2020 IEEE 22nd International Conference on High Performance Computing and Communications; IEEE 18th International Conference on Smart City; IEEE 6th International Conference on Data Science and Systems (HPCC/SmartCity/DSS)

2020 IEEE 22nd International Conference on High Performance Computing and Communications; IEEE 18th International Conference on Smart City; IEEE 6th International Conference on Data Science and Systems (HPCC/SmartCity/DSS) | 978-1-7281-7649-9/20/$31.00 ©2020 IEEE | DOI: 10.1109/HPCC-SmartCity-DSS50907.2020.00158

A Novel Developer Portrait Model based on Bert-Capsule Network

1st Pengyu Yu
College of Computer and Information
Hohai University
Nanjing, China
jadesperwalker@qq.com

2nd Yirui Wu*
College of Computer and Information
Hohai University
Nanjing, China
wuyirui@hhu.edu.cn

3nd Benze Wu
College of Computer and Information
Hohai University
Nanjing, China
1277550118@qq.com

***Abstract*—In order to ensure code quality, it's necessary to construct portraits for developers, which could analyze their behavior to provide personalized programming suggestions. However, most of the existing developer portrait algorithms only use global features and ignore local features extracted from log texts, which leads to the lack of comprehensive personality analysis. To solve this problem, the proposed method proposes a novel developer portrait model, which could describe developers' programming styles more accurately with both global and local information extracted from texts. The proposed model firstly collects the log data produced in the process of continuous integration development. Afterwards, the proposed method proposes the personality portrait model based on BERT-Capsule network, which successfully combines global semantic features and local emotional features. The experimental results show that the proposed BERT-Capsule model can effectively extract the contextual information and the local emotional information of the text, thus improving classification performance of the developer portrait model.**

***Index Terms*—Bert-Capsule Network; Developer Portrait Model; Global Information;**

I. Introduction

With the increasing scale of software products, more and more enterprises adopt the continuous integration development model. The importance of code quality becomes more and more prominent. To ensure code quality, the companies often use continuous integration tools to test their code, and rarely analyze the causes of developers' errors.

In order to provide personalized advice to developers, it is necessary to analyze their personality characteristics and propose different suggestions according to different personalities. However, there are still two difficulties in the study of developer portrait. One is the lack of data sets for building developer portraits from multiple dimensions. Secondly, the analysis of the causes of the defects in the existing portrait models is not solid enough. At the same time, the existing personality classification algorithms often directly extract the global features and ignore the local emotional features in the text, which leads to the lack of comprehensive personality analysis.

To solve the above problems, this thesis collects the log data in the process of continuous integration development, analyzes the developer's work activity and attitude through statistics and clustering and proposes the personality portrait model BERT-Capsule combining global semantic features and local emotional features, so as to build the developer portrait model. The main research work and contents are as follows:

(1) For the lack of constructing developer portrait dataset from multiple dimensions, this thesis collects and marks the log data in the process of continuous integration.

(2) In view of the incomplete analysis of the causes of defects and the neglect of partial emotional features of texts by existing personality classification algorithms, the proposed method proposes the BERT-Capsule character portrait model. The local emotional features extracted from BERT model are fused with global features extracted from BGRU-Capsule model to obtain text emotional features with rich context information. The softmax classifier is used to output personality classification results.

II. Related Work

User information is mined and calculated, and a set of tags that can describe user characteristics is abstracted. use The core of household portraits is through data mining [1] And other technologies, feature extraction of user information from multiple dimensions [2]. The label is usually a highly refined feature identification specified by man, which can be based on The label abstracts the whole picture of a user. The main job of user portrait is to analyze and understand users So that the platform can provide personalized services based on user behavior. When constructing user portraits, It is mainly divided into the construction of portrait model based on statistical clustering and the construction of portrait model based on machine learning algorithm. Kind of way. We believe user portrait model could be improved by quantity of novel methods, such as edge computing [3], data-driven intelligence [4] or big data processing [5].

The construction of the portrait model based on statistical clustering mainly uses numerical statistics, probability distribution, and simple clustering Method to conduct statistical analysis on the collected data, and then obtain label information that can describe the user. Syskill Et

978-1-7281-7649-9/20/$31.00 ©2020 IEEE
DOI 10.1109/HPCC-SmartCity-DSS50907.2020.00158

al [6].By collecting users' satisfaction with the website pages, the user's interest is constructed based on statistics model. Chen Yuejuan and others [7].By tracking the behavior of users using the library, the number of logins and browsing time The length, visits, and number of document downloads were counted, and personalized reading services were tailored for readers. Wang Na and others the user's portrait is constructed by counting the user's ratings and labeling of film and television works. Then use the portrait to investigate the user's satisfaction with the film and television works, and get the user's favorite film and television categories, and finally Recommend to users their favorite works.

