

文章编号: 2095-2163(2020)05-0125-08

中图分类号: 929.5

文献标志码: A

神经网络模型在显式与隐式特征下的情感分类应用研究

张乐江¹, 黄宇涵², 李晓坤¹, 张伟华¹, 赵中英², 陈虹旭¹

(1 黑龙江恒讯科技有限公司国家博士后科研工作站, 哈尔滨 150090; 2 山东科技大学 电子信息工程学院, 青岛 266590)

摘要: 将 BLSTM、GRU 和 BGRU 三种神经网络模型应用于中文文本情感分类, 系统研究三种神经网络模型在不同特征下的参数选择和情感分类结果。以具有显式和隐式情感特征的中文评论数据为研究对象, 将具有不同情感特征的中文文本进行向量化表示, 并将其作为 BLSTM、GRU 和 BGRU 三种神经网络情感分析模型的输入; 通过模型设置和参数优化, 实现情感分类的预测和对比分析。实验结果表明, 基于神经网络的情感分析模型对显式情感特征的评论文本数据的准确率达到 87% 以上; 相反, 对隐式情感特征的评论文本数据的准确率则只有 79% 左右, 说明基于神经网络的情感分析模型在隐式情感特征分析中的表现有待提高。神经网络模型在文本情感分类任务中的性能良好, 验证了神经网络对复杂文本任务具有很强的学习能力, 但其对隐式情感特征的分析尚需进一步深入研究。

关键词: 中文文本; 情感特征; 神经网络; 情感分类

The Application of Neural Networks on Sentiment Classification with Explicit and Implicit Emotional Features

ZHANG Lejiang¹, HUANG Yuhan², LI Xiaokun¹, ZHANG Weihua¹, ZHAO Zhongying², CHEN Hongxu¹

(1 Heilongjiang Hengxun Technology Co. Ltd. Postdoctoral Programme, Harbin 150090, China;

2 College of Electronic and Information Engineering, Shandong University of Science and Technology, Qingdao 266590, China)

[Abstract] Three neural network models, BLSTM, GRU and BGRU, were applied to the emotion classification of Chinese text, and the parameter selection and emotion classification results of the three neural network models under different characteristics were systematically studied. Taking Chinese comment data with explicit and implicit affective characteristics as the research object, Chinese text with different affective characteristics is vector graphically represented and used as the input of BLSTM, GRU and BGRU neural network affective analysis models. Through model setting and parameter optimization, the prediction and comparative analysis of emotion classification are realized. The experimental results show that the accuracy of the neural network based emotion analysis model is over 87% for the explicit emotion characteristics of the comment text data. On the contrary, the accuracy rate of the comment text data on implicit emotion characteristics is only about 79%, which indicates that the performance of the emotion analysis model based on neural network in implicit emotion characteristics analysis needs to be improved. The performance of neural network model in text emotion classification task is good, which proves that neural network has strong learning ability to complex text task, but its analysis of implicit emotion characteristics still needs further research.

[Key words] Chinese text; emotional features; neural network; sentiment classification

0 引言

随着 Web2.0 技术的广泛应用, 互联网中存在着海量的文本数据, 其中与人们的日常生活最为相关的是各类评论数据, 例如酒店评论数据、商品评论数据、热门事件评论数据等。这些数据具有内容简短、观点情感色彩强烈等特点, 同时也存在着情感特

征不明显的隐式情感表达的数据。对这些评论数据进行情感分析有助于评论阅读者对酒店口碑、商品好坏、事件性质等评论主体做出判断和舆情分析, 如果单纯以人工的方式从海量的评论数量和内容中到得评论文本的情感导向十分困难, 利用自然语言处理领域中的相关技术如情感分析可以解决这类问

基金项目: 中小企业创新基金(2017FF1GJ023); 专利优势示范企业基金(2017YBQCZ029); 国家自然科学基金(81273649); 国家自然科学基金(61501132); 中央高校基本科研业务费专项资金(3072019CFT0603); 国家自然科学基金(61672181); 黑龙江省自然科学基金联合引导基金(LH2019F049, LH2019A029); 中国博士后科学基金(2019M650069); 黑龙江省基础科研科技创新基金(KJXC201805); 黑龙江省基础科研青年创新团队基金(RCYJTD201805)。

作者简介: 张乐江(1988-), 男, 硕士, 高级工程师, CCF 会员, 主要研究方向: 物联网、云计算、人工智能等; 李晓坤(1979-), 男, 硕士, 研究员级高级工程师, 教授, CCF 高级会员, 主要研究方向: 虚拟化、人工智能、生物特征识别等; 张伟华(1993-), 女, 本科生, 主要研究方向: 智慧医疗、人工智能、生物特征识别等。

