iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 10 **Issue:** VIII **Month of publication:** August 2022

DOI: https://doi.org/10.22214/ijraset.2022.46380

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 10 Issue VIII August 2022- Available at www.ijraset.com

Language Translation Using Artificial Intelligence

Rupayan Dirghangi[1], Koushik Pal[2], Sujoy Dutta[3], Arindam Roy[4], Rahul Bera[5], Manosijo Ganguly[6], Dipankar Pariary[7], Karan Kumar[8]
[1, 2, 3, 4, 5, 6, 7, 8]*Department of Electronics and Communication Engineering, Guru Nanak Institute of Technology*

Abstract: ***Language is the unit of our whole communication system. As we all know there are many languages in our world depending on their geographical location, so to remove this language differences or language barrier ,here we explain language translator system in this paper.***
It focuses on the translator system from English to Bengali and vice versa using machine translation.
Bangla is widely spoken language and it is the fifth most spoken native language and the seventh most spoken language by the overall variety of speaker in the world. But there are only a few researches in Machine Translation (MT) for Bangla have been seen in the literature previously. Therefore this paper represents to explain a MT system for English-Bangla and Bangla-English translation. Our focus is to help the students to feel a comfort to understand their lesson in different languages.
Machine Translation is a revolutionary technology , and day by day it improves. The newest phase of machine translation is Neural Machine Translation (NMT), which is based on the development of artificial intelligence. In this paper we are proposed a type of English-Bangla language translator which based on the Encoder-Decoder model of NMT.
Keywords: Machine Translation , Artificial Intelligence, Computer Aided Translation, Neural Machine Translation, Encoder-Decoder model, LSTM.

I. INTRODUCTION

As we all know there are 7.97 billion people in our world, so the differences of communicating languages are seen in the different areas. Thus there are many types of language barrier in different regions.That's why the communication process are going to be difficult for the different regions people. It is clear that translation is important to spread information. Translation is the process which is converting one language to another and now a days the automated translation system plays a vital role as a translator. There are two varieties of translation , one is translation by a single word and another one is translation by a single sentence. In this paper we are mainly focused on to translate from English to Bengali language and vice versa using machine translation.
Now, Machine Translation is a sub field of computational linguistics that refers to translate texts or documents from the source language to required language without any human intervention.
Our paper is especially based on the translator system using machine translation for the students , because in West Bengal , many students feel this difficulty when they are going to fulfill their higher educational activity in any other state in India and foreign countries also, as they are already habituated to speak and continue their study in bengali. So, during a short time if they get the chance to go somewhere except bengal to study further then they will find the difficulty to habituated in other language.
Bangla is the fifth most spoken native language and the seventh most spoken language by the overall variety of speaker in the world. So, the main tusk or aim is to translate from the English language to Bangla language and vice versa with approximately 228 million native speaker and another 37 million as a second language speaker.
In the field of translation , machine translation is revolutionary technology which ignited a reform and transformation. It has three stages, from early dictionary-matched machine translation to corpus-based statistical computer-aided translation(CAT), and then to neural machine translation with artificial intelligence as its core technology in recent years. Neural Machine Translation (NMT) is the newest technology of machine translation using artificial intelligence and it has made major breakthroughs in machine translation.
Among the various MT approaches, neural MT (NMT) is the most up to date also as promising methodology that relies on the development of artificial intelligence. NMT is based on the data driven technique which consists of special neural networks (NN).

II. STAGES OF MACHINE TRANSLATION(MT, CAT AND NMT)

Machine Translation (MT) is an automated translation to translate text or speech from one language to another performed by a computer. It need not any human involvement and it is totally based on the computer algorithms. The development of computer technology, information theory, linguistics and other disciplines have closely linked with the development of machine translation.

MT requires complex cognitive operations to perform a seemingly mundane task: decoding a source text and recording into the target language. Rule-based MT (RBMT), statistical MT (SMT), and example-based MT (EBMT) are the three common methods of machine translation.

