

基于Bi-LSTM的不平衡样本文本分类模型

王欣羽, 李 薇

燕山大学理学院, 河北 秦皇岛

收稿日期: 2023年9月28日; 录用日期: 2023年10月25日; 发布日期: 2023年11月3日

摘 要

情感分类任务通常是将有情感倾向的样本分为积极和消极两类。在大多数的理论模型中, 这两类样本的数量都被假定是平衡的, 而事实上, 这两类样本在现实生活中一般是不平衡的。为解决这一问题, 提出一种基于Focal损失的Bi-LSTM神经网络模型。首先, 采集并标注了24,190条旅游评论作为该模型的数据集, 其中积极样本远多于消极样本。为达到更好的分类结果, 首先将样本数据集分为核心样本和非核心样本, 并剔除非核心样本, 提高数据质量; 其次, 用基于Focal损失的Bi-LSTM神经网络模型对数据进行训练; 最后, 对测试集进行验证并得到最终分类结果。通过准确率(accuracy)、F1、召回率(recall)和特异度指标(specificity)这四个评价指标判断模型优劣。一系列的实验结果显示, 基于Focal损失的Bi-LSTM神经网络模型能够更好的解决样本不平衡的问题, 与传统的LSTM模型分类方法相比, 其分类性能更好。

关键词

不平衡数据, LSTM, 情感分类, 特异度指标, Focal损失

Bi-LSTM Based Text Classification Model for Imbalanced Samples

Xinyu Wang, Wei Li

School of Science, Yanshan University, Qinhuangdao Hebei

Received: Sep. 28th, 2023; accepted: Oct. 25th, 2023; published: Nov. 3rd, 2023

Abstract

In general, the task of sentiment classification usually divides samples with emotional tendencies into two categories: positive and negative. In most theoretical models, the number of samples in these two categories is assumed to be balanced, while in fact, the two categories are generally unbalanced in real life. In this paper, a Bi-LSTM network model based on Focal loss is proposed to classify sentiment for unbalanced sample data. Firstly, 24,190 travel reviews were collected and

labeled as the dataset of the proposed model, whose positive samples were much more than negative samples. In order to achieve better classification results, the sample dataset is first divided into core and non-core samples, and the non-core samples are eliminated to improve the data quality; secondly, the data were trained with a Bi-LSTM neural network model based on Focal loss; finally, the test set is validated and the final classification results are obtained. Four evaluation metrics, accuracy, F1, recall and specificity, are used to judge the model merits. A series of experimental results show that the Bi-LSTM neural network model based on Focal loss can better solve the problem of sample imbalance and has better classification performance compared with the traditional LSTM model classification method.

Keywords

Imbalanced Data, LSTM, Sentiment Classification, Specificity Indicator, Focal Loss

Copyright © 2023 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

近年来,情感分类任务一直是国内外热门的研究课题之一,它主要通过分析文本内容中所表达的意义和情感信息,将其分为积极或消极倾向的两种或多种类型。情感分类的主要目的就是识别用户对事物或人的主观看法和态度,所以它也被称为意见抽取、情感挖掘等[1]。随着互联网技术的普及与发展,越来越多的网络用户会在社交平台或购物平台上发表带有主观情感的评论。提取这些评论的重要信息并对其进行分析预测,推断用户情感以及需求的倾向性成为这些平台关注的重点内容之一。而在情感分类任务中,将文本情感倾向划分为积极和消极两类是最为常见的,因此本文研究的是二分类问题。而在目前已有的研究方法中,基于神经网络模型的有监督学习方法是目前的主流方法。杜启明等[2]提出了一种结合上下文和依存句法信息的中文短文本情感分析模型,该模型不仅利用双向长短期记忆网络提取文本的上下文语义,而且引入了一种基于依存关系感知的嵌入表示方法,以加强文本表示中的情感特征。实验表明,该模型在 SWB, NLPCC2014 和 SMP2020-EWEC 数据集上能够有效融合语句中的语义以及句法结构信息,在中文短文本情感二分类以及多分类中均取得了较好的效果。王娅丽等[3]针对目前结合图卷积网络的研究忽略方面术语本身含义以及方面术语与上下文之间的交互的问题,提出了基于交互注意力和图卷积网络的模型 IAGCN,该模型结合 BiLSTM 和修正动态权重层对上下文进行建模,并在句法依存树上使用图卷积网络对句法信息进行编码,然后利用交互注意力机制学习上下文和方面术语中的注意力,重构上下文和方面术语的表示,最后通过 softmax 层获取给定方面术语的情感极性。与其他模型相比,所提模型在 5 个数据集中的准确率和 F1 值都有了显著的提高。

