3)中的随机欠采样方法从各个情绪类别的剩余样本中得到 5 组平衡数据作为训练数据,使用本文方法进行分类。

(7) 随机欠采样+单通道 CNN,使用方法(3)中的随机欠采样方法从各个情绪类别的剩余样本中得到 1 组平衡数据作为训练数据,使用单通道 CNN^[3]进行分类。

(8) 随机欠采样+单通道 CNN+多次训练,多次使用方法(3)中的随机欠采样方法从各个情绪类别的剩余样本中得到 5 组平衡数据作为训练数据,使用单通道 CNN^[3]进行分类(按顺序单独使用每一组平衡数据输入 CNN 训练)。

(9) 完全训练+单通道 CNN,把各个情绪类别的剩余样本作为训练数据,使用单通道 CNN^[3]进行分类。

图 4 比较了方法(1)、方法(2)和方法(3)的分类性能。其中,方法(3)的分类性能取自文献[3]中的实验结果。从图 4 可以看出,方法(1)的分类性能最差。其主要原因是完全训练使得应用单通道 LSTM 神经网络得到的分类结果严重偏向样本数量较多的情绪类别,从而使得样本数量较少的情绪类别的召回率非常低(甚至为 0);方法(3)比方法(2)在 Accuracy 和 G-mean 上分别提高了 2.4%和 8.6%。其主要原因是随机过采样方法使得数据出现冗余,模型训练复杂度增大,而且容易造成模型过拟合问题。图 4 实验结果表明,随机欠采样方法和随机过采样方法都能够在一定程度上消除不平衡情

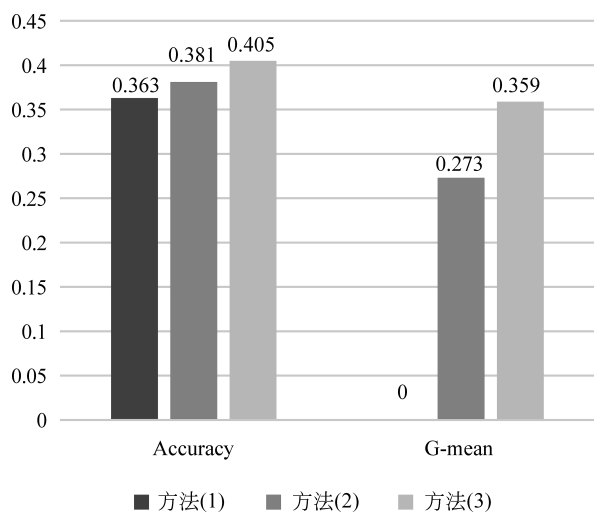


图 4 传统不平衡分类方法分类性能比较

绪分类问题。其中,随机欠采样方法在实验数据集上的分类性能要优于随机过采样方法。

图 5 比较了方法(4)、方法(5)和本文方法的分类性能。其中,方法(4)和方法(5)的分类性能取自文献[3]中的实验结果。从图 5 可以看出,本文方法比方法(4)在 Accuracy 和 G-mean 上分别提高了 3.4%和 4.9%。其主要原因是方法(4)采用传统机器学习方法作为分类器,本文方法采用 CNN 方法作为分类器。与传统机器学习方法相比较,CNN 可以提取出语句中类似 n -gram 的关键信息,能够更好地捕捉局部相关性,从而有助于提高分类性能;本文方法比方法(5)在 Accuracy 和 G-mean 上分别提高了 1.9%和 2.1%。其主要原因是方法(5)采用多通道神经网络作为分类器,本文方法采用单通道神经网络作为分类器。与多通道神经网络相比较,单通道神经网络模型的参数更少,可以避免由于模型参数过多导致的模型过拟合问题。图 5 实验结果表明,本文方法对不平衡情绪分类非常有效,该方法在实验数据集上的分类性能要优于传统不平衡情绪分类方法。

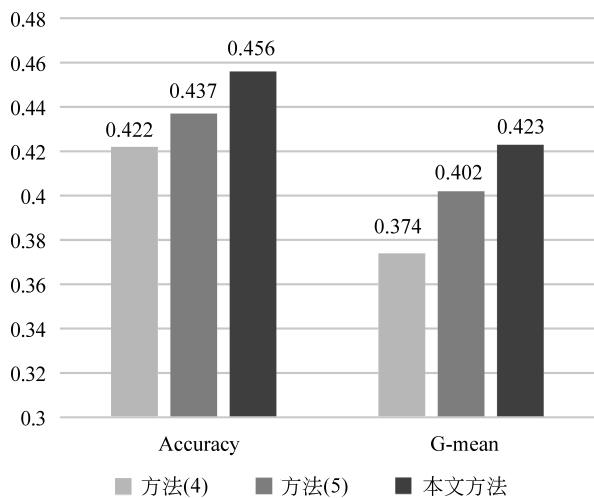


图 5 本文方法和传统不平衡情绪分类方法的分类性能比较

图 6 比较了方法(7)、方法(8)和本文方法的分类性能。其中,方法(7)的分类性能取自文献[3]中的实验结果。从图 6 可以看出,方法(7)和方法(8)的分类性能相当。其主要原因是使用新数据输入 CNN 训练会覆盖其在旧数据上所学的知识;本文方法比方法(7)在 Accuracy 方面和 G-mean 方面分别提高了 5.6%和 6.1%。其主要原因是方法(7)只学习了 1 组平衡数据,本文方法学习了 5 组平衡数据。与方法(7)相比较,本文方

