内容列表可在[ScienceDirect](https://www.sciencedirect.com)上找到

## 信息处理和管理

杂志主页: [www.elsevier.com/locate/infoproman](http://www.elsevier.com/locate/infoproman)

一种用于从文本中检测个性的

情感

感知的深度学习方法

任占成<sup>a</sup>, 沈强<sup>b</sup>, 刁晓磊<sup>c</sup>, 徐浩<sup>a, b, \*</sup><sup>a</sup> 吉林大学人工智能学院, 长春130012, 中国<sup>b</sup> 吉林大学计算机科学与技术学院, 长春, 130012。<sup>c</sup> 特伦托大学信息工程和计算机科学系, 38123, 意大利我的朋友们, 你们知道吗?  
?

关键词: 性格检测, 深

度学习

BERT

多标签分类

## ABSTRACT

基于用户生成的文本内容分析的个性检测对信息科学有很大的影响, 例如, 信息搜索。然而, 现有的基于深度学习的方法有两个主要限制。首先, 他们只提取关键词进行个性检测, 缺乏对情感信息和心理语言学特征的分析。其次, 关于上下文和多义词的形成被忽略了。为了解决这些问题, 我们提出了一种基于神经网络的新型多标签个性检测模型, 它将情感和语义特征结合起来。具体来说, 我们利用变形器的双向编码Representation (BERT) 来生成句子级的嵌入, 用于文本语义提取。此外, 为了考虑情感信息, 情感字典被用于文本情感分析。最后, 我们将上述语义信息和情感信息输入神经网络, 以构建一个自动人格检测模型。该模型的性能已经在两个公共人格数据集上进行了评估。实验表明, 与最先进的技术相比, 我们在Myers-Briggs Type Indicator (MBTI) 和Big Five数据集上获得的平均准确率分别提高了6.91%和6.04%。

## 1. 简介

## 1.1. 背景介绍

人格特征被描述为个体相当固定的特征, 它表明个体的偏好并可能影响个体的决策, 并被纳入信息寻求、网络安全、临床心理学、经济学和政治理论等的方法论中 (Al-Samarraie, Eldenfria & Dawoud, 2017; Russell, Weems, Ahmed & Richard, 2017)。人类的人格特征已被证明会影响生活中的一些基本行为, 如知识处理、情绪调节和信息寻求 (Taramigkou, Apostolou & Mentzas, 2018)。通过社交媒体文本分析了解用户的个性特征, 可以被认为是一项信息分类或信息处理的任务。自动人格分类系统可以有效地帮助专业心理学家快速做出判断和决定, 并帮助企业选择更合适的职位候选人。

\* 通讯作者。

电子邮件地址: [xuhao@jlu.edu.cn](mailto:xuhao@jlu.edu.cn) (H. Xu).

2020年9月12日收到；2021年1月10日收到修订版；2021年1月22日接受  
可于2021年2月2日上网

0306-4573/© 2021 Elsevier Ltd.。保留所有权利。

现代特质理论 (John, Robins & Pervin, 2010) 试图通过设定几个分类维度并构建问卷来测量人格 (Matthews, Deary & Whiteman, 1998)。一个例子是著名的大五人格 (Digman, 1990), 包括经验开放性、自觉性、外向性、神经质和合群性。迈尔斯-布里格斯类型指标 (MBTI) (Briggs Myers & Kirby, 2000) 将人格特征分为四个二元维度: 外向与内向, 感觉与直觉, 思考与感觉, 以及判断与感知。然而, 传统的问卷调查具有时效性低、成本高、难以适应时代发展的特点, 这就导致了现代人格特质检测方法的出现。在过去的几十年里, 人格特质已经被证明与用户的社交媒体行为密切相关 (Azucar, Marengo & Settanni, 2018)。许多算法使用社交媒体信息来预测人格特质。主要有两种类型的人格特征检测方法。第一种是利用词的频率和词的类别对文本进行量化, 并结合传统的机器学习方法 (支持向量机和Naive Bayes) 进行人格分析。代表性的方法是语言学调查和单词计数 (LIWC) (Pennebaker, Francis & Booth, 2001)。第二种是利用深度神经网络提取文本内涵, 用于人格特质检测。代表性的方法包括卷积神经网络 (CNN) (Fukushima, Miyake & Ito, 1988)、循环神经网络 (RNN) (Elman, 1990)。

传统的机器学习和深度学习技术可以有效地检测出用户的个性特征, 而这些方法证明, 基于语言的模型可以准确有效地识别用户的个性特征。然而, 仍然存在一些挑战。1) 为了深入实现对文本语义的理解, 大多数现有的文本编码模型都是基于词向量的静态表示, 例如Word2Vec (Mikolov, Chen, Corrado & Dean, 2013) 和GloVe (Pennington, Socher & Manning, 2014), 它们不能解决诸如一个词的多义性等问题。Boxman-Shabtai等人证明了文本中一个词的多义性可能导致一些模糊的个性特征 (Boxman-Shabtai & Shifman, 2014)。2) 在这一领域出现的一个挑战性问题是, 目前的人格检测方法远非最优。因此, 本文的目的是要回答以下问题。如何才能更准确地从在线开放源码文本中进行自动人格检测? 这主要包括以下几点。1) 社交媒体文本中语义和情感信息的挖掘。2) 神经网络模型的设计。3) 情感信息和语义嵌入的结合。

## 1.2. 本文的动机和贡献

本研究的主要目的是提出一种基于社交媒体文本信息的新型人格检测模型, 准确理解文本的文字信息。我们使用双向编码器表示法 (Bidirectional Encoder Representation From Transformers, BERT) (Devlin, Chang, Lee & Toutanova, 2018) 对文本矢量进行动态编码。利用多头关注机制, 文本中的每个词都能充分考虑到其他位置的词向量信息, 从而产生语义文本编码。此外, 一些研究发现, 人格与情感之间存在着联系 (S. Zhao等, 2019)。该模型应考虑某些文本情感特征, 这些特征在文本编码中是缺失的。我们尝试结合外部情感词典, 让文本信息承载一定的情感信息, 从而得到更准确的人格检测模型。为了验证我们工作的有效性, 我们利用了两个公开的数据集, 即来自Twitter推文的MBTI数据集和来自论文的大五数据集, 每条推文或文本都被贴上了心理学家注释的人格类型。

本文提出了一个使用多文本语义结合情感特征进行多标签分类的人格检测模型。首先, 与之前的word2Vec和GloVe相比, 我们使用BERT来嵌入文本信息。BERT可以通过微调的BERT模型有效生成文本的向量表示。我们特别将多个文本向量连接起来, 基于句子层面的表征。句子层面的表征是通过自我注意机制和加权求和令牌嵌入产生的。其次, SenticNet<sup>51</sup>情感字典 (Cambria, Poria, Hazarika & Kwok, 2018) 被用来提取句子的情感极性, 然后将情感极性映射到向量空间, 最后与文本语义向量相结合, 作为神经网络的输入。第三, 与传统的基于单文本的人格检测模型相比, 我们将之前的多个文本向量与情感向量的连接串联起来, 输入到不同的神经网络中进行训练和学习。使用公共人格数据集进行验证, 结果表明我们的人格检测模型可以有效提高人格检测的准确性。

总之, 我们通过分别使用BERT模型和SentiNet5字典来进行句子层面的语义和情感特征的提取。然后, 我们将这些特征应用于各种神经网络架构 (GRU、LSTM、CNN) 以进一步提取特征和分类。本文的主要贡献如下。

