

基于深度学习的垃圾邮件检测

俞莹妹, 禹素萍, 许武军, 范红

东华大学信息科学与技术学院, 上海

收稿日期: 2023年3月17日; 录用日期: 2023年4月14日; 发布日期: 2023年4月21日

摘要

邮件是日常生活中的一种通讯工具, 但垃圾邮件对用户造成严重困扰, 因此改进垃圾邮件识别技术、提升其准确率和效率具有重要现实意义。在文本分类领域, 深度学习有很好的应用效果。故文章提出了一种基于CNN的BiGRU-Attention模型, 旨在充分利用CNN的特征提取能力和BiGRU的全局特征提取能力。引入注意力机制能够突出显示重要文本, 前后共经过两层双向门控循环单元, 从而更全面地提取邮件文本特征。实验数据选取Trec06c数据集, 并与其他分类模型对比, 结果表明, 检测准确率达到91.56%。

关键词

垃圾邮件, 文本分类, 深度学习, 双向门控循环单元, 注意力机制

Spam Detection Based on Deep Learning

Yingmei Yu, Suping Yu, Wujun Xu, Hong Fan

College of Information Science and Technology, Donghua University, Shanghai

Received: Mar. 17th, 2023; accepted: Apr. 14th, 2023; published: Apr. 21st, 2023

Abstract

Email is a communication tool in daily life, but spam has caused serious problems for users, As a result, it is crucial to improve spam identification technology and improve its accuracy and efficiency. In the field of text classification, deep learning has a good application effect. In order to fully utilize CNN's feature extraction capabilities and BiGRU's global feature extraction capabilities, this article suggests a CNN-based BiGRU-Attention model. The introduction of the attention mechanism can highlight important text, which passes through two layers of two-way gated loop units before and after, so as to extract more comprehensive features of email text. The experimental data is selected from Trec06c dataset and compared with other classification models. The results show that the detection accuracy reaches 91.56%.

Keywords

Spam, Text Classification, Deep Learning, BiGRU, Attention Mechanism

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

当今社会, 互联网的快速发展使得电子邮件在人们的日常生活中发挥了很大的功能, 既可以提高工作效率、节约成本, 又可以促进人们之间的交流和沟通。但同时, 也有不少不法商人、不法之徒通过电子邮件做广告、散布不良信息, 给人们的生活和工作带来了不必要的麻烦和困扰。用户经常会收到一些垃圾邮件, 其内容毫无营养, 有些包含一些恶意链接, 可能会造成电脑中毒、网银被盗刷等危害用户安全的情况, 扰乱网络安全环境。因此, 如何快速准确地检测邮件是否为垃圾邮件是值得研究的重大课题。

迄今为止, 国内外学者对垃圾邮件的识别技术进行了大量的研究, 也取得了良好的效果。当前的研究方法有两大类, 一类是基于邮件发送来源检测的技术, 通过分析判断邮件的发送者地址和网络 IP 地址进行[1], 包括黑白名单识别查询技术和反向 DNS 技术等。第二种是根据邮件内容信息进行检测的技术, 将邮件的内容特征提取出来, 然后将这些特征转化为向量形式, 利用分类器对其进行训练, 从而获得一个分类模型, 例如朴素贝叶斯[2]、SVM、K 近邻[3]等算法。未知邮件再通过训练的分类模型得出该邮件的分类标签。近几年, 深度学习发展迅速, 在垃圾邮件检测领域, 对卷积神经网络(Convolutional Neural Networks, CNN)模型的优化改进已经有了很大的进展。黄鹤等人[4]提出了基于 Skip-gram 的 CNN 模型, 并结合了 Highway 网络, 用低维度的特征向量反映文本特征, 使得邮件分类模型的准确率得到提高。彭毅等人[5]提出了一种基于 BERT_DPCNN 的文本分类模型。通过 BERT 预训练模型获得文本向量, 随后将其输入到 DPCNN 模型中, 以获得更多的语义信息, 避免了梯度消失, 从而提高了模型的性能。目前, 尽管已经有了很多垃圾邮件检测的方法, 但是仍在某些方面需要改进, 例如检测速度、检测准确率等。

