Applied Intelligence (2021) 51:7705–7717
https://doi.org/10.1007/s10489-021-02277-7

Semantic-enhanced sequential modeling for personality trait recognition from texts

Xia Xue[1,2] · Jun Feng[1,3] · Xia Sun[1]

Accepted: 15 February 2021 / Published online: 15 March 2021

Abstract

The automatic recognition of personality traits from texts has attracted significant attention. Existing studies typically combine linguistic feature engineering with traditional models, use five various neural networks to predict personality traits with multiple labels, and fail to achieve the best performance on each label. To this end, in this paper, we propose a novel semantic-enhanced personality recognition neural network (SEPRNN) model, which has a goal of avoiding dependence on feature engineering, allowing the same model to adapt to detecting five various personality traits with no modification to the model itself, and employing deep learning based methods and atomic features of texts to build vectorial word-level representation for personality trait recognition. Specifically, to precisely recognize multi-labeled personality traits, we first propose a word-level semantic representation for texts based on context learning. Then, a fully connected layer is used to obtain higher-level semantics of texts. Finally, the experimental results demonstrate that the proposed approach achieves significant performance improvement for multi-labeled personality traits compared with several baselines.

Keywords Personality trait recognition · Deep learning · Word-level representation

1 Introduction

Personality reflects an individual's particular behavior pattern. There are many ways to describe personality traits of a person, but Big Five model [9] have been widely used and accepted in the literature [55–57]. The Big Five model identified five personality traits, namely extroversion, neuroticism, agreeableness, conscientiousness, and openness, to describe the personality traits of human beings. These five traits are binary (yes/no) values. Extroversion (EXT) means that the individual is outgoing, talkative, assertive, and energetic, or shy and solitary. Neuroticism (NEU) indicates whether the person is sensitive and anxious or secure and confident. Agreeableness (AGR) indicates whether the person is trustworthy, friendly, and modest, or unreliable, antagonistic, and boastful. Conscientiousness (CON) refers to whether the person is efficient, organized, prudent, and self-disciplined, or inefficient, disorganized, careless, and sloppy. Openness (OPN) means the person is inventive and curious, or unimaginative and cautious.

The five personality traits play an important role in a person's social life. The personality of an individual has a significant impact on their job choices, preferences, and lives. Previous studies [1–3] have shown that there is a strong connection between a person's personality traits and the behavior observed from texts. Automatic personality trait recognition from texts has various important practical applications. In talent management, personality traits can affect one's fitness for a job. Personality traits are the main driving force of behavior and are related to work-related traits [4, 5]. Specific personality traits will affect job performance and even determine job success [6–8]. Also, it has been applied in a range of other tasks, such as social

✉ Jun Feng
fengjun@nwu.edu.cn

✉ Xia Sun
raindy@nwu.edu.cn

Xia Xue
xuexia201606@163.com

1 School of Information Science and Technology, Northwest University, Xi'an 710127, China

2 Maths and Information Technology School, Yuncheng University, Yuncheng 044000, China

3 State-Province Joint Engineering and Research Center of Advanced Networking and Intelligent Information Services, Northwest University, Xi'an, 710127, Shaanxi, China

network analysis [44], individual recommender systems [45, 46], word polarity detection [47].

With the rapid development of social media, the study of personality trait recognition has been of great interest to many researchers. In the literature, personality trait recognition has been well studied from different perspectives. Existing study approaches can be grouped into three categories, which are usually correlation analysis methods, traditional machine learning-based methods, and deep learning-based methods. The existing correlation analysis methods based on the significant correlation of personality traits and behavior will no longer suffice.

Previous research shows that machine learning methods are capable of making accurate personality recognition by employing digital records of human behavior [53]. Machine learning-based personality judgments are more accurate than human judgments [54]. A recent method relied on text mining techniques to identify Big Five personality traits automatically. Previous works showed that such method can provide convincible results on the task of personality trait recognition [2, 10, 46, 48–52]. Recently, several researchers [10–12] focused attention on predicting personality traits using various machine learning methods. These traditional methods first extract features from texts, and then feed them into different classifiers to detect personality traits. Although these studies have achieved good performance, one drawback of this human-designed features approach is that the performance depends mainly on the manual inspection of features from domain experts, meaning that tremendous effort can be spent on feature engineering.

With the resurgence of deep learning, some studies in this field turn to use deep learning approaches. The work of Majumder et al. [13] employed deep convolutional neural network (CNN) model concatenated 84 Mairesse features to recognize personality traits. It is worth noting that their work employed hand-crafted features, which largely depended on domain expertise. Another issue is that linguistic features like Mairesse make the adaptation to a deep learning model more time-consuming. Moreover, for five personality traits, they adopted five different neural network models. Sun et al. [15] exploited bidirectional long short-term memory (BiLSTM) and CNN model to detect personality traits. However, this work did not consider the polysemy phenomenon in textual data, which would influence the performance of this task to some degree. More importantly, previous approaches have not achieved the best performance on each label.

To this end, in this paper, we propose a novel semantic-enhanced personality recognition neural network (SEPRNN) model. As stated in Lai et al. [14], a recurrent structure can be used to capture contextual information. Specifically, we first use context learning method based on bidirectional GRU (BiGRU) to dynamically capture the contextual information for obtaining more precise word meanings of social texts. To further capture the higher semantics of texts, we exploit a fully connected layer to fuse the semantic information of personality traits. Then, we adopt binary cross-entropy loss function to recognize multi-labeled personality traits. Finally, we evaluate our SEPRNN model with several state-of-the-art baselines on personality recognition data. Extensive experimental results clearly validate the effectiveness of the multi-labeled personality trait recognition task. One advantage of our model is the lack of feature engineering which allows the same SEPRNN model to adapt to recognizing five various personality traits without modification to the model itself. Also, our proposed model does not need to combine linguistic features like Mairesse, which can only flexibly detect the personality traits using an atomic representation of textual data. Meanwhile, our SEPRNN model also effectively alleviates the influence of polysemy on multi-labeled personality trait recognition tasks.

The contributions in this paper can be summarized as follows:

1) Our model works well based purely on social texts without specific linguistic features created, and allows the same model to adapt to the prediction of five various personality traits with no modification to the model itself;
2) We propose a word-level representation method based on context learning to precisely obtain a word semantic in multi-labeled personality trait recognition tasks;
3) To the best of our knowledge, this is the first work to achieve the best performance on each label. Our experiments on two benchmarks validate that the proposed SEPRNN model provides a new level of state-of-the-art performance.

The rest of this paper is structured as follows. In Section 2, we discuss some related works of our study. In Section 3, the technical details of our Semantic-enhanced Personality Recognition Neura Network will be introduced. In Section 4, we comprehensively evaluate the performance of the proposed model. Finally, in Section 5, we conclude the paper.

2 Related work

The related works in this paper can be grouped into the following two categories, i.e., personality recognition and deep learning-based text mining.

2.1 Personality recognition

Recently, a rising trend of using advanced artificial intelligence (AI) technologies to tackle business issues has emerged [16, 18, 61]. Regarding the problem of personality recognition, how to correctly detect personality traits is always a critical concern in successful social living.

