

基于多通道 LSTM 的不平衡情绪分类方法*

殷昊, 李寿山, 贡正仙, 周国栋

苏州大学自然语言处理实验室, 苏州, 215006

E-mail: hyin@stu.suda.edu.cn, {lishoushan, zhxgong, gdzhou}@suda.edu.cn

摘 要: 情绪分类方法是自然语言处理问题中的重要研究问题之一。情绪分类旨在对文本包含的情绪进行自动分类, 该任务是情感分析的一项基本任务。然而, 已有的研究都假设各情绪类别的样本数量平衡, 这与实际情况并不相符合。本文的研究主要面向不平衡数据的情绪分类任务。具体而言, 本文提出了一种基于多通道 LSTM 神经网络的方法来解决不平衡情绪分类问题。首先, 该方法使用欠采样方法获取多组平衡训练语料; 其次, 使用每一组训练语料学习一个 LSTM 模型; 最后, 通过融合多个 LSTM 模型, 获得最终分类结果。实验结果表明我们的方法明显优于传统的不平衡分类方法。

关键词: 情绪分类; 不平衡分类; LSTM

Imbalanced Emotion Classification Based on Multi-channel LSTM

Hao Yin, Shoushan Li, Zhengxian Gong, Guodong Zhou

Natural Language Processing Lab of Soochow University, Suzhou, 215006

E-mail: hyin@stu.suda.edu.cn, {lishoushan, zhxgong, gdzhou}@suda.edu.cn

Abstract: Emotion classification is an important research issue in natural language processing. It aims at classifying the type of emotion automatically. However, most existing studies assume that the labeled data of each emotional category is balanced, which may not be true in reality. In this paper, we aims at the emotion classification on imbalanced data. Specifically, this paper proposes a novel approach called multi-channel LSTM which combines several LSTM models to solve the problem of imbalanced emotion classification. First, the approach get some groups of balanced training corpus by under-sampling method. Then, it trains several LSTM models with each training corpus. Third, it yields the results through joining all the LSTM models. Evaluation on SINA micro-blog shows that the proposed approach performs better than several existing imbalanced classification methods.

Keywords: Emotion classification; Imbalanced classification; LSTM

1 引言

随着社交网站的兴起, 越来越多的人习惯在网上发表自己的言论, 这些言论中大部分都含有用户的情感信息。分析这些带有情感的言论不仅有助于进行问答系统、舆情监控等技术应用, 还可以用来帮助心理学专家检测用户的心理状态。因此, 近些年来情感分析受到了自然语言处理领域研究者的密切关注, 现已成为一项基本的热点研究任务^[1]。

***作者简介:** 殷昊 (1994—), 男, 江苏, 硕士研究生, 研究方向为自然语言处理。

基金项目: 国家自然科学基金重点项目 (No. 61331011), 国家自然科学基金 (No. 61375073 和 No. 61273320), 国家青年科学基金项目 (No. 61305088), 江苏省产学研联合创新资金—前瞻性联合研究项目 (No. BY2014059-16) 资助。

情感分析又称意见挖掘、观点分析等，是通过计算机帮助用户快速获取、整理互联网上的海量的主观评价信息，对带有情感色彩的主观性文本进行分析、处理、归纳和推理的过程^[2]。文本情绪分类是情感分析的一项基本任务，该任务具体是针对文本表达的个人情绪（如：高兴、伤心、惊讶等）进行自动分类的方法^[3]。表 1 给出了一些含有情绪的微博文本样例。

表 1 新浪微博中带有情绪的微博样例
Tab.1 Emotional text in sina micro blog

情绪类别	微博文本
惊讶	哇塞！好奇怪啊，竟然有这样的事情。
伤心	话说连续三天晒单都没有中奖，感觉好伤心呢！？
高兴	哈哈刚刚下课到家~~~[哈哈][哈哈][爱你]

虽然文本情绪分类的研究已经开展多年，但是目前大部分的研究都假设各种情绪类别的样本分布平衡，即各情绪类别的样本个数相同。然而，现实情况往往并非如此。在实际收集到的语料中，不管是产品的评论文本还是微博文本，各个情绪类别样本的分布往往会非常不平衡。样本分布的不平衡会使得应用传统的机器学习分类方法得到的分类结果严重偏向多样本类别（样本数目多的类别），从而大大地降低分类性能^[4]。