- 1) 我们提出了一个新颖的多标签个性检测模型, 它结合了预先训练好的BERT模型和神经网络, 可以更好地理解句子语义, 处理社交媒体文本数据中的信息。
- 2) 我们提出了一种结合语义和情感特征的人格检测方法, 这种方法增加了一些解释人格的能力, 有助于对人格特征的分析。
- 3) 在MBTI和Big Five数据集上的实验结果表明, 我们的模型在人格检测方面优于最先进的技术。我们的模型可以从社交媒体文本中自动进行性格检测, 帮助各种社交软件进行用户信息挖掘。此外, 我们的模型也可以通过一些简单的修改应用于其他人格分类 (如Cattell的16种人格因素) 任务。

<sup>51</sup> <http://www.sentic.net>

本文的其余部分组织如下。在下一节中，第2节介绍了人格检测的相关工作。第3节提出了我们的个性检测方法。在第四节中，我们介绍了两个用于个性检测的数据集，并进行了实验和评估模型的性能。第5节讨论了本研究和未来研究的前景。最后，第6节总结了本研究。

## 2. 相关工作

本文的相关工作包括两个方面。第一是使用心理学词典的人格检测，第二是使用神经网络的人格检测。

### 2.1. 利用心理学词典进行人格检测

最初的方法是使用关键词的语言特征构建的人格特质检测模型。在过去的十年中，许多研究将语言的使用与广泛的心理学相关因素联系起来。例如，Park等人（Park et al., 2015）建立了一个基于用户语言的人格预测模型，这表明基于语言的评估可以成为一个有效的人格测量。早期的研究者Pennebaker和King发现写作风格（如词频和语篇）和人格之间有可靠的关联（Pennebaker & King, 1999），然后他们提出了语言调查和词数（LIWC）（Pennebaker等人，2001）方法。它首先将单词分为各种与心理有关的类别，统计每个类别中的单词频率，然后借助一些机器学习方法预测文本的个性特征。之后，LIWC被用来构建人格检测的网络应用和应用编程接口（API），以进行商业活动（Golbeck, 2016；Sewwandi等人，2017）。他们提出的技术可以快速检测出一个人的个性，而且准确率很高。由于互联网的快速发展，从社交媒体文本中进行人格检测已经成为最新的趋势。例如，从社交媒体的语义和状态更新中进行个性检测（Akhtar等人，2018；Panicheva, Ledovaya & Bogolyubova, 2016），这可以帮助企业选择合适的员工（Gaddis & Foster, 2015）。同样地，Yin等人研究了负面文本信息和大五人格特征之间的关系（C. Yin, Zhang & Liu, 2020）。他们发现，用户与微博的某些人格特征的互动有助于预测负面微博信息的传播。以上所有的研究表明，文本特征和人格特质之间存在着联系。然而，由于社交媒体平台使用的语言具有动态的抽象性，对结果的解释也需要有良好的心理学语言背景。因此，构建一个基于文本的自动人格检测模型是具有挑战性的。

2013年，Poria等人发现，使用带有情感和情绪信息的常识性知识，可以提高人们的认知。现有框架的有效性（Poria, Gelbukh, Agarwal, Cambria & Howard, 2013）。他们使用了SenticNet（Cambria等人，2018），这是一个流行的工具，用于从文本中提取常识性知识以及相关的情绪极性和情感标签。这些特征被输入到一个分类器中进行个性检测。同时，SenticNet提供了四个情感维度（愉悦度、注意力、敏感度和能力）的情感值以及从1到1的情感极性值（其中1为极度负面，1为极度正面）。Darliansyah等人构建了一个基于情绪的人格检测系统。他们的结果表明，具有相同个性特征的人倾向于以相同的方式表达情绪，这证明了个性和情绪之间的相关性（Darliansyah, Naeem, Mirza & Pears, 2019）。情绪作为一个人在某个时期的状态，与人格特征的长期表达密切相关。因此，SenticNet提取的情感特征可以有效地用于人格检测的分析。

### 2.2. 使用神经网络的人格检测

最新的方法主要使用深度神经网络进行语义理解来分析人格特征。一些评论文章给出了该领域的详细总结（Mehta, Majumder, Gelbukh & Cambria, 2019; Remaida, Abdellaoui, Moumen & Idrissi, 2020）。

#### 2.2.1. 深度学习

2014年之后，端到端的深度神经网络架构和模型被开发出来，以提取特征的方法更具成本效益。Liu等人（Liu, Preotiuc-Pietro, Samani, Moghaddam & Ungar, 2016）使用基于Bi-RNN的词向量生成模型，生成词向量表征，输入到前馈神经网络并用于预测五大特征的特征。后来，Majumder等人（Majumder, Poria, Gelbukh & Cambria, 2017）使用单词嵌入技术和卷积神经网络，自动深入挖掘2467篇论文的文本语义。Rahman等人比较了卷积神经网络中不同激活函数在个性检测中的效果。他们发现，在从文本中进行个性检测时，tanh的整体性能优于sigmoid和leaky ReLU（Rahman, Al Faisal, Khanam, Amjad & Siddik, 2019）。人格特征是通过CNN检测的，但这种方法没有考虑文本信息的连续性。Hernandez（Hernandez & Knight, 2017）和Scott试图使用长短时记忆（LSTM）（Hochreiter & Schmidhuber, 1997）。发现LSTM与普通的RNN、GRU和bi-LSTM相比，效果更好。Sun等人（Sun, Liu, Cao, Luo & Shen, 2018; J.-h. Zhao, Zeng, Xiao, Che & Wang, 2020）引入了基于紧密联系的句子的抽象特征组合（Celli, 2012），他们称之为潜伏句子组，

这种组合能有效提高准确率 ([Drexel, 2019](#))。然后, 作者使用双向LSTMs与CNN串联, 利用文本的结构检测用户的个性特征。

### 2.2.2. 词嵌入和转化器

词嵌入是将一些词转换为 $N$ 乘 $D$ 维矩阵的技术，其中 $N$ 是词的数量， $D$ 是维度输出。在早期的研究中，Word2Vec被用于单词嵌入。在Word2Vec算法中，首先，一个前馈神经网络被训练来预测给定的特定单词和它之前的上下文。然后，前馈神经网络内的隐藏层被用来将单词转换为相关的向量，每个单词用Word2Vec表示为一个固定长度的特征向量，而句子则表示为一个可变数量的单词向量。Majumder等人（Majumder et al., 2017）使用Word2Vec嵌入单词，然后应用深度CNN进行文档中的个性检测。Han等人（Han, Huang & Tang, 2020）也利用Word2Vec进行文本编码，然后提出一个可解释的人格模型。Word2Vec算法还被应用于训练多个大规模的词嵌入，主要集中在Twitter和Wikipedia等大型数据集上，并产生了像GloVe这样的预训练嵌入，这是一套由在线数据源上超过50亿篇文章训练而成的词向量。GloVe也被广泛用于单词表示。例如，Kumar等人使用GloVe单词嵌入技术来嵌入单词，然后提出了一个融合了语言特征的人格特征分类系统（Kumar & Gavrilova, 2019）。

最近的研究提出了一种新的单词嵌入方法：转化器。它没有使用递归神经网络，而是使用大量的多头关注机制（一种收集单词上下文的方法）和前馈神经网络来改善当前的序列到序列任务。2018年，谷歌提出了一种特殊类型的转化器：预训练的BERT模型（Devlin等人，2018）。BERT通过应用一些新颖的机制（如屏蔽LM、下句预测）来学习来自Transformers的双向编码器表示，最终生成更复杂的单词特征来比较单词语义。这种方法可以很好地将知识转化为矢量。BERT模型是一种语言表征模型，需要超大的数据、巨大的模型和巨大的计算开销来进行训练。2019年，Keh等人（Keh & Cheng, 2019）利用BERT模型建立了一个个性检测模型，进行词的嵌入，他们发现基于BERT模型的文本个性检测可以有效提高准确率。由于BERT学习的语境嵌入比传统的词嵌入具有理论和经验上的优势，因此我们在人格检测模型中使用BERT模型构建了一个新的句子级嵌入。

