

面向手机动画的基于注意力机制和卷积神经网络的短信情感分析^①

宗玉英, 段晓芳

(北京工业大学 信息学部, 北京 100124)

摘要: 2008年, 中科院张松懋研究员提出了将3D动画自动生成技术应用在手机短信中. 短信情感分析是手机3D动画自动生成系统的一个重要环节. 目前系统中使用的方法是传统的机器学习方法, 准确率较低, 无法达到实用的目的. 而近几年, 深度学习在情感分析任务中取得了较好的效果, 卷积神经网络可以自动提取短信中的语义情感特征, 且注意力机制可以自动为词加权获取信息. 为此, 本文提出将深度学习中的注意力机制和卷积神经网络相结合应用于手机短信自动生成系统中的情感分类. 实验表明, 基于注意力机制的卷积神经网络比之前的方法准确率、召回率和F值都有明显的提高.

关键词: 短信情感分析; 词向量; 注意力机制; 卷积神经网络

引用格式: 宗玉英, 段晓芳. 面向手机动画的基于注意力机制和卷积神经网络的短信情感分析. 计算机系统应用, 2018, 27(8): 226-231. <http://www.c-s-a.org.cn/1003-3254/6468.html>

Sentiment Analysis of Short Messages Based on Attention Mechanism and Convolution Neural Network for Mobile Animation

ZONG Yu-Ying, DUAN Xiao-Fang

(Faculty of Information Science, Beijing University of Technology, Beijing 100124, China)

Abstract: In 2008, ZHANG Song-Mao, a researcher of Chinese Academy of Sciences, proposed the application of 3D animation automatic generation technology for mobile phone SMS. The sentiment analysis of SMS is an important part of the 3D animation automatic generation system. At present, the method used in the system is a traditional machine learning method, which has a low accuracy and cannot achieve a practical purpose. In recent years, deep learning has achieved good results in the task of sentiment analysis. Convolutional neural network can automatically extract the semantic and sentiment features of text messages, and attention mechanism can automatically obtain weighting information for words. Therefore, this study proposes to apply the attention mechanism and convolutional neural network in deep learning to the classification of sentiment analysis in the system of SMS automatic generation. Experiments show that the convolutional neural network based on attention mechanisms has significantly improved the accuracy, recall rate, and F-value than the previous methods.

Key words: sentiment analysis of short messages; word vector; attention mechanism; convolutional neural network

1 引言

2008年, 在审视了3G通信技术的发展趋势之后, 中科院张松懋研究员创造性地提出了将3D动画自动

生成技术应用到手机短信上的想法^[1], 并带领课题组实现了手机3D动画自动生成系统. 手机3D动画自动生成技术将动画自动生成技术应用于中文手机短信领域,

^① 收稿时间: 2017-12-07; 修改时间: 2017-12-27; 采用时间: 2018-01-08; csa 在线出版时间: 2018-07-28

不仅立足于一个崭新的应用角度,并尝试研究和解决过程中出现的问题,这在人工智能领域具有一定的研究意义和价值。

手机3D动画自动生成系统的处理过程大致分为四个阶段:短信信息抽取、情节定性规划、场景定量规划和网络渲染。短信信息抽取处于手机3D动画自动生成系统首要和关键的位置,在这一过程中,短信内容的理解,即从短信文本中抽取出用于指导动画生成的信息是最基础、也是最关键的步骤^[2]。短信信息抽取又包括语义分析、否定识别、命名实体识别、主客观分析和情感分析等。而情感分析作为信息抽取领域的重要研究课题,对于引导动画的自动生成有着关键的作用。但是,中文短信情感分析的研究目前可利用的资源还不多,目前尚没有一个覆盖面较广的情感词库,且短信字数少,信息量少,结构随意^[3],可利用的上下文信息有限,以及自由开放的特点,给短信情感分类研究带来了一定的难度。

课题组在做情感分析之前,首先做的是主客观的分类。如果主客观分类的结果是主观,则会进行下一步情感分类;否则,不做情感分析。而主客观分析使用传统的机器学习方法得到的效果较好。但是,目前课题组采用传统NLP方法实现的情感分类准确率不高,系统人工的成本较高,系统的泛化和迁移能力差^[4],不能达到实用的目的。而深度学习方法能较好地弥补传统NLP方法的缺陷^[4]。鉴于此,本文采用基于深度学习的算法对手机动画自动生成系统的短信文本进行情感分类。一般的情感分析是将文本分为积极、消极两类,但是,考虑到系统中动画可以表现的情感形式,本文将短信文本中的情感分为喜、怒、哀、惧四类。