在一般的情感分类模型研究中,通常假定数据集各类别样本数量是相同的,即积极样本和消极样本的数目是均衡的,但在现实生活中遇到的实际问题通常不能满足这个条件。也就是说在实际收集到的数据中,积极样本和消极样本的数量往往是不平衡的,因此如何对不平衡样本进行分析就显得尤为重要,这就产生了进一步的研究[4]。当传统的机器学习模型被用来对不平衡的样本进行情感分类时,最终的分

类结果会倾向于样本量大的类别,导致模型失真和分类结果准确率的下降。因此,研究不平衡样本的情感分类任务是很有必要的。之前已经有很多学者对不平衡样本情感分类进行了研究。郭朝有等[5]融合了 Canopy 和 K-means 两种聚类算法,设计了 C-K-SMOTE 改进算法处理样本不平衡问题。实验表明

C-K-SMOTE 算法可有效平衡不平衡数据集。梁越等[6]提出了基于虚拟对抗训练(VAT)的类别不平衡小样本控申文本分类方法,针对信访信件数量偏少且存在对抗样本的情况,在模型训练时引入 VAT 进行优化。任金金等[7]设计了新的 BP NNS 非均衡数据分类方法,即采集处理不平衡数据信息降低数据分类难度;计算不平衡数据分类子类样本数量提高数据分类敏感性;构建不平衡数据分类模型提高数据分类的流畅性,从而实现不平衡数据高精度分类。Tanha [8]在多类不平衡常规数据集和大数据集上,CatBoost 算法和 LogitBoost 算法分别优于其他增强算法。此外,在多类不平衡数据域中,MMCC 是比 MAUC 和 G 均值更好的评价指标。

本文通过改进机器学习模型中常用的交叉熵损失函数,将 Focal 损失应用到 Bi-LSTM 网络中,从模型上处理不平衡分类问题,达到提高分类器性能的目的。

本文的主要贡献概括如下:

- 1) 构建中文多平台旅游评论数据集,该数据集包含 24,190 条旅游评论,分为积极和消极两类。
- 2) 将两类样本分为核心样本和非核心样本,即情感倾向性强的样本和情感倾向性弱甚至没有情感倾向性的样本,对两类样本的非核心样本进行删除处理,以提高样本质量。
- 3) 为了有效的解决数据集样本情感倾向类别不平衡对分类效果的影响,提出基于 Focal 损失的 Bi-LSTM 神经网络模型,该模型包括输入层,编码层,Bi-LSTM 层,Attention 层和输出层五个部分。
- 4) 基于构建的评论数据集进行了相关实验,实验结果表明:不论训练集和测试集是否为同一平台数据,该模型在评价指标上都有优异表现,体现了该模型的泛化性和有效性。

2. 模型设计与实现

2.1. 构建多平台旅游评论数据集

情感分类任务的语料库大多是英文语料,中文的语料很少。主要是因为一些平台的数据涉及到用户的隐私,所以并不会公开大量的评论,这使得研究中文文本分类任务变得困难。本文通过八爪鱼采集器采集了携程网和去哪儿网的旅游评论数据共 24,190 条,构建了一个多平台的旅游评论数据集[9]。评论的情感分类分为积极情感和消极情感两类,在进行人工标注情感极性时发现数据集具有样本不平衡的特点。为进一步提高样本质量,将两类样本分为核心样本和非核心样本,核心样本是指情感倾向性强的样本,如“这里的景色很美,像画一样,还能感受到历史的厚重,值得一游”。非核心样本是指情感倾向性弱甚至没有情感倾向性的样本,如“该用户默认好评”。清洗掉两类样本中的非核心数据,得到情感倾向性较强的数据集作为本文的实验数据集。最终得到的数据集中积极情感的评论数据集约为消极情感评论数据集的 3 倍。所以接下来本文将对不平衡样本的情感分类任务进行研究。

2.2. 基于 Bi-LSTM 模型的文本分类研究

在本节中,本文将详细介绍 Bi-LSTM-Att-Focal loss 模型。如图 1 所示,本文提出的模型包含五个部分:

- 1) 输入层: 输入预处理后的旅游评论;
- 2) 编码层: 将句子中的每个词映射到低维向量;
- 3) Bi-LSTM 层: 利用 Bi-LSTM 网络模型获取高级特征;
- 4) Attention 层: 生成权重向量,将每个单词级特征合并成句子级特征向量;
- 5) 输出层: 使用句子级特征向量进行情感分类。这些部分都将在本节中详细介绍。

