Research and Application of Chinese Composition Online Intelligent Evaluation System

Jian-She ZHOU¹,²,a and Tie-Jun SHAO¹,²,³,*

¹School of Literature, Capital Normal University, Beijing, China
²Research Center for Language Intelligence of China, Beijing, China
³North China Institute of Science and Technology, Yajiao, Hebei, China

a zhoujs@cnu.edu.cn,* shaotj825@ncist.edu.cn

Keywords: Natural language processing; Language intelligence; Intelligence evaluation system; Chinese essay evaluation.

Abstract. Based on the methods of natural language processing, language intelligence, network technology to construct an online Chinese essay intelligent evaluation system. This intelligent essay evaluation system composed of essay word segmentation, keyword extraction, the polymerization degree of the theme and etc. to evaluate Chinese essay automatically and intelligent.

Introduction

With the great development and wide range of applications of natural language processing and language intelligence, the corresponding application range is also expanding. Education is one of the many applications of these theories and technologies. The main application forms in intelligent education are intelligent education software, online intelligent education system, and intelligent evaluation system. At present, the research and application of the student's essay intelligent evaluation system is mainly for English writing, such as the use of PiGai.net; meanwhile the research and application of the Chinese essay intelligent evaluation system is very few, the most famous Chinese essay evaluation system is inKeTang. There are many reasons for this, including the difficulty of Chinese natural language processing which is more difficult than English, such as Chinese word segmentation problem, ambiguity resolution problem, semantic recognition problem, etc. All of these problems have brought certain obstacles to the research and application of Chinese essay intelligent evaluation. This paper will take inKeTang as an example to discusses the research and application of a Chinese essay intelligent evaluation system based on natural language processing and language intelligence.

Related Theories

Natural Language Processing

Natural language processing is a science that integrates linguistics, computer science, and mathematics. Natural language processing is an important direction in the field of computer science and artificial intelligence. NLP studies various theories and methods that can realize effective communication between human and computer in natural language. Natural language processing contains speech recognition, natural language understanding, and natural language generation.

Chinese Word Segmentation

Words are the smallest language units that can be used independently. Chinese word segmentation refers to dividing a sequence of Chinese characters into individual words [1]. The word segmentation is the basic input of the upper natural language processing task, and the quality of the word segmentation will directly affect the effect of the upper NLP layer application’s effects. The existing word segmentation algorithms are mainly divided into three categories: word segmentation methods based on string matching; word segmentation methods based on
understanding and word segmentation methods based on statistics. In the intelligent Chinese essay evaluation system mentioned in this paper, the word segmentation is the first operation to be completed in the evaluation process, and accurate word segmentation is the basic preparation for the subsequent evaluation in the system.

**Keyword Extraction**

In natural language processing, keyword extraction is an important content which task is to automatically identify the term that best describes the subject of the document. Key phrases, key terms, key word segments, or keywords are terms that represent the most relevant information to the subject of the document. Although these terms mentioned above are different but their functions are the same that is to describe the subject of the document. Keyword extraction is an important research issue in text mining, information retrieval and natural language processing. The extraction of keywords in the intelligent Chinese essay evaluation is used to do preparation for the subsequent topic aggregation calculation.

**Language Intelligence Evaluation**

Language intelligence evaluation refers to use computer for evaluating essay. The main idea of language intelligence evaluation is based on a large number of essay examples and teachers’ corresponding comments and scores to do deep learning, that is using the corresponding method of natural language processing to extract the text features under essays, and then input these text features into the neural network to learn the relationship between features and comments as well as scores.

**Theme Aggregation Calculation**

The theme is the object referred to by the text that is the natural component of the content of the text. The theme is the extension or domain of the text what the text wants to reflect the things, facts, events, nature, and relationships in the real world. The degree of aggregation of chapter theme is an important indicator to measure the pros and cons of an article. The theme aggregation degree refers to the closeness between the extension of the theme displayed by the language form in the text. The theme aggregation degree of the chapter is completed by calculating the expressive power of various themes.

The dimensions of the theme aggregation degree evaluation are divided into different levels such as chapter level, paragraph level, sentence group level, complex sentence level and word level. The design goal of theme aggregation degree calculation is to realize the automation of text comprehension and evaluation by means of linguistic intelligence; the value of theme aggregation degree calculation lies in providing scientific, accurate, fast and standardized evaluation core indicators for essay; the formal indicators of theme aggregation degree calculation includes the expressiveness of the theme from the syntactic subject [3], expressiveness of the theme from the syntactic object [4] and expressiveness of the theme from the syntactic predicate[5].

**Automated Essay Scoring**

The automated essay scoring (AES) [2] system uses a specific computer program to score students’ essay. It is a teaching evaluation system in which natural language processing is applied. The automated essay scoring system is a concrete application of the automation and intelligence of the intelligent teaching system, and it is also an application of artificial intelligence in the field of education.