2 相关工作

情感分析又称为意见挖掘 (opinion mining),是借助计算机帮助用户快速获取、整理和分析相关评论信息,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,它包含了一系列的子任务,如情感信息抽取、情感倾向分类、情感信息检索和情感信息归纳等^[5]。本文的短信情感分析是将输入的短信进行情感分析,最终分为喜、怒、哀、惧四类。得到的情感分类结果,可以为手机短信动画后续的表情、动作、灯光等提供有力的参考,从而使生成的动画更加丰富,更加生动地表现短信中的信息。

目前,手机动画系统中已有两版情感分析,第一版使用的是SVM算法,第二版使用是集成算法,但是准确率都不高,不能达到实用的目的。

而深度学习中在自然语言处理领域已经取得了显著成效,2011年,Collobert等在文献^[6]中训练得到C&W模型,去完成自然语言处理的词性标注、命名实体识别、语义角色标注、情感分析等,有不错的成效。考虑到句子中词的上下文信息,2014年, Kim等^[7]提出使用卷积神经网络 (Convolutional Neural Network, CNN) 进行句子建模,做情感分类任务,并在多个数据集上均取得了很好的结果。而后, Dos Santos等提出CharSCNN (Character to Sentence Convolutional Neural Network) 模型^[8]为英语单词增加了字符的特征表示层,可以提高对于英文Twitter中不规则单词的表示能力,取得了不错的成效。Bahdanau等人将注意力机制和循环神经网络 (Recurrent Neural Network, RNN) 结合以提高机器翻译任务的准确率^[9],使注意力机制成功应用到自然语言处理领域,这也是首次将注意力机制应用到自然语言处理领域。此外,随着研究的深入, Yin等人在2015年提出一种基于注意力机制的卷积神经网络,并将该网络用在句子对建模任务中^[10],并取得了比之前更高的准确率。

基于注意力机制的卷积神经网络可以通过一个注意力矩阵计算出句子相应的注意力特征,在模型训练时让神经网络高度关注某些信息而注意力机制可以自动对文本矩阵进行语义加权^[11]。为此,提出在卷积神经网络中,引入注意力机制,使模型准确地表示出每个词的重要性,以获取更准确的信息。

本文提出将卷积神经网络和注意力机制相结合的模型应用到短信文本的情感分类中。经实验验证,引入注意力机制的卷积神经网络模型比单纯的卷积神经网络模型更有效。

3 卷积神经网络和注意力机制卷积神经网络

为了提高手机短信动画系统情感分析的准确率,本文提出将卷积神经网络和注意力机制相结合的模型结构,模型的整体流程为:首先对输入的短信使用哈工大分词系统^[12]进行分词,然后对分词的结果进行词向量编码,得到词向量矩阵,将矩阵和注意力机制相结合,得到注意力词向量矩阵,输入到卷积层进行卷积操作,得到特征表示,将特征表示和注意力机制结合,通过全

连接层之后使用分类器得到情感概率矩阵,进而得到情感标签.

3.1 词向量 word2vec

对于短信数据集 D , 其中包含短信文本 $X = \{x_1, x_2, \dots, x_m\}$ 以及每条短信对应的情感标签 $L = \{l_1, l_2, \dots, l_m\}$, 其中, 每条短信 X_i 包含 n 个词, 表示为 $X_i = \{x_{i1}, x_{i2}, \dots, x_{in}\}$. 首先使用哈工大分词系统^[12]进行分词, 然后, 将分词的结果输入到 word2vec 模型, 得到语料中每个词的向量表示. 本文中的 word2vec 模型中使用 skip-gram 模型.

3.2 卷积神经网络

卷积神经网络在自然语言处理中已经颇有成效, 因其无需进行文本的大量预处理工作, 可以使用卷积

操作自动对文本进行特征提取, 受到了广大研究者的青睐. 基于卷积神经网络的情感分类模型结构, 输入是短信中词向量表示的矩阵, 输出是预定义的情感类别中的一种. 该模型为四层的结构, 分别为输入层、卷积层、池化层以及分类层, 如图 1 所示. 输入层就是 Word embedding 层, 根据 word2vec 得到短信中每个词的向量表示, 得到词向量矩阵, 对于给定长度为 n 个词的短信, 维度为 k 的词向量, 则, 输入矩阵为 $n \times k$ 维矩阵. 卷积层对词向量矩阵进行卷积操作, 得到卷积之后的特征矩阵. 池化层对卷积之后的特征矩阵进行池化, 获取最重要的特征. 池化层的结果和注意力机制结合输入到全连接层, 经过激活函数处理, 就可以得到情感概率矩阵, 并得到最终的分类结果.

