KAPITEL 7 / *CHAPTER 7* [7]
NEURAL MACHINE TRANSLATION: STRENGTHS AND WEAKNESSES

DOI: 10.30890/2709-2313.2023-18-04-019

Introduction

Neural Machine Translation is a new and advanced method of translation of written texts from one natural language to another by using neural networks. These large networks directly transform sentences from the source language into the target language.

Some key features of Neural Machine Translation (NMT) and the role of neural networks in the translation process are described herein. The mechanism of neural machine translation, its specific features, its differences from other machine translation systems and the system limitations are also analyzed.

7.1. NMT – a different approach to translation

NMT systems use artificial neural networks that are trained on a large number of pairs of parallel sentences ('parallel corpora'). These networks can read a word or a sentence in the source language and translate them into a target language. However, word matching and breakdown into phrases is no longer needed . This seems to be the main difference between the NMT system and other machine translation systems, such as Rule-based or Statistical MT. In order to create a NMT system it is required to provide the availability of several million pairs of sentences translated by human translators All modern NMT systems are equipped with encoder-decoder and 'attention' mechanisms. The unique role of the 'attention' mechanism is to predict subsequent words during the translation process. While focusing on one or more words of the original sentence, it adds this information to the encoded full text. This process is similar to the behavior of a human translator who first reads the entire sentence and then looks at individual source words and phrases already translated or yet to be translated. In spite of its advantages, like fluency, NMT systems have also a number of drawbacks. The most frequent are adequacy errors, as well as omissions and additions of content. Transfer of semantic content from the source to the target language often produces mistranslations. The source phrases need to be very clear, coherent and void of ambiguity to prevent translation errors.

[7] ***Authors:*** *Shevchenko Olga Mykolaivna, Ogurtsova Olga Leonidivna*

For decades, scientists have been trying to develop new translation methods to improve translation efficiency. The idea of using neural machine translation method appeared as far back as in the early 1950s. But it was not until 2013 that this method found its further development and practical application.

In 2013 N.Kalchbrenner and P.Blunsom from Oxford University proposed a new translation method which can be regarded as the birth of the Neural Machine Translation (NMT). In 2016, Google introduced the Neural Machine Translation System (GNMTS) to improve the performance of Google's translation service. Google started using NMT, which replaced Phrase-based Machine Translation system (a variety of Statistical Machine Translation (SMT)) previously used by the Google Translate service.

The Neural Machine Translation System (NMTS) uses a new approach which is different from previous methods of machine translation and is very similar to the work of the human brain. NMTS uses a deep learning technology as part of Machine Learning and Artificial Intelligence (AI). This technique allows the system to imitate human behavior through learning. The NMT system has clearly demonstrated its capability of self-learning which resulted in further improvement of the efficiency of neural machine translation.

In recent years a great number of researchers have been actively engaged in studying and developing Neural Machine Translation as a new approach to machine translation. This method was first proposed by N.Kalchbrenner and P.Blunsom, and later by I.Sutskever and D.Bahdanau. It differed from previously known popular MT methods such as Statistical Machine Translation (SMT) which is based on the analysis of existing translations from bilingual text corpora and Rule-based Machine Translation (RbMT) which is a matching translation system based on linguistic information about the source and target languages.

7.2. NMT translation model

How does the NMT system work? All NMT systems comprise basically three main components: the encoder, the attention mechanism, and the decoder. The decoder-encoder mechanism is an indispensable part of the system and is used to predict subsequent words with certain properties according to the sentence structure. The input sentence is first encoded into an abstract set of numbers and the sentence is analyzed in terms of grammar (context) and segments (words) contained therein. Then the decoder, having "read" the sentence to the end and simultaneously from left to right

and from right to left, begins to decode and predict words. Each predicted word is used to predict the next word and so on. Finally, the target sentence is generated. During this process, the attention mechanism plays an important role as it helps the decoder to analyze different parts of the input sentences and helps the system to memorize long sentences. The following diagram (Fig.1) shows the neural machine translation model:

