Heaven’s Light is Our Guide

DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING

Rajshahi University of Engineering & Technology, Bangladesh

Comparative Analysis of Machine Translation Models for Bangla Language: Evaluating Performance and Potential

Author

Md. Tariqul Islam & Md. Khatami

Roll No. 1803165 & 1803180

Department of Computer Science & Engineering

Rajshahi University of Engineering & Technology

Supervised by

Dr. Md. Nazrul Islam Mondal

Professor

Department of Computer Science & Engineering

Rajshahi University of Engineering & Technology

ACKNOWLEDGEMENT

First of all, I would like to thank almighty Allah, for his grace and blessings as well as for providing me with the diligence and enthusiasm along the way to accomplishing my thesis work.

I also want to express my sincere gratitude, admiration and heartfelt appreciation to my supervisor **Dr. Md. Nazrul Islam Mondal**, Professor, Department of Computer Science & Engineering, Rajshahi University of Engineering & Technology, Rajshahi. Throughout the year, he has not only provided me with the technical instructions and documentation to complete the work, but he has also continuously encouraged me, offered me advise, assisted me, and cooperated sympathetically whenever he deemed necessary. His constant support was the most successful tool that helped me to achieve my result. Whenever I was stuck in any complex problems or situation he was there for me at any time of the day. Without his sincere care, this work not has been materialized in the final form that it is now at the present.

I am also grateful to respected **Prof. Dr. Md. Ali Hossain**, Head of the Department of Computer Science & Engineering and all the respective teachers of Department of Computer Science & Engineering, Rajshahi University of Engineering & Technology, Rajshahi for their valuable suggestions and inspirations from time to time.

Finally, I would like to convey my thanks to my parents, friends, and well-wishers for their true motivations and many helpful aids throughout this work.

April 24, 2024 Md. Tariqul Islam & Md. Khatami

RUET, Rajshahi

Heaven's Light is Our Guide

DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING

Rajshahi University of Engineering & Technology, Bangladesh

CERTIFICATE

This is to certify that this thesis report entitled ***"Comparative Analysis of Machine Translation Models for Bangla Language: Evaluating Performance and Potential "*** *submitted by* ***Md. Tariqul Islam & Md. Khatami, Roll:1803165 & 1803180*** *in fulfillment of the requirement for the award of the degree of Bachelor of Science in Department of Computer Science & Engineering of Rajshahi University of Engineering & Technology, Bangladesh is a record of the candidate own work carried out by them under my supervision. This thesis has not been submitted for the award of any other degree.*

Supervisor

____

Dr. Md. Nazrul Islam Mondal
Professor
Department of Computer Science & Engineering
Rajshahi University of Engineering & Technology
Rajshahi-6204

External Examiner

____

-
-
Department of Computer Science & Engineering
Rajshahi University of Engineering & Technology
Rajshahi-6204

ABSTRACT

The fifth most common language in the world is Bangla. Worldwide, 237 million people speak Bangla. When attempting to connect with members of other groups, this enormous population encounters many challenges. One crucial difficulty that divides them is language. An effective machine translation model is quite useful in removing this barrier. finding a translation model that works well. Four different translation models—Encoder-Decoder, Transformer, GRU, and Marian-based models—are compared in this work. We investigate the performance evaluation of these models using BLEU score and accuracy measures, driven by the necessity to evaluate the effectiveness of various architectures in capturing translation nuances. The comparison of four translation models, namely Encoder-Decoder, Transformer, GRU, and Marian-based models, sheds light on their respective performance metrics. The accuracy rates for the first three models stand at 65%, 82%, and 90%, respectively. Correspondingly, their BLEU scores are recorded as 9.61, 11.82, 11.63, and 12.53 respectively. Our comparison reveals that the Transformer and GRU architectures exhibit remarkable proficiency in linguistic processing, outperforming the traditional Encoder-Decoder model and even surpassing the Marian-based model in both accuracy and BLEU score. These results highlight the revolutionary potential of advanced machine translation models such as Transformers and GRU. Our work contributes to a better understanding of these models' capabilities and guides future advancements in machine translation research and applications by offering a thorough examination of their performance.

CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter 1

Introduction

1.1 Introduction

Today our world is called a global village[6], because of the development of communication technology. Communication between different communities increases for this development. But, one of the challenging parts of communication is different people from different communities use different languages. The demand for language translators has increased dramatically due to this growth of communication between different people who use different languages. However, the number of available translators is insufficient to do all the required translating[7]. Therefore, many linguists and technologists want to improve machine translation[7]. However, machine translation originates in the 1960s[8], machine translation doesn't perform so well due to the lack of computation capability. But twenty years ago the development of the computation capability of computers started to increase exponentially. As the computation power increases the performance of machine translation increases which can perform like a human translator. Nowadays machine translation is a crucial part of translation processes in many professional environments, such as many online meetings. Due to this increasing demand for machine translation[7], Many machine translation models have developed. Some of them are the Encoder-decoder-based model, transformer-based model, and LLM. Between the last 20 years, machine translation has become so popular that, between 2010 and 2017, one-third of Internet users regularly utilized one[9].

1.2 Machine Translation Impact on People

Nowadays Machine translation decreases the barrier of language between people from different communities. Any person can freely communicate with any person around the world. language can't create any barrier anymore. Of certainly! Grand View Research has released research indicating that the machine translation market is expanding significantly on a worldwide scale. In 2022, the machine translation market was valued at USD 978.2 million[7]. According to this research machine translation is not only used for personal use but also for Military and defense, automotive, and healthcare industries. people use this Machine translation so often that the Google translator translates thirty trillion sentences annually across over 100 languages[10]. As Machine translation has many benefits it has some disadvantages also. These days, this technology is employed not only in private settings but also in high-risk ones like courts and hospitals to get over language difficulties. In a recent instance, consent to conduct a police search was acquired via Google Translate, raising questions about the legality of the consent. As a result, the evidence was thrown in court[10]. An analysis of potential Machine translation mistakes in a medical context showed that the phrase ”your child is fitting” may have, in one instance, been mistranslated to Swahili as ”your child is dead”[10]. So, Machine translation has a very good effect as well as some disadvantages.

1.3 Translation Model

A large number of engineers and linguists are working to improve Translation Models due to the increasing need for machine translation, which interprets like a person doing the translating. Warren Weaver's 1949 Memorandum on Translation introduces the new field of ”machine translation”[8]. Yehosha Bar-Hillel, the first researcher in the field, started working at MIT in 1951[8]. Following (1951), a Georgetown University research team demonstrated their method to the public in 1954[8]. The United States is promoting machine translation as a way to monitor Russian[8]. It's also among the earliest computer applications that aren't numerical[8]. Russia and Japan began offering machine translation research programs in 1955, and London hosted the first machine translation conference in 1956[8]. As the National Academy of Sciences established a committee (ALPAC) to research machine translation (1964) and the Association for Machine Translation and Computational Linguistics was founded in the United States (1962), researchers kept entering the field[8]. In this development process, many kinds of translation models were introduced:

- Rule-based Machine Translation (RBMT):
- Statistical Machine Translation (SMT)
- Neural Machine Translation (NMT)

Early in the development process, the Rule-based Machine Translation model is introduced. A classic method of machine translation that uses dictionaries and linguistic rules to translate text between languages. Comprehensive language knowledge, including lexicons, dictionaries, syntactic structures, and grammar rules for both the source and target languages, is usually required for Rule-based Machine Translation systems. Later Statistical Machine Translation was introduced. An automatic translation method that uses statistical models is called statistical machine translation. It learns patterns and connections between words and phrases in several languages by examining vast volumes of translated text data. It then creates translations for new sentences using this knowledge. The last technology created was Neural Machine Translation. Due to its significant improvements in translation quality, it has quickly replaced statistical machine translation.

1.4 Overview

Due to the increasing demand for machine translation it's necessary to build a good translation model or find the best developed one among the all developed models that is best fit for this task. In this study, two models are utilized and compared to find the optimized translation model for Bangla to English translation.

1.5 Motivation

Bangla is the fifth largest language in the world. Around 237 million people in the world speak in Bangla[11]. This huge number of people face many problems when they try to communicate with people from other communities. Language creates a vital barrier among them. To remove this barrier a good translator is needed. But a human translator is not always. Also, the number of translators is not enough for this huge number of people. That's why a good machine translation model is necessary. Which is available to the user all the time and does the translation like a human translator. It will make communication easier for the Bangladeshi people.

1.6 Thesis Objectives

As mentioned earlier Bangla is the fifth largest language in the world[11]. Around 237 million people in the world speak in Bangla. An optimized and effective English-to-Bangla translation model is very important to open a window for these 237 million to communicate with other people saying different languages. The main Objectives of this work are as follows:

- Evaluate the effectiveness of four models in accurately translating English text to Bangla, considering factors such as translation quality, fluency, and adequacy.
- Investigate the computational efficiency of the two approaches, analyzing their training and inference times, memory requirements, and scalability.

1.7 Challenges of Translation Model

The challenges of conducting a comparative study between Encoder-Decoder and T5 models for Bangla to English language translation are multifaceted and include::

- Linguistic Complexity: The Bangla language features rich morphology, agglutination, and syntactic structures that differ significantly from English. Handling these linguistic complexities poses a challenge for both models in accurately capturing meaning and preserving fluency during translation.
- Resource Constraints: Limited access to computational resources, including hardware and software infrastructure, may impede the scalability and reproducibility of experiments, particularly for training large-scale models like T5.
- Evaluation Metrics: Selecting appropriate evaluation metrics to assess translation quality, fluency, and adequacy is crucial.
- Model Training and Tuning: Training and fine-tuning these models require significant computational resources and expertise.

The suggested translator model overcame these challenges and produced an effective model. The upcoming chapters address it and provide details of how it was resolved.

1.8 Thesis Organization

The report is divided into six chapters, including this chapter: ***Introduction***, where every relevant topic that is necessary to comprehend the study effort is included. The remainder of the works' overview is arranged as follows:

Chapter 2

Topic - Literature Review

Several machine translation works are discussed in this chapter along with their Result, advantages, and disadvantages.

Chapter 3

Topic - Background Study

This chapter covers the use of machine translation for user pleasure, provides specific information on machine translation, and explores the benefits of using machine translation in commercial contexts.

