Information Sciences 596 (2022) 460–471

Contents lists available at ScienceDirect

Information Sciences

journal homepage: www.elsevier.com/locate/ins

Attention-based BiLSTM models for personality recognition from user-generated content

Lixin Zhou [a], Zhenyu Zhang [b,*], Laijun Zhao [a], Pingle Yang [a]

[a] Business School, University of Shanghai for Science and Technology, Shanghai 200093, China
[b] School of Automation, Nanjing University of Science and Technology, Nanjing 210094, China

ARTICLE INFO

Article history:
Received 27 December 2020
Received in revised form 7 February 2022
Accepted 6 March 2022
Available online 10 March 2022

Keywords:
Emojis
Attention-based Bi-LSTM
Personality traits
User-generated content

ABSTRACT

Emojis have been widely used in social media as a new way to express various emotions and personalities. However, most previous research only focused on limited features from textual information while neglecting rich emoji information in user-generated content. This study presents two novel attention-based *Bi-LSTM* architectures to incorporate emoji and textual information at different semantic levels, and investigate how the emoji information contributes to the performance of personality recognition tasks. Specifically, we first extract emoji information from online user-generated content, and concatenate word embedding and emoji embedding based on word and sentence perspectives. We then obtain the document representations of all users from the word and sentence levels during the training process and feed them into the attention-based *Bi-LSTM* architecture to predict the Big Five personality traits. Experimental results show that the proposed methods achieve state-of-the-art performance over the baseline models on the real dataset, demonstrating the usefulness and contribution of emoji information in personality recognition tasks. The findings could help researchers and practitioners better understand the rich semantics of emoji information and provide a new way to introduce emoji information into personality recognition tasks.

1. Introduction

Recent psychological studies have identified individual personalities as the motivation for individual behaviors such as sharing status, posting photos, and updating other types of user-generated content (UGC) on social network platforms [1]. Personality traits are commonly used to analyze the online behavior of users and make better management decisions in various domains, including recommender systems and personalized advertising [2], depression and suicidal behavior detection [3], marketing and business strategy [4], and product innovation [5].

The Big Five personality model is one of the most widely used models in psychological research [6]. It defines individual'spersonality traits into five distinct dimensions: *neuroticism*(NEU), *extraversion*(EXT), *openness*(OPN), *agreeableness* (AGR), and *conscientiousness*(CON). Individuals who have a high sense of *extraversion* are described as self-confident and communicative, they enjoy being around others and prefer to make friends online [7]. People with a high level of *agreeableness* are characterized by having positive and healthy relationships. *Openness* is also described as *open to experience*, it is characterized by high cognitive flexibility, intelligence, and verbal ability. People with a strong sense of *neuroticism* are always anxious, have trouble handling stressful situations or keeping their emotional stability, and tend to throw negative

* Corresponding author.
E-mail address: zhangzhenyu7@foxmail.com (Z. Zhang).

https://doi.org/10.1016/j.ins.2022.03.038
0020-0255/

emotions in their posts [8]. Recently, many researchers have suggested that a user's personality traits significantly affect his or her online behavior, and have utilized the Linguistic Inquiry and Word Count (LIWC) method to analyze the personality of individuals [9–11]. LIWC is an effective lexicon-based approach for extracting word counts, topic themes, and other statistical features from textual information. Many studies have proposed LIWC-based methods that extract linguistic features and detect different personalities by using traditional machine learning models, but those approaches are hard and costly to implement throughout the feature engineering process [12,13]. To reduce personal subjectivity during feature engineering, many researches have constructed neural network models to predict individual personality, and improved performance through self-learning mechanisms [14,15].

However, the abovementioned approaches only extract limited features from textual data, and neglect the rich emojis information from UGCs. To our knowledge, emojis have been used as a novel and popular ways of expressing emotions in online social networks. For example, 😂 is a very popular emoji (Unicode character is $U+1F602$), and it is associated with the keyword "face with tears of joy", which is widely used to show something fun or interesting. According to this study, emoji usage patterns are closely related to an individual's inner personality, and people who frequently use emoji usually have a high level of openness and favorable attitudes on social networks [16]. Despite the fact that emoji information from UGC can help increase performance on a variety of NLP tasks, most previous studies in the personality recognition domains have given it insufficient attention [17–19]. In this paper, emoji information is introduced into personality recognition tasks using deep learning architectures. Specifically, we propose two novel attention-based Bi-LSTM models that incorporate textual and emoji information at the word and sentence levels. We then utilize the sequence information to predict the users' personality traits. Experiment results show that our models outperform the baseline algorithms on the real dataset, demonstrating the usefulness of the word-level and sentence-level attention based Bi-LSTM models.

The main contributions of this article are summarized as follows.

(i) This study investigates that emoji information contributes to improving the performance of personality recognition tasks, and it adds rich emoji information into the process of personality recognition.
(ii) This work proposes two novel deep learning models to concatenate textual and emoji information at different semantic levels and going beyond the limits of closed-vocabulary approaches.
(iii) Experiment results demonstrate that the proposed models outperform the baseline methods on the real dataset, indicating the effectiveness of word-level and sentence-level models for personality recognition.

The remainder of this paper is organized as follows. Section 2 briefly reviews the approaches for the personality recognition task. In Section 3, we propose two novel models based on deep learning architectures that incorporate emoji and textual information on different semantic levels. Section 4 presents experimental results. Sections 5 and 6 draw conclusions and discuss the limits of our research.

2. Related work

2.1. Personality Recognition

In recent years, many researchers have attempted to study the personality traits of individuals from UGCs by using vocabulary-based approaches, such as LIWC and NRC. These methods obtain word statistics and linguistic characteristics and then investigate the correlation between individual personality and linguistic features. Golbeck and Robles extracted a total of 77 features from Facebook, including profile information, activities and preferences to predict the Big Five personality traits by using a Gaussian Processes model [9]. Then, Biel et al. extracted the behavioral features of individuals from 442 YouTube videos and discovered that these features were substantially related with specific personality dimensions [20]. Furthermore, sentiments and emotions [11], text topics [21], and social networks [22] were introduced to improve the performance of personality recognition.

Since deep learning techniques have achieved state-of-the-art performance in a variety of natural language processing (NLP) tasks, a growing number of researchers have attempted to apply it to improve the performance of personality detection tasks. More specifically, Majumder and Poria employed the convolutional neural network (CNN) to extract semantic information via word embedding and predict the personality of online users [23]. In addition, multilayer perceptron (MLP) and long short term memory (LSTM) were utilized to increase the performance of personality recognition over the baseline corpus of myPersonality [24,18]. Furthermore, Xue and Wu also concatenated users semantic information and online behaviors characteristics, and fed all features into an RCNN-based model to predict the five dimensions of personality traits [14]. However, most previous personality recognition methods often delete or eliminate emoji symbols during data pre-processing, ignoring the rich semantic information.

2.2. Emoji representation

Previous researches primarily focused on understanding the different roles of emojis in social communication, such as maintaining a conversational connection or gaining close relationships [25]. Emojis are widely used to express rich emotions and feelings in UGCs. Consequently, they bring both new opportunities and challenges in understanding semantics [19,16]. Sentiment lexicons present the fixed sentiment value and the emotion category of each emoji and offer a novel way to handle emotions and conduct sentiment analysis with emojis [26].