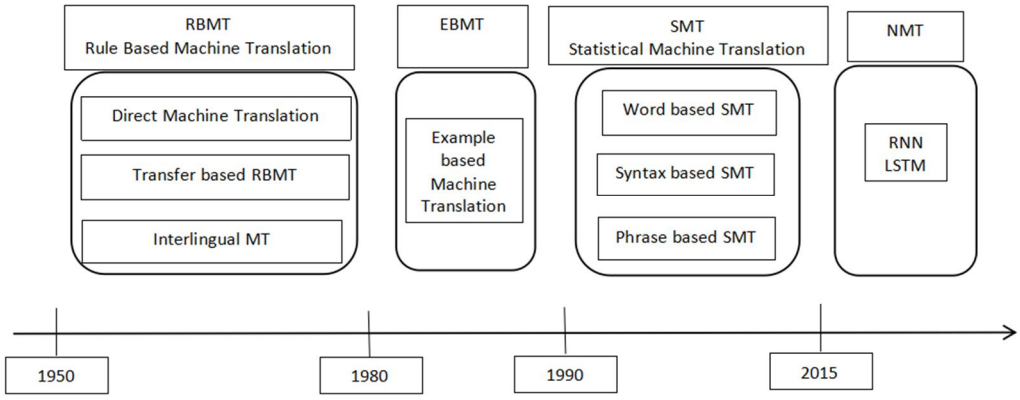


Figure. 1 A Short History of Machine Translation

Previously we have discussed that it has three stages, at first proper machine translation (It has three methods RBMT, EBMT and SMT) was there, then it changes into corpus-based statistical computer-aided translation (It is under the SMT or Statistical Machine Translation) and then it transformed into the neural machine translation which is the best stage of the machine translation in the recent times. Neural Machine Translation or NMT is totally based on the technology of artificial intelligence. Since our paper represents the main focus of the language translation using artificial intelligence , NMT or Neural Machine Translation plays a big role of our paper.

Computer Aided translation (CAT) appeared in machine translation and after digitization of the translation materials this type of technology has been rapidly developed. It is different from previous machine translation software. It does not rely on the automatic translation of machines, for it is the completion of the entire translation process with the participation of human.Computer Aided Translation tools usually depend on translation memory to provide translators with translation suggestions and corrections. Especially , CAT tools help a translator to work faster and more preciously from their previous translation. This is especially helpful for those translators, who focus their translation skills on specialized areas or industries (medicine, agriculture, legal). The core part of computer assisted translation is the memory database . For translating the exact sentence , CAT can produce translation with the help of database. Working mechanism of CAT is shown in the Figure. 2.

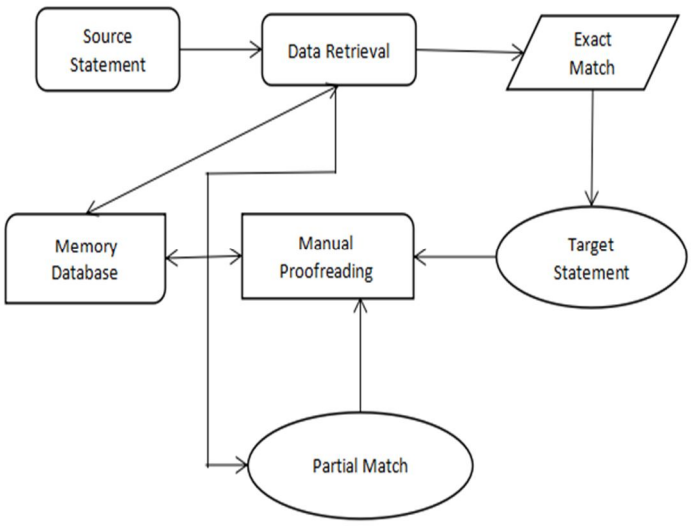


Figure. 2 Working Principle of Computer-aided Translation

Neural Machine Translation is the newest version of machine translation which is totally based on the artificial intelligence. The theory of NMT is based on the theory and techniques of natural language understanding (NLU), natural language processing (NLP), machine translation (MT), translation memory (TM), and statistics-based machine translation (SMT)as well as deep learning. It overcame the barrier of the computer translation technology and achieved high quality machine translation as a new theory.
In 2016, Google introduced their Neural Machine Translation System (GNMT) and GNMT had done the job with very difficult translation production of Chinese-English language pairs, which made a revolution in the translation industry.

III. METHODOLOGY

In this paper , we are mainly focused to help the students who really suffer to grab any other language to fulfill his dream of higher education in any other state and foreign country. This paper represents the English to Bengali translation system and vice versa using the neural machine translation which is the newest phase of the machine translation. This neural machine translation is based on the technology of artificial intelligence. If we want to make a system for the students , we have to implement a technology or software for it. But in this paper we will discuss what type of the technology or what type of model of NMT we are looking for. NMT introduced as a replacement approach with the potential of addressing several shortcomings of traditional machine translation systems. Its design consists of two recurrent neural networks (RNNs), one is use to absorb the input text sequence i.e. encoder and one is use to get translated output text i.e. decoder. As a human, we read the full supplied sentence or text, then understand its meaning, and then provide a translation, NMT or Encode Decoder model follows the same procedure of translation.

