ISSN (E): 2959-3921 *Ethiopian International Journal of Engineering and Technology (EIJET)*
(P): 2959-393X *Volume 3, Issue 2, 2025*
***DOI:** https://doi.org/10.59122/184DFD14*

Kafi Noonoo to English Machine Translation using Deep Learning Approaches

Ashagire Adinew[1]*, Micahel Melese[2]

[1]Bonga University, College of Engineering and Technology, Ethiopia

[2]Addis Ababa University, College of Natural and Computational Science, Ethiopia

*Corresponding Author's Email: ashagire@bongau.edu.et

Abstract

Kafi Noonoo is one of the Ethiopian languages that is spoken by the Kaffa people in the southwestern part of Ethiopia. Additionally, it is a morphologically rich language and has an indigenous name for prestige, cultural place, and cultural dejectedness, which has no equivalent meaning in other languages. Machine translation is a technique that automatically translates text or speech's meaning from one language to another without human involvement to resolve information gaps. Various machine translation studies have been conducted for resource-rich languages like English, French, German, and others. However, the variety of linguistic patterns, the dominance of technologically developed languages, and the lack of machine translation from Kafi Noonoo to English will lead to the disappearance of Kafi Noonoo's indigenous words among native speakers. To tackle such a problem, this article designed a Kafi Noonoo to English and vice versa machine translation solution by using deep learning approaches. The bidirectional long short-term memory, bidirectional gated recurrent unit with and without attention, and transformer were applied. In order to train the model, the bilingual parallel sentences were collected from Kafi Noonoo's linguistic-related sources. Different experiments were applied to find out the optimal value of the proposed model. Based on the experiment's result, the transformer performed better with an accuracy of 89% and a BLEU score of 6.34 and 5.42 for Kafi Noonoo to English and English to Kafi Noonoo, respectively. According to our experiment results, the transformer model was suitable for morphologically rich languages like Kafi Noonoo to English and vice versa, for machine translation. For a better result, there is a necessity to generate parallel corpora in order to conduct comparable research.

Keywords: Kaffa, Kafi Noonoo, Low-resource Machine Translation, Transformer

I. Introduction

Natural language is one of the fundamental aspects of human behavior and a basic element of daily activities, and it is mainly used for exchanging thoughts, feelings, and information through spoken, written, and signed communication [1]. According to [2], distinct natural languages are spoken all over the world.

***Received**: May 28, 2025; **Revised**: June 10, 2025; **Accepted:** July 18, 2025; **Published:** August 2025*
*Corresponding author- **Ashagire Adinew***

ISSN (E): 2959-3921 Ethiopian International Journal of Engineering and Technology (EIJET)
(P): 2959-393X Volume 3, Issue 2, 2025
DOI: https://doi.org/10.59122/184DFD14

Kafi noonoo is one component that is spoken by the Kaffa people (Kafecho) in the southwestern part of Ethiopia. Natural languages are studied in different academic disciplines such as linguistics, psycholinguistics, computational linguistics, and others. Furthermore, each discipline defines a set of problems and has methods to address the identified problems [3]. Kafi Noonoo uses the Latin script writing system, which utilizes a total of thirty-two characters made up of twenty-seven consonant characters and five vowel characters [4][5][6][7]. In morphology, it uses inflectional, derivational, compounding, reduplication, and cases to form another word from existing words and follows the subject-object-verb order to form a grammatically correct sentence structure [8][5][9].

Computational linguistics studies the formal rules of languages and how technological devices such as computers, smartphones. and robots understand and generate natural language. Moreover, it mainly focuses on the computational description of languages as a system, and is applied by natural language processing [10][11][12][13]. Natural language processing (NLP) is a practical and interdisciplinary study domain of linguistics, mathematics, and artificial intelligence (AI) studied under computer science to process natural languages on computational devices [12][14][15]. Accordingly, different natural language processing applications such as machine translation [15][16][17], grammar error detection [18], word sense disambiguation [19], fake account detection [20], sentimental analysis [21], and others were developed.