Nevertheless, the lexical and dictionary-based approaches may lack the ability to calculate the dynamic semantics of emojis when they are in different sentences [27]. Given the success of word embedding algorithms in the NLP fields, word embedding models can be used to explore the textual semantics of different emoji labels, and those methods have been proven effective in sentiment analysis and emotion recognition [25]. Barbieri and Ronzano explored the 2,389 emoji meanings through word embedding from tweet data, and calculated the topical similarity between different emojis [28]. In addition, emoji2vec [29] and EmojiNet [30] built emoji representations from their description texts and provided effective ways to introduce emoji information into sentiment analysis. To better understand emoji semantics, Felbo et al. proposed a transfer model (Deepmoji) for sentiment analysis and achieved better performance over several baseline datasets [31]. According to previous studies, emoji embedding representations have outperformed baseline datasets in a variety of NLP tasks, and are thus useful in personality recognition tasks. As a result, this study emphasizes the importance of emoji information and proposes two novel attention-based Bi-LSTM models that combine textual information and emoji information in the process of personality recognition.

3. Methodology

In this section, we construct two hierarchical models based on the attention mechanism, namely, the word-level attention-based Bi-LSTM and the sentence-level attention-based Bi-LSTM. The two models combine textual and emoji information over different semantic levels, and investigate their personality recognition performance based on a real dataset. The illustrations of our models are shown in Figs. 2 and 3.

3.1. Word Embedding and emoji embedding

Word embedding has become a significant component for many NLP tasks [23,18]. Here, textual information and emoji information are collected from user statuses and subsequently mapped them as word embedding or emoji embedding. Generally, UGC contains large amounts of short texts with emojis. Take the following sentence from an individual's Facebook status as an example: *feels tomorrow will be a busy and very nice day.*😊. The Unicode of 😊 is *U*+1*F*60*A*, which stands for *Smiling Face With Smiling Eyes Emoji*. The sequence words and emojis in the *i*th sentence s_i are presented as $t^{s_i} = \{t_1^{s_i}, w_2^{s_i}, \dots, t_m^{s_i}\}$ and $e^{s_i} = \{e_1^{s_i}, e_2^{s_i}, \dots, e_n^{s_i}\}$. Let $s^{u_i} = \{s_1, s_2, \dots, s_o\}$ describe all of sequence sentences that posted by the i th user u_i. Finally, $D = \{d_1, d_2, \dots, d_j\}$ denote the aggregation of all user $U = \{u_1, u_2, \dots, u_j\}$posts, where l_i is the length of each sentence s_i.

The Word2Vec model has been proven to have high accuracy in many NLP tasks that provided by Google [32]. Hence, we adopt this pre-trained word embedding model to map each word t^{s_i} into vector T^{s_i} in our dataset. Nevertheless, the word embedding model fails to train the emoji representation *E*. Then, we utilize DeepMoji [31] to explain the potential information of emojis in user-generated content. As a state-of-the-art method, DeepMoji outperforms other emoji embedding mod-

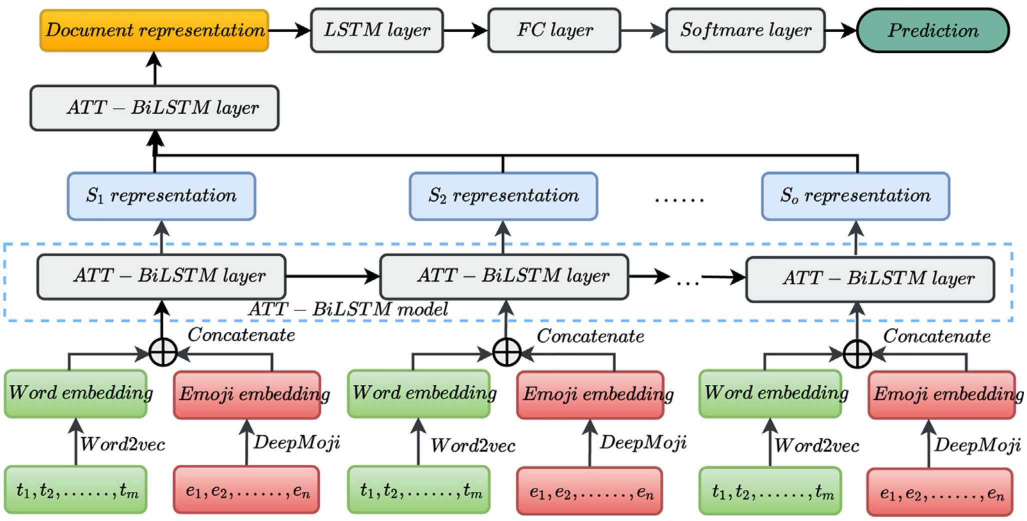

Fig. 2. Illustration of the word-level attention-based Bi-LSTM model.

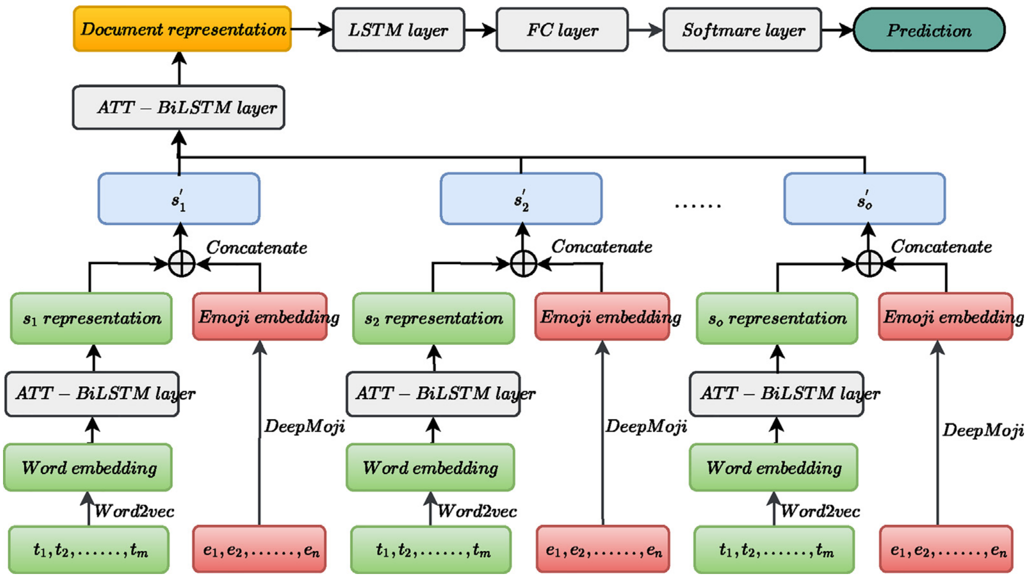

Fig. 3. Illustration of the sentence-level attention-based Bi-LSTM model.

els in terms of emoji representation and emotion detection. The model lists the emoji Unicode, name, possible senses, emotional content and so on. Moreover, the model contains 300-dimensional embeddings for the 2,389 emojis, and it has the same vector dimensions as that of Word2Vec model. The model also provides a suitable tool for calculating emoji's semantic similarity and the sense clustering between emojis and words.

