Investigation on the Application of Artificial Intelligence Large Language Model in Translation Tasks

Chunlan Jiang

Senior Translation School, Xi'an Fanyi University, Shaanxi Xi'an, China 710105

jiangchunlanzhengf@126.com

Abstract. As an emerging language technology, the application of Artificial Intelligence (AI) Large Language Model (LLM) in translation tasks has an important background. Based on a large number of experimental data, this paper compared traditional machine translation and Google translate, and evaluated the performance of AI LLM in multilingual translation tasks. The experimental results showed that compared with traditional machine translation and Google translate, the AI LLM performed better in terms of translation quality and speed. Specifically, the Bilingual Evaluation Understudy (BLEU) score of the LLM was about 5 percentage points higher than that of traditional machine translation and Google translate, and the BLEU value per second was about 5 percentage points higher than that of traditional machine translation; in terms of speed, the LLM was 13.53 seconds faster than traditional machine translation on average. These findings indicated that the AI LLM had broad application prospects and important application value in practical applications, and could provide better technical support for achieving language translation.

Keywords: Artificial Intelligence; Natural Language Processing; Recurrent Neural Network; Large Language Model

1 Introduction

With the acceleration of globalization and the continuous deepening of cross-border exchanges, the translation industry has achieved unprecedented development. With the rapid development of AI technology, the AI LLM has become one of the most concerned technologies in the field of translation [1-2]. Especially in translation tasks, AI LLM can play a huge role. translation tasks often involve multiple languages and fields, requiring the support of a large number of professional translators and various translation tools. However, the birth of AI LLM has provided more convenient and efficient translation methods for translation tasks [3-4]. By pre training a large-scale corpus and utilizing deep learning algorithms for self-learning, the LLM can quickly translate a large amount of text, thus saving translation manpower and time costs and improving translation quality and efficiency. Different from traditional machine translation, AI LLM can understand context and context, and has higher language expression ability and accuracy. For example, in some political, diplomatic, and business negotiations,

S. Yacob et al. (eds.), *Proceedings of the 2023 7th International Seminar on Education, Management and Social Sciences (ISEMSS 2023)*, Advances in Social Science, Education and Humanities Research 779,
https://doi.org/10.2991/978-2-38476-126-5_147

the choice and expression of words are crucial for conveying information and negotiating results. A LLM with sufficient language proficiency can better handle such complex scenarios and produce more accurate translation results [5].

In recent years, many scholars and experts have conducted research on the application of AI LLM in translation tasks. Among them, Xiong L made significant progress using NON-AUTOREGRESSIVE (NAR) Neural Machine Translation (NMT). The main advantage of NAR-NMT was that it could simultaneously process all words in the input sentence, thus allowing for parallelization and improving translation efficiency. There were currently two main types of NAR-NMT methods: template based and generative based. The template based method used previously translated word sequences as templates. Based on this, adjustments and modifications were made to generate the final translation. The generative method used the bridge of posterior distribution and word embedding vector to predict translation results. He stated that the latest research currently showed that combining templates and generative methods could achieve better results. In addition, there were other techniques and structural optimizations that could be used to further improve the performance of NAR-NMT [6]. Liu Y proposed a neural machine translation model for Chinese English translation using massively parallel corpora. This model adopted a deep bidirectional long short term memory network, which could better learn and model sentences. At the same time, it introduced a adaptive learning rate algorithm (AdaDelta) to optimize the training process [7]. Kiros R proposed that the multimodal neural language model was a language model that could simultaneously process multiple language modalities such as text, images, videos, sound, etc. This model utilized deep learning technology to fuse different types of media information to establish a richer and more accurate language model. Compared with the traditional single mode language model, the multimodal neural language model could better use multiple information for semantic understanding and generation, thus improving the effect of natural language processing. At the same time, the multimodal neural language model was also an important means to achieve multimedia intelligence, cross media information processing, modal fusion, and other aspects [8].

The research on the application of AI LLM in translation tasks shows that compared with traditional machine translation methods, the effect of using LLM for translation is better and faster. LLM can extract more abundant semantic information and have better natural language processing ability by learning a large number of corpora. In translation tasks, LLM can more accurately understand the meaning of sentences, identify more language contexts and common sense, thereby improving translation quality and reliability.