To comprehensively identify relevant literature of personality trait recognition, we collected research articles using a combination of keyword searches. With the help of Google scholar search engine, Baidu scholar search engine, and several scholarly online databases (ScienceDirect, Springer, IEEE Xplore, PNAS, etc.), we adopted several search keywords like “deep learning”, “neural network”, “machine learning”, “personality classification”, “personality recognition”, “personality prediction”, etc. We focus on computing methods and leave out psychological studies on personality trait recognition. In view of our focus on personality trait recognition from an intersection of method and domain, so we employed pairs of search terms such as “machine learning” and “personality recognition”, or “deep learning” and “personality prediction” respectively. Single methods such as “SVM” and “CNN” were also used as search terms. The search was further restricted to English publications. We found 137 publications by using this procedure. We only consider personality trait recognition from texts. The abstracts of these research literature must be relevant to both machine learning methods and personality trait recognition field. In doing so, 33 research articles were initially rated as related. The research articles were included only if the three researchers agreed on their topical relevance after reading the abstracts of these research literature. Therefore, 15 research contributions were finally analyzed in detail.

Previous studies can generally be grouped into three categories: correlation analysis, traditional machine learning-based approaches, and deep learning-based methods.

Some early research of the first category has explored correlation analysis methods. For instance, as one of the typical studies in this field, Ross et al. [20] found that there was a correlation between behavior and personality traits. Mohammad et al. [21] found that the use of emotional words is related to personality traits. Markovikj et al. [43] used Pearson correlation coefficient to analyze the correlation between each feature and personality trait.

Different from previous studies, some researchers mainly focused on using machine learning methods to detect personality, which are the mainstream methods. Many methods for detecting personality traits automatically have been designed [58]. An interesting work proposed by Yang et al. [46] is to use text mining technology to identify the game players’ five personality traits, then the games similar to the players’ personality traits were recommended to other players. In [19], Pennebaker et al. used linguistic features to predict personality. One notable work was done by Mairesse et al. [10], who used dictionary resources as features and employed several different machine learning algorithms to identify personality traits. Nguyen et al. [11] extracted emotional features and used SVM to predict personality traits. Tadesse et al. [12] first extracted features from social texts and then used ensemble learning to detect personality traits. To date, these researches have achieved remarkable results against existing work. However, traditional machine learning methods first need to extract features from social texts, and then use a combination of various classifiers and feature engineering to predict personality. The method of human-designed features mainly relies on the strong professional background knowledge of the researchers, which greatly limits the performance of machine learning algorithms. Compared with traditional machine learning algorithms, deep learning methods can automatically generate features suitable for text classification during training.

Recently, deep learning-based methods have attracted significant attention from several researchers. Wei et al. [22] adopted CNN with 1, 2, and 3-grams kernel to learn semantic structure in predicting user personality tasks. Yu et al. [23] used a CNN with average pooling neural network architecture to automatically detect personality from Facebook social networks. Majumder et al. [13] used deep three-dimensional CNN model concatenated 84 Mairesse features extracted from texts to detect personality tasks. This CNN architecture helped to extract monogram, bigram, and trigram features from texts, which achieved remarkable performance in terms of accuracy. It is worth noting that this work used hand-crafted features, which depended heavily on domain expertise. In addition, although this research utilized the same architecture, they trained five different neural networks for five personality traits. Furthermore, linguistic features like Mairesse are mostly language-dependent, making usage of neural network model concatenated Mairesse features more time consuming. Sun et al. [15] employed BiLSTM and CNN to predict the final personality trait, which could efficiently learn text features. However, polysemy is a very common phenomenon in textual data. This work only used word embedding to represent the semantics of texts, but word embedding could not eliminate the polysemy issue entirely, which would influence classification performance to some extent. Although some progress has been made, the personality detection task is still in its early stage, and there is a high potential for achieving remarkable performance improvement.

Different from the above works, in this paper, we propose a novel deep learning method without feature engineering or the need to create specific linguistic features like Mairesse, which makes the presented model traits independent and enable it to work well in various personality trait detection tasks. That is to say, the proposed neural network approach is flexible in predicting personality traits based on an atomic representation of texts. Moreover, our model can effectively alleviate the impact of polysemy in the personality detection task. More importantly, an advantage of our method is to allow the same deep learning model to adapt to the prediction of five different personality traits without modification to the model itself.

2.2 Deep learning-based text mining

In general, the study of personality trait recognition based on textual information can be classified into text mining tasks, which are highly correlated with several natural language processing (NLP) technologies, such as text classification [24] and sentiment classification [25]. Traditional methods of text classification and sentiment classification mainly rely on effective human-designed representations and input features (e.g., word N-gram [26]). In recent years, deep learning has demonstrated advanced performance and flexibility in various application fields, such as computer vision [27], intelligent talent computing like personnel performance prediction [16], [17] and especially text mining [24, 28]. Indeed, many researchers have designed effective deep learning models for text classification and sentiment classification tasks that can learn large-scale textural data without cumbersome and time-consuming feature engineering.

Researchers have adopted two representative and widely-used deep learning architectures, namely convolutional neural network (CNN) [29] and recurrent neural network (RNN) [30], which have been successfully applied to NLP tasks.

Specifically, CNN can efficiently capture local semantical information in textural data. For example, Kalchbrenner et al. [31] presented a dynamic convolutional neural network (DCNN) for modeling sentences, which achieved impressive performance. Ren et al. [25] proposed a CNN-based neural network model for sentiment classification tasks. Zhao et al. [32] proposed a syntax convolutional neural network (SCNN) model to extract drug–drug interactions, which achieved competitive results compared to the existing works.

Meanwhile, RNN-based methods have also received significant attention for many NLP applications and are more beneficial for modeling sequential textual data. For example, as one of the representative works in these fields, Tang et al. [28] used gated RNN to tackle document level sentiment classification. In [33], Zhang et al. introduced a novel deep RNN model for addressing the problem of automatically extracting key phrases. Moreover, the effectiveness of RNN has also been demonstrated in other text tasks, e.g., some educational applications like student performance prediction [34] and machine translation [35].

This work succeeds some excellent ideas mentioned in the aforementioned studies regarding the properties of personality trait recognition. In this paper, we propose a novel Semantic-enhanced Personality Recognition Neural Network model SEPRNN.

3 Method

In this section, we introduce the technical details of the semantic-enhanced personality recognition neural network (SEPRNN). As shown in Fig. 1, SEPRNN mainly includes two components, namely Word-level Representation and Personality Trait Recognition.

Specifically, in Word-level Representation, we first use a sequence of words $w_1, w_2, \cdots, w_n$ as the input of the neural network. The words of social texts are represented by context learning. Then, to capture the higher level of semantic representations from textual data, we feed the context learning into the fully connected layer. Finally, the learned high semantic representations of texts are fed into Personality Trait Recognition to predict the label of personality traits. Here, for characteristic multi-labeled personality trait recognition, we use the binary cross-entropy loss function in the training procedure.

3.1 Word-level representation

Polysemy is a highly universal phenomenon of language. To embed the sequential dependence between words into corresponding representations and obtain a more precise word meaning, we use the context learning method to capture the contexts of textual data in personality trait recognition tasks.

In recent years, word embedding has been successfully applied for various NLP tasks, which refers to the process of mapping words to real value vectors. Various word representation models, such as Word2Vec [36] and GloVe [37], have been presented to learn word embedding. In our experiments, we used the pre-trained word embeddings from GloVe.