We map the sequence word t^{s_i} and emoji e^{s_i} to its embedding $\mathrm{T}^{s_i} \in \mathbb{R}^{d_t}$ and $\mathrm{E}^{s_i} \in \mathbb{R}^{d_e}$, where d_t and d_e are the dimensions of word and emoji embedding. Finally, we input t^{s_i} and e^{s_i} to the next representation layer,and keep the embedded matrices during the training process.

3.2. Sentence representation based on BiLSTM with attention

After constructing the architecture of attention-based Bi-LSTM model, we investigate how the emoji information would contribute to the performance of personality prediction. We utilize the attention-based Bi-LSTM model to leverage the textual and emojis information at different semantics levels: (i) word-level attention-based Bi-LSTM; (ii) sentence-level attention-based Bi-LSTM.

3.2.1. Attention-Based Bidirectional LSTM Networks

The architecture of attention-based Bi-LSTM model is shown in Fig. 1. The model mainly comprised two parts: the Bi-LSTM networks and the attention mechanism. The Bi-LSTM model is adopted to determine which information should be kept or removed in our dataset. Then, the attention mechanism is used to highlight the important emojis and textual information, as they would influence the sentence and document representations in our methodology.

Initially proposed by Graves, the Bi-LSTM network extracts and incorporates contextual information from an input sequence using bidirectional LSTM layers [33]. By introducing the adaptive gating mechanisms, LSTM layer can produce a probability distribution for quantifying the concerns in each position of input, allowing the system to determine which information should be removed or retained. The Bi-LSTM model inputs the sequence into the forward LSTM layer while providing a reversed copy of the sequence into the backward LSTM layer. The hidden state vectors can take advantage of both past and future information to build sentence representation. Therefore, the Bi-LSTM networks are effective in extracting and incorporating contextual information, and widely used to tackle sequence tagging, sentence representation, and question answering [34,35].

Specifically, the architecture of Bi-LSTM contains a bidirectional LSTM layer in which each of the LSTM memory cell c contains three gates: ① input gate i; ② output gate o; ③ forget gate f. At each timestep t, each of the three gates is presented with the input x_t and the previous hidden states h_{t-1}. The input gate i_t specifies which information is added to the cell state. The output gate o_t specifies which information from the cell state is used as the output. The forget gate f_t denotes which information is removed from the cell state. The equations shown below can describe the update process of the memory cells from timestep $t-1$ to t.

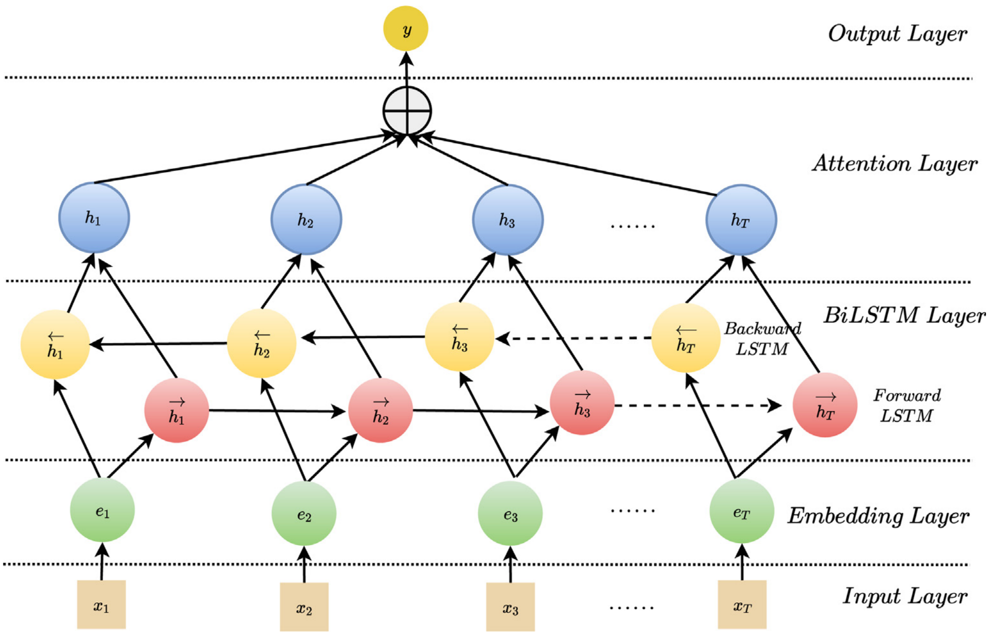

Fig. 1. Architecture of attention-based Bi-LSTM.

$$
\begin{aligned}
i_t &= \sigma(W_{ix}x_t + W_{ih}h_{t-1} + b_i) \\
f_t &= \sigma(W_{fx}x_t + W_{fh}h_{t-1} + b_f) \\
o_t &= \sigma(W_{ox}x_t + W_{oh}h_{t-1} + b_o) \\
g_t &= \tanh(W_{cx}x_t + W_{ch}h_{t-1} + b_c) \\
c_t &= f_t \odot c_{t-1} + i_t \odot g_t \\
h_t &= o_t \odot \tanh(c_t)
\end{aligned} \tag{1}
$$

where h_t denotes the output vector in each LSTM layer. $W_{ix}, W_{ih}, W_{fx}, W_{fh}, W_{ox}, W_{oh}, W_{cx}$ and W_{ch} represent the weight matrices, respectively. b_i, b_f, b_c, and b_o denote the bias vectors. $\odot$ is the element-wise multiplication, and σ is the element-wise sigmoid function.

As shown in Fig. 1, the Bi-LSTM layer inputs the sequence denoted by $x_t, [1, T]$ into the forward LSTM layer. Similarly, a reversed copy of the input sequence is provided as $x_t, [T, 1]$ to the backward LSTM layer. Then, the Bi-LSTM layer combines the output of the forward hidden layer and backward hidden layer, extracting the contextual information from our dataset. Eq. 2 is used to calculate the Bi-LSTM forward hidden state $\overrightarrow{h_t}$, the backward hidden state $\overleftarrow{h_t}$, respectively. Finally, we employ element-wise sum to combine the forward and backward hidden states as the output of the Bi-LSTM layer:

$$
h_t = \left[\overrightarrow{h_t} \oplus \overleftarrow{h_t}\right] \tag{2}
$$

Drawing on the previous work [36], word and emoji information contribute differently to the sentence representation. As a result, we utilize the attention mechanism to detect the importance of words and emojis in each sentence. As shown in Fig. 1, $H = [h_1, h_2, h_3, \ldots, h_T]$ is the input of the attention layer, y denotes the output vector and represents the sentence representation. tanh is the activation function, and the output y is formulated as follows.

$$
\begin{aligned}
M &= \tanh(H) \\
\alpha &= softmax(w^T M) \\
y &= H\alpha^T
\end{aligned} \tag{3}
$$

where $\alpha = [\alpha_1, \alpha_2, \alpha_3, \ldots, \alpha_T]$ is the attention weight. By using α, we can obtain the varying importance of each word in the input sequence.

