

FinBERT-RCNN-ATTACK: 金融文本情感分析模型

段魏诚, 薛 涛

(西安工程大学 计算机科学学院, 陕西 西安 710600)

摘 要: 金融文本包含投资者的情绪及公众对相关事件的态度。近年来, 自然语言处理已广泛应用于金融领域, 对金融文本数据进行情感分析可以得到丰富的投资价值 and 监管参考价值。然而由于金融词汇具有专业性和特殊性, 现有的通用情感分析模型不适合金融领域情感分析任务, 精确度有待提高, 且现有模型易受到对抗样本的干扰导致模型结果出错。为了解决这些问题, 提出了一个 FinBERT-RCNN-ATTACK 模型。利用在金融语料库预训练的 FinBERT 模型进行词嵌入处理, 提取语义特征, 将提取到的特征引入 RCNN 模型进一步挖掘上下文的关键特征, 并且在模型中引入对抗训练, 即在嵌入阶段添加扰动, 提高模型的鲁棒性和泛化性。实验结果表明, 在金融领域数据集上, 提出的模型优于其他情感分析模型, 精准度提升了 3%~35%。

关键词: 金融文本; 情感分析; FinBERT; 循环卷积神经网络; 对抗训练

中图分类号: TP391

文献标识码: A

文章编号: 1673-629X(2024)05-0157-06

doi: 10.20165/j.cnki.ISSN1673-629X.2024.0054

FinBERT-RCNN-ATTACK: Emotional Analysis Model of Financial Text

DUAN Wei-cheng, XUE Tao

(School of Computer Science, Xi'an Polytechnic University, Xi'an 710600, China)

Abstract: The financial text contains investor sentiment and public attitudes towards the events. In recent years, natural language processing has been widely used in financial field, and the emotional analysis of financial text data can get rich investment value and regulatory reference value. However, due to the professionalism and particularity of financial vocabulary, the existing general emotional analysis model is not suitable for the emotional analysis task in the financial field, and the accuracy needs to be improved, and the existing model is vulnerable to the interference of antagonistic samples, leading to the wrong model results. In order to solve these problems, we proposed a FinBERT-RCNN-ATTACK model. The FinBERT model pre-trained in the financial corpus is used for word embedding processing to extract semantic features, and the extracted features are introduced into the RCNN model to further excavate the key features of the context. In addition, adversarial training is introduced into the model, that is, disturbance is added in the embedding stage to improve the robustness and generalization of the model. The experimental results show that the proposed model is better than the other emotional analysis models, and the accuracy is improved by 3%~35%.

Key words: financial text; emotional analysis; FinBERT; recurrent convolutional neural networks; adversarial training

0 引言

自然语言处理的情感分析技术是对带有情感色彩的主观性文本进行分析, 用来处理文本可以从中提取蕴含的语义情感信息。金融领域的文本数据包含对相关事件的情绪, 对金融文本进行情感分析有助于理解投资者的态度, 并影响投资决策和市场走势^[1]。近年

来, 自然语言处理已广泛应用于金融领域。

当情感分析应用于专业语言领域(例如金融)时, 任务会更具挑战性^[2]。基于词典的情感方法以情感词典作为情感倾向判断依据, 该方法虽然能够判断情感极性, 但是对词典依赖较高。基于机器学习的情感分类方法虽然提高了准确度, 但仍需要人工标注金融文

收稿日期: 2023-07-30

修回日期: 2023-11-30

基金项目: 陕西省技术创新引导专项计划资助项目(2020CGXNG-012)

作者简介: 段魏诚(1998-), 男, 硕士研究生, 研究方向为自然语言处理; 通信作者: 薛 涛(1973-), 男, 教授, 博士, 研究生导师, 研究方向为人工智能和知识图谱。

本特征,处理成本较高。还有一些基于神经网络的方法,在金融领域情感分析任务上的应用十分广泛。基于预训练模型的方法^[3]利用大规模未标注文本进行预训练,其词嵌入包含了更多文本信息,得到了更好的效果。然而由于金融领域包含大量专业词汇,通用的情感分析模型不适合于金融领域的任务,传统的方法难以提取到文本更深层次的情感特征^[4]。