CNN 是一种具备特征提取能力和数据降维能力的神经网络。循环神经网络(Recurrent Neural Network, RNN)是一种用于处理序列数据输入的递归神经网络。之后, 研究者们又提出了 RNN 的一种变体: 双向门控循环单元[6] (BiGRU)。BiGRU 层由两个 GRU 组成, 一个用于输入序列的前向处理, 一个用于后向处理。这两个网络均与同一输出端相连, 能够实现更完整的特征提取。GRU 相比 LSTM, 能提供更快的计算和更高的效率, 并且两者的准确率相当。垃圾邮件检测需要从电子邮件中提取关键词, 注意力机制(Attention)可以给电子邮件中的关键词分配高权重, 并提取重要信息。本文所提出的基于 CNN 的 BiGRU-Attention 模型, 综合了以上三种模式各自的功能与优势, 并与其它模型进行了比较, 得到了较好的结果, 数据集选取了公开的 Trec06c 邮件数据集。

2. 中文文本分类

本文中使用的垃圾邮件检测方法是基于对电子邮件文本内容的检测和分类, 即基于文本的分类[7]。对于文本的分类[8]的任务是处理一个未知类别的文件, 并确定它属于哪个或哪些类别。在这个过程中, 使用一些已知的文本数据训练一个模型, 来确定文本特征所对应的类别, 训练好的模型再对其他文本进

行分类。下面给出了一个基本的流程图(图 1)。

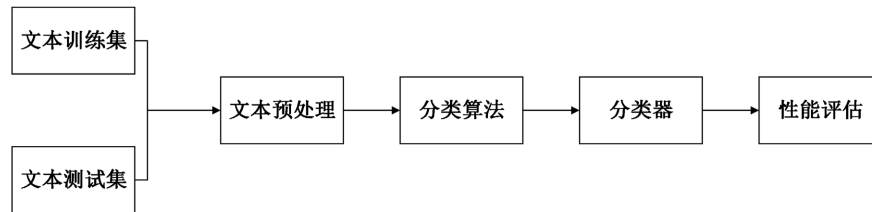


Figure 1. Basic process of Chinese text classification
图 1. 中文文本分类的基本流程

从图 1 可以看出, 文本预处理和分类算法是文本分类的核心部分。

3. CNN-BiGRU-Attention 模型

传统的决策树、贝叶斯等文本分类方法都存在着高维度、高稀疏性、特征表达不强等问题。Kim [9] 在 2014 年提出了 textCNN, CNN 在文本分类问题上取得了比较好的效果。但是, 该 CNN 模型的缺点也很明显: 池化层容易丢失文本序列的有价值信息, 不能解决文本时序特征的问题。RNN 作为序列模型可以解决 CNN 对长文本处理效果不佳的问题, 但无法解决长时依赖问题。因此, 本文在 CNN 的基础上, 引入了 RNN 变体模型: BiGRU。它能够更好地提取出文本的整体结构和特征[10], 而 CNN 则能够有效地提取出文本的局部特征, 再通过引入注意力机制, 突出显示文本的重要特征, 从而提高分类模型的准确性和效率。如图 2 所示, CNN 层从邮件文本数据中提取特征, 然后经过第一个 BiGRU 层, 将 CNN 层提取的特征和原始文本向量进行正反方向的提取组合, 将其输出向量输入进 Attention 层后进入第二个 BiGRU 层, 以全面地获取邮件文本的特征向量, 经过非线性层后最终经过 Softmax 层得到分类结果。

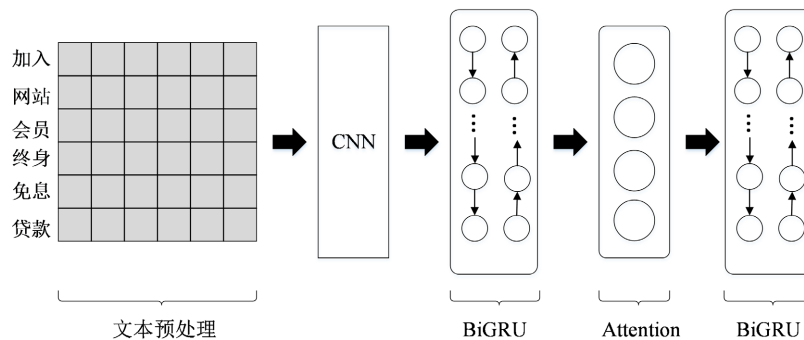


Figure 2. Model architecture
图 2. 模型架构

3.1. 邮件文本预处理

要想计算机能够“理解”电子邮件文本内容, 就要对电子邮件文本的数据进行文本预处理, 最终转换为格式化数据, 为后面的特征提取做准备。预处理包括过滤非法字符、分词、去除停用词、词向量转换等操作。