Then, we adopt the attention-based Bi-LSTM model to learn the contextual information, and build the sentence and document representations in our methodology.

3.2.2. Word-level attention-based Bi-LSTM to combine text and emoji

As shown in Fig. 2, the word-level attention-based Bi-LSTM model concatenates emoji embedding and word embedding as $x^{s_i} = [T^{s_i}, E^{s_i}]$ at the word-level. We denote x^{s_i} as the input of the next representation layer.

In this model, we combine the emojis and words information as input, and build the sentence representation with the attention-based Bi-LSTM layer. First, we denote $x_t^{s_i}$ as the input vectors of the Bi-LSTM layers. Then, the following equations update f_t, i_t, and o_t in the LSTM layer at each timestep t.

$$\begin{aligned}
i_t &= \sigma\left(W_{ix}x_t^{s_i} + W_{ih}h_{t-1}^{s_i} + b_i\right)\\
f_t &= \sigma\left(W_{fx}x_t^{s_i} + W_{fh}h_{t-1}^{s_i} + b_f\right)\\
o_t &= \sigma\left(W_{ox}x_t^{s_i} + W_{oh}h_{t-1}^{s_i} + b_o\right)\\
g_t &= \tanh\left(W_{cx}x_t^{s_i} + W_{ch}h_{t-1}^{s_i} + b_c\right)\\
c_t &= f_t \odot c_{t-1} + i_t \odot g_t\\
h_t^{s_i} &= o_t \odot \tanh(c_t)
\end{aligned} \tag{4}$$

We calculate the forward hidden state $\overrightarrow{h_t^{s_i}}$ and encode the sentence s_i. The backward hidden state $\overleftarrow{h_t^{s_i}}$ encode the reversed sentence s_i in the Bi-LSTM layer. Then, we employ element-wise sum to combine the forward and backward hidden states as $h_t^{s_i} = \left[\overrightarrow{h_t^{s_i}} \oplus \overleftarrow{h_t^{s_i}}\right]$, to summarizes the information of emojis and texts at timestep t.

Then, we calculate the forward hidden state $\overrightarrow{h^{s_i}}$ encode the sentence s_i, the backward hidden state $\overleftarrow{h_t^{s_i}}$ encode the reversed sentence s_i in the Bi-LSTM layer. According to the architecture of attention-based Bi-LSTM model, we input $h_t^{s_i}$ to the attention layer, and denote the sentence representation $s_i^{u_i}$ as output. The mathematical representation of the sentence vector $s_i^{u_i}$ is

$$\begin{aligned}
u_t^{s_i} &= \tanh\left(W_{wh}h_t^{s_i} + b_w\right)\\
\alpha_t^{s_i} &= \frac{\exp\left(\left(u_t^{s_i}\right)^{\top} u_g\right)}{\sum_{l_i} \exp\left(\left(u_t^{s_i}\right)^{\top} u_g\right)}\\
s_i &= \sum_{l_i} \alpha_t^{s_i} u_t^{s_i}
\end{aligned} \tag{5}$$

where $u_t^{s_i}$ denotes the activation function of h^{s_i}, W_{wh} is a weight parameter that need to be trained. $\left(u_t^{s_i}\right)^{\top}$ is a transpose. u_g denotes the context vector is randomly initialized and jointly learned during the training process.

Finally, we utilize $u_t^{s_i}$ to quantify the importance scores of texts and emojis x^{s_i}, and obtain the sentence representation s_i with attention-based Bi-LSTM model.

3.2.3. Sentence-level attention-based Bi-LSTM to combine text and emoji

In this section, we propose sentence-level attention-based Bi-LSTM architecture to learn the document representation. Specifically, we apply word embeddings as the input to learn the sentence representation. Then, we concatenate sentence vectors and emoji embeddings as the new sentence input and train the document representation for the Bi-LSTM layers with the attention mechanism.

As illustrated in Fig. 3, we replace the input x^{s_i} with t^{s_i} in Bi-LSTM layers. The forward hidden state $\overrightarrow{h^{s_i}}$ and the backward hidden state $\overleftarrow{h^{s_i}}$ can be computed as follows:

$$\begin{aligned}
\overrightarrow{h^{s_i}} &= \overrightarrow{LSTM}(t^{s_i})\\
\overleftarrow{h^{s_i}} &= \overleftarrow{LSTM}(t^{s_i})
\end{aligned} \tag{6}$$

In addition, we concatenate the final hidden states vectors as $h^{s_i} = \left[\overrightarrow{h^{s_i}} \oplus \overleftarrow{h^{s_i}}\right]$ to feed into the attention layer. Since the attention-based Bi-LSTM layer also provides the different attention weights on the input, allowing our model to self-select which hidden vector should be prioritized. Then, we compute the importance values of information from sentences through Eq. 5 and obtain the sentence representation s_i'.

Therefore, we can calculate the new sentence representations se_i by concatenating the origin sentence vectors with the emoji vectors.

$$se_i = \left[s_i', e^{s_i}\right] \tag{7}$$

where the dimension of s_i' and $e_s i$ are the same 300, while $\mathrm{se_i} \in \mathbb{R}_{se}^{d}, d_{se} = d_{s_i'} + d_{e_s i} = 600$. Unlike the sentence representation of s_i, the modified sentence vectors of se_i includes the emoji vector and sentence vector at the sentence level.

3.3. Document representation based on attention mechanism

We obtain the document representation d_i using the same hierarchical architecture with the input of sentence representation s_i and se_i. the Bi-LSTM layer was employed and input the sentences $s^{u_i} = [s_1, s_2, ..., s_o]$ posted by ith user u_i, where o denotes the number of sentences.

From Eq. 6, we consider the sequence s_1 to s_o to be the inputs and assess the forward hidden state based on $\overrightarrow{h^{u_i}} = \overrightarrow{LSTM}(s^{u_i})$. Then, this study calculates the backward hidden state $\overleftarrow{h}^{u_i} = \overleftarrow{LSTM}(s^{u_i})$, which reads the sentences s^{u_i} from s_o to s_1, respectively. Afterward, we compute the hidden state h^{u_i} by combining the $\overrightarrow{h^{u_i}}$ with $\overleftarrow{h}^{u_i}$ in the following Equation.

$$h^{s_i} = \left[\overrightarrow{h^{s_i}} \oplus \overleftarrow{h}^{s_i}\right] \tag{8}$$

Let $H_{ui} = \left[h^{s_1}, h^{s_2}, h^{s_3}, \ldots, h^{s_o}\right]$ denote all hidden states by the i th user u_i, which will be fed into the attention layer. Then, we train the document representation u_i by a weighted sum of the different sentences based on weights and jointly learn them in the entire training process in the attention layer as follow:

$$\begin{aligned} M &= \tanh(H_{ui}) \\ \alpha &= softmax(w^T M) \\ r &= H_{ui}\alpha^T \end{aligned} \tag{9}$$

Then, we calculate the importance scores of different sentences for the similarity of α,and normalize the scores via the softmax function. Finally, s_i and se_i are entered into the next layer with the same architecture. Here, d denotes the document representations that can extract all sentence information in the attention layer. We utilize d as the input to the final layer for personality prediction.

