Error Classification of Modal Verbs in English Writing Based on Meta Cognitive Strategies

Shutao ZHOU[1], Hui YU[2]

[1]*College of Humanities and Social Sciences,* Heilongjiang Bayi Agricultural University, Daqing, China
[2]*College of Science,* Heilongjiang Bayi Agricultural University, Daqing, China

***Abstract* — In order to improve the performance of model of error classification of modal verbs in writing, a type of meta-cognitive logical strategy is proposed via feature extraction based on convolutional neural network (CNN). Firstly, feature extraction of modal verb sequence based on convolutional neural network is performed against the feature expression problem of modal verbs and text feature matrix is converted into feature vector V in utilization of the trained convolutional neural network model. Then, knowledge acquisition, inference and knowledge base of modal verb error in writing are constructed in utilization of feature vector V first and then meta-cognitive logical classification strategy is established. Finally, comparison validation is made for the performances of English word segmentation test.**

Keywords - convolutional neural network; feature extraction; modal verb; error classification; meta-cognition

I. INTRODUCTION

Because there is no space between words in English, English words segmentation becomes an important step in English information processing, but errors caused by new words in world segmentation exceed 60%. In the aspect of error analysis of emotion verbs in writing, newly created words also bring a big challenge to English verb writing. In addition, the appearance of new words also greatly puzzles naming body identification, machine translation, question-answering system and other tasks. There are automatic identification methods based on rules, statistics and machine learning in the aspect of emotion verb error in writing. Word association measurement is widely used in method based on statistics. Literature [1] puts forward a kind of discovery algorithm for emotion verb error in writing based on pointwise mutual information and this method can measure the degree of interdependence between two words but fails to have an ideal achievement with regard to the correlation degree between two asymmetric co-occurrence words. Literature [2] improves PIM and obtains enhanced mutual information of emotion verb error in writing in order to measure optimized correlation of co-occurrence of two words. In addition, Literature [3] proposed a kind of distance measurement between cooperative words, namely multi-word expression distance and this kind of measurement between words does not prescribe a limit to the combination length and achieves error classification of emotion verbs in writing. Literature [4] proposed a kind of measurement method for emotion verb error in writing on word formation degree of new words, namely neologism probability, in utilization of the probability of word formation by single word and this method can remove the candidate words that occur frequently but are not new words. Literature [5] identifies new words in utilization of degree of association between words and their left and right flexibility. There are also some methods which not only utilize word co-occurrence information, but also integrate its context information into the identification method of emotion verb error in writing. Literature [6] considers the identification problem of emotion verb error in writing as a binary classification problem and uses support vector machine to identify new words in the method based on machine learning and the disadvantage of this method is that it requires a mass of manpower to select features and mark a large number of linguistic data.

In order to further improve the error classification and identification ability of modal verbs in writing, a kind of meta-cognitive logical strategy is put forward for modal verb error in writing via feature extraction based on convolutional neural network (CNN). On the basis of feature extraction for emotion sequence of emotion verbs based on convolutional neural network, error classification in writing is achieved in further utilization of meta-cognitive logical strategy.

II. MODAL VERB FEATURE EXTRACTION MODEL BASED ON CONVOLUTIONAL NEURAL NETWORK

A. *Convolutional Neural Network Algorithm*

The standard convolutional neural network (CNN) is a special case to deep neural network. The structure includes input layer, the output layer, the connection layer, and multi group pool layer and roll laminated, pool layer and convolutional layer visual for the special structure of the hidden layer. The basic operation of the convolution neural network can be defined as follows: The paper maps all words into a k –dimensional 0,1 vector space, namely $x \in R^k$, where k represents the feature number of the words, the numerical value of each dimension is represented by 0 or 1, 0 means not having such feature and 1 means having such feature. For the sentence given, it contains n words x_i with $1 \le i \le n$, which constitute a $n \times k$ feature matrix. The Thesis records the phrase segment constituted by the words in the sentence from i to j as $X_{[i,j]}$. In a similar way, sentence containing n words is recoded as $X_{[1:n]}$.