3.1.1. 过滤非法字符

因为数据的采集来自于不同的网页, 因此有可能出现大量的非法字符, 比如标点符号、表情符号、

链接、乱码或者其它的无意义的符号。在文本分类中，这些文字是没有任何用处的，所以必须对其进行过滤筛选，从而为提取邮件文本信息提供便利。

3.1.2. 分词

按照一定的规则将一个句子分成若干个词组的过程称为分词。分词可分为英文和中文分词[9]。在英文中，空格起到了自然分词的作用。然而，在中文文本中，词与词之间没有这样的天然分词器，这使得中文的分词比英文的分词更加困难和复杂。

本研究中的模型使用了“结巴[11] (Jieba)”中文分词库，具有方便、快速、准确率高等优点。为了提高邮件正文的分词精度，使用了自定义词典来处理 Jieba 库无法处理的句子分词。目前，jieba 是使用最广泛的分词方法之一[12]。

3.1.3. 词向量转换

经过上述步骤，图 3 中的数据已转换为图 4 所示的结构，从而产生更规整的电子邮件文本内容。

```
<div id="mailContent"><p>加入$$网站
~会员:终身免息贷款 ¥ ¥ </p></div>
```

Figure 3. Original email text
图 3. 原始邮件文本

加入 网站 会员 终身 免息 贷款

Figure 4. Preliminary processing of email text
图 4. 邮件文本初步处理

但是，要让电脑“理解”电子邮件的文字，就必须把它转化成实数向量。词袋模型是一种传统的转换方法。该模型假定文档中每个单词的出现都是独立的，通过统计每个单词的出现频率作为分类器的特征，但是这种方法不考虑语法和单词出现的先后顺序。one-hot [13]将是否出现该词作为词的权重。但一般在表达中，词语之间是有一定的相似性的，而 one-hot 无法判断这种相似性，造成语义缺失。此外，现实生活中，词语的集合非常庞大，而每个向量的维度是和词语集合中词语的数量相匹配的，所以如果使用 one-hot 方法，需要用很高的维度来表示一个词。另外，若一个向量中仅有一个维度是非零的，该向量明显是过于稀疏的。

为解决 one-hot 模型的局限性，Word2Vec 模型[14]用一个固定长度的向量表示每个词，再用这些向量表达单词间的相似程度。该模型包括跳字模型(Skip-gram)和连续词袋模型(CBOW)两种学习方法。CBOW 假定在文本序列中的某个词是基于其前后的词产生的，Skip-gram 模型假定的是根据文本序列中的某个词来产生其前后的词[15]。

Skip-gram 模型具备训练时间短且效果好[4]的优点，故本文选择 Skip-gram 模型来构造 Word2Vec 模型，如下图 5 所示。

训练目标为：给定一个训练词序列 w_1, w_2, \dots, w_N ，最大化公式(1)中 L 的值：

$$L = \frac{1}{N} \sum_{n=1}^N \sum_{-c \leq i \leq c, i \neq 0} \log p(w_{n+i} | w_n) \quad (1)$$