The same word may have different meanings in different contexts, but relying only on word embedding fails to accurately express the actual semantic. Therefore, we combine word embedding with contextual information to obtain precise semantics for words. Figure 2 illustrates the context learning method in Word-level Representation, which utilizes BiGRU to extract the left and right context semantics of words. We define $c_l(w_i)$, $e_w(w_i)$ and $c_r(w_i)$ as the left context vector of word w_i, the word embedding of word w_i, and the right context vector of word w_i. Formally, the left context embedding and right context embedding can be formulated as follows:

$$c_{f\_l}(w_i) = f(W_{f\_l} c_{f\_l}(w_{i-1}) + W^{se}_{f\_l} e_w(w_i)) \tag{1}$$

$$c_{b\_l}(w_i) = f(W_{b\_l} c_{b\_l}(w_{i-1}) + W^{se}_{b\_l} e_w(w_i)) \tag{2}$$

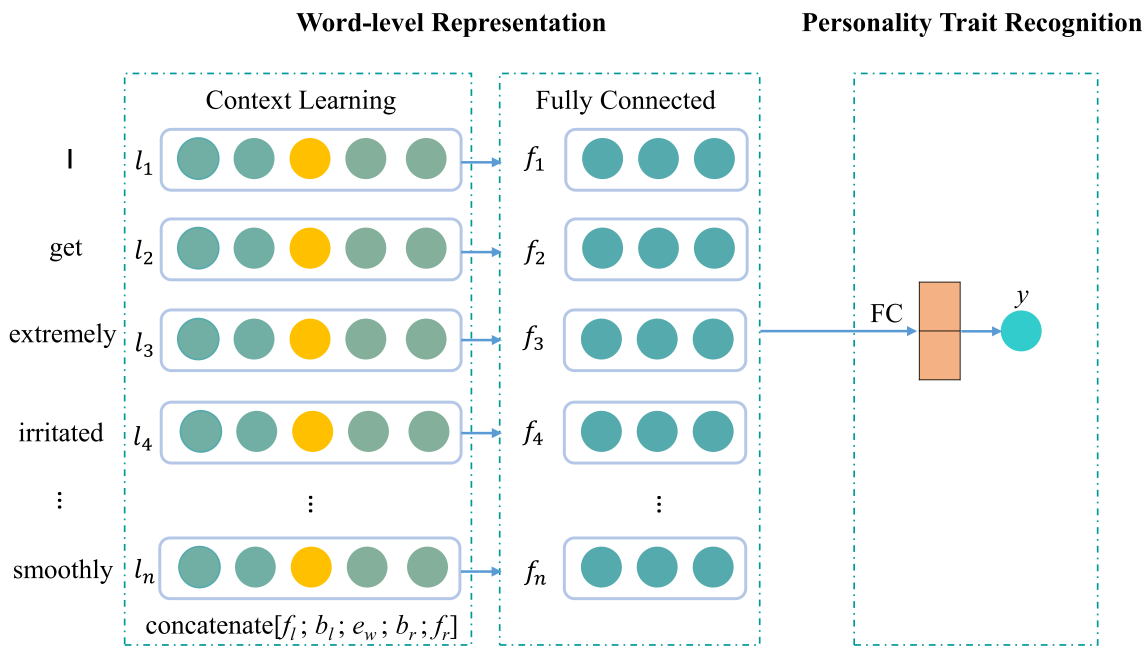

Fig. 1 An illustration of the proposed semantic-enhanced personality recognition neural network (SEPRNN), which can be separated into two components, namely Word-level Representation and Personality Trait Recognition. The figure shown is an example of the sentence "I get extremely irritated when things don't operate smoothly", and the subscript indicates the corresponding position of the word in this sentence, where y expresses classes of the textual data

$$c_l(w_i) = [c_{f\_l}(w_i); c_{b\_l}(w_i)] \quad (3)$$

where W_{f_l} and W_{b_l} are input weight matrices, $W^{se}_{f\_l}$ and $W^{se}_{b\_l}$ are matrices of the recurrent structure which are used for combining the current word's semantic embedding with the previous word's left contextual information, and f is the non-linear activation function.

Additionally, we adopt a similar way to calculate the right context $c_r(w_i)$, as depicted in Equations (4)–(6).

$$c_{b\_r}(w_i) = f(W_{b\_r}c_{b\_r}(w_{i+1}) + W^{se}_{b\_r}e_w(w_i)) \quad (4)$$

$$c_{f\_r}(w_i) = f(W_{f\_r}c_{f\_r}(w_{i+1}) + W^{se}_{f\_r}e_w(w_i)) \quad (5)$$

$$c_r(w_i) = [c_{b\_r}(w_i); c_{f\_r}(w_i)] \quad (6)$$

The semantic vector l_i of word w_i is the concatenation of the left context vector $c_l(w_i)$, the word embedding $e_w(w_i)$ and the right context vector $c_r(w_i)$. Specifically, the formulas l_i can be represented as:

$$l_i = [c_l(w_i); e_w(w_i); c_r(w_i)] \quad (7)$$

To obtain higher-level semantic features, here we utilize a fully connected layer to fuse the current word and the left and right contextual features. Therefore, the final semantic vector can be computed by (8):

$$y(w_i) = Wl_i + b \quad (8)$$

3.2 Personality trait recognition

Using the process of Word-level Representation, we can better learn the representations for texts of personality traits. We treat the task of personality trait recognition as a binary classification problem. Thus, in the final component of SEPRNN, we adopt a fully-connected layer with sigmoid activation function to learn a two-dimensional vector for predicting the trait label y. If the predicted label y is 1, it means that the sample has this trait, and vice versa. Specifically, to detect personality traits with multiple labels, we minimize the binary cross-entropy loss function to train our SEPRNN model in this step.

$$y = sigmoid(Wy(w_i) + b) \quad (9)$$

4 Experiments

In this section, we perform extensive experiments to answer the following questions: 1) Does our model outperform the existing works in personality trait recognition? 2) Is word-level representation learning useful for identifying semantic information of multi-labeled personality traits?

4.1 Dataset

In this paper, to evaluate the effectiveness of the proposed model, we conducted experiments using the following

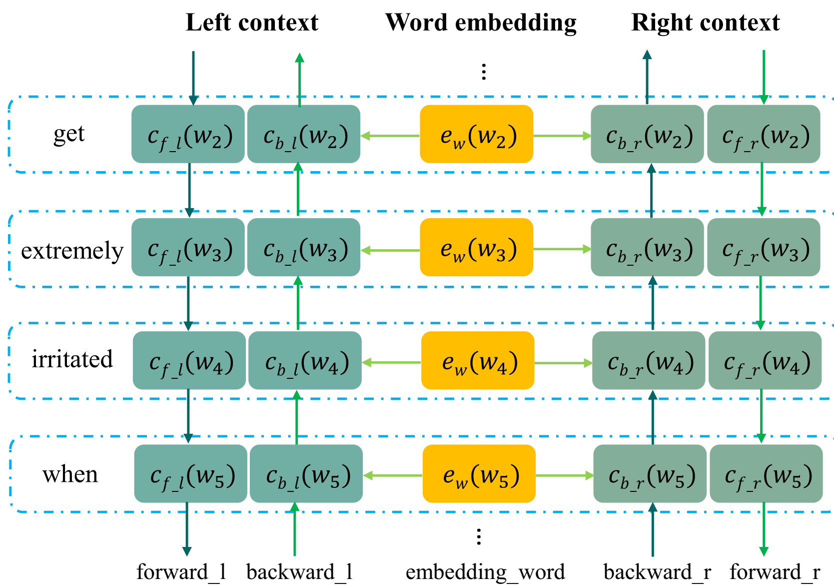

Fig. 2 The structure of context learning method. This figure is a partial example of the sentence "I get extremely irritated when things don't operate smoothly". The subscript expresses the corresponding position of the word in the sentence as described above

two datasets: stream-of-consciousness essays (SoCE) [19] and YouTube (YoTB) [38]. Table 1 provides detailed information about these datasets.

