Design of English Translation System Based on Deep Learning

Yang Ting*

School of Foreign Language Studies, Shaanxi University of Chinese Medicine, Xianyang, China

*Corresponding author: yang_ting@sntcm.edu.cn

Abstract. Intelligent translation software is one of the important tools for people to learn, but the existing intelligent translation system usually needs to be further improved in terms of the flexible application of phrases and the accurate expression of professional vocabulary. The translation system based on deep learning designed in this paper mainly includes professional vocabulary database module, memory analysis module, English translation module and correction module. The professional vocabulary module accurately classifies words related to various disciplines. The memory analysis module compares the professional vocabulary usage in published documents on the Internet to analyze and memorize, so as to recommend relatively accurate professional vocabulary to enter the translation module. Need to enter the correction module to complete the correction of grammar and phrases, and finally complete the translation task with higher accuracy.

1. Introduction

Machine translation (MT) is a comprehensive discipline built on a multi-disciplinary basis. The development of modern theoretical linguistics, the progress of computer science, the application of informatics and probability statistics have had an important impact on the development and evolution of machine translation. The basic idea of machine translation is to use computers to translate natural language [1]. However, various machine translation systems adopt different technologies and concepts, and their translation effects are also quite different.

Translation software usually provides English translation results by searching semantic vocabulary of the entire web based on its own algorithm settings. However, the reliability of these English translation results is poor and cannot be used directly, and a lot of manual proofreading is required later. Deep learning is to learn the inherent laws and representation levels of sample data [2]. The information obtained in the learning process is of great help to the interpretation of data such as text, images and sounds. Its ultimate goal is to allow machines to have the ability to analyze and learn like humans, and to recognize data such as text, images, and sounds [3]. Deep learning is a complex machine learning algorithm, and the results achieved in machine translation far exceed previous related technologies [4]. The intelligent translation system proposed in this article is based on deep learning theory to achieve accurate expression of English translation.
2. Principle

The translation system based on deep learning proposed in this paper is mainly composed of four modules, namely database module, memory analysis module, correction module and translation module. The working process of the system is actually an English translation process. Through the translation of English sentences, the incorrect parts of the original translation results are replaced to realize intelligent proofreading and obtain the correct English translation results as possible. The system searches the internet for relevant translation information according to the characteristics of the sentence to be proofread, and saves it in the memory analysis module.

![Figure 1. The overall architecture of the translation system based on deep learning.](image)

The memory analysis module is the basis for intelligent proofreading of English translation. When the proofreading command is issued, the memory analysis module will receive the search link from the database module. The English translation module analyzes the vocabulary features of the sentences to be proofread and compares them on the network. Degree ranks the translation results, and finally selects the most realistic translation results.

In the process of English translation, how to obtain objective data and use these data for correct analysis is very important. In the data extraction stage, the original data needs to be preprocessed to eliminate the interference of useless data to the entire analysis process. Due to the partial lack and omission of the original data, it is incomplete, it is necessary to fill in the filling process according to certain standards, and then input the processed data into the neural network for analysis. Furthermore, it is necessary to determine the topology of the neural network model, the number of nodes in the input layer, the number of nodes in the output layer, and the number of nodes in the hidden layer as Figure 2 shows.

![Figure 2. Neural network model.](image)
The model used in deep learning is a deep neural network (DNN) model, which contains multiple hidden layers (Hidden Layer, also known as hidden layers). Deep learning uses the hidden layers in the model to transform the original input into shallow features, middle features, and high-level features layer by layer through feature combination to the final task goal. Deep learning is a series of algorithms born by the human brain to process data and create patterns for making decisions, and expand and improve a single model structure called artificial neural network. Like the human brain, neural networks also include many neurons. Each neuron receives the input signal, then multiplies it by the corresponding weight, sums it and inputs it into a non-linear function. These neurons are stacked on top of each other and organized in layers. Assuming that every behavior and every system can finally be represented by a mathematical function (some may be very complex functions), if we can find such a function, we can understand everything about the system. In this system, various vocabulary and phrases are used as neurons as a single layer, and different types of translation criteria and skills in different fields are used as multiple hidden layers to train the neural network model, and finally output an ideal translation result.