B. Neural Network Feature Extraction Based on Word Vector Convolution Neural Network Model

This is different from the word2vec and other real vector representation. In this paper, each word is mapped to a K dimensional 0,1 vector space, Network text, such as twitter, product reviews, due to the limitation of text length, content tend to be more streamlined, bias in the colloquial, emotion is single, so text related emotional expression sequences often represent the sentiment orientation of text. In recognition of the tendency of the text, if we can extract the sequence fragments related to the emotion expression, it will be helpful to accurately determine the emotional polarity of the text. Based on the observation of the rules of the emotional expression sequence of the text, this paper proposes a method of using English word resources to construct word vector.

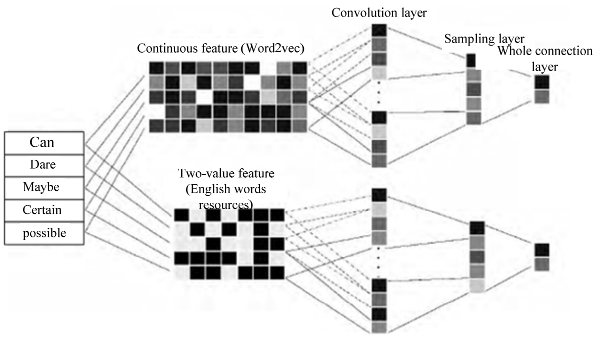

Figure 1. Convolution neural network model

The Thesis performs convolution operation for the input features in utilization of the filter with size of $h \times k$, namely:

$$c_i = f(w \cdot X_{i:i+h-1} + b) \tag{1}$$

Where c_i represents the feature value of No. i in the feature map, $f(\cdot)$ is convolution kernel function and $\omega \in R^{hk}$ is filter; h is sliding window size and b is offset value; $X_{i:i+h-1}$ represents local feature matrix formed by the rows from i to $i+h-1$. Therefore, feature map C is:

$$C = [c_l, c_2, c_3, \cdots, c_{n-h+l}] \tag{2}$$

Down-sampling layer uses max-over-time pooling method to sample features and the feature value obtained is $\hat{c}$:

$$\hat{c} = \max\{C\} \tag{3}$$

Convolution layer and down-sampling layer constitute feature extraction layer, several different types of feature extraction layers is paralleled (different values are taken for h) and there are m feature extraction layers of each type, so the feature vector V of whole connection layer is:

$$V = [\hat{c}_{1,h_1}, \cdots, \hat{c}_{m,h_1} \cdots, \hat{c}_{1,h_j}, \cdots, \hat{c}_{l,h_j}, \cdots] \tag{4}$$

Where $\hat{c}_{l,h_j}$ is the feature l produced by the filter of type j. It is expected to make it possible to further extract words and expressions sequence features related to positive and negative emotion labels via such network on the basis of the feature expression of words and expressions put forward by the Thesis for final purpose of emotion classification via such network structure.

Feature vector output by the down-sampling layer serves as the input of the whole connection layer and gradient update is made for model parameters by counter propagation algorithm in utilization of the classification result output by soft max and in accordance with actual classification labels of training data. There is:

$$P(y \mid V, W_s, b_s) = soft\max y(W_s \cdot V + b_s) \tag{5}$$

Where $y \in \{+1, -1\}$, $W_s \in R^{|V|}$ and b_s are offset items. Finally, text feature matrix of emotion verb error in writing is converted into feature vector V in utilization of the trained convolutional neural network model and meta-cognitive strategy is used to carry out emotion verb error classification.

