

结合金融领域情感词典和注意力机制的细粒度情感分析

祝清麟^{1,2}, 梁斌¹, 刘宇瀚¹, 陈奕¹, 徐睿峰^{1,2,†}, 毛瑞彬³

1. 哈尔滨工业大学(深圳) 计算机科学与技术学院, 广东 深圳 518055
 2. 哈工大理光联合实验室 广东 深圳 518055
 3. 深圳证券信息有限公司 广东 深圳 518028
- †. 通讯作者, Email: xuruifeng@hit.edu.cn

摘要

针对在金融领域实体级情感分析任务中, 往往缺乏足够的标注语料, 以及通用的情感分析模型难以有效处理金融文本等问题。本文构建一个百万级别的金融领域实体情感分析语料库, 并标注五千余个金融领域情感词作为金融领域情感词典。同时, 基于该金融领域数据集, 提出一种结合金融领域情感词典和注意力机制的金融文本细粒度情感分析模型。该模型使用两个LSTM网络分别提取词级别的语义信息和基于情感词典分类后的词类级别信息, 能有效获取金融领域词语的特征信息。此外, 为了让文本中金融领域情感词获得更多关注, 提出一种基于金融领域情感词典的注意力机制来为不同实体获取重要的情感信息。最终在构建的金融领域实体级语料库上进行实验, 取得了比对比模型更好的效果。

关键词: 细粒度情感分析; 金融文本; 情感词典

Attention-based Recurrent Network Combined with Financial Lexicon for Aspect-level Sentiment Classification

Qinglin Zhu^{1,2}, Bin Liang¹, Yuhan Liu¹, Yi Chen¹, Ruifeng Xu^{1,2,†}, Ruibin Mao³

1. Department of Computer Science, Harbin Institute of Technology (Shenzhen), Shenzhen, Guangdong 518055, China;
 2. HIT-RICON Joint Lab, Shenzhen, Guangdong 518055, China;
 3. Shenzhen Securities Information Co. Ltd., Shenzhen, Guangdong 518028 China;
- †. Corresponding author, Email: xuruifeng@hit.edu.cn

Abstract

To address the lack of sufficient annotated corpus and the poor performance of common sentiment analysis models for the task of entity-level sentiment analysis of financial texts. This paper builds a multi-million level corpus of sentiment analysis of financial domain entities and labels more than five thousand financial domain sentiment words as financial domain sentiment dictionary. Based on this financial domain dataset, we propose an Attention-based Recurrent Network Combined with Financial Lexicon, called FinLexNet. FinLexNet model uses LSTM to extract category-level information based on financial domain sentiment dictionary and another LSTM to extract semantic information at the word-level, which can effectively obtain information about the characteristics of financial domain words. In addition, in order to get more attention to the financial sentiment words, an attention mechanism based on the financial domain sentiment dictionary is proposed. Finally, experiments were conducted on the dataset we constructed, which shows that our model has achieved better performance than the comparative models.

Keywords: Fine-grained sentiment analysis, Financial texts, Financial Sentiment Lexicon

1 引言

随着互联网和金融行业的快速发展,在金融领域目前不断出现大量专业的股评报告、研究报告等信息,以及个体投资者的个人看法和分析。无论新闻报道还是针对相关主题与公司的评论信息,往往都包含有对相关事件与公司的评价与态度,具有丰富的投资和监管参考价值。对这些评价信息的全面把握,有助于投资者更好的了解市场,辅助投资决策。同时,对于金融市场监管者,有助于及早从评价中发现潜藏的问题,对于掌握市场动态,消除市场风险也有着重要意义。为此,金融文本的情感分析研究正在成为当前研究和应用热点。

金融领域的文本实体级细粒度情感分析研究尚处于初级阶段,也是细粒度情感分析重要的子任务(Pang and Lee, 2008),面临着诸多挑战。首先缺乏高质量、大规模的金融领域文本情感标注语料,导致文本处理底层技术缺乏数据支撑。现有的通用文本情感分析模型缺乏对金融领域文本特点的分析利用,没有考虑金融领域词性特征,缺乏对金融领域情感先验知识的利用,因此在金融文本上表现欠佳。

针对现有金融领域语料库匮乏的问题,本文构建了金融领域细粒度情感分析语料库。首先对各大金融新闻网站进行数据爬取与清洗,之后按照字级别对所爬取的数据进行实体标注和实体情感标注。总共标注了5206篇新闻稿,整理出3325个实体和对应的9240条情感语句,并构建了包含5047个词的金融领域情感词典。

针对现有方法对金融领域知识利用的不足,本文设计并实现了结合金融领域情感词典和注意力机制的细粒度情感分析模型(Attention-based Recurrent Network Combined with Financial Lexicon, FinLexNet),该模型使用一个LSTM提取词级别的文本信息,并基于金融情感词典将文本中的词分成“积极”,“消极”,“中立”,“金融实体”,“其他”五个类别对文章进行表示,使用另一个LSTM提取金融领域词性特征,这样不仅让模型关注到不同类型词语的特殊性,从而更好的理解上下文的语义信息,还能作为对词级别较为细粒度的信息的补充,获取更宏观的文本信息。模型还使用了金融领域情感词典指导注意力机制,使得注意力机制更加关注金融领域情感词,在构建的数据集上达到了同类模型的最佳效果。

