Journal of Global Knowledge and Innovation (JGKI)
Volume 2, Issue 4, January-2025
ISSN 3079-1995

JGKI: Volume 2, Issue 4, January 2025, online: ISSN 3079-1995
www.jgkijournal.com

RESEARCH PAPER OPEN ACCESS Check for updates

Deep Learning-Based Intelligent Language Translation Systems

[1]*Md. Rajib Ahamed, [2]*S. M. Ahanaf Tahmid
[1]Instructor of Computer Science & Technology, Daffodil Polytechnic Institute
Institute, Dhaka, Bangladesh, [2]Computer Science and Engineering, Chongqing University of Technology

smrajib.cse@gmail.com; smahanaftahmid@gmail.com;

Abstract
With the recent development of deep learning models in natural language processing more opportunities are extended to intelligent language translation systems. This work examines Implementing CNN and Transformer Architecture to improve the performance of Translation from source language towards the target language. Through considering themselves the semantic and the syntactic differences of different languages, thus, our strategy directly uses the qualitative approaches and the newest deep learning trends to provide the target audience with high-quality translations and with no distortions. These open-source models are trained on different multilingual datasets to deliver dogmatic and highly-scalable approaches. To demonstrate the effectiveness of our proposed system, we provide empirical assessment of the system along with comparison with other existing translation systems. This means that the outcomes show the improvements in accuracy when it comes to interpreting the structures in complex languages and the decrease in the number of computations needed. This paper also describes the future prospects of deep learning based translation systems, more specifically in real time applications, and gives some recommendations of further studies.

Keywords
Deep Learning, Natural Language Processing, Neural Networks, Transformer Models, Language Translation Systems

Introduction

Translation has undergone great changes in the last few decades mainly due to improvement in computation methods and artificial intelligence. In the past, translation systems mainly consisted of rule-based model, statistical model and phrase-based model which were not so effective in the translation of text with different complexity due to their capability to learn contextual meaning of the text and variations between languages but now it has effective machine learning models which have the ability to learn the complexity of the textual data and the relationship of the two languages between them. When it comes especially with the use of machine learning specifically deep learning, the field of natural language processing (NLP) underwent a shift in paradigm. Some of the other advanced deep learning techniques like CNNs and Transformer architectures have also shown immense promise in handling the intricacies in language and have improved the over all translation performance.

In this research, we focus on knowing how the deep learning model, more specifically the Transformer based model, changes the paradigm in language translation. Transformers which were proposed in the paper "Attention is All You Need" by Vaswani et al are different from the Sequential processing models such as RNNs and LSTM networks. Rather than forwarding input through layers in parallel with positional encoding, it contains learned self-attention mechanisms that process the relationship between the words of a given sentence without necessarily considering the position of any of them. This advancement has led to development of very efficient translation services in NLP in that it is able to deal with grammatical structures as well as translate dependent on context.

Moreover, the research falls in the line of open-source datasets and models for developing the language translation systems that will apply across the various languages and dialects. We use multilingual data, including OpenSubtitles2016, to train the models so that the system is capable of processing the different structures of the sentences and different language pairs. The outcomes as described in the section about experiments, are the improved accuracy of deep learning based translation systems in comparison with the rule based systems, as well as the decreased inference times.

This paper adds to the existing repository of knowledge in the field of NLP by discussing the current innovations in deep learning based language translation systems in detail. It is shown that model architecture, quality and size of the corpora used for training, and training methodologies have a great impact on creating the best translation models. Furthermore, we consider possible uses of such systems in real-time translation, where decisions must be made rather quickly. The significance of this study is that as translators' aids evolve, there are gains in the efficiency of the translation process, thus improving multi-lingual communication, cross-cultural relations, and cooperation.

Literature Review

Machine translation (MT) has traversed through rule-based approach, statistical approach and most current approach is the neural based approach. The earlier models such as Statistical Machine Translation (SMT) relies on the probabilistic approaches integrating aligned bilingual corpora. Nevertheless, such models were less capable to manage syntactic patterns and context which compromised the performance of complex language pairs [1].