不平衡分类问题在自然语言处理领域是一个经典的研究问题。目前主流的不平衡分类方法是基于欠采样的机器学习分类方法。该方法的主要思想是使用欠采样技术得到平衡的样本，再通过机器学习的分类方法来对样本进行分类。由于欠采样方法使得多类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。为了充分利用已有样本，本文提出了一种基于多通道长短时记忆（Long Short-Term Memory, LSTM）神经网络^[5]的不平衡分类方法用来解决微博情绪分类中的情绪类别样本分布不平衡问题。LSTM 神经网络可以学习长期依赖关系，能够有助于提高文本情绪分类的性能。具体而言，首先，我们使用欠采样方法获取多组平衡训练语料；其次，使用每一组训练语料学习一个 LSTM 模型；最后，通过 Merge 层来融合多个 LSTM 模型，获得最终分类结果。实验结果表明，该方法与传统不平衡分类方法相比较，能够进一步提升情绪分类性能。

本文结构安排如下：第二节介绍了与本文相关的一些工作进展；第三节介绍本文提出的基于多通道 LSTM 神经网络的不平衡分类方法；第四节给出实验结果及相关分析；第五节对本文做出总结，并对下一步工作进行展望。

2 相关工作

2.1 情绪分类

目前，针对社交媒体中的文本情感分析方法的研究大都是面向情感极性的（如正面情感、中性情感、负面情感），而针对细粒度的情绪分类方法的研究还比较缺乏。

计算语言学领域著名的语义评估会议 SemEval 在 2007 年设立了一个评测任务，用来对新闻标题进行情绪分类，该评测任务提供了一个包含 1250 个句子的数据集。为了更好地理解情绪分析问题，该任务强调对情绪进行词法语义分析^[6]。国内的相关会议也组织了中文情绪分析相关的评测任务。如 NLP&CC-2013 中文微博情绪识别任务。该评测任务以新浪微博文本作为基础语料进行标注^[7]，目前已完成 14000 条微博，45431 条句子的情绪标注，

构建了一个规模较大的中文情绪语料库。该语料库为相关科研工作人员分析微博文本的情绪表达提供了支持，有效地促进了相关领域的研究发展。Li 等^[8]提出利用句子的标签因子图和上下文标签因子图，进行句子级的情绪分类，很好的解决了数据稀疏和情绪的多标签问题。Rana R 等^[9]将神经网络方法应用于带有噪音的文本情感分类，实验证明了该方法能够很好地处理噪音问题。梁军等^[10]将 LSTM 扩展到基于树结构的递归神经网络上，并根据句子前后词语间的关联性引入情感极性转移模型。

已有的情绪分析研究基本都是基于样本分布平衡的假设，不平衡数据的情绪分析方法研究还很缺乏。据我们所知，本文是首次针对情绪分类问题提出的不平衡分类方法。

2.2 不平衡分类

不平衡分类问题具有一系列传统模式分类方法所没有考虑到的特点，所以传统模式分类方法难以很好解决不平衡分类问题。不平衡分类问题在机器学习、模式识别等领域均受到广泛关注，是众多实际任务中共同具有的挑战性问题。

主流的不平衡分类方法主要以过采样技术和欠采样技术为主。具体而言，过采样技术通过重复少类样本使得少类样本数和多类样本数平衡；欠采样技术通过减少多类样本使得两类样本平衡。目前针对不平衡语料的情感分析研究还很少，王中卿等^[4]针对不平衡数据的中文情感分类，提出了一种基于欠采样和多分类算法的集成学习框架。Yan 等^[11]提出了一种两阶段分类框架，使得不平衡数据的分类更准确。Li 等^[12]将监督学习、主动学习和半监督学习方法引入不平衡情感分类问题，取得了很好的分类效果。

3 基于 LSTM 神经网络的不平衡情绪分类方法

3.1 情绪分类中的不平衡分布情况

为了更好的理解情绪分类中的不平衡问题，我们分析了 NLP&CC-2013 中文微博情绪分析评测提供的微博语料。我们根据微博的主要情绪将语料分为七个情绪类别，并统计了这些类别的分布情况。这七个情绪分别是高兴、喜好、愤怒、悲伤、恐惧、厌恶和惊讶。

表 2 有情绪微博中各情绪分布情况
Tab.2 Emotions distribute in microblog

情绪类别	高兴	喜好	愤怒	悲伤	恐惧	厌恶	惊讶	合计
样本数量	1460	2203	669	1173	148	1392	362	7407

表 2 给出了七个情绪类别的样本分布情况。从表中可以看出，七个情绪类别的样本分布非常不平衡，最多的类别（喜好类）和最少的类别（恐惧类）之间样本数量比高达到 15 左右。从表中可以看出，恐惧情绪的样本数量最少，惊讶和愤怒情绪的样本数量较少，高兴、悲伤和厌恶情绪的样本数量较多，喜好情绪的样本数量最多。

3.2 基于单通道 LSTM 的不平衡情绪分类方法

LSTM 神经网络使用记忆单元来避免反向传播过程中的梯度消失和梯度爆炸问题，并且可以学习长期依赖关系、充分利用历史信息。Alex Graves 于 2013 年对 LSTM 进行了改良和推广^[13]，使得 LSTM 被广泛应用于自然语言处理、语音识别等领域中。

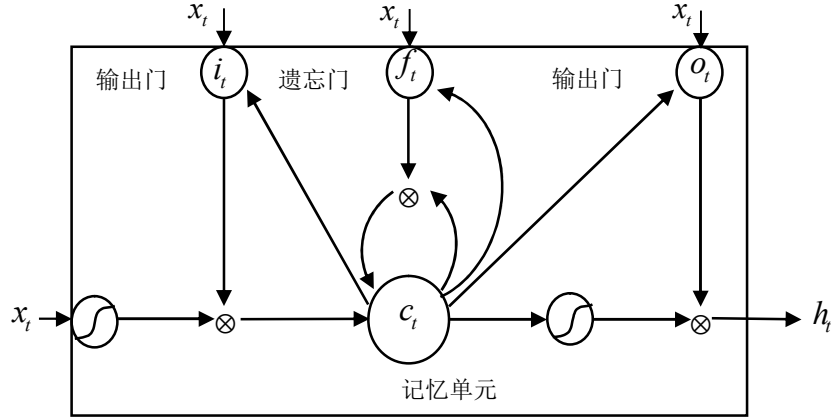


图 1 LSTM 单元

Fig.1 LSTM unit

如图 1 所示，LSTM 单元设置了记忆单元 c 用于保存历史信息。历史信息的更新和利用受到 3 个门的控制：输入门 i 、遗忘门 f 和输出门 o 。LSTM 单元在 t 时刻的更新过程如下：

$$i_t = \sigma(W_i x_t + U_i h_{t-1} + V_i c_{t-1}) \quad (3.1)$$

$$\tilde{c}_t = \tanh(W_c x_t + U_c h_{t-1}) \quad (3.2)$$

$$f_t = \sigma(W_f x_t + U_f h_{t-1} + V_f c_{t-1}), \quad (3.3)$$

$$c_t = f_t \square c_{t-1} + i_t \square \tilde{c}_t, \quad (3.4)$$

$$o_t = \sigma(W_o x_t + U_o h_{t-1} + V_o c_t) \quad (3.5)$$

$$h_t = o_t \square \tanh(c_t), \quad (3.6)$$

其中 x_t 为记忆单元的输入数据， σ 为 logistic sigmoid 函数，符号 \square 为向量之间的点乘运算， W_i ， W_f ， W_c ， W_o ， U_i ， U_f ， U_c ， U_o 和 V_o 为权重矩阵。 i_t ， o_t ， f_t ， c_t 分别为输入门、输出门、遗忘门和记忆单元在 t 时刻的值， \tilde{c}_t 为记忆单元的候选记忆状态值， h_t 为 t 时刻 LSTM 单元的输出。

针对不平衡情绪分类，我们首先使用随机欠采样方法得到各情绪类别平衡的样本，然后采用单通道 LSTM 神经网络作为分类方法。图 2 为单通道的 LSTM 神经网络分类器框架图，该分类模型只有一个 LSTM 层，第一个虚线框部分为单通道 LSTM 模型内部结构，第二个虚线框部分为不平衡样本的处理过程。LSTM 模型的输入为训练样本的词向量表示，词向量具有良好的语义特征，是表示词语特征的常用方式^[14]。输入特征向量经过 LSTM 层得到高维向量，可以学习到更深层次的特征，这些特征能够更好的描述样本。全连接层类似于传统多层感知机的隐藏层，接收来自上一层的全部输出，给这些输出向量加权求和，加权后的输出经过激励函数并传播到 Dropout 层。本实验中该层使用 Relu 作为激励函数，Relu 减少了参数之间的相互依存关系，更接近生物学的激活模型，激励函数如公式 (3.7) 所示：

$$g(x) = \max(0, x) \quad (3.7)$$

其中 x 为输出向量，Relu 函数将小于 0 的值全部置 0，具有引导适度稀疏的能力。

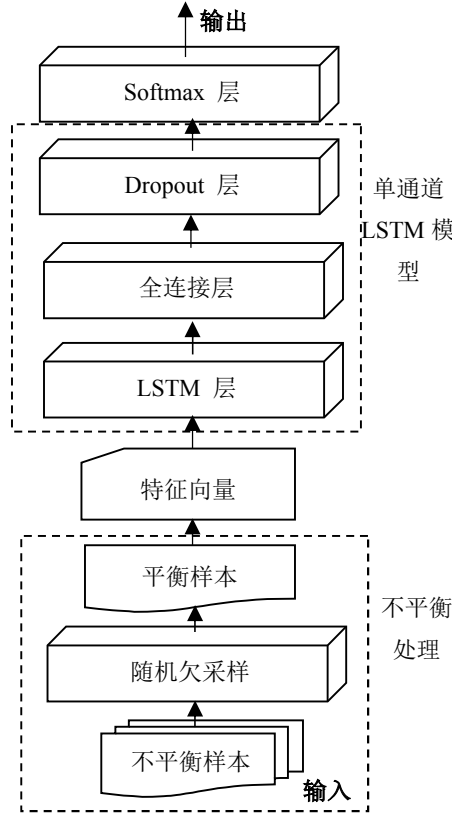


图 2 单通道 LSTM 框架图

Fig.2 The framework of single-channel LSTM

Dropout 层在训练和预测时随机让网络中某些隐含层节点不工作，减少了特征个数，有效的防止了网络过拟合。Dropout 层作为 LSTM 神经网络模型中的隐藏层出现：

$$g = h^* \cdot D(p) \quad (3.8)$$

D 表示 dropout 操作符， p 是一个可调的超参（保留隐层单元的比率）。

最后，单通道 LSTM 模型的输出通过 softmax 输出层来对样本进行分类。我们选择后验概率最大的类别作为预测标签：

$$label_{pred} = \arg \max_i P(Y = i | x, W, U, V) \quad (3.9)$$

其中 x 为上一层输出向量， i 为标签预测值， W ， U ， V 为 LSTM 更新方法中的系数矩阵， $label_{pred}$ 为后验概率最大的预测标签。

3.3 基于多通道 LSTM 的不平衡情绪分类方法

应用随机欠采样和单通道 LSTM 神经网络进行不平衡情绪分类存在一个明显的缺点：由于欠采样只是从多类中选择部分样本，使得大量未选中的样本在后面的分类过程中未能发挥作用，从而丢失了很多可能对分类有帮助的样本。因此，为了充分利用已标注样本，提高分类器性能，我们提出了一种基于多通道 LSTM 神经网络的分类方法。该方法在不平衡样本中多次欠采样得到多组平衡样本，使用每一组平衡样本学习一个 LSTM 模型，通过

Merge 层对多个 LSTM 模型进行联合学习，得到最终的分类结果。

基于多通道 LSTM 神经网络的分类器框架图如图 3 所示。我们首先使用随机欠采样的方法对各类别样本进行 n 次欠采样，每次欠采样的个数为最少类的样本数。在得到的 n 组平衡样本中，每组对应位置的样本的情绪类别相同，可以将该 n 组不同的样本看作是七种情绪类别所对应的 n 组不同的特征表示。我们将该 n 组训练样本的特征向量同时作为网络的输入，且每组输入分别用来训练一个 LSTM 模型。在图 3 中，LSTM_ n 表示由第 n 组训练样本学习得到的 LSTM 模型，该模型的输出是输入特征的更好表示。本实验中 n 取 5，这样可以保证已标注样本中绝大多数样本都能被取到。

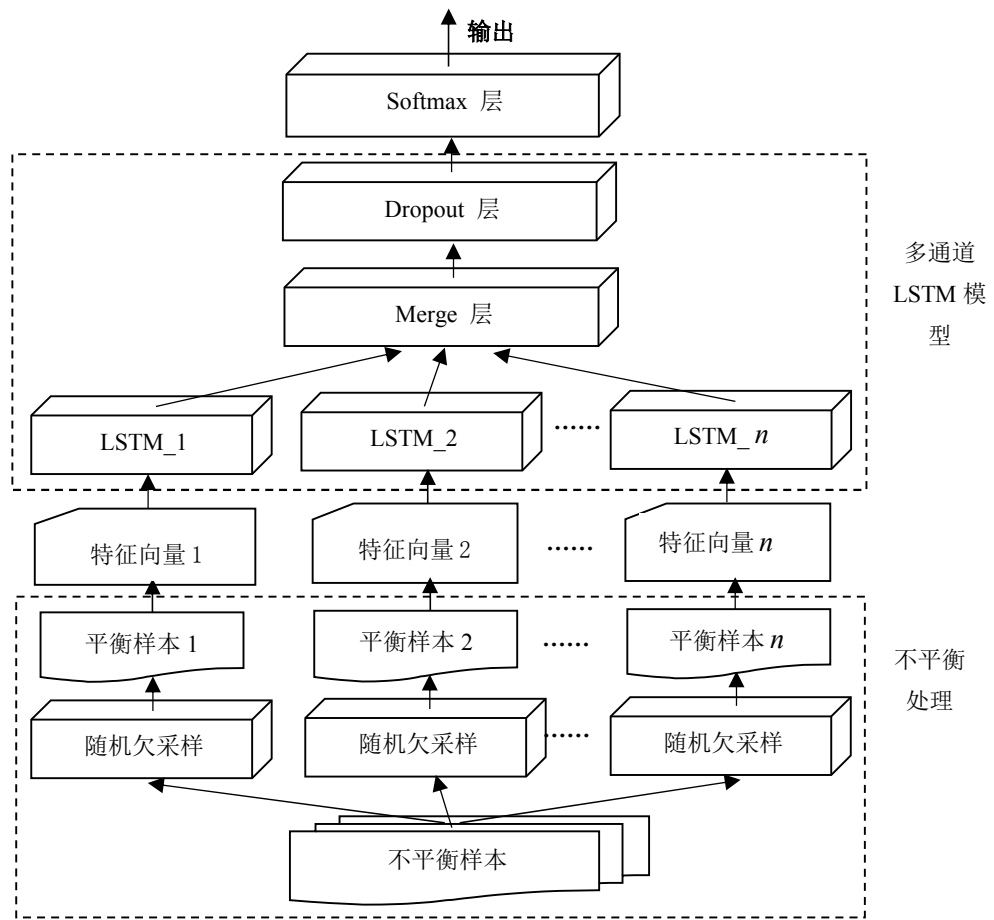


图 3 多通道 LSTM 框架图

Fig.3 The framework of multi-channel LSTM

Merge 层将上述 n 组 LSTM 模型的输出特征进行相加融合，并通过反向传播算法 (back propagation) 来对网络参数进行更新。Dropout 层接收 Merge 层的输出作为输入，该层的功能与其在单通道 LSTM 神经网络中的功能相同。网络的最后一层是 Softmax 输出层，该层用来输出网络模型的预测标签 $label_{pred}$ 。在模型训练的过程中，我们选择最小化交叉熵误差作为损失函数：

$$J = -\sum_{i=1}^N \sum_{l=1}^m [1\{t_i = l\} \log(y_l)] + \frac{\lambda}{2N} \sum_{k=1}^n \left(\sum_{\varepsilon \in \omega} \|W_{\varepsilon}^k\|_F^2 + \sum_{\varepsilon \in \mu} \|U_{\varepsilon}^k\|_F^2 + \sum_{\varepsilon \in \nu} \|V_{\varepsilon}^k\|_F^2 \right) \quad (3.10)$$

其中， N 是训练样本的个数， m 是目标类别的数量， y 是 softmax 层输出的每个类别的预测概率， t_i 是第 i 个训练样本的真实标签。 $\|\cdot\|_F$ 表示 Frobeniu 范数， n 是通道的个数， $\omega = \{i, f, o, c\}$ ， $\mu = \{i, f, o, c\}$ 和 $\nu = \{i, f, o\}$ 表示不同门的集合（分别为 W ， U ， V ）， λ 是用来指定惩罚权重的超参。

在损失函数中，除了极小化负数对数似然，还增加了 W ， U ， V 的 L2 正则化。原因是 softmax 函数的参数存在冗余，也就是极小点不唯一，增加正则项可以将极小点唯一化。惩罚因子 λ 调节正则项的权重，取值越大，对大参数的惩罚越大。

4 实验

4.1 实验设置

本文选用 NLP&CC-2013 中文微博情绪分析评测提供的微博语料作为实验语料。该语料中共有七种情绪类别，具体情绪类别及样本分布可参考 3.1 节的表 2。由于语料中恐惧情绪的样本数量太少，根据其样本数生成的测试集得到的实验结果具有较大的偶然性，因此我们选取第二少类别（惊讶情绪）里面样本数的 20%（即 $362 \times 20\% \approx 72$ ）作为各类别的测试样本数。训练样本则根据不同的分类方法从各类别的剩余样本中抽取。各情绪类别的测试样本数和剩余样本数如表 3 所示。

表 3 各情绪类别测试样本数和剩余样本数

Tab.3 The number of test samples and remaining samples

情绪类别	高兴	喜好	愤怒	悲伤	恐惧	厌恶	惊讶	合计
样本总量	1460	2203	669	1173	148	1392	362	7407
测试样本	72	72	72	72	72	72	72	504
剩余样本	1388	2131	597	1101	76	1320	290	6903

表 4 LSTM 神经网络中的参数设置

Tab.4 Parameters setting in LSTM

参数表述	参数值
输入向量长度	120
LSTM 层输出维度	128
全连接层输出维度	64
Dropout rate	0.5
迭代次数	20

实验中用到的分类算法包括最大熵和 LSTM 神经网络。其中最大熵使用 MALLET 机器学习工具包[†]；LSTM 神经网络使用深度学习开源框架 Keras 搭建[‡]。在进行实验之前我们首先采用复旦大学公布的分词工具 FudanNLP[§]对中文文本进行分词。在使用最大熵分类器

[†] <http://mallet.cs.umass.edu/>

[‡] <http://keras.io/>

[§] <http://code.google.com/p/fudannlp/>

时，我们选取词的 Uni-gram 作为特征，得到文本的向量表示。最大熵分类器所有的参数都使用它的默认值。使用 LSTM 神经网络分类器时，我们首先使用 Python 工具包 gensim**来生成样本的词向量模型。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。LSTM 神经网络模型的具体参数设置如表 4 所示。

实验采用正确率（Accuracy）和几何平均数（G-mean）作为衡量分类效果的标准。几何平均数的计算方法为式（4.1）：

$$G-mean = \sqrt[n]{\prod_{i=1}^n Recall_i} \quad (4.1)$$

其中 $Recall_i$ 表示类别 i 的召回率， n 为类别个数，本实验 n 取 7。

4.2 实验结果

在实验中，我们实现了以下几种处理不平衡情绪分类的方法。这些方法如下：

- 1) **完全训练+最大熵（FullT+Maxent）**，各个类别的全部剩余样本均作为训练样本，采用最大熵分类器。
- 2) **随机过采样+最大熵（OverS+Maxent）**，设最多类（喜好类）的剩余样本数为 n_{\max} ，使用随机过采样技术从各类别的剩余样本中抽取 n_{\max} 个样本作为训练样本，采用最大熵分类器。
- 3) **随机欠采样+最大熵（UnderS+Maxent）**，设第二少类（惊讶类）的剩余样本数为 n_{\min} ，使用随机欠采样技术从各类别的剩余样本中抽取 n_{\min} 个样本作为训练样本，采用最大熵分类器。
- 4) **随机欠采样+单通道 LSTM 神经网络（UnderS+LSTM）**，使用 3）中的采样方法得到训练样本，分类器使用单通道的 LSTM 神经网络。
- 5) **随机欠采样+集成学习（Ensemble-Maxent）**，多次使用 3）中的采样方法得到多组训练样本（本实验取 5 组），并建立多个基分类器。最后通过融合这些基分类器结果进行集成学习^[4]，其中基分类器选择最大熵分类器。
- 6) **随机欠采样+多通道 LSTM 神经网络（Multi-LSTM）**，使用 5）中的采样方法得到多组训练样本（本实验取 5 组），分类器使用多通道（5 通道）的 LSTM 神经网络。