Machine translation (MT) is utilized to convert text or speech from one human language to another, maintaining the natural language context using a computer algorithm. Using MT, anyone can share his/her knowledge, culture, tradition, history, and religious and philosophical writings from one language to another, and the accessibility of documents written in one human language to another language is easily translated [1][17]. Accordingly, to design and develop a machine translation system, rule-based (involving direct, transfer, and interlingua) translations, statistical translations, example-based translations, deep neural network (DNN) translations, and hybrid approaches were formulated [22]. In the field of MT, DNN is a newly emerging approach that has proven to achieve excellent performance. Its translation is enabled by huge amounts of structured data; the translations will be processed with a much higher accuracy rate, more scalable than traditional approaches, and identify patterns at a deeper level that can capture better way [23]. Various machine translation researches were done for Amharic-Wolaytta [24], Amharic-Ge'ez [25], English to Wolaytta [26], English to French [27], English to German [28] by understanding each language's limitations and linguistic patterns [1]. Furthermore, as our world becomes increasingly connected, language translation provides a critical cultural, political, and economic bridge between people from different countries and ethnic groups.

Received: May 28, 2025; **Revised**: June 10, 2025; **Accepted:** July 18, 2025; **Published:** August 2025
Corresponding author- **Ashagire Adinew**

ISSN (E): 2959-3921 Ethiopian International Journal of Engineering and Technology (EIJET)
(P): 2959-393X Volume 3, Issue 2, 2025
DOI: https://doi.org/10.59122/184DFD14

However, technologically developed languages' word dominance, incompatibility of existing algorithms, and the lack of computational linguistic applications like Kafi noonoo to English machine translation will lead indigenous Kafi noonoo words like prestige names (*kateemeraashoo shooderaashoo, maaceraashoo*), cultural place names (moogo, gutoo), and cultural dejectedness names (*gommo, shoosho, hichoo*) to disappear from native speakers. To overcome this problem, this article was designed for Kafi Noonoo for English machine translation (MT) by using deep learning approaches.

II. Literature Review

Mengistu and Kinfe [29] investigated the deep learning technique in bi-directional Amharic to Kistanigna machine translation. Parallel corpora from the Kistanigna dictionary to Amharic and the Holy Bible, like Saint Matiwos, Saint Marikos, and Saint Lukas, were gathered by the researcher. The researcher used a variety of deep learning algorithms, including transformer models, bidirectional long short-term memory (Bi-LSTM), long short-term memory (LSTM) with attention, and long short-term memory LSTM. Ultimately, the Transformer model produced BLEU scores of 22.4 from Kistanign to Amharic and 21.3 from Amharic to Kistanigna.

Agerie Belete [30] carried out deep learning-based bi-directional English-Awngi machine translation. According to the paper, no research was done comparing Awngi and English. The researcher compiled parallel corpora from social media documents, religious texts, and educational resources and obtained BLEU scores of 22.34 for Awngi to English and 24.94 for English to Awngi, respectively.

Amdework Asefa [25] performed a Bidirectional neural machine translation (MT) for Ge'ez and Amharic using deep learning techniques. The goal of the researcher is to demonstrate the deep learning models' proficiency in MT tasks for those morphologically complex languages. Utilizing the praise of Saint Mary, the Mass media, and other Ethiopian Orthodox Church religious texts and the Bible, the researchers created a dataset. Using a transformer, the researcher assessed the results of an experiment and obtained BLEU scores of 22.9 for Ge'ez to Amharic and 29.7 for Amharic to Ge'ez.

Abdu Wahid [31] carried out their research on the deep learning-based MT system for English to Urdu. The dataset was gathered from news articles and everyday utterances included in the English-Urdu parallel corpus. They created a new parallel corpus and experimented with various training settings for the English to Urdu MT model, which is an LSTM encoder-decoder.

Levi Corallo et al. [32] used a gated recurrent unit recurrent neural network (RNN) to perform German-English machine translation. The researchers' framework aims to construct apps, facilitate future work in

Received: May 28, 2025; **Revised**: June 10, 2025; **Accepted:** July 18, 2025; **Published:** August 2025
Corresponding author- **Ashagire Adinew**

the field, and act as a pilot method for translating strings from German news sources into English phrases. They obtain statistics from the WMT2021 and create a framework that might be helpful in creating mobile applications for rapid translation, where efficiency is essential. By classifying RNN models with various hyperparameters, they finished two tests. The final training and validation accuracy of Model I were 0.655 and 0.653, respectively, while the final validation accuracy of Model II was 0.645 and 0.649. Model I's final validation loss is 2.78, while its final training loss is 2.78 (as depicted in Fig. 2).