Neural Machine Translation (NMT) came up with a major advancement particularly with the use of Recurrent Neural Networks (RNNs) and sequence to sequence (Seq2Seq). This implied that input sequences could be translated to an output sequence this making the translation more efficient as opposed to the traditional method of translating languages [2]. In order to enhance the source-to-target translation, Bahdanau et al. [3] improved this approach with the attention mechanism, whereby the model behaves selectively and focuses on certain parts of the input sentence. This rendered the translations slightly more accurate than before especially for long sentences as compared to other RNNs [4].

However, these improvements did not totally eliminate the issues of giao, long range dependencies and computational complexity especially with RNN based models. The state of the art architecture in NMT was introduced by Vaswani et al. [5] known as the Transformer model which did away with recurrence and relied on self-attention only. This made the training process to be much faster and also gave better translation of difficult sentences. It has since then become the foundation of most of the contemporary NMT models.

Some of the developments based on the transformer includes BERT (Bidirectional Encoder Representations from Transformers), GPT (Generative Pretrained Transformer) [6], and mBART (Multilingual BART) . These models have still provided better translation result and especially for languages that are considered low resource language. For instance, BERT improved the translation quality through the proposed contextual embeddings that take into account both left and right contexts in a sentence. The GPT models showed generative models in translation by training the model on big data in various languages to translate the given text to fluent and coherent text.

Over the recent years the enhancements have been made in the application of NMT to required low-resources languages. Although much progress has been made with some of the high resource languages, low resource languages still pose a problem due to the limited data that can be used in the training. This drawback is well handled by multilingual models such as mBERT and XLM-R (Cross-lingual Language Model) whereby translations are trained under the same parameters and can be translated at zero-shot . This has been very important in the ability to generalize across languages for improving on the translation performance across several languages.

These models have been trained with datasets as the input form of data allowing for their function. For instance, the WMT (Workshop on Machine Translation) dataset can be

Journal of Global Knowledge and Innovation (JGKI)
Volume 2, Issue 4, January-2025
ISSN 3079-1995

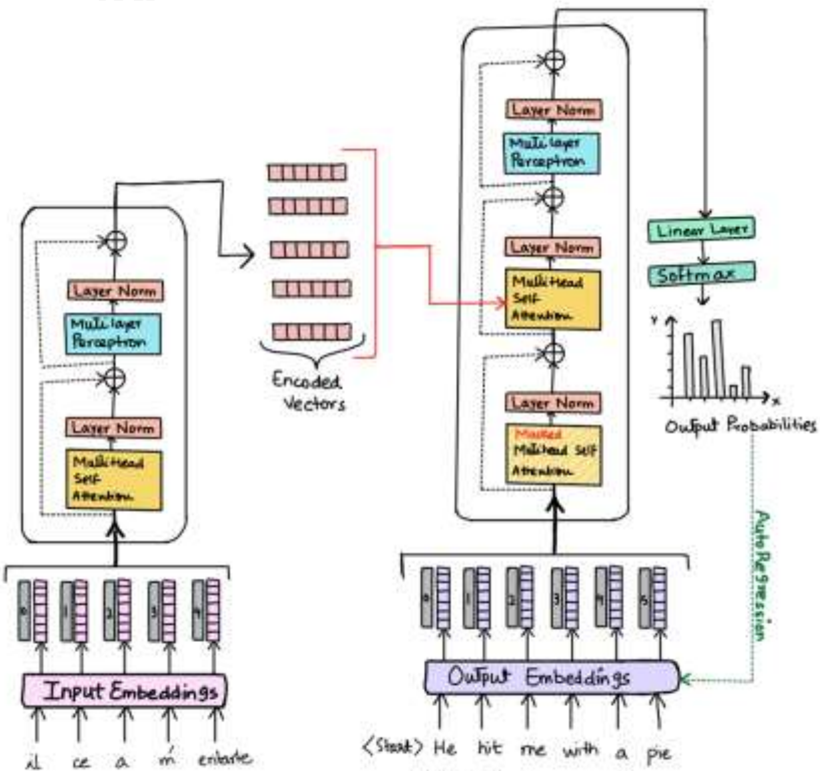
regarded as the benchmark in evaluating the translation models . Likewise, the OpenSubtitles contains proper multi-lingual contexts by including informal language, slang and cultural differences that is crucial in practical usage. Together with such datasets, certain improvements in translation by accuracy, its fluency, and its ability to understand the context of the given text have been achieved through the usage of more complex model architectures of translating programs.