## 3. 建议的方法

在本节中，我们提出了一个基于预训练的BERT模型和神经网络的多标签个性检测模型，该模型结合了文本中的语义和情感特征。如图1所示，该模型由以下三个部分组成。

- (1) 句子嵌入。它采用了BERT模型进行句子级嵌入。
- (2) 情绪分析。我们使用SenticNet5字典进行特征提取。
- (3) 神经网络分类。用于分类的神经网络方法包括CNN和RNN。

### 3.1. 句子嵌入

为了学习更多与个性相关的特征，我们利用个性数据集来训练和微调适应该领域的BERT模型。然后，我们使用微调后的BERT模型将样本中的每个单词编码为矢量表示。在标记级编码表示之后，每个句子被表示为一个 $(N, 768)$ 向量表示，其中 $N$ 是句子中的标记数量。2019年，Keh等人使用单词级嵌入的BERT模型来训练个性检测模型，但这种方法忽略了连续社交媒体文本之间的联系。受此启发，为了将连续的数据文本关联起来，我们利用基于BERT模型的句子级嵌入编码的方法。

首先，我们重新考虑了人格数据集。我们发现，每个样本数据都有多条推特推文或句子，而目前的研究大多使用单条推特推文或句子进行人格检测。人格是一个长期稳定的特征，所以使用多个连续的Twitter推文或句子的文本进行人格检测是一个更好的选择。数据集集中的每个用户都有多个连续的Twitter推文或句子。因此，我们计划使用 $K$ （具体数值将

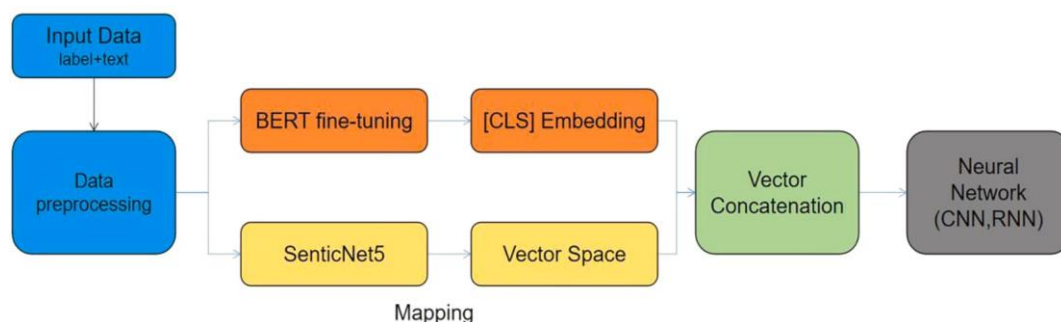


图1.实验流程图。

在随后的实验中得到验证)的推文或句子进行个性检测。然而, 预训练的BERT模型可以接受多达512个标记序列的文本内容, 因此, 使用这个模型, 来自K个连续推文的文本太大, 无法进行文本组合。为了解决这个问题, 我们从BERT模型的倒数第二层隐含层中提取一个名为CLS的向量来表示每条微博的文本。我们将多个文本的CLS向量连接起来, 而不是用标记嵌入的方式, 因为标记嵌入需要太多的空间。在训练期间, BERT模型在每个句子的开头添加一个CLS向量 $X^{R768}$ 。CLS的权重是通过使用自我注意将所有的标记嵌入相加来实现的。因此, 我们利用CLS向量作为可以代表推文文本的语义表示。第二, 我们提取句子序列级别的表示。假设一个样本有K个句子。如果有多于K句子, 取前K句子。如果少于K个句子, 768维的向量将被填充为零。最后的结果是 $X^{RM \times K \times 768}$ , 其中M代表样本的数量, K代表句子的最大长度, 768代表向量的尺寸。用于序列分类任务的BERT模型的主要结构由12个隐藏层组成, 其中有1个隐藏层。隐层大小为768。我们可以简单地理解为, 句子中的每个词都被转换为768维的向量表示。到目前为止, 我们已经将数据集转换为矢量表示, 下一步是训练模型。图2和图3。

### 3.2. 情绪分析

为了进一步挖掘和利用信息, 文本的情感知识是一个合适的选择, 由于目前的人格数据集没有任何情感标签, 所以很难使用BERT模型进行情感特征提取。传统的借助于外部词典的情感特征提取方法是一种有效的方法。LIWC在心理学上有很高的可信度, 但由于它不开放, 本文使用了另一种情感词典SenticNet5, 这是一种流行的工具, 用于从文本中提取常识性知识以及相关的情感极性和情感标签。因此, 我们准备通过结合文本的情感和语义特征来训练人格模型。具体来说, 我们使用SenticNet5情感词典来检测文本词。SenticNet5词典可以有效地提取某些词的情感倾向(积极/消极), 并给出相应的极性分数 $x1, 1$ 。如果一个词在情感词典中, 我们就对它进行情感分析。我们分别计算一个句子中的正负极性值, 并取其平均值来代表该句子的情感极性。首先, 我们将提取的情感分数作为一个一维的特征向量附加到之前的768维语义向量上。通过实验, 我们发现, 添加一维情感特征和不添加情感特征的结果基本相同。与768维语义特征相比, 一维情感特征对于神经网络来说太少了。我们决定将情感特征映射到更多的维度。于是我们尝试将情感极性得分离散化, 得分被分为从消极到积极的20个评级维度(得分每增加0.1就有一个新的情感等级), 然后将这20个等级的情感得分映射到20维的特征向量 $x^{R20}$ 。我们将每个情感等级映射成一个20维的向量, 只在其对应的位置取值为1, 其他位置取值为0, 然后将结果作为这个句子的情感极性的指示, 具体方式如公式。

(1)和公式(2)。

$$E_{class} = \frac{\lfloor \frac{E_{score} + 1}{0.1} \rfloor}{20} \quad E_{score} \in (-1, 1) \quad (1)$$