情感分析的进展也引起了人们对该领域对抗性攻击的新兴趣,研究表明现有的模型具有鲁棒性不足、易受到微小的干扰等问题。对抗性攻击可以欺骗模型^[5],需要研究更强大、更可靠的情绪分析方法来解决这些漏洞。因此,如何更准确地挖掘金融文本中的情感信息,是目前迫切需要解决的问题。

针对现有方法的不足,该文提出了一个 FinBERT-RCNN-ATTACK 模型,该模型用于金融领域文本的情感分析。模型采用 FinBERT 提取金融文本上下文联系和语义信息,情感分类任务的关键在于特征包含的信息,得到金融文本的语义特征后,引入 RCNN 模型提取关键信息进行情感分类。为了增强模型的鲁棒性和泛化性,在嵌入层加入扰动,通过对抗训练进一步提高模型情感分类的准确性。通过对比分析和消融实验,验证了该模型的有效性。

1 相关工作

1.1 基于词典的方法

情感分析是一种分类任务,识别文本中表达的是积极情绪还是消极情绪。在情感分析的早期研究中,大多数研究使用基于词典的方法。情感词典的方法是首先人工构造一个情感词典,词典的内容足够丰富,就可以获得较好的情感分析效果。Loughran 等^[6]基于词典的方法进行金融文本分析的工作。但是依靠手工提取特征的情感分析方法,在面对复杂的文本特征时,仍存在一定的局限性。并且维护词典需要很大的成本^[7],通用情感词典通常无法考虑特定领域的内涵和对先验知识的大量引用。

1.2 基于机器学习和深度学习的方法

基于机器学习的情感分析方法是使用传统的机器学习算法提取文本的特征。与使用情感词典的方法相比,机器学习不依赖于人工构建,并且通过数据库可以及时更新词典,节省了大量人力。文献[8]中通过对金融新闻进行语义标注、提取特征,然后分别使用朴素贝叶斯、支持向量机和 k 邻近算法进行分类。基于机器学习的情感分类方法与构建情感词典相比有一定的进步,但仍需要人工标注金融文本特征,人为主观因素会影响最终结果。

近年来随着深度学习的不断发展,使用深度学习

模型的金融文本情感分析任务取得不错的效果。文献[9]收集了财务数据、财务报表,对文本数据进行情感分析,以每只股票的情绪评分进行评估,使用词级极性和半监督学习进行了信用风险预测。文献[10]基于 Attention-Seq2Seq 的多任务学习模型,实现了文本情感倾向判决。文献[11]采用 Recurrent Convolutional Neural Networks (RCNN) 处理分类任务,RCNN 的优势在于可以较为均匀地利用单词的上下文信息,既能够解决 RNN 后面的单词比前面的单词影响力更大的缺点,也不像 CNN 一样需要通过窗口大小来设定对上下文的依赖长度,与传统的神经网络相比,会减少噪声。文献[12]情感分析模型使用半监督学习方法处理金融市场数据。尽管基于机器学习和深度学习的方法在情感分析中取得了很大的成功,但是由于缺乏大量标记的数据集,很难充分发挥神经网络在情感分析任务上的潜力。

1.3 基于预训练模型的方法

随着文本预训练思想的不断发展与普及,以 BERT 为代表的预训练模型发展迅速。BERT 模型在 Transformer 的基础上利用大规模未标记文本进行预训练,其词嵌入包含了更多文本信息。文献[13]针对金融文本特点提出了截断抽取法,利用阅读理解技术的微调方法扩展现有 BERT 模型。文献[14]提出了一种基于 BERT-BiGRU 多模集成学习的深层情感语义识别方法,判断社交文本的情感倾向性。文献[15]提出了一种基于 BERT 的情感分析和关键实体检测方法,该方法应用于社交媒体的在线金融文本挖掘和舆论分析。虽然此类模型在情感分析任务上效果良好,但是在金融文本分类上依然具有局限性。因为此类预训练模型通常是在一般语料库上进行预训练,其中几乎不包含金融文本,由于金融语料库和一般领域语料库之间的词汇表达存在较大差异,因此处理金融领域的特定词汇时,无法保证高效的实验结果^[16]。