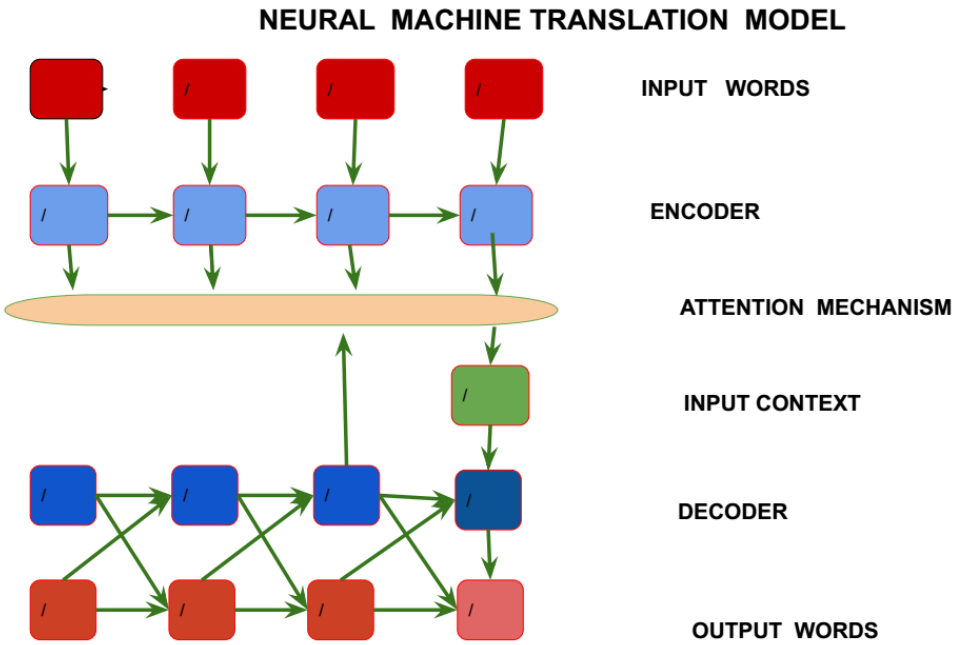

Fig.1

In order to create an effective NMT system it is required to have a huge bilingual corpus that contains text in the source language with the equivalent translation in the target language.

In the process of translation the source word is analyzed in terms of its morphological structure, grammatical category and grammatical role in a sentence. After that the system finds the same word representation in the target language.

NMT system is trained on a large number of pairs of sentences and also it uses 'knowledge' obtained previously. Word matching and breakdown into phrases are no longer needed. Thus, for a more precise choice of a word for translation, the context of the entire original sentence is used, as well as the context with all previously predicted words. And this is more like how a human translation occurs. In the process of training of a neural system, there is a constant comparison of each predicted word with the 'correct' word, i.e. with the word that is used in a target training sentence. .If the predicted word does not coincide with it, then the parameters of the model are updated. This process is repeated multiple times and, as a result, the neural system goes through

a multitude of pairs of sentences which are mixed in random order several times. The process stops when there is no further improvement in word prediction.

According to some estimates NMT outperforms other types of machine translation systems in fluency and accuracy. It also gives better results in terms of inflection and reordering. In spite of some obvious deficiencies, the translations produced by NMT systems contain 50% fewer word order errors, 17% fewer lexical errors and 19% fewer grammatical errors than traditional machine translation systems.

7.3. Some drawbacks of NMT

Along with the obvious advantages NMT also demonstrates certain weaknesses.

The translation quality deteriorates in case the sentence or its segments are too long. NMT may become totally ineffective if a sentence comprises more than 20 words. Satisfactory results in solving this problem can be achieved by using a long-sentence preprocessing technique. The method uses the extraction of bilingual phrases and creation of bilingual corpus based on these phrases.