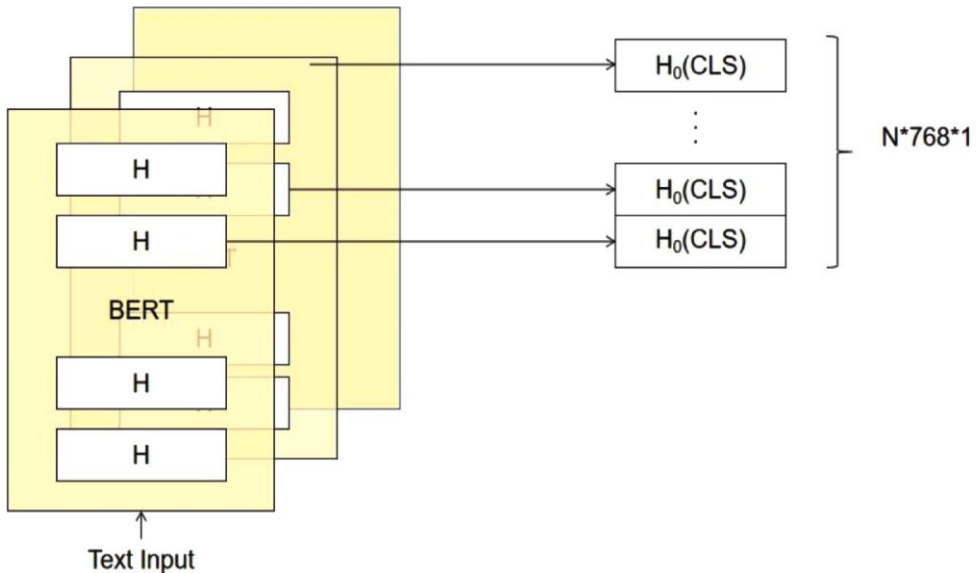




图2. 使用BERT提取文本语义CLS向量。

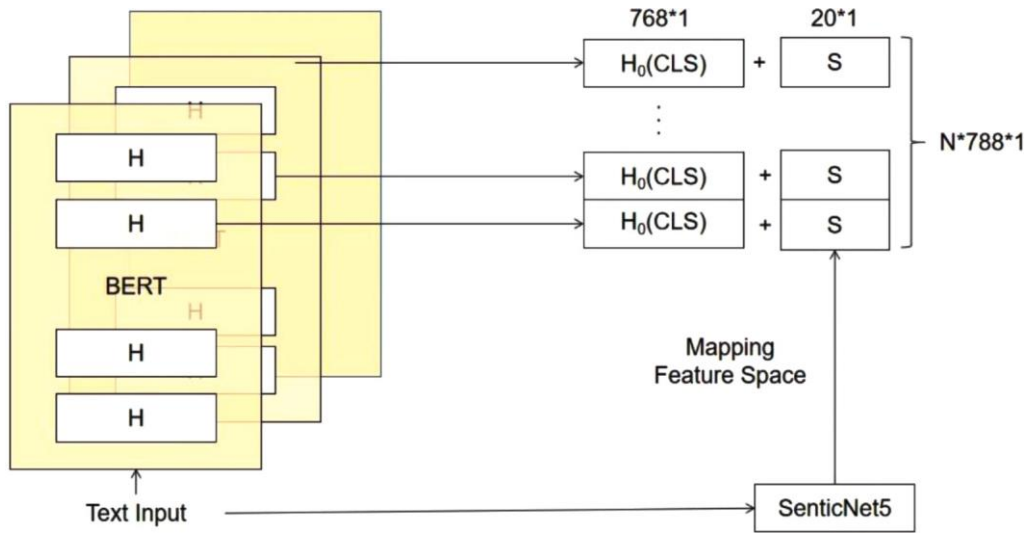


图3.用SenticNet5构建输入向量。

$$X_{E_{class}} = [x^1, x^2, \dots, x^{E_{class}}, \dots, x^{20}] \quad x^{E_{class}} = 1, \text{ else } x^i = 0 \quad (2)$$

其中,  $E_{score}$  是每个情感的得分,  $E_{class}$  是计算的情感水平,  $x_{E_{class}}$  是最终的情感特征表示。

之后, 每个句子的情感极性向量被附加到语义特征向量空间后面, 并输入到神经网络中。在接下来的部分, 我们将介绍我们的分类神经网络。

### 3.3. 神经网络分类

人格特征的检测可以被看作是一项文本分类任务。目前, 最有效的方法是深度学习技术。我们在分类任务中使用了三种不同的神经网络进行实验: CNN、GRU和LSTM。我们将文本样本转换为句子嵌入, 然后将其输入神经网络进行训练。

CNN是一种深度学习算法, 它使用前一层的局部区域视图来生成一个传输的中间表示, 用于网络训练。通过应用相关的过滤器, 可以更好地捕捉相应的输入特征, 涉及到参数数量的减少和权重的可重复使用。在文本的分类任务中, 上下文的顺序很重要, 所以我们也尝试使用LSTM和GRU。长短时记忆 (LSTM) 是一种时间递归神经网络, 它可以通过专门执行恒定的误差流来弥合超过1000个离散时间步长的最小时间延迟 (Hochreiter & Schmidhuber, 1997), 从而反映了时间序列的动态行为。与普通的RNN不同, LSTM是专门处理长期依赖关系的, 因为它采用了通过一系列门的“记忆”机制, 同时减少了梯度消失的问题。GRU是基于LSTM的一个变体, 其目的是在保留长序列信息的同时加快训练过程。

我们在此恢复将使用的三种类型的神经网络。

- 1) CNN是一个卷积神经网络, 可以有效地进行特征提取。我们使用了三个连续的卷积层进行特征提取。卷积核的大小为1 3、3 3和3 1, 然后是全连接层。输出单元的数量等于人格维度的数量, 例如, MBTI人格数据集有四个输出单元, 大五人格数据集有五个。
- 2) GRU本质上是RNN的一种类型。为了探索连续句子之间是否存在一些联系, 我们使用了GRU的神经网络结构, 它可以很好地捕捉连续的句子之间的关系。我们使用了两层的GRU堆叠网络结构, 其中所有隐藏层节点的大小为100, 然后是两个连续的全连接层进行训练, 输出单元的数量与上述相同。
- 3) 为了更好地了解文本序列是否会影响分类效果, 我们还使用LSTM进行验证。我们最终实验了一个具有128个隐藏层节点的LSTM结构作为LSTM神经网络, 并使用最后一个节点的输出连接全连接层进行多标签分类。具体的网络结构如图4所示。

我们直接采用多标签分类法来构建模型, 并将输出单元的数量设定为每一个子维度的数量。这三个网络都使用多标签分类器和二元交叉熵。

$y = labels$ 

(3)

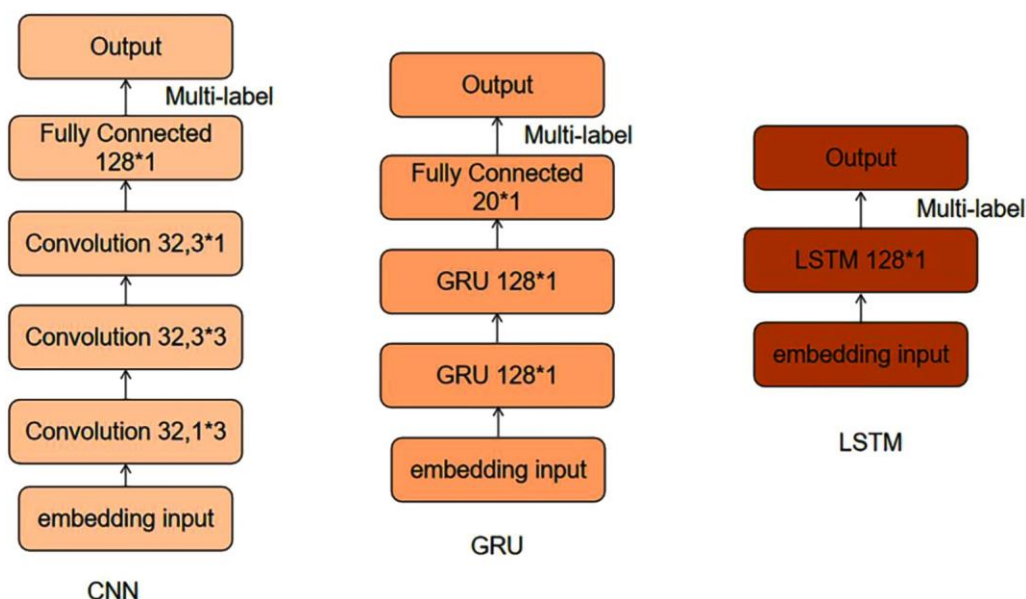


图4. 神经网络模型结构。

$$p_{ij} = \text{sigmoid}(\text{logits}_{ij}) = \frac{1}{1 + e^{-\text{logits}_{ij}}} \quad (4)$$

$$\text{损失}_{ij} = -y_{ij} * \ln p_{ij} + 1 - y_{ij} * \ln(1 - p_{ij}) \quad (5)$$