At present, the famous international automatic essay evaluation systems include IEA (INTELLIGENT ESSAY ASSESSOR), E-RATER (ELECTRONIC ESSAY RATER) and PEG (PROJECT ESSAY GRADER). These automatic essay evaluation systems have rather amount of advantages, they are designed mainly for people who speak English as the first language rather than for those who learn English as a second language.

The well-known civil English automatic essay evaluation system includes the PiGai network and REALSKILL. In them, the PiGai network has revised more than 300 million English essays all
around the country, which has eased the burden of English teachers. Students upload essays in real time and perform automatic evaluate, the PiGai network has achieved significant results in practical applications. These automatic essay evaluation systems are all automatic evaluation systems for English essay. At present, the online products for Chinese automatic essay evaluation system is very few.

Chinese Essay Intelligent Evaluation System

Related Work

It is an important teaching goal to cultivate students' writing ability in Chinese teaching. The evaluation and correction of essay is an indispensable important relief in essay teaching that is an important means to improve students' writing level. The improvement of the writing level is of great help to improve the language ability of students in all aspects, such as language expression ability and communication ability. Therefore, in the essay teaching, students are encouraged to write more and practice more to improve the writing ability. However, in the actual teaching, teachers will spend a lot of energy and time on in time and meticulous revision of a large number of essays. At the same time, teachers should also take into account other teaching and scientific research tasks, so that correcting essays, especially correcting essays in time, becomes a heavy burden for teachers. Students write more and practice more and teachers overburden to form sharp contradictions.

Advantages of the Essay Intelligent Evaluation System

The online essay intelligent evaluation system is not only a supplement to the traditional teaching essay mode, but can be regarded as an extension of the traditional mode. The evaluation of the essay extends from the school to the outside of the school, and extends from the offline classroom to the online network. Using network technology, language intelligence, natural language processing and other methods to construct online Chinese essay intelligent evaluation system, this system provides students with real-time online essay evaluation by breaking boundary such as the time, the space and other constraints.

Convenience. The essay intelligent evaluation system takes the network as platform for essay evaluation. Students can do essay writing level evaluations anytime and anywhere through the Internet and mobile phones, computers and other terminals without the limitation of time and space.

Intelligence. In the traditional essay evaluation, due to the limited energy and time of teachers, it is impossible to evaluate and feedback each student's essay work in time. The online essay intelligent evaluation system incorporates advanced natural language processing technology, through essay segmentation, keyword extraction, and difficulty analysis of chapter and other aspects of the essay from the multiple levels of intelligent evaluation. And the online essay intelligent evaluation system would highlight the wonderful place of essays, give comments in paragraphs, combined with the evaluation points to give the evaluation scores of the composition, to achieve the evaluation of intelligence.

Real-time. The online essay intelligent evaluation system is based on natural language processing technology running on outstanding computer that real-time feedback of the evaluation results can be realized, so that students can see the evaluation results in time.

Individuation. The online essay intelligent evaluation system is an intelligent evaluation system that performs separate evaluations based on the essays uploaded by students. The evaluation system would give individual evaluation scores and reviews to each essay, thereby achieving one-to-one evaluation that leads to individual evaluation.

Evaluation Process

The online essay intelligent evaluation system inKeTang consists of essay segmentation, essay keyword extraction, Chinese character complexity evaluation, word complexity evaluation, text difficulty assessment, chapter structure fluency assessment, chapter subject conformity assessment, etc. Each step involves methods such as word segmentation technology, keyword extraction
technology, text complexity calculation, topic aggregation degree calculation, and text structure analysis. The evaluation process is as shown in the Fig. 1:

![Evaluation Steps Diagram]

**Aspects of Evaluation**

The online essay intelligent evaluation system evaluates the essay based on the essay title, the essay topic and the student grade from two aspects: content and expression. The content evaluation includes whether the essay conforms to the given meaning, whether the essay is prominent, whether the composition is substantial, and whether the essay is ideological health and essays are true feelings; expression evaluation includes whether the essay is standardized, whether the essay meets the requirements, whether the essay structure is rigorous, and whether the essay is fluent in language.

**Evaluation Result Measure**

There are two methods for measuring changes in learning performance. One is the longitudinal variation of the same sample, such as the change of two tests with the same content (ie, pretest and posttest); the other is the horizontal comparison of different samples, such as differences in grades in the synonym test in the experimental class and the reference class. The in class's essay evaluation system uses a vertical calculation method for student performance changes, that is, comparing the scores of two or more evaluations of the same essay topic to show the changes in the student's essay writing level.

**Conclusion**

In this paper we discuss an intelligent essay evaluation system for Chinese essay based on natural language processing, language intelligence in order to provide a real time evaluation system to evaluate essay intelligent, effective. We take inKeTang as an example that has been applied widely and get good effects. The further research will be concentrated on semantic analysis of essay to evaluate essay more intelligent.

**References**