The construction of the portrait model based on the machine learning algorithm [8] will first perform feature extraction on the collected data, and then use the existing label information to predict the unknown label to obtain the category information to which the user belongs, and finally Predict user behavior information. Torres Valencia et al [9] use support vector machine to analyze the user's emotional characteristics Analysis. Kuzma et al [10]extract user preference features based on neural network model. Muelle et al. pass more This feature of word structure differentiates the gender of Twitter username information. Slanzi et al [11] use logic Logistic regression algorithm analyzes the content browsed by users to help the website provide more content that users prefer. Lu et al [12]based on the fuzzy C-mean algorithm, from the temperature sensitivity, electricity price sensitivity and electricity consumption stability three Dimension analyzes the user's electricity consumption behavior and constructs a portrait of the user's electricity consumption.

Chen Ye et al [13]proposed that when constructing user portraits, the timeliness and timeliness of user data should also be considered.There are two aspects to the dynamics of portrait data. User's behavior will change over time.When drawing a portrait, if you select some older user data and do not update the portrait, then draw the value of the image is difficult to reflect. Therefore, when constructing user portraits, you need to choose an appropriate time period to ensure the validity of the portrait. Franca et al [14]Based on research in psychology, it emphasizes user personalization

III. The Proposed Method

In order to improve the work activity and attitude of developers, it is necessary to analysis. According to the different personality characteristics of the developers, develop from two aspects: work activity and work attitude Personnel provide personalized suggestions, so that developers can better accept the suggestions, make timely changes, and improve the generation Code quality.

In order to analyze the personality of developers, this article chooses a five-factor personality model as the analysis theory Based on the analysis and modeling of the log feedback information of the developers. Due to the short text in the log feedback There are mostly sentences. When using traditional word vector models to extract the semantic features of the text, the vector space will be sparse The problem.

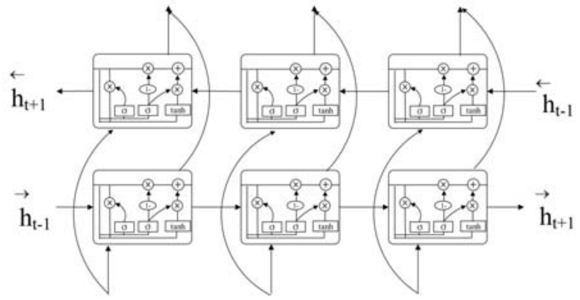

Fig. 1. BERT-Capsule model framework.

In response to the above problems, this paper proposes a personality sentiment classification model BERT-Capsule for short texts, The BERT word vector model is selected to extract the local emotional semantics in the feedback information, and the BGRU model and The Capsule model is combined to extract the global semantic features in the feedback information, and then the local features and the global The local features are integrated to obtain a more comprehensive emotional semantic feature of the short text, and then to the developer Character traits.

A. Structure Design

Since the five-factor personality model contains the positive and negative tendencies of five personalities, this article proposes 10 classified BERT-Capsule model, used to judge whether the input feedback information belongs to a certain personality positive Tendency or negative tendency. The BERT-Capsule model is shown in Figure 1, which mainly contains local semantic features There are three parts: extraction, global semantic feature extraction, feature fusion and classification.

(1) Local semantic feature extraction: In order to increase the weight of emotional words and obtain more obvious emotion Sensing characteristics, this article first based on the BERT model, introducing sentiment dictionary and masked word prediction tasks, will continue The integrated log feedback information is input into the BERT model and retrained to obtain the BERT emotional model. Then, use the trained BERT emotional model to map the log feedback information into a word vector matrix, which serves as the bureau Partial semantic features;

(2) Global semantic feature extraction: In global semantic feature extraction, first input the word vector matrix In the BGRU model, the contextual semantics of the short text are acquired, and then the extracted feature information is input Enter the Capsule network to further extract semantic information such as the word position and

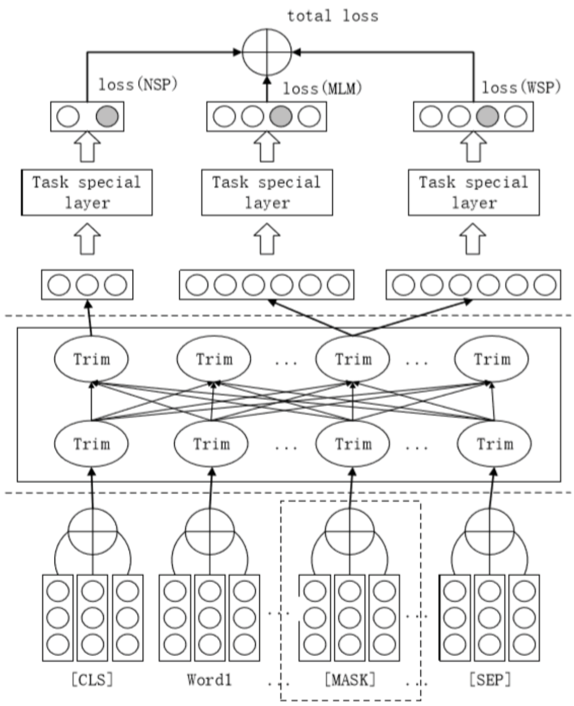

Fig. 2. BERT affective model structure.

