Classification of Classroom Teachers' Speech Intention Based on Deep Learning

Xilin Zhang[1], Jiaqi Wang[2] , Zhenhong Wan[3] and Zuying Luo[4*]

[1] College of Artificial Intelligence, Beijing Normal University, Beijing, 100875, China

[2] Beijing Qingfan Technology Co., Beijing, 100193, China

[3] College of Artificial Intelligence, Beijing Normal University, Beijing, 100875, China

[4] College of Artificial Intelligence, Beijing Normal University, Beijing, 100875, China

[*]Corresponding author's e-mail: luozy@bnu.edu.cn

Abstract. Teachers use language to guide classroom teaching activities. The automatic classification of teacher speech according to intention is helpful for the quantitative analysis and evaluation of classroom teaching process. Teachers' speech in real classroom teaching of middle school Chinese and mathematics is used to construct a corpus, and deep convolutional neural network (CNN) is trained to classify teachers' speech and identify three kinds of teacher-led teaching activities, including teaching, questioning and classroom management. The experimental data show that :(1) compared with the classical shallow network classification algorithm SVM, the classification accuracy of CNN is increased by 10% to 95.5%, which can meet the requirements for accuracy of automatic analysis of classroom teaching process; (2) Classifying and statistical analysis of classroom teaching behaviors by using CNN classification algorithm can provide useful ideas for classroom analysis and research.

Keyword. teaching behaviors; text classification; teacher Discourse; CNN

I. Introduction

Classroom teaching behavior is easy to observe and is the main basis of quantitative evaluation of classroom teaching process [1]. The speech interaction between teachers and students is the main classroom teaching behavior. Teachers use speech acts to guide teaching activities. A large number of scholars use classroom language to classify classroom teaching behavior and conduct behavioral analysis and research on classroom teaching process [1-4].

As early as 1970, Flanders divided classroom language interaction behavior into three categories, namely teacher language, student language and invalid language, altogether 10 kinds of teaching behavior, and proposed the Flanders interactive analysis system (FIAS) [2] for the evaluation of classroom teacher-student interaction. Subsequent scholars add or subtract this system to meet the needs of different classroom analysis. For example, Gu Xiaoqing et al. designed the interactive analysis system (ITIAS), which meets the requirement of teaching with information technology, by increasing the types of teaching behaviors [3]. In order to reduce coding intensity, enhance objectivity and facilitate automatic coding, S-T analysis system simplifies teacher-student interaction behavior into teacher behavior T and student behavior S [4].

With the continuous development of natural language processing (NLP) technology, there have been many researches on the automatic classification of classroom teaching behavior using NLP technology in the domestic and abroad [5-6]. Compared with the study of automatic classification of classroom teaching behavior using English NLP in foreign countries [5], Chinese NLP is relatively late to carry out relevant research in China [6]. In the literature [6], teachers' speech was used for automatic classification of teaching intention, and the shallow network classification algorithm SVM (support vector machine model) was used to classify "teaching", "questioning" and "instruction". At the same time, with the rapid development of deep learning technology (DL), foreign scholars have introduced deep convolutional neural network (CNN) technology into teacher speech classification [5].

In this paper, CNN [7] is used to carry out the research on the classification algorithm of teaching intention based on teachers' speech, and a sample library of 3000 teachers' speech is constructed by collecting teachers' speech in real classroom, which is used to train and test the classification effect of CNN. Compared with the traditional SVM [6], the accuracy of CNN classifier is improved by 10%, reaching 95.5%, which is fully able to meet the classification research of teachers' speech. In the end, this paper tries to study the teaching characteristics of 16 middle school Chinese demonstration classes by using CNN classifier.

II. Methods

A. Construction of Sample Database

First, the automatic speech recognition commercial software was used to transliterate the 31 middle school Chinese/mathematics demonstration classes recommended online by Beijing Normal University, and the subtitle adding software was used to manually correct and proofread the class language to obtain the teacher's speech [1]. Then, typical samples are extracted from teachers' speech for deletion and integration. After that, a small and medium-sized sample database is constructed with 1000 language samples from each of the three teaching intentions of "teaching", "questioning" and "instruction".

B. CNN Network Training for the Three Categories of Teaching Intention

Deep Convolutional Neural Network (CNN) is a typical deep learning model [7]. With its good performance, it can effectively solve the problem of text classification [8]. The

CNN model used in this paper takes short passages of different lengths as matrix inputs and extracts key information in sentences with multiple filters of different sizes. The horizontal and vertical axes of sentence input are word vector dimensions and text length respectively. Different features of texts are extracted with convolution kernels of different sizes.