YoTB: This dataset contains 404 samples. Their Big Five personality scores range from 1 to 7. Personality traits are extroversion (EXT), neuroticism (NEU), agreeableness (AGR), conscientiousness (CON), and openness (OPN). The class of each trait is viewed as 1 if personality trait scores are greater than four. Otherwise the label of each trait is regarded as 0.

SoCE: This dataset consists of 2467 samples labeled. The labels are EXT, NEU, AGR, CON and OPN. In the context of personality recognition, each trait is 1 if the sample belongs to this trait. Otherwise each trait is 0.

4.2 Experimental setup

In this subsection, we introduce experimental settings. All experiments were conducted on Python 3.5 with TensorFlow 1.10 and Keras 2.2. The computer we employed was comprised of an Intel(R) Core(TM) i7-8550U CPU running at 1.80 GHz and a NVIDIA GeForce MX150 GPU with the cuDNN library.

Both datasets were separated into a training set, validation set, and testing set. We split 90% of the two datasets into a training set, used the remaining 10% as the test set, and 20% of the training set was used as the validation set. We did not delete any symbols in the texts. For the two datasets, we utilized the same architecture and parameters. We set the vector size of the word embedding as $|e| = 100$ and the size of the context vector as $|c| = 200$. The dimension of the dense layer was 200. The learning of our model was conducted during 10 epochs with a mini-batch size of 32 samples. Moreover, we adopted RMSprop as an optimizer and set the learning rate as 0.001. Table 2 shows our hyper-parameter setting.

Table 1 A summary of the datasets, including the number of multi-labeled instances and the number of classes

Datasets	Multi-labeled Instances	Classes
YoTB	404	7
SoCE	2, 467	2

4.3 Baseline methods

To evaluate the effectiveness of our SEPRNN model, several state-of-the-art models were selected as baselines. Baseline methods include Mairesse, TF-IDF+Bayes, 2CNN, CNN+Mairesse, LSTM, BiLSTMs, 2CLSTM, self-attention, AC-BiLSTM, and ABCDM. The details of the methods are as follows:

– **Mairesse**: Mairesse baseline is a well-known baseline in personality recognition which uses different combinations of feature sets [41].

Table 2 Parameter settings of the neural network

Parameters	Value
The size of the word embedding	100
the size of the context vector	200
dimension of dense layer	200
batch size	32
epochs	10
the learning rate	0.001

- **TF-IDF+Bayes**: TF-IDF+Bayes is a common method in text classification tasks and is a Bayesian classifier based on TF-IDF features. This model is used for personality detection.
- **2CNN**: 2CNN is a two-dimensional convolution network model and is used for predicting personality [22].
- **CNN+Mairesse**: This CNN+Mairesse model is used for personality detection modeling tasks and employs convolution kernels for three-dimensional words, sentences and documents [13].
- **LSTM**: LSTM [39] is a variant of recurrent neural network (RNN), and is a typical model in text classification tasks.
- **BiLSTMs**: Bidirectional LSTM [42] also belongs to RNN and can capture contextual information well in text classification.
- **2CLSTM**: This deep learning framework is used to detect personality traits. 2CLSTM first employs bidirectional LSTM concatenated with the current word vectors to encode sentence embeddings. Then, a group of convolutional layers and max-pooling layers learn latent sentence groups [15].
- **Self-attention**: Self-attention [40] is a state-of-the-art deep learning architecture that is widely used for text tasks. It can also capture global information by flexibly selecting features according to the contributions of different features.
- **AC-BiLSTM**: AC-BiLSTM model is designed for text classification task. It uses one dimension convolutional layer with different filter sizes to extract the features from the word embedding vectors. The output of the convolutional layer is fed into a bidirectional LSTM layer. And then attention mechanism is used to focus on the important information of the bidirectional LSTM output layer. Finally, softmax is used as the classifier [59].
- **ABCDM**: ABCDM deep model is used for sentiment classification task. This model employs two parallel bidirectional LSTM and bidirectional GRU layers to extract the sequence information. The attention mechanism is then applied on the bidirectional outputs layers to pay attention to different words. Finally, ABCDM model uses the convolution and pooling layer to reduce the dimensionality of features and capture local features [60].
- **SEPRNN**: Our proposed method.

4.4 Evaluation metrics

Personality trait recognition tasks in this paper can be treated as binary classification problems. Thus, we used accuracy, precision, recall, and F1-scores (F1) as evaluation metrics to measure the performance of the SEPRNN model in the experiments. F1 is a comprehensive evaluation index of precision and recall. The indexes, including accuracy, precision, recall, and F1, are defined by true positive (TP), false negative (FN), false positive (FP), and true negative (TN). The formula of accuracy, precision, recall, and F1 is as follows:

$$accuracy = \frac{TP + TN}{TP + FN + FP + TN} \tag{10}$$

$$precision = \frac{TP}{TP + FP} \tag{11}$$

$$recall = \frac{TP}{TP + FN} \tag{12}$$

$$F1 = \frac{2 \times precision \times recall}{precision + recall} \tag{13}$$

The indexes accuracy, precision, recall, and F1 are in the range of [0, 1]. The bigger the value is, the better the classification performance is.

4.5 Overall performance

To demonstrate the effectiveness of the proposed model, we used the previous baselines and SEPRNN model to evaluate the performance on personality trait recognition datasets. This paper compared the approaches of Mairesse, TF-IDF+Bayes, 2CNN, CNN+Mairesse, LSTM, BiLSTMs, 2C-LSTM, self-attention, AC-BiLSTM, and ABCDM. The previous work (e.g., TF-IDF+B-ayes) was based on traditional machine learning methods, while 2CNN, CNN+Mairesse, LSTM, BiLSTMs, 2CLSTM, self-attention, AC-BiLSTM, ABCDM, and our SEPRNN model used deep learning. Here, our SEPRNN model was used to recognize the personality traits based only on texts without extra features. The experimental results using accuracy, precision, recall and F1 metrics are reported in Tables 3, 4, 5 and 6. The ten baselines and our SEPRNN model used in the comparison are listed in the first column of the tables. We employ bold font to emphasize the top model for all metrics.