文献[17]提出了 FinBERT 模型,它是在大规模金融语料库上预训练的特定领域语言模型,用来处理金融领域的 NLP 任务,可以更好地捕获语言知识和语义信息。文献[18]提出一种混合框架进行股票市场价格预测和情绪分析,对财经新闻数据使用 FinBERT 模型进行分析。以上方法取得了一定的成功,但是缺少捕获关键信息和深层语义的过程,且没有处理对抗样本的扰动影响,容易造成模型错误的分类结果。

1.4 基于对抗训练的方法

神经网络由于其线性的特点,很容易受到线性扰动的攻击。目前很多模型缺乏正确分类对抗性样本的能力,甚至当存在很小的对抗性扰动时,模型的分类结果会受到很大影响。对抗训练最初是由生成式对抗网

络提出的,它是一种在图像领域防御恶意攻击的方法,以提高模型的鲁棒性^[19]。在实际应用中,对抗样本可用于攻击和防御,对抗训练属于其中的一种“防御”方式,通过在数据中添加扰动生成对抗样本并将对抗样本放入模型中训练。将对抗性训练应用于自然语言处理领域的文本分类任务,是一种提高模型分类性能的正则化方法^[20]。

结合对抗训练和预训练模型,既可以具有很强的特征表示能力,也可以进一步强化预训练模型在特定任务中的表现,尤其是在面对对抗攻击或数据扰动时。

2 模型设计

一方面,由于 FinBERT 预训练模型花费了大量数据、大量金融数据进行了训练,因此它更适合金融相关任务。该文使用 FinBERT 嵌入,以捕获语义特征提供给模型。另一方面,循环卷积神经网络(RCNN)能够尽可能多地捕获上下文信息,提取文本中关键特征。因此,采用该模型来进一步提取文本特征。同时创建对抗性示例,使用损失函数的梯度创建扰动,将扰动添加到嵌入中,不断输入新类型的对抗样本进行训练,能够提高模型的鲁棒性和精准度。

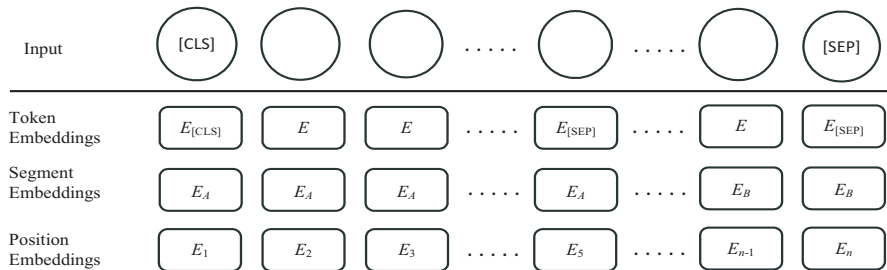


图2 嵌入

(1) Token Embeddings: 将词转为固定维度的向量;

(2) Segment Embeddings: 辅助 FinBERT 区别句子对中的两个句子的向量表示;

(3) Position Embeddings: 辅助 FinBERT 学习到输入的顺序属性。

最后,将3个向量的加和作为模型输入,可以保留更多有效信息及情感倾向,利于后续任务。包含:金融文本数据的词向量、token 向量、文本向量、位置向量。

$$E_{\text{word}} = E_{\text{tok}} + E_{\text{seg}} + E_{\text{pos}} \quad (1)$$

将得到的词向量作为输入量,通过一个包含了多层多个 Transformer 的特征提取器可以提取到特征信息 $T_1 T_2 \cdots T_n$ 。

2.2 RCNN 模块

情感分类任务的关键在于特征包含的信息。虽然 FinBERT 模型可以更好地处理上下文联系和语义信息,但是其缺少关键信息的提取。直接将 FinBERT 获

通过将 FinBERT 嵌入与 RCNN 相结合,并融入对抗训练,该文构建了一个 FinBERT-RCNN-ATTACK 模型,提取金融文本的特征进行情感分析。其具体结构如图1所示。