其中, c 是前后文的词数, N 是训练样本的数量。

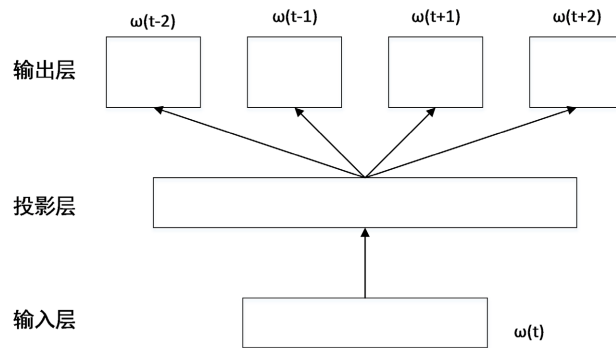


Figure 5. Skip-gram
图 5. 跳字模型

3.2. 卷积神经网络

在图像识别领域, 卷积神经网络(CNN)作为一种前馈神经网络, 有着良好的表现。由于其出色的提取局部特征的能力, CNN 也被广泛用于文本分类领域[16]。

CNN 的卷积层采用不同尺寸的卷积核, 得到了不同的列向量。池化层可以从卷积层中抽取最大列向量, 并能有效地克服因句子长度的差异而造成的差别。最后, *Softmax* 整合池化数据, 完成全连接层的操作, 以完成分类任务。在文本分类方面, CNN 能够很好地从文本中抽取出局部特征。本文中的 CNN 架构负责从文本数据中提取特征, 并整合分类, 得到最终结果。

3.3. BiGRU

循环神经网络是一种对序列数据进行处理的神经网络, 其内部结构设计使得网络可以捕捉序列之间的关系, 并提取上下文的关系。但在训练长序列数据时, 梯度消失、梯度爆发等问题时常出现[17]。

为了克服上述问题, Hochreiter 等人[18]基于 RNN 提出了长短期记忆(LSTM)网络这一特殊的变体; Cho 等人[19]又在其基础上提出了门控循环单元(GRU), 它们都通过引入门控机制来控制梯度的流动, 从而使得模型更容易训练和更好地捕捉序列之间的关系。GRU 与 LSTM 模型效果相似, 但是 GRU 计算复杂度相比 LSTM 要小, 内部结构也相对较简单[20]。为了增强语义关联并提高分类精度, 本文引入 BiGRU 双向门控循环单元模型, 相比于单向模型, 它更能有效地捕捉文本中的语义关系, 从而获得更高的分类精度; 并且在引入注意力机制后又引入了一个 BiGRU, 目的是为了更全面地获得文本全局特征。

双向 GRU 结构如图 6 所示, 模型在每个时刻都有前向和后向的 GRU 作为输入, 而输出则由这两个输入决定。这种结构可以更好地捕捉序列数据中的双向依赖关系, 从而提高模型的性能和准确率。

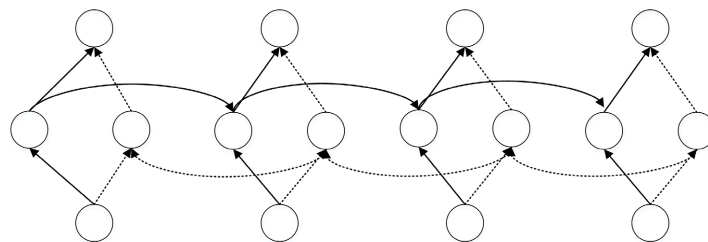


Figure 6. BiGRU
图 6. 双向门控循环单元

如上图所示, t 时刻的输入 x_t 以及 $(t-1)$ 时刻的前向和反向隐层状态的输出 \vec{h}_{t-1} 、 \overleftarrow{h}_{t-1} 共同决定了 BiGRU 在该时刻的隐层状态。如公式(2)中所示, 对当前时刻的前向和反向隐层状态 \vec{h}_t 、 \overleftarrow{h}_t 进行加权求和得到 BiGRU 在该时刻的隐层状态 h_t [21]。

$$\begin{aligned}\vec{h}_t &= GRU(x_t, \vec{h}_{t-1}) \\ \overleftarrow{h}_t &= GRU(x_t, \overleftarrow{h}_{t-1}) \\ h_t &= w_t \vec{h}_t + v_t \overleftarrow{h}_t + b_t\end{aligned}\quad (2)$$

3.4. 注意力机制

文本向量在经过第一层 BiGRU 层后, 邮件信息已经较全面地被提取, 但邮件中的需要重点关注的信息没有被突出强调。Attention 机制将文本数据转化为由一系列的<Key, Value>数据对组成的 Source 数据, 并基于 Query 与 Key 的相似度得出各个 Key 对应的不同 Value 的权值, 最后对每个 Value 进行加权求和得出 Attention 数值, 以实现的关键信息进行更精准的提取[22]。