2.2.1. Word2vec 词向量方法

文本表示是将非结构化数据、不可计算的文本转换为计算机可以识别的数据信息[10],简单来说就是

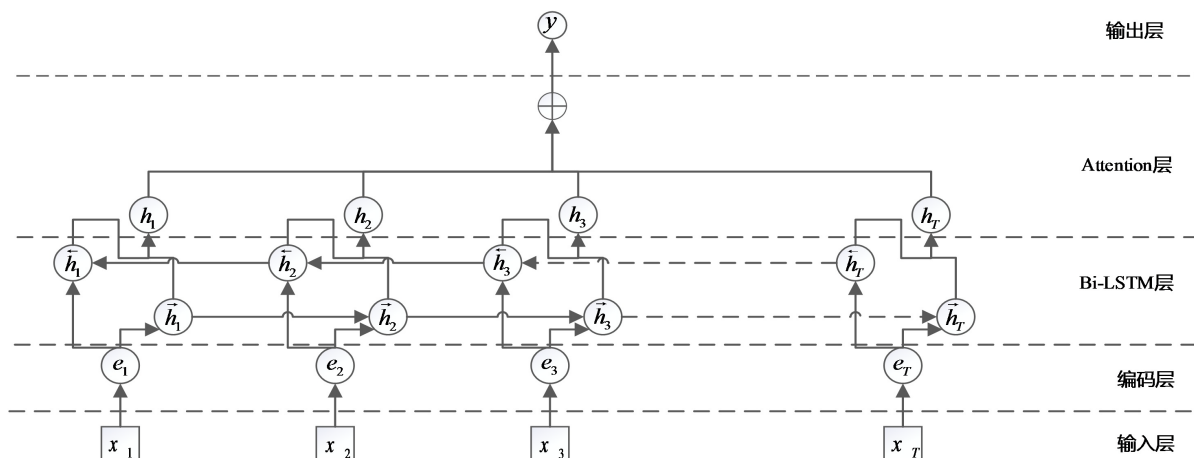


Figure 1. Structure of the neural network model

图 1. 神经网络模型结构图

不把文本看作字符串，而是看作在数学上处理起来更为方便的向量。常见的文本表示方法有 one-hot 独热编码、词袋模型、词嵌入等等。Word2vec 和 Glove 是词嵌入的两种主流算法。在提出词嵌入技术之前，one-hot 独热编码被广泛应用以处理文本信息，这种编码方式简单直接，一个词用一个长度为词典长度的向量表示，词向量中仅一个元素为 1，其它均为 0。这种方式的缺点在于编码无法反映词与词之间的关系，且维度较多计算量较大，使得该编码方式储存效率低下[11]。同样地，词袋模型也有类似独热编码的缺点。而词嵌入技术作为自然语言处理的重要突破之一，比上述两种编码方式有更多的优点：

- 1) 该方式不会像独热编码一样维度很长，转而通过低维度的向量来表示文本；
- 2) 文本中相近词意的在空间上距离也较为相近；
- 3) 该编码方式普遍适用性强，用途广泛[12]。

本文所采用的是词嵌入技术的其中一种，Word2vec。Word2vec 有两种训练模式：

- 1) 连续词袋模型(Continuous Bag-of-Words Model, CBOW)，该模型是将一个词本身作为输出，通过上下文来预测这个词和它的意思，适用小型训练数据；
- 2) 连续跳字模型(Continuous Skip-gram Model, Skip-gram)，该模型是将一个词本身作为输入，通过这个词来预测可能出现的上下文的词，适合训练大型的数据。

上述两者的结构如图 2 所示。

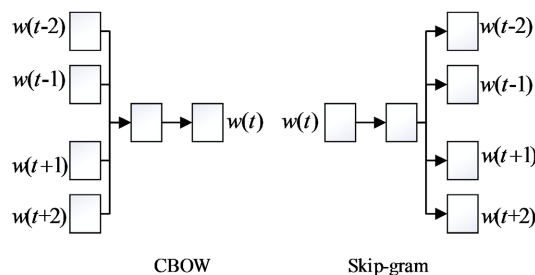


Figure 2. Word2vec structure diagram

图 2. Word2vec 结构图

较其他的词嵌入方法相比，Word2vec 进行文本表示的效果更好[13]，它所依靠的模型会联系上下文关系，从而使词向量维度大大降低，处理速度也更快。并且 Word2vec 具有高度的通用性，适用于广泛的