通讯作者: 李晓坤 Email: li.xiaokun@163.com

收稿日期: 2020-03-05

题。具体可以分为基于情感词典(特征)的情感分析^[1-4],基于机器学习方法的情感分析^[5-11]和基于深度学习的情感分析^[12-18]3个方面的研究。要解决这个问题核心是构建情感词典。目前常用的情感词典很多,比如知网词典 HowNet^[19]、台湾大学 NTUSD^[20]、清华大学李军中文褒贬义词典^[21]、否定词词典^[22]。随着深度学习方向的发展,使用基于深度学习的方法进行情感分析成为了当前研究的热点,基于深度学习的句子级分布式表示算法,能够通从词语分布式表示复合得到句子或段落分布式表示,从而能够将句子或段落分布式表示应用于文本情感分析任务。卷积神经网络模型^[14]、循环神经网络模型^[15]和递归神经网络模型^[23]开始在情感分析中广泛应用。

针对不同情感特征的文本数据,本文以具有显式和隐式情感特征的中文评论数据作为研究对象,将具有不同情感特征的中文文本进行向量化表示,并将其作为基于 BLSTM、GRU 和 BGRU 三种神经网络情感分析模型的输入;通过模型设置和参数优化实现情感分类的预测和比较分析。

1 相关工作

现有的文本情感分析技术可以分为三类:基于词典(特征)的情感分析、基于机器学习方法的情感分析和基于深度学习的文本情感分析技术。以上三类技术已经相对成熟,本文将从文本情感分析的3个方向进行相关的工作。

(1)基于词典(特征)的情感分析。Chao 等人使用汉字中的偏旁部首作为单词特征,并通过单字特征和双字特征点互信息比较,完成对文本的情感极性分类;El 等人构造了一个动态变化的单词情感急性字典,通过引入新的数据对字典中各单词表达的情感倾向进行微调;Yu 等人改进文本中每个单词的词向量,使其在情感词典中更接近语义上和情感上相似的词,并远离在情感表达方面不相似的词;Zhang 等人将情感词典与模糊卡诺模型相结合,获得用户的有关产品的不同态度方面的细粒度的情绪分析。基于词典的情感分析的关键是如何构建一个符合数据特点的情感词典。

(2)基于机器学习方法的情感分析。Alkubaisi 等人使用混合朴素贝叶斯分类器判断推特中和股票相关的文本的情感极性。Asgar 等人设计四个分类器:俚语分类器、表情符号分类器、感知网络分类器和改进的领域分类器对推特文本进行情感极性分类;Fang 等人针对传统的机器学习方法不能很好地

表达中文文本观点的问题,提出了一种语义模糊的多策略情感分析方法,对文本中的情感极性进行判断。彭云等人从单词的句法、语义、语境等层面入手,学习单词之间的语义关系,然后将这些关系作为先验知识嵌入到 LDA 中,学习文本中单词层面的情感表达;Liu 等人基于潜在语义分析识别出文本中的特征,并使用感知器,根据得到的特征对文本进行情感分析;Bang 等人设计出一种句子依存树结构来减轻词义的歧义,并解决词义的固有多义性,同时,句子依存树还可以对句子的情感极性进行识别;He 等人基于一个改进的非参数贝叶斯模型来估计能够完美解释当前时间片的最佳主题数量,并同时分析这些潜在的主题及其情感极性。这类方法是传统的数据挖掘思想在情感分析领域的应用,相关研究比较多。

(3)基于深度学习的文本情感分析技术。Pham 等人使用多层知识表示体系结构,表示文本中的不同情感级别,并将这些情感表示集成到神经网络中,完成对文本中各情感表达方面的情感评价;Zhao 等人研究了对文本情感分析进行建模的低层网络结构,即卷积特征提取和长短时记忆,通过使用神经网络学习文本的一个高级表示(嵌入空间)的形式完成对文本中表达的情感极性的判断;Xu 等人使用两种预训练嵌入(通用嵌入和领域特定嵌入)对文本中的单词特征进行扩展,基于两种嵌入方式设计双嵌入层的 CNN 模型,推理出文本的情感极性;Zhao 等人设计关系嵌入和子树嵌入对单词特征进行表示,并使用 RNN 对两种嵌入进行建模,学习句法路径中每个词的情感搭配,最后通过得到的情感搭配完成情感分析工作;Cong 等人基于双向 LSTM 对文本中的情感极性进行建模;Wang 等人提出了一种基于注意力的 LSTM 模型进行情绪分类。当不同的情感表达作为输入时,注意机制可以将模型的注意力集中该情感中;Ma 等人设计了结合特征表示和单词嵌入的组合策略来增强注意力机制,并设计基于特征的复合内存网络,解决细粒度的情感识别问题。随着深度学习研究的发展,采用深度神经网络进行情感的分析成为了研究人员关注的热点之一。