Elias Asefa [33] used the neural network approach in conjunction with RNN in order to create a unidirectional English-to-Dawurootsuwa MT model. He finished creating the algorithm to identify and examine patterns in data sequences and forecast the output text using the text data that was entered. Additionally, a total of 20,345 pairs of corpora that were gathered from various sources were able to acquire a BLEU score of 0.5187.

Mekdes Melese [26] worked on English-Wolaita's attention-based neural machine translation (NMT). The researcher's goal is to use attention mechanisms to create an NMT for English-Wolaita and gathered 27351 sentences from a comparable corpus. Additionally, the data was preprocessed for appropriate NMT usage. An LSTM encoder and LSTM decoder architecture with an attention mechanism has been used in the Seq-to-Seq idea to construct the English-Wolaita NMT system model. The researcher attempted to compare the model with the non-attention model and the attention mechanism in order to gauge performance by the BLEU score for measuring the effectiveness of the attention mechanism. Overall, the researcher was assessing the effectiveness of the model investigations. Finally, it was acknowledged that the attention mechanism translated more accurately, achieving a BLEU score of 5.16 and an accuracy of 88.65% for English-Wolaytta. According to literature reviews, there is no study that has been done on MT between Kafi Noonoo and English. This article designed a Kafi Noonoo to English machine translation using deep learning approaches.

III. Research Methodology

A. *Data Source*

This study requires a dataset to train and evaluate the model for translating Kafi Noonoo texts into English and vice versa. Nevertheless, there is no published dataset in bilingual content. Therefore, in order to gather the bilingual text, this study planned, identified, and gathered bilingual text from authorized Kafi Noonoo linguistics-related data sources. We have collected 36,000 sentences in two columns, one for Kafi Noonoo and one for English.

***Received**: May 28, 2025; **Revised**: June 10, 2025; **Accepted:** July 18, 2025; **Published:** August 2025*
*Corresponding author- **Ashagire Adinew***

ISSN (E): 2959-3921 Ethiopian International Journal of Engineering and Technology (EIJET)
(P): 2959-393X Volume 3, Issue 2, 2025
DOI: https://doi.org/10.59122/184DFD14

As mentioned in Table I, the corpus we have collected is from various sources such as academic documents, the holy bible, and the existing Kafi Noonoo dictionary in order to meet the objective of creating an acceptable dataset for bilingual text from Kafi Noonoo to English.

Table I: Collected dataset

Bilingual sources	Amount
Academic document	7,000
Holy bible	10,000
Kafi Noonoo to the English dictionary	19,000
Total	36,000

B. Model Architecture

The first phase, which is carried out by the Kafi Noonoo to English translation MT architecture (as mentioned in Fig. 1), is preprocessing. Under this phase, different activities, such as data cleaning, which is used mainly to remove irrelevant symbols from the collected dataset, are done. The normalization techniques were applied to standardize all symbols in a unique form to improve dataset quality. Tokenization is the process of converting plain text into a series of tokens. Two main procedures are performed sequentially in that process. We used the tokenized words to create a word index, which is mainly focused on word identification in the dataset. The word sequence to numeric sequences conversion was also used to convert sequences of sentences into sequences of numeric values because deep learning algorithms do not understand raw text. The sequence normalization techniques were applied to normalize the sequence in matrix form. Next to this step, we designed an encoder and decoder that are mainly used to translate sentences. The encoder-decoder model has two parts. The first one accepts vectorized form source language sentences, i.e., Kafi noonoo, with a deep learning algorithm, bidirectional long short memory (BiLSTM), bidirectional gated recurrent (BiGRU) with and without attention, and transformer with fixed input sequence lengths, targeted language sequences, i.e., English, with softmax activation function. Encoder processes input sequences token by token by updating its hidden state at each time step and finally produces a model. The second decoder takes the encoder vector and generates a target sequence one at a time. The decoder performs a time step each time, as input, previous target sequences generated tokens from the decoder, as well as the previous hidden state. The hidden state is updated depending on the previous state and previous target token, and used to generate the next target token. This process continues until the decoder brings about an end-of-sequence token, indicating that the decoder has generated the entire target sequence.