本文构建了一个百万级的金融领域实体级细粒度情感分析语料库,并在此基础上提出了一种结合金融情感词典和注意力机制的情感分析模型。不仅促进围绕金融领域文本的情感分析研究的深入,具有很好的科学意义,同时,可以很好服务于面向金融领域的舆情分析、市场判断和监管协调,具有突出的应用价值。

2 相关研究

细粒度情感分析是情感分析的一个热门且具有重要应用价值的领域(赵妍妍 et al., 2010),侧重于对细粒度情感信息的挖掘。对于金融领域,实体级的细粒度情感分析是分析出金融文本中出现的金融实体的情感,常用的方法有基于情感词典,机器学习和深度学习的细粒度情感分析方法。

情感词典是识别文本情感的有效工具,有不少学者研究构建情感词典的方法(Meng et al., 2012; 梅莉莉 et al., 2016),形成了如Word Net(Fellbaum, 2012)等具有代表性的英文情感词典和董振东等人编制的中文知网情感词典How Net。基于情感词典的细粒度情感分析方法主要是利用句式词库和情感词典去分析文本语句的特殊结构及情感倾向词,如Wu等(2006)根据情感词情感强度的不同而赋予不同的情感权重,然后进行加权求和。Lipenkova等(2015)提出了预建立的词典和通用语言规则相结合的方式,其在中文方面级情感分析任务上取得了较好的效果。

基于机器学习进行细粒度情感分析也是主流的方法之一,在早期的研究中,细粒度情感分析被当作一般情感分类任务,使用情感词典、文本语义等特征等提取文本特征来建立细粒度情感分类模型。Kiritchenko等(2014)引入了产品的总体评分和情感词库两个外部知识,并和SVM分类器相结合一起,在SemEval2014年竞赛中取得了最佳性能。Ramesh等(2015)提出使用马尔科夫随机场解决在线课程MOOC中的方面级情感分类问题。郝志峰等(2015)提出把情感对象识别看作一个序列标记问题,通过在传统的CRF序列标记模型上增加情感对象的全局节点,有效地结合上下文信息、句法依赖以及情感词典,从而可以识别出微博中的情感对象。然而传统的机器学习方法通常需要依赖大量的人工筛选特征,这需要耗费大量的时间和精力。

随着深度学习技术的发展,研究人员设计了一系列的神经网络自动生成对象和内容的低维度表示方法,并且在细粒度情感分类任务中得到了较好结果。Tang等(2016a)提出一种基于目标的长短期记忆网络(TD-LSTM),依据目标词的位置将输入的文本切分成左右两个部分并分别送入LSTM,较传统LSTM模型性能有所提升。注意力机制(Attention Mechanism)源于对人类视觉的研究,近年来,随着注意力机制的深入研究,很多学者基于注意力机制提出了一系列的方法进行细粒度情感分析。赵冬梅等(2018)提出一种利用协同过滤算法计算得到用户的情感分布矩阵,再使用注意力机制提取文本信息,从而进行实现情感分类。曾峰等(2019)提出了一种基于注意力机制的LSTM神经网络模型,模型从词级别和句子级别两个层面进行语义提取,从而获取不同词语和句子的重要性。吴小华等(2019)使用字向量对文本进行字级别的表示,并使用双向的LSTM网络和注意力机制提取上下文之间的关系。

金融领域细粒度情感分析研究较少,Cortis等(2017)讨论了SemEval-2017会议“金融微博和新闻的情感分析”任务三十余位参赛者的方法和工具,其中最多人使用的是基于传统机器学习模型SVM和SVR的方法。Do等(2019)指出金融领域数据标注需要广泛的领域专业知识,进行专业标注会很昂贵,所以构建的数据较少。Maia等(2018)发布了一个非常小的数据集(FiQA),包含了金融领域的文本实例和文本中提到的实体,并给每个实体的情感打分。Yang等(2018)基于ELMo模型提出了ULMFiT方法分析FiQA数据集上的金融实体情感。Salunkhel等(2019)提出了一种用于方面分类的迁移学习方法和一种用于金融数据的情感预测的回归方法,迁移学习方法利用了BERT,并使用了不同的回归方法,其中线性支持向量回归法的效果最好。

细粒度情感分析的方法较多,但是在金融领域实体级细粒度的情感分析研究较少,尤其是缺乏数据集的情况下使得金融领域的研究更难以开展。我们针对语料库匮乏的问题构建了金融领域细粒度情感分析语料库。针对现有模型缺乏对金融领域知识利用的问题,构建了提出了结合金融领域情感词典的细粒度情感分析方法,用金融领域情感词指导注意力机制,并结合金融领域词性特征,取得了同类模型的最佳性能。

3 金融领域实体级细粒度情感分析语料库构建

针对金融领域情感分析语料库匮乏的问题,我们设计并构建了金融领域实体级细粒度情感分析语料库。考虑到新闻文本信息丰富,更新速度快且较为正规,我们爬取了各大金融数据网站(21世纪经济报道⁰、财新网¹、每经网-公司板块²、生意社³、人民网⁴)作为数据来源,采用Scrapy框架共计爬取22681篇新闻文本,并对文章进行了删除特殊符号,利用正则匹配剔除一些无关信息等预处理。

首先我们进行了金融实体的标注。对于金融实体,我们标识出文本中的公司名,人名和品牌名称。实体名基于长匹配的原则进行标注,并通过天眼查辅助确定公司名、品牌名称等。

例如:“乐融致新和乐视网业务发展的颓势仍没有出现明显的好转。”