自然语言处理应用。同时,它能够比较准确的计算词语的相似度。所以本文通过 Word2vec 对数据预处理之后的样本进行训练,得到相应的词向量。

2.2.2. 神经网络模型

为解决梯度爆炸和梯度消失问题, Hochreiter S. 和 Schmidhuber J. 次提出了 LSTM 网络模型[14], 其主要思想是引入一种自适应门控机制, 它可以有选择的将信息去除或增加到当前的单元状态上, 从而更好地提取特征信息。之后就有很多学者提出了大量的 LSTM 变体。本文采用了其中一种变体[15]。通常, 由三个部分合成基于 LSTM 的递归神经网络:

1) 输入门 i_t 与相应的权重矩阵 W_{xi} 、 W_{hi} 、 W_{ci} 、 b_i ;

2) 遗忘门 f_t 与相应的权重矩阵 W_{xf} 、 W_{hf} 、 W_{cf} 、 b_f ;

3) 输出门与相应的权重矩阵 W_{xo} 、 W_{ho} 、 W_{co} 、 b_o 。设置好所有这些相应的部分后, 由当前输入 x_t 、前一步生成的状态 h_{t-1} , 以及前一个单元状态 c_{t-1} , 决定是否进行输入, 遗忘之前存储的信息, 并记忆新生成的信息。正如下面的方程式所示:

$$i_t = \sigma(W_{xi}x_t + W_{hi}h_{t-1} + W_{ci}c_{t-1} + b_i) \quad (1)$$

$$f_t = \sigma(W_{xf}x_t + W_{hf}h_{t-1} + W_{cf}c_{t-1} + b_f) \quad (2)$$

$$g_t = \tanh(W_{xc}x_t + W_{hc}h_{t-1} + W_{cc}c_{t-1} + b_c) \quad (3)$$

$$c_t = i_t g_t + f_t c_{t-1} \quad (4)$$

$$o_t = \sigma(W_{xo}x_t + W_{ho}h_{t-1} + W_{co}c_{t-1} + b_o) \quad (5)$$

$$h_t = o_t \tanh(c_t) \quad (6)$$

因此, 当前单元状态 c_t 是由前一个单元保留的信息和当前单元生成的新的信息计算加权和来生成的。对于情感分类任务来说, 我们需要所使用的网络模型可以通过捕捉双向的语义以更完整的提取文本的特征信息。然而, 基础的 LSTM 网络随时间推移对序列信息按顺序进行处理, 因此它们忽略了下文可能对当前状态产生的影响。而 Bi-LSTM 网络由两个输出方向相反的 LSTM 隐藏层组成, 能够同时利用上下文的信息对当前状态进行计算, 能够更准确、更完整的提取特征信息。

在本文中, 我们使用了 Bi-LSTM 网络, 该网络包含了两个针对左右序列上下文的子网络, 由一个向前传递的网络和一个向后传递的网络组成。第个字的输出为:

$$h_t = [\bar{h}_t \oplus \tilde{h}_t] \quad (7)$$

在这里, 我们使用元素求和来组合正向和反向传递输出。

2.2.3. 注意力机制

本节主要介绍分类任务的注意机制, 具体结构如图 3 所示。设 H 是一个由 Bi-LSTM 层产生的输出向量 $[h_1, h_2, h_3, \dots, h_T]$ 组成的矩阵, 其中 T 是句子长度。句子的表示 r 由这些输出向量的加权和组成, 相关公式为:

$$M = \tanh(H) \quad (8)$$

$$\alpha = \text{softmax}(\omega^T M) \quad (9)$$

$$r = H\alpha^T \quad (10)$$

其中 $H \in \mathbf{R}^{d^o \times T}$, d^o 是词向量的维数, ω 是一个训练好的参数向量, ω^T 是一个转置子。 ω , α , r 的维

数分别是 d^o , T , d^o 。最终用于分类的公式为:

$$h^* = \tanh(r) \tag{11}$$

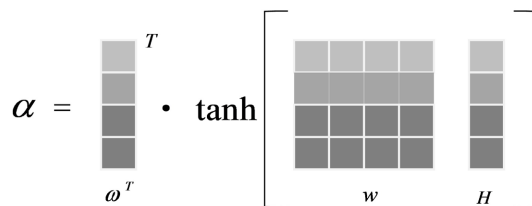


Figure 3. Structure of the attentional mechanism

图 3. 注意力机制结构图

2.2.4. Focal 损失

在文本分类任务中, 大多数的分类器都会使用交叉熵损失函数作为模型的损失函数, 但它并不能改善不平衡样本对分类模型造成的消极影响。为进一步提升分类器的性能, 本文用 Focal 损失[16]代替传统的交叉熵损失函数, 从而优化模型以更好的解决样本不均衡问题, 并取得了不错的效果。