在最后的网络预测阶段，我们使用sigmoid函数将输出的标量数字转换为0,1的范围。如果它大于或小于一个概率阈值（通常是0.5），就被认为属于一个类别。

#### 4. 实验

在本节中，我们介绍了实验的细节。实验中使用了两个不同的文本人格数据集，并通过与最先进的模型进行比较给出了最终结果。

##### 4.1. 人格数据集

现有的带有标准人格标签的人格数据集很少，而且由于涉及到隐私问题，收集这些数据集很困难，此外，寻找专业心理学家做标签的成本也很高。Akrami等人证明了基于小型高可靠性数据集的模型比基于大型低可靠性数据集的模型表现更好（Akrami, Fernquist, Isbister, Kaati & Pelzer, 2019），所以我们决定使用一些目前比较常用的公共人格数据集。

我们在实验中使用了两个数据集，即MBTI数据集<sup>2</sup>和大五数据集，以检验我们的个性表现。检测模型。对于MBTI数据集，最近的一些研究仍在使用该数据（Hernandez & Knight, 2017）。不幸的是，大五人格领域最大的数据集MyPersonality在2018年因隐私泄露而无法使用。所以我们从Majumder等人的研究中提取了数据集（Kumar & Gavrilova, 2019; Majumder等人, 2017）。基本数据集信息见表1。

我们在此恢复将使用的两种类型的数据集。

- (1) MBTI数据集。MBTI性格类型数据集是一个标准数据集，它是目前最大的公开发表的性格数据集。Twitter的MBTI性格数据集，分为4个不同的维度，是通过Personality Café论坛收集的。该数据集包括来自8675名志愿者的50条推文和他们的性格标签。这给了我们422,845个总标签点（帖子文本，MBTI类型）。该数据集可在Kaggle网站上找到<sup>3</sup>，并获得许可已经为其在学术研究中的使用提供了支持。

---

<sup>2</sup> <https://www.kaggle.com/datasnaek/mbti-type>

<sup>3</sup> <https://www.kaggle.com>

表1

数据集介绍。			
	数据集Personality的	DatasetSource	大小尺寸 描述
	MBTITwitter8675×50	四维的。	每个项目由用户的50个连续的Twitter文本组成 每个项目包括多个
	(帖子文本, 类型	I-E, S-N, T-F, 和J-P	
	)	五维的。	
大	五行论文	EXT, NEU, AGR, CON, 和OPN	用户的句子
2467×50	(作文文本, 类型		
	)		

(2) 大五数据集。我们使用的大五数据集是James Pennebaker和Laura King的意识流文章数据集 (Pennebaker & King, 1999)。它包含了2468篇匿名文章的内容, 其中有作者标记的大五人格特征。EXT, NEU, AGR, CON, 和OPN。由于实验数据集中有一篇文章写着 "Err:508'", 所以实验是用剩下的2467篇数据进行的。

#### 4.2. 数据预处理

首先, 我们删除了数据文本中以 "http://" 或 "https://" 开头的URL链接, 然后使用NLTK<sup>4</sup>软件包删除了一些停止词。最后, 我们删除了一些非字母、数字或标点符号的特殊符号。然后, 我们对数据集进行了一些初步统计, 结果如图5和图6所示。在MBTI数据集中, 很明显EI和SN维度之间存在着严重的数据不平衡。因此, 我们通过降低采样率来平衡数据集中每个维度的样本。相反, 大五数据集中的类比MBTI数据集更加平衡。

在数据预处理之后, 我们的实验主要产生了两种可用的数据形式进行训练。第一种形式的数据是为了微调Bert模型的领域适应性。由于每个数据样本存在多条文本数据, 使得BERT模型难以训练, 我们将样本中的每条文本作为一个独立的数据单元 (文章中每条以句号、问号或感叹号结尾的推文或句子作为一个新样本), 数据标签保持不变。第二种数据形式是将一个人的连续推文或文本 (具有相同的标签) 作为一个样本, 除了微调Bert模型外, 所有的实验都使用这种形式。

#### 4.3. 实验设置

##### 4.3.1. 伯特微调

我们的实验是基于BERT模型的。使用的预训练的BERT模型是BERT-BASE<sup>5</sup>模型 (网络层数 $L$  12, 隐藏层维度 $H$  768, 注意12, 参数总数超过110M)。在这里, 我们只是用我们的个性数据对模型进行微调。

我们使用上述第一种形式的数据 (每条推文或句子作为一个例子) 来微调BERT模型 然后, 我们将数据集按6:2:2的比例分为训练集、验证集和测试集。

之后, 我们将数据输入BERT模型, 该模型基于TensorFlow<sup>6</sup> 1.14框架, 训练参数如表2所示。

##### 4.3.2. 调整连续文本

在本节中, 我们需要确定连续文本的数量, 也就是第3.1节中提到的 $K$ 值。为了验证具体的 $K$ , 我们在两个数据集中进行了实验。在实验中,  $K$ 从10开始, 每次增加10, 最大值为50 (在数据集中一个人的最大连续样本数为50)。 $K$ 个连续的句子由Bert模型编码。然后, 它们被输入到一个简单的神经网络 (全连接层分类输出) 进行多标签分类, 同时在神经网络中加入一个3×3卷积核以加速特征提取。因此, 一个BERT CNN模型被用来进行调整连续文本的实验。结果显示在图7中。

在MBTI数据集中, 最高的准确率是在 $K$ 20和30之间实现的。之后, 我们对 $K$  25进行了实验, 结果如表3所示。因此, 我们在这个数据集上采取了 $K$  25。

在五大数据集中, 从图中可以看出, 随着 $K$ 值的增加, 准确性不断提高。因此, 我们在这个数据集上取 $K$ =50。

##### 4.3.3. 皮尔逊相关系数

为了研究人格维度的相关性, 我们计算了实验数据中各人格维度的皮尔逊相关系数。皮尔逊相关系数是一种用来衡量两个变量之间相关程度的统计方法。当皮尔逊相关系数在0.3到0.3之间时 (Mukaka, 2012), 我们一般可以认为它是不相关的。在表4和表5中

2.  $R^2$  中，我们可以清楚地看到，人格维度之间没有相关性（最高为0.16和人格维度之间为-0.16）。

---

<sup>4</sup> <http://www.nltk.org>

<sup>5</sup> [https://storage.googleapis.com/bert\\_models/2020\\_02\\_20/uncased\\_L-12\\_H-768\\_A-12.zip](https://storage.googleapis.com/bert_models/2020_02_20/uncased_L-12_H-768_A-12.zip)