在这个文本中“乐融致新”和“乐视网”为我们标注的实体。

对于金融实体情感标注,我们将金融实体的情感极性标注为三大类:无情感、消极、积极,每一类指定的标注准则如下:

(1)积极情感

对于积极情感的标注,如果文本中出现了有利于公司经营的事实,以及一些人为的积极评价,则标注为积极。

例如:“伴随着近年来白酒行业复苏,水井坊业绩也水涨船高。”

(2)中立情感

对于中立情感的标注,如果文本中出现的信息为与公司经营相关,但无法判断是有利还是不利的情况标注为中立,包括一下情况:

①一些与公司经营相关的事实性的陈述,包括(但不限于):公司人事变更、子公司或者下属经营企业的设立与关闭、公司财务或投资操作等等。

②既有有利事实也有不利事实(句中不存在转折词例如尽管、然而、虽然、但是等表达情感偏向的副词)。

⁰<http://news.21so.com/chanye/>

¹<http://companies.caixin.com/news/>

²<http://www.nbd.com.cn/columns/346>

³<http://news.toocle.com/list/c-3511-1.html>

⁴<http://industry.people.com.cn/GB/413887/index.html>

③一些与公司经营相关的中性人为表述与评价。

例如：“电商是未来发展的方向，所有的企业都在发力，华为也不例外，但目前来看，这一动作的成效需要检验。”

(3)消极情感

对于消极情感的标注，若文本的信息不利于公司经营，标注为消极。包括一些不利于公司经营的事实，以及一些人为的消极评价。

例如：“由于游戏收入下滑，热门游戏进入周期末尾，近期市场对腾讯的评估本来就不太乐观。”

为了构建金融领域细粒度情感分析数据集，从爬取的22681篇新闻文本中选取了5206篇进行标注。首先由4名标注人员进行预标注2000条，在标注过程中分别对各自的标注结果进行比对收集差异与有歧义的地方，制定对各类实体以及针对模糊和有冲突的语境制定相应的标注准则。在标注过程中，每一段新闻文本由至少两名标注者独立标注，即标注过程中标注者之间彼此没有交流，完全依赖先前制定好的标注准则。独立标注完成后，对于有差异或有错误的标注结果，一名额外的标注者会参与讨论，直到所有的标注者意见统一后，对以标注数据进行人为修改，最终完成标注。

最终整理出3325个金融实体，每个金融实体对应一个或多个语句，共计有9240条对应的情感语句，共108.7万字。在9240个情感语句中，金融实体情感是积极的有4189条，中性的有3202条，消极的有1627条。具体的统计结果如表1所示。

极性	积极	中性	消极
数量	4179	3202	1627
百分比	46.39%	35.55%	18.06%

表 1. 金融实体情感数据统计

通过分析金融文本数据，根据经验判断出了哪些词汇会影响对与实体情感极性的判断，从而构建了一个金融领域情感词典，其中包含了2079个积极词，1070个中立词和1898个消极词。金融领域情感词典的具体统计信息如表2所示。

极性	积极	中性	消极
数量	2079	1070	1898
百分比	41.19%	21.20%	37.61%

表 2. 金融领域情感词统计

为了计算带标注的语料库与标注者之间的一致性，计算了Cohen’s Kapp(Cohen, 1960)值与Fleiss’ Kappa(Fleiss, 1971)值。Fleiss’ Kappa值为0.6686，实验标注结果数据具有较好一致性。Cohen’s Kapp值达到0.7210，这说明标注者可以在给定文本的情况下可靠地识别目标实体的情感。

4 结合金融领域情感词典和注意力的细粒度情感分析模型

本文提出的一种结合金融领域情感词典和注意力的情感分析模型框架图如图1所示。为提取细粒度的语义信息，使用LSTM提取词级别的语义信息（见模型右半部分）；为了让模型关注到不同类型词语的特殊性，并获取更宏观的文本信息作为对词级别信息的补充，使用另一个LSTM提取词类级别的语义信息（见模型左半部分）。其中词类级别的表示是指将文章词分成5个类别：Pos, Neg, Neu, Entity, Other，即“积极”，“消极”，“中立”，“金融实体”，“其他”五个类别。然后使用Word2Vec模型对文本进行训练，从而获取每一类词语词向量的平均值来表示该类词向量。为了更关注到与预测情感极性相关度高的词语，使用金融实体与金融文本进行词级别的注意力。为了让模型更加关注金融领域情感词，模型还使用了金融领域情感词典去指导注意力机制，从而使模型更加关注金融情感词所在的位置，提升情感分析的准确度。