本文基于深度学习的情感分析技术,以深度神经网络中的 BLSTM、GRU 和 BGRU 三种模型为基础,构建深度神经网络情感分析模型,并以显式和隐式情感特征文本情感数据为对象,深入比较和分析模型在两种情感特征文本情感分析方面的效果。

2 基于深度神经网络的情感分析模型

深度神经网络可以发现文本中单词之间的关系和潜在的语义特征,使用神经网络对文本中的单词特征进行提取,动态地学习输入序列的特征,并保持一定的记忆能力,即前一时刻的记忆能力可以融入到当前时刻的计算过程中,保留了单词序列之间的联系。

本文使用 Word2vec 训练词向量,并将其作为输入,探究了基于 BLSTM、GRU 和 BGRU 三种神经网络的情感分析模型,并对这些模型的优缺点进行比较。首先,对文本进行预处理,即对文本进行分词和对停顿词等无用词汇进行去除;其次,根据单词表对单词进行索引化表示,并构建映射矩阵,使用词嵌入模型对单词进行向量化表示;最后,将得到的词向量作为神经网络的输入序列,学习文中单词表达出的情感特征,并经过激活函数得出文本的情感分类。

2.1 文本预处理

在使用神经网络进行文本情感分析时,需要将非结构化的文本数据转变为可供神经网络计算的矩阵数据。本文通过对文本进行分词、索引化转换等步骤为神经网络的训练准备合适的输入序列。

(1)分词。单词分词的精度对词向量的生成有着直接的影响,同时也影响着文本情感分析的结果。中文文本单词存在的交叉歧义和组合歧义的特点,给分词工作带来的困难,本文使用 jieba 分词对中文文本进行分词,jieba 分词使用大量语料对中文单词进行训练,结合字典树对中文单词进行分类,从而达到分词的效果。

另外,为了降低停顿词、标点符号等对情感分析无实际影响的单词或字符对词向量生成过程的干扰,在对文本进行分词时,首先使用 re.sub() 函数去除停顿词、标点符号等无效字符,之后使用 jieba 分词器对文本进行分词,并将其转化为 list 的数据结构存储数据。

(2)索引化。对文本进行索引化,将文本中的句子转换为一段数字索引序列,使得计算机可以识别文本数据。但是本文使用的文本数据类型为评论数据,评论数据随意性的特点导致文本数据长度差别较大,索引序列预设的长度过长会增加数据冗余、浪费计算机资源;过短会造成数据缺失,影响训练效果。因此,本文对文本数据进行分析,为文本索引化寻找合适的索引长度。基于本文使用的文本数据,对数据中的语句长度进行统计,统计出索引长度涵盖的样本范围,随着索引长度的变化而产生的变化

情况,统计结果(部分)如表 1 所示。由表 1 可知,当索引长度取值为 236 时,涵盖的样本数量最多,因此本文选取 236 作为词向量训练使用的索引长度。

表 1 索引长度涵盖样本范围随索引长度变化的变化

Tab. 1 Index length covers the variation of sample range with index length

索引长度	涵盖样本范围/%
234	95.55
235	95.50
236	95.65
237	95.60
238	95.58

2.2 基于 BLSTM 的情感分析

BLSTM(Bi-directional LongShort-Term Memory, BLSTM)由前向 LSTM 和后向 LSTM 组合而成,常常被用来建模上下文信息^[24-26]。BLSTM 针对 LSTM 无法从后向前对上下文信息进行编码的缺点,通过双向编码的方式更好地捕捉文本中的上下文信息。