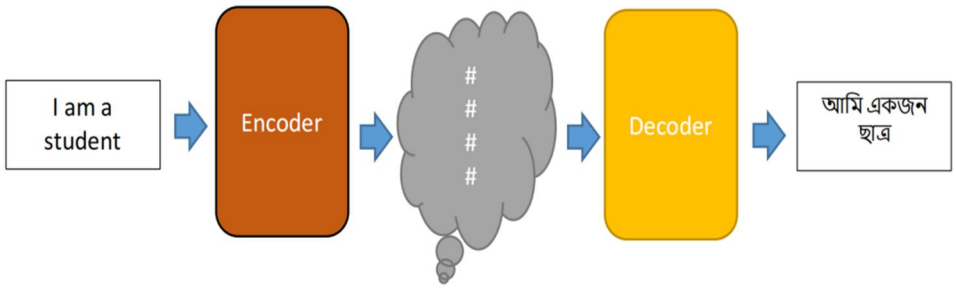


Figure.3 Encoder-decoder architecture

Encoder-Decoder Workflow in NMT (Neural Machine Translation)

A. Procedure

1) Encoder LSTM(Long Short-Term Memory) outputs: We only keep the state outputs of encoder LSTM layer as it will contain all the information about the input data.
2) This states of encoder LSTM will be used to initialized the decoder LSTM. Also, the [start] token will be provided as first word as we are performing teacher forcing.
3) The output of this decoder LSTM layer will be passed through Dense layer to predict the output word.

B. Workflow

1) Encoder side : Input -> Encoder LSTM -> Encoder States.
2) Decoder side :
a) Encoder States + [start] -> Decoder LSTM -> Word + Decoder States
b) Decoder_States + Word -> Decoder LSTM -> Word2 + Decoder States2
c) Word2 + Decoder States2 -> Decoder LSTM -> Word3 + Decoder States3 …..so on.
d) The process stop when end token is predicted.

IV. ROLE OF ARTIFICIAL INTELLIGENCE

Artificial Intelligence (AI) allows devices or software to carry out tasks without the human help. There are various types of AI, but using processing power to detect patterns in data sets, and comparing different uses of patterns to find solutions is mostly use to do some specific task.

It is known to all why AI applies to translation. Languages are mainly large data sets of words and meanings, and they are governed by rules determining how those words are used.

AI translation is nothing but a machine learning technique over the languages. It is more applicable than older day's machine translation as it is translate every word without any misunderstanding and also understand the phrases, complex sentences structures,tones of voice and even jokes or slang.

Especially AI is used on the NMT, which is the newest phase of machine translation.In 2016 Google introduced the NMT using artificial intelligence with us, which is called GNMT or Goggle's Neural Machine Translation. It is a revolutionary job in the field of language translator as well as artificial intelligence.

Advantages & Disadvantages of using AI translation

AI translators have some major advantages and disadvantages also, including:

A. Advantages

1) *Time:* Machine translation is much faster than human translation. The human translator can translate the 2500 words in a day but the machine translator can translate 1000 words in a minute.
2) *Cost:* In human translation the cost is mere fraction of the machine translation.
3) *Adaption:* Machine translation can recognize the key terms. This helps to translation that are very consistent to something that's more difficult when using multiple human translators.

B. Disadvantages

1) *Cultural misunderstandings:* Humans are more intelligent and knowledgeable than the machine so accurate translations can still be misleading to local people and humor is something that computers aren't Well-equipped to analyze.
2) *Accuracy:* Machine cannot translate to perfectly accurate on a regular basis. The technology has improved day by day but still its work is going on. So it means the original documents are not perfectly accurate.
3) *Blandness:* Computers are no doubt good at repetitive tasks with clear outputs, but they aren't good at making texts engaging and fun to read.Texts made using AI translation can often be bland and neutral that is a good thing but branding needs to be a personality that can provide a humans.