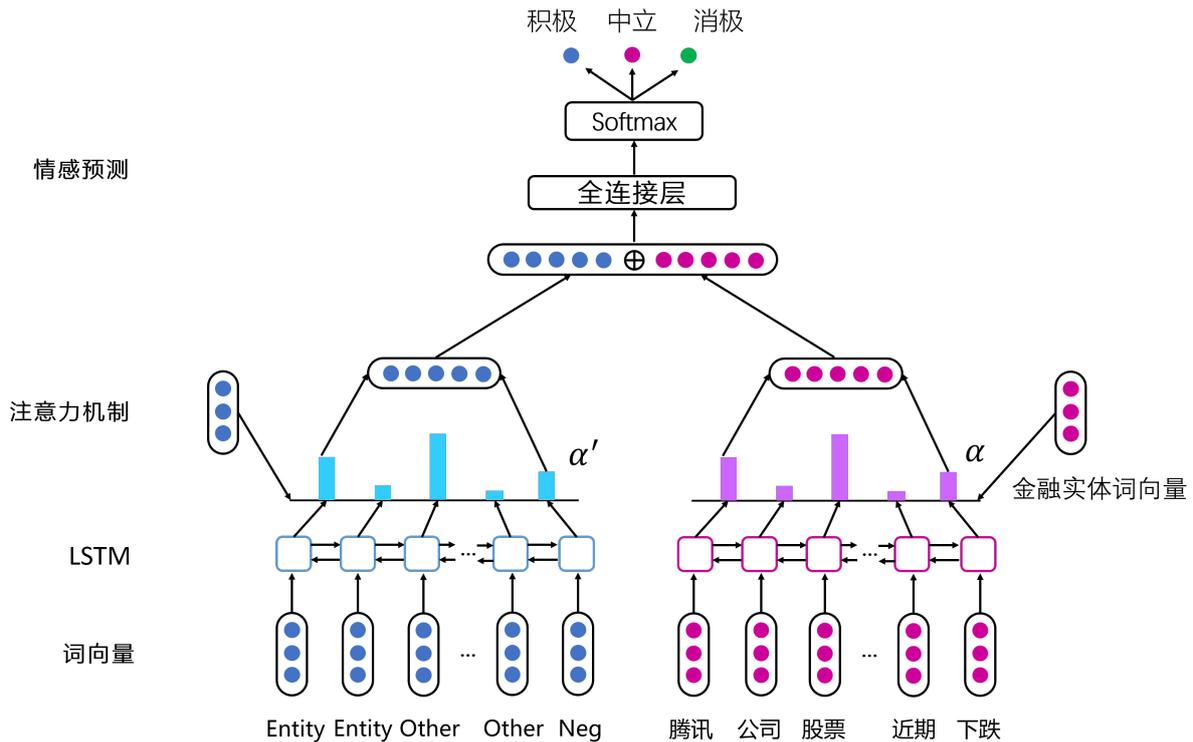


图 1. 结合金融领域情感词典和注意力的情感分析模型

4.1 基于LSTM的词级别编码器

为了提取文本的语义，使得其特征表示更加符合当前语境信息，我们采用了长短记忆网络 (Long-Short Term Memory, LSTM) 提取词级别的语义特征，将整条金融文本先经过分词和词嵌入之后输入LSTM，如公式 (1) 所示。

$$\vec{h}_t = \overrightarrow{LSTM}_w(x_{it}), t \in [1, k] \quad (1)$$

其中，一共有*i*个金融文本，每个金融文本*d*中包含*k*个词语： $\{w_{i1}, w_{i2}, \dots, w_{ik}\}$ 。 x_{it} 为*t*时刻。通过LSTM就可以获得每个词语的向量表示为 $[h_1^i, h_2^i, \dots, h_c^i]$ 。

4.2 基于LSTM的词类级别编码器

上一节中提取的文本的是词级别较为细粒度的信息，但不能注意到不同类别的词汇。在本节中使用词类级别的表示方法，本文基于领域情感词典和已经标注好的金融实体，将文章词分成5个类别：Pos, Neg, Neu, Entity, Other，即“积极”，“消极”，“中立”，“金融实体”，“其他”五个类别。其中“积极”，“消极”，“中立”来自标注的情感词典，“金融实体”为标注的金融实体，“其他”为其他词汇或是未登录词。然后使用Word2Vec模型对文本进行训练，从而获取每一类词语词向量的平均值来表示该类词向量。通过词类级别的表示不仅可以给模型提供不同词类级别的信息，让模型关注到不同类型词语的特殊性，从而更好的理解上下文的语义信息，而且还能获取更宏观的文本信息，作为对词级别较为细粒度的信息的补充。

为了具体说明如何进行词类级别的表示，在此举一个简单的例子：经过分词后“腾讯 公司 股票 近期 下跌”这句话中有5个词语，其中“腾讯”和“公司”两个词语为金融实体，“股票”和“近期”属于其他词汇，“下跌”为消极类金融领域情感词。则这句话的词类级别表示为 $[Entity \ Entity \ Other \ Other \ Neg]$ ，每一类对应着相同的词向量。

LSTM的隐状态输出序列 $[h_1, h_2, \dots, h_t]$ 可以作为当前文本的特征表示，其中的 h_t 对应于文本序列中第*t*个词的特征。

为了提取词类级别的语义特征，本文使用另一个LSTM网络作为文本的特征提取器，将之前介绍的经过词嵌入表示的词类级别文本输入 $LSTM'$ ，如公式(2)。

$$\vec{h}_t^i = \overline{LSTM}_w'(x_{it}), t \in [1, k] \quad (2)$$

4.3 词级别的注意力机制

通过两个LSTM建模得到的文本表示，会给每一个词分配相同的权重，因而无法准确把握语义的重点。注意力机制的思想是不同的情境下不用文本的重要程度不同，在计算过程中将文本的语义根据分配的权重进行加权求和，获得与任务更相关的文本的表示。为了进一步提升情感分析的准确度，借助注意力机制，建模实体情感与各个词语之间的关系，为子句的词序列语义特征分配不同的权重，使得更重要的词语得到更多的关注。由公式(3)与公式(4)为词语的注意力权重计算方式：

$$\alpha_{it} = \frac{\exp(\gamma(h_c^i, e^E))}{\sum_{j'} \exp(\gamma(h_c^j, e^E))} \quad (3)$$

$$\gamma(h_c^j, e^E) = \tanh(h_c^i \cdot w_m^T \cdot e^{ET} + b_a) \quad (4)$$

其中经过LSTM后每个词语的向量表示为 $[h_1^i, h_2^i, \dots, h_c^i]$, w_m 是权值矩阵, b_a 为偏移量, e^E 为金融实体的词向量 (如果金融实体被切分为多个词则 e 为这些词向量取均值后的结果); α_{it} 为词 w_{it} 相对于金融实体 e^E 的注意力权重。