BLSTM 对每一个输入序列进行前向运算和后向运算,无论前向还是后向传播,其计算过程都与 LSTM 的过程相同。即在前向 LSTM 和后向 LSTM 中,对于输入序列中的每一个单词 w_t ,输入当前 t 时刻 w_t 的词向量 x_t 、 $t-1$ 时刻的隐藏状态 h_{t-1} 和记忆状态 c_{t-1} ,按照公式(1)~(6)得到当前 t 时刻的记忆状态 $c_t \in R^n$ 和隐藏状态 h_t :

$$i_t = \sigma(W_i x_t + U_i h_{t-1} + b_i), \quad (1)$$

$$f_t = \sigma(W_f x_t + U_f h_{t-1} + b_f), \quad (2)$$

$$o_t = \sigma(W_o x_t + U_o h_{t-1} + b_o), \quad (3)$$

$$g_t = \tanh(W_g x_t + U_g h_{t-1} + b_g), \quad (4)$$

$$c_t = f_t \odot c_{t-1} + i_t \odot g_t, \quad (5)$$

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

其中, σ 表示损失函数。当 BLSTM 用于情感分析时,首先使用单向 LSTM 按照上述公式分别从前和后对输入序列进行编码,将得到的两种编码序列进行拼接得到输入的 BLSTM 序列,最后使用激活函数对输入序列的情感极性进行判断。

本文使用 BLSTM 进行情感分析模型的训练,结果如图 1 所示,BLSTM 模型有 78,055,349 个参数。BLSTM 模型第二层为双向结构的 LSTM 层,LSTM 层的输出结果进入单向的 LSTM 层以及激活函数层,通过激活函数得到最终的情感分析结果。

2.3 基于 GRU 的情感分析

GRU 模型针对 LSTM 复杂的记忆单元结构进行改进,GRU 模型在记忆单元的设计中将输入门和

遗忘门组合为更新门,GRU 更新门的设计不但保留了文本中的上下文信息,同时简化了模型结构^[27-28]。

Layer (type)	Output Shape	Param #
embedding (Embedding)	(None, 280, 300)	7764000
gru (GRU)	(None, 280, 94)	49498
gru_1 (GRU)	(None, 280, 16)	2382
gru_2 (GRU)	(None, 1)	2382
dense (Dense)	(None, 1)	5

Total params: 77,999,477
Trainable params: 34,577
Non-trainable params: 77,964,900

图1 BLSTM 情感分析模型结构

Fig. 1 BLSTM Structure of emotion analysis model

GRU 使用一个门控回归单元,使得每一个单元可以自动捕获不同时间和不同尺度的依赖关系。同时,GRU 中的参数按照公式(7)~(10)进行参数更新:

$$r_t = \sigma(W_r x_t + U_r h_{t-1}), \quad (7)$$

$$z_t = \sigma(W_z x_t + U_z h_{t-1}), \quad (8)$$

$$\hat{h}_t = \tanh(W_x x_t + U_x (r_t \odot h_{t-1})), \quad (9)$$

$$h_t = ((1 - z_t) h_{t-1} + z_t \hat{h}_t). \quad (10)$$

其中, σ 表示损失函数; r_t 表示 GRU 中的重置门; z_t 表示更新门; \hat{h}_t 表示中间隐藏层; \odot 表示元素相乘。使用 GRU 进行情感分析时,更新门和重置门学习输入序列中的文本特征,并通过编码的方式对输入序列中的文本特征进行表示,使用激活函数完成对输入序列的情感极性分类。

本文使用的 GRU 模型结构训练结果如图 2 所示。由图 2 可知,GRU 模型共使用了 77,999,477 个参数。使用三层 GRU 结构搭建基于 GRU 的情感分析模型,每层 GRU 结构的神经元个数呈下降趋势。

Layer (type)	Output Shape	Param #
embedding (Embedding)	(None, 280, 300)	7764000
bidirectional (Bidirectional)	(None, 280, 94)	49498
lstm (LSTM)	(None, 16)	2382
dense (Dense)	(None, 1)	5

Total params: 78,045,340
Trainable params: 34,427
Non-trainable params: 77,984,900

图2 GRU 情感分析模型结构

Fig. 2 GRU Structure of emotion analysis model

2.4 基于 BGRU 的情感分析

GRU 在结构上对 LSTM 实现了优化,但是 GRU 单向编码的特点使其不能学习到更多的语义信息,因此,本文使用 BGRU 模型对中文文本进行情感分

析。与 BLSTM 设计思路一致,BGRU 采用双向编码的形式学习输入序列的上下文信息,并通过激活函数得到文本的情感分类结果^[29-30]。

前向 GRU 从左至右读取输入序列并通过若干个 GRU 单元获取输入序列的前向隐藏状态: $(h_{11}, h_{12}, h_{13}, \dots, h_{1n})$ 。同样地,后向 GRU 从右至左获取输入序列的另一种隐藏状态: $(h_{r1}, h_{r2}, h_{r3}, \dots, h_{rn})$ 。其中, h_{1n} 和 h_{rn} 的更新策略同公式(7)~(10),最后,BGRU 将得到的前向隐藏状态 h_{1n} 和后向隐藏状态 h_{rn} 进行串联得到 BGRU 的输出隐藏状态:

$$h_n = [h_{1n}, h_{rn}]. \quad (11)$$

本文使用的 BGRU 模型结构训练结果如图 3 所示。由图 3 可知,BGRU 模型共有 78,037,061 参数,BGRU 模型将第二和第三层的 GRU 结构改为双向 GRU 结构,神经元个数逐层下降。