syntactic structure of the short text. Obtain global semantic features;

(3) Feature fusion and classification: Firstly, the global semantic features and local semantic features are fused to obtain To the more comprehensive emotional semantic information of the short text, and then input the fused feature information through the fully connected layer Go to the softmax classifier to get the final classification result. Figure 1 BERT-Capsule model architecture

B. Local Semantic Feature Extraction based on BERT Emotion Model

In the local feature extraction module, this article first uses the Jieba word segmentation tool and NLTK word segmentation tool to preprocess the log feedback information. After the processed log feedback information data is obtained, the text information needs to be mapped to a high-dimensional vector space to obtain a word vector that can describe the text information. Since the log feedback information is mainly short text, traditional word vector models such as Word2Vec and Glove are prone to loss of emotional semantics and inability to deal with polysemous words when extracting semantic features of short texts, while the BERT model is extracting the semantics of short texts. When characteristic, it has a good effect and can solve the situation of ambiguity. Therefore, the BERT model can be selected to extract local semantic features.

In response to the above problems, this paper uses the BERT word vector model to map the word vector of the log feedback information, and extract the emotional features in the text information as local features. However, in the log feedback information, the proportion of emotional words is relatively small and the hidden emotional features are difficult to be noticed by the model. To this end, based on the BERT model, this article first replaces the implicit emotional information in the log feedback information with words with relatively single and distinct emotions through the emotional dictionary, and then selects all emotional words to form an emotional word set, which is added when calculating the loss Larger sample weight. Among them, this article selects HowNet as the external emotion dictionary introduced. At the same time, this paper proposes to introduce the Covered Word Sentiment Prediction Task Model (WSP) into the BERT model, and use the semantic information of the context to predict the emotion of the covered word to obtain more emotionally biased semantic features.

The model structure is shown in Figure 2. The model is mainly composed of three parts: input layer, Transformer layer and task layer. In the input layer, the results of word embedding, position embedding, and sentence embedding are added to obtain the final input. In the Transformer layer, the semantic features of the context are extracted from the input word vector. In the task layer, the BERT model mainly uses the next sentence prediction task model (NSP), the masked word language task model (MLM) and the masked word sentiment prediction task model (WSP) to extract text emotion features. Among them, the NSP task model and the MLM task model are the basic models for extracting text features.

In order to be able to give more weight to the emotional words in the sentence, this article retains the NSP task model, modifies the MLM task model, and uses the WSP task model as an auxiliary task of the MLM task model. The final loss of the model can be obtained by equation (1).

$$\begin{aligned} loss(total) = \lambda_1 * (loss(NSP) + loss(MLM)) \\ + \lambda_2 * loss(WSP) \end{aligned} \tag{1}$$

For the WSP task model, the sentiment of the masked word needs to be labeled first, and then combined with the semantic features of the context, the feature of the masked word is classified using the softmax function to classify the sentiment tendency of the masked word:

$$P_S\left(a \mid x_{[MASK]}\right) = \operatorname{softmax}\left(W_S^T \cdot x_{[MASK]}\right) \tag{2}$$

Among them, $P_S\left(a \mid x_{[MASK]}\right)$ represents the probability that the masked word's feature vector $x_{[MASK]}$ belongs to category a, and W_S^T is a random initialization matrix. Therefore, the loss of all masked words to be predicted is:

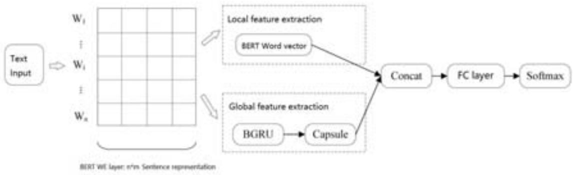

Fig. 3. BGRU model structure.

$$\text{loss}(WSP) = -\sum_{i}\sum_{a} L\left(x^{i}_{[MASK]}, a\right) \log\left(P^{i}_{S}\left(a \mid x^{i}_{[MASK]}\right)\right) \quad (3)$$

After constructing the BERT emotion model, the feedback information needs to be mapped into word vector form through the BERT emotion model. The output of the BERT emotion model mainly has two forms: character-level and sentence-level. The character-level vector output form corresponds to a vector for each character in the text, while the sentence-level uses the vector of the first special word case [CLS] in BERT as the vector of the entire sentence. In order to analyze the sentiment tendency in the log feedback information, this paper uses character-level vector representation. In the BERT model, the [CLS] word example mark is set at the beginning of the sentence, the [SEP] word example mark is set at the end of the sentence, and the [MASK] mark is used in the sentence to hide some words. For each word in the sentence, the BERT model performs three operations: word embedding, position embedding, and sentence embedding, and then splicing the result vector to obtain the BERT word vector.