III. META-COGNITION LOGICAL STRATEGIES

A. Theoretical Framework Description

Fig.2 shows the system structure drawing for knowledge acquisition and inference in utilization of knowledge of emotion verb error in writing put forward by the Thesis and the system is divided into three major modules which are respectively knowledge acquisition, inference and knowledge base, where arrow 1 and 4 represent that initiative logic uses a large number of knowledge to plan movement or tasks so as to lead the learning of error classifier of emotion verbs in writing for environment; arrow 3 and 2 represent that new observation is obtained via learning and perception and theorem is converted into the knowledge in the form that can be identified by the inference machine and then is put into knowledge base. Consistency of the knowledge in the knowledge base is required to be inspected.

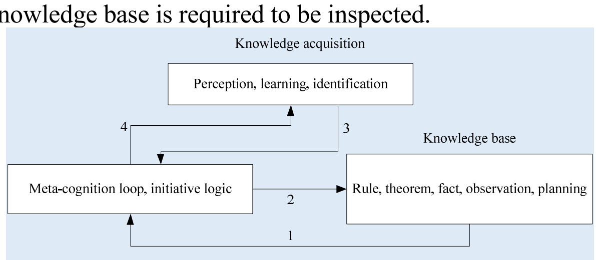

Figure 2. Knowledge inference structure of modal verb error

If the error classifier of emotion verb in writing needs to complete task in complex dynamic environment, it is not enough to only have inference and learning. Abnormity response also requires meta-knowledge, which means that the error classifier of emotion verb in writing is required to have knowledge about itself and such self- introspective cognition that can make it know about itself is called meta-cognition. If the error classifier of emotion verb in writing has meta-cognition ability, when classifier is out of order, it will detect this abnormal condition and then find out method to deal with such abnormity, such as classification calibrator. Based on it, the Thesis introduces closed loop exception handling mechanism based on meta-cognition loop as shown in Fig.3 and its theory is: meta-cognition loop firstly produces expectation and then make comparison and observation to detect whether there is abnormity and will make evaluation and response and then adjust the expectation in case of abnormity.

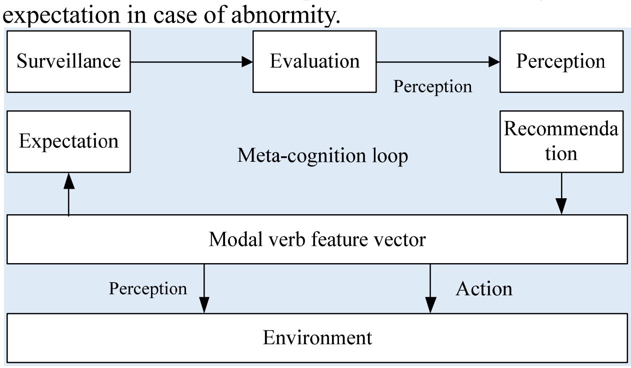

Figure 3. Meta-cognition loop for error classification of modal verbs

Initiative Logic

1) Single step execution. All rules or knowledge of imitative logic have time parameter. For example, the current step i and the next step $i+1$ are respectively expressed as:

$$\begin{cases} i : now(i) \\ i+1 : now(i+1) \end{cases} \tag{6}$$

2) Deductive inference. Step i has knowledge A and rules, B can be inferred from A and then B is added into the knowledge base at step $i+1$, and the deductive inference process is as follows:

$$\begin{cases} i : A, A \to B \\ i+1 : A, A \to B, B \end{cases} \tag{7}$$

$$\frac{i : \cdots, p_1 a, \cdots, p_n a, \forall x[(p_1 x \wedge \cdots \wedge p_n x) \to Qx]}{i+1 : \cdots, Qa} \tag{8}$$

Formula (3) is the expansion of the deductive inference of Formula (2), which means that many preconditions come to one conclusion.