Layer (type)	Output Shape	Param #
embedding_1 (Embedding)	(None, 280, 300)	7764000
bidirectional_1 (Bidirectional)	(None, 280, 94)	49498
bidirectional_2 (Bidirectional)	(None, 280, 94)	49498
gru_1 (GRU)	(None, 1)	2382
dense_1 (Dense)	(None, 1)	5

Total params: 78,037,061
Trainable params: 74,381
Non-trainable params: 77,964,680

图3 BGRU-情感分析模型结构

Fig. 3 BGRU-Structure of emotion analysis model

3 深度神经网络情感分析模型在显式情感特征数据上的分析

本文利用公开数据集对三种神经网络情感分析模型进行实验和对比分析。

3.1 数据集

本文选用中科院计算所谭松波老师标注的酒店评论语料库作为三种神经网络情感分析模型训练和测试使用的数据集。选用的酒店评论语料库共有 4 000 条语料,包含正向评论 2 000 条和负向评论 2 000 条。语料库样本示例,如表 2 所示。

3.2 实验设置

本文使用 Tensorflow 和 Keras 作为神经网络模型的训练平台。Tensorflow 具有高度的灵活性和丰富的算法库。Keras 具有简洁易懂的 API 和用途广泛的模块函数。本文从 keras.model 导入 Sequential 类,同时将 model 设置为 Sequential,之后使用 add() 函数为本文使用的 BLSTM、GRU 和 BGRU 三种神经网络添加网络层。

本文基于如下分析对神经网络的训练参数进行设置。

表 2 酒店评论语料示例

Tab. 2 Examples of Hotel Review corpus

语料	情感极性
这地方千万别去住了,绣花枕头一个。样子好看。别的没什的	0
房间还算干净整洁,服务也可以,以这个价格来说,不错了	1
建议携程不要和这家酒店合作,名曰三星,要我看准星级都勉强!	0
酒店环境很好,地点方便,服务好,下回继续住!	1
订了酒店房间保险链是坏的,半天也没有服务员来,还是自己修的!	0
最滥的酒店,什么都不行。这种酒店千万别去。80年代的 3 星	0
地理位置不错,在市中心繁华地段,去天津街、清泥洼都很方便。	1
98 年这家酒店开业时住过感觉很好今年去青岛还是选择住在这里	1

(1)最佳 batchsize 和 epochs 个数。batch size 指的是一个 batch 中存在的样本总数,一个 batch 表示的是将数据集均分成的子集中的一个子集,表示训练集显示的神经网络运行次数。本文以 20 步长为单位,从 10~100 的数据规模中评估不同 batchsize 对神经网络模型的影响,结果如图 4 所示。当 batchsize 为 100,epochs 为 20 时,模型的训练效果最佳。因此,本文将 batchsize 参数调整为 100,epochs 参数调整为 20。

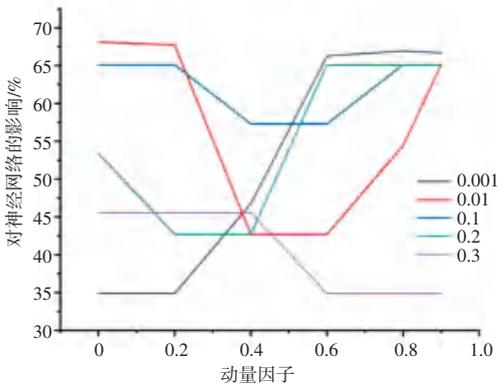


图 4 不同 batch size 和 epochs 对神经网络模型的影响

Fig. 4 Influence of different batch size and epochs on neural network model

(2)优化算法的选择。为帮助神经网络获得更快的收敛速度,选取 Adam 优化器,结合 AdaGrad+RMSProp 两种优化算法的优点,能训练较为复杂的神经网络。Adam 优化器实现简单、参数