V. LITERATURE SURVEY

A. In Paper 1

English to Bangla Machine Translation Using Recurrent Neural Network . In 2, June 2020, this paper was published by Shaykh Siddique, Tahmid Ahmed, Md. Rifayet Azam Talukder, and Md. Mohsin Uddin in International Journal of Future Computer and Communication. They said in their paper was that an architecture of English to Bangla machine translation system. The system has been implemented with the encoder-decoder RNN (Recurrent Neural Network). It helps to find out the structure of the machine translation of English to Bangla and the efficient activation functions from the paper.

B. In Paper 2

Artificial Intelligence Based Language Translator Platform. In 2021 , this paper was published by Manjur Kolhar and Abdalla Alameen. It was published form Prince Sattam Bin Abdulaziz University, Wadi Ad Dawaser, 11990, Saudi Arabia. They said in their paper was that a machine translation (MT) system is developed as alearning technology. The proposed system can be linked to a digital podiumand projector to reduce multitasking.

C. In Paper 3

A Review on Machine Translation in Indian Languages. In 2018, this paper was published by Deepti Chopra, Nisheeth Joshi, Iti Mathur in Engineering, Technology & Applied Science Research. They said in their paper was that different approaches of MT, the problems faced in MT in Indian languages, the limitations of some of the current existing MT systems, and present a review of the work that has been done until now in MT in Indian language perspective.

D. In Paper 4

Neural Machine Translation System for English to Indian Language Translation Using MTIL Parallel Corpus In March 20,2019 , this paper was published by Premjith B., M. Anand Kumar, Soman Kp. They said in their paper was that proposed a neural machine translation(NMT) system for Indian languages. They also collected sentences from different sources and cleaned them to make four parallel corpora for each of the language pairs, and then used them to model the translation system.

E. In Paper 5

Transformer Deep Learning Model for Bangla-English Machine Translation.In June ,2021, this paper was published by Argha Dhar, Arna Roy, Md. Ahsan Habib, M.A.H. Akhand in International Conference on Artificial Intelligence: Advances and Applications (ICAIAA 2021)At: India.This paper aims at developing an MT system for Bangla-English translation. They used modern deep learning-based transformer model for this language pair as it worked well for other language pairs.The proposed model is tested on a benchmark of Bangla-English corpus, which out-performed some prominent existing MT methods.

VI. FUTURE AIM

The success rate of Language Translation is improved day by day but in the case of most sweetest language 'Bangla', it is less explored than any other languages till now . In today's life, we find many online free language translator website but their we see that language translation accuracy level is very much low. Many complicated words in Bengali languages can't be expressed properly in English languages by the software. It also can't understand the different feelings , gender difference and still have many problem with uses of preposition. So we need to look at this problem and In future we are trying to make a translator software which is totally based on the AI, which can minimize the accuracy of all this problem and then in this translator we are also trying to add more different types of global languages for making a language barriers less world.