3) Negation. At time step i, it is known that there is no β and $\alpha \to \beta$ is recalled at time step $i+1$, and then it can be worked out that:

$$\frac{i : \cdots, \neg\beta_1, (\alpha \to \beta)}{i+1 : \cdots, !\alpha} \tag{9}$$

4) Inheritance. Knowledge of the previous time step is inherited at the next time step.

$$\frac{i : \cdots, \alpha}{i+1 : \cdots, \alpha} \tag{10}$$

5) Inspection for contradictory knowledge. In case of conflict between the newly produced knowledge (obtained by inference or observation) and old knowledge in knowledge base, then improper knowledge shall be removed to eliminate contradiction. There are two contradictory knowledge P and $\neg P$ at step i. This contradiction is detected out and marked at step $i+1$. The contradiction can be dealt with when indirect contradiction converts into direct contradiction, which means that the contradiction is dealt with at step $i+2$.

$$\begin{cases} i : P, \neg P \\ i+1 : contra(P, \neg P) \end{cases} \tag{11}$$

IV. EXPERIMENTAL ANALYSES

A. *Experimental Settings*

The evaluation data set consists of 10000 groups of English emotion verbs, which are derived from the twitter data set. Data set of emotion verb error in writing is divided into training data set and testing data set. Training data come from the same one topic and there are totally 6000 pieces, 1003 of which have positive emotion and 3216 have negative emotion. Testing data come from three different topics such as TENSE, VOICE, FIXED COLLOCATION, and there are totally 4000 pieces, 2247 of which have positive emotion and 1753 have negative emotion.

TABLE 1. PARAMETER SETTING OF CONVOLUTIONAL NEURAL NETWORK

Adjustable parameter	Value
Convolution kernel function	Rectified linear function
Sliding window size of filter h	2,3,4,5
Number of filter m	100
Proportion of randomly updated parameters	0.5
Training iterations	50

In the aspect of data preprocessing, word segmentation tool is used to complete word segmentation and part-of-speech tagging for experimental data set of emotion verb error in writing. Word vector is produced in training by Skip-gram model of word2vec open-sourced by Google in utilization of 20 million pieces of twitter linguistic data. Dimensionality of word vector is 50-dimension, containing 330 thousand words, word coverage rate on experimental data set is 90.08 %, and for convolutional neural network model, the Thesis uniformly uses the adjustable parameter settings of convolutional neural network model as shown in Table 1.

B. Comparison of Experimental Results

The number of training data set is selected under the range of 1000~6000, the classification accuracy of the proposed algorithm, W2VCNN algorithm and NBSVM, recall rate and F-Score index are compared and verified. The experimental results are shown in Table 2. Performance of model proposed in the Thesis and other existing models is compared via experiment.

TABLE 2. EXPERIMENTAL RESULTS OF THE CONSTRUCTED DATA SET

index	Algorithm	1000	3000	6000
Classification accuracy	Algorithm in this paper	0.8158	0.8531	0.8856
	W2VCNN	0.6217	0.7023	0.7543
	NBSVM	0.7318	0.7769	0. 8126
Recall index	Algorithm in this paper	0.8194	0.8443	0.8769
	W2VCNN	0.6842	0.7108	0.7361
	NBSVM	0.7015	0.7637	0.8017
F-Score index	Algorithm in this paper	0.8235	0.8721	0.9126
	W2VCNN	0.7036	0.7532	0.7936
	NBSVM	0.7582	0.8035	0.8216

It can be learned from the data in Table 2 that, about precision index, algorithm proposed in the Thesis can maintain within 0. 8158-0.8856 and is better than the two selected comparison models: W2VCNN and NBSVM. For Recall index, algorithm proposed in the Thesis can maintain within 0.8294-0.8769 and is also better than the two selected comparison models: W2VCNN and NBSVM. For F-Score index, algorithm proposed in the Thesis can maintain within 0.8235-0.9126 and is still better than the two selected comparison models: W2VCNN and NBSVM. The above experiment results verify the superiority of the method proposed.