注意力机制(Attention)可以从一堆复杂的输入信息中, 找到对当前输出更重要的部分。在文本分类中, 注意力机制被广泛应用于强化关键信息的权重, 从而提高分类效果。本文为突出邮件的关键信息, 引入了注意力机制, 其实现过程如下图 7 所示。Attention 机制将文本数据转化为由一系列的<Key, Value>数据对组成的 Source 数据, 并基于 Query 与 Key 的相似度得出各个 Key 对应的不同 Value 的权值, 最后对每个 Value 进行加权求和得出 Attention 数值, 以实现的关键信息进行更精准的提取[22]。

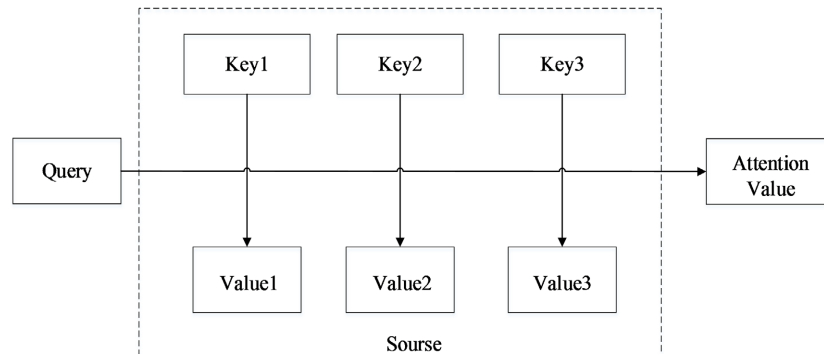


Figure 7. Implementation process of attention mechanism
图 7. 注意力机制实现过程

Attention 机制将文本数据转化为由一系列的<Key, Value>数据对组成的 Source 数据, 并基于 Query 与 Key 的相似度得出各个 Key 对应的不同 Value 的权值, 最后对每个 Value 进行加权求和得出 Attention 数值, 以实现的关键信息进行更精准的提取[22]。

4. 实验与结果

4.1. 实验环境与超参数设置

本文实验环境为 Intel(R) Core(TM) i7_9750 CPU@2.60 GHz 以及 16 GB 内存, 操作系统为 Windows 10, 显卡(GPU)为 RTX2080Ti。模型超参数的设置如表 1 所示。

Table 1. Hyperparameters of the model**表 1.** 模型超参数

模型超参数	参数数值
词向量维度	200
神经元个数	256
Epoches	10
Sequence Length	200
Dropout 值	0.5

4.2. 实验数据

本文选取了国际文本检索会议提供的公开垃圾邮件语料库 trec06c 作为实验数据集, trec06c 分为英文和中文两类邮件数据。正常邮件有 21766 条数据, 垃圾邮件 42854 数据。实验采用的验证方法为十折交叉验证, 即把 trec06c 数据集分成十份, 其中一份作为验证集, 另外九份则作为训练集, 循环进行 10 次, 以保证模型训练和评估的公正性和可靠性。

4.3. 评价指标

准确率(Accuracy)表示垃圾邮件检出率, 以此来衡量邮件分类模型分类的准确程度; 精确率(Precision)表示邮件识别的精确程度; 召回率(Recall)表示垃圾邮件检出率、F1 值是调和召回率与准确率的一个评价价值, 具体计算公式如下:

$$\text{Accuracy} = \frac{A + D}{N} \times 100\% \quad (3)$$

$$\text{Precision} = \frac{A}{A + B} \times 100\% \quad (4)$$

$$\text{Recall} = \frac{A}{A + C} \times 100\% \quad (5)$$

$$\text{F1} = \frac{2 \cdot \text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}} \quad (6)$$

其中, A 表示实际为垃圾邮件, 系统判定也为垃圾邮件的电子邮件数量; D 表示实际为正常电子邮件, 系统判定也为正常电子邮件的电子邮件数量; B 表示实际为正常电子邮件, 但系统判定为垃圾邮件的电子邮件数量; C 表示实际为垃圾邮件, 但系统判定为正常的电子邮件数量; N 为邮件总数。