Compared with Windows system, Linux system is more conducive to the use of CNN model, so Linux system is selected to build the environment. From the three types of corpus samples, 800 samples are selected as the training set, 100 samples as the test set and 100 samples as the verification set. The CNN model consists of an input layer, a convolution layer, an activation function, a pooling layer and a full connection layer. It completes the tasks of corpus input, feature extraction, non-linear element addition, compression and extraction of main features, and connection of all features successively, and finally sends the output value to the classifier. We set the dimension of word vector as 64, sequence length (i.e. text length) as 600, the number and size of convolution kernel as 256 and 5 respectively, full-connection layer neurons as 128. This setting is convenient for model creation and debugging. The dropout ratio is 0.5 to avoid over fitting.

III. Experimental Results

A. Verification of Classification Accuracy of CNN Instructional Intent Classification Algorithm

In order to verify the accuracy of CNN's instructional intent classification algorithm, we use the classic shallow network classification algorithm SVM as a comparison algorithm to carry out a comparison experiment.

As a classic shallow network classification algorithm for machine learning, SVM has the advantages of global optimization and simple structure. It has been widely studied and applied to text classification [9]. And it has a unique advantage in dealing with small-scale texts. The team of Central China Normal University used SVM to automatically classify the teaching intention of "teaching", " questioning " and " instruction ". By combining TF-IDF feature based on chi-square test with Word2vec word vector feature, a classifier of instructional intention based on SVM is constructed, and the accuracy rate was 86% [6]. Similarly, this paper builds its own SVM teaching intention classifier based on SVM.

Based on our self-built teaching behavior sample database, we first used CNN classification algorithm to conduct classification tests on 200 samples of each of the three teaching intentions. As shown in Table 1, the results of CNN classification were obtained. The classification accuracy of all CNN classification algorithms is more than 90%, with an average of 95.5%, indicating that the CNN teaching intention classification algorithm has a high classification accuracy and can meet the accuracy requirements of classroom teaching

behavior process analysis.

Table 1. The Results of CNN Classification

Teaching Intention	Teaching	Questioning	Instruction	Accuracy
Teaching	181	1	18	90.5%
Questioning	0	200	0	100.0%
Instruction	11	1	188	94.0%

Then we further compared the classification accuracy of SVM and CNN for different teaching intentions. As shown in Table 2, it can be found that CNN completely surpasses SVM in classification accuracy, reflecting the performance advantage of deep neural network over shallow network. In terms of average classification accuracy, CNN's classification accuracy is 95.5%, which is not only 20 percentage points higher than the 75.3% accuracy of SVM trained by ourselves, but also nearly 10 percentage points higher than the 86% accuracy of SVM classifier in literature [6]. CNN classification algorithm based on deep neural network can well meet the accuracy requirements of classroom teaching behavior analysis.

Table 2. Comparison between SVM Model and CNN Model

Teaching Intention	SVM	CNN
Teaching	75.5%	90.5%
Questioning	68.0%	100.0%
Instruction	82.5%	94.0%

B. The Application of CNN Classifier in the Evaluation of Chinese Classroom Teaching in Middle School

We use the CNN instructional intention classifier to conduct statistical analysis of teacher speech acts in the classroom teaching process. As shown in Figure 1, the pie chart of teachers' speech acts in a Chinese course of a middle school is obtained, in which teaching acts account for 41.59%, directive acts account for 31.22%, and questioning acts account for 27.20%. It shows the teaching characteristics of this teacher, who mainly teaches, supplemented by instructions and questions, and balanced teaching.