3.4. Personality prediction

As the document representation d contains the information of words and emojis in each sentence, we employ d as the feature to predict individual Big Five personality traits.

In the present work, we utilize the LSTM layer and the fully connected layer to map the document representation d into five target classes, and choose Rectified Linear Units ($ReLU$) services as activation function of the fully connected layer. Subsequently, the softmax layer is constructed to determine the probability distribution of a personality for each dimension:

$$\begin{aligned} p &= W_c d + b_c \\ p_C &= \frac{\exp(p_c)}{\sum_{k=1}^{C} \exp(p_k)} \end{aligned} \tag{10}$$

where C denotes the number of personality classes. We employ p_C to represent the predicted probability of a particular class of personality. D stands for the document representation for all users. Since we are dealing with the task of personality recognition, the $cross\_entropy\_error$ is employed as the loss function for optimization during the training process. Finally, we obtain each dimension of the personality traits.

$$Loss = -\sum_{d \in D}\sum_{c=1}^{C} p_c^g(d) \cdot \log(p_c(d)) \tag{11}$$

4. Experiments

In this section, we describe in detail the dataset, baseline algorithms and parameter settings in the personality experiments.

4.1. Dataset

To evaluate the effectiveness of our proposed models, our experiments employ the public standard dataset of MyPersonality that provided by The Workshop on Computational Personality Recognition [14]. The dataset is collected from user generated content and profile data on Facebook, including user id, status, likes, social network relationships and other information. The benchmark labels of personality traits are obtained through the online Big Five personality test, the five classes of each user are shown as follow: EXT, NEU, OPN, AGR, and CON. In total, the dataset collected 9917 posted statuses containing words and emojis from 250 users. The average number of status per user is 39.67, and the average of the maximum status length per user is 70. As shown in Fig. 4, the frequently used emojis are the Heart 💖 and Smile Face 😄 and the

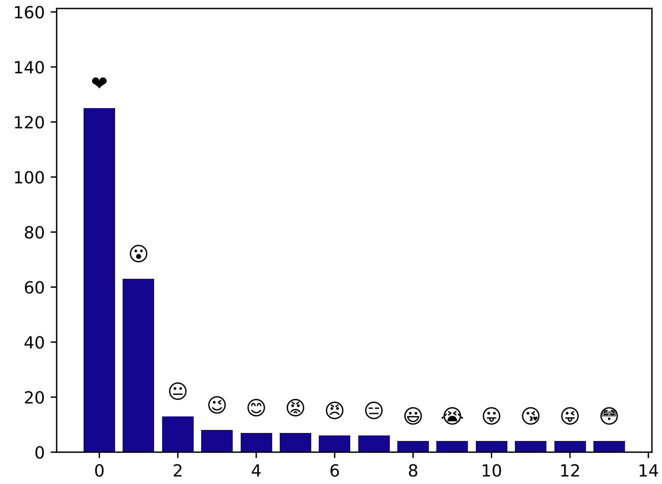

Fig. 4. Emojis description of all sentences in the dataset.

Pensive Face 😳, which represent happiness and sorrow emotions, respectively. Then, we feed each dimension of personality traits into the deep learning architectures.

In the preprocessing, all url, e-mail, hashtag and punctuation in the status are removed. Then, we adopt the spell correction and lemmatization methods to tackle the informal language in online social media. Moreover, all texts are converted into lowercase format, and all emojis are identified in the sentences. Afterward, we tokenize each status into a sequence of words and emojis, remove the users with fewer than 20 posted status to retain enough information for the neural networks.

4.2. Parameter Settings

This study employs *TensorFlow* to implement and evaluate our hierarchical model. The dataset is randomly divided into three parts: 80% training set, 10% validation set, and 10% testing set, which are used for model training, validation and testing in the experiments. *GridSearchCV* is a widely used hyperparameter tuning tool, which can help exhaustively search parameter values and find the optimal parameters with the best performance [37]. Therefore, The hyperparameters of our model are obtained through *GridSearchCV*: The total number of words and emojis per sentence is set to 70, and the size of user documents per user is 40. In the embedding layers, this study uses word vectors pre-trained Word2vec and Deepmoji to initialize the word embedding and emoji embedding vectors, the dimensions of word embedding and emoji embeddings are set to $d_t = d_e = 300$, which make sure word and emoji vectors can be combined at different semantic levels. The size of Bi-LSTM unit in each attention layer is 512, which produce 512*2 hidden state about forward and the backward hidden vectors. The dimension of FC layer was 100, and the batch size of our model is set to 8. In addition, dropout regularization was set 0.2 in each layer to address overfitting problems. The stochastic gradient descent Adam, with a learning rate of 0.001, is employed as the optimizer to minimize the *Cross_Entropy_with_Logits_Loss* to update all parameters during the training process. Afterward, this study adopts the metric of *Accuracy* and *F1-score*, which are typically used to evaluate the performance of personality recognition. Finally, the parameters are selected based on the performance of the validation dataset, and we utilize the trained model to predict the Big Five personality traits on the testing dataset.

4.3. Baseline algorithms

In recent years, *LIWC*, *Doc2vec* and *BERT* have become the common methods for personality recognition on social media [38–40]. Therefore, we choose the above mention models to benchmark our hierarchical models. In addition, we utilize *GridSearchCV* to train our baseline models and select the optimal parameter of the baseline models in the present study.

LIWC + SVM without Emoji: This method adopts the *LIWC* to obtain language statistical features, and use the Support Vector Machine (*SVM*) to predict personality traits. Also, the kernel of *SVM* is set to 'rbf' during the training process of the baseline model.

LIWC + SVM with Emoji: As the *LIWC* method fails to handle emoji information, we utilize text description to express the emoji symbols. For example, we map 😄 to its name *Smiling Face With Smiling Eye*. Both text and emoji information are regarded as the input, and subsequently entered into the *LIWC* model to predict the individual traits.

Doc2vec without Emoji: *Doc2vec* is a traditional unsupervised algorithm for large text presentations in the NLP fields, in which distributed memory models can be build from sentences, paragraphs, and documents [39]. Therefore, all of the words in our dataset are treated as an input to learn the document vector during the training process. We predict the personality traits by employing machine learning-based methods, including random forest(*RF*), *SVM*, K-nearest neighbors (*KNN*) algorithms. In addition, for the baseline model of *RF* and *KNN*, the optimal parameters for the baseline models are shown as: (1) max_depth = 5, max_features = 9, estimators = 80, bootstrap = Ture; (2) algorithm = 'auto', n_neighbors = 6, weights = 'uniform'.

Doc2vec with Emoji: As the *Doc2vec* model fails to handle the emoji embedding directly, this study converts emoji information of each sentence into textual descriptions, and input all text information to build the document representation. Then, we use *RF*, *SVM*, and *KNN* algorithms to evaluate the performance of personality recognition task.