The following observations can be obtained from Tables 3–6. Table 3 shows the performance of the baselines and SEPRNN on personality trait recognition data in terms of accuracy (%). In all datasets, our method outperforms all the baselines for all five traits. For the SoCE dataset, the accuracy of our SEPRNN model on EXT, NEU, AGR, CON, and OPN labels is 58.91%, 59.51%, 57.49%, 57.49%, and 63.16%, respectively. For the YouTube dataset, the accuracy of our SEPRNN model on the five trait labels is 70.73%, 78.05%, 73.17%, 65.85%, and 65.85%, respectively. We compared our SEPRNN to well-designed feature sets of the Mairesse method. The experimental results validate that our SEPRNN model outperforms the handcrafted features-based method. We believe that the

Table 3 The performance of SEPRNN and baselines for personality trait recognition in terms of accuracy (%)

Model	SoCE					YoTB				
	EXT	NEU	AGR	CON	OPN	EXT	NEU	AGR	CON	OPN
Mairesse [41]	55.13	58.09	55.35	55.28	59.57	-	-	-	-	-
TF-IDF+Bayes	54.25	52.23	48.58	52.23	57.89	53.66	58.54	58.54	53.66	46.34
2CNN	-	-	-	-	-	-	-	-	-	-
CNN+Mairesse [13]	58.09	57.33	56.71	56.71	61.13	-	-	-	-	-
LSTM	48.58	46.15	53.04	52.02	54.05	63.41	71.95	57.32	63.41	53.66
BiLSTMs	53.44	54.05	56.68	55.47	57.89	53.66	60.98	63.41	53.66	56.10
2CLSTM	49.80	53.85	53.44	54.25	54.66	63.41	60.98	58.54	58.54	53.66
Self-attention	54.66	59.11	53.85	55.47	60.73	65.85	56.10	62.20	54.88	56.10
AC-BiLSTM	56.48	56.68	50.20	55.26	61.54	68.29	68.29	70.73	63.41	60.98
ABCDM	56.68	51.82	55.87	48.99	61.74	63.41	62.20	67.07	63.41	59.76
SEPRNN (Ours)	**58.91**	**59.51**	**57.49**	**57.49**	**63.16**	**70.73**	**78.05**	**73.17**	**65.85**	**65.85**

Bold highlights the best performance for each trait

SEPRNN model can capture long-distance patterns. An advantage of the SEPRNN model is that it does not require handcrafted features. This means that the SEPRNN model may be useful in low-resource languages.

In Table 3, compared with the traditional machine learning methods (e.g., TF-IDF+Bayes), the accuracy improvements range from 4.66–8.91% and 12.19–19.51% for the five traits on SoCE and YouTube datasets, respectively. The experimental results demonstrate that our deep learning method is superior over the traditional approaches (e.g., TF-IDF+Bayes) for the two datasets. This proves that our SEPRNN model can effectively learn more contextual information of texts compared to traditional approaches based on the TF-IDF model. Also, our SEPRNN model may suffer from the data sparsity issue less.

We also compared the accuracy values of all the deep learning models on the two datasets with five different traits (Table 3). Obviously, SEPRNN outperforms other neural networks on all the datasets. For example, compared to the CNN+Mairesse model, the accuracy of SEPRNN on five different labels increases by 0.82%, 2.18%, 0.78%, 0.78%, and 2.03% for the SoCE dataset. The CNN+Mairesse model

Table 4 The performance of SEPRNN and baselines for personality trait recognition in terms of precision (%)

Model	SoCE					YoTB				
	EXT	NEU	AGR	CON	OPN	EXT	NEU	AGR	CON	OPN
Mairesse	-	-	-	-	-	-	-	-	-	-
TF-IDF+Bayes	50.62	53.90	49.12	52.48	61.22	48.57	60.53	62.16	54.84	45.71
2CNN	53.21	48.59	50.29	49.23	47.73	49.63	57.96	50.83	50.71	49.54
CNN+Mairesse	50.20	50.47	50.55	48.10	50.60	52.08	59.94	50.46	51.85	51.33
LSTM [15]	49.72	49.26	58.42	51.81	53.63	61.85	53.25	55.38	61.17	57.91
LSTM	49.63	50.00	58.78	53.59	53.39	61.11	53.54	56.10	61.62	58.06
BiLSTMs [15]	53.79	53.17	57.43	54.49	56.54	52.20	56.13	55.47	55.67	51.83
BiLSTMs	53.90	53.87	57.35	55.06	56.19	52.06	55.48	55.48	55.56	51.59
2CLSTM [15]	55.64	56.77	58.87	53.52	54.19	67.69	61.28	58.02	61.17	55.46
2CLSTM (our experiments)	55.63	56.91	58.57	53.82	54.78	68.05	60.98	58.54	61.56	55.56
Self-attention	55.16	56.39	52.97	52.86	62.76	71.43	64.71	67.00	57.69	62.86
AC-BiLSTM	57.25	57.53	50.33	56.32	61.16	62.36	60.14	65.74	63.48	59.45
ABCDM	56.82	55.74	56.73	51.87	63.03	69.44	46.10	50.03	61.76	60.98
SEPRNN (Ours)	**58.48**	**59.71**	**61.20**	**58.40**	**64.43**	**71.79**	**74.47**	**75.00**	**64.71**	**68.75**

Bold indicates the best performance for each trait used five different neural networks for five traits, but the apparent advantage of our model is to use the same model to predict various personality traits. This proves that our proposed approach is more flexible and efficient than CNN+Mairesse. Compared with the 2CLSTM model, the accuracy of SEPRNN on five different labels for SoCE dataset increases by 9.11%, 5.66%, 4.05%, 3.24%, and 8.50%, while the accuracy of SEPRNN on five different tags for the YouTube dataset improves by 7.32%, 17.07%, 14.63%, 7.31% and 12.19%, respectively. We believe that the main reason for these results is that the word-level representation method helps us to gain a more precise word semantic and avoid the impact of polysemy in the personality trait recognition task. The experimental results demonstrate the effectiveness of the proposed approach.

Table 4 lists the performance of the baselines and SEPRNN on personality trait recognition data in terms of precision (%). In all datasets, our approach outperforms all the baseline methods on five trait labels. For the SoCE dataset, the precision of our SEPRNN model on EXT, NEU, AGR, CON, and OPN labels is 58.48%, 59.71%, 61.20%, 58.40%, and 64.43%, respectively. For the YouTube dataset, the precision of our SEPRNN model on the five trait labels is 71.79%, 74.47%, 75.00%, 64.71% and 68.75%, respectively. We also compared the precision values of the neural networks on the two datasets with five different traits. Results show that our model SEPRNN performs better than all other neural network methods on all the datasets. For instance, we compared SEPRNN to 2CNN and CNN+Mairesse, and found that the SEPRNN performs better than 2CNN and CNN+Mairesse in all datasets. We believe that the reason is that the context learning structure in the SEPRNN method captures more contextual information than the window-based structure in 2CNN and CNN+Mairesse methods. These results validate the effectiveness of the proposed approach.

In Table 4, compared to 2CLSTM model, the precision of SEPRNN on five different labels increases by 2.84%, 2.94%, 2.33%, 4.88%, and 10.24% for the SoCE dataset, while the precision of SEPRNN improves by 4.10%, 13.19%, 16.98%, 3.54%, and 13.29% for the YouTube dataset, respectively. We find that the proposed SEPRNN method has a significant performance improvement for detecting personality traits on each dataset in terms of the precision evaluation metric. Compared with AC-BiLSTM model, the precision improvements range from 1.23–10.87% and 1.23–14.33% on SoCE and YouTube datasets, respectively. The results demonstrate that our SEPRNN method is superior over AC-BiLSTM model in terms of precision. Due to polysemy being a common phenomenon in language, we fuse the word itself and the context information to obtain a more complete semantic representation. We believe that word-level representation helps to achieve a significant performance improvement.

Table 5 displays the performance of the baselines and SEPRNN on personality trait recognition data in terms of recall (%). In all datasets, our SEPRNN approach is superior over all the baselines on five trait labels.