的更新不受梯度影响、步长的更新具有较弱的波动性,因此本文选用 Adam 优化器作为神经网络优化算法。

(3)优化学习率与动量因子。对神经网络进行训练时,当一个 batch 结束时,学习率会控制神经网络权重的更新,动量因子决定上一个 batch 权重对当前 batch 权重产生的影响程度。为了选取最佳的优化学习率和动量因子,本文探究了不同优化学习率和动量因子对神经网络的影响,结果如图 5 所示。由图 5 可知,学习速率在 0.01、动量因子在 0.0~0.2 之间时,模型效果最佳。

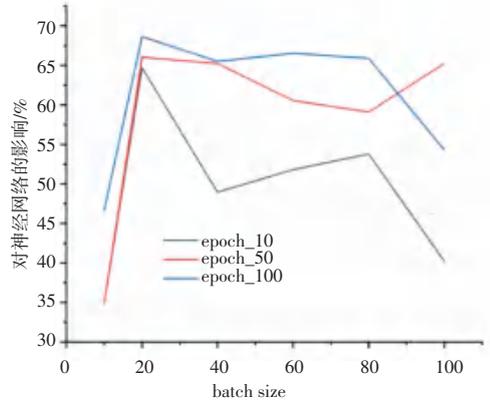


图 5 学习率和动量因子对神经网络模型的影响

Fig. 5 The influence of learning rate and momentum factor on neural network model

当神经网络完成对输入序列的编码时,需要使用激活函数判断输入序列的情感极性,情感极性取值为正向情感和负向情感。本文选用目前较为流行的 Sigmoid 函数作为本文神经网络判断情感极性时使用的激活函数。

3.3 实验结果分析

本文使用酒店评论语料库对三种神经网络情感分析模型的性能进行评价,将 2000 条正向评论的情感倾向标记为 1,表示好评;2000 条负向评论的情感倾向标记为 0,表示差评。从酒店评论语料库中随机选取 90%的样本数据作为训练集,余下 10%的评论数据作为测试集,使用准确率作为三种模型的评价指标。

(1)三种神经网络情感分析模型结果比较。表 3 展示了三种神经网络情感分析模型在酒店评论语料中的表现,其中 Early Stopping 时刻表示模型是否需要运行完所有的预设 epochs(预设值为 20),若为否,则输出 Early Stopping 时刻。通过表 3 所示结果可以得出,BLSTM 和 GRU 在情感分析方面的表现相差不大,但是 BLSTM 使用了 20 个 epoch 进行训

练,而 GRU 在第 10 个 epoch 时已经获得较好的准确率,GRU 简化了 LSTM 中的结构,使得 GRU 具有比 BLSTM 模型更高的效率。同时,使用双向结构的 BGRU 获得了比 GRU 更丰富的单词上下文信息,因此,BGRU 模型具有比 GRU 模型更好的情感倾向预测能力。

表 3 三种神经网络情感分析模型结果比较

Tab. 3 Comparison of the results of three neural network emotion analysis models

模型名称	准确率/%	Early Stopping 时刻
BLSTM	87.75	无
GRU	88.00	第 10 个 epoch
BGRU	90.25	第 11 个 epoch

(2) 训练集与测试集的比例对准确率产生的影响。为了验证训练集规模对模型预测能力的影响,本文使用 9:1 的比例对实验所用酒店评论语料进行训练集和数据集规模划分的同时,又使用 8:2、7:3、6:4、5:5 的比例分别对酒店评论语料进行划分,验证在不同训练集和测试集规模的影响下,三种神经网络情感分析模型对情感倾向预测的准确率的变化,结果如图 6 所示。由图 6 所示结果得出,训练集数据规模的增加有助于模型学习到更多的单词上下文信息,使得模型可以对样本的情感倾向做出更准确的判断。

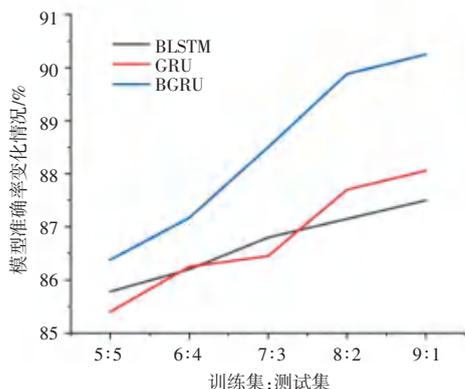


图 6 不同训练集与测试集比例下三种模型的准确率变化

Fig. 6 The change of accuracy of three models under different training and test set ratios

4 深度神经网络情感分析模型在隐式情感特征数据上的分析

人类的语言拥有复杂和多样的特点,从日常的情感表现方式上来说,大体将情感分析分为显式以及隐式的情感分析。在实际的生活中,通常会使用不同的修辞手法、说反话、反问、隐喻的方式来隐晦地表达情感,这样的隐式情感文本往往没有情感词