LSTM提取的带有注意力加权的子句文本特征表示如公式(5)所示：

$$o_i = \sum_t \alpha_{it} h_{it} \quad (5)$$

将词类级别的送入LSTM'后，同样使用注意力机制确定与实体相关的上下文语义信息。相似的，得到注意力得分为 α'_{it} ，得到的特征向量为 o'_i 。

4.4 基于金融领域情感词典的注意力指导

注意力机制能够更好的关注到重要的词汇从而提高模型识别的准确率，但不一定能够准确的识别哪些词语是对结果有较大影响的金融领域情感词。为了解决这个问题，我们使用构建的金融领域情感词典去指导注意力机制，使得金融情感词的获得更大的关注。

为了使用构建的金融领域情感词典，对于一个输入的分词后的句子，构建了一个与分词后的句子长度相同的情感词向量，称为 Vec_{Lex} ，并初始化为0。遍历输入金融文本中的词语若其出现在金融领域情感词典中，则在情感词向量中将对应位置设为1。为了更方便的理解金融领域



图 2. 金融领域情感词向量示意图

情感词向量的概念，举一个简单的例子例如图 2 所示，假设输入的金融文本为“腾讯 公司 股票 近期 下跌”，首先初始化一个情感词向量 $[0, 0, 0, 0, 0]$ ，遍历输入的句子发现“下跌”这个词出现在金融领域情感词典中，属于消极词，便把“下跌”这个词在情感词向量对应的位置设置为1，则该句话的情感词向量为 $[0, 0, 0, 0, 1]$ 。

为了使得注意力机制更加的关注金融领域情感词，我们修改了损失函数，在交叉熵损失后又加入了一项 $\lambda(\alpha - Vec_{Lex})^2$ 。其中 λ 是确定情感词典损失重要性的超参数, α 为注意力机制的得分, Vec_{Lex} 为情感词典向量。从而使得注意力机制得分 α 去拟合金融情感词向量，从而使模型更加关注输入金融文本金融情感词。

之后将词类级别的注意力表示和词级别的注意力表示相结合，将两个带有注意力加权句子文本特征表示向量拼接起来，最后经由softmax层得到模型的概率输出，如公式(6)所示：

$$p'_i = \text{softmax}(o_i \oplus o'_i) \quad (6)$$

其中 \oplus 为向量拼接操作， o_i 经过LSTM的注意力机制的词级别表示， o'_i 是经过LSTM'的注意力机制的词类级别表示。

模型的最终的损失函数为公式(7):

$$L = - \sum_{i \in D} y_i \log p_i + \lambda (\alpha_{norm} - Vec_{Lex}) \quad (7)$$

其中， D 为样本集合， y_i 为子句真实标签， p_i 为模型的预测结果， λ 是确定情感词典损失重要性的超参数， α_{norm} 为LSTM词级别注意力得分 α 和经过LSTM'的词类级别注意力得分 α' 的平均值。

5 实验

5.1 数据集

实验数据集采用构建的金融领域实体级细粒度情感分析语料库，将数据集分成测试集，验证集与训练集，具体的划分如表3所示。

划分	积极	中性	消极	合计
训练集	2771	2157	1072	6000(64.94%)
验证集	564	461	215	1240(13.42%)
测试集	907	725	368	2000(21.64%)

表 3. 金融领域情感文本训练集测试集数据统计

5.2 评价指标

本文使用准确率(Accuracy)和Macro-F1值作为评价标准。

5.3 实验设计

(1)词向量

实验的词向量采用腾讯AI Lab公开的中文词向量数据(Song et al., 2018)，该数据集涵盖面广，囊括了800余万个中文词语，数据集的维度为200维。该词向量的训练使用了腾讯研制的Directional Skip-Gram (DSG)算法，相比于广泛采用的词向量训练算法Skip-Gram (SG)，DSG算法额外考虑了词对的位置信息，以从而能更准确的表示词汇的语义。具有词语覆盖率全，新鲜度高，词向量准确率高的特点。在训练的过程中词向量不冻结，参数随训练一起更新。

(2)超参设置

数优化采用 Adam(Kingma and Ba, 2014)优化算法，学习率设置为 0.0001。对词向量矩阵以及不同LSTM层之间的连接采用Dropout(Srivastava et al.,)，对LSTM层内部与隐状态相关的权重矩阵采用 DropConnect(Cho et al., 2014)。Batch-size设置为128，Dropout为0.2，DropConnect为0.1，LSTM的隐藏层为200维，LSTM Attention的输出为200维，LSTM' Attention的输出为50维，情感词典损失重要性的超参数 $\lambda = 0.035$ 。

(3)对比模型介绍

对比模型包括基础的Bi-LSTM模型和近些年在方面级的情感分析 (Aspect Based Sentiment Analysis) 领域的深度学习模型进行对比，参与对比的模型有以下几种:

- Bi-LSTM: Bi-LSTM是Bi-directional Long Short-Term Memory的缩写，是由前向LSTM与后向LSTM组合而成，使用Bi-LSTM模型提取文本的语义信息之后，直接送入softmax层进行分类。
- TD-LSTM(Tang et al., 2016a): 基于目标的长短期记忆网络，根据特定目标单词的所在位置，将训练语句拆分左右两部分，通过LSTM获取左右部分两个隐层的输出，输入分类器，获取分类结果。

- IAN(Ma et al., 2017): 该模型改进了传统的分类模型中将两者分开独立建模或者只针对内容建模的方法, 该模型先让内容和目标分别通过不同的LSTM后, 利用注意力机制实现两者的信息交互, 从而提升模型的准确度。
- AOA(Huang et al., 2018): 该模型建模了目标和文本的交互关系, 分别将文本和目标经过双向的LSTM, 并使用隐藏层的输出接着计算两者的交互矩阵, 将该矩阵得到的信息送入softmax实现对情感的分类。
- MemNet(Tang et al., 2016b): 该模型利用了注意力机制的QA系统中的深度记忆网络, 将方面词的上下文信息作为存储器中存储的内容, 实现了一个针对方面级的情感分析模型。
- ATAE-LSTM(Wang et al., 2016): 该模型利用了注意力机制来获取上文下信息与目标词信息之间的关系, 结合了LSTM神经网络与注意力机制提取句子语义, 从而提升情感分类的准确度。

5.4 实验结果与分析

(1) 总体性能

所有实验均采用 NVIDIA GeForce GTX 2080Ti 显卡进行计算加速, 并在单张显卡下完成。在自标注的数据集进行了实验, 总体性能的实验结果如表4。

模型	Precision	Macro-F1
Bi-LSTM	0.7040	0.6860
TD-LSTM	0.7132	0.6847
IAN	0.7115	0.6853
AOA	0.7285	0.6971
MemNet	0.7225	0.6954
ATAE-LSTM	0.7165	0.6913
FinLexNet(Ours)	0.7425	0.7147

表 4. 总体实验性能结果图

从实验结果可以看出, 我们提出的模型FinLexNet取得了74.25%的准确度和0.7147的Macro-F1值, 均达到了对比模型的最佳效果。基础模型Bi-LSTM的效果最不理想, 是因为只能获取总体的文本信息, 并不能对实体进行建模。TD-LSTM提取实体前后语句语义的综合, 性能有所提升。IAN和AOA实现了实体与模型之间的交互, 更好的理解到了实体在文中的语义信息, 同Bi-LSTM相比也有不小提升。ATAE-LSTM使用注意力机制对实体和文本进行建模, 但是我们认为注意力机制没有准确把握关键词导致性能没有明显提升。我们认为提出的FinLexNet模型性能较好的原因是结合了金融领域词性信息并用金融情感词指导注意力机制, 能让模型获得的信息更加丰富并使得注意力更好的把握关键词。

(2) 消融实验

为了考察模型框架中各组件的贡献程度, 本文设置了模型中不同结构的消融实验。

- LSTM-ATT: 使用LSTM去提取文本信息, 并使用注意力机制。
- LSTM-ATT-Lex: 使用LSTM提取文本信息, 并使用标注的情感词典指导注意力机制。
- Double-LSTM-ATT: 使用两个LSTM分别提取文本和词类表示的文本信息, 并使用注意力机制后输入到softmax层, 不使用标注的情感词典。
- Double-LSTM-ATT-Lex: 使用两个LSTM分别提取文本和词类表示的文本信息并结合注意力机制, 并使用标注的情感词典指导注意力机制。

实验结果如表5所示, 从实验结果来看, 使用金融领域情感词典指导注意力机制对实验结果具有较大的提升, 说明金融领域情感词典中的词对判断实体的情感极性有较大的帮助, 而通过

模型	Precision	Macro-F1
LSTM-ATT	0.7255	0.6914
LSTM-ATT-Lex	0.7295	0.6903
Double-LSTM-ATT	0.7303	0.6964
Double-LSTM-ATT-Lex	0.7425	0.7147

表 5. 消融实验结果

修改损失函数可以有效的指导注意力机制着重关注金融领域情感词，从而达到提升实验效果的目的。

(3)注意力可视化

为了探究模型注意力机制关注的内容，对三个输入样例的注意力权重进行了可视化，颜色表示一个词在给定句子中的重要性，颜色越深越重要。如图3所示。

派思股份与自贡华燃之间的交易便是根据辽宁众华出具的上述评估报告确认的评估结果为依据由双方协商确定《每日经济新闻》记者注意到对三家估值溢价率极高的标的未来经营状况和收益状况的预测无疑成为辽宁众华重要的评估依据

对于业绩下滑智慧松德给出的解释是承担公司主要智能装备业务的全资子公司深圳大宇精雕科技有限公司(以下简称大宇精雕)目前正在对产品结构进行调整导致产品验收周期延长未能及时确认收入和利润