<sup>6</sup> <https://tensorflow.google.cn/>

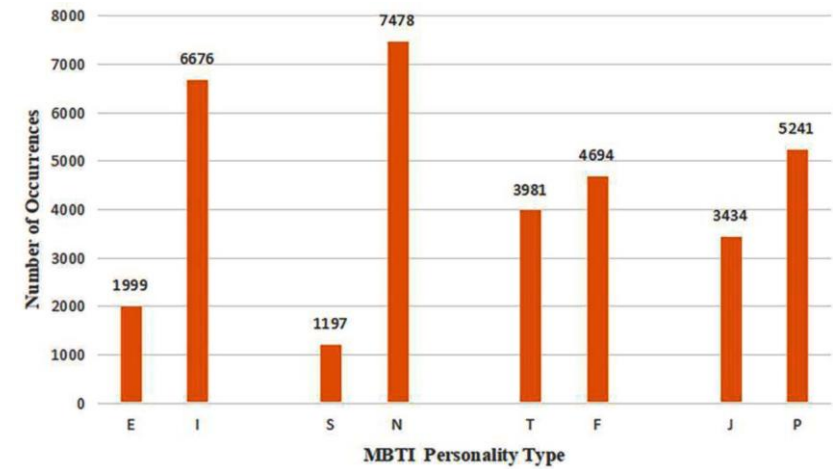


图5. MBTI性格类型的统计图。

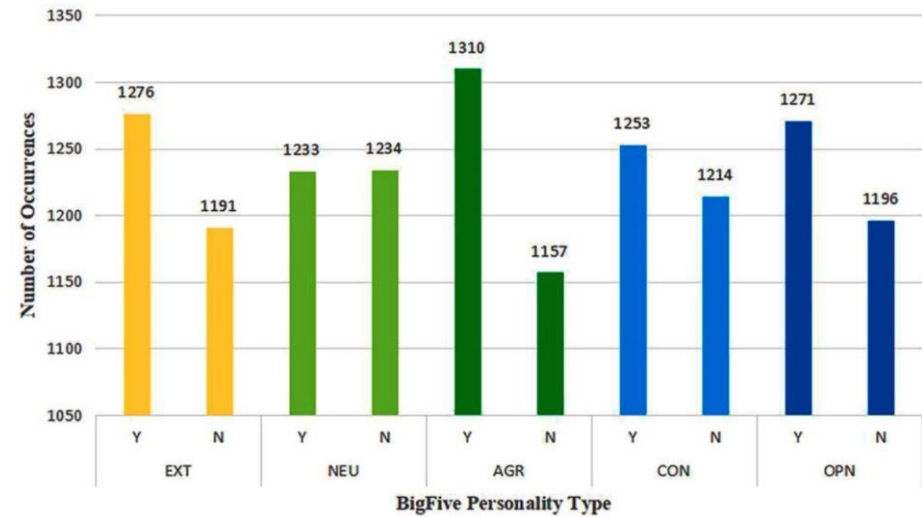


图6.大五人格类型的统计图。

表2  
微调贝尔特模型的训练参数。

数据集	学习率	批量大小	训练纪元	培训步骤
MBTI	0.00002	16	10	93,404
五巨头	0.00002	16	10	26,557

4.3.4. 培训

在个性检测领域，交叉验证在实验中普遍使用。交叉验证（k 4）被用来衡量我们分类器模型的性能。训练参数如表6所示。

4.4. 与最先进的技术比较

我们使用以下两个不同的数据集进行实验，并分别比较不同数据集的结果。

1) MBTI数据集。我们将我们的方法与以下最先进的技术进行比较。(a) 我们使用Keh等人在2019年提出的模型，这是目前唯一使用BERT模型进行人格检测的实验。(b) 关于最近的基于语言特征的分类技术，2019年，Pavan Kumar K. N等人提出了一个整合了计数



矢量化（例如通过TF-IDF）和GloVe单词嵌入的个性检测系统，这一点正在被使用。

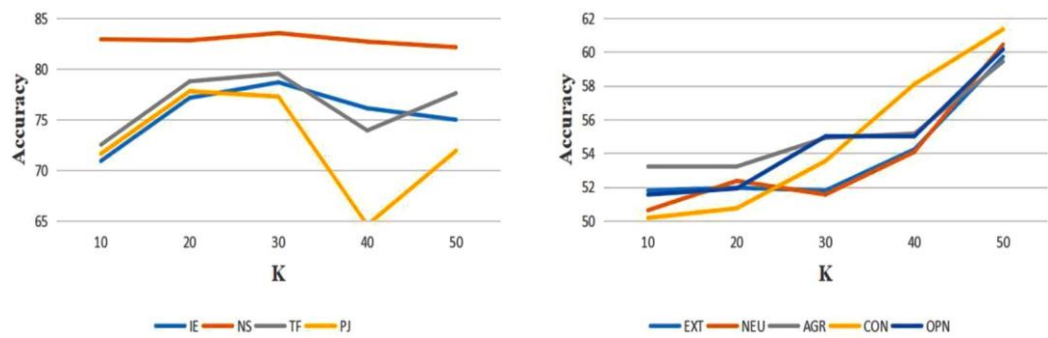


图7. MBTI数据集和大五数据集的不同K。

表3  
MBTI数据集中的不同K。

	IE	ÄÄÄ	TF	PJ
K = 20	77.13	82.81	78.76	77.79
K = 25	78.84	83.90	78.50	77.88
K = 30	78.67	83.52	79.52	77.26

表4  
MBTI-Pearson系数。

	IE	ÄÄÄ	TF	蒋介石
IE	1	-0.046	-0.070	0.160
		NS1	-0.081	0.015
			TF1	-0.00447
斐济				1

表5  
大五-佩尔森系数。

	EXTNEUAGRCONOPN	EXT1		-
0.160.120.13	0.079			
		NEU1	-0.089	-0.047
			-0.148	0.018
			AGR10.134	-0.027
			CON1	1
OPN				

表6  
分类神经网络的训练参数。

神经网络学习	率批量	大小训练纪元
	CCN/GRU/LSTM0.00132	20

表7  
BERT+CNN结构的MBTI中的单标签和多标签结果。

性格特征	E-I	S-N	T-F	J-P
BERT (单标签)	0.7262	0.7433	0.7502	0.7039
(K = 1)				
BERT (多标签)	0.7379	0.7360	0.7706	0.7172
(K = 1)				
BERT (单标签)	0.7842	0.7854	0.7750	0.7135
(K = 25)				
BERT (多标签)	0.7884	10 0.8390	0.7217	0.7203

( $K = 25$ )

---

2) 五大数据集。由于最近最大的MyPersonality数据集由于隐私泄露而无法使用，基于这个数据集的一些研究成果不能作为比较的基础。因此，我们使用以下模型。(使用Majumder等人在2017年提出的深度卷积神经网络的人格分类技术。

为了评估最终的分类结果，我们使用了准确率。准确率在每个人格维度的人格特征检测中是最常用的，我们也用它和最终结果一样的方式进行相关比较。

从表7和表8可以看出，与使用单标签的BERT模型进行分类检测相比，使用多标签的BERT对结果几乎没有影响。但是，当K值不同时，我们可以看到准确率有了一定程度的提高，因此可以推断出连续文本对人格特征的分析是有帮助的。此外，与传统的多二进制分类模型相比，我们只需要训练一个使用多标签分类的人格检测模型。表9和表10显示了我们的最终结果。与传统的神经网络相比，我们的模型从一开始就使用BERT模型来获得文本的嵌入向量，从而充分考虑了文本的上下文内容，与传统的编码相比取得了更好的结果。随后，我们尝试在BERT模型之后加入三种不同的神经网络进行更深层次的特征提取，结果如表9和表10所示，BERT + CNN获得的结果最好，比单独使用BERT要好。

带有神经网络的BERT可以有效提高个性检测的准确性。在三个神经网络中，CNN得到了最好的结果。

最后，在BERT + CNN的基础上，我们通过使用BERT模型的嵌入向量与SenticNet5的情感信息相结合，取得了更好的效果。将外部领域知识与文本语义相结合，可以提高对文本内容的理解，提高准确性。

该表显示了我们的结果与该领域其他主要工作中报告的各种性能指标的比较。与BERT模型相比，我们的模型在两个数据集上分别获得了6.91%和6.04%的平均精度提升。我们的研究表明，基于语义表征和词汇统计的个性特征检测模型的性能要优于传统的基于语言特征和神经网络的方法。

## 5. 讨论和未来工作

### 5.1. 结果分析

通过对结果的分析，表明我们的个性检测模型比最先进的模型效果更好。与传统的神经网络相比，我们的模型使用BERT模型来获得文本的嵌入向量。与丰富的BERT词嵌入相比，传统的编码方式显得苍白无力，BERT词嵌入是利用转化器、编码器和注意力机制的联盟来训练的。此外，为了提高模型的效率，我们将单词嵌入改为句子嵌入，以便获得更好的文本表示。同时，我们将分类器改为多标签分类器。与传统的多二元分类模型相比，我们可以利用多标签分类模型同时获得多个维度的个性检测结果。最后，我们将外部领域知识与文本语义相结合，通过将BERT模型的嵌入向量与SenticNet5的情感信息相结合，在模型的输入向量中携带了更多的个性相关信息，从而获得更好的结果。