Table 6 shows the performance of the baselines and SEPRNN on personality trait recognition data in terms of F1 (%). For the SoCE dataset, the F1 of our SEPRNN model on EXT, NEU, AGR, CON, and OPN labels is 71.50%, 62.36%, 71.92%, 63.46%, and 67.84%, respectively. For the YouTube dataset, the F1 of our SEPRNN model on the five trait labels is 82.35%, 79.55%, 83.08%, 77.20%, and 78.57%, respectively. In all datasets, our method achieves

Table 5 The performance of SEPRNN and baselines for personality trait recognition in terms of recall (%)

Model	SoCE					**YoTB**				
	EXT	NEU	AGR	CON	OPN	EXT	NEU	AGR	CON	OPN
Mairesse	-	-	-	-	-	-	-	-	-	-
TF-IDF+Bayes	71.30	58.91	67.74	52.15	47.62	94.44	59.00	88.46	77.27	84.21
2CNN	-	-	-	-	-	-	-	-	-	-
CNN+Mairesse	-	-	-	-	-	-	-	-	-	-
LSTM	53.17	51.13	55.40	64.57	65.67	95.65	73.17	58.54	77.78	75.00
BiLSTMs	60.00	56.28	64.10	59.32	64.18	95.45	80.00	78.57	68.18	80.00
2CLSTM	90.43	53.44	58.99	62.60	52.99	65.52	82.76	76.92	90.91	83.33
Self-attention	50.36	63.56	78.40	62.71	67.91	86.21	78.57	72.41	65.22	81.48
AC-BiLSTM	51.36	56.81	50.10	56.33	60.60	62.93	54.12	67.10	60.75	55.51
ABCDM	91.24	53.00	57.09	50.73	62.94	86.21	65.85	70.73	91.30	65.85
SEPRNN (Ours)	**91.97**	**65.25**	**87.20**	**69.49**	**71.64**	**96.55**	**85.37**	**93.10**	**95.65**	**91.67**

Bold highlights the best result for each trait

Table 6 The performance of SEPRNN and baselines for personality trait recognition in terms of F1 (%)

Model	SoCE					YoTB				
	EXT	NEU	AGR	CON	OPN	EXT	NEU	AGR	CON	OPN
Mairesse	-	-	-	-	-	-	-	-	-	-
TF-IDF+Bayes	59.21	56.30	56.95	50.59	53.57	64.15	59.76	73.02	64.15	59.26
2CNN	-	-	-	-	-	-	-	-	-	-
CNN+Mairesse	-	-	-	-	-	-	-	-	-	-
LSTM	51.34	50.56	57.04	58.57	58.90	74.58	61.83	57.29	68.76	65.45
BiLSTMs	56.79	55.05	60.54	57.11	59.92	67.37	65.52	65.04	61.23	62.73
2CLSTM	68.88	55.12	58.78	57.88	53.87	66.76	70.22	66.48	73.41	66.67
Self-attention	52.65	59.76	63.23	57.36	65.23	78.13	70.97	69.60	61.22	70.97
AC-BiLSTM	54.15	57.17	50.21	56.32	60.88	62.64	56.97	66.41	62.09	57.41
ABCDM	70.03	54.34	56.91	51.29	62.98	76.92	54.23	58.61	73.68	63.32
SEPRNN (Ours)	**71.50**	**62.36**	**71.92**	**63.46**	**67.84**	**82.35**	**79.55**	**83.08**	**77.20**	**78.57**

Bold indicates the best result for each trait

the best F1, which outperforms the other baselines on five trait labels. Indeed, this proves that the presented method is more efficient than the baselines. In comparison with baseline self-attention, the F1 of SEPRNN on five different labels increases by 18.85%, 2.60%, 8.69%, 6.10%, and 2.61% for the SoCE dataset, while the F1 of SEPRNN improves by 4.22%, 8.58%, 13.48%, 15.98%, and 7.60% for the YouTube dataset, respectively. We consider that self-attention may not be suitable for this sequential data. Next, we compare against the recent baseline ABCDM. As Table 6 shows, compared to ABCDM model, the F1 of SEPRNN on five labels increases by 1.47%, 8.02%, 15.01%, 12.17%, and 4.86% for the SoCE dataset, while the F1 of SEPRNN improves by 5.43%, 25.32%, 24.47%, 3.52%, and 15.25% for the YouTube dataset, respectively. The SEPRNN method outperforms window-based ABCDM model. This illustrates that SEPRNN method that does not depend on the window size can obtain more precise word semantics. We believe that the reason is that the word-level representation structure in the SEPRNN method learns more contextual information than the window-based structure in ABCDM method.We believe the SEPRNN model has a strong learning ability. The results indicate that the SEPRNN model can make full use of the texts, benefiting the prediction.

In summary, a detailed comparison of baselines with accuracy, precision, recall, and F1 is shown in Table 3–6, and indicates that our SEPRNN model outperforms the other baselines on all the indicators. The experimental results demonstrate that the SEPRNN model has an excellent ability to detect multi-labeled personality traits.

4.6 Ablation study

To further validate the effectiveness of the strategy, we report the ablation study of our SEPRNN method. Our main approach performs better than its sub-module, although the sub-module may still be useful when applied to some other tasks. We applied our SEPRNN model and its sub-module, which are without context learning, to the two datasets, respectively. The experimental results for the ablation study are reported in Table 7. The results of the "- context learning" row refers to the results without context learning method.

Table 7 Ablation study of SEPRNN with various modules on personality trait recognition

	Strategy	Accuracy									
		EXT	Δ	NEU	Δ	AGR	Δ	CON	Δ	OPN	Δ
SoCE	SEPRNN (Our Model)	58.91	-	59.51	-	57.49	-	57.49	-	63.16	-
	-context learning	55.06	−3.85	51.62	−7.89	54.86	−2.63	51.21	−6.28	59.11	−4.05
YoTB	SEPRNN (Our Model)	70.73	-	78.05	-	73.17	-	65.85	-	65.85	-
	-context learning	65.85	−4.88	70.73	−7.32	71.95	−1.22	54.88	−10.97	60.98	−4.87

The experimental results shown in Table 7 indicate: 1) Our main model outperforms the variant (- context learning) under the same settings; 2) As expected, the experimental results for the simplified models all drop a lot in the two datasets. For the sub-module "- context learning", the drop of accuracy, precision, recall, and F1 shows that the context learning method is useful. Specifically, the context learning method helps to eliminate the influence of polysemy, which can obtain a more accurate word semantic. This clearly proves the effectiveness of this strategy, which is like the context learning method, significantly boosting the performance of personality trait recognition with multiple labels.

5 Conclusion

Automatic and accurate personality trait recognition can have an impact in many ways. For instance, recruiters can fit the most suitable candidates for a certain job based on their personality. Products and services can adjust the behavior to better match each user's personality. Moreover, people may also rely on automatic personality recognition to make decisions about their career paths and lives. In view of the significant role of personality trait recognition in society, it is an important challenge to use the behavior footprint observed from texts to assess psychological traits accurately and automatically.

In this paper, we developed a novel semantic-enhanced personality recognition neural network (SEPRNN) for recognizing personality traits with multiple labels. Specifically, we first designed a word-level semantic representation for texts of the individuals based on the context learning method. We further used a fully connected layer to learn the higher semantics of texts. Then, binary cross-entropy loss function was used to predict multi-labeled personality trait tasks. Finally, extensive experiments conducted on personality trait data clearly validated the effectiveness of our SEPRNN model compared with several baseline methods.

In the future, we would like to focus on other personality model, e.g., HEXACO and DSM-5. We also intend that our solutions be generalized for other similar text tasks that handle multiple labels.

Acknowledgements This work was sponsored by the Key Research and Development Program in Shaanxi Province of China (No.2019ZD LGY03-10) and the National Natural Science Foundation Projects of China (No.61877050).