来为模型作为指导,而且与语境有关,很难统一判断标准,显然会比显式情感更加难以判断。使用基于深度神经网络 BLSTM、GRU、BGRU 构建的三种情感分析模型进行隐式文本的训练,测试模型判断隐式情感的准确率。

4.1 隐式情感特征数据集

模型训练和测试使用的隐式情感文本来自新浪微博、去哪儿网、携程、大众点评、京东、淘宝等,主题包括了生活、娱乐、节日等方面。人工对这些评论数据进行预处理和情感倾向标注后,得出评论数据中共有 7785 条。其中,正面隐式情感 3828 条,负面隐式情感 3957 条,隐式情感分析语料示例如表 4 所示。

表 4 隐式情感分析语料示例

Tab. 4 Examples of implicit affective analysis corpus

语料	情感极性
节目看完,感同身受,很多那个时代下的背景元素	1
别以为政治与你无关,有人送我蒙牛的产品我会以为对方要害我!	0
来动物园的人不是来看动物的,是来看人的	0
初到成都,被这里的雾霾深深的吓到了[衰]来给成都人民吸霾来了	0
该湾湾长沙细,岸上绿树如带,构成三亚海滨风景区美丽动人的风景线。	1
位于西安市中心,钟楼鼓楼遥想相望,声闻于天。	1
不要吓领导,吓坏了以后新闻里只报道领导看戏	0
锁定@辽宁卫视 @江苏卫视,引爆你的笑点!	1

4.2 实验设置

与显式情感特征数据中的实验设置类似,使用 Tensorflow 和 Keras 作为 BLSTM、GRU、BGRU 三种神经网络模型的训练平台。为了使显式特征数据和隐式特征数据的实验结果具有对照性,在隐式特征数据的实验中使用和显式情感特征数据实验阶段相同的训练参数,即:batchsize = 100; epochs = 20; 学习速率 = 0.01; 动量因子 = (0.0, 0.2]; 优化器选用 Adam 优化器; 激活函数选用 Sigmoid 函数。

4.3 实验结果分析

图 7 显示了三种神经网络情感分析模型对主题评论文本的隐式情感分析的结果,同时加入三种模型在显式情感特征数据上的实验结果作为对照。通过图 7 展示的结果得出,在隐式情感特征数据中,三种模型表现出相似的预测能力,证明对于隐式评论文本,BLSTM、GRU、BGRU 三种神经网络具有相似的单词语义捕捉能力。另外,将隐式和显式情感特征数据的实验结果进行比较,三种模型在显式文本

的预测精度都高于隐式文本,证明 BLSTM、GRU、BGRU 三种神经网络对表达清晰的文本具有更好的提取能力。三种模型在隐式和显式情感特征数据中的差值分别为 9.19%、10.08%、11.69%,证明在 BLSTM、GRU、BGRU 三种神经网络中,BGRU 更适用于分析情感倾向明显的评论文本。

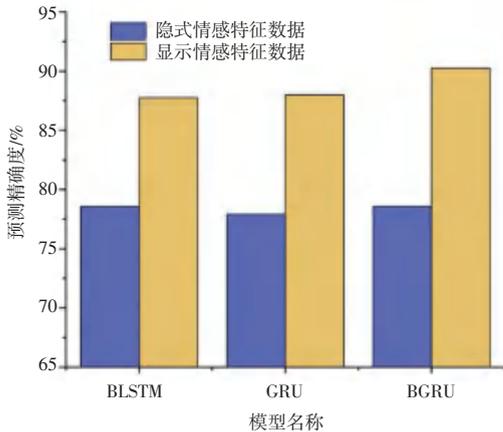


图7 在显式和隐式特征数据中,三种神经网络情感分析模型的预测精度

Fig. 7 Prediction accuracy of three neural network emotion analysis models in explicit and implicit feature data

5 结束语

本文使用神经网络对中文文本中显式情感特征和隐式情感特征进行分析。分别使用 BLSTM、GRU、BGRU 三种神经网络对中文评论文本进行情感极性的预测,分析三种基于不同神经网络的情感分析模型之间的优点和差异。BLSTM 和 BGRU 通过双向编码的形式获取单词丰富的上下文信息,GRU 简化了 BLSTM 中的结构。通过三种深度神经网络情感分析模型在显式和隐式特征文本数据的对比实验分析发现,基于神经网络的情感分析模型对于隐式评论文本的预测能力还有待提高,下一步工作将对隐式评论文本中的情感极性进行更深入的研究。

参考文献

[1] CHAO A F Y, YANG H L. Using Chinese Radical Parts for Sentiment Analysis and Domain-Dependent Seed Set Extraction [J]. Computer Speech & Language, 2018, 47: 194-213.