问及信达生物此前主动撤回信迪单抗注射液上市申请是否影响其港股上市多位投资人士一致表示“没有影响”原因在于生物医药板块投资情绪强劲加上信达生物基本面向好

图 3. 注意力机制可视化图

如在第一段话中，金融实体为“派思股份”，情感极性为积极。从注意力可视化看出“溢价极率高”的颜色最深，对照了“派思股份”积极的情感极性。

在第二个金融文本中，金融实体为“智慧松德”，在文本中该公司对业绩下滑进行解释，情感极性为消极。从注意力可视化可以看出“业绩下滑”最能体现出情感极性，颜色最深。而连词“导致”往往用于不好的结果，也被模型准确的识别出来。

第三个金融文本中，金融实体为“信达生物”，文本说了撤回上市申请对公司的业务没有影响，所以情感极性为中立。在可视化结果中着重强调了“主动撤回”和“没有影响”，较为准确的找到了判断情感极性的关键词。

通过以上可视化的结果可以说明注意力机制较好的注意到了关键词和和金融领域情感词典，有助于模型判断金融实体的情感极性。

(4)错误分析

为了更好的改进模型，选取了一些错误案例进行分析。为了更方便的进行分析，将金融文本中的实体进行了加粗表示，消极的语句加上了下划线，积极的语句用波浪线标识。

例如：“市场认为，从财务数据来看，**宣亚的收购是划算的。**宣亚2016年的营业收入为4.67亿元，净利润为5871.01万元。8月15日，**宣亚国际发布2017年中报**，报告期内，公司实现营业收入2.10亿元，同比下降6.74；净利润为2722.00万元，同比增长4.22。而**映客直播**的同期营收达到了43.47亿元，归母净利润更是高达4.8亿元，远远高于上市公司。”

这个例子中，“映客直播”的情感极性为正面，而模型判断为负面。分析原因是模型可能没

有找准映客直播对应的语句，同“宣亚国际”的营收下降产生了混淆。

例如：“中信银行向佳兆业伸出援手始于佳兆业陷入债务危机之时。彼时中信银行深圳分行对危机中的佳兆业施以援手，提供大约300亿元资金助其解困，100亿元用于置换佳兆业位于上海、杭州等地的8个优质资产项目债务；另有100亿元将作为佳兆业的后续开发贷款。此后平安银行也与佳兆业达成全方位的战略合作，签约金额为500亿元，用于支持佳兆业的未来发展。”

该例子中，“佳兆业”的情感极性为正面，而模型判断为中立。分析原因是模型注意到了消极观点“陷入债务危机”，同时也注意到中信银行伸出援手，提供了300亿资金纾困，“支持佳兆业的未来发展”。从而认为是中立。而如果进一步的推理可以知道，这是一个利好的消息，所以情感极性为积极。这说明虽然模型学习到了很多情感词，但是由于缺乏对这种褒贬都存在的情况的进一步推理，导致分类错误。

6 总结与展望

本文构建了一个金融领域实体级细粒度情感分析语料库，并提出了一种结合金融领域情感词典和注意力的细粒度情感分析模型。为了利用金融领域词性信息并结合粗细粒度的文章信息，本文使用两个LSTM网络分别提取词类级别和词语级别的语义。为了让模型有针对性地关注对情感结果影响较大的词语，本文使用金融领域情感词典对注意力机制进行修正。最后，在本文标注的金融领域细粒度情感分析语料库上进行实验，实验结果表明本文提出的结合金融领域情感词典和注意力模型能有效提升细粒度情感分析的准确性。未来的工作可以针对文本中有正负两面评价金融实体的情感进行研究，并考虑如何充分利用文本中的金融数字信息。

致谢

国家自然科学基金 61876053, 61632011, 深圳市技术攻关项目JSGG20170817140856618, 深圳市基础研究学科布局项目JCYJ20180507183527919, JCYJ20180507183608379, 广东省新冠肺炎疫情防控科研专项项目 2020KZDZX1224

参考文献

- Kyunghyun Cho, Bart Van Merriënboer, Caglar Gulcehre, Dzmitry Bahdanau, Fethi Bougares, Holger Schwenk, and Yoshua Bengio. 2014. Learning phrase representations using RNN encoder-decoder for statistical machine translation. *arXiv preprint arXiv:1406.1078*.
- Jacob Cohen. 1960. A coefficient of agreement for nominal scales. *Educational and psychological measurement*, 20(1):37–46. Publisher: Sage Publications Sage CA: Thousand Oaks, CA.
- Keith Cortis, André Freitas, Tobias Daudert, Manuela Huerlimann, Manel Zarrouk, Siegfried Handschuh, and Brian Davis. 2017. SemEval-2017 Task 5: Fine-Grained Sentiment Analysis on Financial Microblogs and News. In *Proceedings of the 11th International Workshop on Semantic Evaluation (SemEval-2017)*, pages 519–535, Vancouver, Canada, August. Association for Computational Linguistics.
- Hai Ha Do, PWC Prasad, Angelika Maag, and Abeer Alsadoon. 2019. Deep Learning for Aspect-Based Sentiment Analysis: A Comparative Review. *Expert Systems with Applications*, 118:272–299, March.
- Christiane Fellbaum. 2012. WordNet. *The encyclopedia of applied linguistics*. Publisher: Wiley Online Library.
- Joseph L. Fleiss. 1971. Measuring nominal scale agreement among many raters. *Psychological bulletin*, 76(5):378. Publisher: American Psychological Association.
- Binxuan Huang, Yanglan Ou, and Kathleen M. Carley. 2018. Aspect level sentiment classification with attention-over-attention neural networks. In *International Conference on Social Computing, Behavioral-Cultural Modeling and Prediction and Behavior Representation in Modeling and Simulation*, pages 197–206. Springer.
- Diederik P. Kingma and Jimmy Ba. 2014. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*.