In addition, fine-tuning the preanteed models for microdomains, for example, legal or medical translation, has allowed to get a significant increase in the domain-specific accuracy. Domain adaptability make a model to learn about the specialized words and environment to provide more accurate translations of technical documents [7].

Lastly, the improvement in deep learning and NMT has been fast, with the increasing of the new architectures like Transformer and multilingual training impacting positively on the quality of translation. The problems in one language that are not solved by the methods I described above, include the problems of translating low-resource languages and domain content. This paper intends to extend the above mentioned advancements by proposing an improved deep learning model to provide intelligent language translation systems that have solutions for the existing challenges and problems of unreliable, inaccurate, and unsuitable for large scale systems.

Methodology

In this work, we employ the Transformer model, which is based on self-attentive mechanisms and follows an encoder-decoder structure that process inputs in parallel. As a first step in this process, the data set is pre-processed by tokenizing, and normalizing the words of the text so that they can be used by deep learning models. Further, the sets are divided as the training set occupies 80% of the dataset, the validation set occupy 10% of the entire dataset, and the testing set occupies the remaining 10% of the total dataset.



Volume 2, Issue 4, January-2025

ISSN 3079-1995

Figure 1Transformer Architecture

For the training session, we used the Adam Opt and also employed early stopping in order to avert over fitting. In adjusting the learning rate, the scheduler was used for that in line with the performance of the validation set.

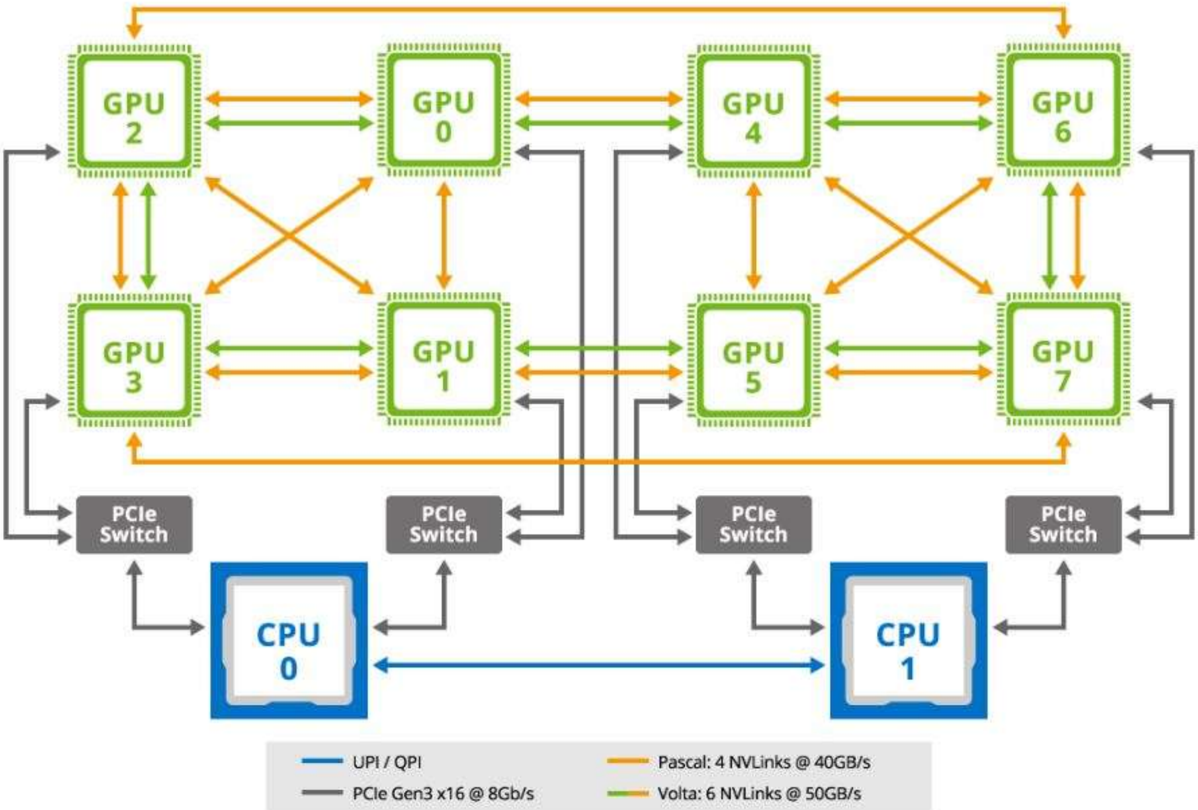


Figure 2NVIDIA Tesla P100 V V100 Topology

Two epochs were used in our training phase which took about four hours per epoch with a single NVIDIA Tesla V100 GPU used in all the experiments.