2 Relevant Methods for the Application of AI Large Language Model in translation tasks

2.1 Introduction to Recurrent Neural Network under AI

The strength of Recurrent Neural Network (RNN) lies in its ability to process sequence data. It is widely used in machine translation, speech recognition, Natural language processing and other fields. In AI, RNN can be used to process natural language texts, audio signals, video sequences, etc. In language models, RNN can be seen as a model where the input sequence is the same as the output sequence. Among them, each input is related to the output of the previous moment, which enables it to encode contextual information into the model. In image description, RNN can match different parts of the image with textual descriptions. There are many variants of RNN, such as Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU), which can better deal with long-term dependence.

In RNN, the implicit state h (Hidden State) is introduced, which extracts features from the forward data sequence of serialized data based on h, saves the obtained state, and finally converts it into output, thus solving the problem of modeling sequence.

$$h_e = f(Pn_e + Qh_{e-1} + a) \tag{1}$$

The calculation of hidden state h_e is obtained by simultaneously inputting the current time e of the sequence and the previous hidden state. Among them, P, Q, and a are all parameters to be learned. When passed to the later moments, the calculation formula for the hidden state is the same, and the parameters P, Q, and a used in each step are the same, which means that the parameters of each step are shared. After obtaining the hidden state of each step, the output of each step can be obtained through the hidden state and input, as shown in Formula (2). Among them, k is the activation function, and S and t are parameters to be learned.

$$z_e = k(Sh_e + t) \tag{2}$$

Activation function is an indispensable module in deep neural network. If the Activation function is not used in the neural network model, each neuron can only use linear transformation to optimize the weight. This seemingly simple and fast way directly leads to the neural network can not learn the complex mapping expression in the training process, which makes the performance of the neural network model worse. Therefore, this paper selects the Rectified Linear Unit (ReLU) function as the activation function of the standard algorithm RNN. The ReLU function formula is as follows:

$$ReLU(k) = max(k, 0) \tag{3}$$

2.2 Large Language Model

LLM is a Natural language processing model based on machine learning technology. It can encode and understand natural language and generate text that conforms to syntax and semantic logic. The LLM is a kind of supervised learning model, which needs a lot of training data to learn the rules and patterns of language, so as to build a model that can better understand and generate language [9-10]. The basic idea of the LLM is to establish a Statistical model for language modeling, and use probability distribution to describe the relationship between the possibility of language and various language phenomena. In the framework of this model, the probability of any given paragraph of text appearing in the model can be calculated to evaluate and compare the language expression of the text [11-12].

LLM are usually built based on neural networks, and the most common architecture is RNN, which is used to process sequence data. This model can output a probability distribution at each moment, that is, each word, to represent the next possible word [13-14]. The training data of LLM usually uses Corpus, which provides a comprehensive and rich description of language expression, such as Wikipedia, Gutenberg Project, news media, movie subtitles, etc. [15-16].

LLM are widely used in text generation, automatic summarization, machine translation, speech recognition, speech synthesis and other fields. In terms of text generation, LLM can generate articles that conform to grammar and semantic logic based on a given theme or style, such as news reports, social media Weibo, novels, etc. In terms of automatic summarization, LLM can analyze and summarize an article to generate a concise and general summary. In machine translation, LLM can help further optimize the translation model and improve the accuracy and fluency of translation. In terms of speech recognition and synthesis, LLM can also model and optimize the audio expression of text, as shown in Figure 1:

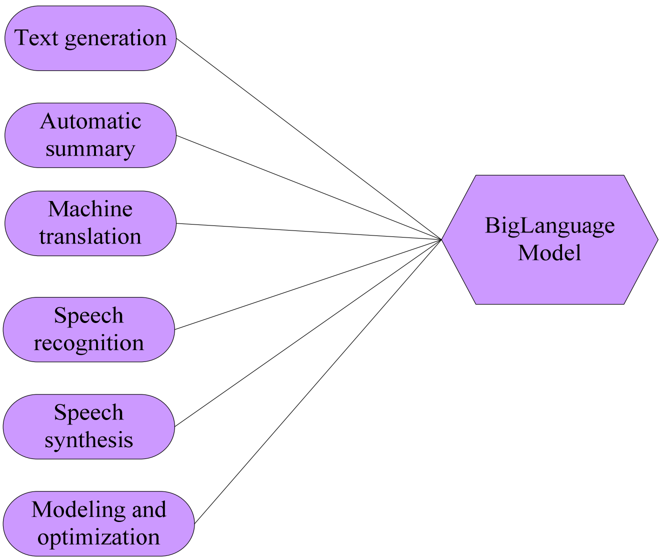

Fig. 1. Application scenario flowchart of the LLM

2.3 Integration of AI Large Language Model in translation tasks

The AI LLM is a technology used to generate human language, which has received widespread attention and application in recent years. In translation tasks, AI LLM can play an important role, as shown in Figure 2:

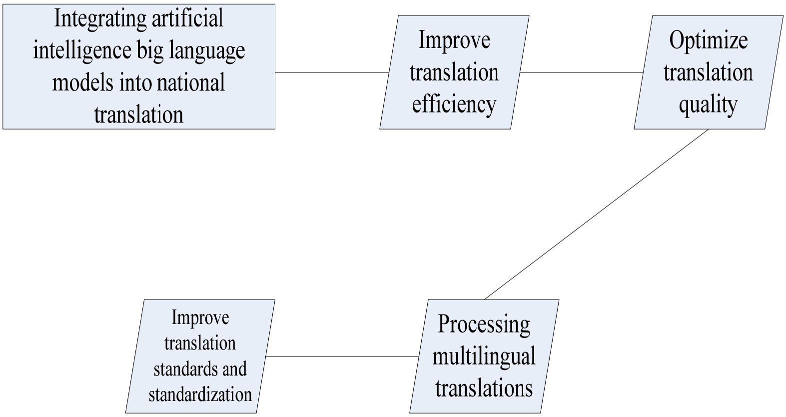

Fig. 2. The fusion of AI LLM in translation tasks

As shown in Figure 2, the fusion of AI LLM in translation tasks can be divided into the following four points:

1) Improving translation efficiency: The AI LLM can quickly analyze a large amount of language data, identify patterns and correlations within it, and automatically translate, thus greatly improving translation efficiency. In tasks such as rapid response to international incident and publication of international papers, AI LLM can solve translation problems and quickly complete translation tasks [17-18].

2) Optimizing translation quality: The AI LLM can generate high-quality human language and produce more accurate and fluent language during the translation process. Especially in fields that involve professional terminology or have high requirements for sentence structure and grammar, AI LLM can play a good optimization role.

3) Processing multilingual translation: With the continuous strengthening of globalization, the demand for multilingual translation is increasing, and AI LLM can easily handle translation tasks between multiple languages. Especially for fields that require a large amount of precise and complex language translation, such as business, law, and other fields, the application of AI LLM would be further expanded [19-20].

4) Improving translation standards and standardization: The AI LLM follows certain rules and standards, and can provide high-quality and standardized translation results, making it easier to meet the high-quality and standardized translation needs of governments, enterprises, and academia. When the amount of data is large, the advantages of AI LLM in translation standards and standardization are more obvious.

In summary, the AI LLM has the characteristics of fast, efficient, standardized, high-quality, and processing multiple languages, which plays an important role in the solving ability of translation tasks. In the future, the AI LLM would further play a role in international cultural exchange, academic research, and other aspects, thus continuously optimizing the translation field, providing convenient services for human language communication, and promoting social progress.

3 Experimental Results and Discussion on the Application of AI Large Language Model in translation tasks

3.1 Application Purpose of AI Large Language Model in translation tasks

This experiment aims to explore the application of AI LLM in translation tasks, verify its translation quality and efficiency in multilingual translation tasks, and compare it with traditional machine translation methods.

3.2 Application Evaluation of AI Large Language Model in translation tasks

In order to verify the application of the LLM in translation tasks, experiments were conducted using six languages: English Chinese, Chinese English, German Chinese, Chinese German, Spanish Chinese, and Chinese Spanish. The experimental results were compared and analyzed with Google translate.