***Received**: May 28, 2025; **Revised**: June 10, 2025; **Accepted:** July 18, 2025; **Published:** August 2025*
*Corresponding author- **Ashagire Adinew***

ISSN (E): 2959-3921
(P): 2959-393X
DOI: *https://doi.org/10.59122/184DFD14*
Ethiopian International Journal of Engineering and Technology (EIJET)
Volume 3, Issue 2, 2025

The encoding and decoding processes are trained jointly using a Seq2Seq loss function, optimizer, and learning rate epochs. During training, the input sequence is supplied to the encoder, and the decoder generates the target sequence token by token. The final encoder hidden layer context vector and target embedding sequence are accepted by the decoder model during training in order to predict the target output. Thus, the context vector is passed through to the decoder to get an output sequence.

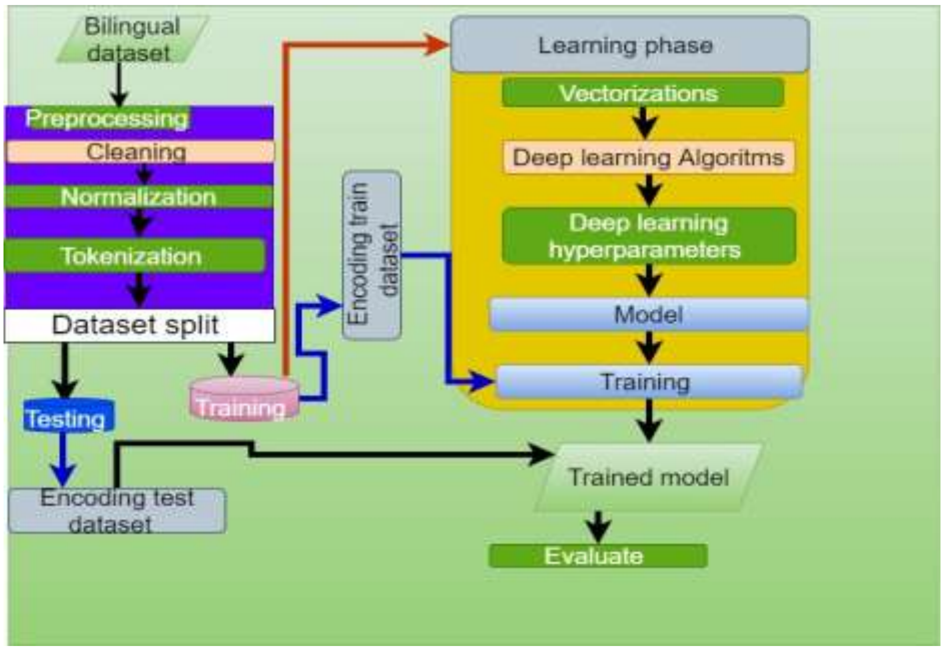

Fig. 1: Model Architecture

C. Model Performance Evaluation Metrics

1) Accuracy

The accuracy metric is used to monitor the model's performance during training. It indicates how successfully the model can categorize or predict the right output given the input data. We also used this measurement of accuracy to assess the quality of the translation in the translation process. It is computed computationally by adding the errors for each sample throughout the training set. Each batch of data is used to calculate training loss, which is then plotted as a curve [34].

2) BLEU (Bilingual Evaluation Understudy)

BLEU is a reference-based metric that measures the similarity between the system-generated translations to evaluate the quality of our translations and human-generated reference translations. It calculates common elements of n-grams, continuous sequences of matched words between the system output and the reference

***Received**: May 28, 2025; **Revised**: June 10, 2025; **Accepted:** July 18, 2025; **Published:** August 2025*
*Corresponding author- **Ashagire Adinew***

ISSN (E): 2959-3921 Ethiopian International Journal of Engineering and Technology (EIJET)
(P): 2959-393X Volume 3, Issue 2, 2025
DOI: https://doi.org/10.59122/184DFD14

translations. The core idea behind the BLEU score is that a good translation should contain similar n-grams to the human reference translations. Performance evaluation is important after training models; there are various methods used in MT translation accuracy. BLEU computes the precision of n-grams, including unigrams, bigrams, trigrams, etc. The effectiveness of MT is determined by assessing how closely it aligns with human translation. When the machine-generated text closely resembles the human translation, it is considered to be of higher quality and more successful in achieving accurate translation [35].

IV. Results and Discussion

A. Results

The proposed deep learning-based Kafi Noonoo to English machine translation model has been aimed to translate Kafi Noonoo to English and vice versa. Finally, the artefact was designed and developed by using BiLSTM and BiGRU with attention and without attention, and transformer deep learning algorithms. In addition, different experiments were conducted by the identified deep learning algorithms to determine the effect of deep learning approaches to translate Kafi Noonoo to English and vice versa, and evaluated each experiment's results by using accuracy and BLEU (as depicted in Fig. 3). In the first experiments, we got BLEU scores of 5.18 and 4.37 from the Kafi Noonoo to English translation and English to Kafi Noonoo, respectively, by using BiGRU.

In the second trial, we added attention mechanisms to BiGRU, and the model's training time and the effectiveness of the model with the same data size and hyperparameters were improved. The training times in the analysis are compared with the BiGRU model training time, with attention to the training time without attention in the prior experiment. An attention layer has been allowed in the model for capturing the context and ignoring the ignore word. For this reason, encoder-decoders without attention mechanisms cannot handle large numbers of sentences. The other outcome of attention methods can reduce the amount of training time needed for the model to finish. The effectiveness of the BiGRU model with attention includes an accuracy of 86%, a loss of 0.11 from Kafi Noonoo to English, and the impact of attention on the metrics compared to previous experiments. We see better accuracy, loss, and training results compared to previous experiments. In this phase, the BLEU score from Kafi Noonoo to English has taken 404 seconds, and from English to Kafi Noonoo, it has taken 420 seconds, and the BLEU score has obtained 5.82 and 4.74, respectively.

Our third experiment was done using bidirectional long short-term memory. It contains double long and short-term memory (LSTM) cells on the encoder side, which take up a lot more memory than a single LSTM. Here we have used the same parameters as we did with BiGRU and BiGRU with attention. We

Received: May 28, 2025; *Revised*: June 10, 2025; *Accepted:* July 18, 2025; *Published:* August 2025
*Corresponding author- **Ashagire Adinew***

obtained accuracy results of 87.89% and 88.5% and loss values of 0.39 and 0.12 from Kafi Noonoo to English and English to Kafi Noonoo. It took 365 seconds and 378 seconds to get BLEU 5.89 and 4.83 from Kafi Noonoo to English and English to Kafi Noonoo, respectively.

In the fourth trial, we conducted the Bidirectional long short-term memory with attention, and we used similar parameters to those we used in the others. In the BiLSTM with attention model, we analyzed the result of varying the number of epochs on the model and explored how increasing the number of epochs influences the model's ability to learn from the training dataset, improve its performance, and decrease the value of loss. It took 265 seconds and 300 seconds to get BLEU 5.99 and 4.93 from Kafi Noonoo to English and English to Kafi Noonoo, respectively.

Our transformer model does not have an additional attention mechanism. Here we have fed both token embedding and position embedding vectors to our transformer model. Because the transformer model accepts all inputs in parallel. We have used 6 layers in our transformer model. Like the BiGRU and BiLSTM without attention and with attention, our transformer models use only the last encoder context vector passed to the decoder directly without additional attention mechanisms, because the transformer model has its attention. The encoder-transformer model has a feed-forward network and self-multi-head attention. The decoder model has a masked multi-head focused and a feed-forward neural network (NN). And we have used the RLU feed-forward network in both the encoder and decoder models. It took 100 seconds and 120 seconds to get BLEU 6.34 and 5.42 from Kafi Noonoo to English and English to Kafi Noonoo, respectively.

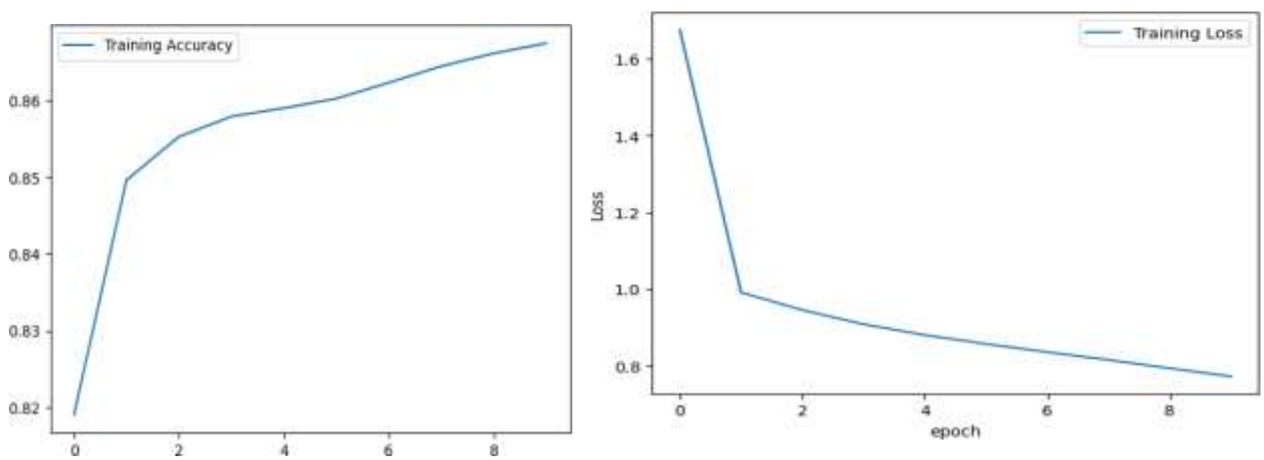

Fig. 2: Training accuracy and loss graph

***Received**: May 28, 2025; **Revised**: June 10, 2025; **Accepted:** July 18, 2025; **Published:** August 2025*
*Corresponding author- **Ashagire Adinew***

ISSN (E): 2959-3921
(P): 2959-393X
***DOI:** https://doi.org/10.59122/184DFD14*
Ethiopian International Journal of Engineering and Technology (EIJET)
Volume 3, Issue 2, 2025

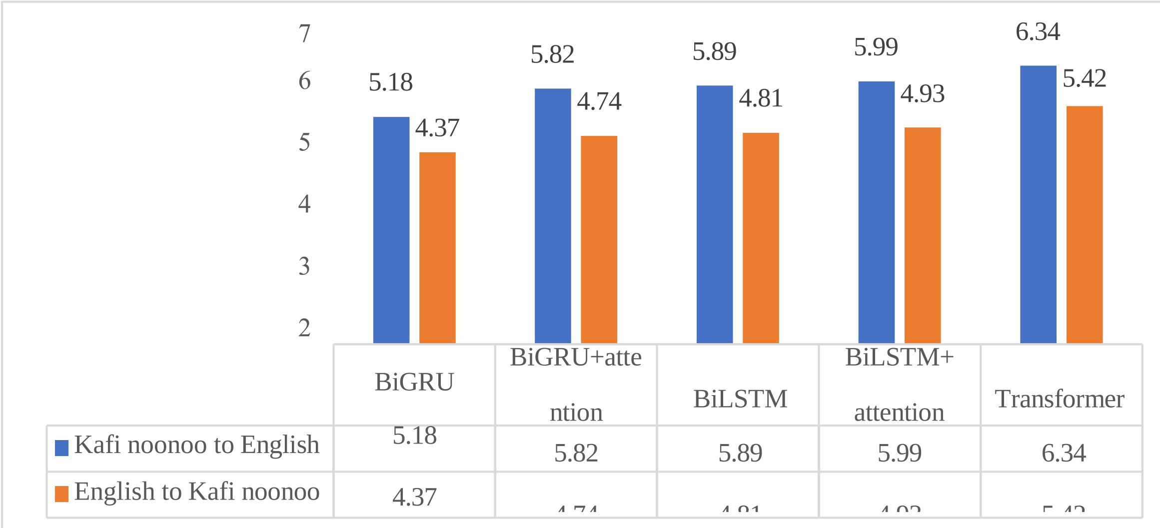

Fig. 3: Comparison of the deep learning algorithm's BLEU score

B. Discussion

To translate Kafi Noonoo to English and vice versa, different preprocessing steps were done. The encoder-decoder architecture was built with attention or without attention using BiLSTM, BiGRU, and transformers. The encoder reads all of the data into a single real-valued vector, which is then sent to the decoder, which uses the vector to produce the target translation. To identify a word by a specific index at the moment it is encountered in the data, each word in the sentences must be given a new identity number as it is visited. Different experimentation setups with different deep learning algorithm hyperparameters were designed to find out the optimal result measured by accuracy and BLEU score. We obtained different results in each experiment. Based on the result of the experiment on BiLSM and BiGRU with attention, the BLEU score shows that the attention-based approach is better than the one without attention for Kafi Noonoo-English MT models, and a better outcome is obtained when the transformer model is used instead of BiLSTM and BiGRU with and without attention.

Nevertheless, it is difficult to obtain a more effective translation model, given the short amount of data employed in this article. Even with a small dataset, the highest possible BLEU score of 6.34 and accuracy of 89% is recorded by adjusting different hyperparameters while experimenting with this small corpus. From our experimental findings, we have seen that deep learning has become a successful MT technique as a result of increased processing capacity. Sentences can be translated by a built model with remarkable accuracy using the Encoder-Decoder architecture. Finally, considering all the experiments, a better BLEU score is achieved or documented when Kafi Noonoo is used as the source language and English as the target language using transformer deep learning approaches.

***Received**: May 28, 2025; **Revised**: June 10, 2025; **Accepted:** July 18, 2025; **Published:** August 2025*
*Corresponding author- **Ashagire Adinew***

V. Conclusion

Based on the experiments' results, the transformer model outperforms Bi-LSTM and BiGRU with and without attention models in terms of accuracy, BLEU score, and training time. The transformer also demonstrates better performance than other models examined in these studies when translating more complex sentences. In both translation directions, the transformer model has a faster training time than the other models used in the experiment.

Overall, our research confirms that the transformer model provides good translation accuracy in both directions between the language pairs. Consequently, there is a need to generate parallel corpora and add different dialect corpora to conduct comparable research and include more phonetic variation in Kafi Noonoo.

References

1. A. N. M. Translation and W. W. Gaga, "Translation techniques and challenges in scientific texts: The case of the College of Natural Sciences," *B.A. thesis*, Dept. Foreign Lang. & Lit., Addis Ababa Univ., Addis Ababa, Ethiopia, Oct. 2020.
2. "Language statistics & facts," *Vistawide*. [Online]. Available: https://www.vistawide.com/languages/language_statistics.htm. Accessed: Mar. 8, 2023.
3. I. A. Bolshakov and A. Gelbukh, Computational Linguistics: Models, Resources, Applications, 1st ed. Mexico: Instituto Politécnico Nacional, 2004.
4. A. Asfaw, Kafi Noonoo Dictionary for Printing. Unpublished.
5. S. Takili, Kafi Noonee Doyee Inde IIqqeena'o. Unpublished manuscript.
6. K. M. Macho and Addis, "Documentation and description of noun morphology in Kafinoonoo," *J. Afr. Lang. Stud.*, vol. 25, no. 46, pp. 1–120, 2006.
7. T. Kassa, "Case system of Kafinoonoo," *M.A. thesis*, Dept. Linguist, Addis Ababa Univ., Addis Ababa, Ethiopia, 2012.
8. K. Meshesha, "Kafi noonoo noun morphology," *M.A. thesis*, Dept. Linguist, Addis Ababa Univ., Addis Ababa, Ethiopia, 2007.
9. R. Theil, "Kafa phonology," *J. Afr. Lang. Linguist.*, vol. 28, no. 1, pp. 1–24, 2007, doi: 10.1515/JALL.2007.010.
10. J. Allen, *Natural Language Understanding*, 2nd ed. Redwood City, CA: Benjamin Cummings, 1995.
11. R. Mitkov, Ed., *The Oxford Handbook of Computational Linguistics*. Oxford, UK: Oxford Univ. Press, 2003.

***Received**: May 28, 2025; **Revised**: June 10, 2025; **Accepted:** July 18, 2025; **Published:** August 2025*
*Corresponding author- **Ashagire Adinew***

12. F. Katamba, *Morphology*. London, UK: Macmillan, 1993.
13. A. K. Joshi, *Computational Linguistics: An Introduction*. Cambridge, UK: Cambridge Univ. Press, 2001.
14. R. Hausser, *Foundations of Computational Linguistics*. Berlin, Germany: Springer, 1999, doi: 10.1007/978-3-662-03920-5.
15. Y. Goldberg, *Neural Network Methods in Natural Language Processing*. San Rafael, CA: Morgan & Claypool, 2017.
16. W. J. Hutchins, Machine Translation: Past, Present, Future. *Chichester*, UK: Ellis Horwood, 1986.
17. J. R. Chaudhary and P. A. C. Patel, "Bilingual machine translation using RNN-based deep learning," *Int. J. Eng. Technol.*, vol. 4, no. 4, pp. 1480–1484, 2018.
18. A. Adinew, "Kafi Noonoo grammar errors detection using deep learning approaches," in *Proc. ICT4DA*, 2024, pp. 154–159, doi: 10.1109/ICT4DA62874.2024.10777277.
19. T. T. T. Ayele, "Word sense disambiguation for Wolaita language using machine learning approach," *M.S. thesis, Adama Sci. Technol. Univ.*, Adama, Ethiopia, 2021.
20. F. Benabbou, H. Boukhouima, and N. Sael, "Fake accounts detection system based on bidirectional gated recurrent unit neural network," *Int. J. Electr. Comput. Eng.*, vol. 12, no. 3, pp. 3129–3137, 2022, doi: 10.11591/ijece.v12i3.pp3129-3137.
21. P. Wu et al., "Sentiment classification using attention mechanism and bidirectional long short-term memory network," *Appl. Soft Comput.*, vol. 112, p. 107792, 2021, doi: 10.1016/j.asoc.2021.107792.
22. S. F. Ayegba, O. E. Osuagwu, and N. D. Okechukwu, "Machine translation of noun phrases from English to Igala using the rule-based approach," *J. Lang. Technol. Comput. Linguist*, vol. 11, no. 1, pp. 18–28, 2014.
23. P. Goyal, S. Pandey, and K. Jain, *Deep Learning for Natural Language Processing*. Berkeley, CA: Apress, 2018.
24. S. T. Abate et al., "Parallel corpora for bi-lingual English-Ethiopian languages statistical machine translation," in *Proc. Interspeech*, 2018, pp. 3102–3111.
25. A. A. Belay, "A bi-directional Ge'ez-Amharic neural machine translation: A deep learning approach," *M.S. thesis, Bahir Dar Univ.*, Bahir Dar, Ethiopia, 2021.
26. M. Melese, "Attention-based neural machine translation from English to Wolaytta," M*.S. thesis, Addis Ababa Univ.*, Addis Ababa, Ethiopia, 2023.
27. T. Tian et al., "A French-to-English machine translation model using a transformer network," *Procedia Comput. Sci.*, vol. 199, pp. 1438–1443, 2022, doi: 10.1016/j.procs.2022.01.182.

***Received**: May 28, 2025; **Revised**: June 10, 2025; **Accepted:** July 18, 2025; **Published:** August 2025*
*Corresponding author- **Ashagire Adinew***

DOI: *https://doi.org/10.59122/184DFD14*

28. L. Corallo et al., "A framework for German-English machine translation with GRU RNN," in *Proc. IEEE Conf. Comput. Linguist*, 2022, pp. 1–8.
29. M. K. Negia, R. M. Tamiru, and M. Meshesha, "Amharic-Kistanigna bi-directional machine translation using deep learning," *Proc. ICT4DA*, 2022, pp. 61–65, doi: 10.1109/ICT4DA56482.2022.9971194.
30. A. Belete, "Bi-directional English-Awngi machine translation using deep learning," *M.S. thesis, Bahir Dar Univ.*, Bahir Dar, Ethiopia, 2022.
31. A. Wahid, "Machine translation system using deep learning for English to Urdu," *J. Nat. Lang. Process.*, vol. 12, no. 3, pp. 45–60, 2022.
32. L. Corallo et al., "A framework for German-English machine translation with GRU RNN," *IEEE Trans. Comput. Linguist*, vol. 15, no. 2, pp. 112–125, 2022.
33. E. Asefa and H. Seid, "Developing English to Dawurootsuwa machine translation model using RNN," *Ethiop. J. Sci. Technol.*, vol. 69, no. 6, pp. 50–56, 2021.
34. C. D. Manning, P. Raghavan, and H. Schütze, *Introduction to Information Retrieval*. New York, NY: Cambridge Univ. Press, 2008.
35. A. Vaswani et al., "Attention is all you need," in *Adv. Neural Inf. Process. Syst.*, vol. 30, 2017, pp. 5999–6009.

***Received**: May 28, 2025; **Revised**: June 10, 2025; **Accepted:** July 18, 2025; **Published:** August 2025*
*Corresponding author- **Ashagire Adinew***