4.4. 对比实验

为了测试本文的模型在垃圾邮件分类中的有效性, 本文使用 SVM 模型[23]、CNN 模型、CNN-BiGRU 模型[24]、LSTM-Attention-CNN 模型[25]、基于 CNN 的 BiGRU-Attention 模型进行了 5 个对比实验, 并使用了相同的数据集和实验环境, 以确保实验的公平性。

4.5. 实验结果

5 组对比实验结果如表 2 所示。

Table 2. Results of the experiment**表 2.** 实验结果

模型	准确率/%	精确率/%	召回率/%	F1 值/%
SVM	70.20	80.21	72.88	76.54
CNN	81.98	88.47	83.89	91.04
CNN-BiGRU	89.13	93.36	90.03	92.83
LSTM-Attention-CNN	91.34	95.23	91.56	93.04
CNN-BiGRU-Attention	91.56	95.52	93.92	95.8

从上面表格中可以看出, SVM 模型准确率最低只有 70.20%, 传统分类方法分类效果较差; 与 SVM 相比, CNN 模型准确率大幅上升, 数值提高至 81.98%。CNN 引入 BiGRU 之后, 准确率进一步提高至 89.13%; CNN 引入 LSTM 的 Attention 机制后, 准确率也大幅提升, 达到 91.34%; CNN 引入 BiGRU 的 Attention 机制后, 准确率达到 91.56%。虽然这两组模型的准确率差异并不大, 但在召回率、F1 值上, 基于 CNN 的 BiGRU-Attention 模型都高于基于 CNN 的 LSTM-Attention 模型, 并且基于 CNN 的 BiGRU-Attention 模型的训练时间也短得多, 即基于 CNN 的 BiGRU-Attention 模型的训练速度更快。

5. 结语

本文提出了基于深度学习的垃圾邮件检测技术的模型, 具体为 CNN-BiGRU-Attention 模型, 实验结果显示, 该模型在训练速度、准确性等方面比传统的文本分类模型以及其他一些垃圾邮件检测算法的效果更好。

在本次实验中, 模型训练与测试时, 仅仅截取了邮件数据集中的内容部分, 而没有把邮件的收件人和发件人地址等划入研究范围, 但是这些部分对垃圾邮件的检测也有较大意义, 具有研究价值。另外, 目前已经出现许多图片型垃圾邮件, 图片内容无用甚至非法, 仅仅依靠对内容进行文本分类的方法还不够全面。综上, 对垃圾邮件的检测和过滤的研究需要进一步改进、加强。

致 谢

在完成本篇论文的过程中, 我得到了来自各方面的支持和帮助。我想向所有为我提供帮助和支持的人表达我的感激之情。

首先, 我要特别感谢我的导师禹素萍教授。她给了我在研究领域深入思考的机会, 并在我的学术研究中提供了无私的指导和帮助。她的耐心和专业知识, 使我能够克服在研究过程中的难题, 同时对我的研究充满信心。我将永远感激她对我的帮助和支持。

其次, 我要感谢我的朋友们在我生活中的陪伴, 给予我鼓励和支持, 使我更加坚强和勇敢地面对生活中的挑战。没有他们的支持, 我将无法完成这篇论文。

最后, 我要感谢我的男朋友宗某。他是我在学习和生活中的坚实后盾, 他给了我无尽的爱和支持, 鼓励我克服所有的困难和挑战。他的陪伴和支持是我完成本篇论文的最大动力。

在此, 我再次向所有为我提供帮助和支持的人表达我的感激之情。我将永远感激你们对我的关注和支持。

参考文献

- [1] 冯军军, 李力. 垃圾邮件检测与绕过技术的研究[J]. 电脑知识与技术, 2021, 17(7): 36-37.