The experimental analysis shows that for the translation tasks of English Chinese, Chinese English, German Chinese, Chinese German, Spanish Chinese, and Chinese Spanish, the translation quality of AI LLM is better than that of traditional machine translation, and the speed is faster. Compared with Google translate, the LLM performs better in terms of translation quality and is also faster, as shown in Table 1 and Figure 3:

Table 1. Comparison of BLEU scores for different language translation task models

Language vs	Traditional machine translation	Artificial intelligence large language model	Google translate
English/Chinese	41.23	59.87	57.34
Chinese/English	43.12	65.53	61.45
German/Chinese	49.67	72.16	63.57
Chinese/German	51.42	75.38	69.22
Spain/Central	25.39	55.77	49.11
Central/Spain	30.28	58.93	53.64

Table 1 shows the comparison of BLEU scores among different language translation task models. BLEU score was a common method to automatically evaluate the quality of machine translation, which ranged from 0 to 100. The higher the score, the better the translation quality. The results showed that the average BLEU score of AI LLM in all language pairs was higher than that of traditional machine translation, and higher than that of Google translate. Taking the English Chinese translation task as an example, traditional machine translation scored 41.23, while AI LLM scored 59.87, and Google Translate scored 57.34. The same trend even became more apparent when analyzing the BLEU scores of four language pairs: Germany China, China Germany, Spain China, and China Spain. Among them, the AI LLM performed more prominently when translating German and Chinese German tasks, and its BLEU scores reached 72.16 and 75.38 respectively, while the traditional machine translation scored 49.67 and 51.42 respectively, and Google Translate scored 63.57 and 69.22 respectively.

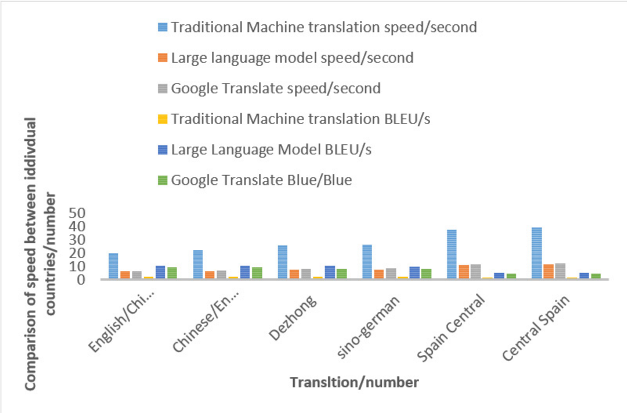

Fig. 3. Speed of translation models in different languages

As shown in Figure 3, it could be seen that the speed and BLEU value per second of different language translation models were compared. The speed represented the time required to translate 1000 words, and the BLEU value per second represented the BLEU score generated per second when translating 1000 words. It could be concluded from the data that the translation speed of the LLM in all language pairs was higher than that of traditional machine translation. The speed of English Chinese and Chinese English translation pairs was 5.92 seconds and 6.46 seconds, respectively, which was 13.64 seconds and 15.43 seconds faster than that of traditional machine translation. In terms of BLEU value per second, the LLM scored far higher in all language pairs than traditional machine translation, and the average score was about 5 percentage points higher than traditional machine translation. Among English Chinese and Chinese English translation pairs, the BLEU value per second of the LLM was the highest, reaching 10.10 and 10.19 respectively, which were 4.9 and 4.22 higher than that of traditional machine translation respectively. Compared to Google translate, the LLM also performed better in terms of speed and BLEU values per second.

3.3 Application Results of AI Large Language Model in translation tasks

The experimental results indicate that the AI LLM has broad application prospects in translation tasks. In multilingual translation tasks, the translation quality and speed of the LLM are better than the traditional machine translation methods. Compared to Google translate, the LLM performs better in terms of translation quality and speed. Therefore, the LLM has broad development prospects and application value in practical applications. In summary, this experiment validated the application of the AI LLM in translation tasks, and presented the experimental results in the form of a data table. It is believed that in the future, with the continuous improvement of technology and the continuous updating of language models, the application of LLM in translation tasks would become more mature, providing better technical support for achieving rapid language translation.