法能够更加充分利用已有样本；本文方法比方法(8)在 Accuracy 方面和 G-mean 方面分别提高了 4.9% 和 5.7%。其主要原因是本文方法引入了 EWC 算法避免了 CNN 中的灾难性遗忘。图 6 实验结果表明，引入 EWC 算法克服 CNN 中的灾难性遗忘在本文方法中起到了关键性的作用，并且本文方法在实验数据上的分类性能要优于基于单通道 CNN 的不平衡情绪分类方法。

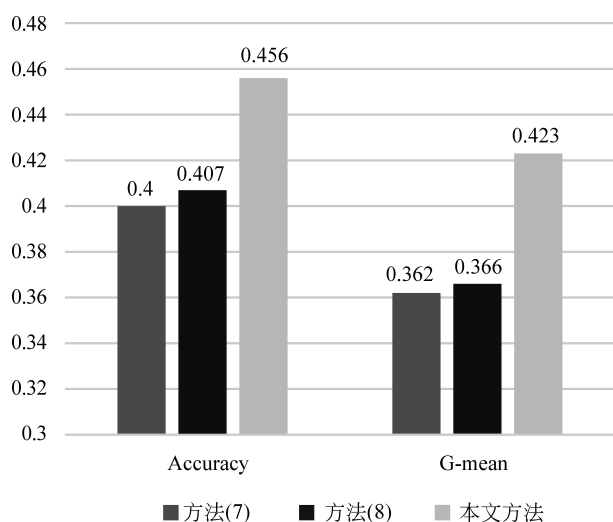


图 6 基于 CNN 的不平衡情绪分类方法分类性能比较

表 4 给出了方法(9)、方法(7)和本文方法在各个情绪类别上的召回率。从表 4 可以看出，方法(9)

在惊讶和恐惧情绪类别上的召回率分别为 4.2% 和 0，其在喜好情绪类别上的召回率为 77.8%。从整体上看，方法(9)在样本数量越多的情绪类别上的召回率越高。其主要原因是完全训练使得应用单通道 CNN 神经网络得到的分类结果严重偏向样本数量较多的情绪类别，从而使得样本数量较少的情绪类别的召回率较低；方法(7)比方法(9)在惊讶和恐惧情绪类别上的召回率分别提高了 21.4% 和 13.9%，其比方法(9)在喜好情绪类别上的召回率降低了 20.6%。主要原因是随机欠采样方法能够在一定程度上消除不平衡情绪分类问题，从而提高了样本数量较少的情绪类别的召回率。但由于丢失了很多样本数量较多的情绪类别的样本，从而降低了样本数量较多的情绪类别的召回率；本文方法在惊讶和恐惧情绪类别上的召回率分别为 30.3% 和 17.2%，其在喜好情绪类别上的召回率为 62.8%。从整体上看，本文方法在样本数量越多的情绪类别上的召回率越高。其可能的原因是，在本文方法中样本数量越少的情绪类别样本的重复率越高，越容易使得模型学习到的信息过于特别而不够泛化。表 4 实验结果表明，与方法(9)相比较，本文方法不仅能够提高样本数量较少的情绪类别的召回率，而且能在一定程度上保持样本数量较多的情绪类别的召回率；与方法(7)相比较，本文方法在各个情绪类别上的召回率均有所提高。

表 4 基于 CNN 的不平衡情绪分类方法在各个情绪类别上的召回率

情绪类别	喜好	高兴	厌恶	悲伤	愤怒	惊讶	恐惧
方法(9)	0.778	0.472	0.458	0.423	0.373	0.042	0
方法(7)	0.572	0.542	0.5	0.488	0.304	0.256	0.139
本文方法	0.628	0.549	0.541	0.532	0.469	0.303	0.172

本文方法还存在以下不足之处：

(1) 由于在使用随机欠采样方法得到的多组平衡数据中存在冗余数据，从而容易造成模型过拟合问题。另外，如果在冗余数据中存在“噪声”数据，就会放大“噪声”数据对模型的影响。