NMT system has difficulty with translating rare and low-frequency words. Also, its overall performance remains relatively slow and requires post-editing.

In many cases NMT remains domain-specific which means that the system is trained to translate texts from specific domains, e.g. business, e-commerce, legal, medical, etc. Besides the need to prepare large amounts of data for training, a significant computing power is required to train the system.

Neural machine translation aims at building a single neural network which can be trained to achieve the best translation results.

An important prerequisite for starting the NMT process is the availability of a dataset (parallel corpora) in both source and target languages.

The neural machine translation can be of two types: the classical NMT and NMT with attention mechanism. Both of them were studied and the results showed that NMT with the attention mechanism has significantly outperformed the classical NMTin its performance.

Nowadays NMT is widely used to increase fluency and accuracy in translation. However, adequacy still remains a major problem in NMT In spite of its obvious advantages over other types of translation systems. NMT generates numerous adequacy errors during the translation process. It means that a great deal of post-editing work on the part of human translators is still required. There can be different types of adequacy errors made by the NMT system including omissions, additions or

mistranslations. Some typical examples of adequacy errors produced by the NMT system are given in Table 1.

The following translations were generated by the neural machine translation services, such as Google Translate, DeepL and Bing Microsoft Translator.

Table 1. NMT Adequacy Errors (Translations generated by Google Translate, DeepL, Bing Microsoft Translator)

Category	Source sentence	Google Translate	DeepL	Bing Microsoft Translator
Verb Tenses,	Я іду на кухню (eng. I go to the kitchen)	I'm going to the kitchen.	I'm going into the kitchen.	I go to the kitchen
	Вчора він дві години писав статтю (eng. Yesterday he was writing an article for two hours)	Yesterday he wrote an article for two hours	Yesterday he spent two hours writing an article	Yesterday he wrote an article for two hours
Passive Voice	The girl was danced by Anthony Brown (укр. З дівчиною танцював Ентоні Браун)	Дівчину танцював Ентоні Браун	Дівчину танцював Ентоні Браун	Дівчину танцював Ентоні Браун
	The boys were slept in a small room (укр. Хлопців поклали спати в маленькій кімнаті)	Хлопці спали в маленькій кімнаті	Хлопці спали в маленькій кімнаті	Хлопці спали в маленькій кімнаті
Idiomatic expressions	He had a narrow escape (укр. Він ледве втік)	Йому вдалося втекти	Йому вдалося врятуватися	У нього була вузька втеча

Category	Source sentence	Google Translate	DeepL	Bing Microsoft Translator
	To win the match today will be a piece of cake (укр. Виграти сьогодні матч буде простіше простого)	Перемогти в сьогоднішньому матчі буде просто шматком пирога	Виграти матч сьогодні буде простіше простого	Виграти матч сьогодні буде шматок пирога
	Did you understand what he said? – No, it was all Greek to me (укр. Ви зрозуміли, що він сказав? – Ні, для мене це була китайська грамота)	Ви зрозуміли, що він сказав? – Ні, для мене все було грецьким	Ви розуміли, що він говорив? – Ні, для мене це була грецька мова	Ви зрозуміли, що він сказав? – Ні, для мене це все було грецькою мовою
Gender	The child is lovely. I like it (укр. Дитина мила. Вона мені подобається)	Дитина мила. Мені це подобається	Дитина чудова. Мені подобається	Дитина прекрасна. Мені подобається
	The box was closed. He hoped to try to open it (укр. Коробка була закрита. Він сподівався спробувати її відкрити	Ящик був закритий. Він сподівався спробувати її відкрити	Коробка була закрита. Він сподівався спробувати її відкрити	Коробку закрили. Він сподівався спробувати відкрити його

Conclusions

It should be noted that NMT in spite of its growing popularity is still facing some problems which have to be overcome to make this method more effective and flawless. The major factors for future development of NMT are the growing volumes of big data and online content which requires an efficient and robust machine translation system able to cope with the emerging challenges.