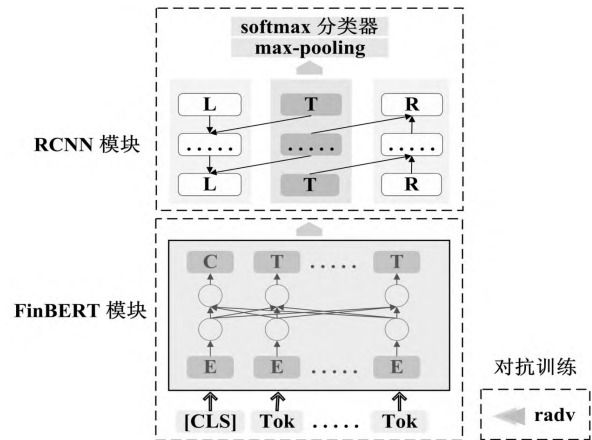


图1 模型结构

2.1 FinBERT 模块

为了更有效地提取金融文本的语义特征,首先利用 FinBERT 预训练模型对预处理后的金融文本数据进行嵌入处理,即 Token Embeddings、Segment Embeddings 和 Position Embeddings(见图2)。

得的特征用于分类的效果有限,其模型性能仍有提高的空间。利用 FinBERT 模型提取语义特征后,再结合其他神经网络模型可以进一步提升整体分类精度。因此,该模型中引入 RCNN,可以更好地提取特征信息进行情感分类。

RCNN 模型将传统卷积神经网络的卷积层替换为内部具有递归结构的循环卷积层,并按前馈连接建立深度网络结构,循环卷积层结构如图3所示。

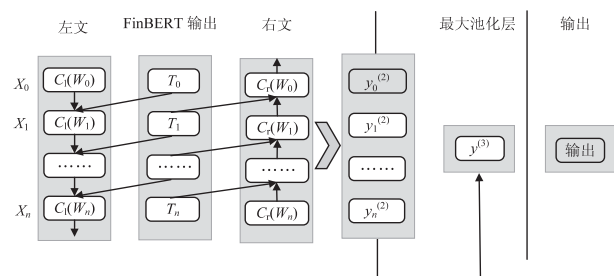


图3 循环卷积层结构

每个单词会生成两个向量,首先产生从左到右的

学习,再产生从右到左的学习。 x 包含了字符及其上下文向量,在 RCNN 中将其输入池化层,通过 max-pooling 取得文本中最具有代表性的关键特征,再用全连接层输出,最后通过 softmax 函数获得分类的结果。

2.3 对抗训练

为提升模型的鲁棒性与泛化性,提升文本分类模型的分类效果,实验中在文本表示模型生成的词向量中加入对抗扰动。对抗样本的特点是:相对于原始输入,对抗扰动很微小,不影响模型分类结果。

为了创建对抗性示例,使用文献[21]的公式,其中使用损失函数的梯度创建扰动:

$$r_{\text{adv}} = \arg \min_{r, \|r\| \leq \epsilon} \log p(y | x + r; \hat{\theta}) \quad (2)$$

假设 $p(y | x; \hat{\theta})$ 是给定输入 x 和模型参数 θ 的标签 y 的概率。其中 r 表示输入上的扰动, $\hat{\theta}$ 是 θ 的常数副本,限制 r 在一个范围内。寻找最坏的扰动,并能够最小化模型的损失。通过线性化在 x 附近。将以下扰动添加到输入嵌入中,以在嵌入空间中创建新的对抗性句子。

3 实验与分析

3.1 数据集

选取 2 个公开数据集对提出的模型和其他现有模型进行实验。数据集 1 采用“CCF 金融信息负面及主体判定”比赛中所使用的数据集,训练集 10 000 条,验证集 5 000 条,测试集 5 000 条,包括文本内容、情绪极性(积极、消极)。数据集 2 采用公开数据集,含有 5 000 条数据,包含训练集 3 000 条,验证集 1 000 条,测试集 1 000 条,情感标签包含消极的,中立的或积极的。

3.2 评估指标

实验使用精确率(Precision)、召回率(Recall)和 F1 分数(F1-Score)作为衡量模型分类性能的标准,计算公式如下所示。

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad (3)$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (4)$$

$$\text{F1} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (5)$$

其中,TP(True Positive)表示被正确预测的正例。即该数据的真实值为正例,预测值也为正例的情况;FP(False Positive)表示被错误预测的正例。即该数据的真实值为反例,但被错误预测成了正例的情况;FN(False Negative)表示被错误预测的反例。即该数据的真实值为正例,但被错误预测成了反例的情况。