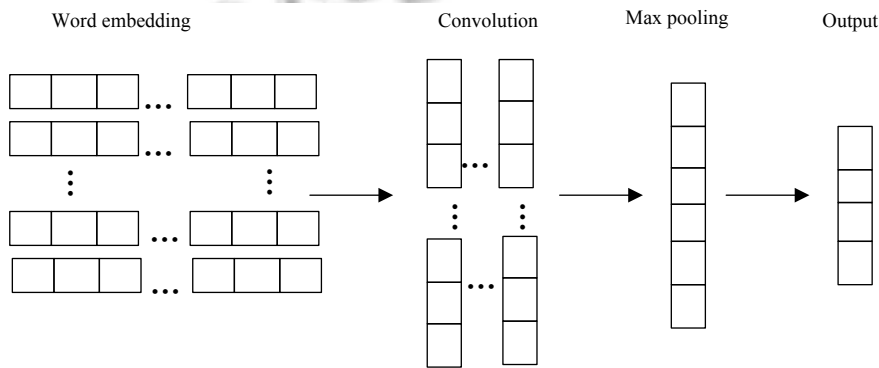


图 1 卷积神经网络模型结构图

3.3 注意力机制

Bahdanau 等人将注意力机制和循环神经网络结合应用到提高机器翻译领域^[9], 取得了显著的功效. Yin 等人将注意力机制的卷积神经网络用在句子对建模任务中^[13], 并取得了比之前更高的准确率. 文献^[11]提出将多注意力机制和词向量结合做特定目标的情感分类, 其实验表明多注意力机制可以有效地提高情感分类的准确率. 文献^[14]提出将注意力机制和卷积之后的特征表示相结合, 实验证明此模型是有效的. 为了进一步探索注意力机制的有效性, 为此, 本文提出将注意力机制和词向量以及卷积之后的特征表示相结合, 并应用到手机短信动画情感分类系统中, 来进一步提高分类的准确率, 以达到实用的目的.

注意力机制可以在训练的过程中知道哪一部分数据是重要的, 从而使模型重点关注这些重要信息, 以提高注意力. 对于词级别的分类而言, 词层面的信息是最

重要的, 将注意力机制和词向量输入矩阵相结合, 即词向量注意力机制, 可以得到注意力矩阵, 此矩阵中包含每个词的重要程度, 然后使用卷积操作, 再使用注意力机制等操作, 可以得到更好的模型分类性能.

对于短信 $m = \{w_1, w_2, \dots, w_i, \dots, w_n\}$ 而言, 提取出目标词 w_i 的词向量 x_i 作为注意力矩阵, 将注意力矩阵和短信 m 的词向量矩阵进行运算操作, 可得到注意力特征矩阵 A^c ,

$$A_{i,i} = t_i \cdot x_i \tag{1}$$

$$A_{i,i}^c = \frac{\exp(A_{i,i})}{\sum_{j=1}^n \exp(A_{i,j})} \tag{2}$$

利用计算得到的注意力矩阵 A^c 和原来的词向量使用向量拼接运算, 可分别得卷积神经网络的输入矩阵:

$$z_i^c = x_i \oplus A_{i,i}^c \tag{3}$$

接着将 z_i 矩阵输入到卷积层进行卷积操作. 此外, 在经过卷积之后, 本文也使用了注意力机制, 以获取不同重要程序的特征信息, 从而提高分类的准确率. 此时的注意力文本用 a_i 表示:

$$u_i = \text{fun}(x_{ij}, L_i) \tag{4}$$

$$\alpha_i = \frac{\exp(u_i)}{\sum_{i=1}^m \exp(u_i)}, 1 \leq i \leq m \tag{5}$$

$$\text{att}_i = \sum_{i=1}^m \alpha_i \cdot x_i \tag{6}$$

其中, x_i 代表一条短信, L_i 代表此短信对应的标签, fun 代表含一个隐层的前向网络函数, α_i 与 att_i 代表文本中每个词的重要度信息.

$$L_{pre} = \text{soft max}(w \cdot (c \otimes a_i) + b) \tag{7}$$

其中, L_{pre} 表示预测得到的情感标签, w 为全连接层的权重矩阵, c 为池化层的特征表示, att_i 为注意力文本, b 为偏置项.

引入注意力机制的卷积神经网络模型, 如图 2 所示.