[2] ELALAOUI I, GAHI Y, MESSOUSI R, et al. A Novel Adaptable Approach for Sentiment Analysis on Big Social Data [J]. Journal of Big Data, 2018, 5(1): 12-29.

[3] YU L C, WANG J, LAI K R, et al. Refining Word Embeddings Using Intensity Scores for Sentiment Analysis [J]. IEEE/ACM Transactions on Audio, Speech, and Language Processing, 2017, 26(3): 671-681.

[4] ZHANG J, CHEN D, LU M. Combining Sentiment Analysis With a Fuzzy Kano Model for Product Aspect Preference Recommendation [J]. IEEE Access, 2018, 6: 59163-59172.

[5] ALKUBAISI G A A J, KAMARUDDIN S S, HUSNI H. Stock

Market Classification Model Using Sentiment Analysis on Twitter Based on Hybrid Naive Bayes Classifiers [J]. Computer and Information Science, 2018, 11(1): 52-64.

[6] ASGHAR M Z, KUNDI F M, AHMAD S, et al. T-SAF: Twitter Sentiment Analysis Framework Using a Hybrid Classification Scheme [J]. Expert Systems, 2018, 35(1): 12233-12251.

[7] FANG Y, TAN H, ZHANG J. Multi-strategy Sentiment Analysis of Consumer Reviews Based on Semantic Fuzziness [J]. IEEE Access, 2018, 6: 20625-20631.

[8] 彭云, 万常选, 江腾蛟, 等. 基于语义约束 LDA 的商品特征和情感词提取 [J]. 软件学报, 2017, 28(3): 676-693.

[9] LIU C L, HSAIO W H, LEE C H, et al. Movie Rating and Review Summarization in Mobile Environment [J]. IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews), 2011, 42(3): 397-407.

[10] BANG T S, SORNLERLAMVANICH V. Sentiment Classification for Hotel Booking Review Based on Sentence Dependency Structure and Sub-Opinion Analysis [J]. IEICE TRANSACTIONS on Information and Systems, 2018, 101(4): 909-916.

[11] HE R, LEE W S, NG H T, et al. An Unsupervised Neural Attention Model for Aspect Extraction [C]//Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers). 2017: 388-397.

[12] PHAM D H, LE A C. Learning Multiple Layers of Knowledge Representation for Aspect Based Sentiment Analysis [J]. Data & Knowledge Engineering, 2018, 114: 26-39.

[13] ZHAO W, GUAN Z, CHEN L, et al. Weakly-Supervised Deep Embedding for Product Review Sentiment Analysis [J]. IEEE Transactions on Knowledge and Data Engineering, 2017, 30(1): 185-197.

[14] XU H, LIU B, SHU L, et al. Double Embeddings and CNN-Based Sequence Labeling for Aspect Extraction [C]//Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics. 2018: 592-598.

[15] ZHAO Y, QIN B, LIU T. Encoding Syntactic Representations with a Neural Network for Sentiment Collocation Extraction [J]. Science China Information Sciences, 2017, 60(11): 110101.

[16] CONG D, YUAN J, ZHAO Y, et al. A Joint Model for Sentiment Classification and Opinion Words Extraction [M]//Chinese Computational Linguistics and Natural Language Processing Based on Naturally Annotated Big Data. Springer, Cham, 2018: 337-347.

[17] WANG Y, HUANG M, ZHU X, et al. Attention-based LSTM for Aspect-Level Sentiment Classification [C]//Proceedings of the 2016 conference on empirical methods in natural language processing. 2016: 606-615.

[18] MA R, WANG K, QIU T, et al. Feature-Based Compositing Memory Networks for Aspect-Based Sentiment Classification in Social Internet of Things [J]. Future Generation Computer Systems, 2019, 92: 879-888.

[19] DONG Z, DONG Q, HAO C. HowNet and Its Computation of Meaning [C]//Proceedings of the 23rd International Conference on Computational Linguistics: Demonstrations. Association for Computational Linguistics, 2010: 53-56.

[20] KU L W, CHEN H H. Mining Opinions from the Web: Beyond Relevance Retrieval [J]. Journal of the American Society for Information Science and Technology, 2007, 58(12): 1838-1850.