Neural Machine Translation - Unravel the possibilities

*M. Nithila and V.R. Bithiah Blessie

Department of Computer Science, Holy Cross College (Autonomous), Nagercoil - 629004.

Affiliated to Manonmaniam Sundaranar University, Tirunelveli - 627012.

*Corresponding Author - Email: nithiladavid79@gmail.com

ABSTRACT

Neural Machine Translation (NMT) has emerged as a powerful approach for automating the translation of text between different languages. Unlike traditional statistical machine translation systems, which relied on complex linguistic rules and hand-engineered features, NMT utilizes deep neural networks to learn the translation patterns directly from large amounts of bilingual data. This paradigm shift has significantly improved the quality of more fluent and accurate machine translations.

Introduction

Neural machine translation revolutionizes the translation process by utilizing artificial intelligence to automatically convert text from one language to another, eliminating the need for human involvement.

The main objective of NMT is to produce accurate translations that capture the essence of the source language and convey it effectively in the target language. Traditionally, human translators have been relied upon for this task, but now machine translation has made remarkable advancements and is making a positive impact worldwide. The days of depending on someone else to translate a text into your native language or any other language are diminishing. Thanks to smartphone applications, anyone can easily comprehend the content on signboards, books, and more.

Machine translation employs various methodologies to achieve its goals

Rule-based Machine Translation: RBMT systems rely on explicit linguistic rules to perform translations. These systems follow predetermined instructions on how words or phrases in the source language should be rendered in the target language. However, this approach demands significant manual effort and poses challenges when scaling to multiple languages.

Statistical Machine Translation: This model relies on a large language resource containing extensive and structured sets of texts. These models are trained to identify the most probable translation based on the frequency of word sequences.

Neural Machine Translation: Neural machine translation is an approach that utilizes artificial neural networks to predict the probability of a sequence of words. It has emerged as the predominant model in modern machine translation.

Principles of Neural Networks

In a neural network, there exist three layers: the Input Layer, Hidden Layers, and Output Layer. The Input Layer consists of independent variables, denoted as X, which serve as predictors. These inputs are gathered from external sources like textual data, images, audio, or video files. In a neural network, these Xs represent information received from sensory organs.

The Output Layer produces the result of the neural network. In regression problems, it can be a numerical value, while in classification problems, it can be a binary or multi-class label.

Additionally, the output can include tasks such as handwriting recognition, audio or image classification, or text categorization.

Apart from the Input and Output layers, Neural Networks incorporate Hidden Layers that extract features for the model.
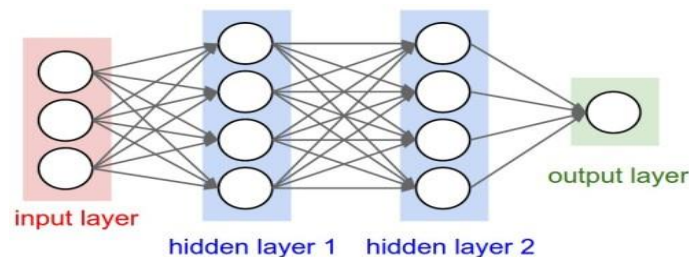
Deep Learning is an advanced technology founded on neural networks that aims to mimic the workings of the human cortex. Recurrent Neural Networks (RNN) serves as fundamental network architectures upon which other deep learning models are constructed. RNNs encompass a diverse range of deep learning architectures and possess the ability to process sequences of varying lengths using their internal state (memory). Let's consider that RNNs possess memory, capturing, storing, and utilizing processed information to compute the outcome.

Currently, two types of RNN can be identified:

❖ Bidirectional RNN: These operates in two directions, enabling the output layer to simultaneously gather information from past and future states.

❖ Deep RNN: These consists of multiple layers, allowing the deep learning model to extract more hierarchical information.

RNNs are primarily employed in text translation.

In addition to RNNs, there exist other models in neural networks, including Long Short-Term Memory, Convolutional Neural Networks, Deep Belief Network, and Deep Stacking Network.

Frameworks for constructing NMT applications

Several popular frameworks are commonly utilized for building neural machine translation (NMT) applications. Here are a few examples:

❖ TensorFlow

❖ PyTorch

❖ OpenNMT

❖ Marian

❖ Fairseq

❖ Nematus

These frameworks represent only a subset of the commonly used options for NMT applications. The choice of framework often depends on specific requirements and familiarity with the framework.