Focal 损失的提出源自图像领域中目标检测任务中样本数量不平衡性的问题, 并且这里所谓的不平衡性跟平常理解的是有所区别的, 它还强调了样本的难易性。因为 Focal 损失[17]是在交叉熵损失函数的基础上提出的, 具体公式如下:

$$FL(p_i) = -(1 - p_i)^\gamma \log(p_i) \tag{12}$$

其中 γ 为调节因子, 取值为[0, 5], 当 $\gamma = 0$, 就等同于 CE 函数[18]; γ 值越大, 表示模型在难易样本上聚焦的更厉害。但是上面的 Focal 损失公式只是体现了难易样本的区分, 没有区分正负。这样就引出了完整版的 Focal 损失表达形式:

$$FL(p_i) = -\alpha_i (1 - p_i)^\gamma \log(p_i) \tag{13}$$

这样 Focal 损失既能调整正负样本的权重, 又能控制难易分类样本的权重。本文将该损失函数应用到文本情感分类任务中处理样本不平衡问题, 得到较好的效果。

3. 实验结果与分析

3.1. 数据集采集标注及预处理

本实验的评论语料是由作者通过八爪鱼采集器在携程网和去哪儿网等平台采集了 24,190 条中文旅游评论构建而成。获取数据集后由 3 名研究人员以盲的形式对数据集进行情感标注, 将表达积极情感的语句标签设置为 1, 消极情感的语句标签设置为 0。经过此处理的评论标签数据才可以作为模型训练的数据。其中包含训练数据集 21,244 条, 测试数据集 515 条。部分实验数据如表 1 所示, 词向量表示如表 2 所示。

Table 1. Partial experimental data

表 1. 部分实验数据

评论	标签
好玩的地标。	1
景色很美, 趣味性很强, 性价比很高, 值得欣赏。	1
又一个徒有虚名的景点, 一个破池子而已。	0

Table 2. Word vector representation of selected data
表 2. 部分数据的词向量表示

词向量(维度 50)				
0.079931212	0.061140408	0.541640046	0.884781944	-0.573372103
0.972455561	0.063708975	-0.058366775	0.631188750	0.108583219
0.159094184	-0.231376662	0.017554214	-0.182645156	-0.251457571

3.2. 评论文本预处理

由于采集的数据集是不平衡样本, 积极评论远多于消极评论, 所以对数据集进行了预处理并构建字典和数字特征列表, 得到模型训练数据, 具体过程如下:

- 1) 对数据集进行预处理, 筛除去部分不具有情感倾向的样本, 如“该用户评价默认好评”。
- 2) 对评论数据集进行中文分词, 去掉评论文本中的表情包, 标点符号, 以及停用词等, 本文选择的分词工具是 jieba 中文分词库, 最终得到只含有汉字的评论数据集, 并输出每条评论语句相对应的分词列表。
- 3) 构建字典, 字典中包含评论数据集中所有的词汇集。
- 4) 得到评论数据集的字典后, 利用 Word2vec 将数据集中的每条评论文本对应的分词列表转换为数字特征列表。处理后得到的部分 Word2vec 词向量如表 2 所示。
- 5) 规定最长文本数据长度, 该数据集文本数据的最长长度为 169, 再截长补短使得所有样本的长度一致。
- 6) 利用 one-hot 独热编码对评论标签进行转换, 经过这六步处理的评论文本数据才可以作为模型训练的数据。

3.3. 评论情感分类结果及分析

3.3.1. 评价指标

本文采用准确率(accuracy)、F1、召回率(recall)、特异度(specificity)、四个指标作为模型的测评指标, 具体公式如下。其中, F1 和特异度(specificity)是判定样本不均衡模型优劣的重要指标。

$$\text{accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \quad (14)$$

$$\text{F1} = \frac{2\text{TP}}{2\text{TP} + \text{FP} + \text{FN}} \quad (15)$$

$$\text{recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (16)$$

$$\text{specificity} = \frac{\text{TN}}{\text{TN} + \text{FP}} \quad (17)$$

损失函数采用 Focal 损失, 公式为:

$$\text{FL}(p_i) = -\alpha_i (1 - p_i)^\gamma \log(p_i) \quad (18)$$

3.3.2. 分类结果及分析

本文将 Bi-LSTM-Att-Focal loss 分类模型的与 Bi-LSTM 分类模型和 Bi-LSTM-Att 分类模型进行对比, 通过测试集的评价指标判断模型效果的优劣, 另外, 为检测模型的泛化性, 本文还针对不同平台数据进

行了对比实验, 具体结果如下:

- a) 表 3 为训练集与测试集为不同平台数据的结果;
- b) 表 4 为训练集与测试集为同一平台数据的结果。

本实验通过柱状图更直观的表现出三个模型评价指标的优劣, 具体结果见图 4 和图 5。

Table 3. Evaluation indicators for different platform data test sets

表 3. 不同平台数据测试集评价指标

	accuracy	F1	Recall	specificity
Bi-LSTM	82.81	82.92	88.89	78.38
Bi-LSTM-Att	87.5	87.49	90.63	84.38
Bi-LSTM-Att-Focal loss	93.75	93.75	93.33	94.12

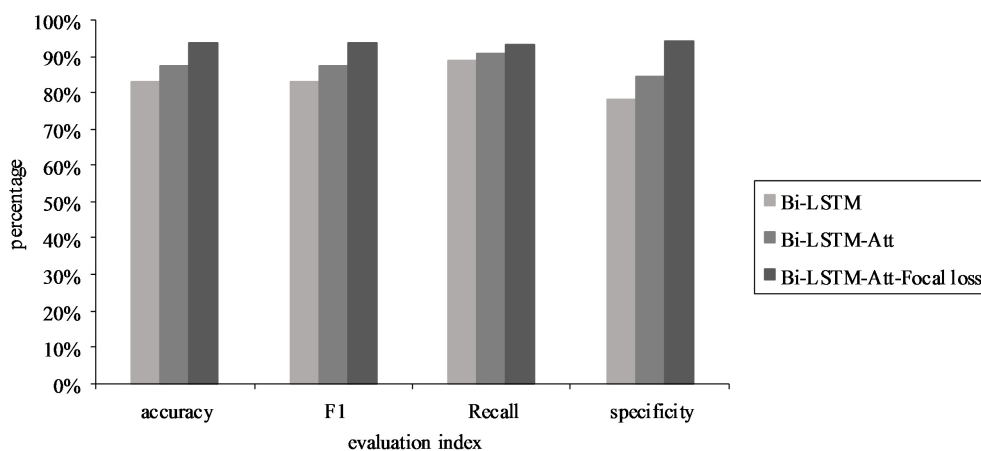


Figure 4. Histogram of evaluation indicators for different platform data test sets

图 4. 不同平台数据测试集评价指标柱状图

Table 4. Evaluation indicators for the same platform data test set

表 4. 同一平台数据测试集评价指标

	accuracy	F1	Recall	specificity
Bi-LSTM	90.63	90.63	92.11	85.19
Bi-LSTM-Att	91.06	92.95	92.5	88.89
Bi-LSTM-Att-Focal loss	95.31	95.32	93.55	96.97

从表 3 和表 4 的模型评价指标来看, 与传统的 LSTM 分类模型相比, Bi-LSTM-Att-Focal loss 分类模型有更优的评价指标。尤其在判断样本不平衡模型优劣的特异度(specificity)指标上, Bi-LSTM-Att-Focal loss 分类模型较其他两个模型有显著提高, 可见该模型确实能改善样本不平衡给分类模型造成的消极影响。另外, 在模型的泛化能力上, 以测试集的评价指标为准, 我们可以看到当训练集与测试集不是同一平台的数据时, 模型的准确率基本在 93%以上, 而当训练集与测试集是同一平台的数据时, 模型的准确率基本在 95%以上, 说明不同数据在该模型下的准确率都是比较高的, 该模型的泛化能力是不错的。