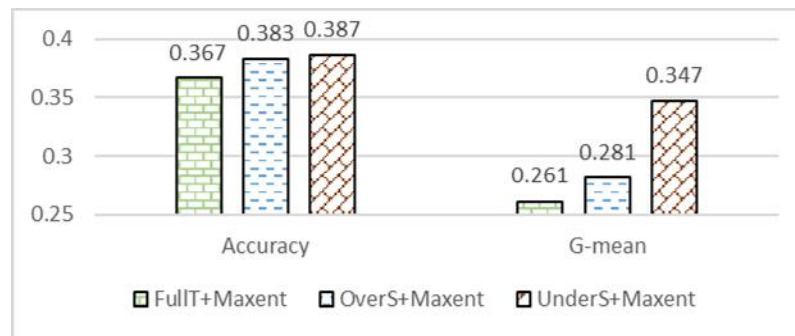


图 4 传统不平衡分类方法分类性能比较

Fig.4 Results of traditional imbalanced classification methods

图 4 比较了完全训练、随机过采样和随机欠采样方法在基于不平衡数据的情绪分类中

** <http://radimrehurek.com/gensim/install.html>

的分类效果。我们可以看出随机欠采样的分类性能明显优于前两者，其优势在 G-mean 值上表现得尤为突出。该现象的主要原因是在完全训练和随机过采样方法中，分类算法严重趋向样本数量较多的类别，使得样本数量较少的类别的召回率非常低。

接下来我们比较最大熵和 LSTM 神经网络在随机欠采样方法下的情绪分类性能。从图 5 可以看出单通道 LSTM 神经网络的分类性能要优于最大熵的分类性能，在 Accuracy 和 G-mean 上分别提高了 1.8 和 1.2 个百分点。我们分析其主要原因是 LSTM 神经网络能够充分利用历史信息，可以学习到样本之间的长期依赖关系。

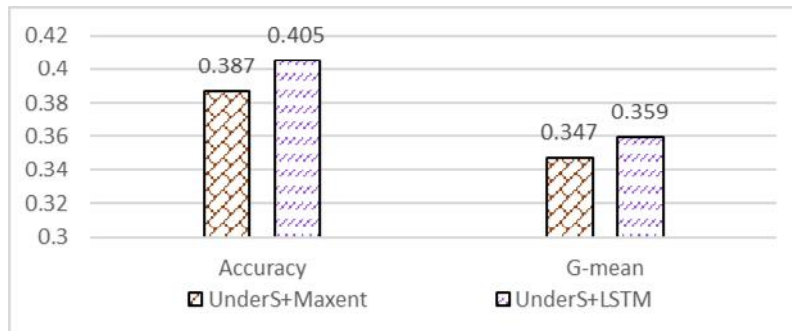


图 5 最大熵和 LSTM 神经网络的分类性能比较

Fig.5 Performances of Maxent and LSTM

在不平衡分类问题中，为了充分利用所有标注样本，但又保持训练样本之间的平衡，基于欠采样的集成学习的方法表现得较为理想。接下来我们将比较基于欠采样的集成学习的分类方法和我们提出的基于多通道 LSTM 神经网络的分类方法，二者的分类性能如图 6 所示。