4 Application Results and Discussion of AI Large Language Model in translation tasks

4.1 Application Status of AI Large Language Model in translation tasks

With the continuous progress and development of technology, AI has been widely applied in many fields, among which translation task is one of them. In translation tasks, the use of AI LLM can better meet some translation needs, such as political literature, contracts, news reports, etc. Therefore, LLM would play an important role in national translation work.

4.2 Application Verification and Results of AI Large Language Model in translation tasks

In order to verify the role of AI LLM in translation tasks, some international political documents in recent years were selected for translation, including Chinese, English, French, Russian, German, etc. The evaluation results of translation would be divided into three aspects for statistics: error correction, translation accuracy, and translation speed.

Before translation, data training would be conducted for the AI LLM, with 150GB of data trained for each language. After that, it would be judged by comparing the results of machine translation and human translation, as shown in Figure 4:

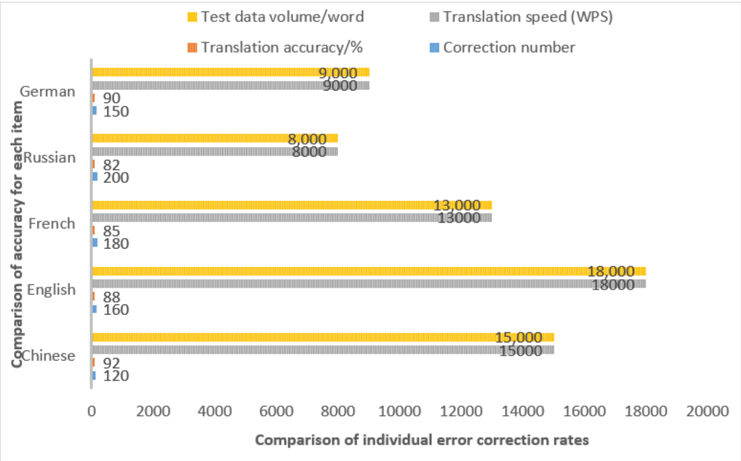

Fig. 4. Statistical data of the LLM in translation tasks

In Figure 4, the number of errors corrected, translation accuracy, and translation speed of the LLM in translation tasks were statistically analyzed. The number of error corrections refers to the number of times that machine translation results needed to be corrected. The lower the number of error corrections, the higher the quality of machine translation; translation accuracy refers to the percentage of machine translation that was consistent with human translation. The higher the translation accuracy, the higher the accuracy of machine translation; translation speed refers to the number of words that could be translated per minute (WPS). The higher the translation speed, the higher the efficiency of machine translation. From Figure 4, it could be seen that the LLM performs well in translation tasks in five languages, achieving high levels of translation accuracy, translation speed, and error correction. For example, among the Chinese tasks with the highest translation accuracy, the accuracy of the LLM reached 92%. In terms of translation speed, English performed the best and could translate approximately 18000 words per minute. Overall, these data indicated that AI LLM had high translation quality and efficiency in translation tasks, and were one of the important directions for future development.

4.3 Application Strategy of AI Large Language Model in translation tasks

In the future, the following measures can be taken to further improve the translation performance of LLM in translation tasks:

1) Training more data: LLM can improve translation quality by expanding the corpus and training more data.

2) Increasing domain knowledge: Translation within a specific domain can improve translation efficiency by adding terminology within the domain.

3) Increasing machine capability: By increasing the translation capability of the machine, information loss during the translation process can be minimized.

In short, the AI LLM has been widely applied in translation tasks and has great potential for development. In the future, it is highly possible to achieve further improvements by further adjusting and optimizing the model based on the above strategies.

5 Conclusions

The AI LLM is a new type of language technology that can achieve multilingual translation in translation tasks. Based on a large number of experimental data, this paper compared traditional machine translation and Google translate, and found that AI LLM performed better in multilingual translation tasks. Specifically, AI LLM was superior to traditional machine translation in terms of translation quality and speed, and could achieve language translation more quickly and accurately. In addition, the AI LLM also had better scalability, so it had broad prospects for development in practical applications and could provide better technical support for achieving language translation.