3.3 实验环境

实验基于 Pytorch 框架,训练时所使用的 GPU 为 NVIDIA GeForce RTX 3090。模型参数如表 1 所示。

3.4 对比实验

为了验证提出的模型在金融领域情感分析上的效果,本节对以下模型进行了对比实验(见图 4 和图 5)。BiLSTM: BiLSTM 是该领域的基线模型,是在 LSTM^[22]的基础上进行改进和扩展的;TextCNN:文献[23]使用 TextCNN 处理情感分类任务,具有速度快,准确率高的特征;BERT_CNN:在 BERT 模型下游引入 CNN 提取词级特征;BERT_wwm:文献[24]等针对金融领域文本,基于 BERT 提出全词覆盖与特征增强的预训练模型,提高了金融文本情感分类的精度;FinBERT:针对金融领域的 NLP 任务提出的预训练模型,可以更好地捕获语言知识和语义信息;Our_Model:该文提出的模型。

表 1 模型参数

参数	数值
词向量维度	768
Dropout	0.1
学习率	5e-5
Batch Size	32
Pad Size	64

表 2 展示了各个模型在 2 个数据集上的实验结果。可以看出,该文提出的模型在 Precision, Recall 以及 F1-score 上均优于其他模型。

表 2 对比实验 %

数据集	模型	Precision	Recall	F1
数据集 1	BiLSTM	71.25	72.71	71.97
	BERT_CNN	78.32	82.42	80.31
	BERT_wwm	86.38	87.28	86.82
	FinBERT	91.29	91.68	91.48
	Our_Model	94.31	94.83	94.56
数据集 2	BiLSTM	68.49	69.51	68.99
	TextCNN	69.92	71.03	70.47
	BERT_CNN	75.60	80.75	78.09
	FinBERT	90.64	91.45	91.04
	Our_Model	93.76	93.59	93.67

首先,与 BiLSTM 相比,文中模型在两个数据集的 F1 值分别提高了 31 个百分点和 35 个百分点。与 TextCNN 相比,提升了 35 个百分点左右。改进主要是由于预训练模型提供的上下文嵌入。

其次,与 BERT_CNN 和 BERT_wwm 相比,性能

提升了 8 个百分点到 17 百分点。这是由于 FinBERT 是在大规模金融语料库上预训练的特定领域语言模型,利用它处理金融领域任务,使性能得到了明显提升。

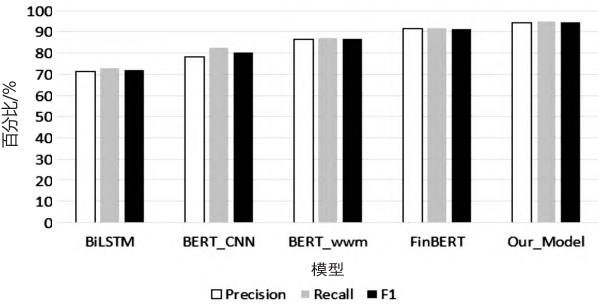


图 4 数据集 1 对比实验

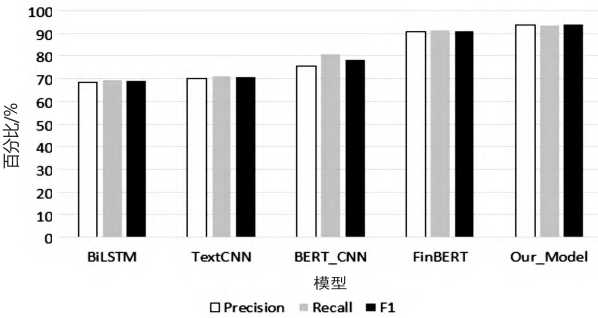


图 5 数据集 2 对比实验

相比于 FinBERT 模型,文中模型在两个数据集上提升了 3 个百分点左右。这是由于提取到语义特征后,进一步捕获了关键信息,更好地挖掘文本深层情感信息,并在嵌入层加入扰动,增加了模型的泛化性和鲁棒性。

3.5 消融实验

为了考察模型中各个模块的有效性,设置了模型中不同结构的消融实验(见图 6 和图 7),模型具体细节如下:FinBERT:金融领域的预训练模型;FinBERT-RCNN:预训练模型下游引入 RCNN 模型提取关键信息;FinBERT_对抗训练:预训练模型引入对抗训练;Our_Model:该文提出的模型。