- Svetlana Kiritchenko, Xiaodan Zhu, Colin Cherry, and Saif Mohammad. 2014. NRC-Canada-2014: Detecting Aspects and Sentiment in Customer Reviews. In *Proceedings of the 8th International Workshop on Semantic Evaluation (SemEval 2014)*, pages 437–442, Dublin, Ireland, August. Association for Computational Linguistics.
- Janna Lipenkova. 2015. A system for fine-grained aspect-based sentiment analysis of Chinese. In *Proceedings of ACL-IJCNLP 2015 System Demonstrations*, pages 55–60, Beijing, China, July. Association for Computational Linguistics and The Asian Federation of Natural Language Processing.
- Dehong Ma, Sujian Li, Xiaodong Zhang, and Houfeng Wang. 2017. Interactive attention networks for aspect-level sentiment classification. *arXiv preprint arXiv:1709.00893*.
- Macedo Maia, Siegfried Handschuh, André Freitas, Brian Davis, Ross McDermott, Manel Zarrouk, and Alexandra Balahur. 2018. WWW'18 Open Challenge: Financial Opinion Mining and Question Answering. In *Companion of the The Web Conference 2018 on The Web Conference 2018 - WWW '18*, pages 1941–1942, Lyon, France. ACM Press.
- Xinfan Meng, Furu Wei, Ge Xu, Longkai Zhang, Xiaohua Liu, Ming Zhou, and Houfeng Wang. 2012. Lost in Translations? Building Sentiment Lexicons using Context Based Machine Translation. In *Proceedings of COLING 2012: Posters*, pages 829–838, Mumbai, India, December. The COLING 2012 Organizing Committee.
- Bo Pang and Lillian Lee. 2008. Opinion Mining and Sentiment Analysis. *Foundations and Trends® in Information Retrieval*, 2(1–2):1–135, July. Publisher: Now Publishers, Inc.
- Arti Ramesh, Shachi H. Kumar, James Foulds, and Lise Getoor. 2015. Weakly Supervised Models of Aspect-Sentiment for Online Course Discussion Forums. In *Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*, pages 74–83, Beijing, China, July. Association for Computational Linguistics.
- Ashish Salunkhe and Shubham Mhaske. 2019. Aspect Based Sentiment Analysis on Financial Data using Transferred Learning Approach using Pre-Trained BERT and Regressor Model. 06(12):5.
- Yan Song, Shuming Shi, Jing Li, and Haisong Zhang. 2018. Directional Skip-Gram: Explicitly Distinguishing Left and Right Context for Word Embeddings. In *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 2 (Short Papers)*, pages 175–180, New Orleans, Louisiana, June. Association for Computational Linguistics.
- Nitish Srivastava, Geoffrey Hinton, Alex Krizhevsky, Ilya Sutskever, and Ruslan Salakhutdinov. Dropout: A Simple Way to Prevent Neural Networks from Overfitting. page 30.
- Duyu Tang, Bing Qin, Xiaocheng Feng, and Ting Liu. 2016a. Effective LSTMs for Target-Dependent Sentiment Classification. In *Proceedings of COLING 2016, the 26th International Conference on Computational Linguistics: Technical Papers*, pages 3298–3307, Osaka, Japan, December. The COLING 2016 Organizing Committee.
- Duyu Tang, Bing Qin, and Ting Liu. 2016b. Aspect Level Sentiment Classification with Deep Memory Network. In *Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing*, pages 214–224, Austin, Texas, November. Association for Computational Linguistics.
- Yequan Wang, Minlie Huang, Xiaoyan Zhu, and Li Zhao. 2016. Attention-based LSTM for Aspect-level Sentiment Classification. In *Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing*, pages 606–615, Austin, Texas, November. Association for Computational Linguistics.
- Chung-Hsien Wu, Ze-Jing Chuang, and Yu-Chung Lin. 2006. Emotion recognition from text using semantic labels and separable mixture models. *ACM Transactions on Asian Language Information Processing*, 5(2):165–183, June.
- Steve Yang, Jason Rosenfeld, and Jacques Makutonin. 2018. Financial Aspect-Based Sentiment Analysis using Deep Representations. *arXiv:1808.07931 [cs]*, August. arXiv: 1808.07931.
- 吴小华, 陈莉, 魏甜甜, and 范婷婷. 2019. 基于Self-Attention和Bi-LSTM的中文短文本情感分析. *中文信息学报*, 33(6):100–107.

- 曾锋, 曾碧卿, 韩旭丽, 张敏, and 商齐. 2019. 基于双层注意力循环神经网络的方面级情感分析. 中文信息学报, 33(6):108–115.
- 梅莉莉, 黄河燕, 周新宇, and 毛先领. 2016. 情感词典构建综述. 中文信息学报, 30(5):19–27.
- 赵冬梅, 李雅, 陶建华, and 顾明亮. 2018. 基于协同过滤Attention机制的情感分析模型. 中文信息学报, 32(8):128–134.
- 赵妍妍, 秦兵, and 刘挺. 2010. 文本情感分析. 软件学报, 21(8):1834–1848.
- 郝志峰, 杜慎芝, 蔡瑞初, and 温雯. 2015. 基于全局变量CRFs模型的微博情感对象识别方法. 中文信息学报, 29(4):50–58.

JCL2020