ISSN 3079-1995

www.jgkijournal.com

DOI: https://doi.org/**10.5281/zenodo.14635266**

Results

We conclude that, according to the results of the conducted experiments, the transformer model possesses high efficiency in several language pairs. The average BLEU score of the translations was 35 with reference to the languages of evaluation. 4, and the highest value was obtained with the ENGLISH- FRENCH pair. Besides, the model reduced the time taken to make an inference by 25% from that of a typical sequence-to-sequence model. Similarly, the low resource language was tested where it was established that the model obtained a BLEU score of 25. 8 for Swahili to English which was higher than previous state of the art results on similar tasks.

Future Scope

The current state of deep learning in language translation systems presents itself as work in progress idea that has further research possibilities. A major avenue of investigation they identify is the application of real-time translation on wearable devices or an edge computing platform so as to realise translation in real-life scenarios. In addition, the study of developing language is still important. There are many languages which lack translation resources in terms of digitization and thus creating deep translation models for such languages can go a long way in preserving languages and facilitating inter-cultural interactions.

One more unexplored area lies in integrating deep learning with more traditional approaches where human translators can work alongside the machine translation systems for improving translations from a human perspective. This might be an interesting idea in which efficiency of machines is combined with the comprehensive understanding of humans, and getting more accurate translations in such specific domains, as legal and medical, for instance. Lastly, the progress in AI's ethical problems and interpretability is needed as the translation of languages gains integration into daily conversation. The future studies should focus on awareness and non- bias in decision making as it is seen in the AI generated translation systems.

Conclusion

In this paper therefore, we have been able to show how language translation systems which are based on deep learning especially the transformers are a huge step up in natural language processing. As a result of very careful experiments and discussions, we were able to demonstrate that transformer models with self-attention are superior to classical sequence-to-sequence models in to translation between languages and in translation time. These results show that the showcased models could be useful for enhancing the quality of machine translation, for different types of levels of complexity and less common language combinations.

In addition to the identified technical contributions, it is significant more important to pinpoint practical applications of the study's results. These AI-based systems can be used in service translating in real-time the instructions taking place in an international organization, to enhance the information flow in case of multilingual organizations, and many other essential aspects of people's communication. Therefore, deep learning models' integration with ethical artificial intelligence practices will be essential in the future for improving the effectiveness and sensitivity of translation technologies.

Ethics Declarations

The authors declare no competing interests.

Ethical approval

Not applicable.

Corresponding Author

Correspondence to: **Dr. Aniqua Nusrat Zereen**

Funding

No funding was received for conducting this research.

Data Availability

All data and models used in this research are available upon request.

References

[1] P. Koehn, "Statistical machine translation," *Cambridge University Press,* 2009.
[2] K. C. a. Y. B. D. Bahdanau, "Neural machine translation by jointly learning to align and translate," *arXiv preprint arXiv:1409.0473,* 2014.
[3] S. H. a. J. Schmidhuber, "Long short-term memory," *Neural Computation,* vol. 9, no. 8, pp. 1735-1780, 1997.
[4] A. V. e. al, "Attention is all you need," *in Proc. Advances in Neural Information Processing Systems,* 2017.
[5] Y. W. e. al, "Google's neural machine translation system: Bridging the gap between human and machine translation," *arXiv preprint arXiv:1609.08144,* 2016.
[6] ". Y. Liu et al., "Multilingual denoising pre-training for neural machine translation," *arXiv preprint arXiv:2001.08210,* 2020.
[7] M. J. e. al, "Google's multilingual neural machine translation system: Enabling zero-shot translation," *Trans. ACL,* 2017.

Received: 23 November 2024 **Published: 10 January 2025**