The pseudo code corresponding to the computer intelligent proofreading method based on the improved phrase translation model designed by the system is as follows:

Start
Terms Form = NEW semantics blurry;
// (Restore the English to be proofread)
ashy relevance = NEW semantics theme
// (To restore the proofreading vocabulary weight set)
Vocabulary = English translation (Vocabulary Form)
// (Extract the basic meaning of vocabulary and subject information)
Simultaneously (phrasal language environment Is Not Empty. ){
IF (Vocabulary semantics blurry best)
Vocabulary = folds cooperate (phrasal message);
// (Cannot be obtained from the vocabulary weight set, change the weight value to query in all directions)
Otherwise
Crunode = phrasal message (Scientific);
// (Correctly analyze translation results)
Analog = assort compute (Vocabulary, Crunode);
// (Measure the semantic similarity of words)
Analog _ Form. Place (lauygdbg of (Sfrg));
// (Find the most reasonable vocabulary corresponding to the vocabulary to be proofread from the sorted results)
Vocabulary. free surface (If(simregrfist). Nbtjuke);
// (Induction for word segmentation)
END}

3. Experiment and result
In order to verify the effectiveness of the system in the English translation computer intelligent proofreading, the English translation proofreading test was carried out on the system, and the relevant data during the experiment was recorded to analyze the system performance. In the experiment, the number of word proofreading is 400 characters, the number of short text proofing is 500, the proofreading speed is 15 Kb/s, and the word recognition speed is 25 Kb/s. The accuracy of the English translation results after the proofreading by the system of this paper is compared with the accuracy before the proofreading is shown in Table.1.
Table 1. Accuracy of English translation results before and after using the proposed system

| experiment number | accuracy of English translation results before (percentage) | accuracy of English translation results after (percentage) |
|-------------------|----------------------------------------------------------|-----------------------------------------------------------|
| 1                 | 58.6                                                     | 98.7                                                      |
| 2                 | 64.2                                                     | 99.1                                                      |
| 3                 | 61.5                                                     | 99.5                                                      |
| 4                 | 68.7                                                     | 98.8                                                      |
| 5                 | 56.8                                                     | 99.2                                                      |

It can be seen from Table 1 that the translation accuracy rate before using this system does not exceed 70%, and the translation accuracy rate after using this system can almost reach 99%, the accuracy rate has increased by almost 29%.

All manuscripts must be in English, also the table and figure texts, otherwise we cannot publish your paper. Please keep a second copy of your manuscript in your office. Please keep a second copy of your manuscript in your office. When receiving the paper, we assume that the corresponding authors grant us the copyright to use the paper for the book or journal in question. Should authors use tables or figures from other Publications, they must ask the corresponding publishers to grant them the right to publish this material in their paper.

Please keep a second copy of your manuscript in your office. When receiving the paper, we assume that the corresponding authors grant us the copyright to use the paper for the book or journal in question. Should authors use tables or figures from other Publications, they must ask the corresponding publishers to grant them the right to publish this material in their paper.

4. Conclusions
The intelligent English translation system designed in this paper uses the database module to complete the precise selection of vocabulary in the target to be translated; the software part uses the computer intelligent proofreading method based on the improved phrase translation model to realize the intelligent proofreading of English translation. Experiments show that the system in this paper improves the accuracy of English translation results by 29%, and can effectively proofread the incoherent context in English translation results. Compared with similar systems, the designed system has high accuracy and contextual coherence. Performance advantage. The design of the system provides a new means for intelligent proofreading of English translation, greatly reducing the cost of manual proofreading, and improving the efficiency of English translation proofreading.

Acknowledgments

References
[1] Feng Zhiwei. Phrase-based and syntax-based statistical machine translation [J]. Journal of Yanshan University. 2015. 39(16): 546-554.
[2] Chen Hui. A study of reformation of formative evaluation system for translation courses of English major [J]. Journal of Sichuan International Studies University. 2017. 33(4):133-136.
[3] Wang Wenhu, WU Minhua, LUO Liming, et al. Design and implementation of English intelligent question-answering system based on similarity algorithm [J]. Computer applications and software. 2017. 34 (6): 62-68.
[4] Li Yegang, LIANG Lijun, SUN Fuzhen, et al. Machine translation model integrated with bilingual maximal-length noun phrase [J]. Application research of computers. 2017.34 (6): 1316-1320.