V. CONCLUSIONS

The Thesis proposes a kind of meta-cognitive logical strategy for modal verb error in writing via feature extraction based on convolutional neural network (CNN), converts text feature matrix into feature vector *V* in utilization of the trained convolutional neural network model, and constructs knowledge acquisition, inference and knowledge base for modal verb error in writing and establishes meta-cognitive logical strategy, and the experiment result verifies the effectiveness of the proposed method.

ACKNOWLEDGMENTS

This work is supported by Philosophy and Social Science Research Program in Heilongjiang Province (Grant No. 16TJE01).

REFERENCES

[1] Kim B, Park H, Baek Y. Not just fun, but serious strategies: Using meta-cognitive strategies in game-based learning[J]. Computers & Education, 2009, 52(4):800-810.

[2] Arafat A E R. The Effectiveness of a Training Program Based on Meta-Cognitive Strategies on Developing Vocabulary Acquisition for Secondary Stage Saudi Students[J]. Learning Languages, 2013, 2(2):919-927.

[3] Hwang G J, Chiu L Y, Chen C H. A contextual game-based learning approach to improving students' inquiry-based learning performance in social studies courses[J]. Computers & Education, 2015, 81:13-25.

[4] Jing W U. Study on Using the Meta-Cognitive Strategies to Improve the Internet-based Autonomy Learning in College English[J]. Journal of Qiannan Normal College for Nationalities, 2008.

[5] Zarrabi F. The Impact of Listening Strategy Training on the Meta-Cognitive Listening Strategies Awareness of Different Learner Types[J]. English Language Teaching, 2016, 9(5).

[6] Ma W. An Exploration of English Teaching for Mechanical Engineering from the View of Meta-cognitive Strategy Based on Multimedia Network[J]. International Journal of Emerging Technologies in Learning, 2016, 11(2).

[7] Zhu J. Research on the Method of Improving Efficiency in Adult Education Based on Metacongnitive Theory[J]. Adult Education, 2015.

[8] Rahim R A, Noor N M, Zaid N M. Meta-analysis on Element of Cognitive Conflict Strategies with a Focus on Multimedia Learning Material Development[J]. International Education Studies, 2015, 8(13).

[9] Jinyu Hu and Zhiwei Gao. Distinction immune genes of hepatitis-induced heptatocellular carcinoma[J]. Bioinformatics, 2012, 28(24): 3191-3194.

[10] Jiang, D., Ying, X., Han, Y., & Lv, Z. (2016). Collaborative multi-hop routing in cognitive wireless networks. Wireless personal communications, 86(2), 901-923.

[11] Lv, Z., Tek, A., Da Silva, F., Empereur-Mot, C., Chavent, M., & Baaden, M. (2013). Game on, science-how video game technology may help biologists tackle visualization challenges. PloS one, 8(3), e57990.

[12] Jiang, D., Xu, Z., & Lv, Z. (2015). A multicast delivery approach with minimum energy consumption for wireless multi-hop networks. Telecommunication systems, 1-12.

[13] Lv, Z., Chirivella, J., & Gagliardo, P. (2016). Bigdata Oriented Multimedia Mobile Health Applications. Journal of medical systems, 40(5), 1-10.

[14] Hassan Seifeddin A, Zakareya Ahmed S, Eman Y M E. A Program Based on English Digital Stories to Develop the Writing Performance and Reflective Thinking of Preparatory School Pupils.[J]. Online Submission, 2015, 8.

[15] Rosner A, Schuller B, Kostek B. Classification of Music Genres Based on Music Separation into Harmonic and Drum Components[J]. Archives of Acoustics, 2015, 39(4):629-638.

[16] Mohammadi S M, Zahir A T, Ghourchian N G, et al. Surveying the Effect of Emotional Intelligence and Meta-cognition on Psychological Capital of High School Teachers in Boushehr Province[J]. 2015.

[17] Semerci Ç, Batdi V. A Meta-Analysis of Constructivist Learning Approach on Learners' Academic Achievements, Retention and Attitudes.[J]. Journal of Education & Training Studies, 2015, 3(2):171-180.

.