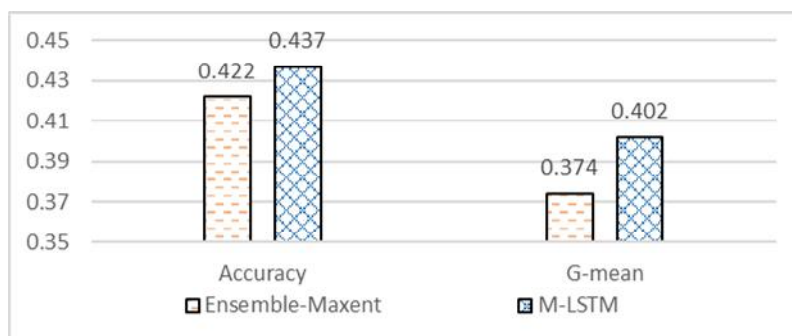


图 6 集成学习和多通道 LSTM 神经网络的分类性能比较

Fig.6 Performances of ensemble learning and multi-LSTM

对照图 4 和图 6，我们可以发现基于欠采样的集成学习的分类方法在性能上要优于以往所有的分类方法，其原因是该方法既能够保持各类别训练样本之间的平衡，又能够充分利用已有样本的信息。图 6 结果表明我们提出的基于多通道 LSTM 神经网络的分类方法相对于其他方法有明显优势，Accuracy 较集成学习方法相比提高了 1.5 个百分点，G-mean 则提高了 2.8 个百分点。这些结果说明了基于多通道 LSTM 神经网络的分类方法对不平衡情绪分类非常有效。该方法不仅可以在各类别样本分布不平衡时充分利用已有样本的信息，而且可以考虑文本上下文之间的关系。

5 结语

本文针对情绪分类任务中的数据不平衡问题,提出了一种基于多通道 LSTM 神经网络的分类方法。该方法首先使用随机欠采样方法获取多组平衡训练语料;其次,使用每一组训练语料学习一个 LSTM 模型;最后通过融合多个 LSTM 模型,获得最终分类结果。实验结果表明该方法能够充分利用训练样本,性能上明显优于传统的不平衡分类方法。

在下一步工作中,我们将收集其它领域的语料,比如贴吧、QQ 空间说说等,并在其上应用我们的方法进行实验,验证方法的有效性。此外,我们将探索如何改进欠采样技术,使得多类样本在采样过程中分布的更合理。今后的工作我们将着力解决上述问题,以便找出性能更佳的不平衡数据情绪分类方法。

参 考 文 献

- [1] Jiang L, Yu M, Zhou M, et al. Target-dependent Twitter sentiment classification[C]// Meeting of the Association for Computational Linguistics. 2011:151-160.
- [2] 赵妍妍, 秦兵, 刘挺. 文本情感分析[J]. 软件学报, 2010, 21(8):1834-1848.
- [3] Aman S, Szpakowicz S. Identifying Expressions of Emotion in Text[M]// Text, Speech and Dialogue. Springer Berlin Heidelberg, 2007:196-205.
- [4] 王中卿, 李寿山, 朱巧明,等. 基于不平衡数据的中文情感分类[C]. 中国计算语言学研究前沿进展. 2011:33-37.
- [5] Hochreiter S, Schmidhuber J. Flat minima.[J]. Neural Computation, 1997, 9(1):1 - 42.
- [6] Strapparava, Carlo, Mihalcea, Rada. Learning to Identify Emotions in Text[J]. Unt Scholarly Works, 2008, 43(3):254-255.
- [7] 姚源林, 王树伟, 徐睿峰,等. 面向微博文本的情绪标注语料库构建[J]. 中文信息学报, 2014, 28(5):83-91.
- [8] Li S., L. Huang, R. Wang, and G. Zhou. Sentence-level Emotion Classification with Label and Context Dependence[C].In Proceedings of ACL-15, 2015, 1045-1053.
- [9] Rana R. Emotion Classification from Noisy Speech - A Deep Learning Approach[J]. 2016.
- [10] 梁军, 柴玉梅, 原慧斌,等. 基于极性转移和 LSTM 递归网络的情感分析[J]. 中文信息学报, 2015, 29(5):152-159.
- [11] Yan Y, Liu Y, Shyu M L, et al. Utilizing concept correlations for effective imbalanced data classification[C]// IEEE, International Conference on Information Reuse and Integration. IEEE, 2014:561-568.
- [12] Li S, Zhou G, Wang Z, et al. Imbalanced sentiment classification[C]// ACM Conference on Information and Knowledge Management, CIKM 2011, Glasgow, United Kingdom, October. 2011:2469-2472.
- [13] Graves A. Supervised Sequence Labelling with Recurrent Neural Networks[J]. Studies in Computational Intelligence, 2012, 385.
- [14] Turian J, Ratinov L, Bengio Y. Word representations: a simple and general method for semi-supervised learning[C]// Meeting of the Association for Computational Linguistics. Association for Computational Linguistics, 2010:780-1.