在连续文本实验中，我们发现，文本的语义连接有助于获得个性的结果。在具体的 "K " 的分析中，我们发现在Twitter文本中，前25条推文会提高准确率，而之后添加更多的文本连接会降低准确率，这可能与文本的周期性有关。在五大数据集中，随着K值的增加，准确性也在提高。在论文数据集中，连续的语义表达了更清晰的含义，所以越多的语义连接会带来更好的结果。由此，我们可以得出结论，语义的连续性有助于提高分类结果的准确性。

在CNN和RNN中，哪种架构在文本分类任务中表现更好，取决于对整个序列的语义理解有多重要，对于情感分类这样的分类任务，选择CNN是有意义的，因为情感通常由一些关键短语决定 (W. Yin, Kann, Yu & Schütze, 2017)。在Bert模型和不同的神经网络架构的组合中，CNN的表现最好。在连续的语义连接中，整体语义信息更为重要。CNN算法擅长提取局部和位置不变的特征，而RNN通常擅长预测序列中接下来会发生什么，一些关键的语义信息对个性检测来说比时间序列更重要。CNN可以识别一些区域的主要特征，这使得它可以利用一些关键信息对任务进行分类。此外，由于我们结合了不同的特征（语义、情感），输入向量携带了丰富的信息。GRU和LSTM可以学习与时间序列相关的信息，但它们在提取一些关键特征方面不如CNN好。CNN可以有效地提取文本的核心特征，从而获得更好的结果。

另一个有趣的成果是，MBTI总体上优于大五，这也被Celli等人证实 (Celli & Lepri, 2018)。在MBTI的结果中，EI和SN两个维度更加可靠。在大五人格数据结果中，各维度的结果差异不大。结果显示，我们的模型在MBTI数据集中的表现比大五人格的表现更好。

### 5.2. 对研究的影响

我们的模型在个性检测方面取得了重大成果。我们获得了目前研究的以下重要意义：(1) 在整体语义相关性较大的文本中，更多的连续文本连接可以带来有希望的结果。

(2) 词语嵌入和情感信息整合有助于人格检测的有效性。(3) 情感

表8

BERT+CNN结构的大五中的单标签和多标签结果。

EXTNEUAGRCONOPN				个性特征
BERT (单标签)				
(K = 1)		0.62900.73820.65290.69290.7131		
BERT (多标签)				
(K = 1)		0.69430.68230.69710.70440.6604		
BERT (单标签)				
(K = 50)	0.	71800.		70160.72680.
BERT (多标签) (K		77130.7016		
= 50)	0.	71100.		74310.72090.
		70730.7485		

表9

MBTI数据集结果 (K = 25)。

	个性特征E-I	S-N	T-F	J-P
BERT (2019年)	0.7583	0.7441	0.7575	0.7190
tf-idf+glove+svm (2019)	0.6830	0.8500	0.7730	0.7560
BERT (单标签)	0.7842	0.7854	0.7750	0.7135
BERT (多标签)	0.7884	0.8390	0.7217	0.7203
BERT+GRU (多标签)	0.7842	0.8853	0.5839	0.6163
BERT+LSTM (多标签)	0.7842	0.8853	0.6716	0.6328
BERT+CNN (多标签)	0.8175	0.9075	0.7931	0.7876
BERT+Sentic+CNN (多标签)	0.8146	0.9251	0.8357	0.8236

表10

大五数据集结果 (K = 50)。

性格特征	延伸	NEU	AGR	锥心之痛	OPN
CNN(2017)	0.5809	0.5938	0.5671	0.5730	0.6268
BERT (单标签)	0.7180	0.7016	0.7268	0.7713	0.7016
BERT (多标签)	0.7110	0.7431	0.7209	0.7073	0.7485
BERT+GRU (多标签)	0.5253	0.5258	0.5406	0.4983	0.4948
BERT+LSTM (多标签)	0.5290				
BERT+CNN (多标签)	0.7630	0.7601	0.52530.53990.53910.5730	0.7634	0.7682
BERT+Sentic+CNN (多标签)	0.7994	0.8014	0.8030	0.8023	0.8035

信息与人格特质密切相关，可以尝试用来对人格特质进行进一步的解释。我们使用有效的外部知识来补充信息，使我们的输入文本承载更多的信息。这是一个新的尝试，深度学习和情感或心理词汇的结合是一个值得探索的未来方向。(4) 我们的研究引入了一个基于BERT句子表示和情感检测的全局模型，它比普通的单词嵌入能够捕捉到更多的上下文元素，这应该是未来使用深度学习进行文本个性检测的基础。

这项研究对实践也有重要意义。社交媒体，如Facebook、Twitter和微博。已经成为传播信息的最快方式之一。对每个用户进行单独识别和定位是至关重要的。了解用户的personality特征可以应用于各种下游任务，如信息的准确推荐，信息的寻求等。

在几种研究中已经证明，人格特质会影响信息搜索行为。我们可以用我们的模型来进行人格检测，识别不同的人格特征，然后将其应用于人工智能系统的个性化搜索设置。我们还可以将人格特质与其他信息结合起来进行推荐，比如对具有相同人格特质和偏好的人进行聚类。因此，一个准确的人格检测模型是必要的。

### 5.3. 未来的研究

在我们的实验中，数据集比较小，未来大数据分析的应用是我们应该考虑的方向。很多因素都会影响到人格检测，我们只是用文字信息来分析人格。事实上，其他辅助信息如音频和视频会帮助我们做出更准确的预测。要设计一个完全准确的人格检测系统是很有挑战性的。在未来，多模态融合将成为这一领域的主导技术。我们将在后续的研究中结合社交媒体的图像信息进行进一步的人格分析。

## 6. 总结

在本文中，我们提出了一个基于情感和语义特征的新型人格检测模型，该模型只需使用少量的社交媒体文本就可以进行accurate人格检测。我们首先利用预训练的BERT模型，这是NLP领域的一个前沿技术方法，用于句子级的文本语义提取。然后，我们使用情感字典，它增加了一些解释个性的能力，来分析每个文本的情感，并将其映射到向量空间。最后，我们将一个用户的多个文本内容结合起来进行组合分析。这两个向量，即情感向量和语义向量，被输入到一个神经网络分类模型中，使用多标签分类器计算结果。我们的模型在MBTI和大五数据集上得到了验证，可以产生卓越的人格检测结果。因此，使用NLP技术提取语义特征与传统的心理学词典（如情感分析）相结合是很有前途的。

## 竞争性利益声明

作者证明，本稿件所讨论的主题没有利益冲突。

## CRedit作者贡献声明

任占成：概念化，方法学，软件，数据整理，写作--原稿，写作--审查和编辑。

沈强：视觉化，调查。刁晓磊：软件，验证。徐浩：监督。

## 鸣谢

该研究得到了国家自然科学基金（62077027）、中华人民共和国科学技术部（2018YFC2002500）、吉林省发展和改革委员会（2019C053-1）的支持。中国吉林省教育厅（JJKH20200993K），中国吉林省科学技术厅（2020000801002GH），欧盟地平线2020 FET Proactive项目“WeNet-The Internet of us”（No 823783）。感谢吉林大学哲学社会科学国际创新团队的一些支持。