C. Global semantic feature extraction based on BGRU-Capsule model

In the global semantic feature extraction module, this paper inputs the word vector of the log feedback information into BGRU-In the Capsule model, the context information is extracted as a global semantic feature.

When the traditional GRU model processes text sequences, it mainly uses forward reading to learn textThe above features of the text, but did not consider the information features below the text, and the BGRU model used in this articleThe type can be read bidirectionally, combining the forward information and backward information of the text,The context information of the column is extracted. Although the CNN model can extract text through convolutional layers and pooling layers information characteristics, but the operation of taking the maximum value or taking the average value is rough, which is easy to cause text information characteristics.In the case of loss, this paper uses the Capsule network to replace the CNN model. In the Capsule model,The feature extraction is mainly implemented in the way of "vector input-vector output", and each output vector represents a class,features, and can pass the current calculation results to the next layer through the dynamic routing mechanism, so as to achieve theWhen extracting the semantic information, location information, and grammatical structure of words, it can reduce the loss of features and improve the accuracy of the feature information. Therefore, this article combines the BGRU model with the Capsule model to extract the global emotional semantic features of short texts. The model structure of BGRU is shown in Figure 3.

For time t, the forward GRU model can obtain the above information at time t, and can output the current state $\overrightarrow{h_t}$, and the backward GRU model can obtain the following information, and output the current state $\overleftarrow{h_t}$. This article outputs the results of BGRU through connection. For a given n-dimensional word vector $(x_1, x_2, \ldots, x_n)$, the output at time t is $\text{BG} = \{bg_1, bg_2, \ldots, bg_n\}$, where $bg_t = \left(\overrightarrow{h_t}, \overleftarrow{h_t}\right)$. For $\overrightarrow{h_t}$ and $\overleftarrow{h_t}$ can be obtained by formula (4) and formula (5).

$$\overrightarrow{h_t} = \sigma\left(W_{x\vec{h}} x_t + W_{\vec{h}\vec{h}} \vec{h}_{t-1} + b_{\vec{h}}\right) \quad (4)$$

$$\overleftarrow{h_t} = \sigma\left(W_{x\overleftarrow{h}_t} + W_{\overleftarrow{h}\overleftarrow{h}} \overleftarrow{h}_{t-1} + b_{\overleftarrow{h}}\right) \quad (5)$$

Among them, $\overrightarrow{h_t}$ represents the forward output of BGRU,$\overleftarrow{h_t}$ represents the backward output of BGRU, σ is the activation function, W is the weight matrix, and b is the bias vector.

D. Feature Fusion and Classification

In order to make a more accurate classification of the developer's personality, this paper combines the local emotional semantic features obtained from the BERT model with the global semantic features obtained from the BGRU-Capsule model, and comprehensively considers the local semantics and global semantic information in the log feedback information.

In this paper, the output vector of the BERT model in the local feature is spliced and the output vector of the BGRU-Capsule model again through the splicing method. A fully connected layer inputs it into the final classification vector , and outputs the classification result through the softmax classifier. Among them, the calculation formula of is shown in formula (6).

$$V_{con} = Dense(Concat(V_b ert, V_b g)) \quad (6)$$

E. Training methods

In order to obtain the emotional semantic features of log feedback information, this paper adopts synchronous training BERT-Capsule method.The first is to train the Bert model, and then to train the whole BERT-Capsule model.

In the training of the best model, the MLM task model in the Bert model is modified, and then the WSP task model is introduced as the auxiliary task of the MLM task model. Finally, the manually annotated log feedback information is input into the Bert emotion model for training, and the Bert emotion model is obtained. After getting the Bert emotion model, the original log feedback information is vectorized by the Bert emotion model, and the Bert word vector with emotional features is obtained, and then it is input into the Bert capsule model for further training.

In the training of the Bert capsule model, the local feature extraction and global feature extraction of the Bert word vector are firstly carried out. Because of the emotional semantic features of the best word vector, it can be directly used as the local feature in the local feature extraction module. In the global feature extraction module, the global features of the Bert word vector can be extracted by bgru capsule model. Then, local features and global features are fused. Finally, the soft Max classifier is used to train the Bert capsule model.

IV. Experiments

A. Data set settings

In order to evaluate the performance of the BERT-Capsule model, this article selects IMDB sentiment data set and continuous integration log information data set (CILDDB) for experiments. Since the CILDDB data set has unbalanced categories before expansion, this article uses the expanded sample data of the SMOTE algorithm for the personality analysis experiments of developers. Among them, 80% of the data is selected as the training set to train the model, 10% of the data is selected as the test set to test the performance of the trained model, and 10% of the data is selected as the validation set to adjust the experimental parameters of the model.