(2) 由于使用随机欠采样方法得到的每一组平衡数据的样本数量较少，从而容易造成模型过拟合问题。

4 结语

针对文本情绪分类任务中的不平衡数据分类问题，本文提出了一种融合 CNN 和 EWC 算法的不平

衡文本情绪分类方法。首先，该方法使用随机欠采样方法从不平衡数据中获取多组平衡数据；其次，按顺序单独把每一组平衡数据输入 CNN 训练，同时在训练过程中引入 EWC 算法以克服 CNN 中的灾难性遗忘；最后，把使用最后一组平衡数据输入 CNN 训练得到的模型作为最终分类模型。实验结果表明，该方法在分类性能上明显优于基于欠采样和多分类算法的集成学习框架，且该方法比基于多通道 LSTM 神经网络的不平衡情绪分类方法在 Accuracy 和 G-mean 上分别提高了 1.9% 和 2.1%。

在下一步工作中，我们将在其他领域的不平衡数据上评估本文方法，验证本文方法的有效性；我们将探索如何过滤训练数据中的“噪声”数据，从而降

低“噪声”数据对模型的影响;我们将探索如何改进 EWC 算法,使得当使用新数据输入神经网络训练时神经网络能够保留更多其在旧数据上所学的知识。今后,我们将主要围绕上述问题展开工作,以此找到分类性能更佳的不平衡文本情绪分类方法。

参考文献

- [1] 李然,林政,林海伦,等.文本情绪分析综述[J].计算机研究与发展,2018,55(01): 30-52.
- [2] 王中卿,李寿山,朱巧明,等.基于不平衡数据的中文情感分类[J].中文信息学报,2012,26(03): 33-37,64.
- [3] 殷昊,李寿山,贡正仙,等.基于多通道 LSTM 的不平衡情绪分类方法[J].中文信息学报,2018,32(01): 139-145.
- [4] 赵楠,张小芳,张利军.不平衡数据分类研究综述[J].计算机科学,2018,45(S1): 22-27,57.
- [5] Krizhevsky A, Sutskever I, Hinton G. ImageNet classification with deep convolutional neural networks[C]//Proceedings of Advances in Neural Information Processing Systems, 2012: 1097-1105.
- [6] Kim Y. Convolutional neural networks for sentence classification[J]. arXiv preprint arXiv: 1408.5882, 2014.
- [7] Xi Ouyang, Pan Zhou, Cheng Hua Li, et al. Sentiment analysis using convolution neural network[C]//Proceedings of 2015 IEEE International Conference on Computer and Information Technology, Ubiquitous Computing and Communications, Dependable, Automatic and Secure Computing, IEEE, 2015: 2539-2564.
- [8] 李德玉,罗锋,王素格.融合 CNN 和标签特征的中文文本情绪多标签分类[J].山西大学学报(自然科学版), 2019: 1-12.
- [9] 张璐,沈忱林,李寿山.基于情绪特定词向量的情绪分类算法[J].计算机科学,2019,46(S1): 93-97.
- [10] 郝苗苗,徐秀娟,于红,等.基于中文微博的情绪分类与预测算法[J].计算机应用,2018,38(S2): 89-96.
- [11] 黄磊,李寿山,周国栋.基于句法信息的微博情绪识别方法研究[J].计算机科学,2017,44(02): 244-249.
- [12] 温雯,吴彪,蔡瑞初,等.基于多类别语义词簇的新闻读者情绪分类[J].计算机应用,2016,36(08): 2076-2081.
- [13] 乌达巴拉,汪增福.一种基于组合语义的文本情绪分析模型[J].自动化学报,2015,41(12): 2125-2137.
- [14] 林怀逸,刘箴,柴玉梅,等.基于词向量预训练的不平衡文本情绪分类[J].中文信息学报,2019,33(05): 132-142.
- [15] Tripathi S, Acharya S, Sharma R D, et al. Using deep and convolutional neural networks for accurate emotion classification on DEAP dataset[C]//Proceedings of the 29th IAAI Conference, 2017.
- [16] 周安众,罗可.一种卷积神经网络的稀疏性 Dropout 正则化方法[J].小型微型计算机系统,2018,39(08): 1674-1679.
- [17] McCloskey M, Cohen N J. Catastrophic interference in connectionist networks: The sequential learning problem[M]. Psychology of Learning and Motivation, Academic Press, 1989, 24: 109-165.
- [18] T Xiao, J Zhang, K Yang, et al. Error driven incremental learning in deep convolutional neural network for large-scale image classification[C]//Proceedings of the 22nd ACM International Conference on Multimedia, ACM, 2014: 177-186.
- [19] Rebuffi S, Kolesnikov A, Sperl G, et al. iCaRL: Incremental classifier and representation learning[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017: 2001-2010.
- [20] Kirkpatrick J, Pascanu R, Rabinowitz N, et al. Overcoming catastrophic forgetting in neural networks[J]. Proceedings of the National Academy of Sciences, 2017, 114(13): 3521-3526.



程艳(1976—),通信作者,博士,教授,主要研究领域为情感分析、机器学习。

E-mail: chyan88888@jxnu.edu.cn



项国雄(1965—),本科,教授,主要研究领域为信息技术、教育信息化。

E-mail: 11840088@qq.com



朱海(1994—),硕士研究生,主要研究领域为自然语言处理、机器学习。

E-mail: zhuhai1994@jxnu.edu.cn