为进一步判断分类模型的优劣, 本实验绘制了三个模型的 loss 随 epoch 变化的折线图, 如图 6 所示。

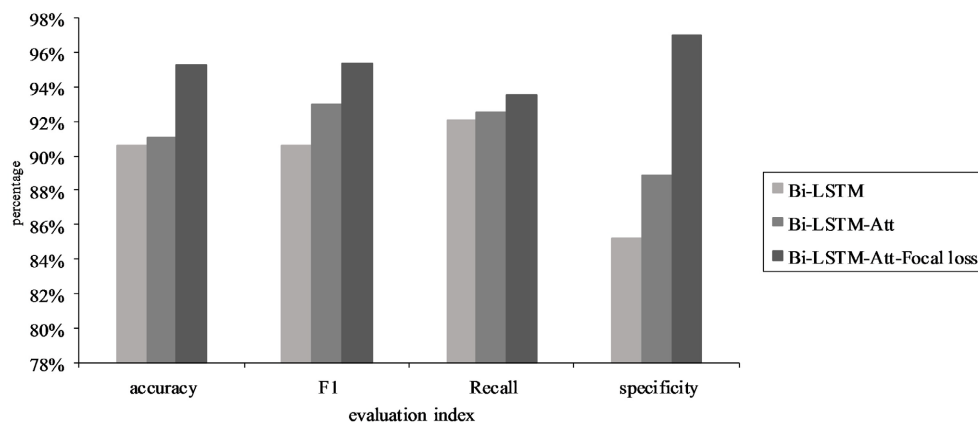


Figure 5. Histogram of evaluation indicators for the same platform data test set
图 5. 同一平台数据测试集评价指标柱状图

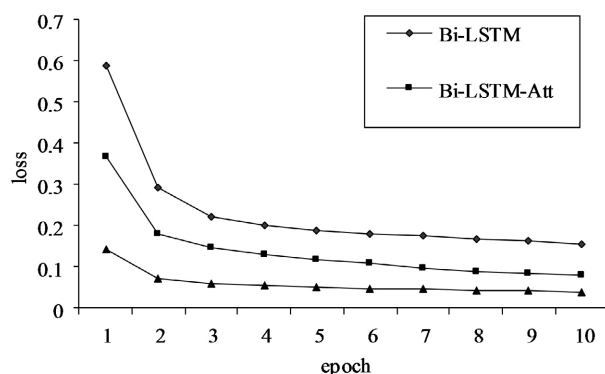


Figure 6. Folding graph of loss with epoch
图 6. loss 随 epoch 变化折线图

从图 6 来看, Bi-LSTM-Att-Focal loss 分类模型与传统 LSTM 分类模型相比有更稳定的 loss 变化, 整体 loss 值较其他两个分类模型显著变小, 可见 Bi-LSTM-Att-Focal loss 分类模型的分类效果更好。通过上述实验反映出的结果来看, 本文所提出的模型指标较另外两个模型有显著提升, 为更好地判断该模型与另外两个模型准确率(accuracy)的均值是否有显著差异, 本文对三个模型进行了假设检验, 假设检验结果如表 5 所示。

Table 5. Hypothesis testing results
表 5. 假设检验结果

	SS	df	MS	F	P-value
Bi-LSTM	0.011	1	0.011	14.44	0.008
Bi-LSTM-Att-Focal loss	0.172		0.17	0.12	0.968
Bi-LSTM-Att	0.007	1	0.007	24.00	0.002
Bi-LSTM-Att-Focal loss	0.8125		0.813	0.384	0.713

从表 5 中可以看出 P 值都小于 0.05, 即拒绝原假设, 说明 Bi-LSTM-Att-Focal loss 模型的准确率较 Bi-LSTM 模型和 Bi-LSTM-Att 模型相比有显著差异。即 Bi-LSTM-Att-Focal loss 模型的性能更优。

4. 结论

本文提出一种基于 Focal 损失的 Bi-LSTM 网络模型, 来处理常见的不平衡数据问题。方法中首先构建了旅游评论数据集, 采用欠采样方法对数据进行处理, 然后利用基于 Focal 损失的 Bi-LSTM 网络模型对训练集进行训练, 最后使用训练好的模型对测试集进行测试并得到分类结果。最终的实验结果表明, 文中提出的方法确实能够更好地处理不平衡样本的情感分类任务, 该方法相比于传统 LSTM 分类模型性能有显著提高。

本文只是提出了一种从模型上处理不平衡样本情感分类问题的方法, 还可以从其他角度出发来解决不平衡样本的情感分类问题。未来工作中, 会考虑从改进词嵌入[19]或样本增广等方面处理不平衡数据情感分类问题。我们会在接下来的时间里研究这个问题, 力求提升情感分类器的性能。