## 参考文献

- Akhtar, R., Winsborough, D., Ort, U., Johnson, A., Chamorro-Premuzic, T. J. P., & Differences, I. (2018). 使用社交媒体状态更新检测人格的黑暗面。 *Personality and Individual Differences*, 132, 90-97.
- Akrami, N., Fernquist, J., Isbister, T., Kaati, L., & Pelzer, B. (2019).从文本中自动提取个性。挑战与机遇。在2019年IEEE大数据（Big Data）国际会议上发表的论文。
- Al-Samarraie, H., Eldenfria, A., & Dawoud, H. (2017). 人格特征对用户信息搜寻行为的影响。 *信息处理与管理*, 53 (1), 237-247.
- Azucar, D., Marengo, D., & Settanni, M. (2018). 从社交媒体上的数字足迹预测五大人格特质。A meta-analysis. *Personality and Individual Differences*, 124, 150-159.
- Boxman-Shabtai, L., & Shifman, L. (2014).回避的目标。破译中介幽默中的多义词。 *通信杂志*, 64, 977-998.
- Briggs Myers, I., & Kirby, L. K. (2000).类型介绍。了解你的迈尔斯-布里格斯类型指标结果的指南。欧洲英文版本。Cambria, E., Poria, S., Hazarika, D., & Kwok, K. (2018).论文发表在第三十二届AAAI人工智能会议上。Paper presented at the.Celli, F. (2012).In Paper presented at the Proc. of sixth international conference on digital society.
- Celli, F., & Lepri, B. (2018).Is big five better than MBTI? a personality computing challenge using twitter data.在CLiC-it会议上发表的论文。
- Darliansyah, A., Naeem, M. A., Mirza, F., & Pears, R. (2019).SENTIPEDE: A Smart System for Sentiment-based Personality Detection from Short Texts. *J. UCS*, 25(10), 1323-1352.
- Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2018).Bert : 用于语言理解的深度双向变换器的预训练。Digman, J. M. (1990).人格结构。五因素模型的出现。 *Annual Review of Psychology*, 50(1), 417-440.
- Drexel, I. B. (2019).载于国际信息管理会议上的论文。Elman, J. L. (1990).Finding Structure in Time. *Cognitive Ence*, 14(2), 179-211.
- Fukushima, K., Miyake, S., & Ito, T. (1988).Neocognitron:一个用于视觉模式识别机制的自组织神经网络模型。 *IEEE Transactions on Systems Man and Cybernetics*, SMC-13 (5), 826-834.
- Gaddis, B. H., & Foster, J. L. (2015).全球领导者的阴暗面人格特征和关键工作行为的元分析。研究结果和对领导力发展和高管教练的影响。 *应用心理学*, 64 (1), 25-54.
- Golbeck, J. A. (2016).从社交媒体文本预测人格。 *AIS Transactions on Replication Research*, 2 (1), 2.
- Han, S., Huang, H., & Tang, Y. (2020).词语的知识。一种可解释的社交媒体个性识别方法。 *Knowledge-Based Systems*, Article 105550.
- Hernandez, R., & Knight, I. (2017).在第31届神经信息处理系统会议的论文中。Hochreiter, S., & Schmidhuber, J. (1997).长短期记忆。 *Neural Computation*, 9(8), 1735-1780.
- John, O. P., Robins, R. W., & Pervin, L. A. (2010).人格手册》。理论与研究》。Guilford出版社。
- Keh, S.S., & Cheng, I.T. (2019).使用预先训练的语言模型进行Myers-Briggs人格分类和特定人格的语言生成。arXiv preprint arXiv.06333.
- Kumar, K. N. P., & Gavrilova, M. L. (2019).在2019年第16届IEEE高级视频和信号监控国际会议（AVSS）上发表的论文。
- Liu, L., Preotiu-Pietro, D., Samani, Z. R., Moghaddam, M. E., & Ungar, L. H. (2016).通过社交媒体上的图片选择来分析个性。论文发表于 the ICWSM。

Majumder, N., Poria, S., Gelbukh, A., & Cambria, E. (2017). 基于深度学习的文档建模，用于从文本中检测个性。 *IEEE 智能系统*, 32 (2), 74-79.

Matthews, G., Deary, I., & Whiteman, M. (1998). *人格特质*. 英国剑桥：剑桥大学出版社。

Mehta, Y., Majumder, N., Gelbukh, A., & Cambria, E. (2019). Recent Trends in Deep Learning Based Personality Detection. *Artificial Intelligence Review*. Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2013). Efficient estimation of word representations in vector space. *Computer ence*.



- Mukaka, M. (2012).统计学角。在医学研究中适当使用相关系数的指南。《马拉维医学杂志：马拉维医学协会杂志》，24（3），69-71。
- Panicheva, P., Ledovaya, Y., & Bogolyubova, O. (2016).俄语脸书文本中黑暗三合会人格特征的词汇、形态和语义相关因素。论文，在2016年IEEE人工智能和自然语言会议（AINL）。
- Park, G., Schwartz, H. A., Eichstaedt, J. C., Kern, M. L., Kosinski, M., & Stillwell, D. J. (2015)。通过社交媒体语言进行自动人格评估。《人格社会心理学杂志》，108（6），934。
- Pennebaker, J. W., Francis, M. E., & Booth, R. J. (2001).语言学调查和字数。LIWC 2001.Mahway:Lawrence Erlbaum Associates, 71（2001），2001。
- Pennebaker, J. W., & King, L. A. (1999).语言风格。语言使用是一种个体差异。《人格社会心理学杂志》，77（6），1296。
- Pennington, J., Socher, R., & Manning, C. D. (2014).在2014年自然语言处理的经验方法会议上发表的论文（EMNLP）。
- Poria, S., Gelbukh, A., Agarwal, B., Cambria, E., & Howard, N. (2013).在墨西哥国际人工智能会议上发表的论文。Rahman, M. A., Al Faisal, A., Khanam, T., Amjad, M., & Siddik, M. S. (2019)。In Paper presented at the 2019 1st International Conference on Advances in Science, Engineering and Robotics (ICASERT)。
- Remaida, A., Abdellaoui, B., Moumen, A., & Idrissi, Y. E. B. E. (2020)。在2020年第一届应用科学、工程和技术创新研究国际会议（IRASET）上的论文。
- Russell, J. D., Weems, C. F., Ahmed, I., & Richard, G. G. (2017)。自我报告的安全和不安全的网络行为。因素结构和与人格的关联因素。《网络安全技术杂志》，1-12。
- Sewwandi, D., Perera, K., Sandaruwan, S., Lakchani, O., Nugaliyadde, A., & Thelijagoda, S. (2017)。在2017年第六届全国技术和管理会议（NCTM）上发表的论文。
- Sun, X., Liu, B., Cao, J., Luo, J., & Shen, X. (2018)。在2018年IEEE国际通信会议（ICC）上发表的论文。
- Taramigkou, M., Apostolou, D., & Mentzas, G. (2018)。利用个性特征和互动环境的探索性搜索。《信息处理与管理》，54（4），609-629。
- Yin, C., Zhang, X., & Liu, L. (2020).微博上的负面信息转贴。人格特征是否重要？《信息处理与管理》，57(1), 1-18.Yin, W., Kann, K., Yu, M., & Schütze, H. (2017).用于自然语言处理的Cnn和Rnn的比较研究。arXiv预印本arXiv:01923。
- Zhao, J.-h., Zeng, D.-L., Xiao, Y., Che, L.-p., & Wang, M. (2020).基于话题偏好和情感分析的LSTM模型的用户个性预测。《Pattern Recognition Letters》，138, 397-402。
- Zhao, S., Gholaminejad, A., Ding, G., Gao, Y., Han, J., & Keutzer, K. (2019)。通过对生理信号的个性感知高阶学习实现个性化的情绪识别。《ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM)》，15(1), 1-18。