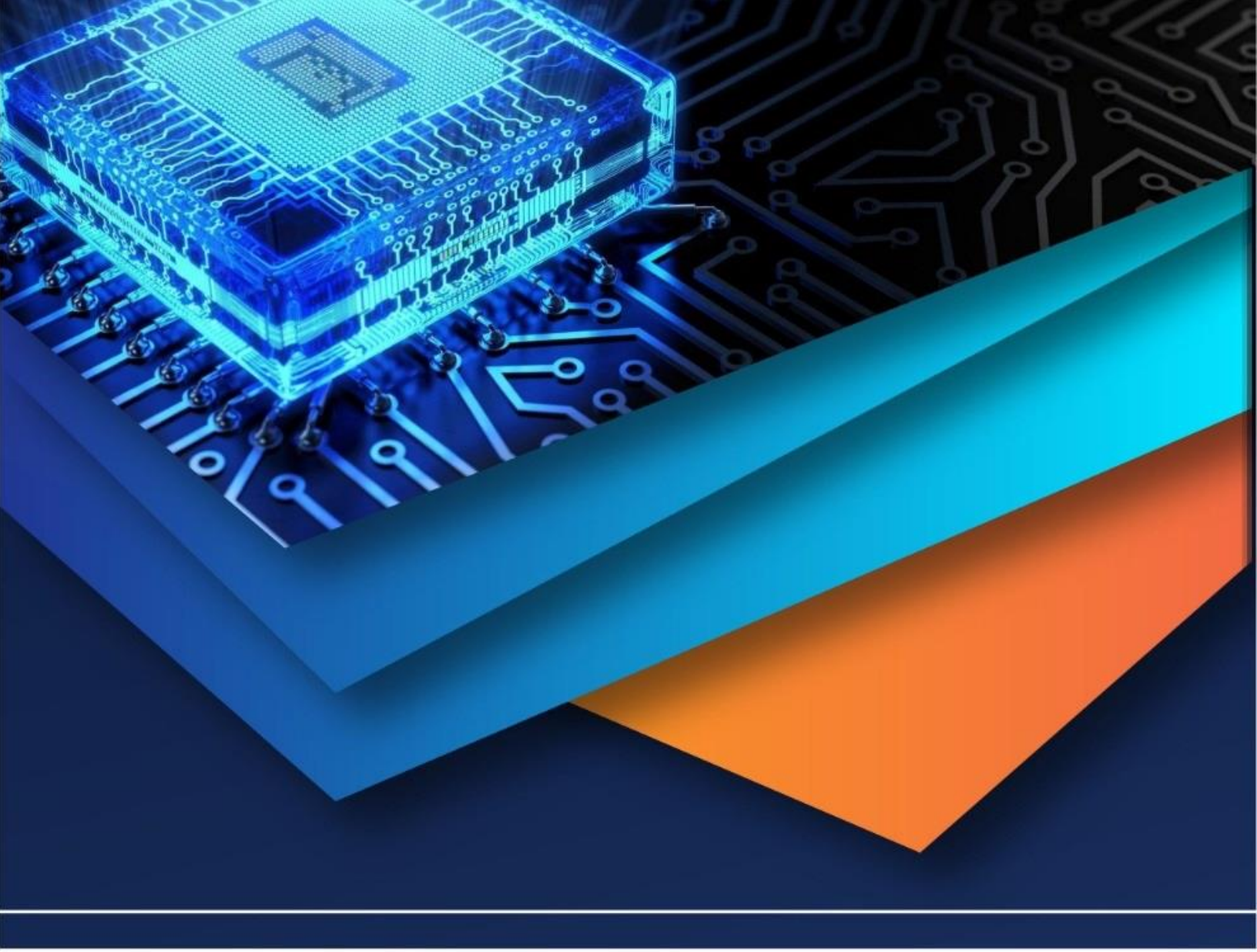
VII. CONCLUSIONS

As we seen that AI has played an important role for translation from English to Bengali and vice versa. AI translator has huge amounts of potential when it comes to language learning, understanding the different languages and etc etc. Students can be benefited by it, because some students usually goes to foreign or other state of their country for higher studies, so they have to face difficulties and a language barrier. Language translator using AI will be reducing their difficulties to understand their language and habituated with it.

It is not only beneficial for the students but also it has global aspects. In few countries , people does not understand English, they only know their local language. So removing this problem, it is very important to implement the language translator system.

REFERENCES

[1] Argha Dhar, Arna Roy, Md. Ahsan Habib, M.A.H Akhand, Transformer Deep Learning Model for Bangla-English Machine Translation, ICAIAA, June 2021.
[2] Manjur Kolhar, Abdalla Alameen, Artificial Intelligence Based Language Translation Platform, Intelligent Automation and Soft Computing, January 2021.
[3] Zhaorong Zong, Research on Relations Between Machine Translation and Human Translation,Journal of Physics Conference Series, September 2018.
[4] Abinash Dutta, Partha Pakray, Shivaji Bandyopadhyay, Sahinur Rahman Laskar, Neural Machine Translation: English to Hindi, IEEE Conference on Information and Communication Technology, December 2019.
[5] Shamsun Nahar, Mohammad Nurul Huda, Md. Nur-E-Arefin, Mohammad Mahbubur Rahman, Evaluation of Machine Translation Approaches to Translate English to Bengali, ICCIT, December 2017.
[6] Shaykh Siddique, Tahmid Ahmed, Md. Rifayet Azam Talukder, and Md. Mohsin Uddin, English to Bangla Machine Translation Using Recurrent Neural Network, International Journal of Future Computer and Communication, June 2020.
[7] LingoHub Academy. [Online]. Available:https://lingohub.com/academy/glossary/machine-translation
[8] LingoHub Academy. [Online]. Available:https://lingohub.com/academy/glossary/cat-vs-machine-translation
[9] Wikipedia. [Online]. Available:https://en.wikipedia.org/wiki/Machine_translation
[10] Getblend. [Online]. Available:https://www.getblend.com/blog/ai-translation/
[11] http://codesign.blog/wp-content/uploads/2018/05/a443f-1d-if6wcvycwfdlkghpjvkw.png



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY