PJETS
Pakistan Journal of Engineering Technology and Science (PJETS)
Volume 11. Issue. 1, PP. 29-44, November 2023
E-ISSN 2224-2333
https://doi.org/10.22555/pjets.v11i1.1002
P-ISSN 2222-9930

A REVIEW OF EXISTING MACHINE TRANSLATION APPROACHES, THEIR CHALLENGES, AND EVALUATION METRICS FOR THE LOW-RESOURCE LANGUAGES

Naseer Ahmed [1],* Mansoor Ahmed Khuhr[2], Mazhar Ali Abro [3]

[1] Department of Computer Science, Sindh Madresatul Islam University, Karachi, Pakistan.
[2] Department of Artificial Intelligence and Mathematical Sciences, Sindh Madressatul Islam University, Karachi.
[3] Center for Research on Microgrids (CROM), AAU Energy, Aalborg University, 9220 Aalborg, Denmark

*Corresponding author: makhuhro@smiu.edu.pk
naseerbajoi@gmail.com

Abstract: **The Machine translation is the process of translating a natural language into another language. The primary goal of machine translation is to bridge the linguistic gap between languages. A significant job that may be utilized is to extract information from texts written in different languages through machine translation (MT). Due to the remarkable developments in Machine Translation over the past several years, we have entered the era of Neural Machine Translation (NMT). A review of MT is essential for a better understanding of MT in the domain of Natural Language Processing. This review significantly deals with the challenges and evaluation metrics associated with low-resource languages as compared to previous reviews where attention was given to general machine translation techniques. This paper focuses on Statistical, Corpus, Example, Rule-based, Transfer-based, direct-based, and Neural Machine Translation approach, their challenges, and existing evaluation metrics such as BLEU score and Human evaluation metrics.**
Keywords: **Machine Translation, Neural Machine Translation, BLUE metrics, Human Evaluation metrics, Natural Language Processing Statistical Machine Translation, Machine Translation approaches and Machine Translation Process Cycle.**

I. INTRODUCTION

Recent years have seen a rise in machine translation (MT) because of the rapid growth of artificial intelligence [1]. Language is a source of communication through which people may express their ideas, feelings, desires, and other emotions, which is a powerful and organized way of communication. There are around 6,800 active languages in the world [2]. There are two types of machine translation (MT): paraphrase and metaphrase [3]. Machine Translation[4]. The LSTM neural network type has been successfully used for machine translation. It can handle sequential data types like text [5]. Many scarce, less-resourced languages are still not included in the evaluation of MT. It is much more challenging to develop a corpus for such languages [6]-[7] and [8]. However, numerous studies have demonstrated that primary MT research methodologies, such as models and evaluation metrics, have limited application when translating languages with different linguistic [9]. While other reviews mainly focus on general machine translation approaches, our review specifically tackles the challenges and evaluation metrics associated with low-resource languages. In another research, it has been discussed that A Hybrid, Example-based, and Statistical Machine Translation approach was used to perform MT on English to other languages on Parallel corpus using evaluation metrics such as

E-ISSN 2224-2333
P-ISSN 2222-9930
Pakistan Journal of Engineering Technology and Science (PJETS)
Volume 11. Issue. 1, PP. 29-44, November 2023
https://doi.org/10.22555/pjets.v11i1.1002

BLEU (Bilingual Evaluation Understudy) [10]. The authors suggested A Phrase-based MT Transliteration for English to Assamese to enhance machine translation. Wordnet of Assamese is utilized on phrases in both English and Assamese to enhance MT output [11]. Phrase-based MT is also implemented for English to Bangla languages [12]. The impacts of sentence simplification in the Hindi-to-English machine translation system are being investigated [13]. The authors implemented a new method for machine translation that provided noticeable results. The BERT has achieved great success in language understanding tasks[14]. The authors also introduced another approach through which RoBERT has gained more attention in machine learning and natural language processing. The models are used for the downstream tasks by giving context-aware embedding of an input sequence after being pre-trained on many unlabelled data to capture rich input representation [15].
In the same way, various evaluation metrics are also introduced to evaluate the quality and accuracy of a machine translating information from one Language to another[16]-[17]. For evaluating the Machine Translation, the BLEU score is widely used [18]. Machine Translation systems are also evaluated using the Human evaluation metric[19]. Machine translation has significant objectives when translating the source Language into the target language [2].

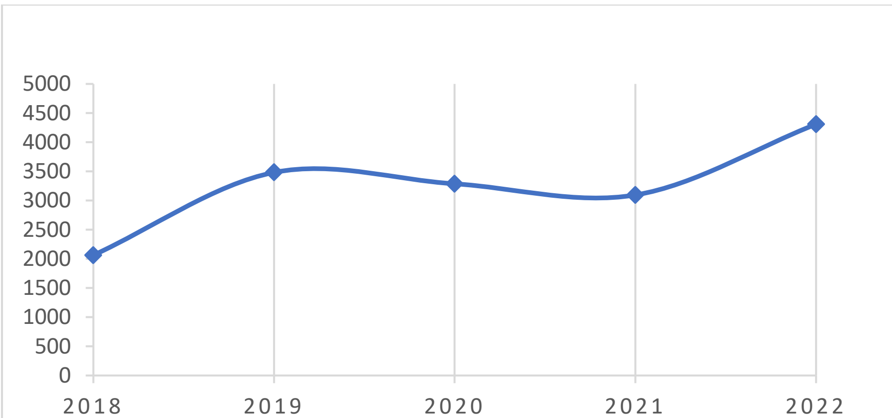

Figure 1. Machine Translation Trends 2018 to 2022 [20].

The MT trend in 2020 and 2021 is modest, and several potential causes exist as shown in Figure 1. It's possible that the COVID-19 epidemic had a detrimental effect on the MT sector. Since the pandemic, there has been less demand for MT services since there has been less travel and tourism. The availability of parallel corpora, which are text databases translated into two or more languages, also decreased due to the epidemic. Because parallel corpora are crucial for training MT systems, creating new MT systems has become increasingly challenging as their availability has decreased. It's also possible that the MT market was already saturated, which would explain the downward trend in MT in 2020 and 2021. Many well-known MT businesses, including Google Translate, Microsoft Translator, and Babylon, existed by 2020. It was challenging for new entrants to compete since these businesses had previously made significant investments in creating MT systems.
However, in 2022, the trend in MT has risen. There are several potential explanations for this. One reason for the rise in travel and tourism is that the COVID-19 outbreak has receded. As a result, there is now more demand for MT services. Parallel corpora have also become more widely available in recent years. New MT system development has become simpler as a result, which has increased the number of MT systems. Instead, the machine translation still faces some major issues, as shown in Figure 3. The usage of MT technology is becoming more common, and one reason for this is that it is now more accessible, which is why Literacy of machine translation is increasing daily [20].

A. Machine Translation Approaches

Over the years, there have been methods for overcoming language obstacles with machine translation as shown in Figure 2. Conventional rule-based techniques depend on dictionaries and rules, whereas transfer-based machine translation uses an intermediate language to facilitate multilingual translation. However, direct-based machine translation

PJETS
E-ISSN 2224-2333
P-ISSN 2222-9930
Pakistan Journal of Engineering Technology and Science (PJETS)
Volume 11. Issue. 1, PP. 29-44, November 2023
https://doi.org/10.22555/pjets.v11i1.1002

streamlines the procedure by translating between the source and target languages. While corpus-based machine translation uses methodologies, example-based systems use parallel corpora to discover examples. Furthermore, the translations produced by example-based machine translation rely on training data examples. Neural network-based machine translation models, such as Transformers, are popular today. Statistical machine translation uses models, whereas based machine translation is considered a state-of-the-art approach that significantly enhances the quality of translations. These different approaches offer a range of techniques to tackle the complexities of machine translation.

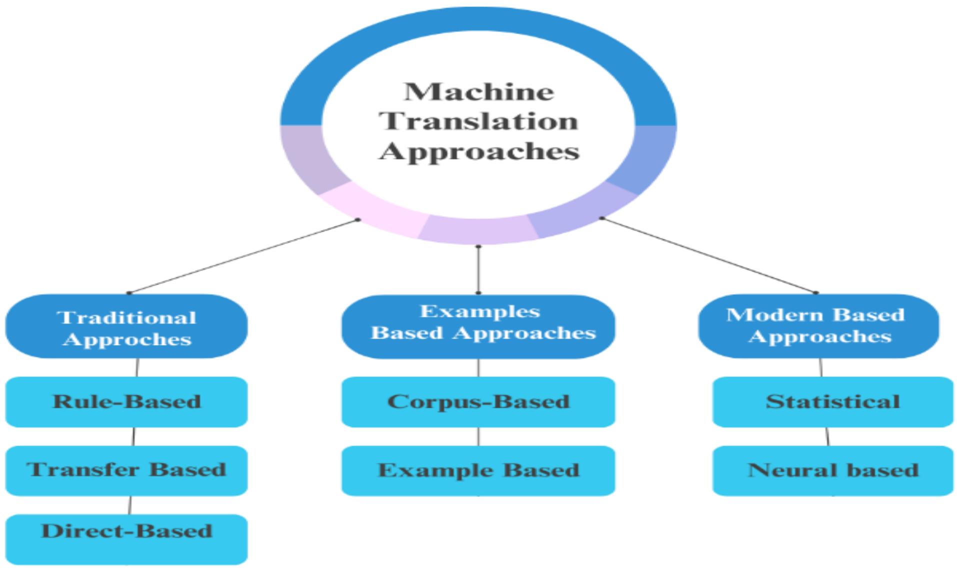

Figure 2. Machine Translation Approaches

B. Machine Translation Process Cycle

In Figure 3, we can see the cycle of the machine translation process when translating a sentence from Brahui to English. The process begins by feeding the sentence into a machine translation system. The system first analyses the meaning of the sentence. Then, it combines that meaning with the grammar and syntax of the Language to generate a new sentence. Finally, the machine translation system produces the translated sentence as its output. There are several approaches to machine translation [21]. AI chatbots assist with language translation by acting as partners in conversation [22].

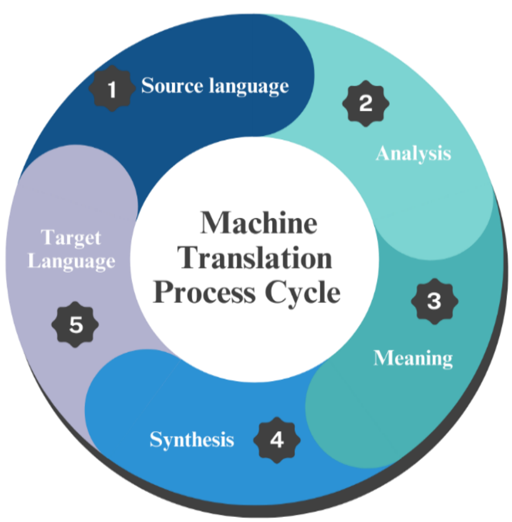

Figure 3. Machine Translation Process Cycle

Pakistan Journal of Engineering Technology and Science (PJETS)
Volume 11. Issue. 1, PP. 29-44, November 2023
https://doi.org/10.22555/pjets.v11i1.1002
E-ISSN 2224-2333
P-ISSN 2222-9930

C. Selection criteria of papers for review

Figure 4 presents the selection criteria for papers for this review, focusing on machine translation challenges and evaluation metrics related to low-resource languages. The searching and selection criteria are divided into four levels. Level 1. Initial Search: The search for relevant research papers thoroughly used Google Scholar. The terms machine translation, human evaluation metrics, BLEU score, and low-resource language MT were among the search terms. Three hundred (300) publications were obtained for review. Level 2. Duplicate Removal: Duplicate publications were found and removed to reduce redundancy and concentrate on original research viewpoints, getting a set of 100 papers. Level 3. Full-Text Availability: Only publications with easily accessible full texts were considered, further narrowing the list to 70 articles. Level 4. Analysis and Categorization: A thorough study was conducted to extract important information from the remaining 70 papers. Significant difficulties in low-resource language machine translation. The effectiveness of different evaluation criteria, such as the human evaluation and the BLEU score. The frequency of research on machine translation was presented at the ACL conference from 2018 to 2022. This systematic methodology ensured that the review paper had a strong and relevant base, ensuring that the analysis and conclusions were founded on a reliable selection of recent research.

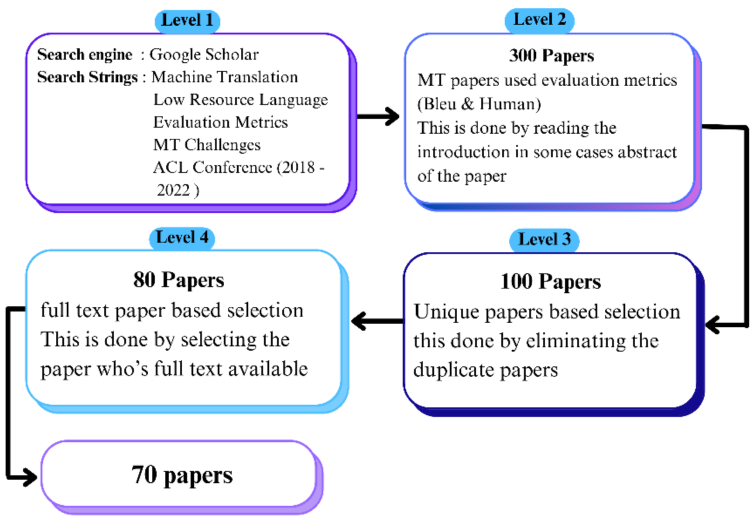

Figure 4. Paper selection criteria

D. Statistical Machine Translation

The SMT approach is a useful machine translation approach, as shown in Table 1, although its accuracy will depend on the language pair and dataset used. The variation in dataset sizes used for each language pair may cause the accuracy mismatch between English-Hindi and English-Urdu translations. It can be seen from Table -1 that six languages have been selected because they are low-resource languages commonly used in the literature.

E. Corpus-Based Machine Translation

The efficiency of corpus-based machine translation (MT) methods for translating closely related languages has been researched in two studies as shown in Table 2. Both research studies established the effectiveness of corpus-based MT methods for translating closely related languages. It is important to remember that the dataset size and quality can impact the performance of MT systems.

PJETS
E-ISSN 2224-2333
P-ISSN 2222-9930
Pakistan Journal of Engineering Technology and Science (PJETS)
Volume 11. Issue. 1, PP. 29-44, November 2023
https://doi.org/10.22555/pjets.v11i1.1002

Table 1. Statistical Machine Translation

Ref	Description	Dataset	Evaluation Metrics and accuracy	Experimental work /results	Language	Author/s
[23]	A hybrid method integrating SMT and EBMT is used to conduct MT.	54K sentences in English and Hindi. 53K sentence drills. Examining: 100	The accuracy through the BLEU score is 0.432.	Nil	English to Hindi MT	Ambati
[24]	The SMT method is used for MT	private	Human Evaluation	Nil	Translating Texts from Persian to English	Saffari M
[25]	The SMT method is used for MT	private	Nil	Nil	English to Arabic	Ibrahim Jibreel
[26]	The SMT method is used for MT,	120153 words are utilized for Training, while 8557 words are for Testing.	Nil	Nil	English to Hindi MT	Ramanathan, A., Bhattacharyya
[27]	The SMT approach is used for MT on English and Hindi letters	Private SRILM toolkit for Training.	Mean F-Measure obtained: 0.876.	46.3% accuracy using GIZA++,	Transliteration of English to Hindi	Rama & Gali
[28]	The SMT approach is used for MT in English and Urdu.	IRSTLM and Moses translation setup were used to train the model on Trainset, where 20173 sentence pairs are used for Training, 5059 for fine-tuning, and 1590 for Testing.	BLEU score	37.10 %	English-Urdu	Ali, A., Hussain
[29]	Statistical MT was used to translate the Telugu language into English.	dataset of 10 lack words was used	Nil	Nil	Telugu to English	Raju

PJETS
E-ISSN 2224-2333
P-ISSN 2222-9930
Pakistan Journal of Engineering Technology and Science (PJETS)
Volume 11. Issue. 1, PP. 29-44, November 2023
https://doi.org/10.22555/pjets.v11i1.1002

Table 2. Corpus-Based Machine Translation

Ref	Description	Dataset	Evaluation Metrics and accuracy	Experimental work /results	Language	Author/s
[11]	Corpus-based MT approach used to perform MT.	Where 14,371 sentences of English and Assamese were used as dataset; Testing: 500 sentences	Nil	Wordnet of Assamese is used to improve MT output.	English to Assamese	Barman et al.
[30]	A corpus-based MT approach translated Bhagavad Gita from Sanskrit to Hindi.	145,34,215 sentences are used as a dataset; Training sentences are 12698 and Testing sentences are 192679.	RBMT	The author claimed the results of the proposed MTS are 24% better than those of other Translation systems.	Sanskrit to Hindi	Singh et al.

F. Example-Based Machine Translation

Table 3 shows that the usefulness of example-based machine translation (MT) methods for translating languages that are closely related to each other was examined in three studies. To train MT systems, example-based MT approaches used a corpus of parallel texts. The authors did not show any experimental results in their research. Moreover, it's worth mentioning that all three studies used their private dataset.

Table 3. Example-Based Machine Translation

Ref	Description	Dataset	Evaluation Metrics and accuracy	Experimental work/results	Language	Author/s
[10]	The example-based MT approach used performance alignment.	Private	BLEU score	Using a manual and statistical dictionary built from GIZA++.	English to Indian Languages	Ambati
[31]	The primary focus of this investigation is the Example-based approach in comparing sentences, which can help translate them effectively.	Private	Nil	Nil	Sindhi to Hindi	Khan et a
[32]	Example-based approach with a rule-based used to translate sentences. An attempt was made to fuse modern artificial intelligence techniques with the classical Paninian framework based on Sanskrit grammar.	Private	Nil	Nil	English to Hindi	Sinha & Jain

PJETS

Pakistan Journal of Engineering Technology and Science (PJETS)
Volume 11. Issue. 1, PP. 29-44, November 2023
E-ISSN 2224-2333
https://doi.org/10.22555/pjets.v11i1.1002
P-ISSN 2222-9930

G. Rule-Based Machine Translation

As shown in Table 4, High levels of MT accuracy can be achieved using rule-based MT approaches. However, the accuracy depends on the language dataset used. For instance, the Sinhala-English MT system obtained a BLEU score of 77.33%, while the Tamil-English MT system received a score of 85%. Rule-based MT approaches can be used for languages with little relevant data. This is because rule-based MT techniques require less data than other MT techniques, such as statistical MT. Therefore, the rule-based translation technique can be used for low-resource languages such as Brahui and Balochi, as well as other regional languages of Pakistan.

Table 4. Rule-Based Translation Method

Ref	Description	Dataset	Evaluation Metrics and accuracy	Experimental work/results	Language	Author/s
[33]	The authors used a rule-based method for transliteration, which developed CLIR systems for Hindi and Marathi to English.	Their approach utilizes bi-lingual dictionaries for query translation.	Nil	Nil	Hindi and Marathi to English	Chinnakotla et a
[32]	Sentences are translated using a rule-based approach	Private	Nil	90%	English to Hindi	Sinha & Jain
[34]	A rule-based text simplification technique is adopted to improve English to Tamil MT.	Testing: 200 sentences. Accurate results in 115 sentences.	BLEU	28%	English to Tamil	C et al.
[35]	Rule-based MT was used to translate the Sinhala language into English.	for 300 words were collected from social media because there is no dataset available for	BLEU	77.33%	Sinhala Language	Nandathilaka

PJETS
Pakistan Journal of Engineering Technology and Science (PJETS)
Volume 11. Issue. 1, PP. 29-44, November 2023
E-ISSN 2224-2333
https://doi.org/10.22555/pjets.v11i1.1002
P-ISSN 2222-9930

		Sinhala language				
[36]	The rule-based MT approach was used on the Tamil language to translate into English, for which authors used their dataset where accuracy achieved was evaluated through the score.	private	BLEU	85%	Tamil to English	Kumar & Singh

H. Transfer-Based Machine Translation

Table 5 shows that both transfer-based MT and direct MT have been applied to various languages, but the authors have not shown any experimental results. Moreover, neither machine translation method has been used compared to the other methods. It can be illustrated from the table that the method's accuracy depends on the corpus.

Table 5. Transfer-based and direct MT

Ref	Description	Dataset	Evaluation Metrics and accuracy	Experimental work /results	Language	Author/s
[37]	A comparative analysis of Transfer-based MT and corpus-based MT was done	Private	Nil	Nil	Hindi to English	Gehlot et al
[38]	The Direct MT approach translated Marathi to English.	Private Data Set Dictionary	Nil	Nil	Marathi to English	Todase

I. Neural-Based Machine Translation

Table 6 shows that Neural MT methods use artificial neural networks to translate text from one Language to another. Neural networks are a type of machine learning algorithm that can learn to recognize patterns in data. The BLEU score metric is used to evaluate the accuracy of machine translation systems. A large dataset is required to train MT systems properly. This is because neural networks need a lot of data to learn the patterns in Language. The neural method has resulted in more significant improvements in MT quality[44].

PJETS
E-ISSN 2224-2333
P-ISSN 2222-9930
Pakistan Journal of Engineering Technology and Science (PJETS)
Volume 11. Issue. 1, PP. 29-44, November 2023
https://doi.org/10.22555/pjets.v11i1.1002

Table 6. NMT

Ref	Description	Dataset	Evaluation Metrics	Experi-mental work /results	Language	Author/s
[39]	The neural MT approach was used to translate Tamil into English.	183451 training corpus used 1000 sentences for validation and 2000 sentences for Testing.	BLEU source	80% of the data used is encoded in UTF 8.	Tamil to English	Choudhary
[40]	The NMT approach is used to translate Kannada text to English through LSTM.	For which No of training sentences for English 40500 and Kannada 40500 for testing purposes, 500 English and 500 Kannada sentences were used	BLEU source	86.32 %.	Kannada to English	Nagaraj et a
[41]	The NMT approach is used to translate English to Hindi through Bi-directional LSTM.	the data set is used, which is available at https://www.clarin.eu/resource-families/parallel-corpora canting 3 lakh Sentences	BLEU source	21.97%	English to Hindi	Gogineni et a
[42]	The NMT approach translates English to Hindi sentences through Quantum Neural Network.	Private corpus canting 2600 sentences and 11500 words used	BLEU source	91.30%.	English to Hindi	Narayan et al.
[43]	The NMT approach translates English to Hindi sentences through a Deep Neural Network. As per	Private bilingual corpus canting 20000 words used	Nil	Nil	English to Hindi	Singh et a

II. CHALLENGES

The majority of NMT models assume that there are sufficient multilingual training data [45]. Designing a better MT framework with high accuracy is significantly challenging [46] Recently, the idea of using NMT to complete the MT task was offered. [47]-[48]. In a different study, authors offered a new encoder-decoder architecture called RNMT+ that combines the benefits of RNN-based and Transformer-based translation models to address the issues with machine translation [49]-[50]. Figure 5 shows the challenges with machine translation.

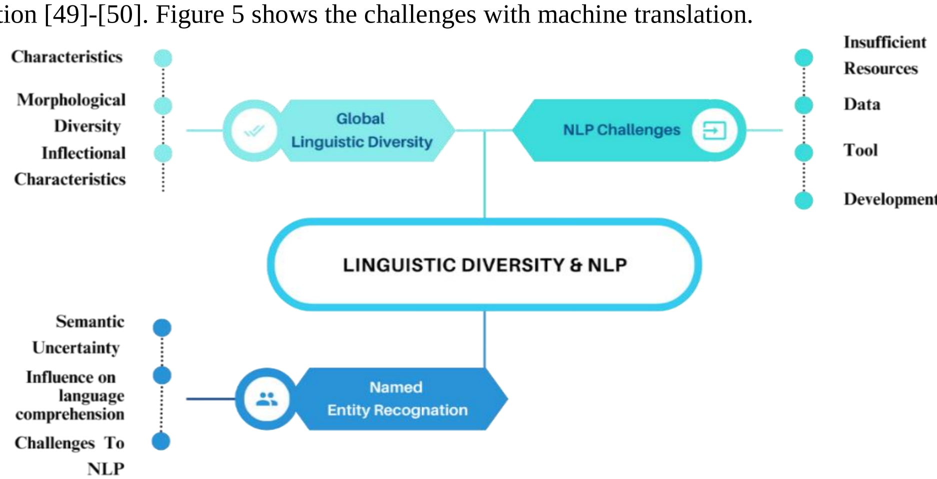

Figure 5. Challenges with machine translation

E-ISSN 2224-2333
P-ISSN 2222-9930
Pakistan Journal of Engineering Technology and Science (PJETS)
Volume 11. Issue. 1, PP. 29-44, November 2023
https://doi.org/10.22555/pjets.v11i1.1002

The Figure 5 illustrates the relationship between diversity and Natural Language Processing (NLP). It presents three elements. Firstly, it emphasizes the array of languages worldwide, showing their unique grammatical and structural features. Secondly, it highlights the difficulty NLP faces in focusing on a language with rich linguistic resources while neglecting others that lack sufficient data for developing NLP tools. Lastly, it underscores the role played by Named Entity Recognition (NER). However, NER becomes complex due to forms of nouns in related languages, leading to ambiguity and challenges for NLP systems. In another investigation, the focus was on designing a much deeper Transformer model; after a deep analysis, the authors traced the various challenges in machine translation [51]. The common challenges in every translation approach are discussed below in Table 7.

Below, Figure 6 represents the issues with machine transition in Natural Language Processing (NLP); techniques and processes are used to understand and work with human languages. One such technique is Morphological Analysis, which helps break down words into their root forms, making it easier to identify those forms based on word-level information. Part of Speech Tagging is another technique that assigns labels, like nouns, verbs, and adjectives, to words in sentences, helping us to understand how sentences are structured. Chunking is a process that groups words into phrases, such as noun or verb phrases, giving us an organized view of sentence content. Parsing involves the use of Parse Trees. Provides a detailed analysis of sentence structure by combining part of speech tagging and chunking data. Lastly, Word Sense Disambiguation helps resolve the meanings of words in context when they have interpretations. These various processes work together to ensure accurate language understanding in NLP applications.

Usually, automatic machine translation evaluation metrics are used [52]. Machine translation is not feasible without automated metrics[53]. Several machine translation evaluation metrics include Human evaluation, TER, TESLAM, and METEOR[54]. Twenty papers that used BLEU and human evaluation metrics were reviewed. The last five years of papers selected for review are often considered general criteria [55]. It is important to evaluate MT systems to determine how accurate translation techniques are[56].

Table 7. MT approaches and their challenges

Translation Approach	Challenges
Statistical-based MT system	Developing a parallel corpus from the copyrighted resources provided is quite challenging.
	Ineffective for languages with irregular word ordering.
Corpus-Based MT system	This MT fails for Language with less resources.
	Building high-quality training corpora requires a lot of effort.
Example-Based MT system	The quality of the bilingual corpus influences the accuracy of the translation. Developing multilingual corpora with a high degree of similarity takes a lot of effort.
Rule-based MT system	The unavailability of multilingual dictionaries, both qualitative and quantitative.
Transfer Based MT system	Results could be unexpected depending on the quality of the accessible resources.
Direct-Based MT system	As direct translation mostly concentrates on word-to-word translation, there is a chance of mistranslation at the lexical, word order, and word meaning levels.
Neural-Based MT system	This MT fails for language pairs with limited resources.
	The fundamental issue is the acute shortage of parallel corpora for languages with minimal resources.
	A challenge is selecting the appropriate deep-learning algorithm.

E-ISSN 2224-2333
P-ISSN 2222-9930

Pakistan Journal of Engineering Technology and Science (PJETS)
Volume 11. Issue. 1, PP. 29-44, November 2023
https://doi.org/10.22555/pjets.v11i1.1002

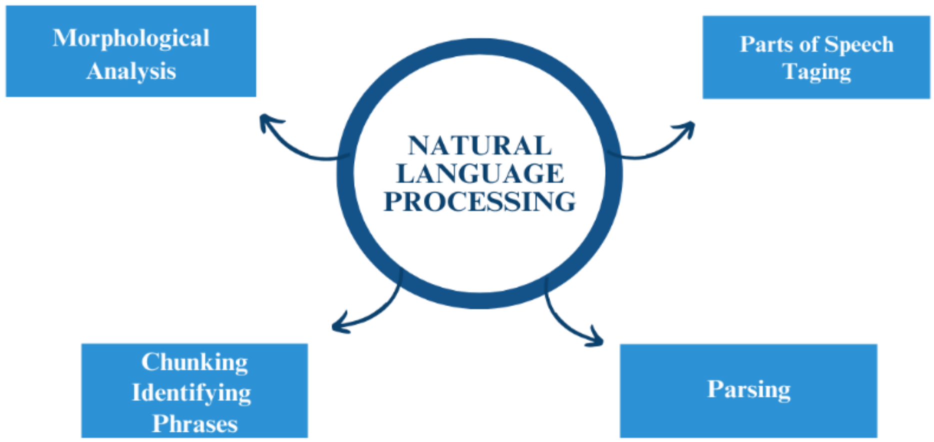

Figure 6. Issues with machine translation

III. MACHINE TRANSLATION EVALUATION METRICS

A. Bilingual Evaluation Understudy (BLEU)

For more than 15 years, the BLEU metric has been frequently used in NLP to assess NLP systems, particularly in machine translation and natural language generation [57]. Evaluating the translation closeness is the fundamental concept of BLEU [52]. Compared to human evaluation, automatic methods to evaluate translation quality have several advantages. With an existing reference translation, they are quick and almost free to use. Additionally, their results are repeatable. A five-point scale is used to evaluate the machine translation quality, as shown in Figure 7. Adequacy and fluency scores may be calculated using five-point scales in Graphs 8,9 and 10.

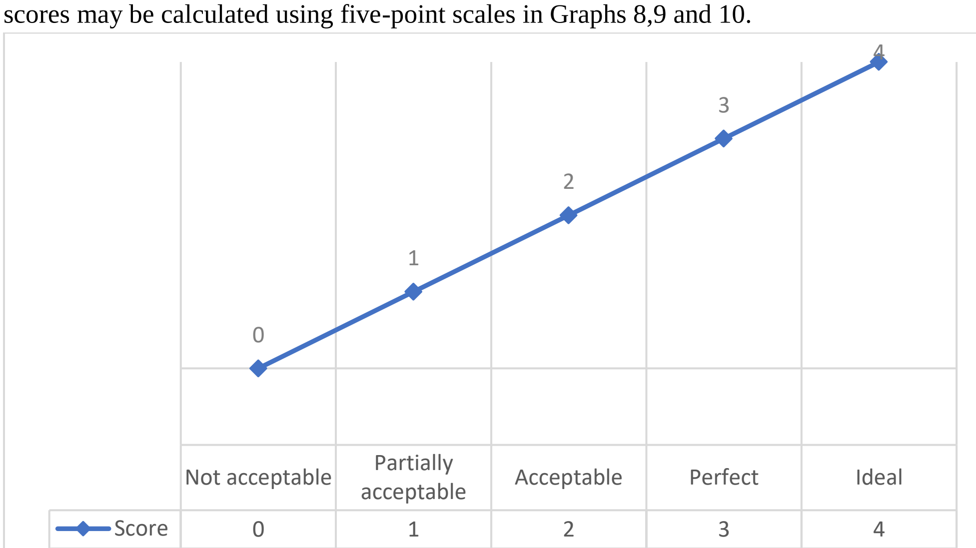

Figure 7. A five-point rating scale for evaluation of machine translation quality

Figure 7 shows a rating scale for the evaluation of machine translation quality. The scores range from 0 to 4, with 0 being with not acceptable and 4 being ideal. According to Figure 7 the highest level of translation quality is the level 4. This means that the translation is so well done that it is virtually indistinguishable from a text translated by a human. While also quite good, the perfect level may have some minor errors. On the other hand, the acceptable level is understandable. It may contain some potentially confusing errors. The partially acceptable level falls in the middle. It

PJETS
E-ISSN 2224-2333
P-ISSN 2222-9930
Pakistan Journal of Engineering Technology and Science (PJETS)
Volume 11. Issue. 1, PP. 29-44, November 2023
https://doi.org/10.22555/pjets.v11i1.1002

can be moderately difficult to comprehend. Lastly, the unacceptable level represents the quality of translation, where it becomes completely incomprehensible and unusable.

Figure 8 shows a scoring system often used to evaluate the amount of information provided in a text. The scores range from 0 to 4, with 1 being the least information and 4 being the most complete information. Within this system, each score indicates the comprehensiveness of the information contained in a given text. The top rating, "Complete Information," indicates that the text effectively presents all the information. As we go down the scale, "Most Information" implies that while the text includes the required details, there might be some gaps or omissions. "Much Information" suggests that although the text provides specifics, some details may still be missing. Continuing this "Little Information," the text offers some information that may not be sufficient to understand. Lastly, at the bottom of the spectrum is "None," which describes a text that fails to provide all required information. This scoring system serves as a tool to evaluate and categorize content based on its value.

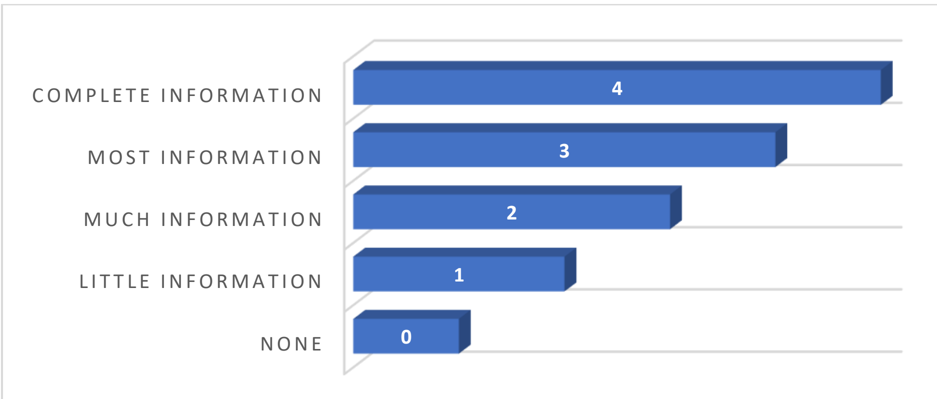

Figure 8. A five-point measure to assess the adequacy of machine translation

In Figure 9, a scoring system is used to assess the fluency of machine translation. The scores range from 0 to 4. These scores are significant in determining the text quality generated by machine translation. The top level, "Ideal," represents text that may not flow perfectly but somehow provides the information. Below is the "Good " indicating fluent text with only a few small mistakes. Moving down the scale, we encounter "Non-Native," meaning understandable text. It may contain noticeable errors leading to potential ambiguity.

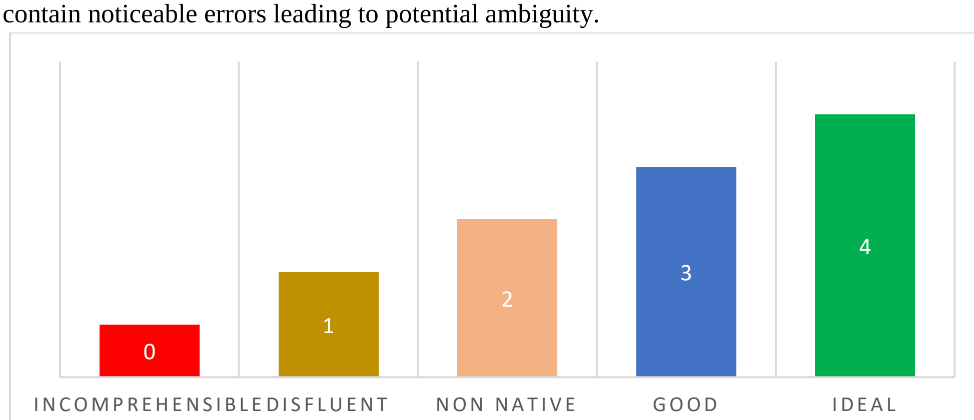

Figure 9. A five-point scale to assess the fluency of machine translation

Conversely, "Disfluent" suggests that the text might be challenging to comprehend due to grammar and sentence structure. Finally, at this point, we have "Incomprehensible," which describes completely unclear text. These various rating levels have been essential for assessing and categorizing the quality of machine-translated text processes of significance over the last o decades[58].

PJETS
E-ISSN 2224-2333
P-ISSN 2222-9930
Pakistan Journal of Engineering Technology and Science (PJETS)
Volume 11. Issue. 1, PP. 29-44, November 2023
https://doi.org/10.22555/pjets.v11i1.1002

B. Bilingual Evaluation Understudy (BLEU)

The Document-level evaluation is as important as the document-level MT methods and booster of MT technology. [59] For human evaluation, some linguistic features are used [60]. Are shown in the Figure 10 below:

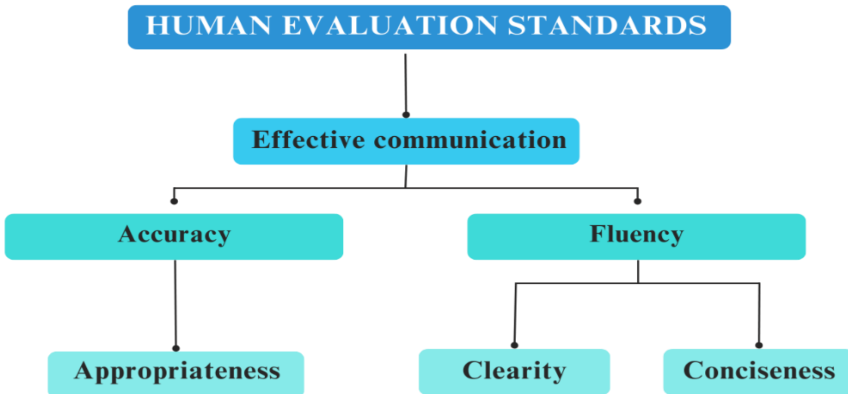

Figure 10. Human evaluation standards

Figure 10 shows a set of criteria for evaluating communication. It begins with "Accuracy," highlighting the significance of providing information. This is followed by "Fluency," which emphasizes the importance of delivering messages. Next, "Appropriateness" underscores the need to tailor communication to the intended audience. Then comes "Clarity," emphasizing the importance of ensuring comprehension. Lastly, we have "Conciseness," stressing the value of keeping messages brief and to the point. Additionally, Figure 9 displays the Comparative study of the Human Evaluation Metrics and BLEU from 2018 to 2022. The MT research papers presented at *ACL conferences from 2018 to 2022 were manually annotated. We looked for the phrases "MT" or "translation" in their names on the ACL Anthology website to find MT works, and we examined two machine evaluations. The research that used BLEU and Human evaluation metrics for machine translation is shown in the Figure below.

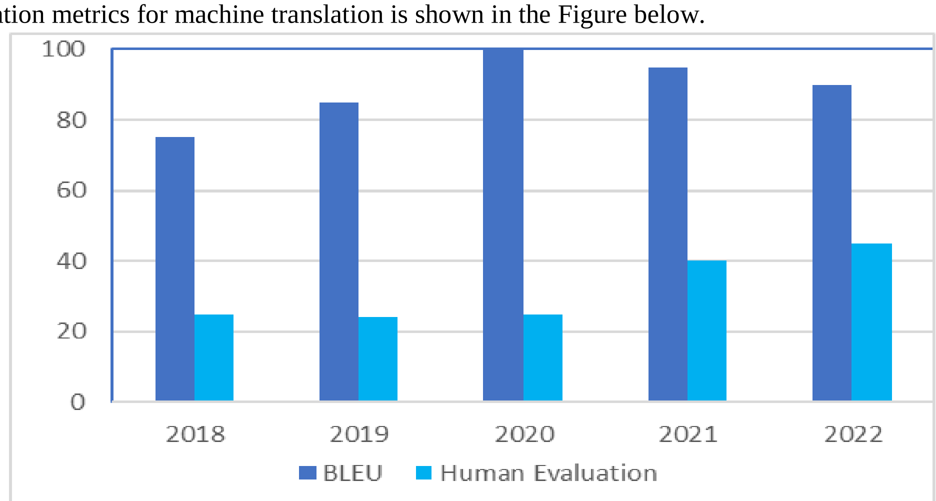

Figure 11. %Age of papers using BLEU and Human Evaluation metrics per year used in over twenty publications.

Figure 11 shows that the BLEU score has consistently outperformed human evaluation methods over the past five years. The BLEU score is more accurate, efficient, and reliable than human evaluation methods for assessing translation quality.[61] because of that, BLEU is used more frequently than human evaluation measures. According to extensive research, human translations are preferred over machine translation MT when applying more effective evaluation techniques [62].

PJETS

Pakistan Journal of Engineering Technology and Science (PJETS)
Volume 11. Issue. 1, PP. 29-44, November 2023
E-ISSN 2224-2333
https://doi.org/10.22555/pjets.v11i1.1002
P-ISSN 2222-9930

IV. DISCUSSION AND FUTURE DIRECTIONS

This study carried out a thorough analysis of previous reviews on machine translation of low-resource languages. However, the previous reviews primarily focused on the topic of machine translation. The challenges of translating languages with limited resources and evaluation standards are examined in this study. The paper sheds light on machine translation by analyzing assessment techniques and addressing issues, especially when translating low-resource languages. Consequently, it contributes to the present state of machine translation review. The use of machine translation (MT) systems has recently increased, which has been advantageous for improving communication and information access by automatically translating text between languages. Studies have additionally shown the usefulness of machine translation in learning a language. [63]-[64]. There are still issues that need to be resolved in order to improve the accuracy and performance of MT systems. Managing the absence of parallel corpora, which is necessary for these systems' testing and training, is one of the major challenges. It should be noted that developing alternative corpora might be challenging because not all languages may have access to them.

Machine translation has difficulties when handling ambiguities. There are many different ways to interpret words and sentences. Determining the correct meaning relies heavily on the context in which they are used. Despite these challenges, machine translation has seen advancements. The introduction of approaches, like Neural machine translation, has greatly enhanced its accuracy. There are three techniques under Neural Machine Translation (NMT), Recursive NMT, Recurrent NMT, and Convolution NM [65]. However, more work must be done to improve MT systems' performance, such as Quantum neural networks, a new type of neural network being developed for MT. Quantum neural networks have the potential to achieve even higher levels of accuracy than traditional neural networks.

V. CONCLUSION

In this paper, we discussed Machine Translation approaches, evaluation metrics, and Challenges faced in Machine Translation on low-resource languages. Progress has been made in machine translation up to this point for languages with few resources. As per this conducted review, our findings are consistent with the existing knowledge that the quality of machine translation systems is directly proportional to the corpus size and translation approach. It is also found that the two most widely used methods are statistical and neural machine translation. However, a huge amount of data is required for these machine translation systems, which is unavailable for many languages that are less resourced and addressed in this study.

Moreover, it is apparent from the survey that further work is needed in MT as a whole to produce intelligible translations to improve translation accuracy. While numerous evaluation metrics exist for assessing the quality of machine translation (MT) outputs, our review shows that BLEU and Human Evaluation Metrics stand out as two dominant methodologies. Different approaches have their advantages and play a role in evaluating the quality of machine translation (MT) systems. In this study, we focused on comparing the effectiveness of BLUE score and human evaluation metrics using twenty published studies over five years. The findings reveal that the BLEU score has proven more reliable than human evaluation methods over the past five years.

Moreover, our study also suggests that the BLEU score offers a dependable way of assessing the quality of translations when compared to Human evaluation metrics. Machine translation encounters challenges that impact its effectiveness. These challenges encompass scarcity of data, linguistic complexities, and limited resources. The efficient approach to tackle these language challenges, neural machine translation (NMT), uses learning techniques to mitigate the challenges posed by data scarcity. However, NMT can translate languages with limited resources and the specific language being translated. As research in machine

PJETS
Pakistan Journal of Engineering Technology and Science (PJETS)
Volume 11. Issue. 1, PP. 29-44, November 2023
E-ISSN 2224-2333
https://doi.org/10.22555/pjets.v11i1.1002
P-ISSN 2222-9930

translation progresses, it becomes crucial to address these challenges to fully harness the potential of NMTs to enhance accessibility and overcome obstacles faced by low-resource languages.

This review is useful for researchers working with low-resource languages and investigating evaluation matrices, challenges, and pitfalls leading towards developing machine translation systems for the low-resource languages.

REFERENCES

[1] X. Deng and Z. Yu, "A Systematic Review of Machine-Translation-Assisted Language Learning for Sustainable Education," *Sustainability (Switzerland)*, vol. 14, no. 13. MDPI, Jul. 01, 2022. doi: 10.3390/su14137598.

[2] D. Padhya and J. Sheth, "A Review of Machine Translation Systems for Indian Languages and Their Approaches," *Advances in Intelligent Systems and Computing*, vol. 841, pp. 103–110, 2019, doi: 10.1007/978-981-13-2285-3_13.

[3] D. Chopra, N. Joshi, and I. Mathur, "A Review on Machine Translation in Indian Languages," *Engineering, Technology & Applied Science Research*, vol. 8, no. 5, pp. 3475–3478, 2018, doi: 10.48084/etasr.2288.

[4] A. Venugopal, M. Madanan, and T. Kadarkarai, "Analysis of Fusion of Machine Learning Tools in Education," *Fusion: Practice and Applications*, vol. 12, no. 2, pp. 88–97, 2023, doi: 10.54216/FPA.120207.

[5] S. Khepra, "A Survey of Punjabi Language Translation using OCR and ML," pp. 136–144.

[6] J. Mirzakhalov *et al.*, "A Large-Scale Study of Machine Translation in the Turkic Languages," *EMNLP 2021 - 2021 Conference on Empirical Methods in Natural Language Processing, Proceedings*, pp. 5876–5890, 2021, doi: 10.18653/v1/2021.emnlp-main.475.

[7] B. Ji, Z. Zhang, X. Duan, M. Zhang, B. Chen, and W. Luo, "Cross-lingual pre-training based transfer for zero-shot neural machine translation," *AAAI 2020 - 34th AAAI Conference on Artificial Intelligence*, pp. 115–122, 2020, doi: 10.1609/aaai.v34i01.5341.

[8] J. J. Zhang and C. Q. Zong, "Neural machine translation: Challenges, progress and future," *Sci China Technol Sci*, vol. 63, no. 10, pp. 2028–2050, 2020, doi: 10.1007/s11431-020-1632-x.

[9] J. Libovický and A. Fraser, "Findings of the WMT 2021 Shared Tasks in Unsupervised MT and Very Low Resource Supervised MT," *WMT 2021 - 6th Conference on Machine Translation, Proceedings*, pp. 726–732, 2021.

[10] V. Ambati, "A Hybrid Approach to Example based Machine Translation for Indian Languages," no. January, 2006.

[11] A. K. Barman, J. Sarmah, and S. K. Sarma, "Assamese wordNet based quality enhancement of bilingual machine translation system," *GWC 2014: Proceedings of the 7th Global Wordnet Conference*, pp. 256–261, 2014.

[12] M. Z. Islam, J. Tiedemann, and A. Eisele, "English to Bangla phrase-based Machine Translation," *EAMT 2010 - 14th Annual Conference of the European Association for Machine Translation*, no. May, 2010.

[13] K. Mishra, A. Soni, R. Sharma, and D. Sharma, "Exploring the effects of Sentence Simplification on Hindi to English Machine Translation System," pp. 21–29, 2015, doi: 10.3115/v1/w14-5603.

[14] J. Zhu *et al.*, "Incorporating BERT into Neural Machine Translation," pp. 1–18, 2020.

[15] Y. Liu *et al.*, "RoBERTa: A Robustly Optimized BERT Pretraining Approach," no. 1, 2019.

[16] S. Lee *et al.*, "A Survey on Evaluation Metrics for Machine Translation," *Mathematics*, vol. 11, no. 4, pp. 1–22, 2023, doi: 10.3390/math11041006.

[17] J. S. White, T. A. O'Connell, and L. M. Carlson, "Evaluation of machine translation," no. Bamberg, p. 206, 1993, doi: 10.3115/1075671.1075717.

[18] G. Wentzel, "Funkenlinien im Röntgenspektrum," *Ann Phys*, vol. 371, no. 23, pp. 437–461, 1922, doi: 10.1002/andp.19223712302.

[19] M. Hämäläinen and K. Alnajjar, "The Great Misalignment Problem in Human Evaluation of NLP Methods," *Human Evaluation of NLP Systems, HumEval 2021 - Proceedings of the Workshop, as part of the 16th Conference of the European Chapter of the Association for Computational Linguistics, EACL 2021*, pp. 69–74, 2021.

[20] L. Bowker, "Chinese speakers' use of machine translation as an aid for scholarly writing in English: a review of the literature and a report on a pilot workshop on machine translation literacy," *Asia Pacific Translation and Intercultural Studies*, vol. 7, no. 3, pp. 288–298, 2020, doi: 10.1080/23306343.2020.1805843.

[21] S. Saini and V. Sahula, "A survey of machine translation techniques and systems for Indian languages," *Proceedings - 2015 IEEE International Conference on Computational Intelligence and Communication Technology, CICT 2015*, no. Census 2001, pp. 676–681, 2015, doi: 10.1109/CICT.2015.123.

[22] H. Kim, H. Yang, and J. Ho Lee, "Design principles and architecture of a second language learning chatbot." [Online]. Available: http://hdl.handle.net/10125/73463

[23] V. Ambati, "A Hybrid Approach to Example based Machine Translation for Indian Languages," no. January, 2006.

[24] M. Saffari, M. Pourhaji, S. Fathi-Alishah, S. Sajjadi, and M. Mohammadi, "Translating Medical Texts from Persian to English: Accuracy of Machine Translation. Archives of Advances in Biosciences," *AAB*, vol. 14, 2023, doi: 10.22037/aab.v14i.

[25] I. Jibreel, "Online Machine Translation Efficiency in Translating Fixed Expressions Between English and Arabic (Proverbs as a Case-in-Point)," *Theory and Practice in Language Studies*, vol. 13, no. 5, pp. 1148–1158, May 2023, doi: 10.17507/tpls.1305.07.

[26] A. Ramanathan, P. Bhattacharyya, J. Hegde, R. M. Shah, and M. Sasikumar, "Simple Syntactic and Morphological Processing Can Help English-Hindi Statistical Machine Translation," *IJCNLP 2008 - 3rd International Joint Conference on Natural Language Processing, Proceedings of the Conference*, vol. 1, no. i, pp. 513–520, 2008.

[27] T. Rama and K. Gali, "Modeling machine transliteration as a phrase based statistical machine translation problem," *NEWS 2009 - 2009 Named Entities Workshop: Shared Task on Transliteration at the Joint Conference of the 47th Annual Meeting of the ACL and the 4th International Joint Conference on Natural Language Processing of the AFNLP, ACL-IJCNLP 2009*, no. August, pp. 124–127, 2009, doi: 10.3115/1699705.1699737.

[28] A. Ali, A. Hussain, and M. Kamran Malik, "Model for English-Urdu statistical machine translation," *World Appl Sci J*, vol. 24, no. 10, pp. 1362–1367, 2013, doi: 10.5829/idosi.wasj.2013.24.10.760.

[29] B. N. V. N. Raju and M. S. V. S. B. Raju, "Statistical Machine Translation System for Indian Languages," *Proceedings - 6th International Advanced Computing Conference, IACC 2016*, pp. 174–177, 2016, doi: 10.1109/IACC.2016.41.

[30] M. Singh, R. Kumar, and I. Chana, "Corpus based Machine Translation System with Deep Neural Network for Sanskrit to Hindi Translation," *Procedia Comput Sci*, vol. 167, no. 2019, pp. 2534–2544, 2020, doi: 10.1016/j.procs.2020.03.306.
[31] I. A. Khan, A. Khan, B. Nazir, S. S. Hussain, F. G. Khan, and I. A. Khan, "Urdu Translation: the Validation and Reliability of the 120-Item Big Five IPIP Personality Scale," *Current Psychology*, vol. 38, no. 6, pp. 1530–1541, 2019, doi: 10.1007/s12144-017-9706-5.
[32] R. M. K. Sinha and A. Jain, "AnglaHindi: an English to Hindi machine-aided translation system," *MT Summit IX, New Orleans, Louisiana, USA*, pp. 1–5, 2003.
[33] M. K. Chinnakotla, S. Ranadive, P. Bhattacharyya, and O. P. Damani, "Hindi and Marathi to English cross language information retrieval at CLEF 2007," *CEUR Workshop Proc*, vol. 1173, pp. 111–118, 2007.
[34] P. C, D. V, A. Kumar M, and S. K P, "Rule based Sentence Simplification for English to Tamil Machine Translation System," *Int J Comput Appl*, vol. 25, no. 8, pp. 38–42, 2011, doi: 10.5120/3050-4147.
[35] M. Nandathilaka, S. Ahangama, and G. T. Weerasuriya, "Language," *2018 3rd International Conference on Information Technology Research (ICITR)*, pp. 1–5, 2018.
[36] A. Kumar and A. K. Singh, "Transformer-based neural machine translation system for Tamil - English," *WAT@EMNLP-IJCNLP 2019 - 6th Workshop on Asian Translation, Proceedings*, pp. 171–174, 2021.
[37] A. Gehlot, V. Sharma, S. P. Singh, and A. Kumar, "Hindi to English Transfer Based Machine Translation System," vol. 19, no. 19, 2015.
[38] J. A. Todase and S. Shelke, "Script Translation System for Devnagari to English," *2018 International Conference on Information, Communication, Engineering and Technology, ICICET 2018*, pp. 1–4, 2018, doi: 10.1109/ICICET.2018.8533734.
[39] H. Choudhary, A. K. Pathak, R. R. Shah, and P. Kumaraguru, "Neural Machine Translation for English-Tamil," *WMT 2018 - 3rd Conference on Machine Translation, Proceedings of the Conference*, vol. 2, pp. 770–775, 2018, doi: 10.18653/v1/W18-64086.
[40] P. K. Nagaraj, K. S. Ravikumar, M. S. Kasyap, M. H. S. Murthy, and J. Paul, "Kannada to english machine translation using deep neural network," *Ingenierie des Systemes d'Information*, vol. 26, no. 1, pp. 123–127, 2021, doi: 10.18280/isi.260113.
[41] S. Gogineni, G. Suryanarayana, and S. K. Surendran, "English to Hindi Language," no. Icosec, pp. 209–214, 2020.
[42] R. Narayan, S. Chakraverty, and V. P. Singh, "Quantum neural network based machine translator for English to Hindi," *Applied Soft Computing Journal*, vol. 38, pp. 1060–1075, 2016, doi: 10.1016/j.asoc.2015.08.031.
[43] S. P. Singh, A. Kumar, H. Darbari, L. Singh, A. Rastogi, and S. Jain, "Machine translation using deep learning: An overview," *2017 International Conference on Computer, Communications and Electronics, COMPTELIX 2017*, pp. 162–167, 2017, doi: 10.1109/COMPTELIX.2017.8003957.
[44] S. Castilho and N. Resende, "Post-Editese in Literary Translations," *Information (Switzerland)*, vol. 13, no. 2, Feb. 2022, doi: 10.3390/info13020066.
[45] Y. Cheng, *Joint Training for Pivot-Based Neural Machine Translation*. 2019. doi: 10.1007/978-981-32-9748-7_4.
[46] M. X. Chen, G. Foster, and N. Parmar, "Combining Recent Advances in Neural Machine Translation," 2014.
[47] D. Bahdanau, K. H. Cho, and Y. Bengio, "Neural machine translation by jointly learning to align and translate," *3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings*, pp. 1–15, 2015.
[48] K. Cho *et al.*, "Learning phrase representations using RNN encoder-decoder for statistical machine translation," *EMNLP 2014 - 2014 Conference on Empirical Methods in Natural Language Processing, Proceedings of the Conference*, pp. 1724–1734, 2014, doi: 10.3115/v1/d14-1179.
[49] Q. Wang *et al.*, "Learning Deep Transformer Models for Machine Translation," 2018.
[50] B. Zhang, I. Titov, and R. Sennrich, "Improving Deep Transformer with Depth-Scaled Initialization and Merged Attention," 2019.
[51] Y. Li, Q. Wang, T. Xiao, T. Liu, and J. Zhu, "Neural Machine Translation with Joint Representation".
[52] C.-Y. Lin and F. J. Och, "ORANGE: a Method for Evaluating Automatic Evaluation Metrics for Machine Translation."
[53] N. Mathur, T. Baldwin, and T. Cohn, "Tangled up in BLEU: Reevaluating the Evaluation of Automatic Machine Translation Evaluation Metrics," Jun. 2020, [Online]. Available: http://arxiv.org/abs/2006.06264
[54] C. Liu, D. Dahlmeier, and H. T. Ng, "Better Evaluation Metrics Lead to Better Machine Translation."
[55] B. Marie, A. Fujita, and R. Rubino, "Scientific Credibility of Machine Translation Research: A Meta-Evaluation of 769 Papers," Jun. 2021, [Online]. Available: http://arxiv.org/abs/2106.15195
[56] E. Chatzikoumi, "How to evaluate machine translation: A review of automated and human metrics," *Nat Lang Eng*, vol. 26, no. 2, pp. 137–161, Mar. 2020, doi: 10.1017/S1351324919000469.
[57] A. Gravano, "Turn-taking and affirmative cue words in task-oriented dialogue," *Dissertation Abstracts International, B: Sciences and Engineering*, vol. 70, no. 8, p. 4943, 2010, doi: 10.1162/COLI.
[58] B. Marie, A. Fujita, and R. Rubino, "Scientific credibility of machine translation research: A meta-evaluation of 769 papers," *ACL-IJCNLP 2021 - 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing, Proceedings of the Conference*, pp. 7297–7306, 2021, doi: 10.18653/v1/2021.acl-long.566.
[59] H. Hassan *et al.*, "Achieving Human Parity on Automatic Chinese to English News Translation," 2018.
[60] B. R. G. Noll, "S Ponsored H Uman E Mbryonic," vol. 2, no. 1985, pp. 21–36, 2006.
[61] Chris Callison-Burch Miles Osborne Philipp Koehn, "Callison-Burch, Ch., Osborne, M., & Koehn, Ph. (2006). Re-evaluating the role of bleu in machine translation. In Proceedings the 11th Conference of the European Chapter of the Association for Computational Linguistics (pp. 249-256). Trento, Italy.," *Proceedings the 11th Conference of the European Chapter of the Association for Computational Linguistics*, p. (pp. 249-256), 2006.
[62] L. Barrault *et al.*, "Findings of the 2020 Conference on Machine Translation (WMT20)," *5th Conference on Machine Translation, WMT 2020 - Proceedings*, pp. 1–55, 2020, doi: 10.18653/v1/w19-5301.
[63] L. W. Rowe, "Google Translate and Biliterate Composing: Second-Graders' Use of Digital Translation Tools to Support Bilingual Writing," *TESOL Quarterly*, vol. 56, no. 3, pp. 883–906, Sep. 2022, doi: 10.1002/tesq.3143.
[64] N. Bin Dahmash, "I Can't Live Without Google Translate: A Close Look at the Use of Google Translate App by Second Language Learners in Saudi Arabia," *Arab World English Journal*, vol. 11, no. 3, pp. 226–240, Sep. 2020, doi: 10.24093/awej/vol11no3.14.
[65] T. P. Nagarhalli, V. Vaze, and N. K. Rana, "A Novel Framework for Neural Machine Translation of Indian-English Languages," *Proceedings of the 5th International Conference on Inventive Computation Technologies, ICICT 2020*, pp. 676–682, 2020, doi: 10.1109/ICICT48043.2020.9112513.