By using the distribution map of teacher speech acts, we can not only compare the teaching characteristics among different teachers, but also carry out other teaching characteristics research. As shown in Figure 2, we compared the proportion of teaching behaviors in 16 Chinese lessons in middle school. We found that compared with the other 13 senior high school Chinese classes, the teaching behaviors of teachers in the 3 middle school Chinese classes "Case 9", "Case 13" and "Case 16" were significantly less, which indicated that the Chinese teaching in junior high school paid less attention to the teaching efficiency than the teaching in senior high school, and paid more attention to the teaching interaction such as instruction and question-and-answer.
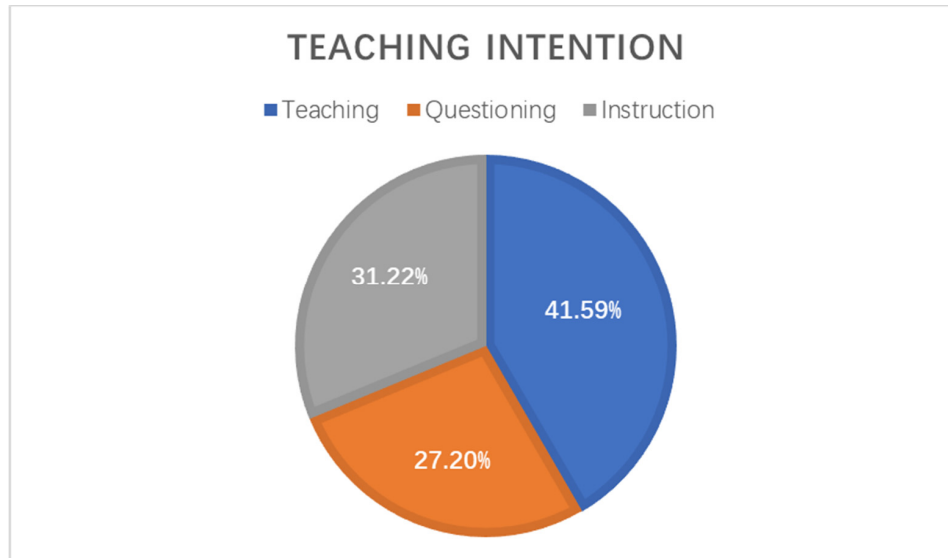
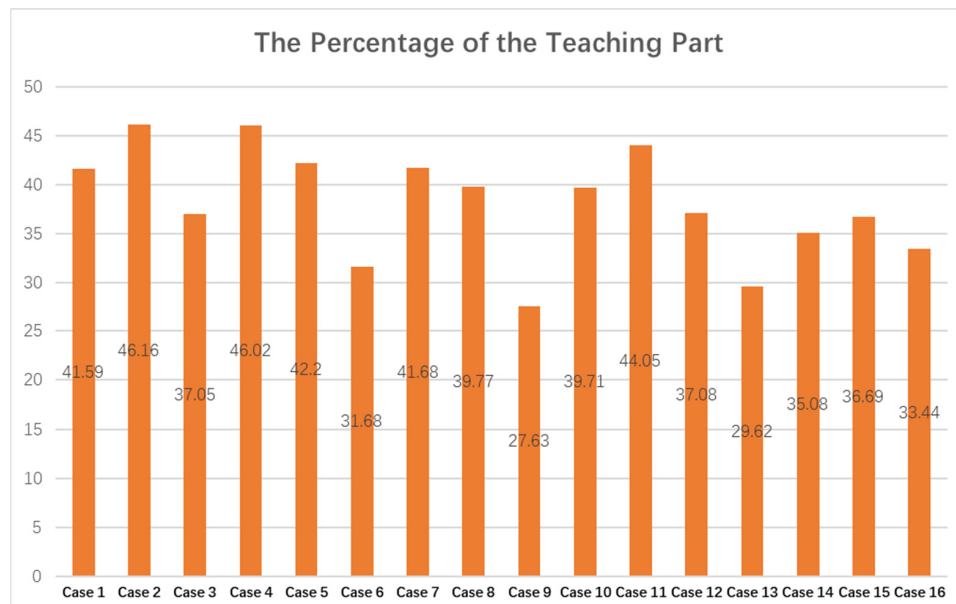Figure 1. The Classification Results of Teachers' Speech Intent in the Classroom Teaching

Figure 2. The Comparison Chart of the Percentage of the Teaching Part of the Chinese Demonstration Class

IV. Conclusions

With the rapid development of artificial intelligence technology, commercial software can quickly transcribe high precision teacher speech text from teacher near-field speech. In this paper, CNN text classification algorithm is used to achieve the classification of teaching intent based on teachers' speech texts, which has a high classification accuracy and can carry out teaching research such as the analysis of teaching characteristics of teachers, providing a meaningful research perspective for the analysis of fully automatic classroom teaching behavior process.

References

[1] Luo Z. (2018) Automatic Evaluation of Classroom Teaching and Learning: from theory to applications. Beijing Normal University Publishing Group, Beijing.

[2] Ned. F. (1970) Analyzing Teaching Behavior. ＭＡ. Addision—— Wesley Publishing 122. Company, New York.

[3] Gu X, Wang W. (2004) New Exploration of Classroom Analytical Techniques for Supporting Teacher Professional Development. China Educational Technology, 7:18-21.

[4] Cheng Y, et al. (2016) Application Research of Improved S-T Analysis Method based on Video. E-education Research, 37(06): 90-96.

[5] Donnelly, P. J., et al. (2016) Automatic Teacher Modeling from Live Classroom Audio.In: 24th ACM Conference on User Modeling, Adaptation and Personalization.45-53.

[6] Xu L, et al. (2019) Automatic Classification of Discourse in Chinese Classroom Based on Multi-feature Fusion. In: 2019 International Conference on Computer, Information and Telecommunication Systems (CITS). Beijing. 1-5.

[7] Kim Y. (2014) Convolutional Neural Networks for Sentence Classification. Eprint Arxiv.

[8] Suresh A, et al. (2018) Using deep learning to automatically detect talk moves in teachers' mathematics lessons. In: IEEE International Conference on Big Data. Seattle. 5445-5447.

[9] Liu Z. (2016) An improved hybrid kernel support vector machine text classification method. Industrial Control Computer, 29(06): 113-114+117.