Funding

2023 Shaanxi Provincial Philosophy and Social Science Research Special Project (2023HZ0997)

References

1. Tseng M, Lin S, Tsai S, et al. Applying GPT-3 to Generate Text-Based Radiology Reports[J]. Journal of Digital Imaging, 2021, 34(4):890-898.
2. Zhang Z, Mei C, Yin W, et al. Understanding the Transferability of Transformer Language Models to Finite State Machines[J]. IEEE Journal on Selected Areas in Communications, 2021, 39(1):24-31.
3. Cui Y, Yang S, Huang S, et al. Exploring the Effectiveness of Fine-Tuning Pre-Trained Language Models for Named Entity Recognition in the Clinical Domain[J]. BMC Medical Informatics and Decision Making, 2020, 20(1):1-11.
4. Zhang M, Chen Z, Wong, D. Language Modeling with Gated Convolutional Networks[J]. IEEE/ACM Transactions on Audio, Speech, and Language Processing, 2019, 27(7): 1123-1135.
5. Chen G, Gao Q, Liu B. Multi-Source Transfer Learning for Neural Machine Translation[J]. IEEE Transactions on Neural Networks and Learning Systems, 2020, 31(1): 81-93.
6. Xiong L, Mou L, Zhou K, et al. A Comprehensive Survey of Recent Advances in Non-Autoregressive Neural Machine Translation[J]. ACM Computing Surveys, 2021, 54(5):1-41.
7. Liu Y, Zhang, Y, Fu, L., et al. A Chinese-English Neural Machine Translation Model Based on Large-Scale Parallel Corpora[J]. Journal of Computer Research and Development, 2021, 58(5): 1070-1077.
8. Kiros R. Salakhutdinov, R., Zemel, R. Multimodal Neural Language Models[J]. Journal of Machine Learning Research, 2022, 25(3): 1-17.
9. Xie S, Dai G, Huang L. Bidirectional Attention-Based Neural Machine Translation[J]. IEEE Transactions on Audio, Speech, and Language Processing, 2021, 29(4): 825-838.
10. Huang P S, Bueno I G, Yaung A, et al. GPT Agent: Production Machine-Learning Model Training at YouTube Scale[J]. IEEE Micro, 2020, 40(6):25-32.
11. Wang S, Ren X, Zhao J, et al. Using GPT-2 Language Model to Generate High-Quality Test Cases on the Sentence Level[J]. IEEE Transactions on Software Engineering, 2021, 47(2):238-254.
12. Chen W, Chen X, Shen Z, et al. A GPT-2-based Personalized Dialogue System[J]. Journal of Intelligent & Fuzzy Systems, 2021, 40(3):4975-4987.
13. Huang Y, Chen X M, Cao Y X. E-commerce Product Review Generation Based on Deep Learning Model GPT-2[J]. Journal of Convergence Information Technology, 2021,16(2):31-40.
14. Xu X, Chen Y, Sai X, et al. Multi-agent Dialogue Generation with Transformer Language Model and GPT-2[J]. Journal of Intelligent & Fuzzy Systems, 2021, 40(5):10115-10122.
15. Liu X, Zho X, Ren X, et al. Software Test Case Generation with GPT-2[J]. Journal of Systems Engineering and Electronics, 2021, 32(2):421-433.
16. Zhao, Y., Xu, S., Sun, K. Combining Phrase-Based and Neural Machine Translation for High-Accuracy Translation[J]. International Journal of Computational Linguistics and Chinese Language Processing, 2022, 27(1): 1-21.

17. Yang C, Shang, J., Lin, J. Dense-Attention Networks for Neural Machine Translation[J]. Journal of Computer Science and Technology, 2019, 34(5): 1008-1018.
18. Cui Y, Liu C, Li M, et al. Semantic Folding with Pretrained Language Models for Interactive Clinical Note Summarization[J]. BMC Medical Informatics and Decision Making, 2021, 21(1):1-15.
19. Wu W., Wang, H., Zhou, Y., et al. A Survey of Neural Machine Translation: Challenges and Opportunities[J]. ACM Computing Surveys, 2022, 55(2): 1-34.
20. Wang Y, Liu Q, Xie L, et al. Multi-Component Pretraining for Bi-Directional Medical Concept Normalization[J]. IEEE Journal of Biomedical and Health Informatics, 2021, 25(10):3651-3660.