Chapter 4

Topic - Methodology & Implementation

The dataset and suggested methods are covered in this chapter. It also provides a thorough explanation of the suggested architecture, model training, and data pre-processing.

Chapter 5

Topic - Result & Performance Analysis

This chapter compares the proposed architecture with comparable works and examines the experimental results and performance of the design. This chapter also includes a description of the measures that were used to assess our models.

Chapter 6

Topic - Conclusion

The research has come to an end with this chapter. The research findings are summarized in this article. Additionally, an attempt was made to draw attention to the work's shortcomings and possible areas for future development.

1.9 Conclusion

An outline of the upcoming study and a preview of the planned work was given in this chapter. Insights into the motivation, objectives, and research problems that will be further investigated in later chapters were covered in the conversation.

Chapter 2

Literature Review

2.1 Introduction

In recent years, significant progress has been made in machine translation (MT) due to advancements in deep learning and neural network techniques. This review paper seeks to investigate and contrast different MT models, emphasizing their advantages, limitations, and effectiveness in translating various language pairs. Below are some influential papers that have inspired and influenced my research.

2.2 Related Works

Isidora Stevanović and Luka Radičević [12] The authors provided a concise comparative analysis of the two most widely used translation models in their research. Their objective was to pinpoint the key advantages and disadvantages of statistical and neural machine translation, offering a fresh perspective on the field. They explored four primary approaches to machine translation: direct, rule-based, corpus-based, and knowledge-based. Subsequently, they conducted a study on statistical and neural machine translation systems, revealing approximately 200 errors in SMT compared to 150 errors in neural machine translation. This investigation was based on 196 sentences from a well-balanced evaluation set. To round off their theoretical examination of machine translation systems, they evaluated and compared four distinct machine translators: GUAT, Amebis Presis, Microsoft Bing Translator, and Google Translator. Ultimately, they concluded that neural machine translation surpasses its statistical predecessor in sophistication. Depending on the specific translation needs, any of the systems evaluated could be a suitable choice.

Research by Maria Stasimioti Vilelmini, Sosoni Despoina, Mouratidis, and Katia Kermanidis [13] compared three machine translation systems: a customized neural machine translation system, a generic statistical machine translation system, and a tailored-NMT system. The study focuses on the English-to-Greek language pair. They showed that the tailored NMT system surpassed both the statistical machine translation and the neural machine translation systems. The total number of errors in SMT is 699, 550 in NMT, and 414 in tailored NMT systems. They also show that SMT has 40% grammatical errors, NMT has 31%, and tailored NMT has 30% grammatical errors. Regarding the variations between the SMT and NMT outputs, their research demonstrates that the NMT systems provide translations of a better quality, which supports the conclusions of earlier studies on a range of language pairings. Specifically, the study shows that both the tailored NMT and the generic NMT outputs score higher when it comes to human and automatic assessment criteria; the tailored NMT output does even better than the generic NMT output. Furthermore, the tailored NMT output was evaluated better for both adequacy and fluency, making it the top output.

Mentioning the importance of quality estimation of machine translation Sugyeong Eo, Chanjun Park, Hyeonseok Moon, Jaehyung Seo, and Heuiseok Lim [14] published their work, where they undertake comparison tests and analyses using cross-lingual language models (XLMs), multilingual BERT, and XLM-RoBERTa, and show QE utilizing the as-yet-unutilized multilingual BART model. The experiment results allowed us to demonstrate that the XLM-TLM model outperformed the other models on both sub-tasks and that pre-training's induction of language alignment learning had a beneficial effect.

In The seminal paper "Attention Is All You Need," [2] the authors introduced the Transformer, a novel neural network architecture that eschews the recurrent and convolutional layers commonly found in sequence transduction models in favor of an attention mechanism. This shift allows for increased parallelization, reducing training time significantly1. The Transformer achieves state-of-the-art performance on English-to-German and English-to-French translation tasks, surpassing previous models and ensembles2. The paper also explores the model's application to English constituency parsing, demonstrating its versatility and potential for generalization to other tasks.

Using morphosyntactic divergence as a lens, Jiaming Luo, Colin Cherry, and George Foster [15] did a fine-grained comparative study of machine translations (MTs) against human translations (HTs). They determine that MT is more conservative than HT, with less morphosyntactic variation, more convergent patterns, and more one-to-one alignments, through assessments on both the aggregate level and the individual pattern level. We also see that for the less common source patterns, MT tends to be less comparable to HT.

Özdemir, Özgüra Akın, Emre Salihb Velioğlu, Rızac, and Dalyan, Tuğbaa[16] did a comparison analysis of Turkish language neural machine translation models. The studies were carried out on two distinct architectures. First, a translation job is run using both the Sequence to Sequence (Seq2Seq) architecture and a version that makes use of an attention mechanism. Second, a thorough comparison is carried out using an architecture called Transformer, which is entirely based on the self-attention mechanism. The experiments are carried out on two distinct datasets: the WMT18 News dataset, which is provided by The Third Conference on Machine Translation (WMT) for shared tasks on various aspects of machine translation, and the TED Talks dataset, which is one of the well-known benchmark datasets for NMT, particularly among morphologically rich languages like Turkish. The results of the examination of translations from Turkish to English show that the Transformer model, which combines BPE and Gumbel Softmax, scored 22.4 BLEU on TED Talks and 38.7 BLUE on the WMT18 News dataset. The actual findings corroborate the claim that the Gumbel Softmax distribution enhances translation quality for both systems.

Adam Roberts, Noam Shazeer, and Colin Raffel A team of Google researchers led by Katherine Lee, Sharan Narang, Michael Matena, Yanqi Zhou, Wei Li, and Peter J. Liu conducted ground-breaking research[17]. They begin by introducing the T5 model, also referred to as the text-to-text Transfer Transformer model. Their work opens many roads of NLP tasks such as translation, Sentence completion, Paraphrasing, Sentiment analysis, and so on. Their introduced T5 model is pre-trained on large amounts of data. Which later was fine-tuned on an interesting downstream assignment.

2.3 Conclusion

The majority of previous literature reviews have centered on comparing different translation models, with many studies utilizing various machine translation models and technologies to compare. To continue this research on the Bangla language we use two translation models one of them is an encoder-decoder-based model and the second one is the based translation model.

Chapter 3

Background Study

3.1 Introduction

The use of machine learning in translation technologies has transformed the way that language barriers are overcome and allowed for easy interaction between speakers of different languages. Because of this revolutionary combination of artificial intelligence and computational linguistics, complex algorithms that can improve translation accuracy by learning from large data sets have been developed. Machine learning models are able to translate text with ever-increasing precision, frequently in real-time, by analyzing patterns and subtleties in language. This development has not only increased the effectiveness of translation procedures but also increased worldwide information accessibility, fostering cross-border cooperation and understanding. The promise of machine learning's continued advancement lies in its potential to further bridge linguistic, cultural, and social divides through its incorporation into translation systems.

3.2 Machine Translation

Machine translation (MT) is a fascinating field that sits at the intersection of linguistics, computer science, and artificial intelligence. It involves the use of software to translate text or speech from one language to another[18]. At its core, MT aims to break down language barriers and facilitate communication across different languages without human intervention. The divisions and categories of MT can be broadly classified into three types[8]:

- Rule-based Machine Translation (RBMT)

- Statistical Machine Translation (SMT)
- Neural Machine Translation (NMT)

RBMT converts text by applying linguistic concepts. These guidelines are developed by carefully analyzing the grammar, syntax, and semantics of the source and destination languages[19]. RBMT frameworks need large lexicons and collections of grammatical rules, and they are usually customized for particular language pairings. They fall into three categories: interlingual, transfer-based, and direct translation. While transfer-based systems use intermediate representations to extract the meaning of the source text before translating it, direct translation systems simply replace words with equivalents. On the other hand, interlingual systems condense the source text into a representation that is independent of language and then use that representation to produce the target text.
Machine Translation (SMT), on the other hand, makes use of statistical models built by analyzing large amounts of multilingual text[20]. Taking into account the context in which words or phrases appear, these models make predictions about the chance that they are accurate translations. SMT is categorized into several categories, including syntax-based, phrase-based, and hierarchical phrase-based translation. While hierarchical phrase-based SMT takes into account layered phrases to capture more complicated language patterns, phrase-based SMT focuses largely on translating word sequences. On the other hand, syntactic parses of sentences are employed by syntax-based SMT to improve understanding of the grammatical structure and provide more accurate translations.

The emergence of neural machine translation (NMT), a subset of SMT, has led to notable advances in the field. Using deep learning techniques, NMT creates a coherent, complex neural network that contains the whole translation process. Contextual knowledge, translation accuracy, and fluidity have all improved as a result of this strategy.

Hybrid machine translation, which includes components of both RBMT and SMT to capitalize on each technique's advantages, is an additional category to the ones already stated. Moreover, machine translation systems may be customized for certain fields like law or medicine, where accurate translations and specialist vocabulary are essential. A crucial part of MT system evaluation is its assessment, which includes both human and automatic measures (e.g., BLEU and METEOR) to measure translation accuracy. The development of computer-aided translation (CAT) technologies, which support human translators by giving them access to translation memory, glossaries, and other resources, is another result of MT advancements.

The advancement of machine translation is being driven by ongoing research, to lower biases, increase accuracy, and address challenges associated with low-resource languages. The potential for seamless, instantaneous translation is becoming closer as technology develops, offering a promising future in which language boundaries will no longer prevent people from communicating globally.

3.3 Foundations of Neural Machine Translation

A significant change in the method for translating text between languages is represented by neural machine translation (NMT)[21]. Neural machine translation (NMT) uses sophisticated deep learning algorithms instead of rule-based and statistical methods to improve translation accuracy. The core of NMT is its ability to evaluate large amounts of multilingual text data, which allows the system to predict the probability that a set of words in one language would match a set of words in another. By considering the context and overall meaning of phrases instead of just substituting words for words, this method makes it easier to produce translations that are more genuine and logical.

An encoder-decoder framework is commonly used in the construction of NMT systems[22]. The encoder processes the input sentence and condenses the information into a context vector that captures the essential meaning of the phrase. The translated sentence is then produced in the target language by the decoder using this vector. Unlike previous methods that required several different systems that were each customized separately, this single neural network simplifies the translation process.

Compared to standard machine translation techniques, Neural Machine Translation (NMT) has the benefit of being able to handle the subtleties and complexities of human language without the requirement for pre-established rules or linguistic experience. By being exposed to a variety of instances, NMT systems learn how to translate, making them more adaptable to varied languages and dialects. NMT has advantages, but it also has drawbacks, such as how training data amount and quality affect translation accuracy and how much processing power is needed for training and running the system. These issues might prevent NMT from being widely used. Furthermore, because deep neural networks are complicated, there are issues with the transparency of NMT models.

Fundamentally, deep learning and artificial intelligence advancements provide the foundation of neural machine translation. By introducing a more accurate and efficient method of language translation, neural machine translation (NMT) has revolutionized the field of machine translation. Improvements in NMT systems are expected with continued research[2]. The development of NMT shows a persistent attempt to get over linguistic barriers and enhance global communication by making it more useful and accessible.

3.3.1 Rule Based Machine Translation (RBMT)

In computational linguistics, Rule-Based Machine Translation (RBMT) is a new technique where computers translate text according to an extensive collection of linguistic rules[19]. The grammatical, syntactic, and semantic characteristics of the source and destination languages serve as the foundation for the laws. When translating text from one language to another, rule-based machine translation (RBMT) systems—which were created in the early stages of machine translation technology—mainly rely on dictionaries and grammatical rules. The three primary steps of this process are typically: interpreting the original text, translating pertinent linguistic elements into the target language, and creating the final target text.

Rule-Based Machine Translation (RBMT) analyzes the source text by processing it to identify its grammatical structure. The approach entails utilizing morphological principles to ascertain the roles of distinct words and phrases in sentences. The detected structures are then mapped to corresponding structures in the destination language using a predetermined set of rules during the transfer step. These guidelines take into account the differences in syntax and grammar between the two languages to make sure the translation accurately captures the original meaning. The translated text is then rebuilt in the target language while following its syntactic and grammatical rules throughout the generation phase.

Accuracy and consistency in translation are two of RBMT's key benefits[19], as RBMT systems operate by adhering to clear rules rather than machine learning models or statistical probabilities. When working with technical or specialist materials that require precision and adherence to certain language, this method works well. However, there are several restrictions with RBMT systems. They may have trouble comprehending the nuances and context-dependent parts of language that depart from conventional rules, and they need a great deal of manual labor to build and maintain the rule sets.

3.3.2 Statistical Machine Translation (SMT)

Statistical Machine Translation (SMT) is a system that uses statistical models obtained from the analysis of large bilingual text corpora[20], which represents a significant shift from rule-based methods requiring extensive linguistic knowledge. The core concept of statistical machine translation (SMT) is the estimation of the likelihood that a given text string in one language is an accurate translation of a text string in another language through the use of probability distributions. Researchers at IBM's Thomas J. Watson Research Center first developed this approach in the late 1980s and early 1990s, expanding on ideas first proposed by Warren Weaver in 1949.

Sentence synchronization within a bilingual corpus serves as the training data for the statistical models and is a component of the SMT process. These models often operate based on the Bayes Theorem, in which the language model assesses the likelihood of seeing a certain string in the target language, and the translation model determines the likelihood that a string in the source language is the translation of a string in the target language. Then, the search for the best translation is conducted by selecting the string that has the highest probability. This task is completed by a machine translation decoder that makes use of several heuristics to efficiently traverse the large search space.

3.3.3 Neural Machine Translation (NMT)

With the introduction of neural machine translation (NMT), machine translation technology has undergone a dramatic transition from statistical and rule-based approaches to models that make use of artificial neural networks and deep learning[2]. The capacity of neural machine translation, or NMT, to improve translation quality—especially for languages with copious linguistic resources—has led to its rapid rise in popularity. NMT works by estimating the likelihood of a string of words, frequently by modeling entire sentences in a single model. The basic architecture of NMT systems typically consists of an encoder-decoder framework, in which the encoder parses the input sentence and creates a vector representation of it, and the decoder creates the translated sentence word by word by utilizing the source sentence's representation as well as the words that have been predicted.

Although NMT has a long history dating back to early research in the late 1980s[23], its full potential wasn't realized until the introduction of more advanced neural network models, such as Recurrent Neural Networks (RNNs) and Convolutional Neural Networks (CNNs). By increasing translation accuracy and speed, the transformer model—which makes use of self-attention mechanisms—further transformed neural machine translation (NMT). Large multilingual text datasets are used to train these models, which then pick up on patterns and subtleties in the input to translate.

Even with its advancements, Neural Machine Translation (NMT) still has certain challenges. When translating texts that don't match the training data or when working with languages with few linguistic resources, the quality of the translation may suffer. Furthermore, NMT algorithms often produce translations that are quite literal and may not have the same natural flow or idiomatic precision as human translations.

3.4 Encoder-Decoder Based Model

An important concept in the field of neural networks—more especially, natural language processing—is the encoder-decoder mechanism. Its goal is to process data sequences, such as sentences in a translation job, by encoding the input sequence into a fixed-dimensional representation and then decoding this representation to produce the output sequence[24]. 1. The encoder part of the model processes and compresses the incoming data into a context vector. The context vector functions as a simplified depiction of the input sequence and is later employed by the decoder to build the output sequence piecemeal. The context vector and pre-generated elements serve as the basis for each element's generation.

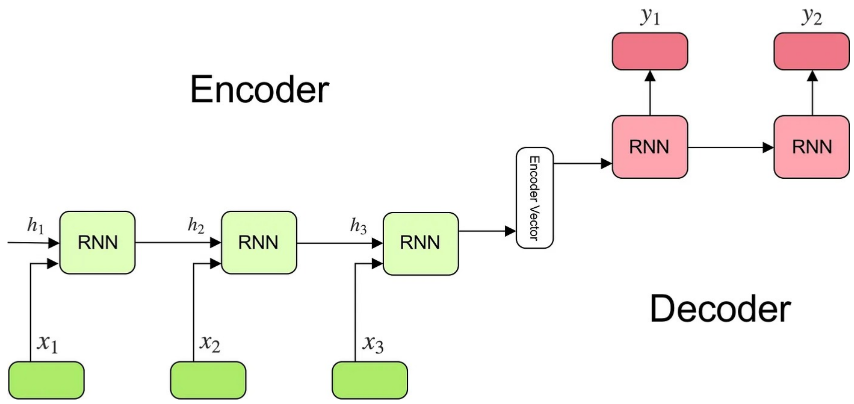

Figure 3.1: Encoder-Decoder Architecture [1].

3.5 Transformar Based model

The Transformer model's different design, which processes data sequences by self-attention mechanisms, has completely transformed the field of neural machine translation (NMT)[2]. The Transformer model achieves a significant improvement in translation speed and efficiency by operating on input data concurrently, unlike previous models that depended on recurrent or convolutional layers. This architectural strategy has improved the performance of NMT systems and proven useful for a variety of additional tasks in the fields of computer vision, natural language processing, and other related fields. Scholarly publications and resources provide in-depth and thorough insights into the complexities of the Transformer model in the context of NMT, as well as its related issues and broader applications.

3.5.1 Attention

As each piece of the output sequence is generated, the model may dynamically focus on various portions of the input sequence thanks to a key development in this mechanism: the application of attention[2]. The attention technique allows the decoder to access the whole input sequence directly, avoiding the need for the context vector, which can be a performance-limiting issue, especially for long sequences. As a consequence, the model will be able to focus more on the relevant portions of the input sequence for every output element, producing outputs that are more accurate and coherent.

The ‘MultiHeadAttention‘ layer in the code uses several heads to do attention calculations, incorporating the attention mechanism. The model is able to capture a fuller representation of

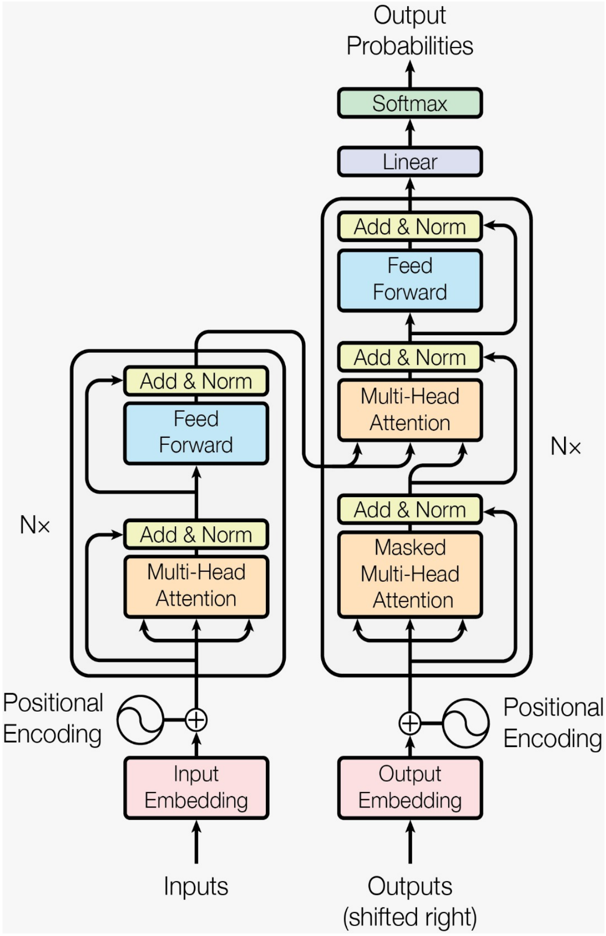

Figure 3.2: Transformer Architecture [2].

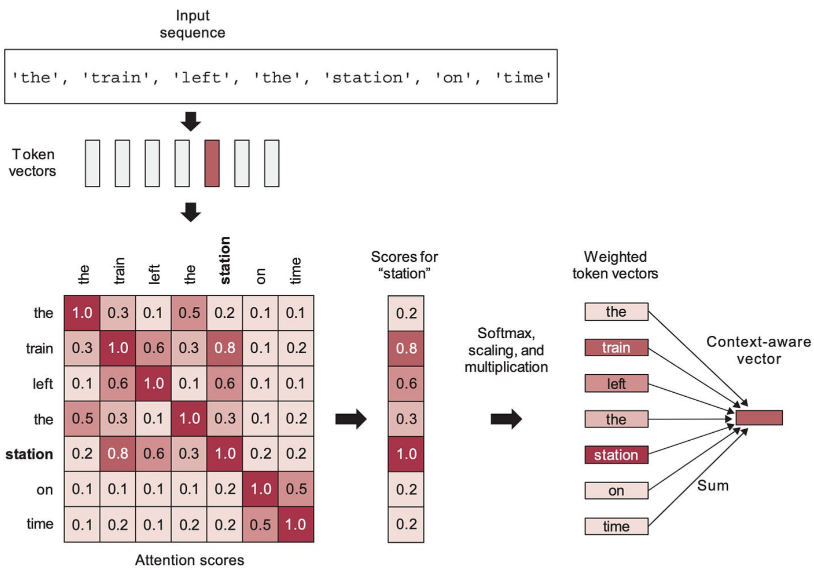

Figure 3.3: Attention Mechanism [3].

the input sequence because each head separately listens to information from various representation subspaces at different points. The attention outputs from each head are then merged and processed. Deep network training is made quicker and more stable by employing the 'Layer Normalization' layer after it to stabilize the learning process.

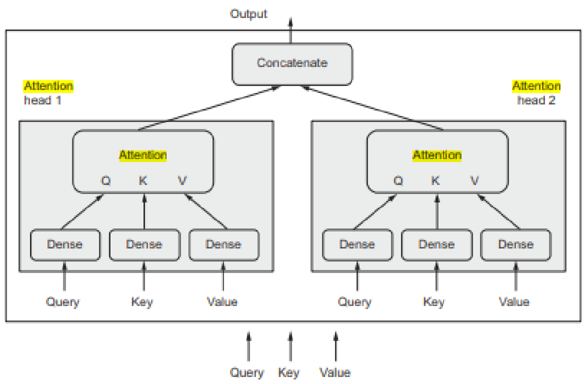

Figure 3.4: The MultiHeadAttention layer[4].

The inclusion of long short-term memory (LSTM) layers in the encoder and decoder is essential for identifying temporal relationships in the data. A particular kind of recurrent neural network that is particularly good at learning long-term dependencies is called an LSTM network, or long short-term memory network. This capacity is especially important for jobs like translating across languages, where the context may span long word sequences. The input sequence and the attention-enhanced representation are processed by the LSTM layers in the encoder to produce the context vector. In contrast, the context vector and the outputs from earlier time steps are used as inputs by the LSTM layers in the decoder to create the output sequence.

Attempts to improve the model's generalization and reduce overfitting are seen in the inclusion of 'Dropout and 'Batch Normalization' layers in the design. In order to achieve a stable learning process and a considerable reduction in the number of training epochs required for deep network training, batch normalization normalizes the inputs to a layer for each mini-batch. By randomly ignoring neurons during training, dropout, on the other hand, is a regularization technique that improves the model's resilience and lessens its dependence on the exact weights of individual neurons.

As the code reveals, the encoder-decoder architecture with attention is a complex design that works well for complex sequence-to-sequence operations. The LSTM layers capture the important temporal relationships, and the attention mechanism increases the model's capacity to focus on relevant portions of the input sequence. Layer normalization and dropout are included to ensure that the model is robust and robust, demonstrating a careful consideration of the challenges involved in training deep neural networks. This architectural structure is proof of the ongoing advancements in the field of artificial intelligence and how it is applied to the understanding and production of human language.

3.6 GRU Model

A significant advance in the field of computational linguistics is the Neural Machine Translation (NMT) model based on Gated Recurrent Units (GRU)[25]. By presenting a unique GRU-gated attention mechanism, it overcomes the drawbacks of previous sequence-to-sequence models. With the help of this technique, the model may develop context vectors that are more sensitive to the partial translation that the decoder generates, which results in translations that are more accurate and discriminative. Because of its design, the GRU is especially well-suited for challenging language translation jobs where it can efficiently handle long-term dependencies. Studies have demonstrated that GRU-based NMT models can perform better than conventional models, particularly for tasks that call for sophisticated syntactic and contextual comprehension.

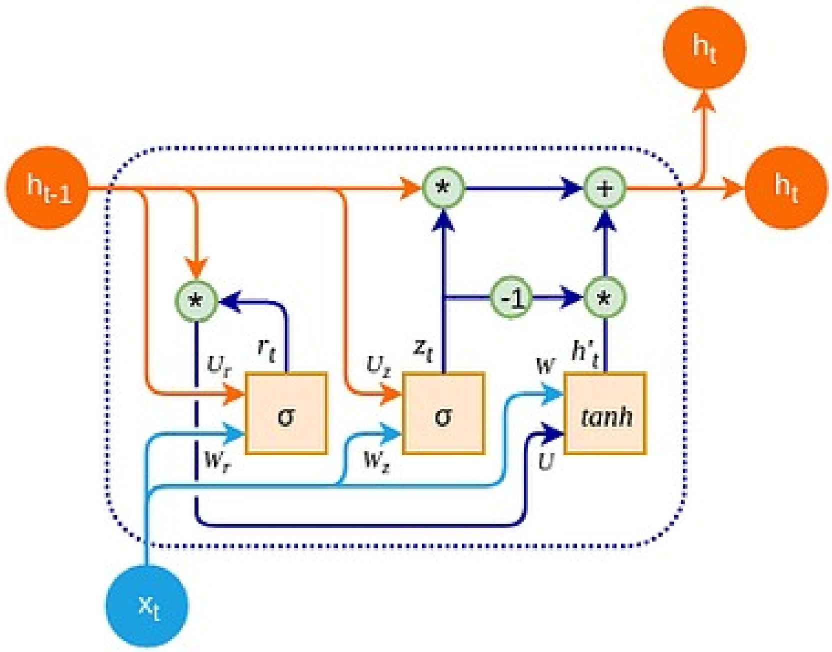

Figure 3.5: GRU Architecture [5].

3.7 Principles of Neural Networks in Translation

With its advanced method for comprehending and interpreting languages, neural networks have completely transformed the area of machine translation[26]. The idea of deep learning, which involves training a model to identify patterns in enormous volumes of data, is central to the concepts of neural networks in translation. By providing the network with instances of text in one language and its translation into another, the training process enables the network to gradually pick up on the subtleties of vocabulary and linguistic structures.

The sequence-to-sequence architecture, which interprets the input text as a series of symbols and produces the translated text as another sequence, is a crucial element of neural machine translation (NMT). This method is different from previous approaches that relied on statistical machine translation, which employed statistical models to translate text based on massive corpora of samples, or rule-based systems, where linguists manually constructed translation rules.

The 2017 release of the Transformer model represents a significant advancement in NMT[2]. It utilizes attention techniques in place of the conventional recurrent layers seen in previous models. Through these techniques, the model can forecast each translated word while giving distinct chunks of the input text different weights. As a result, this method improves translation accuracy and context sensitivity.

Moreover, multi-modal translation assignments are made easier by the fact that neural networks used in translation are not limited to textual information alone. They also handle audio and visual data. One illustration of this is the capability of immediately translating spoken words or text contained inside photographs into another language's text or voice. This adaptability demonstrates neural networks' ability to support a wide range of data types and translation needs.

The extensive use of NMT models for translation services by significant IT businesses is evidence of their efficacy. For example, by using deep learning and a lot of data, Google's Neural Machine Translation system has significantly improved translation quality, producing translations that are more like human language and more natural.

Even with these improvements, problems still exist. In addition to straight translation, NMT must take into account the inherent ambiguity and fluidity of human language, which calls for an awareness of cultural quirks, context, and even inferred meanings. By tackling these issues, NMT will be able to translate not just words but also the meaning and emotion associated with them, making machine translation an even more potent instrument for international communication.

To sum up, the basis of neural networks in translation is their ability to learn from examples and improve over time. Neural networks are a dynamic field of study that is continuously broadening the scope of machine translation via ongoing discovery and improvement.

3.8 Architectures: Transformers and Beyond

The development of neural network designs has been an adventure in creativity and discovery, and the introduction of Transformer models was a major turning point. Because these structures allow models to handle sequential data with unprecedented efficacy, they have revolutionized industries such as computer vision and natural language processing (NLP). The Transformer's primary self-attention mechanism enables it to assess the significance of various input data elements in diverse ways, demonstrating its remarkable capacity to comprehend context and linkages within data[2].

But it continued after the release of the first Transformer model. To increase the Transformer's effectiveness and performance, researchers have persisted in pushing the envelope. To improve Transformer structures and specifically address the computational complexity brought by the model's attention processes, this led to the investigation of Neural Architecture Search (NAS) methodologies. Reducing the amount of computing resources needed without sacrificing the model's performance was the aim.

The suggestion of a novel framework that makes use of NAS to identify the best designs for effective Transformers is one such development. This method has been proven to produce equivalent accuracy to traditional Transformers while greatly increasing computing efficiency on applications such as image classification and machine translation. According to the research, neither effective attention mechanisms nor typical Softmax attention properly balances accuracy and efficiency. However, each has its advantages. More innovation has resulted from this, such as integrating two attention kinds to lessen performance mismatches.

These developments have implications that go beyond computer vision and natural language processing. Because efficient transformers make complex AI models more accessible to those with low computational capabilities, they have the potential to democratize the usage of these models. Additionally, they provide new opportunities for edge computing and real-time applications when processing power is limited.

The investigation of ”Transformers and Beyond” is expected to proceed quickly as we move forward[27]. We are getting closer to developing AI systems that are not just more potent but also more effective and widely available with every new finding. The adventure of architectural invention is far from ended, and what comes next should prove to be just as fascinating as the last one. Research on AI models' quest for the ideal ratio of performance to efficiency is still active and dynamic, with the potential to completely change the way we think about technology.

3.9 The Business Potential of Neural Machine Translation (NMT)

Neural machine translation (NMT) has huge and wide commercial potential[28]. Natural language translation (NMT) is a subfield of artificial intelligence that trains on massive datasets to translate text between languages. The translations produced by NMT systems are often as accurate and fluent as those made by human translators. With the ability to overcome linguistic boundaries that have traditionally hampered international trade, diplomacy, and cooperation, this technology has enormous ramifications for global communication.

NMT provides businesses with an affordable way to localize content on a large scale. By translating websites, product manuals, and marketing materials into other languages, businesses may easily penetrate new markets and broaden their customer base. Additionally, NMT improves multilingual real-time communication, which makes multinational collaborations and negotiations easier.

NMT makes it possible for businesses to offer multilingual customer care, enhancing customer happiness and experience. For software enterprises and internet service providers who serve a worldwide customer base, this is especially advantageous. Furthermore, by incorporating NMT into consumer gadgets like smartphones and smart home appliances, more user-friendly and accessible user interfaces are made possible, which expands the market for consumers.

3.10 Conclusion

In conclusion, the transition from RBMT to NMT demonstrates the significant influence that machine learning has had on translation technology. Language obstacles may be easily solved in the future as machine learning advances and the possibility of developing smooth, real-time translation systems becomes more real. This organized summary lays the groundwork for future research and development in the subject by illuminating the ways in which machine learning has revolutionized translation systems. The ongoing development of these technologies contributes to our global connection by improving communication across linguistic barriers.

Chapter 4

Methodology & Implementation

4.1 Introduction

In recent times, the domain of natural language processing (NLP) has observed notable progressions, primarily driven by the advancement of sophisticated neural network structures. Notably, recurrent neural networks (RNNs), encompassing Gated Recurrent Units (GRUs), and more recently, attention-based models like Transformers, have garnered substantial attention due to their effectiveness in diverse NLP tasks. Concurrently, encoder-decoder architectures have played a pivotal role in tasks such as machine translation and text summarization. The objective of this research is to present a comprehensive comparative analysis of four prominent neural network architectures: GRU, Encoder-decoder, Transformer models, and Marian, within the specific context of language translation. Through a systematic evaluation of their performance across various benchmarks, our aim is to elucidate the strengths and weaknesses inherent in each architecture, thereby shedding light on their applicability and effectiveness in different NLP scenarios.

4.2 Methodology

Our work is mainly concerned with the careful comparison and assessment of different translation models. We have invested significant time and energy into developing three separate models, each precisely designed to capture certain language subtleties, in addition to integrating a popular pre-trained model as a baseline for comparison. In order to optimize performance and facilitate the training process, we have carefully selected a dataset that includes a wide variety of language pairings and contextual settings relevant to our research. The dataset was carefully vectorized and subjected to difficult preprocessing before the training phase began. This thorough preparation made sure the data was free of noise and irregularities and converted it into a format that would work well for training our models. Furthermore, the textual information was carefully encoded into numerical representations using the vectorization process, which helped in speedy calculation and improved the models' capacity to identify significant patterns in the data. We wanted to provide a strong basis for our comparative research by using these strict methods, which would help us understand the advantages and disadvantages of each model we were looking at.

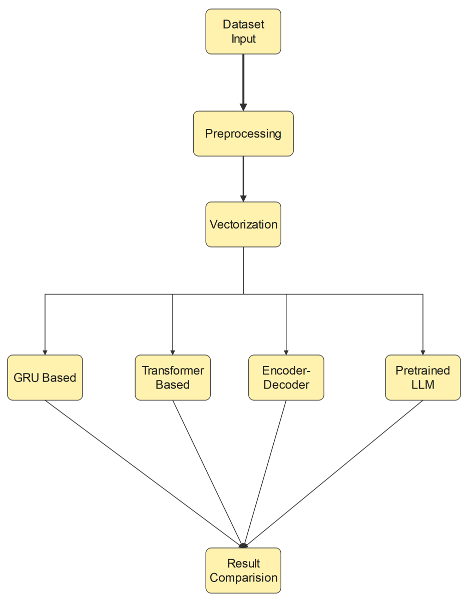

Figure 4.1: Methodology .

4.3 Implementation

During the implementation phase, neural network architectures such as GRU, Encoder-decoder, and Transformer are converted into executable code using deep learning frameworks like TensorFlow, PyTorch, or Keras. The GRU model employs a variant of recurrent neural network (RNN) with gated mechanisms to control the flow of information. The Encoder-decoder structure comprises an encoder and a decoder, which process the input sequence to produce a fixed-length context vector. The GRU cell structure, which includes reset and update gates, is utilized to enable the model to capture long-range dependencies in sequential data while addressing the vanishing gradient issue. The decoder generates the output sequence based on the context vector, often incorporating attention mechanisms to focus on relevant input segments. The Transformer model utilizes self-attention mechanisms to efficiently capture global dependencies across input sequences. The codebase is designed to be modular, well-documented, and easily expandable, which facilitates future experimentation and reproducibility. Each implementation is compared against reference implementations and verified for accuracy.

4.3.1 Dataset Descriptions

For the training purpose a balanced dataset was selected[29]. Prior to utilizing the dataset for machine translation tasks, a series of preprocessing steps were necessary to guarantee the quality and usability of the data. Initially, the text files underwent a cleansing process to eliminate any corrupt or irrelevant data that could have a detrimental effect on the model's performance. This involved removing HTML tags, rectifying encoding errors, and eliminating any non-textual elements. Following this, tokenization was carried out to break down the text into individual words and symbols, which is crucial for the model to grasp the structure of the languages. Subsequently, normalization procedures were implemented for both the English and Bangla texts to standardize the format of numbers, dates, and other entities. Additionally, case normalization was performed by converting all characters to lowercase to prevent any issues related to case sensitivity during the translation process. To further enhance the dataset, a sentence alignment step was undertaken to ensure that each sentence in the source language accurately corresponded to its target language counterpart, thereby preserving the context and meaning across languages. Another crucial measure involved eliminating duplicate and nearly identical sentences to augment the variety within the dataset. This measure serves to mitigate the risk of the model exces-

	ben	eng	eng_length	ben_length
0	সে ঘুরে ল্যাংডনের চোখের দিকে তাকায়।\n	He turned to look Langdon in the eye.\n	8	6
1	সব মানুষের মধ্যে সমতা আনতে হলে কারও না কারও সু...	better o . Guaranteeing that every individual ...	15	13
2	যাই জিজ্ঞেস কর না কেন, সব প্রশ্নের জবাব দিব ডি...	"What do you say I tell you over dinner."\n	9	11
3	ঠিক যখন আমি সভাকে শুভরাত জানিয়ে বিদায় নিতে উদ্...	When I was just about to say good-night to the...	23	18
4	যেতে দিনা!\n	Let me pass."\n	3	2
...	...	...	...	...
2753064	পৃথিবীর সর্ববৃহৎ এ উপসাগরটি পশ্চিমে ভারত ও শ্র...	Bay of Bengal a northern extended arm of the i...	64	30
2753065	প্লাবন ও বন্যা (Inundation and flood) বন্যা বা...	Inundation and flood Some of the devastating n...	29	24
2753066	সামুদ্রিক ও স্বাদুপানির শুঁটকিযোগ্য গুরুত্বপূর...	The important marine and freshwater species th...	37	36
2753067	ক্ষতিকর প্রাণী দমন (pest control) যে কোনো ধরনে...	Pest control any measure deliberately initiate...	19	16
2753068	বাংলাদেশে উপকূলীয় মৎস্যচাষ পুরনো সনাতন ব্যাপার...	Coastal aquaculture is an old traditional prac...	53	38

Figure 4.2: Dataset Snapshot

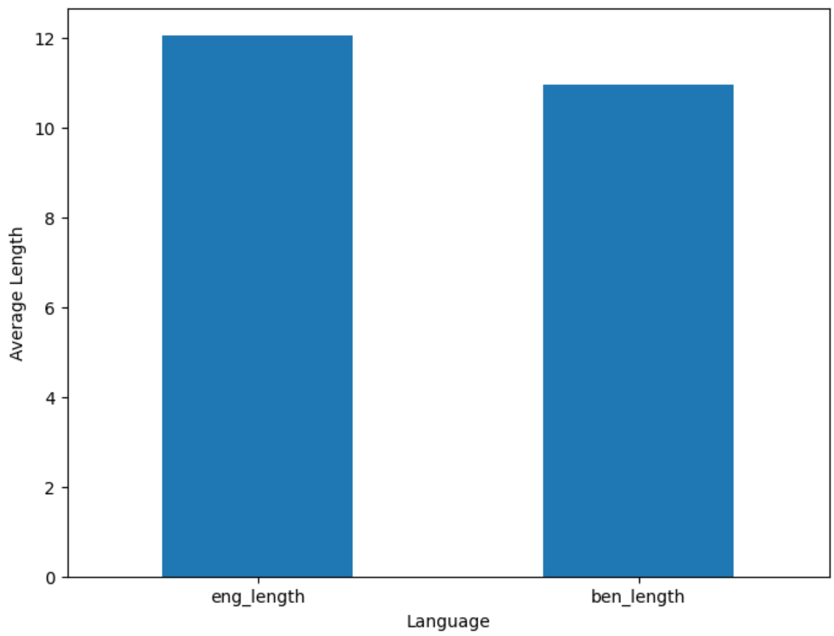

Figure 4.3: Average Text Length

sively fitting repetitive patterns and enhances its capacity to extrapolate from the training data. Furthermore, the dataset underwent length filtering to exclude sentences that were excessively brief or excessively lengthy, as such sentences may either lack substantial content or introduce unnecessary complexity to the training process.

4.3.2 Vectorization & Word Embeddings

Vectorization is a fundamental process in Natural Language Processing (NLP) that transforms text into numerical data, allowing machines to comprehend and analyze language. This technique entails converting words, phrases, or entire documents into numerical vectors, which are essential for executing a wide range of machine-learning tasks. Conventional approaches such as one-hot encoding, count vectorization, and Term Frequency-Inverse Document Frequency (TF-IDF) have certain limitations, such as the inability to capture the semantic significance of words. In response to this challenge, word embeddings were introduced, which map words to a continuous vector space where words with similar meanings are positioned closer together. Models like Word2Vec, GloVe, and FastText utilize neural network architectures to learn word representations from extensive text corpora. Notably, Word2Vec, developed by Google, is recognized for its effectiveness in capturing word associations and syntactic nuances. On the other hand, GloVe combines count-based and predictive methods to generate word vectors that encapsulate global statistics of the corpus. These embeddings have significantly transformed NLP by facilitating the development of more sophisticated and context-aware machine learning models, which are applied in diverse areas such as sentiment analysis and machine translation.

4.3.3 Encoder-Decoder Model

The implementation of the Encoder-Decoder model involves creating a dual-component structure comprising an encoder and a decoder. The encoder processes the input sequence to produce a context vector of fixed length, which captures the semantic information of the input. This process typically includes using recurrent or convolutional layers to encode the input sequence into a compact representation. Subsequently, the decoder utilizes this context vector to generate an output sequence, often incorporating attention mechanisms to concentrate on pertinent sections of the input during decoding. The implementation of the Encoder-Decoder architecture utilizes deep learning frameworks like TensorFlow or PyTorch to define and train the encoder and decoder components independently. Throughout the training process, the encoder and decoder are optimized together to minimize a loss function specific to the task, such as cross-entropy for sequence prediction tasks like machine translation or text summarization. Hyperparameters like learning rate, embedding size, and attention mechanism type are adjusted to enhance the model's performance on validation data.

4.3.4 Transformer Model

The utilization of the Transformer model signifies a departure from conventional recurrent architectures as it relies exclusively on self-attention mechanisms to capture dependencies across input sequences. The model's architecture comprises multiple layers of self-attention and feed-forward neural networks, enabling parallel computation and mitigating the sequential bottleneck associated with recurrent models. To define the Transformer architecture, deep learning frameworks like TensorFlow or PyTorch are employed, incorporating mechanisms such as multi-head attention and positional encoding. Throughout the training process, the model is optimized using standard optimization algorithms like Adam or SGD, with hyperparameters fine-tuned to maximize performance on the designated task. Notably, the Transformer model introduces innovations like scaled dot-product attention and layer normalization, which significantly contribute to its exceptional effectiveness in capturing long-range dependencies and generalizing to unseen data. In summary, the implementation of the Transformer model represents a paradigm shift in sequence modeling, offering unparalleled performance and scalability for a diverse range of NLP tasks.

4.3.5 GRU Model

The implementation of the GRU model involves the creation of a recurrent neural network (RNN) variant that incorporates gated mechanisms to tackle the issue of vanishing gradients commonly observed in traditional RNNs. This process includes designing a network structure consisting of one or more GRU cell layers. Each GRU cell is equipped with reset and update gates, which play a crucial role in controlling the information flow within the network. Typically, deep learning frameworks like TensorFlow or PyTorch are utilized to define the architecture of the GRU cell and integrate it into the overall neural network design. During the training phase, backpropagation through time (BPTT) is utilized to adjust the model parameters and minimize the selected loss function. Hyperparameters such as learning rate, batch size, and dropout rate are fine-tuned to enhance the model's performance and prevent overfitting. The resulting GRU model demonstrates the ability to capture temporal dependencies in sequential data, making it particularly suitable for tasks such as sequence prediction, language modeling, and time series forecasting.

4.3.6 Training

Neural network architectures are trained using optimization algorithms such as stochastic gradient descent (SGD), Adam, or RMSProp, along with carefully selected learning rates, weight decay, and other hyperparameters. To prevent overfitting and address the issue of exploding gradients, regularization techniques like dropout, layer normalization, or early stopping are employed. Hyperparameter tuning is carried out using methods like grid search, random search, or Bayesian optimization to optimize performance metrics on validation datasets while avoiding overfitting to the training data. The training process takes place on high-performance computing infrastructure, making use of GPUs or TPUs to accelerate computation and expedite convergence. Gradient clipping is also utilized as a technique to mitigate the problem of exploding gradients, ensuring stable training dynamics. Model checkpoints are saved at regular intervals to allow for resumption in case of interruptions. Throughout the training process, key metrics such as training loss, validation loss, and performance on held-out test data are continuously monitored. Training is stopped either when performance plateaus or when specific stopping criteria are met, ensuring optimal utilization of resources and preventing overfitting.

4.3.7 Optimization

Stochastic gradient descent (SGD) or its derivatives, such as Adam or RMSprop, are employed for the purpose of optimizing the model parameters. In order to mitigate the issue of overfitting, regularization techniques like dropout are implemented, while learning rate schedules are utilized to dynamically adjust the learning rate during the training process, thereby enhancing the convergence of the model.

4.4 Conclusion

The mentioned code establishes a connection between Bengali and English languages by utilizing neural machine translation, effectively reducing language barriers. By incorporating sophisticated methods like GloVe word embeddings and LSTM networks, it serves as evidence of the impressive capabilities of contemporary computational linguistics. The potential uses of this model are extensive, spanning from enabling instantaneous communication to unlocking a vast amount of knowledge that was previously inaccessible due to linguistic limitations. As the model undergoes training, it not only guarantees translation but also fosters cultural connections, positioning itself as a symbol of technological progress in the field of language translation.

Chapter 5

Result & Performance Analysis

5.1 Introduction

The differences between English and Bangla offer a complex tapestry of opportunities and problems in language and translation studies. The need for precise and effective translation models grows as globalization speeds up the flow of ideas and information across linguistic boundaries. Within this framework, the thesis explores the subtleties of several English-to-Bangla translation models intending to conduct a thorough comparative study that clarifies their advantages, disadvantages, and general effectiveness. This chapter thoroughly reviews the study, the methodology used, and the reasoning behind the approaches used. It acts as a doorway to the subsequent examination of data and analysis. By combining theoretical frameworks, empirical studies, and computational approaches, this research aims to decipher the nuances present in English to Bangla translation problems. Through a critical examination of several translation models—from neural machine translation architectures to rule-based systems—this study aims to reveal the subtleties involved in both language transfer and cultural adaptation. Furthermore, by comparing these models to standards for precision, fluidity, and cultural authenticity, this research hopes to offer insightful information about how machine translation technologies are developing. As we set out on this academic adventure, we must recognize the wider ramifications of this research project, both in terms of developing the theoretical discourse in translation studies. As a result, this introduction lays the groundwork for a thorough investigation of English-to-Bangla translation models and a thorough analysis that aims to significantly contribute to academics and industry.

5.2 Evaluation Metrics

It's important to take into consideration several measures that reflect various facets of translation quality when assessing English to Bangla translation models. The following list of assessment measures is typical in the field

5.2.1 BLEU (Bilingual Evaluation Understudy)

A popular measurement for comparing candidate translations to one or more reference translations is called BLEU. It compares n-grams, or continuous sequences of n elements, between the reference translations and the candidate translations to function. In general, the BLEU (Bilingual Evaluation Understudy) score goes from 0 to 100; a number nearer 100 denotes a better degree of similarity between the reference texts and the applicant translation. It's crucial to remember, as well, that getting a BLEU score of 100 is extremely uncommon and sometimes impractical. A value around 20 is considered a good score. Machine translation systems often use BLEU values that span from 0 to 30 or 40, where higher scores correspond to higher translation quality. However, to have a thorough picture of translation performance, BLEU ratings should always be read in conjunction with other evaluation metrics and qualitative assessments. Furthermore, BLEU ratings can change based on several variables, including the text's difficulty, the language pair, and the accessibility of excellent reference translations. The formula used for computation is:

$$P_n = \frac{\sum_{c \varepsilon candidates} \sum_{n-gram \varepsilon c} count_{clip}(n-gram)}{\sum_{c' \varepsilon candidates} \sum_{n-gram \varepsilon c'} count_{clip}(n-gram')} \tag{5.1}$$

where candidates denote the candidate translation, c denotes each phrase in the candidate translation, and P_n denote the matching degree of the nth order, which is typically the fourth order, Candidates denotes the translation of the candidate, and c denotes each sentence in the candidate translation. Lastly, the BLEU computation technique is:

$$BLEU = BP \bullet exp(\sum_{N}^{n=1} logP_n) \tag{5.2}$$

$$BP = \begin{cases} 1, c > r \\ e^{1-\frac{r}{c}}, c \leq r \end{cases} \tag{5.3}$$

where w_n is the weight coefficient, r, and c are the lengths of the reference and candidate translations, and BP is the length penalty factor.

5.2.2 Accuracy

One of the most important metrics for assessing how well machine learning algorithms work is accuracy. It calculates the ratio of the algorithm's accurate predictions to all of the predictions made. When assessing the model's performance using the available dataset, accuracy is helpful. To evaluate the model's capacity for generalization and prevent overfitting of the training set, methods such as hold-out testing and cross-validation are essential.

$$Accuracy = \frac{CorrectPredictions}{TotalPredictions} \tag{5.4}$$

In summary, these metrics are used to quantify the performance of classification models by comparing their predictions with the actual labels of the data points. By understanding these metrics, we can now proceed to explain each evaluation metric and how they contribute to assessing the model's effectiveness.

5.3 Model Performance

The encoder-decoder, Transformers, GRU-based, and Marian-based pre-trained models were compared in this comprehensive study. These models underwent rigorous training using the described dataset, ensuring a sufficient understanding of the data's complexities and nuances. To accurately facilitate the summary and abstraction of textual information, each model was methodically adjusted and improved to extract meaningful translation from the input sequences. The performance of each model is discussed below.

5.3.1 Encoder-Decode Based Model

Our first evaluation model is the encoder-decoder model. which shows pretty good accuracy at the time of training which is shown in Figure 5.1. The accuracy of the model on the training dataset shows a steady increase, indicating that the model is learning and improving its predictions over time. The model's accuracy on a separate validation dataset increases as well but begins to plateau around epoch 40. Figure 5.2 shows the model loss at the time of training. Training Loss Decreases sharply, showing the model is learning from the training data. Validation Loss Decreases more gradually, indicating the model's performance on unseen data.

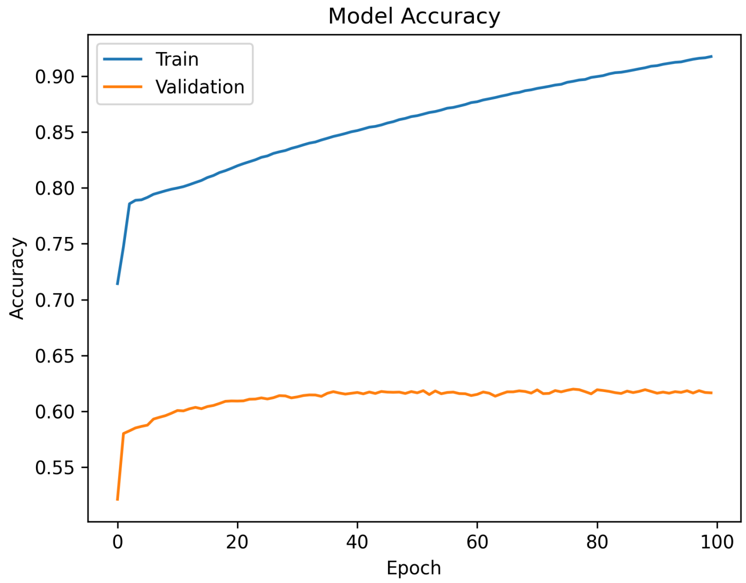

Figure 5.1: Accuracy vs Epoch Graph for Encoder-Decoder Model.

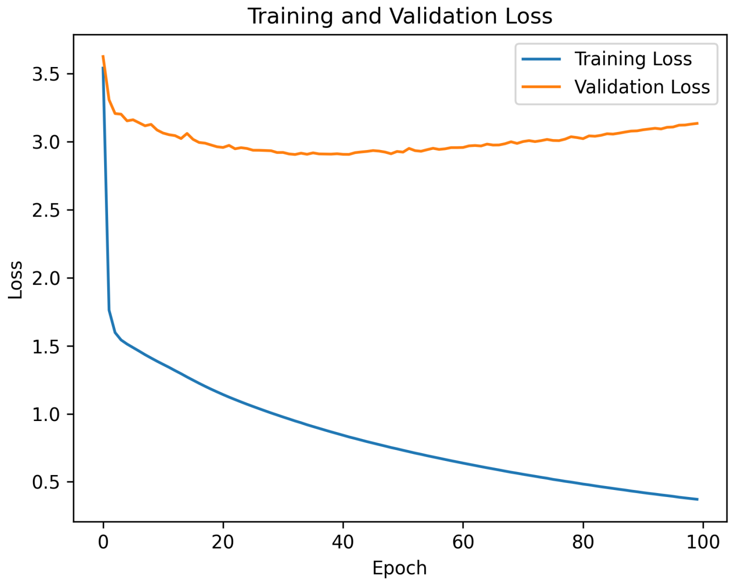

Figure 5.2: Loss vs Epoch Graph for Encoder-Decoder Model.

5.3.2 Transformer Model

Our second evaluation model is the transformer model. which shows outstanding accuracy at the time of training which is shown in Figure 5.3. The accuracy of the model on the training dataset steadily increases, suggesting the model is learning well. validation accuracy, showcases improvement over time, albeit with gentle fluctuations, suggesting the model's adaptability to diverse data patterns. Figure 5.4 shows the model loss at the time of training. Training Loss starts high and decreases as the number of epochs increases, indicating the model is learning from the training data. Validation Loss also starts high but decreases at a slower rate compared to the training loss, suggesting the model's generalization to new, unseen data.

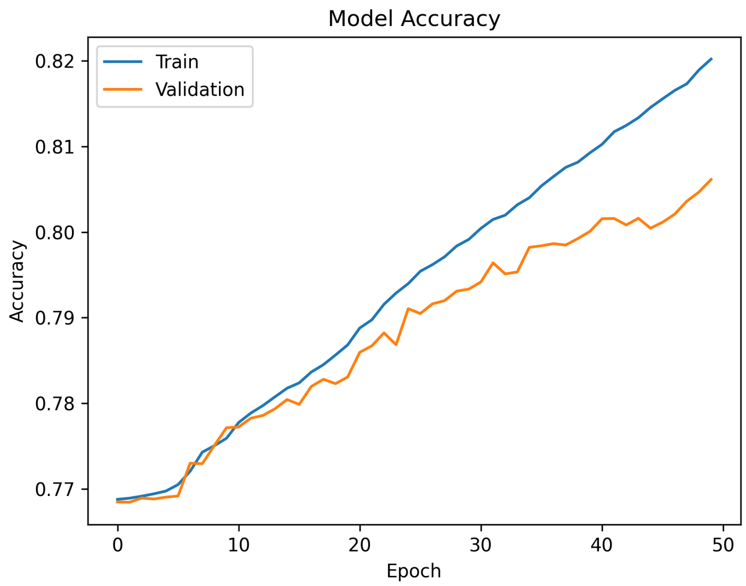

Figure 5.3: Accuracy vs Epoch Graph For transformer model.

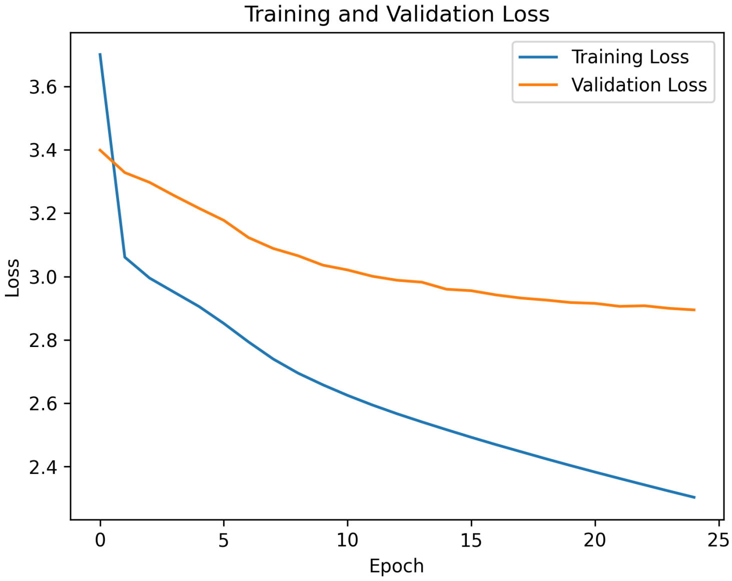

Figure 5.4: Accuracy vs Epoch Graph For transformer model.

5.3.3 Bidirectional GRU Model

Our third evaluation model is the Bidirectional GRU model. which shows better accuracy than our previous two models at the time of training which is shown in Figure 5.3. However, the validation accuracy is very low compared to the previous two models. This indicates the model is stuck in the overfitting problem. Figure 5.6 represents the loss curve of the model. In this curve, training loss is very low. but the validation loss is very high. Which is also a sign of overfitting problem.

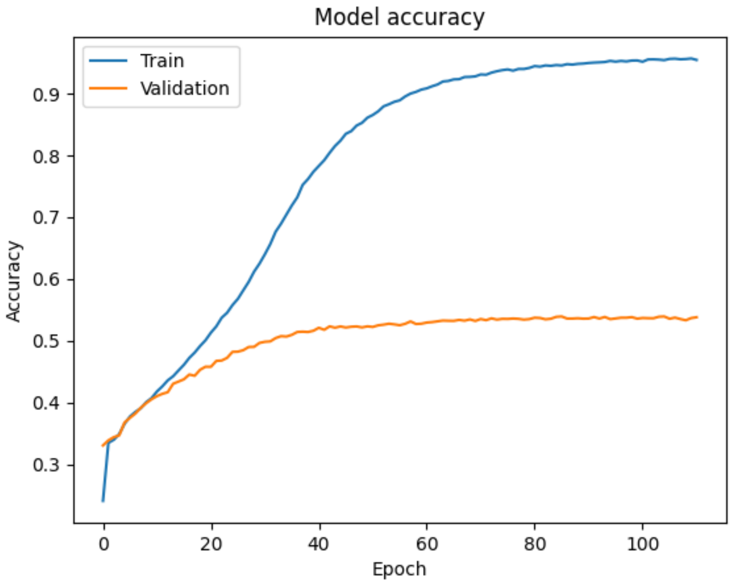

Figure 5.5: Accuracy vs Epoch Graph For GRU Model.

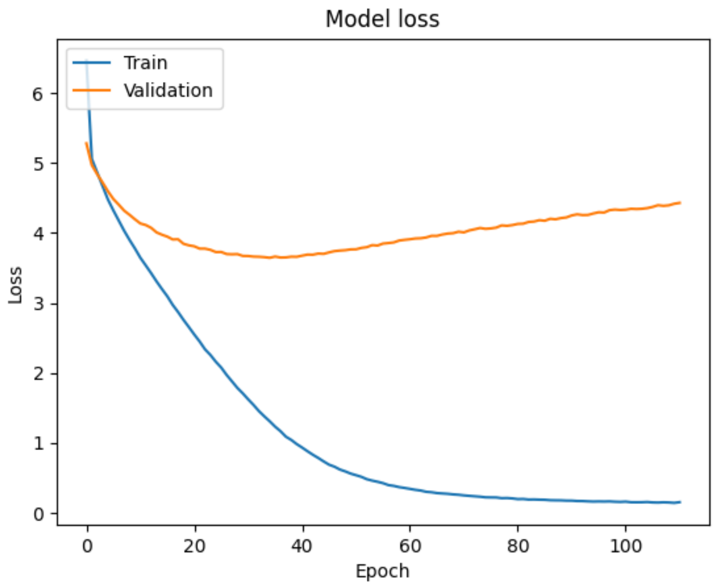

Figure 5.6: Accuracy vs Epoch Graph For GRU Model..

Table 5.1: Model Accuracy

Model	Accuracy
Encoder-Decoder	65%
Transformer	82%
GRU	90%

The table 5.1 summarizes the performance of three developed models. From the table, it is shown that GRU based model has higher accuracy. Thought only based on accuracy we can't assess a model. So, next, we compare their BLUE score and their human assessment.

Table 5.2: Model BLEU Score

Model	BLEU Score
Encoder-Decoder	9.61
Transformer	11.82
GRU	11.63
Marian	12.53

The table 5.2 summarizes the performance of three developed models and compares them with the pre-trained Marian large language model. From the table, it is shown that the GRU and the transformer-based model have a higher BLEU Score.

The table 5.3 summarizes the correctness of three developed models and compares them with the pre-trained Marian large language model. From the table, it is shown that the GRU and the transformer-based model translate nearly the same number of sentences as the Marian Large Language model.

Table 5.3: Model Correctness

Model	Number of Correct Translation between 200 sentences
Encoder-Decoder	113
Transformer	177
GRU	172
Marian	187

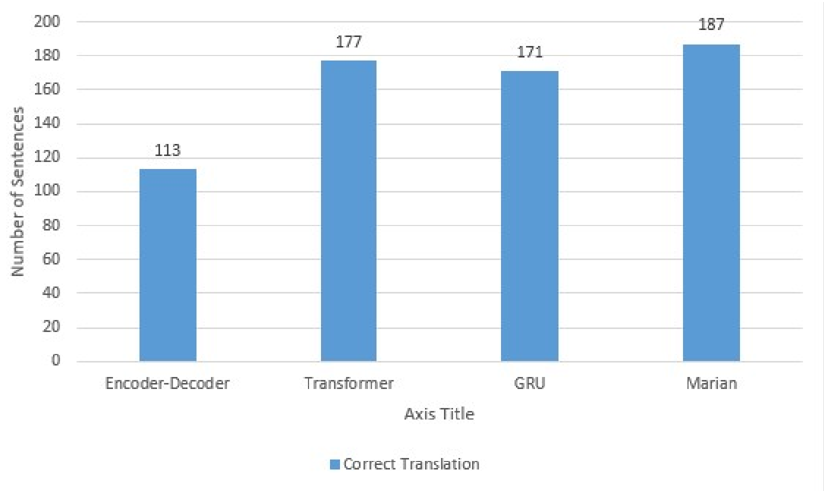

Figure 5.7: Model Number of Correct Translations

5.4 Analysis

The comparison of four translation models, namely Encoder-Decoder, Transformer, GRU, and Marian-based models, sheds light on their respective performance metrics. The accuracy rates for the first three models stand at 65%, 82%, and 90%, respectively. Correspondingly, their BLEU scores are recorded as 9.61, 11.82, and 11.63, respectively. Interestingly, these figures underscore the Transformer and GRU models' substantial efficacy in capturing translation nuances compared to the Encoder-Decoder model. Notably, the Transformer model stands out with an accuracy of 82% and a commendable BLEU score of 11.82, showcasing its superior capability in linguistic processing. The GRU model, though slightly behind the Transformer, still surpasses the Marian-based model in both accuracy and BLEU score, with an accuracy rate of 90% and a BLEU score of 11.63. This observation is particularly noteworthy, considering that the Marian model is a large-scale language model known for its robustness. The competitive performance of the Transformer and GRU models against the Marian-based model highlights their efficacy in capturing semantic nuances and linguistic intricacies. These findings suggest that while large language models like Marian offer substantial linguistic capabilities, models like Transformer and GRU demonstrate a commendable balance between performance and computational efficiency, making them promising candidates for various translation tasks.

5.5 Conclusion

To evaluate each model's performance, this chapter combines the use of the BLEU score, accuracy, and human observation. The results show that the transformer-based and GRU-based models perform better compared to the Marian big language model when compared to each other's performance using relevant architectures.

Chapter 6

Conclusion & Future Works

6.1 Introduction

This chapter offers an overview of the whole study, including the problem's scope, previous research done, the authors' original contributions, the experimental analysis carried out, and the findings drawn. Moreover, a brief overview of possible directions for further investigation is also provided.

6.2 Summary

The research aims to address the growing need for effective machine translation, particularly for the Bangla language, which is spoken by approximately 237 million people worldwide[11]. The study compares four models, the Encoder-Decoder, transformer-based, GRU-based, and LLM to determine the most optimized model for translating Bangla into English. Performance metrics of the four translation models—Encoder-Decoder, Transformer, GRU, and Marian-based models—are clarified by comparing them. The first three models had accuracy rates of 65%, 82%, and 90%. Their corresponding BLEU values are 9.61, 11.82, 11.63, and 12.53, in that order. Since Bangla is the fifth most widely spoken language[11] in the world and millions of Bangladeshis can benefit from improved communication and the removal of language obstacles, this research is essential. The results of the study may offer a useful instrument for smooth translation, improving comprehension and concreteness around the world.

6.3 Conclusion

According to the study, the Transformer model translated Bangla to English with the highest accuracy and BLEU ratings. In terms of translation adequacy and fluency, it fared better than other models.Significance for the Bangla Community: This development might greatly benefit the Bangla-speaking population by reducing obstacles to communication and improving access to worldwide knowledge. Effect on Machine Translation: The Transformer model's performance highlights how machine learning can transform language translation technology by providing more precise and effective solutions. These results support continuing initiatives to enhance machine translation and promote improved interlanguage comprehension. After a thorough comparison of these four models for translation from Bangla to English, the study offers important new information on which model works best for this kind of translation. Through analysis, these models' strengths and weaknesses have been clarified, offering a more complex picture of how well they translate from Bangla to English. While the specifics of the results are limited to the thesis's comparative analysis, it is evident that the study provides a thorough assessment to determine the best model for precise and effective translation from Bangla to English. These findings have significant ramifications for the Bangla-speaking population as well as the larger machine translation technology community. The discovery of a more efficient translation model is a step forward for the Bangla-speaking population in terms of removing language barriers, promoting more seamless communication, and gaining access to information in English, the universal language. This might provide Bangla speakers with greater educational, cultural, and economic possibilities by improving the accessibility and understandability of a wide range of English information.

6.4 Limitation

Considering the enhanced efficacy shown by the proposed methodology, it is important to recognize its significant limitations. In this study, four models are compared and analyzed for an efficient translation model. The limitations of our work are given below:

- Computational Resources: Computational constraints were a barrier to the research, compromising model complexity and data processing and perhaps affecting the breadth and depth of the findings.

- Linguistic Complexities: The study tackled the complexities of language analysis, highlighting the need for advanced natural language processing techniques to fully capture the depth of linguistic subtleties.

- Dataset Size and Scope: The datasets used in our research were limited by availability, diversity, and size, potentially limiting their generalizability and preventing exploration of relevant aspects, thereby limiting the applicability of our findings.

6.5 Future Works

There are many opportunities to broaden the focus of this research in the future. Here are some such task examples:

- The primary goal will be to increase the model's accuracy by thoroughly investigating different model architectures.

- Build a good, generalized, and balanced dataset for model training.

- Employing advanced feature reduction techniques for understanding the linguistic connection between input data.

In the future, if things permit, I hope to contribute to these difficulties to improve the model and make it more applicable and adaptable than ever.

REFERENCES

[1] S. Kostadinov, “Understanding Encoder-Decoder sequence to sequence model,” 12 2021.

[2] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, □. Kaiser, and I. Polosukhin, “Attention is all you need,” *arXiv (Cornell University)*, 1 2017.

[3] F. Chollet, “How I think about LLM prompt engineering,” 10 2023.

[4] F. Chollet, *Deep Learning with Python, Second Edition*. Simon and Schuster, 12 2021.

[5] I. Vasilev, “Advanced Deep Learning with Python.”

[6] D. J. O’Byrne and A. Hensby, *Globalization: The Global Village*. 1 2011.

[7] “Machine Translation Market Size, share Trends analysis Report by application (Automotive, military Defense, electronics, IT, healthcare, others), by technology, by region, and segment forecasts, 2023 - 2030,” 4 2024.

[8] “A (Brief) History of Machine Translation | Smartling.”

[9] Y. Lin, “The relationship between machine translation and human translation in the era of artificial intelligence machine translation,” *Applied and computational engineering*, vol. 5, pp. 133–138, 6 2023.

[10] L. N. Vieira, M. O’Hagan, and C. O’Sullivan, “Understanding the societal impacts of machine translation: a critical review of the literature on medical and legal use cases,” *Information, communication society*, pp. 1–18, 6 2020.

[11] W. contributors, “List of languages by number of native speakers,” 4 2024.

[12] I. Stevanović and L. Radičević, “Comparative analysis of machine translation systems,” *ResearchGate*, 11 2020.

[13] M. Stasimioti, V. Sosoni, K. L. Kermanidis, and D. Mouratidis, “Machine Translation Quality: A comparative evaluation of SMT, NMT and tailored-NMT outputs,” 11 2020.

[14] S. Eo, C. Park, H. Moon, J. Seo, and H. Lim, "Comparative analysis of current approaches to quality estimation for neural machine translation," *Applied sciences*, vol. 11, p. 6584, 7 2021.

[15] J. Luo, C. Cherry, and G. Foster, "To Diverge or Not to Diverge: A Morphosyntactic Perspective on Machine Translation vs Human Translation," *arXiv (Cornell University)*, 1 2024.

[16] □. Özdemir, E. S. Akın, R. Velioğlu, and T. Dalyan, "A comparative study of neural machine translation models for the Turkish language," *Journal of intelligent fuzzy systems*, vol. 42, pp. 2103–2113, 2 2022.

[17] C. Raffel, N. Shazeer, A. Roberts, K. Lee, S. Narang, M. Matena, Y. Zhou, W. Li, and P. J. Liu, "Exploring the Limits of Transfer Learning with a Unified Text-to-Text Transformer," *arXiv (Cornell University)*, 1 2019.

[18] W. contributors, "History of machine translation," 4 2024.

[19] J. Vičič, "A fast implementation of rules-based machine translation systems for similar natural language," *Informatica*, vol. 37, pp. 455–456, 12 2013.

[20] F. Josef and H. Ney, "Statistical machine translation," 06 2001.

[21] D. Bahdanau, K. Cho, and Y. Bengio, "Neural machine translation by jointly learning to align and translate," *arXiv (Cornell University)*, 1 2014.

[22] V. Shah, S. Valiullah, and M. Makwana, "Neural machine translation to local languages," 04 2023.

[23] S. A. Mohamed, A. Elsayed, Y. Hassan, and M. A. Abdou, "Neural machine translation: past, present, and future," *Neural computing applications*, vol. 33, pp. 15919–15931, 7 2021.

[24] F. Liu, W. Ding, Y. Qiao, and L. Wang, "Transfer learning-based encoder-decoder model with visual explanations for infrastructure crack segmentation: New open database and comprehensive evaluation," *Underground space*, vol. 17, pp. 60–81, 8 2024.

[25] J. Chung, □. Gülçehre, K. Cho, and Y. Bengio, "Empirical evaluation of gated recurrent neural networks on sequence modeling," *arXiv (Cornell University)*, 1 2014.

[26] R. Dhayarkar, "Neural Machine Translation — A comprehensive guide - Rishikesh Dhayarkar - Medium," 12 2021.

[27] M. E. Consens, C. Dufault, M. Wainberg, D. Forster, M. Karimzadeh, H. Goodarzi, F. J. Theis, A. M. Moses, and B. Wang, "To Transformers and Beyond large language models for the genome," *arXiv (Cornell University)*, 1 2023.

[28] "Machine Translation Market Size, Share Trends Analysis Report by application (Automotive, Military Defense, electronics, IT, healthcare, others), by technology, by region, and segment forecasts, 2023 - 2030," 10 2023.

[29] "Tab-delimited Bilingual Sentence Pairs from the Tatoeba Project (Good for Anki and Similar Flashcard Applications)."