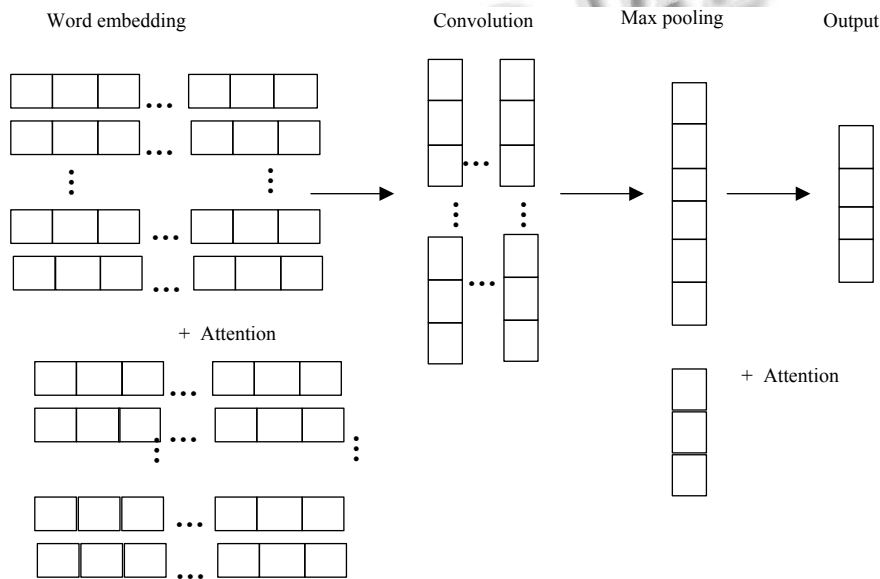


图 2 引入注意力机制的卷积神经网络结构图

4 实验

将传统机器学习方法 SVM、卷积神经网络和注意力机制卷积神经网络分别在数据集上实验, 解决短信动画生成系统的情感分析问题. 通过对比实验, 验证了在本任务中, 注意力机制比 SVM 和单纯的卷积神经网络有更好的情感分析性能. 在本文的实验中, 语料库为 SMS-2008 标注中文短消息库^[15]中的 10 万条短信, 经处理得到含有喜、怒、哀、惧情感的短信共 18163 条, 词向量采用 word2vec 的 skip-gram 训练得到, 其中每个词的维度为 100 维. 本文采用哈工大分词系统^[12]进行分词.

4.1 实验数据准备

本文采用 SMS-2008 标注中文短消息库^[15]中的短

信进行实验, 首先, 去除主客观分类结果为客观, 即不包含喜、怒、哀、惧情感的短信, 然后手工将数据样本的情感标记为喜、怒、哀、惧四类, 表 1 给出本文实验使用的数据并统计各情感极性的语料数.

表 1 各情感极性的语料数

情感极性	数量
喜	4928
怒	4947
哀	6888
惧	1400

4.2 超参数设置

本文使用注意力机制和卷积神经网络相结合的类型完成, 实验中使用多种卷积核对输入矩阵进行卷积

操作,训练中的损失函数采用 `categorical_crossentropy`, 梯度函数使用 `adam` 来更新梯度, 激活函数使用适合多分类的 `softmax`, 其他参数如表 2 所示。

表 2 实验超参数设置

Parameter	Parameter Description	Value
h	Window Size	2,3,4
p	Dropout out	0.5
b	Batch_Size	128
n	Num of Feature Map	128

4.3 对比实验

将本文提出的方法和以下的两种方法在相同的数据集上进行实验:

1) SVM. Svetlana Kiritchenko^[16]提出的 SVM 分类模型, 该模型使用 SVM 对文本进行分类, 并取得了不

错的成效。

2) CNN. Kim 等^[7]提出使用卷积神经网络 CNN (Convolutional Neural Network, CNN), 是最基本的卷积神经网络, 并取得了比其他方法更好的分类效果。

3) ATT_CNN. 本文提出的注意力机制和卷积神经网络相结合的模型。

4.4 实验结果分析

本文将 3 种模型在短信语料库上分别进行实验, 分析出短信的情感极性, 并预测情感极性。实验中, 将语料分为 3 部分, 分别是训练集、验证集和测试集, 采用的比例是 8:1:1。表 3 给出 3 组实验在语料库上分别实验的训练准确率和验证准确率, 得到分类模型之后, 又对预测集进行实验并得到分类的准确率、召回率和 F 值。

表 3 各模型在本语料实验中的预测准确率、召回率和 F 值

Model	Train accuracy	Val accuracy	Predict Accuracy	Predict Recall Rate	Predict F Value
SVM	0.7553	0.7453	0.7405	0.7447	0.7485
CNN	0.8572	0.8064	0.7821	0.7819	0.7851
ATT_CNN	0.8924	0.8426	0.8235	0.8234	0.8257