The IMDB data set is a movie review data set, which contains 50,000 pieces of movie review information with obvious emotional tendencies, and is widely used for sentiment analysis in natural language. The IMDB data set is mainly divided into two categories: positive reviews and negative reviews. The number of positive reviews is 25,000 and the number of negative reviews is 25,000. You can select 20,000 positive reviews and 20,000 negative reviews from the IMDB data set as the training set, 2500 positive reviews and 2500 negative reviews as the test set, and 2500 positive reviews and 2500 negative reviews as the validation set. In the CILDDB data set, it can be divided into five categories of personality data: extraversion, openness, agreeableness, neuroticism, and conscientiousness. Each category contains two situations: positive affective tendencies and negative affective tendencies.

Among them, from the two emotional tendencies of each personality, 10*1743 are selected as the training set, 10*218 are selected as the test set, and 10*218 are selected as the verification set.

TABLE I
Experimental results of different parameters of the BERT-Capsule model under the IMDB data set.

λ_1	λ_2	m	Precision(%)	Recall(%)	F1(%)
1.00	0.00	1.00	85.68	84.36	85.01
0.95	0.05	1.50	86.82	85.49	86.15
0.90	**0.10**	**2.00**	**87.46**	**87.90**	**87.68**
0.85	0.15	3.00	86.96	85.61	86.28
0.80	0.20	5.00	86.53	85.24	85.88
0.70	0.30	10.00	85.12	83.83	84.47

TABLE II
Experimental results of different parameters of the BERT-Capsule model under the CILDDB data set.

λ_1	λ_2	m	Precision(%)	Recall(%)	F1(%)
1.00	0.00	1.00	69.42	70.73	70.07
0.95	**0.05**	**1.50**	**73.25**	**73.85**	**73.55**
0.90	0.10	2.00	72.86	73.21	73.03
0.85	0.15	3.00	70.91	71.44	71.17
0.80	0.20	5.00	68.73	70.38	69.55
0.70	0.30	10.00	64.17	69.80	66.87

B. experiment

In order to get the best effect of the BERT-Capsule model under the IMDB data set, this article selects the IMDB verification set to design multiple sets of experiments to obtain the optimal parameter values of the BERT-Capsule model, that is, the loss weights of the task layer λ_1 and λ_2 and Mask the m value of the word model. Among them, the value of m is a constant greater than 1, and the sum of the value of λ_1 and λ_2 is 1. When the value of λ_1 is 1.00, the value of λ_2 is 0.00, and the value of m is 1.0, it represents the original BERT model. The experimental results of the BERT-Capsule model under different parameters are shown in Table 1.

It can be found that the improved method proposed in this article, as the masked word prediction task weight λ_2 and the masked word model m increase, the accuracy, recall and F1 value of the model will be improved, which means the masked word prediction task The loss also decreased. When λ_1 =0.90, λ_2 =0.10, m=2.00, the experimental results of the model under all indicators have reached the best. However, when the value of ??2 is greater than 0.10 and the value of m is greater than 2.00, the accuracy, recall, and F1 value of the model will gradually decrease, indicating that the loss of the masked word prediction task increases. When λ_1 =0.70, λ_2 =0.30, m=10.00, the effect of the model is lower than that of the original BERT model, which may be due to overfitting. Therefore, this article chooses λ_1 = 0.90, λ_2 = 0.10, and m=2.00 as the training parameters of BERT-Capsule under the IMDB data set.

Experimental results of different parameters of the BERT-Capsule model under the CILDDB data set. In order to get the best effect of the BERT-Capsule model on the CILDDB data set, this article conducted multiple

TABLE III
EXPERIMENTAL RESULTS OF MACHINE LEARNING ALGORITHMS UNDER IMDB DATA SET.

Method	Precision(%)	Recall(%)	F1(%)
LR [15]	74.62	73.48	74.04
SVM [16]	75.36	74.49	74.92
RF [17]	79.35	77.83	78.58
GDBT	78.74	80.58	79.64
Stacking	81.49	81.61	81.54
BERT-Capsule	**87.46**	**87.90**	**87.68**

TABLE IV
EXPERIMENTAL RESULTS OF MACHINE LEARNING ALGORITHMS UNDER CILDDB DATA SET.

Method	Precision(%)	Recall(%)	F1(%)
LR [15]	55.61	54.64	55.12
SVM [16]	56.38	55.27	55.82
RF [17]	58.74	59.62	59.18
GDBT	59.48	58.91	59.19
Stacking	61.17	62.53	61.84
BERT-Capsule	**73.25**	**73.85**	**73.55**

experiments on the CILDDB validation set to obtain the best effect of the BERT-Capsule model on the CILDDB data set. The accuracy comparison of the experimental results is shown in Table 2.