参考文献

- [1] 林夕, 陈孜卓, 王中卿. 基于不平衡数据与集成学习的属性级情感分类[J]. 计算机科学, 2022, 49(S1): 144-149.
- [2] 杜启明, 李男, 刘文甫, 等. 结合上下文和依存句法信息的中文短文本情感分析[J]. 计算机科学, 2023, 50(3): 307-314.
- [3] 王娅丽, 张凡, 余增, 等. 基于交互注意力和图卷积网络的方面级情感分析[J/OL]. 计算机科学: 1-16. <http://kns.cnki.net/kcms/detail/50.1075.TP.20221228.1124.001.html>, 2023-03-29.
Xu, G., Yu, Z., Yao, H., et al. (2019) Chinese Text Sentiment Analysis Based on Extended Sentiment Dictionary. *IEEE Access*, 7, 43749-43762. <https://doi.org/10.1109/ACCESS.2019.2907772>
- [4] Barandela, R., Sanchez, B.J.S., Garcia, V., et al. (2003) Strategies for Learning in Class Imbalance Problems. *Pattern Recognition*, 36, 849-851. [https://doi.org/10.1016/S0031-3203\(02\)00257-1](https://doi.org/10.1016/S0031-3203(02)00257-1)
- [5] 郭朝有, 许喆, 马砚堃, 等. 面向不平衡数据集融合 Canopy 和 K-means 的 SMOTE 改进算法[J]. 科学技术与工程, 2020, 20(22): 9069-9074.
- [6] 梁越, 刘晓峰, 李权树, 等. 面向司法文本的不均衡小样本数据分类方法[J]. 计算机应用, 2022, 42(S2): 118-122.
- [7] 任金金, 汪彪彪. 基于神经网络的不均衡数据分类方法[J]. 长江信息通信, 2022, 35(12): 70-72.
- [8] Tanha, J., Abdi, Y., Samadi, N., et al. (2020) Boosting Methods for Multi-Class Imbalanced Data Classification: An Experimental Review. *Journal of Big Data*, 7, 1-47. <https://doi.org/10.1186/s40537-020-00349-y>
- [9] 杨敏, 李君轶, 徐雪. ICTs 视角下的旅游流和旅游者时空行为研究进展[J]. 陕西师范大学学报(自然科学版), 2020, 48(4): 46-55.
- [10] 史沛卓, 陈凯天, 钟叶珂, 等. 基于 Text CNN 的中国古诗文分类方法研究[J]. 电子技术与软件工程, 2021(10): 190-192.
- [11] Wang, Y.S., Sohn, S., Liu, S.J., et al. (2019) A Clinical Text Classification Paradigm Using Weak Supervision and Deep Representation. *BMC Medical Informatics and Decision Making*, 19, 1-13. <https://doi.org/10.1186/s12911-018-0723-6>
- [12] 李文亮, 杨秋翔, 秦权. 多特征混合模型文本情感分析方法[EB/OL]. 计算工程与应用: 1-12. <https://kns.cnki.net/kcms/detail/11.2127.TP.20210621.1806.004.html>, 2021-07-02.
- [13] 唐明, 朱磊, 邹显春. 基于 Word2Vec 的一种文档向量表示[J]. 计算机科学, 2016, 43(6): 214-217.
- [14] Hochreiter, S. and Schmidhuber, J. (1997) Long Short-Term Memory. *Neural Computation*, 9, 1735-1780. <https://doi.org/10.1162/neco.1997.9.8.1735>
- [15] Zhou, P., Shi, W., Tian, J., et al. (2016) Attention-Based Bidirectional Long Short-Term Memory Networks for Relation Classification. *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics*, 2, 207-212. <https://doi.org/10.18653/v1/P16-2034>
- [16] Mukhoti, J., Kulharia, V., Sanyal, A., et al. (2020) Calibrating Deep Neural Networks Using Focal Loss. *Advances in Neural Information Processing Systems*, 33, 15288-15299.
- [17] Lin, T.Y., Goyal, P., Girshick, R., et al. (2017) Focal Loss for Dense Object Detection. 2017 *IEEE International Conference on Computer Vision (ICCV)*, Venice, 22-29 October 2017, 2999-3007. <https://doi.org/10.1109/ICCV.2017.324>
- [18] Pang, T., Xu, K., Dong, Y., et al. (2019) Rethinking Softmax Cross-Entropy Loss for Adversarial Robustness. arXiv

preprint arXiv: 1905.10626.

- [19] 张小川, 刘连喜, 戴旭尧, 等. 基于词性特征的 CNN_BiGRU 文本分类模型[J]. 计算机应用与软件, 2021, 38(11): 155-161.