从表 3 的结果可以看出, 在本文的情感分类任务中, CNN 和 ATT_CNN 在此语料上的分类效果都比传统法的机器学习 SVM 分类模型好, 原因在于 CNN 和 ATT_CNN 可以进行卷积操作, 从而更好地提取文本特征。而引入了注意力机制后的卷积神经网络模型相比 CNN 模型分类的正确率也有所提高, 原因在于注意力机制可以有针对性地提取出文本特征中的情感特征, 从而使注意力机制模型比 CNN 分类模型在验证集和测试集上都取得更好的分类性能。由以上实验结果和分析可知, 本文提出的注意力机制和卷积神经网络结合是可行的, 同时也是有效的。

5 总结

本文研究的主要内容是在手机 3D 动画自动生成系统中提高情感分析的分类性能, 以达到实用的目的。为此, 提出将注意力机制应用到此系统中。首先使用 `gensim` 得到“词-索引”形式的键值对, 然后使用 `word2vec` 得到词的向量表示, 与注意力矩阵进行运算, 通过卷积神经网络的卷积运算自动提取到文本的特征, 再结合注意力矩阵得到最重要的特征。实验表明, 基于注意力机制的卷积神经网络比之前方法的准确率都高,

也证明了注意力机制的可行性和有效性。下一步工作, 将进一步探讨注意力机制与其他神经网络相结合的深度学习方法。

参考文献

- 1 陆汝钦, 张松懋. 从故事到动画片——全过程计算机辅助动画自动生成. 自动化学报, 2002, 28(3): 321-348.
- 2 吴中彪. 全过程计算机辅助手机 3D 动画自动生成系统的设计与实现[硕士学位论文]. 北京: 北京工业大学, 2011: 11-38.
- 3 Arif MH, Li JX, Iqbal M, *et al.* Sentiment analysis and spam detection in short informal text using learning classifier systems. *Soft Computing*, 2017: 1-11. [doi: 10.1007/s00500-017-2729-x]
- 4 张志华. 基于深度学习的情感词向量及文本情感分析的研究[硕士学位论文]. 上海: 华东师范大学, 2016.
- 5 赵妍妍, 秦兵, 刘挺. 文本情感分析. 软件学报, 2010, 21(8): 1834-1848.
- 6 Collobert R, Weston J, Bottou L, *et al.* Natural language processing (almost) from scratch. *The Journal of Machine Learning Research*, 2011, 12: 2493-2537.
- 7 Kim Y. Convolutional neural networks for sentence classification. *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing*. Doha,

- Qatar. 2014. 1746–1751.
- 8 Dos Santos CN, Gatti M. Deep convolutional neural networks for sentiment analysis of short texts. Proceedings of COLING 2014, the 25th International Conference on Computational Linguistics: Technical Papers. Dublin, Ireland. 2014. 69–78.
- 9 Bahdanau D, Cho K, Bengio Y. Neural machine translation by jointly learning to align and translate. arXiv:1409.0473, 2014.
- 10 Cui H, Mittal V, Datar M. Comparative experiments on sentiment classification for online product reviews. Proceedings of the Twenty First National Conference on Artificial Intelligence. Boston, MA, USA. 2006. 1265–1270.
- 11 梁斌, 刘全, 徐进, 等. 基于多注意力卷积神经网络的特定目标情感分析. 计算机研究与发展, 2017, 54(8): 1724–1735.
- 12 刘挺, 车万翔, 李正华. 语言技术平台. 中文信息学报, 2011, 25(6): 53–62. [doi: [10.3969/j.issn.1003-0077.2011.06.008](https://doi.org/10.3969/j.issn.1003-0077.2011.06.008)]
- 13 Yin WP, Schütze H, Xiang B, *et al.* ABCNN: Attention-based convolutional neural network for modeling sentence pairs. arXiv:1512.05193, 2015.
- 14 冯兴杰, 张志伟, 史金钊. 基于卷积神经网络和注意力模型的文本情感分析. 计算机应用研究, 2018, 35(5).
- 15 马旭, 徐蔚然, 郭军, 等. SMS-2008 标注中文短信息库. 中文信息学报, 2009, 23(4): 22–26. [doi: [10.3969/j.issn.1003-0077.2009.04.004](https://doi.org/10.3969/j.issn.1003-0077.2009.04.004)]
- 16 Kiritchenko S, Zhu XD, Cherry C, *et al.* NRC-Canada-2014: Detecting aspects and sentiment in customer reviews. Proceedings of the 8th International Workshop on Semantic Evaluation. Dublin, Ireland. 2014. 437–442.