BERT without Emojis: *BERT* achieves state-of-the-art performance in many NLP tasks, including question answering, named entity recognition, and natural language inference [40]. Therefore, this study employs *BERT-Base-Uncased* as the baseline model for personality recognition. The model can remove emoji symbols and combine all sentences of each user, and obtain those document representations. Finally, the labels of personality traits are predicted using the dense layer. The BERT model need not re-pretrain, all we have to do is Fine-Tuning according to the specific personality recognition tasks [41].

BERT with Emojis: this method converts emoji information of user sentences into textual descriptions, and combines all sentences into a document for each user. Then, this approach trains the document representations and predict individual personality traits using the *BERT* model.

Basic Attention BiLSTM Model without Emojis: To evaluate the important role of emoji information in sentences, we directly input textual information into the hierarchical model instead of combining the sentence and emoji vectors in the hidden layer of our model.

5. Results and discussion

We evaluate the performance of our models by calculating their *accuracy* and *F1-score*, and utilize the baseline approaches on the same dataset. The results of each dimension of the personality traits are shown in Tables 1 & 2. We then adopt the arithmetic average to measure the overall performance of each model [18]. As shown in Tables 1 & 2, the proposed models generally outperform these baseline methods.

Contribution of emoji information in different personality recognition models. The trends in Table 1 indicate that most models with emoji information perform better than those without emojis. In this study, we combine the emoji vector and word embeddings at different semantic levels using deep learning architectures. As a result, more information is utilized, suggesting higher accuracy. Compared with the basic hierarchical deep learning model, the average accuracy of our models improved by 8.5% and 15.3%, thus demonstrating the usefulness and contribution of emoji information in personality recognition tasks. In terms of the baseline models based on *LIWC*, *Doc2vec* and *BERT*, these models with emojis only convert emoji in sentences into textual descriptions, and then combine the original textual and emoji information as input to train sentence and document representations. In this study, the performance of the above models was better than that of the models without emojis.

The text replacement for emoji information helped to improve the performance of the baseline models. For instance, the average accuracy of the *BERT* model has improved by 6.8% due to the use of emoji information. Similarly, *LIWC* and *Doc2vec* based models also outperform these models without emoji information. As shown in Table 1, the results show that the performance of personality recognition can be improved when emoji information is added to the input.

Effectiveness of word-level and sentence-level emoji concatenation methods As the emoji information is helpful for personality recognition, we evaluate the two methods that introduce emojis at the word and sentence levels. More specifically, the basic attention-based Bi-LSTM model without emojis can achieve accuracies of 58%, 62%, 61%, 50% and 62% for personality recognition. Meanwhile, our word-level model combines the text with emoji at the word level while keeping

Table 1
Accuracy performance of personality recognition.

Model	*Classifier*	*Emoji*	*EXT*	*NEU*	*OPN*	*AGR*	*CON*	*Average*
LIWC	*SVM*	*No*	0.58	0.48	0.60	0.47	0.55	0.54
LIWC	*SVM*	*Yes*	0.61	0.58	0.64	0.55	0.51	0.58
Doc2vec	*RF*	*No*	0.60	0.45	0.50	0.53	0.44	0.51
Doc2vec	*SVM*	*No*	0.57	0.59	0.47	0.47	0.40	0.50
Doc2vec	*KNN*	*No*	0.54	0.53	0.46	0.50	0.43	0.45
Doc2vec	*RF*	*Yes*	0.61	0.65	0.60	0.56	0.53	0.59
Doc2vec	*SVM*	*Yes*	0.67	0.62	0.62	0.43	0.56	0.58
Doc2vec	*KNN*	*Yes*	0.62	0.59	0.50	0.61	0.50	0.56
BERT	*Dense*	*No*	0.61	0.55	0.58	0.52	0.60	0.58
BERT	*Dense*	*Yes*	0.64	0.62	0.63	0.58	0.65	0.62
Basic	*Attention BiLSTM*	*No*	0.58	0.62	0.61	0.50	0.62	0.59
Word Level	*Attention BiLSTM*	*Yes*	0.66	**0.75**	**0.69**	0.57	0.56	**0.64**
Sentence Level	*Attention BiLSTM*	*Yes*	**0.72**	0.64	**0.76**	**0.64**	**0.67**	**0.68**

Table 2
$F1$-*score* performance of personality recognition.

Model	Classifier	Emoji	EXT	NEU	OPN	AGR	CON	Average
LIWC	SVM	No	0.56	0.49	0.61	0.45	0.54	0.53
LIWC	SVM	Yes	0.60	0.53	0.61	0.51	0.57	0.57
Doc2vec	RF	No	0.59	0.59	0.48	0.54	0.48	0.54
Doc2vec	SVM	No	0.63	0.57	0.46	0.46	0.38	0.48
Doc2vec	KNN	No	0.50	0.59	0.47	0.33	0.40	0.46
Doc2vec	RF	Yes	0.64	0.61	0.59	0.53	0.52	0.60
Doc2vec	SVM	Yes	0.58	0.65	0.58	0.38	0.43	0.52
Doc2vec	KNN	Yes	0.60	0.65	0.36	0.57	0.52	0.53
BERT	Dense	No	0.60	0.54	0.52	0.54	0.53	0.55
BERT	Dense	Yes	0.62	0.52	0.55	0.59	0.56	0.57
Basic	Attention BiLSTM	No	0.63	0.61	0.58	0.31	0.53	0.53
Word Level	Attention BiLSTM	Yes	0.65	0.67	**0.76**	0.63	0.65	**0.67**
Sentence Level	Attention BiLSTM	Yes	**0.70**	0.71	0.67	**0.64**	**0.69**	**0.65**

the sequence information of sentences to learn the document representation. The accuracy results of our word-level model are 66%, 75%, 69%, 57% and 56%. Different from the word-level model, our sentence-level model introduces emoji embedding after obtaining sentence representation, and the accuracy ranges between 62% and 76%, which corresponds to the best results among the models. Even though the baseline models can achieve state-of-the-art performance in many NLP tasks, our model outperforms those models in personality recognition tasks in the present work. By utilizing combined emoji and textual information from online social media, the hierarchical architectures of our model can handle the personality recognition tasks at the word, sentence, and document levels. Interestingly, we have observed that the sentence-level model tends to lose some sequence information, but the document representation from each user's data can keep more information about the emojis, hence its high correlation with the Big Five personality traits. Interestingly, we have observed that the sentence-level model may lose some sequence information, but the document representation for each user keeps more information about the emojis, which highly correlated with the Big Five personality traits. Hence, the *accuracy* and $F1$-*score* value of the sentence-level model performs better than the word-level model. In general, the results validate the usefulness of our two attention-based Bi-LSTM models.

Validity of word-level and sentence-level attention-based Bi-LSTM models

Aiming to further demonstrate the effectiveness of our approaches, this study compares our models with the baseline models *LIWC*, *Doc2vec* and *BERT*. As shown in Table 1, the word-level and sentence-level attention-based Bi-LSTM models outperform the baseline methods. The *LIWC* method obtains statistical features to predict users' personality without emojis. Meanwhile, the *Doc2vec-based* model introduces the emoji information through the process of converting it to textual descriptions. The *BERT* model is added to capture document representations and personality traits from UGC, which obtains state-of-the-art performance in many NLP tasks in recent years [40]. As expected, the *accuracy* and $F1$-*score* of the *BERT* models perform better than those of *LIWC* and *Doc2vec*.