- [2] 冯军军, 李力. 机器学习在垃圾邮件过滤中的实现[J]. 电脑知识与技术, 2021, 17(8): 154-155.
- [3] 吴宗卓. 基于图和 K 近邻的文本分类算法[J]. 微型电脑应用, 2021, 37(10): 46-49.
- [4] 黄鹤, 荆晓远, 董西伟, 等. 基于 Skip-Gram 的 CNNs 文本邮件分类模型[J]. 计算机技术与发展, 2019, 29(6): 143-147.
- [5] 彭毅, 姜昕宇. 基于 BERT_DPCNN 文本分类算法的垃圾邮件过滤系统[J]. 电脑知识与技术, 2022, 18(22): 66-69.
- [6] 吴贵珍. 基于词向量和 CNN-BIGRU 的情感分析研究[D]: [硕士学位论文]. 镇江: 江苏科技大学, 2022.
- [7] 顾孟钧, 冯文舟, 陈中兵. 不同长度下中文垃圾邮件分类模型的研究[J]. 工业信息安全, 2022(7): 28-35.
- [8] 贾澎涛, 孙炜. 基于深度学习的文本分类综述[J]. 计算机与现代化, 2021(7): 29-37.
- [9] Kim Y. (2014) Convolutional Neural Networks for Sentence Classification. *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, Doha, 25-29 October 2014, 1746-1751. <https://doi.org/10.3115/v1/D14-1181>
- [10] 邵珊珊, 王立非, 刘智洋. 基于大数据的中外标准中英文本情感分析[J]. 中国标准化, 2019(17): 62-67.
- [11] 徐博龙. 应用 Jieba 和 Wordcloud 库的词云设计与优化[J]. 福建电脑, 2019, 35(6): 25-28.
- [12] 朱昶胜, 康亮河, 冯文芳. 基于自适应鲸鱼优化算法结合 Elman 神经网络的股市收盘价预测算法[J]. 计算机应用, 2020, 40(5): 1501-1509.
- [13] Harris, S. and Harris, D. (2014) 6—Architecture. In: Harris, S.L. and Harris, D.M., Eds., *Digital Design and Computer Architecture*, 2nd Edition, Elsevier, Amsterdam, 294-369. <https://doi.org/10.1016/B978-0-12-394424-5.00006-9>
- [14] Mikolov, T., Sutskever, I., Chen, K., et al. (2013) Distributed Representations of Words and Phrases and Their Compositionality. *Proceedings of the 27th Annual Conference on Neural Information Processing Systems*, Lake Tahoe, 5-10 December 2013, 3136-3144.
- [15] 周航帆, 周莲英. 基于会话记录的 Word2Vec 音乐推荐算法研究[J]. 通信技术, 2019, 52(4): 850-857.
- [16] 薛兴荣, 靳其兵. 基于词典的文本极性计算及分类研究[J]. 网络安全技术与应用, 2020(4): 57-61.
- [17] 杨星鑫, 吕泽均. 基于 LSTM 的无人机轨迹识别技术研究[J]. 现代计算机, 2020(5): 18-22+25.
- [18] 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>
- [19] Cho, K., van Merriënboer, B., Gulcehre, C., et al. (2014) Learning Phrase Representations Using RNN Encoder-Decoder for Statistical Machine Translation. *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, Doha, 25-29 October 2014, 1724-1734. <https://doi.org/10.3115/v1/D14-1179>
- [20] 胡玉琦, 李婧, 常艳鹏, 等. 引入注意力机制的 BiGRU-CNN 情感分类模型[J]. 小型微型计算机系统, 2020, 41(8): 1602-1607.
- [21] 赵晋斌, 王凯, 李盼. 基于深度学习的多维度证据要素关联关系抽取研究[J]. 中国电子科学研究院学报, 2021, 16(12): 1251-1256.
- [22] 赵宇轩, 胡怀湘. 基于 BiGRU-Attention-CNN 模型的垃圾邮件检测方法[J]. 计算机与现代化, 2021(4): 122-126.
- [23] 徐娟, 卞良. 基于 SVM 的中文垃圾邮件预测系统研究[J]. 数字技术与应用, 2020, 38(1): 38-39.
- [24] 吴贵珍, 王芳, 黄树成. 基于词向量与 CNN-BIGRU 的情感分析研究[J]. 软件导刊, 2022, 21(8): 27-32.
- [25] 宋丹. 基于改进的卷积神经网络的垃圾邮件过滤方法[D]: [硕士学位论文]. 淮南: 安徽理工大学, 2021.