It can be found that when λ_1 =0.95, λ_2 =0.05, and m=1.50, the results of each indicator of the model in the CILDDB data set have reached the best. With the increase of λ_2 and m, the accuracy rate, recall rate and F1 value of the model begin to decrease to varying degrees. When λ_1 =0.80,λ_2 =0.20, m=5.00, the accuracy, recall and F1 value of the model are lower than the experimental results of the original model, which should be due to overfitting when judging the personality tendency. Therefore, you can choose λ_1 =0.95, λ_2 =0.05, m=1.50 as the experimental parameters of the BERT-Capsule model in the CILDDB data set.

C. Comparison of classification performance with existing traditional machine learning algorithms

Experimental results of machine learning algorithm in IMDB dataset. In order to verify the performance of the Bert capsule model.In this paper, we compare the BER capsule model with the existing traditional machine learning algorithms. For the existing traditional machine learning algorithms, this paper selects the logical regression algorithm (LR), support vector machine algorithm (SVM), random forest algorithm (RF) and gradient lifting decision tree algorithm (GDBT), and integrates the above traditional machine learning algorithms by stacking algorithm to obtain the stacking integrated classification algorithm.The classification results of traditional machine learning algorithm and Bert capsule algorithm in IMDB dataset are shown in table 3.

TABLE V
EXPERIMENTAL RESULTS OF THE DEEP LEARNING MODEL UNDER THE IMDB DATASET.

model	Precision(%)	Recall(%)	F1(%)
Capsule-A [18]	86.54	86.10	86.32
Capsule-B [18]	87.45	86.67	87.06
CNN-multichannel [19]	86.38	86.14	86.26
LR-Bi-LSTM [20]	86.74	86.38	86.56
CSVM [21]	85.26	86.68	84.97
LSTM-CNN [22]	86.91	85.88	86.39
MLCNN	**89.25**	**88.57**	**88.91**
BERT-Capsule	87.46	87.90	87.68

It can be found that under the IMDB data set, the RF algorithm and gdbt algorithm in the integration algorithm are effective.It is better than SVM algorithm and LR algorithm. The integrated stacking algorithm is better than other traditional machine learning algorithms in accuracy, recall and F1 value, which shows that stacking algorithm can give full play to the advantages of each basic model. Compared with the traditional model algorithm, the proposed model is better than the traditional model in accuracy rate, recall rate and F1 value, which shows that the best capsule model has a good ability in the polarity classification of text emotion.

Experimental results of machine learning algorithm in CILDDB dataset. In order to further evaluate the performance of the BERT-Capsule model in the CILDDB dataset, this paper compares the BERT-Capsule model with the traditional machine learning algorithm again. The experimental results are shown in table 4.

It can be found that compared with SVM algorithm and LR algorithm, ensemble learning is better in CILDDB dataset.The results of RF algorithm and GDBT algorithm are obviously improved in accuracy, recall rate and F1 value. At the same time, in terms of accuracy rate, recall rate and F1 value, the classification effect of Stacking model is better than RF and GDBT ensemble learning algorithm, which shows that stacking algorithm can effectively make up for the deficiency of basic model in some aspects. The proposed model is superior to the traditional machine learning algorithm in accuracy, recall rate and F1 value, which shows that the model is significant in the classification of personality and emotion of log feedback information.

D. Comparison of classification performance with existing deep learning algorithms

Experimental results of deep learning model in IMDB dataset. In order to further verify the proposed method.In this paper, the effectiveness of the best capsule model is compared with Capsule-A, capsule-b, CNN multichannel, LR Bi LSTM, CSVM, lstmcnn and MLCNN in the IMDB dataset. The classification performance of each model in IMDB dataset is shown in table 5, F1 value comparison is shown in Figure (a), and loss comparison is shown in Figure 3 (b).

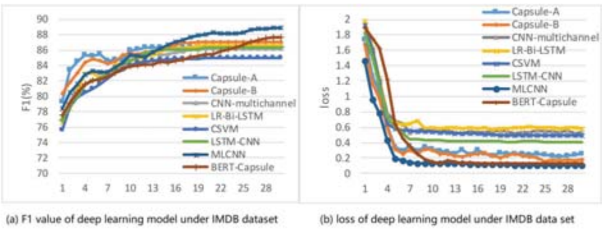

Fig. 4. Experimental results of deep learning model based on IMDB dataset.

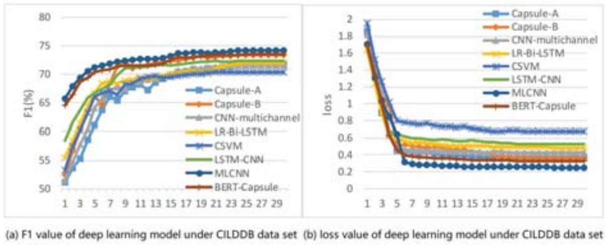

Fig. 5. Experimental results of deep learning model based on CILDDB dataset.