本节进行了消融实验,用来分析各模块的性能,结果记录在表 3。通过消融实验发现模型的每个模块在提高性能方面都起着关键作用。

表 3 消融实验 %

数据集	模型	Precision	Recall	F1
数据集 1	FinBERT	91.29	91.68	91.48
	FinBERT_RCNN	91.89	92.28	92.08
	FinBERT_对抗训练	92.70	93.42	93.05
	Our_Model	94.31	94.83	94.56
数据集 2	FinBERT	90.64	91.45	91.04
	FinBERT_RCNN	93.69	93.16	93.42
	FinBERT_对抗训练	91.43	92.05	91.73
	Our_Model	93.76	93.59	93.67

首先,FinBERT - RCNN 模型对预处理后的金融文本数据进行嵌入处理,采用 FinBERT 语言模型处理上下文联系和语义信息,并引入 RCNN 提取关键信息进行情感分类。相比于未处理的 FinBERT 模型,在 F1 结果上提升了 0.6~2.6 百分点。

下面在模型的基础上,结合了对抗训练,来增强模型的鲁棒性和泛化性,提升对抗线性扰动攻击的能力。与未经处理的 FinBERT 模型相比,数据集 1 在 F1 结果上有 17 百分点的提升,数据集 2 提升 0.7 百分点。数据集 1 由于具有更强的对抗样本,结合了对抗训练进行处理后获得较大的提升。数据集 2 相对提升较小的原因关键在于数据的特性,该数据集的对抗样本扰动较弱,因此更适合 FinBERT - RCNN 提取局部特征和长距离依赖。

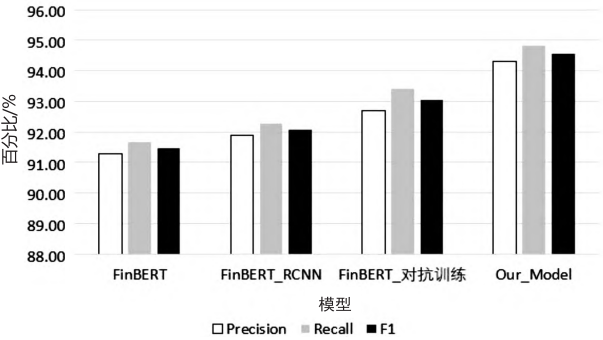


图 6 数据集 1 消融实验

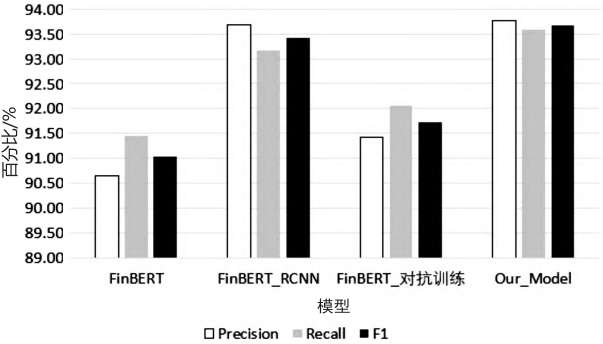


图 7 数据集 2 消融实验

最后,提取 FinBERT 模型的语义特征输出并结合 RCNN 提取特征包含的关键信息,并使用损失函数的梯度创建扰动,在嵌入中添加扰动,得到最终模型。通过对比可知,该文提出的模型相比于未处理的 FinBERT 模型,在 F1 结果上提升 3 个百分点左右,证实了模型的有效性。

4 结束语

该文提出了一种应用于金融领域的情感分类模型。该模型集成了 FinBERT、RCNN、对抗训练,通过使用 RCNN 模块对 FinBERT 嵌入层的输出进行特征提取,捕获关键特征提高模型情感分类的准确性,并在

嵌入阶段添加扰动,提高模型的鲁棒性和泛化性。实验证明,提出的金融领域情感分析模型与其他模型相比在性能上获得明显提升。

参考文献:

- [1] CAROSIA A E O, COELHO G P, SILVA A E A. Analyzing the Brazilian financial market through Portuguese sentiment analysis in social media[J]. *Applied Artificial Intelligence*, 2020, 34(1): 1-19.
- [2] MA Y, MAO R, LIN Q, et al. Multi-source aggregated classification for stock price movement prediction[J]. *Information Fusion*, 2023, 91: 515-528.
- [3] DEVLIN J, CHANG M W, LEE K, et al. Bert pre-training of deep bidirectional transformers for language understanding [C]//Proceedings of the 2019 conference of the North American chapter of the association for computational linguistics; human language technologies. Minneapolis: NAACL, 2019: 4171-4186.
- [4] 胡新荣, 王哲, 刘军平, 等. 基于多层注意力机制的服装电商评论情感分析[J]. *计算机技术与发展*, 2022, 32(1): 67-72.
- [5] 张影, 方贤进, 杨高明. 面向自然语言处理领域的对抗样本生成方法[J]. *计算机技术与发展*, 2023, 33(3): 98-104.
- [6] LOUGHRAN T, MCDONALD B. Textual analysis in accounting and finance: a survey[J]. *Journal of Accounting Research*, 2016, 54(4): 1187-1230.
- [7] VAN ATTEVELDT W, VAN DER VELDEN M A C G, BOUKES M. The validity of sentiment analysis: comparing manual annotation, crowd-coding, dictionary approaches, and machine learning algorithms[J]. *Communication Methods and Measures*, 2021, 15(2): 121-140.
- [8] ALAMRANI Y, LAZAAR M, ELKADIRI K E. Random forest and support vector machine based hybrid approach to sentiment analysis[J]. *Procedia Computer Science*, 2018, 127: 511-520.
- [9] YOON B, JEONG Y, KIM S. Detecting a risk signal in stock investment through opinion mining and graph-based semi-supervised learning[J]. *IEEE Access*, 2020, 8: 161943-161957.
- [10] 金雪敏. 基于深度学习的财经文本情感分析技术研究[D]. 成都: 电子科技大学, 2023.
- [11] 王浩畅, 孙铭泽. 基于 ERNIE-RCNN 模型的中文短文本分类[J]. *计算机技术与发展*, 2022, 32(6): 28-33.
- [12] ALMALIS I, KOULOUMPRIS E, VLAHAVAS I. Sector-level sentiment analysis with deep learning[J]. *Knowledge-Based Systems*, 2022, 258: 109954.
- [13] 高杨帆. 面向金融领域的文本情感分析关键技术研究[D]. 武汉: 华中科技大学, 2023.
- [14] 游兰, 曾晗, 韩凡宇, 等. 基于 BERT-BiGRU 集成学习的情感语义识别[J]. *计算机技术与发展*, 2023, 33(5): 159-166.
- [15] ZHAO L, LI L, ZHENG X, et al. A BERT based sentiment analysis and key entity detection approach for online financial texts[C]//2021 IEEE 24th international conference on computer supported cooperative work in design (CSCWD). Dalian: IEEE, 2021: 1233-1238.
- [16] DAUDERT T. Exploiting textual and relationship information for fine-grained financial sentiment analysis[J]. *Knowledge-Based Systems*, 2021, 230: 107389.
- [17] ARACI D. FinBERT: financial sentiment analysis with pre-trained language models[J]. *arXiv*: 1908.10063v1, 2019.
- [18] SRIJIRANON K, LERTRATANAKHAM Y, TANANTONG T. A hybrid framework using PCA, EMD and LSTM methods for stock market price prediction with sentiment analysis[J]. *Applied Sciences*, 2022, 12(21): 10823.
- [19] GOODFELLOW I J, POUGET-ABADIE J, MIRZA M, et al. Generative adversarial nets[J]. *Stat*, 2014, 1050: 10.
- [20] 刘一廷, 宋珣. 自然语言处理中的对抗攻击与防御研究综述[J]. *现代计算机*, 2020(36): 52-57.
- [21] MIYATO T, DAI A M, GOODFELLOW I. Adversarial training methods for semi-supervised text classification[J]. *Stat*, 2016, 1050: 7.
- [22] HOCHREITER S, SCHMIDHUBER J. Long short-term memory[J]. *Neural Computation*, 1997, 9(8): 1735-1780.
- [23] 程钢, 陈秀明, 于翔. 基于 TextCNN 融合模型的离散情感分析[J]. *科学技术创新*, 2023(21): 124-127.
- [24] 朱鹤, 陆小锋, 薛雷. 基于 BERT 的金融文本情感分析模型[J]. *上海大学学报: 自然科学版*, 2023, 29(1): 118-128.