For the sentence-level model, the *NEU*, *EXT*, *OPN*, *AGR*, and *CON* have achieved state-of-the-art accuracy performance with the values of $72\%, 75\%, 76\%, 64\%, and 62\%$, respectively. In terms of the $F1$-*score*, our models obtain the avenge results of 67% and 65% in Table 2. Results demonstrate the effectiveness of our model in terms of precision and recall. As the calculated values of our personality labels are not balanced, the performance of total accuracy is higher than the $F1$-*score* in a specific dimension. Additionally, since *NEU* has little relationship with emoji usage patterns, which means that the dimensions of *NEU* are hard to handle only using emoji information in the present study [42]. Therefore, our models do not work very well with $F1$-*score* in the dimension of *NEU*. Nonetheless, the results show that our word-level and sentence-level models generally outperform the baseline models, thus validating the overall effectiveness of our proposed models.

6. Conclusion

This study proposed two attention-based Bi-LSTM models that incorporate textual and emoji information at different semantic levels to predict the Big Five personality traits. The word-level attention-based Bi-LSTM model concatenated both emoji and word embeddings while keeping the sequence information of each sentence. The sentence-level attention-based Bi-LSTM model combined the emoji embeddings with sentence vectors, with the purpose of learning the emoji information more effectively. The experiments on the baseline dataset indicated that our hierarchical models outperform the benchmark approaches.

Unlike previous studies, the work utilized the attention-based Bi-LSTM network to highlight the value of emoji information on social media. This study is the first attempt to introduce emoji embedding for personality recognition tasks. Particularly, we proposed two hierarchical neural models that incorporate emoji and textual information at different semantic levels. Our work can be further extended to handle tasks that require multisense embedding inputs. We add to the literature by highlighting the rich semantics of emoji symbols, and improve the performance of personality recognition task.

The findings indicate numerous ways for business firms to make better decisions about recommendation systems, open innovation platforms, and other managerial domains. We suggest that firms make management decisions depending on the personality traits of their users. Drawing on the research of Yang [43], people who share a particular personality trait tend to have similar interests and hobbies. Firms thus should suggest similar products and services to customers who share certain personality traits. In addition, the findings may help the companies in better understanding user behavior and improving the performance of recommendation systems. Because *openness* is a personality trait that positively linked to numerous new product ideas emanating from innovation process [5], the findings of our study will drive firms to build decision support systems as a means of identifying innovative users based on personality traits, thus facilitating product innovation and knowledge dissemination.

However, the present study has some limitations that can be addressed in future research. First, since millions of users have recently shared images and videos on social media, many researchers have investigated the relationship between the unstructured information and individual personality traits [44]. This study only investigated textual and emoji information in experiments, however, visual features should also be considered in future work to improve the performance of personality recognition task. Second, it would be worthwhile to investigate the performance of our models on additional social datasets, such as Twitter and Flickr. Finally, FastText [45], Graph Convolutional Network [46], and other sequence learning models [47] have recently achieved state-of-the-art performance in many NLP domains, and they may be applied to improve the performance of personality recognition tasks.

CRediT authorship contribution statement

Lixin Zhou: Writing - original draft, Conceptualization, Data curation, Software, Validation. **Zhenyu Zhang:** Methodology, Software, Writing - original draft. **Laijun Zhao:** Supervision, Writing - review & editing. **Pingle Yang:** Writing - review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgment

The authors would like to thank the editor and reviewers for the helpful comments and suggestions, which helped to improve the manuscript. This work was supported by grants from the National Social Science Foundation of China (No. 18AZD005), the National Natural Science Foundation of China (No.71874108), China Postdoctoral Science Fund (No.2021M692135), Shanghai Philosophy and Social Science Planning Project (No.2021BTQ003).

References

[1] A. Doniec, S. Lecoeuche, R. Mandiau, A. Sylvain, Purchase intention-based agent for customer behaviours, Inf. Sci. 521 (2020) 380–397.
[2] A. Azaria, J. Hong, Recommender systems with personality, in: Proceedings of the 10th ACM conference on Recommender systems, 2016, pp. 207–210.
[3] T.M. Entringer, J.E. Gebauer, J. Eck, W. Bleidorn, P.J. Rentfrow, J. Potter, S.D. Gosling, Big five facets and religiosity: Three large-scale, cross-cultural, theory-driven, and process-attentive tests, J. Personality Soc. Psychol. 120(6) (2021) 1662..
[4] G. Nobre, G. Demo, F. Scussel, E. Watanabe, The magical world of disney: building relationships with clients from the brand personality, Tourism Manage. Stud. 16 (1) (2020) 39–49.
[5] L. Zhou, Z. Zhang, L. Zhao, P. Yang, How user personality and information characteristics influence the creative information quality on open innovation platforms: an elaboration likelihood model, Kybernetes..
[6] D. Azucar, D. Marengo, M. Settanni, Predicting the big 5 personality traits from digital footprints on social media: A meta-analysis, Personality Individ. Differ. 124 (2018) 150–159.
[7] D. Watson, K. Stanton, S. Khoo, S. Ellickson-Larew, S.M. Stasik-O'Brien, Extraversion and psychopathology: A multilevel hierarchical review, J. Res. Pers. 81 (2019) 1–10.
[8] H.A. Schwartz, J.C. Eichstaedt, M.L. Kern, L. Dziurzynski, S.M. Ramones, M. Agrawal, A. Shah, M. Kosinski, D. Stillwell, M.E. Seligman, et al, Personality, gender, and age in the language of social media: The open-vocabulary approach, PloS One 8 (9) (2013) e73791.
[9] J. Golbeck, C. Robles, K. Turner, Predicting personality with social media, Vol. 27, Springer, 2011, pp. 253–262..
[10] J. Golbeck, C. Robles, M. Edmondson, K. Turner, Predicting personality from twitter, in: 2011 IEEE third international conference on privacy, security, risk and trust and 2011 IEEE third international conference on social computing, IEEE, 2011, pp. 149–156..
[11] M. Kosinski, D. Stillwell, T. Graepel, Private traits and attributes are predictable from digital records of human behavior, Proc. Nat. Acad. Sci. 110 (15) (2013) 5802–5805.
[12] R.T. Proyer, K. Brauer, Exploring adult playfulness: examining the accuracy of personality judgments at zero-acquaintance and an liwc analysis of textual information, J. Res. Pers. 73 (2018) 12–20.
[13] C. Stachl, Q. Au, R. Schoedel, et al, Predicting personality from patterns of behavior collected with smartphones, Proc. Nat. Acad. Sci. 117 (30) (2020) 17680–17687.
[14] D. Xue, L. Wu, Z. Hong, S. Guo, L. Gao, Z. Wu, X. Zhong, J. Sun, Deep learning-based personality recognition from text posts of online social networks, Appl. Intell. 48 (11) (2018) 4232–4246.
[15] Y. Mehta, S. Fatehi, A. Kazameini, C. Stachl, E. Cambria, S. Eetemadi, Bottom-up and top-down: Predicting personality with psycholinguistic and language model features, in: 2020 IEEE International Conference on Data Mining (ICDM), IEEE, 2020, pp. 1184–1189.
[16] S. Liu, R. Sun, To express or to end? personality traits are associated with the reasons and patterns for using emojis and stickers, Front. Psychol. 11 (2020) 1076.

[17] Z. Al-Halah, A. Aitken, W. Shi, J. Caballero, Smile, be happy:) emoji embedding for visual sentiment analysis, in: Proceedings of the IEEE/CVF International Conference on Computer Vision Workshops, 2019.
[18] J. Sun, Z. Tian, Y. Fu, J. Geng, C. Liu, Digital twins in human understanding: a deep learning-based method to recognize personality traits, Int. J. Comput. Integr. Manuf. (2020) 1–14.
[19] M. Fernández-Gavilanes, E. Costa-Montenegro, S. García-Méndez, F.J. González-Castaño, J. Juncal-Martínez, Evaluation of online emoji description resources for sentiment analysis purposes, Expert Syst. Appl. 184 (2021) 115279.
[20] J.-I. Biel, O. Aran, D. Gatica-Perez, You are known by how you vlog: Personality impressions and nonverbal behavior in youtube, in: Proceedings of the International AAAI Conference on Web and Social Media, Vol. 5, 2011, pp. 446–449..
[21] Y. Liu, J. Wang, Y. Jiang, J. Sun, J. Shang, Identifying impact of intrinsic factors on topic preferences in online social media: A nonparametric hierarchical bayesian approach, Inf. Sci. 423 (2018) 219–234.
[22] S. Han, H. Huang, Y. Tang, Knowledge of words: An interpretable approach for personality recognition from social media, Knowl.-Based Syst. 194 (2020) 105550.
[23] N. Majumder, S. Poria, A. Gelbukh, E. Cambria, Deep learning-based document modeling for personality detection from text, IEEE Intell. Syst. 32 (2) (2017) 74–79.
[24] Y. Mehta, N. Majumder, A. Gelbukh, E. Cambria, Recent trends in deep learning based personality detection, Artif. Intell. Rev. 53 (4) (2020) 2313–2339.
[25] Q. Bai, Q. Dan, Z. Mu, M. Yang, A systematic review of emoji: Current research and future perspectives, Front. Psychol. 10 (2019) 2221.
[26] D. Rodrigues, M. Prada, R. Gaspar, M.V. Garrido, D. Lopes, Lisbon emoji and emoticon database (leed): Norms for emoji and emoticons in seven evaluative dimensions, Behav. Res. Methods 50 (1) (2018) 392–405.
[27] G.W. Tigwell, D.R. Flatla, Oh that's what you meant! reducing emoji misunderstanding, in: Proceedings of the 18th international conference on human-computer interaction with mobile devices and services adjunct, ACM, 2016, pp. 859–866.
[28] F. Barbieri, F. Ronzano, H. Saggion, What does this emoji mean? a vector space skip-gram model for twitter emojis, in: Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC, 2016, pp. 3967–3972.
[29] B. Eisner, T. Rocktäschel, I. Augenstein, M. Bošnjak, S. Riedel, emoji2vec: Learning emoji representations from their description, arXiv preprint arXiv:1609.08359..
[30] S. Wijeratne, L. Balasuriya, A. Sheth, D. Doran, Emojinet: An open service and api for emoji sense discovery, in: Eleventh International AAAI Conference on Web and Social Media, Vol. 11, 2017..
[31] B. Felbo, A. Mislove, A. Søgaard, I. Rahwan, S. Lehmann, Using millions of emoji occurrences to learn any-domain representations for detecting sentiment, emotion and sarcasm, arXiv preprint arXiv:1708.00524..
[32] F. Hemmatian, M.K. Sohrabi, A survey on classification techniques for opinion mining and sentiment analysis, Artif. Intell. Rev. 52 (3) (2019) 1495–1545.
[33] A. Graves, A. Mohamed, G. Hinton, Speech recognition with deep recurrent neural networks, in: 2013 IEEE international conference on acoustics, speech and signal processing, IEEE, 2013, pp. 6645–6649..
[34] J. Wang, An intelligent computer-aided approach for atrial fibrillation and atrial flutter signals classification using modified bidirectional lstm network, Inf. Sci. 574 (2021) 320–332.
[35] R. Catelli, V. Casola, G. De Pietro, H. Fujita, M. Esposito, Combining contextualized word representation and sub-document level analysis through bi-lstm+crf architecture for clinical de-identification, Knowl.-Based Syst. 213 (2021) 106649.
[36] Z. Yang, D. Yang, C. Dyer, X. He, A. Smola, E. Hovy, Hierarchical attention networks for document classification, in: Proceedings of the 2016 conference of the North American chapter of the association for computational linguistics: human language technologies, 2016, pp. 1480–1489.
[37] Y. Li, Z. Su, J. Yang, C. Gao, Exploiting similarities of user friendship networks across social networks for user identification, Inf. Sci. 506 (2020) 78–98.
[38] G. Farnadi, S. Zoghbi, M.-F. Moens, M. De Cock, Recognising personality traits using facebook status updates, in: Seventh International AAAI Conference on Weblogs and Social Media, Vol. 7, 2013, pp. 14–18..
[39] D. Jeon, J.M. Ahn, J. Kim, C. Lee, A doc2vec and local outlier factor approach to measuring the novelty of patents, Technol. Forecast. Soc. Chang. 174 (2022) 121294.
[40] V. Lynn, N. Balasubramanian, H.A. Schwartz, Hierarchical modeling for user personality prediction: The role of message-level attention, in: Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics, 2020, pp. 5306–5316.
[41] A. Adhikari, A. Ram, R. Tang, J. Lin, Docbert: Bert for document classification, arXiv preprint arXiv:1904.08398..
[42] W. Li, Y. Chen, T. Hu, J. Luo, Mining the relationship between emoji usage patterns and personality, in: Proceedings of the International AAAI Conference on Web and Social Media, Vol. 12, 2018..
[43] H.-C. Yang, Z.-R. Huang, Mining personality traits from social messages for game recommender systems, Knowl.-Based Syst. 165 (2019) 157–168.
[44] V. Moscato, A. Picariello, G. Sperlí, An emotional recommender system for music, IEEE Intell. Syst. 36 (5) (2020) 57–68.
[45] S. Minaee, N. Kalchbrenner, E. Cambria, N. Nikzad, M. Chenaghlu, J. Gao, Deep learning–based text classification: A comprehensive review, ACM Computing Surveys (CSUR) 54 (3) (2021) 1–40.
[46] S. Zhang, H. Tong, J. Xu, R. Maciejewski, Graph convolutional networks: a comprehensive review, Computat. Social Networks 6 (1) (2019) 1–23.
[47] Q. Liu, J. Lu, G. Zhang, Shen, Domain-specific meta-embedding with latent semantic structure, Inf. Sci. 555 (2021) 410–423.