References

1. Li Y, Zhu T, Li A, Fan Z (2011) Web behavior and personality: A review. In: Web Society
2. Schwartz HA, Eichstaedt JC, Kern ML, Dziurzynski L, Ramones SM, Agrawal M, Shah A, Kosinski M, Stillwell D, Seligman MEP (2013) Personality, Gender, and Age in the Language of Social Media: The Open-Vocabulary Approach. Plos One 8(9):e73791
3. Qiu L, Lu J, Ramsay JE, Yang S, Qu W, Zhu T (2017) Personality expression in Chinese language use. Int J Psychol 52(6):463–472
4. Mccrae RR, Costa PT (1987) Validation of the Five-Factor Model of Personality Across Instruments and Observers. J Personal Soc Psychol 52(1):81–90
5. Roth PL, Bobko P, Van Iddekinge CH, Thatcher JB (2016) Social Media in Employee-Selection-Related Decisions A Research Agenda for Uncharted Territory. J Manag 42(1):269–298
6. Tett RP, Jackson DN, Rothstein M (1991) Personality Measures as Predictors of Job Performance: A Meta-Analytic Review. Person Psychol 44(4):703–742
7. Barrick MR, Mount MK (1991) The big five personality dimensions and job performance: a meta-analysis. Person Psychol 44(1):1–26
8. Judge TA, Higgins CA, Thoresen CJ, Barrick MR (1999) The big five personality traits, general mental ability, and career success across the life span. Person Psychol 52(3):621–652
9. Digman JM (1990) Personality Structure: Emergence of the Five-Factor Model. Ann Rev Psychol 41(1):417–440
10. Mairesse F, Walker MA, Mehl MR, Moore RK (2007) Using linguistic cues for the automatic recognition of personality in conversation and text. J Artif Intell Res 30(1):457–500
11. Nguyen T, Phung DQ, Adams B, Venkatesh S (2011) Towards Discovery of Influence and Personality Traits through Social Link Prediction. In: Proceedings of the Fifth International Conference on Weblogs and Social Media. Barcelona, Catalonia, Spain
12. Tadesse MM, Lin H, Xu B, Yang L (2018) Personality Predictions Based on User Behavior on the Facebook Social Media Platform. IEEE Access 6:61959–61969
13. Majumder N, Poria S, Gelbukh AF, Cambria E (2017) Deep Learning-Based Document Modeling for Personality Detection from Text. IEEE Intell Syst 32(2):74–79
14. Lai S, Xu L, Liu K, Zhao J (2015) Recurrent Convolutional Neural Networks for Text Classification. Proceedings of the Twenty-Ninth AAAI Conference on Artificial Intelligence. 2267–2273. http://www.aaai.org/ocs/index.php/AAAI/AAAI15/paper/view/9745
15. Sun X, Liu B, Cao J, Luo J, Shen X (2018) Who Am I? Personality Detection Based on Deep Learning for Texts. In: 2018 IEEE International Conference on Communications, ICC 2018, Kansas City, pp 1–6
16. Xue X, Feng J, Gao Y, Liu M, Zhang W, Sun X, Zhao A, Guo SX (2019) Convolutional Recurrent Neural Networks with a Self-Attention Mechanism for Personnel Performance Prediction. Entropy 21(12):1227
17. Xue X, Gao y, Liu M, Sun X, Zhang W, Fend J (2021) GRU-based capsule network with an improved loss for personnel performance prediction. Applied Intelligence. pages 1–14
18. Teng M, Zhu H, Liu C, Zhu C, Xiong H (2019) Exploiting the Contagious Effect for Employee Turnover Prediction. In: The Thirty-Third AAAI Conference on Artificial Intelligence, AAAI 2019, The Thirty-First Innovative Applications of Artificial Intelligence Conference, IAAI 2019, The Ninth AAAI Symposium on Educational Advances in Artificial Intelligence, EAAI 2019, Honolulu, pp 1166–1173
19. James W P, Laura A K (1999) Linguistic styles: language use as an individual difference. Journal of personality and social psychology 77(6):1296–1312
20. Ross C, Orr ES, Sisic M, Arseneault JM, Simmering MG, Orr RR (2009) Personality and motivations associated with Facebook use. Comput Hum Behav 25(2):578–586