TABLE VI
EXPERIMENTAL RESULTS OF THE DEEP LEARNING MODEL UNDER THE CILDDB DATA SET.

model	Precision(%)	Recall(%)	F1(%)
Capsule-A [18]	71.46	71.95	71.70
Capsule-B [18]	73.18	73.66	73.42
CNN-multichannel [19]	71.63	71.49	71.56
LR-Bi-LSTM [20]	72.25	71.83	72.04
CSVM [21]	70.57	70.14	70.35
LSTM-CNN [22]	72.33	72.29	72.31
MLCNN	**74.65**	**73.76**	**74.20**
BERT-Capsule	73.25	73.85	73.55

It can be found that the convergence rate of the proposed Bert capsule model on IMDB dataset is similar.For the slow experiment, the accuracy rate, recall rate, F1 value and loss value are lower than that of MLCNN, but better than that of deep learning model except MLCNN. Compared with Capsule-A model, the best capsule model is 1.36% higher than Capsule-A model, 0.62% higher than capsule-b model, 1.42% higher than CNN multichannel model, 1.12% higher than LR Bi LSTM model, 2.71% higher than CSVM model, 1.29% higher than lstm-cnn model and 1.23% lower than MLCNN model The model has a good classification ability in the classification of positive and negative emotional semantics. At the same time, the experimental effect of the Bert capsule model is lower than that of the MLCNN model. The reason may be that MLCNN uses multiple CNN feature fusion mechanism in extracting local features, which can extract the emotional semantics of the text more effectively, and the experimental effect of MLCNN model is better than that of the Bert capsule model in this paper.

Experimental results of deep learning model under CILDDB data set. In order to further verify the effectiveness of the proposed model.Based on the five factor personality model, this paper selects the cilddb dataset for further verification. The experimental results are shown in table 6, F1 value comparison is shown in Figure 4 (a), and loss comparison is shown in Figure 4(b).

It can be found that after the fusion of local feature semantic information and global feature semantic information, the classification effect of Bert-Capsule model is better than that of other deep learning models other than MLCNN model, and better results are obtained in accuracy rate, recall rate, F1 value and loss value, which indicates that the better capsule model has better classification effect The semantic model can extract better emotion information. At the same time, the accuracy, recall rate, F1 value and loss value of MLCNN model are better than the Bert-Capsule model proposed in this paper. The main reason is that the MLCNN model can better analyze the different emotional tendencies of the same personality in the feedback information, and can better mine it The main emotional semantic information in feedback information. Therefore, the classification results of MLCNN model are better than that of Bert capsule model proposed in this paper. In addition, the effect of Capsule-B model is second only to that of Bert-Capsule model. The main reason is that although Capsule-B model can obtain emotional semantic information in text through multi-channel, it is mainly context semantic information, and it does not pay attention to the semantic information of local emotion. It is not comprehensive enough to analyze the personality characteristics of developers. Therefore, the experimental effect is slightly worse than that of the Bert capsule model.

V. CONCLUSION

To solve this problem of the lack of comprehensive personality analysis for program developers, the proposed method proposes a novel developer portrait model. The proposed model firstly collects the log data produced in the process of continuous integration development. Afterwards, the proposed method proposes the personality portrait model based on BERT-Capsule network. The experimental results show that the proposed BERT-Capsule model can effectively extract the contextual information and the local emotional information of the text. Our future work is to further improve the proposed model by designing distinguish features.

ACKNOWLEDGE

This work was supported by National Key R&D Program of China under Grant 2018YFC0407901, the Fundamental Research Funds for the Central Universities under Grant B200202177, the Natural Science Foundation of China under Grant 61702160, the Natural Science Foundation of Jiangsu Province under Grant BK20170892.

REFERENCES

[1] Xue Bai, Chuan Liao, Ming Xu, and Yaru Zheng. Correction of atmospheric model through data mining with historical data of two-line element. *IEEE Access*, 8:123272–123286, 2020.

[2] Xiaolong Xu, Xihua Liu, Zhanyang Xu, Fei Dai, Xuyun Zhang, and Lianyong Qi. Trust-oriented iot service placement for smart cities in edge computing. *IEEE Internet Things J.*, 7(5):4084–4091, 2020.

[3] Xiaolong Xu, Bowen Shen, Xiaochun Yin, Mohammad R Khosravi, Huaming Wu, Lianyong Qi, and Shaohua Wan. Edge server quantification and placement for offloading social media services in industrial cognitive iov. *IEEE Transactions on Industrial Informatics*, 2020.

[4] Lei Ren, Zihao Meng, Xiaokang Wang, Lin Zhang, and Laurence T Yang. A data-driven approach of product quality prediction for complex production systems. *IEEE Transactions on Industrial Informatics*, 2020.

[5] Xiaokang Wang, Laurence Tianruo Yang, Yihao Wang, Lei Ren, and M Jamal Deen. Adtt: A highly-efficient distributed tensor-train decomposition method for iiot big data. *IEEE Transactions on Industrial Informatics*, 2020.