How to Train Machine Translation Engine

A machine translation engine offers many advantages, with reduced translation times and minimized use of human resources being the main benefits. AI-powered machine translation engines leverage data to identify correlations and structures, extracting information from vast amounts of data to tackle problems that would require thousands or millions of hours for a human to process.By incorporating machine learning and deep learning techniques, the capabilities of a machine translation engine are enhanced, allowing for continuous improvement in the results provided. However, achieving enhanced translation quality relies on effective training.

The training process for a machine translation engine can be summarized in four steps:

Incorporation of base data

Training an engine requires high-quality data in the form of translated sentence examples from the source language to the target language.

Data cleaning and normalization

After collecting raw data, it is essential to clean and normalize it, which includes tasks like ensuring the correct use of quotation marks in both languages. Once this is done, the translation engine can be fed with appropriate data.

Possibility of sentiment analysis

Advanced technologies enable translation engines to analyze the sentiment of texts, understanding and considering the true meaning of a text or the speaker's intention during translation. This involves combining machine learning and NLP.

Maintenance

Basic training can span several days, and measures such as a 'stop criterion' allow the engine to automatically determine when it has stopped learning anything new, thus avoiding wastage of time. Additionally, in the case of specialized models for a specific domain, training will be performed using the available data, with the level of specialization determining the aggressiveness or conservatism of the training.

Achieving optimal results necessitates continuous training beyond the initial phase.

Advantages and Disadvantages of NMT

Let's begin with the advantages of using machine translation from the perspective of professional translators.

❖ **Speed:** Provides nearly instant translations.

❖ **Scalability:** Capable of handling vast amounts of content rapidly.

❖ **Flexibility:** One system can translate content into multiple languages.

❖ **User-friendliness:** Easy to use for both language professionals and everyday users.

❖ **Integration:** Language professionals can use machine translation to expedite their workflow.

❖ **Automation:** Automating the initial translation stage accelerates the entire process, making it more affordable for the end client.

These advantages contribute to an efficient translation process, allowing translators to automate the first step and edit the output to deliver the final content faster. This reduces turnaround times and enables professional translators to deliver 100% quality at lower rates compared to translating without machine translation.

The main disadvantages of machine translation are related to output quality:

❖ **Quality:** Even the best AI translation tools are still far from matching the quality of professional translators.

❖ **Consistency:** The quality of machine translation varies significantly based on the complexity of the input language and the linguistic gap between the source and target languages.

❖ **Word-for-word output:** Despite advancements, algorithms still predominantly produce translations that closely resemble word-for-word equivalents.

❖ **Grammar:** Although significant progress has been made in recent years, machine translation still faces challenges with grammar, particularly when dealing with languages that have substantially different grammatical systems.

❖ **Context:** While AI technologies have made significant strides in contextual understanding, the final results are still far from matching the capabilities of humans.

❖ **Nuance:** Algorithms struggle to grasp and replicate the intricate nuances of human language.

Despite the significant advancements in artificial intelligence over the past few decades, the quality of machine translation output has only seen marginal improvements compared to the abilities of professional translators. This demonstrates the complexity of language comprehension and why the human brain remains the most powerful "computer" when it comes to the subtle aspects of communication.

The Future of NMT

Lionbridge's R&D teams estimate that Neural Machine Translation is improving at a rate of 3-7% per year. Current machine translation systems often face difficulties in capturing context and generating translations that account for the surrounding text. Future models are expected to incorporate better contextual understanding and consider the broader context to produce more accurate and contextually appropriate translations.

In the future, machine translation systems have the potential to translate spoken conversations in real-time, translate text within images or videos, and offer broader translation solutions across different forms of media.

Effective cooperation between humans and machines holds the key to the future of machine translation. Human translators will remain essential in tasks such as post-editing, ensuring quality, and fine-tuning translations generated by machines. The development of user interfaces and tools will enable smooth collaboration between human translators and machine translation systems.

Conclusion

Despite its success, NMT still faces challenges such as handling rare or out-of-vocabulary words, translating idiomatic expressions, and handling language pairs with limited parallel data. Ongoing research aims to address these challenges through techniques such as sub word modelling, transfer learning, and reinforcement learning.

References

https://summalinguae.com/language-technology/rule-based-machine-translation-vs-statistical-and-neural-machine-translation/

https://www.geeksforgeeks.org/introduction-to-recurrent-neural-network/

https://en.wikipedia.org/wiki/Recurrent_neural_network

https://towardsdatascience.com/choose-the-right-transformer-framework-for-you-b7c51737d45

https://blog.pangeanic.com/train-machine-translation-engine

https://www.translateplus.com/blog/what-are-advantages-disadvantages-using-machine-translation/

https://www.lionbridge.com/blog/translation-localization/the-future-of-language-technology-the-future-of-machine-translation/