21. Mohammad SaifM., Kiritchenko S (2013) Using Nuances of Emotion to Identify Personality. CoRR, arXiv:1309.6352
22. Wei H, Zhang F, Yuan NJ, Cao C, Fu H, Xie X, Rui Y, Ma W-Y (2017) Beyond the Words: Predicting User Personality from Heterogeneous Information. In: Proceedings of the Tenth ACM International Conference on Web Search and Data Mining, WSDM 2017, Cambridge, pp 305–314
23. Yu J, Markov K (2017) Deep learning based personality recognition from Facebook status updates. In: IEEE 8th International Conference on Awareness Science and Technology, iCAST 2017, Taichung, pp 383–387
24. Wang S, Huang M, Deng Z (2018) Densely Connected CNN with Multi-scale Feature Attention for Text Classification. In: Proceedings of the Twenty-Seventh International Joint Conference on Artificial Intelligence, IJCAI 2018, Stockholm, pp 4468–4474
25. Ren Y, Zhang Y, Zhang M, Ji D (2016) Context-Sensitive Twitter Sentiment Classification Using Neural Network. In: Proceedings of the Thirtieth AAAI Conference on Artificial Intelligence, Phoenix, pp 215–221
26. Wang SI, Manning CD (2012) Baselines and Bigrams: Simple, Good Sentiment and Topic Classification. In: The 50th Annual Meeting of the Association for Computational Linguistics, Proceedings of the Conference, Korea - Volume 2: Short Papers, pp 90–94
27. Krizhevsky A, Sutskever I, Hinton GE (2012) ImageNet Classification with Deep Convolutional Neural Networks. In: Advances in Neural Information Processing Systems 25: 26th Annual Conference on Neural Information Processing Systems 2012. Proceedings of a meeting held December 3-6, Lake Tahoe, pp 1106–1114
28. Tang D, Qin B, Liu T (2015) Document Modeling with Gated Recurrent Neural Network for Sentiment Classification. In: Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing, EMNLP 2015, Lisbon, pp 1422–1432
29. Lecun Y, Bottou L (1998) Gradient-based learning applied to document recognition. Proc IEEE 86(11):2278–2324
30. Elman JL (1990) Finding Structure in Time. Cogn Sci 14(2):179–211
31. Kalchbrenner N, Grefenstette E, Blunsom P (2014) A Convolutional Neural Network for Modelling Sentences In: Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics, ACL 2014, Baltimore, Volume 1: Long Papers, pp 655–665
32. Zhao Z, Yang Z, Luo L, Lin H, Wang J (2016) Drug drug interaction extraction from biomedical literature using syntax convolutional neural network. Bioinform 32(22):3444–3453
33. Zhang Q, Wang Y, Gong Y, Huang X (2016) Keyphrase Extraction Using Deep Recurrent Neural Networks on Twitter. In: Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing, EMNLP 2016, Austin, pp 836–845
34. Su Y, Liu Q, Liu Q, Huang Z, Yin Y, Chen E, Ding CHQ, Wei S, Hu G (2018) Exercise-Enhanced Sequential Modeling for Student Performance Prediction. In: Proceedings of the Thirty-Second AAAI Conference on Artificial Intelligence, (AAAI-18), the 30th innovative Applications of Artificial Intelligence (IAAI-18), and the 8th AAAI Symposium on Educational Advances in Artificial Intelligence (EAAI-18), New Orleans, pp 2435–2443
35. Cho K, van Merrienboer B, Gülçehre Ç, Bahdanau D, Bougares F, Schwenk H, Bengio Y (2014) Learning Phrase Representations using RNN Encoder-Decoder for Statistical Machine Translation. In: Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing, EMNLP 2014, Doha, A meeting of SIGDAT, a Special Interest Group of the ACL, pp 1724–1734
36. Mikolov T, Sutskever I, Chen K, Corrado GS, Dean J (2013) Distributed Representations of Words and Phrases and their Compositionality. In: Advances in Neural Information Processing Systems 26: 27th Annual Conference on Neural Information Processing Systems 2013. Proceedings of a meeting held December 5-8, Lake Tahoe, pp 3111–3119
37. Pennington J, Socher R, Manning CD (2014) Glove: Global Vectors for Word Representation. In: Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing, EMNLP 2014, Doha, A meeting of SIGDAT, a Special Interest Group of the ACL, pp 1532–1543
38. Biel J-I, Tsiminaki V, Dines J, Gatica-Perez D (2013) Hi YouTube!: personality impressions and verbal content in social video. In: Proceedings of the 15th ACM on International conference on multimodal interaction, ACM, pp 119–126
39. Hochreiter S, Schmidhuber J (1997) Long Short-Term Memory. Neural Comput 9(8):1735–1780
40. Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomez AN, Kaiser L, Polosukhin I (2017) Attention is All you Need. In: Advances in Neural Information Processing Systems 30: Annual Conference on Neural Information Processing Systems 2017, Long Beach, pp 5998–6008
41. Mohammad SM, Kiritchenko S (2015) Using Hashtags to Capture Fine Emotion Categories from Tweets. Comput Intell 31(2):301–326
42. Graves A, Schmidhuber J (2005) Framewise phoneme classification with bidirectional LSTM and other neural network architectures. Neural Netw 18(5-6):602–610
43. Markovikj D, Gievska S, Kosinski M, Stillwell D (2013) Mining Facebook Data for Predictive Personality Modeling. In: ICWSM 2013
44. Balmaceda JoseM, Schiaffino SN, Godoy D (2014) How do personality traits affect communication among users in online social networks? Online Inf Rev 38(1):136–153
45. Yin H, Wang Y, Li Q, Xu W, Yu Y, Zhang T (2018) A Network-enhanced Prediction Method for Automobile Purchase Classification using Deep Learning. In: 22nd Pacific Asia Conference on Information Systems, PACIS 2018, Yokohama, pp 111
46. Yang H-C, Huang Z-R (2019) Mining personality traits from social messages for game recommender systems. Knowl-Based Syst 165:157–168
47. Majumder N, Poria S, Peng H, Chhaya N, Cambria E, Gelbukh AF (2019) Sentiment and Sarcasm Classification With Multitask Learning. IEEE Intell Syst 34(3):38–43
48. Shlomo A, Dhawle S, Moshe K, Pennebaker JW (2005) Lexical Predictors of Personality Type. In: 2005 Joint Annual Meeting of the Interface and the Classification Society of North America
49. Oberlander J, Nowson S (2006) Whose Thumb Is It Anyway? Classifying Author Personality from Weblog Text. In: ACL 2006, 21st International Conference on Computational Linguistics and 44th Annual Meeting of the Association for Computational Linguistics, Proceedings of the Conference, Sydney
50. Tausczik YR, Pennebaker JW (2009) The Psychological Meaning of Words: LIWC and Computerized Text Analysis Methods. J Lang Soc Psychol 29(1):24-54
51. Yarkoni T (2010) Personality in 100,000 Words: A large-scale analysis of personality and word use among bloggers. J Res Person 44(3):363–373
52. Poria S, Gelbukh AF, Agarwal B, Cambria E, Howard N (2013) Common Sense Knowledge Based Personality Recognition from Text. In: Advances in Soft Computing and Its Applications - 12th Mexican International Conference on Artificial Intelligence, MICAI 2013, Mexico City, Proceedings, Part II, pp 484–496

53. Kosinski M, Stillwell D, Graepel T (2013) Private traits and attributes are predictable from digital records of human behavior. Proc Natl Acad Sci USA 110(15):5802–5805
54. Youyou W, Kosinski M, Stillwell D (2015) Computer-based personality judgments are more accurate than those made by humans. Proc Natl Acad Sci USA 112(4):1036–1040
55. O"Connor BP (2002) A quantitative review of the comprehensiveness of the five-factor model in relation to popular personality inventories. Assessment 9(2):188–203
56. Goldberg LR, Johnson JA, Eber HW, Hogan R, Ashton MC, Cloninger CR, Gough HG (2006) The international personality item pool and the future of public-domain personality measures. J Res Personal 40(1):84–96
57. Mehta Y, Majumder N, Gelbukh AF, Cambria E (2020) Recent trends in deep learning based personality detection. Artif Intell Rev 53(4):2313–2339
58. Vinciarelli A, Mohammadi G (2014) A Survey of Personality Computing. IEEE Trans Affect Comput 5(3):273–291
59. Liu G, Guo J (2019) Bidirectional LSTM with attention mechanism and convolutional layer for text classification. Neurocomputing 337:325–338
60. Basiri MohammadE, Nemati S, Abdar M, Cambria E, Acharya UR (2021) ABCDM: An Attention-based Bidirectional CNN-RNN Deep Model for sentiment analysis. Fut Gener Comput Syst 115:279–294
61. Sun Y, Zhuang F, Zhu H, Song X, He Q, Xiong H (2019) The Impact of Person-Organization Fit on Talent Management: A Structure-Aware Convolutional Neural Network Approach. In: Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, KDD 2019, Anchorage, pp 1625–1633

Xia Xue received the B.S. and M.S. degrees from Northwest University, China. She is currently pursuing a Ph.D. degree in Computer Science at Northwest University, Xian, China. Her research is in the areas of machine learning, intelligent talent computing, and natural language processing. She has served as a reviewer for Arabian Journal for Science and Engineering and Journal of Intelligent & Fuzzy Systems.

Jun Feng received a Ph.D. from City University of Hong Kong in 2006. She is a Professor in the School of Information Science and Technology at Northwest University. Her research areas include pattern recognition and machine learning, especially in the fields of medical imaging analysis and intelligent education. Recent projects have included medical image analysis with deep learning, and intelligent education based on AI and Brain-Human Interaction. She has reviewed for many journals, including TSP, IEEE Access, JIVP, MTAP, JDIM, CJC, JCAD, OPE, and INFPHY. Conferences she has reviewed for include IEEE-VR,MICCAI, SIGCSE, IWCSE, and CompEd. She is a member of IEEE and ACM, and is co-author of 132 articles and coeditor of three books.

Xia Sun received a Ph.D. from Xian Jiaotong University, China, in 2006. She is a Professor in the School of Information Science and Technology at Northwest University, Xian, China. Her current research interests include natural language processing and machine learning. Recent projects have included causality extraction from educational data, and relationship extraction from bioinformatics text. She has served as a reviewer for IEEE Transactions on Industrial Informatics, IEEE Transactions on Neural Networks and Learning Systems and Chinese Journal of Electronics. She has also reviewed for the IEEE International Conference on Computational Science and Engineering. She is the co-author of 40 articles and is editor or co-editor of four books.