[6] Michael J. Pazzani, Jack Muramatsu, and Daniel Billsus. Syskill & webert: Identifying interesting web sites. In William J. Clancey and Daniel S. Weld, editors, *Proceedings of the Thirteenth National Conference on Artificial Intelligence and Eighth Innovative Applications of Artificial Intelligence Conference, AAAI 96, IAAI 96, Portland, Oregon, USA, August 4-8, 1996, Volume 1*, pages 54–61. AAAI Press / The MIT Press, 1996.

[7] Yang Zhang, Jingjun Zhang, Yuejuan Chen, and Qiaoling Wang. Resolving synchronization and analyzing based on aspect-oriented programming. In *2008 International Symposium on Computer Science and Computational Technology, ISCSCT 2008, 20-22 December 2008, Shanghai, China, 2 Volumes*, pages 34–37. IEEE Computer Society, 2008.

[8] Xiaokang Wang, Laurence T Yang, Liwen Song, Huihui Wang, Lei Ren, and M Jamal Deen. A tensor-based multi-attributes visual feature recognition method for industrial intelligence. *IEEE Transactions on Industrial Informatics*, 2020.

[9] Cristian Torres-Valencia, Mauricio Álvarez-López, and Álvaro Orozco-Gutiérrez. Svm-based feature selection methods for emotion recognition from multimodal data. *Journal on Multimodal User Interfaces*, 11(1):9–23, 2017.

[10] Miron Kuzma and Gabriela Andrejková. Predicting user's preferences using neural networks and psychology models. *Appl. Intell.*, 44(3):526–538, 2016.

[11] Gino Slanzi, Jorge Balazs, and Juan D Velásquez. Predicting web user click intention using pupil dilation and electroencephalogram analysis. In *2016 IEEE/WIC/ACM International Conference on Web Intelligence (WI)*, pages 417–420. IEEE, 2016.

[12] Shixiang Lu, Haoran Jiang, Guoying Lin, Xiaofeng Feng, and Yang Li. Research on creating multi-attribute power consumption behavior portraits for massive users. In *2018 8th International Conference on Power and Energy Systems (ICPES)*, 2018.

[13] Ye Chen, Cheng Lv, Fangting Li, and Tiejun Li. Distinguishing the rates of gene activation from phenotypic variations. *BMC Syst. Biol.*, 9:29, 2015.

[14] Fabrício Olivetti de França, Denise Hideko Goya, and Claudio Luis de Camargo Penteado. User profiling of the twitter social network during the impeachment of brazilian president. *Soc. Netw. Anal. Min.*, 8(1):5, 2018.

[15] Gino Slanzi, Jorge A. Balazs, and Juan D. Velásquez. Predicting web user click intention using pupil dilation and electroencephalogram analysis. In *2016 IEEE/WIC/ACM International Conference on Web Intelligence, WI 2016, Omaha, NE, USA, October 13-16, 2016*, pages 417–420. IEEE Computer Society, 2016.

[16] Cristian A. Torres-Valencia, Mauricio A. Álvarez, and Alvaro Angel Orozco Gutierrez. Svm-based feature selection methods for emotion recognition from multimodal data. *J. Multimodal User Interfaces*, 11(1):9–23, 2017.

[17] Xuewei Li and Hongyun Ning. Chinese text classification based on hybrid model of CNN and LSTM. In *Proceedings of the 3rd International Conference on Data Science and Information Technology, DSIT 2020, Xiamen, China, July 24-26, 2020*, pages 129–134. ACM, 2020.

[18] Min Yang, Wei Zhao, Jianbo Ye, Zeyang Lei, Zhou Zhao, and Soufei Zhang. Investigating capsule networks with dynamic routing for text classification. In Ellen Riloff, David Chiang, Julia Hockenmaier, and Jun'ichi Tsujii, editors, *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing, Brussels, Belgium, October 31 - November 4, 2018*, pages 3110–3119. Association for Computational Linguistics, 2018.

[19] Yoon Kim. Convolutional neural networks for sentence classification. In Alessandro Moschitti, Bo Pang, and Walter Daelemans, editors, *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing, EMNLP 2014, October 25-29, 2014, Doha, Qatar, A meeting of SIGDAT, a Special Interest Group of the ACL*, pages 1746–1751. ACL, 2014.

[20] Qiao Qian, Minlie Huang, and Xiaoyan Zhu. Linguistically regularized lstms for sentiment classification. *CoRR*, abs/1611.03949, 2016.

[21] Huaiguang Wu, Daiyi Li, and Ming Cheng. Chinese text classification based on character-level CNN and SVM. *Int. J. Intell. Inf. Database Syst.*, 12(3):212–228, 2019.

[22] Yue Li, Xutao Wang, and Pengjian Xu. Chinese text classification model based on deep learning. *Future Internet*, 10(11):113, 2018.