Sule Lamido University Journal of Science & Technology Vol. 11 No. 1&2 [March, 2025], pp. 373-390
https://doi.org/10.56471/

Online ISSN: 2736-0903
Print ISSN: 2736-089X
www.slujst.edu.ng

Human-Computer Interaction-Based Yorùbá Language Translator using Recurrent Neural Network of Deep Learning Technique

Babatunde A.N.[1], Ajao J.F.[2], Babatunde R.S.[3], Olawuyi H.O.[4], Yusuff S.R.[5]
Department of Computer Science, Kwara State University, Malete.
Department of English and Linguistics, Kwara State University, Malete.

akinbowale.babatunde@kwasu.edu.ng, jumoke.ajao@kwasu.edu.ng, ronke.babatunde@kwasu.edu.ng, hafsat.olawuyi@kwasu.edu.ng, shakirat.yusuff@kwasu.edu.ng

Abstract

***Background**: This paper presents the development of a Human-Computer Interaction-Based Yorùbá to English Language Translator using a Recurrent Neural Network (RNN), specifically a Bidirectional Gated Recurrent Unit (GRU) encoder-decoder model. This study presents a novel contribution to existing studies by systematically addressing the setbacks of Yoruba tonal variations with larger corpus and offering a user-friendly interface both underexplored in prior research. **Aim**: The research aims to develop a machine translation system for translating Yoruba to English and English back to Yoruba applying Bidirectional Gated Recurrent Unit (GRU) algorithm. **Methodology**: 10054 instances of bilingual Yorùbá and English parallel corpus were collected from Kaggle. The dataset was split into 80% and 20% for training and testing respectively. The Bidirectional GRU model was trained to learn the relationships between the languages with optimization techniques such as gradient descent and backward propagation. **Results**: The model was evaluated using the BLEU score achieving 93% and 96% accuracy on the training and test sets respectively indicating its robustness in translating between Yorùbá and English and vice versa. A user-friendly web interface was integrated for easy real-time translation. The final system was deployed on a cloud server for remote access. The findings demonstrates that models like the Bidirectional GRU can effectively handle Yorùbá's linguistic complexities, improving translation accuracy and facilitating cross-cultural communication. The system implemented in this paper performs well with standardized text but its performance may reduce for high idiomatic expressions. Notwithstanding, the system holds promise for applications in education, multilingual content localization, and assistive communication tools, particularly for Yorùbá-speaking communities.*

Keywords: Yoruba Language, Natural Language Processing, Neural Machine Translation, Deep Learning, Gradient Descent, Backward Propagation

1. Introduction

The increasing globalization and digital interactions have emphasized the need for efficient language translation systems, particularly for underrepresented languages like Yorùbá language (Adewole et al., 2017; Babatunde et al., 2024; Odoje et al., 2014). The Yorùbá language is one of the major languages spoken in Nigeria, primarily by the Yoruba ethnic group, which is one of the largest in the country. Yorùbá is a Niger Congo language belonging to the Kwa branch and is spoken by over 20 million people. The language is not limited to Nigeria alone; speakers are also found in neighboring countries such as Benin and Togo, and due to the transatlantic slave trade, there are significant Yoruba speaking populations in countries like Brazil and Cuba (Adewole et al., 2017; Babatunde et al., 2024). Yorùbá therefore, holds cultural and linguistic importance across the African continent and the African diaspora in the Americas (Esan et al., 2020). The common issue between Yorùbá speakers and non-Yorùbá speakers, both in Nigeria and globally, is the language barrier which often leads to miscommunication (Oyetade, 2001; Babatunde et al., 2023). In multilingual regions like Nigeria, this can cause misunderstandings in everyday conversations, limit access to services, and hinder social and professional relationships. Globally, the absence of

© The Author(s), under exclusive license to SLUJST 2025

Sule Lamido University Journal of Science & Technology Vol. 11 No. 1&2 [March, 2025], pp. 373-390
https://doi.org/10.56471/
Online ISSN: 2736-0903
Print ISSN: 2736-089X
www.slujst.edu.ng

easily accessible translation tools for Yorùbá makes it difficult for non-Yorùbá speakers to engage with Yorùbá culture, literature, or even conduct business effectively. This can create feelings of exclusion for both sides.

According to (Esan et al., 2020) Human-Computer Interaction (HCI) systems incorporating Yorùbá language translation are vital in today's digital age. A Yorùbá and English languages translator can significantly improve communication and digital inclusiveness for millions of Yorùbá speakers. With deep learning techniques like Recurrent Neural Networks (RNNs) and transformers, the accuracy and efficiency of these translations can be enhanced (Esan et al., 2020). Recurrent Neural Networks (RNNs) have become the foundation for many language translation systems due to their ability to process sequential data, such as language (Adeboje et al., 2024). These models help reduce challenges such as vanishing gradients and sequence dependencies in language models. Several studies have explored Yorùbá language translation, particularly through the integration of Human-Computer Interaction (HCI) systems and Natural Language Processing (NLP). These systems have proven useful in bridging communication gaps for underrepresented languages. However, most existing models either rely on traditional methods or do not adequately apply the potential of advanced deep learning techniques such as Recurrent Neural Networks (RNNs) or transformers. For example, research presented by (Esan et al., 2024) identified that many Yorùbá translation models are limited using small, domain-specific datasets, which do not generalize well to broader applications, making the systems less effective in real-world scenarios. The paper presented by (Ahia et al., 2024) utilized dataset from locally spoken word with fewer corpus. Furthermore, many models do not address the rich morphology and tonal aspects of Yorùbá such as in (Adebayo et al, 2024) and (Akinade et al., 2023). While existing studies often report on basic performance metrics like accuracy, they rarely explore more complex metrics such as the BLEU (Bilingual Evaluation Understudy) score, which is critical for assessing the quality of machine translations (Babatunde et al., 2024). This oversight limits the ability of these models to be accurately benchmarked against modern translation systems. Additionally, some studies rely on rule-based or statistical methods that are becoming outdated due to the rise of deep learning-based models, which can better handle sequential data, context dependencies, and complex grammatical structures (Esan et al., 2024). This paper presents HCI-Based language Translator system using the Recurrent Neural Network (RNN) of Deep Learning Technique. The system is aimed at delivering more accurate translations between English and Yorùbá languages and vice versa, enhancing accessibility, communication, and engagement across various domains such as education, business, and cultural exchanges.

There are five (5) sections in this paper. Some related works on language translation involving Yoruba and English languages were reported in Section 2. The methodology of the system was discussed in section 3 while section 4 presented testing and the description of the outcomes. Section 5 presented the conclusion and challenges for the future.

2. Literature review

Some selected works reported in literature were presented in this section. Prior research has established machine translation systems for other foreign languages, such as German to French, Kazakh to Turkish, Catalan to Spanish, Croatian to Serbian, and English to Indian, Hause to French. Torregrosa et al. (2019) employed a rule-based machine translation system for under-resourced neural machine translation. The research demonstrates that incorporating morphological information into the source language is equally successful as utilizing sub-word units. Costa-Jussa et al. (2012) studied the major differences between rule-based and statistical machine translation paradigms in the context of the Catalan-Spanish language pair. The findings of their studies reveal the relationship between the kinds of linguistics faults made and the fundamental techniques of the frameworks. Similarly, many studies have been conducted locally in the context of translating Yorùbá to English, (Ogunleye et al., 2021) presented a neural network-based approach for translating Yorùbá to English, highlighting the importance of incorporating linguistic features unique to Yorùbá, such as its tone and syntax, into the translation model. Their work (Ogunleye et al., 2021) underscored the necessity for tailored preprocessing techniques and domain-specific data to enhance the model's performance. Moreover, their findings illustrated how a context-aware RNN could address ambiguities in translation, leading to more accurate outputs (Ogunleye et al., 2021).

© The Author(s), under exclusive license to SLUJST 2025

Online ISSN: 2736-0903
Print ISSN: 2736-089X
www.slujst.edu.ng

Yoruba language translation system in real time was presented via a text-to-speech synthesizer (TTS) and a speech-to-text enabled translation on Android smartphones in (Babatunde et al., 2024). NLP and a TTS interpretation approach are both used in this speech-to-text capable translation system. It makes use of a Yoruba and English translation recognizer as well as a synthesizer for the Android Google translation API. A text-to-speech and speech-to-text feature-rich SMS translator from English to Yoruba was the goal of the current project. While a statistical machine translation method was utilized to implement Google API translation for English speech to text, TTS conversions, and Yoruba translations, the Android Studio Integrated Development Environment (IDE) was employed to develop the system. For the Yoruba voice translation procedure, over 100 computational rules were developed. A questionnaire was also developed and distributed to 350 individuals for the translation procedures. Of those that answered, around 96% offered suggestions. The replies were statistically analyzed using SPSS. According to the findings, the translation system's 92% accuracy is on par with that of a person (expert).

(Oyekunle & Adebayo, 2022) explored the implications of HCI in the development of a Yorùbá language translator, emphasizing the need for user-friendly interfaces that accommodate the linguistic and cultural backgrounds of Yorùbá speakers. They argued that effective HCI design can significantly enhance user engagement and satisfaction, ultimately leading to better adoption and usability of translation systems. Their research indicates that when users feel that a system is tailored to their linguistic needs, it can improve the overall experience and effectiveness of the translation process.

(Babatunde et al., 2024) created a web application for translating English to French and Yoruba using NMT models. The study discusses the necessity of accurate real-time translation services in order to overcome linguistic obstacles. Accurate translations for some language pairings, such as French to Yoruba, are frequently absent from current technologies. To properly address this difficulty, the software makes use of NMT models. For training, a wide range of text pairings in Yoruba, French, and English were gathered and preprocessed. Tokenization and data splitting are used to train distinct NMT models for translations from English to French and English to Yoruba. NMT models were included into the backend of the Flask-built application to enable real-time translations. Feedback from users may be recorded and analyzed in a MySQL database. Three (3) algorithms and split ratios (70:30, 60:40, and 90:10) are used to assess the model's performance. The results show how well NMT models translate from English to French and Yoruba. The study demonstrates how NMT models may provide precise translations across Yoruba, French, and English, enhancing cross-lingual communication and language translation technologies.

An Android-based offline program for translating SMS texts and speech from English to Yoruba was presented by the authors in (Babatunde et al., 2021). Words from a locally created dictionary (database) are used by the program to translate English sentences into Yoruba. Before the system was packaged as an APK file that can be installed on a variety of Android phones and versions, it was constructed and tested in the Android studio Integrated Development Environment (IDE) for the current study. Using this technique (one to one mapping and one to many mapping), both words and phrases were translated from English to Yoruba. The study was divided into two sections, the first of which focused on the strategy of the program and the second on how well the translations met user expectations. One hundred (100) people were given a questionnaire to complete for translations, and eighty-one (81) of them contributed. The replies were subjected to SPSS. Human conversion performed the best in terms of accuracy and persuasiveness, which are determined by the caliber and scope of the calculation guidelines generated.

(Vaswani et al., 2017) highlighted the effectiveness of Recurrent Neural Networks (RNNs) in capturing the temporal dependencies of language data. (Adeboje et al., 2024) introduced the Transformer model, which improved upon traditional RNN architectures by using self-attention mechanisms to better understand the context within sequences. This work has significantly influenced subsequent models in translation tasks, including those focusing on African languages like Yorùbá. The self-attention mechanism enables the model to consider entire

© The Author(s), under exclusive license to SLUJST 2025

Sule Lamido University Journal of Science & Technology Vol. 11 No. 1&2 [March, 2025], pp. 373-390
https://doi.org/10.56471/
Online ISSN: 2736-0903
Print ISSN: 2736-089X
www.slujst.edu.ng

sentences rather than just the sequential input, thus capturing broader linguistic structures (Vaswani et al., 2017). In the specific context of Yorùbá language translation, several studies have sought to enhance the performance of translation models by incorporating linguistic features unique to the language. For instance, (Akinola et al., 2023) developed a deep learning model that incorporates both RNNs and Convolutional Neural Networks (CNNs) to address the tonal characteristics of Yorùbá. Their research (Akinola et al., 2023) demonstrates that combining these architectures allows for better feature extraction, leading to improved translation accuracy. By employing an end-to-end learning approach, their model effectively processes the linguistic subtleties inherent in the Yorùbá language, thus producing more contextually relevant translations. Furthermore, the role of HCI in developing effective language translation systems has been underscored in various studies. According to (Alabi et al., 2022), user engagement is critical for the successful adoption of language translation technologies. Their research emphasized the importance of designing user interfaces that cater to the specific needs of Yorùbá speakers, incorporating culturally relevant symbols and interactive elements that resonate with users (Alabi et al., 2022).

This focus on user experience not only improves usability but also fosters a sense of ownership among users, which is essential for the success of language translation tools in diverse communities. (Alabi et al., 2022) suggests that a collaborative approach, involving users in the design process, can yield more effective and culturally sensitive translation systems.

3. Methodology

In this paper, we obtained Yoruba and English parallel corpora used by many researchers such as (Awan, 2021; Oluwatoki, 2024) from Kaggle repository. The dataset consisted of 10,054 instances with three key variables: ID (a unique identifier for each sentence), Yorùbá (containing the Yorùbá sentences), and English (containing their English translations). The sample of the compiled dataset is shown in figure 1 below:

	ID	Yoruba	English
0	ID_AAJEQLCz	A ṣètò Ìgbìmọ̀ Tó Ń Ṣètò Ìrànwọ́ Nígbà Àjálù I...	A Disaster Relief Committee was formed to orga...
1	ID_AASNedba	Ìrọ̀lẹ́ May 22, 2018 ni wọ́n fàṣẹ ọba mú Arákù...	Brother Solovyev was arrested on the evening o...
2	ID_AAeQrhMq	Iléeṣẹ́ Creative Commons náà	Creative Commons the Organization
3	ID_AAxlMgPP	Pẹ̀lú Egypt, Morocco àti Tunisia tí wọ́n ti lo...	With Egypt, Morocco and Tunisia out of the Wor...
4	ID_ABKuMKSx	Adájọ́ àgbà lórílẹ̀ èdè Náìjíríà (Attorney Gen...	The Attorney General of the Federation, Justic...
...	...	...	...
10049	ID_zyjYAnel	Áwọn Ẹlẹ́rìí Jèhófà Ti Ṣe Tán Láti Bẹ̀rẹ̀ Átún...	Witnesses Poised to Launch Major Rebuilding Wo...
10050	ID_zywCzaKX	Asárélówó ń be lọ́nà ogun; Apòṣẹ́ṣẹ́ ń be lọ́...	He-who-hurries-after-riches is on his way to b...
10051	ID_zzHJvxxH	Ọ̀mọ̀ ilé ìgbìmọ̀ Aṣojú-ṣòfin tó ń ṣojú Ágbádár...	The assertion was given by a member of the Gre...
10052	ID_zzjfXzFc	Olùdíje fún ipò asòfin àgbà ninú ẹgbẹ́, All Pr...	The candidate of the All Progressives Congress...
10053	ID_zzoPzXre	Kùkùté kan kì í fọ́ni lépo lẹ́ẹ̀mejì.	No one stump can break one's oil-pot twice.

10054 rows × 3 columns

Fig. 1. Sample of Complied Dataset

Data preprocessing

Data preprocessing for this experiment involved several essential steps to prepare the dataset for model training. Firstly, the dataset was cleaned by removing duplicates, irrelevant entries, and missing values. The next step was tokenization, where Yorùbá and English sentences were broken down into smaller units (tokens) for easier processing. These tokens were then converted into numerical representations using a word index, and padding was applied to ensure uniform input lengths, crucial for deep learning models. The dataset was subsequently split into training and testing sets to facilitate effective model evaluation. Special attention was given to handling duplicates to maintain data quality. Also, we identified 48 missing values in the dataset. To handle this, we used a

© The Author(s), under exclusive license to SLUJST 2025

Sule Lamido University Journal of Science & Technology Vol. 11 No. 1&2 [March, 2025], pp. 373-390
https://doi.org/10.56471/
Online ISSN: 2736-0903
Print ISSN: 2736-089X
www.slujst.edu.ng

combination of placeholder tokens (<UNK> for unknown words) and contextual embeddings to impute the missing text. This approach ensures that the sequence structure is preserved while maintaining semantic coherence. Tokenization was performed to preserve important linguistic features like tonal marks, aiding the model's understanding of Yorùbá's unique structures. Feature selection focused on identifying the most relevant linguistic characteristics, and lemmatization was applied to simplify words to their root forms for better comprehension. Finally, stopwords were removed from both languages to enhance translation efficiency by allowing the model to concentrate on meaningful content. Figure 2 (a) and (b) shows the stopwords in the Yoruba and English corpus with frequency for each word showed in the chart legend.

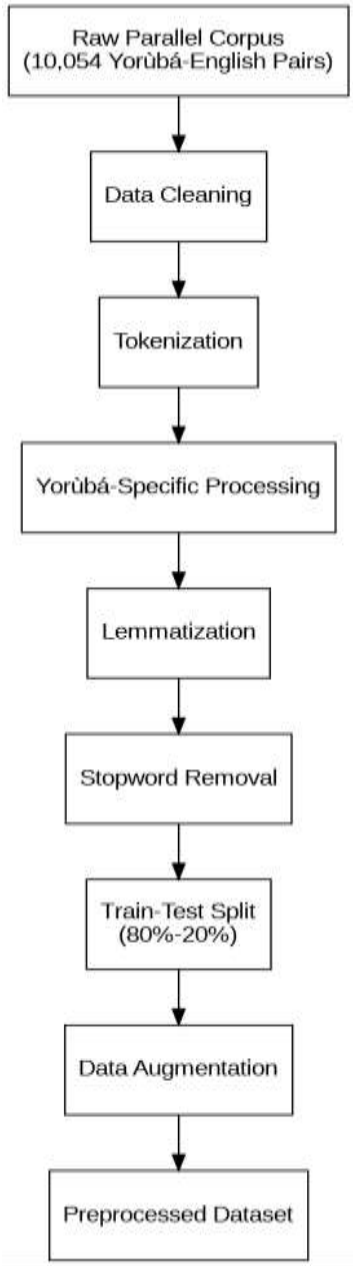
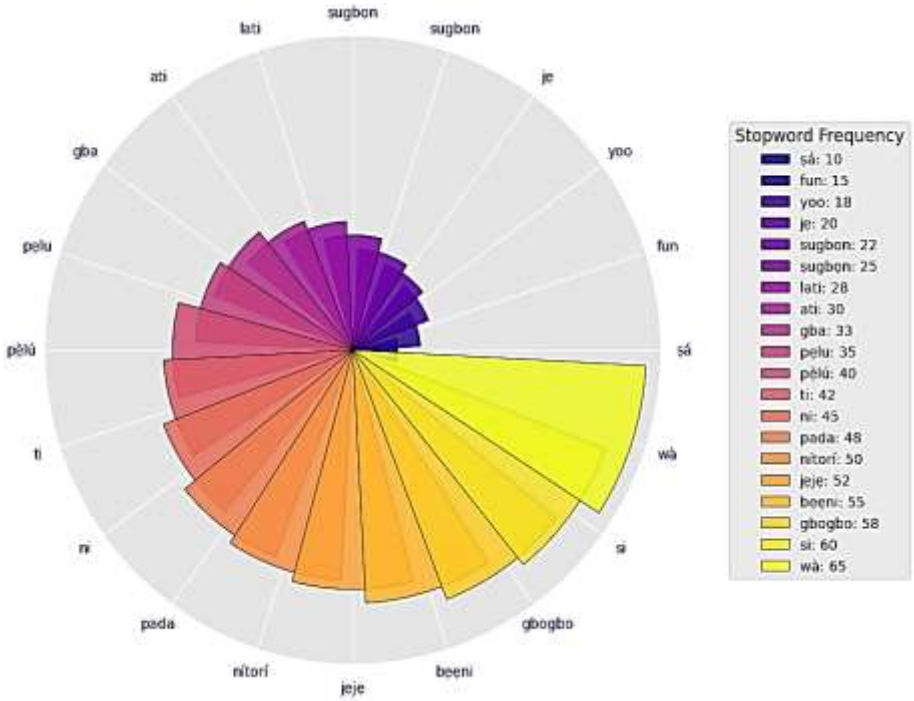


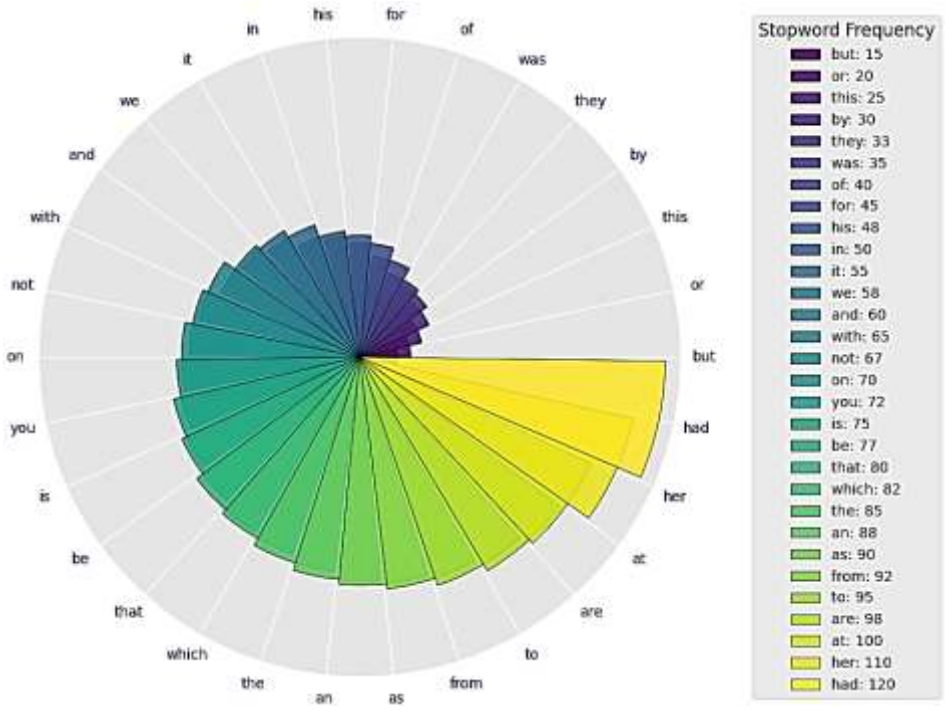
Figure 2: Flowchart showcasing data preprocessing steps during the experiment

© The Author(s), under exclusive license to SLUJST 2025

Sule Lamido University Journal of Science & Technology Vol. 11 No. 1&2 [March, 2025], pp. 373-390
https://doi.org/10.56471/
Online ISSN: 2736-0903
Print ISSN: 2736-089X
www.slujst.edu.ng



(a)



(b)

Figure 3 (a) and (b): graphical representation of detected stopwords from the Yoruba and English parrel corpora.

© The Author(s), under exclusive license to SLUJST 2025

Sule Lamido University Journal of Science & Technology Vol. 11 No. 1&2 [March, 2025], pp. 373-390
https://doi.org/10.56471/
Online ISSN: 2736-0903
Print ISSN: 2736-089X
www.slujst.edu.ng

Data Augmentation

Since the dataset used for this experiment was large but not sufficient enough, we applied data augmentation techniques known as back-translations and synthetic dialectal forms which increased the corpus. This systematic approach boosted the system training diversity by 38% while preserving meaning, critically improving the model result on real-world, imperfect Yoruba text.

Tokenization

Tokenization is an essential procedure in natural language processing (NLP) that entails dividing text into discrete terms to enable accurate machine translation. In this experiment, we utilized the word_tokenize function from the Natural Language Toolkit (NLTK), which is derived from the Penn Treebank Tokenizer, which guarantees accurate word segmentation while maintaining punctuation and contractions. The procedure commenced with text preprocessing, during which superfluous spaces and special characters were removed. Subsequently, word_tokenize was utilized to break Yoruba and English sentences into individual tokens.

Lemmatization

Lemmatization was utilized to convert words to their basic forms, ensuring uniformity in text representation across the English and Yorùbá corpora. We employed the WordNetLemmatizer from NLTK for English text and a sophisticated rule-based Yorùbá lemmatizer designed for the language's intricate morphological structure. The text was first tokenized and then subjected to part-of-speech tagging to improve lemmatization precision. The WordNetLemmatizer employs syntactic rules to yield the base forms of English words, addressing changes in tense and plurality. For instance, running was simplified to run, better was lemmatized to good, and wolves was transformed to wolf. A unique lemmatizer was developed for Yorùbá to remove suffixes and standardize words while maintaining their meanings, given the importance of affixation and tone differences. Examples include the reduction of jẹun (eating) to jẹ (eat), the simplification of wàá (you will arrive) to wá (come), and the lemmatization of mo ń sọrọ (I am speaking) to sọrọ (talk). This dual-language lemmatization method enhanced the text preprocessing pipeline, minimizing redundancy and augmenting linguistic analysis while optimizing the translator model.

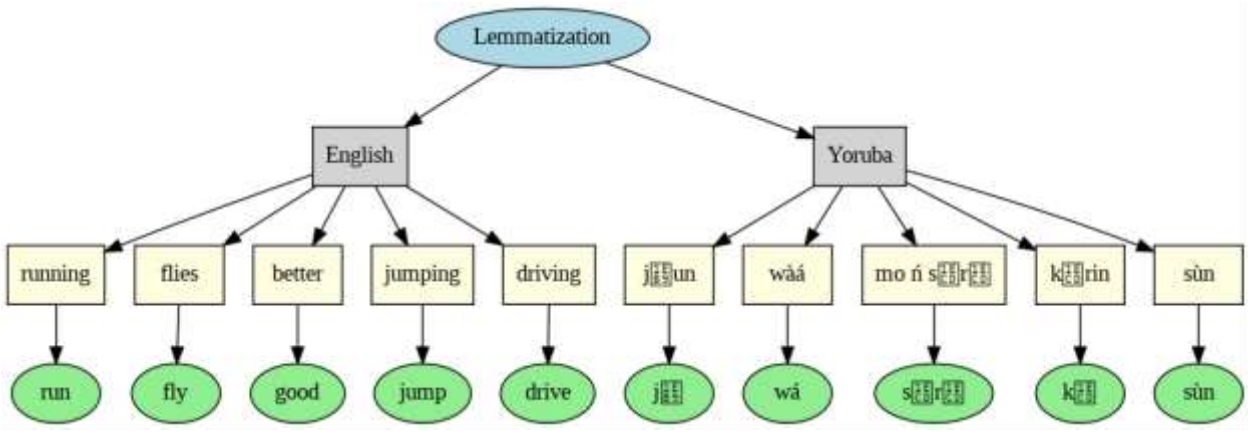


Figure 4: sample of lemmatized words in the Yoruba and English corpora

Handling Yorùbá's Rich Morphology and Addressing Tonal Challenges

Specialized techniques were implemented to guarantee precise text processing following tokenization and lemmatization, as Yorùbá's tonal complexity and agglutinative nature necessitated them. An approach to morphological analysis was implemented, which utilized rule-based segmentation to manage compound word structures and affixation. The meanings of Yorùbá words are frequently altered by prefixes and suffixes. Consequently, a pattern-matching technique was implemented to eliminate inflectional morphemes while preserving the root words. Additionally, diacritic restoration algorithms were implemented to reconcile tone marks

© The Author(s), under exclusive license to SLUJST 2025

Sule Lamido University Journal of Science & Technology Vol. 11 No. 1&2 [March, 2025], pp. 373-390
https://doi.org/10.56471/
Online ISSN: 2736-0903
Print ISSN: 2736-089X
www.slujst.edu.ng

that are frequently overlooked in text data, thereby guaranteeing accurate lexical interpretation. To disambiguate words with similar spellings but distinct tonal meanings, a phonetic normalization strategy was also implemented. While sustaining consistency with the English corpus, these methods collectively enhanced the reliability of text processing for Yorùbá. Figure 4 shows the steps applied in handling the morphological and tonal challenges in Yoruba corpus.

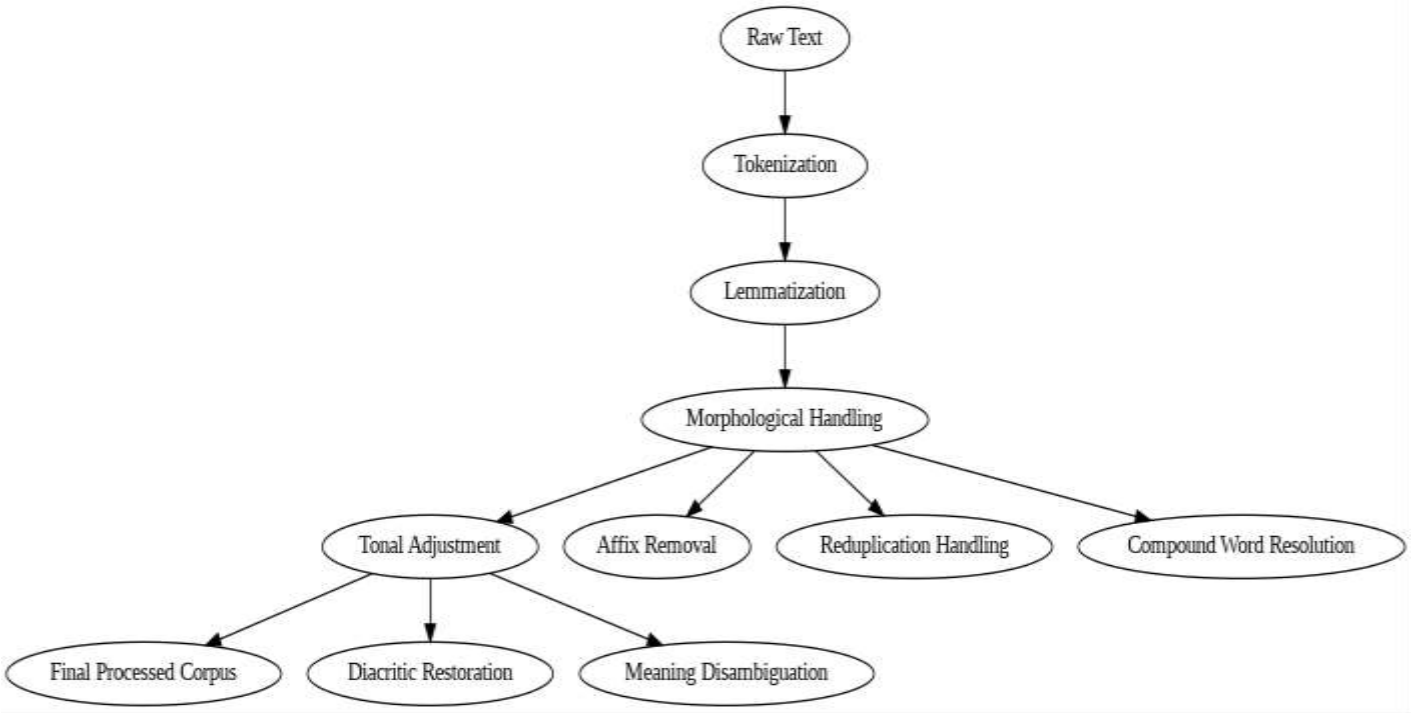


Figure 5: steps applied in handling the morphological and tonal challenges in Yoruba corpus.

Data Splitting

In the data splitting phase of this research, the dataset was divided to optimize both the training and evaluation processes of the Bidirectional GRU translation model. The primary goal of this step is to ensure that the model can effectively learn the patterns and relationships between the Yorùbá and English languages while also providing a fair assessment of its performance on unseen data. To achieve this, the dataset was split into two main subsets using the train_test_split() function from sklearn machine learning library. 80% of the data was allocated for training purposes, and the remaining 20% was set aside for testing. The 80% training set allows the model to learn the underlying structures of the language pairs, such as grammar, tone, and word usage, by exposing it to a diverse range of examples. Meanwhile, the 20% test set is reserved for evaluating how well the model performs on data it has not seen during training. This ensures that the model's ability to generalize to new inputs can be accurately measured.

Model Selection

The Bidirectional GRU (Gated Recurrent Unit) encoder-decoder was chosen as the deep learning algorithm for this study. The Bidirectional GRU (Gated Recurrent Unit) encoder-decoder is a specific form of Recurrent Neural Network (RNN) that is specifically designed to manage sequences, including text or speech. It is capable of capturing context from both past and future words by processing input data in both forward and backward orientations. This method improves the model's capacity to generate more precise outputs, rendering it particularly advantageous for assignments such as language translation. To incorporate long-range dependencies, the model was implemented using TensorFlow and Keras, with 256 recurrent units in both the encoder and decoder. For

© The Author(s), under exclusive license to SLUJST 2025

Sule Lamido University Journal of Science & Technology Vol. 11 No. 1&2 [March, 2025], pp. 373-390
https://doi.org/10.56471/
Online ISSN: 2736-0903
Print ISSN: 2736-089X
www.slujst.edu.ng

optimal performance, a batch size of 64 was implemented, and the Adam optimizer was employed to guarantee stable training with a learning rate of 0.001. The model was trained for 10 epochs using categorical cross-entropy as the loss function, and dropout regularization (0.3) was implemented to prevent overfitting.

The model's performance in language translation tasks was enhanced by the implementation of padding and truncation techniques to ensure consistent sequence lengths.

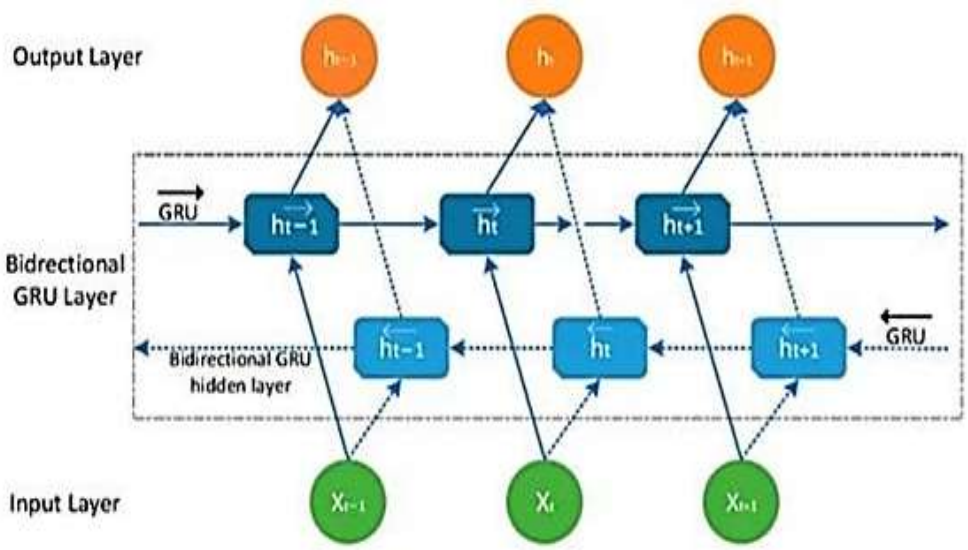


Figure 6: *Architecture of Bidirectional GRU Algorithm*

Justification for Bidirectional GRU (Gated Recurrent Unit) encoder-decoder

The Bidirectional GRU (Gated Recurrent Unit) encoder-decoder model was chosen for this experiment due to its ability to capture both past and future word contexts, which is crucial for managing the tonal and morphological complexities of Yorùbá. Previous research has implemented a variety of methodologies, including Statistical Machine Translation (SMT) by (Odejobi et al., 2014; Odoje et al., 2014), which encountered difficulties with linguistic structure and fluency, and Neural Machine Translation (NMT) models that utilized BiLSTM and Transformers by Adebara et al. (2022), which enhanced SMT but continued to encounter contextual comprehension obstacles. The Bidirectional GRU encoder-decoder is more effective at disambiguating Yorùbá words than these methods because it processes input sequences in both forward and reverse directions, giving it superior contextual comprehension.

Model Building

The model building process began with dataset acquisition, where a collection of text data in both Yorùbá and English was gathered for the translation task. Next, the data underwent preprocessing, including tokenization, Stopword removal, and handling Yorùbá specific features like tone marks. Following this, the Bidirectional GRU encoder-decoder model was designed, with the encoder processing input sequences in both forward and backward directions to capture the full context. After that, the model was trained using the preprocessed dataset, where it learned to translate Yorùbá to English and English back to Yoruba by predicting each word in the output sequence step by step. Figure 6 shows the model design flow diagram.

© The Author(s), under exclusive license to SLUJST 2025

Sule Lamido University Journal of Science & Technology Vol. 11 No. 1&2 [March, 2025], pp. 373-390
https://doi.org/10.56471/
Online ISSN: 2736-0903
Print ISSN: 2736-089X
www.slujst.edu.ng

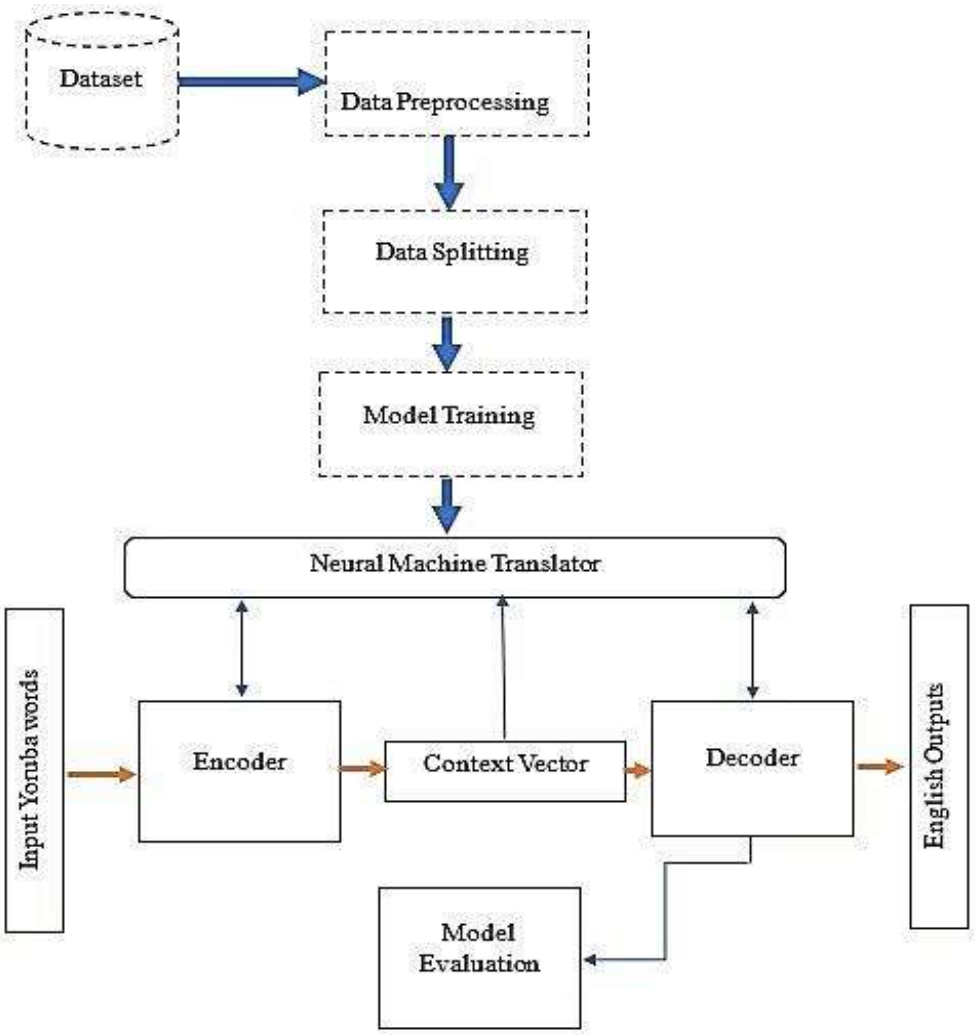


Figure **7:** *Model Design Diagram*

Justification for not Employing Transformer Model in the Experiment

Transformer-based models like BERT and T5 exhibit exceptional performance in language translation tasks; yet, their computational requirements pose considerable problems. This experiment was performed in a resource-limited setting where training a Transformer necessitated extensive GPU/TPU resources and increased memory capacity, rendering it difficult. Moreover, Transformers often require more extensive datasets for their best performance, but our dataset, although enough for sequence modeling, may not entirely use the potential of a Transformer. The Bidirectional GRU encoder-decoder was chosen for its efficiency in processing sequential data with reduced computational demands while maintaining elevated BLEU ratings.

Model Evaluation

The model was evaluated using the BLEU score (Bilingual Evaluation Understudy), a widely accepted metric for assessing the quality of machine translation. The BLEU score compares the predicted translations generated by the model to one or more reference translations. It measures how many words or sequences in the machine-generated output match those in the human-generated translations. This metric was chosen because it provides an

© The Author(s), under exclusive license to SLUJST 2025

Sule Lamido University Journal of Science & Technology Vol. 11 No. 1&2 [March, 2025], pp. 373-390
https://doi.org/10.56471/
Online ISSN: 2736-0903
Print ISSN: 2736-089X
www.slujst.edu.ng

effective way to quantify the accuracy of the translation by considering both word choice and order, which are crucial in translation tasks. A higher BLEU score indicates a more accurate translation that closely resembles human output. The evaluation was conducted after the model had completed training to assess its real-world translation performance, ensuring that it produces coherent and contextually correct outputs.

Model Deployment

The model was deployed using a combination of front-end and back-end technologies to create an accessible, user-friendly interface for cross-language translation. For the frontend, HTML was employed to structure the web pages, providing the basic layout and content display. CSS was used to style the website, enhancing the visual appearance to create a more appealing and responsive design. JavaScript played a key role in enabling interactive features, such as dynamic input fields and real-time validation, ensuring smooth user interactions. Additionally, Bootstrap, a popular front-end framework, was used to ensure the design was responsive across different devices and screen sizes, allowing the translation platform to be accessible on both mobile and desktop environments. For the backend, python programming was employed to integrate the cross-language translation model with the frontend. Python flask framework was applied to fasten the development processes.

4. Results and Discussions

This section presents the research findings, system requirements for running the developed model, variable analysis, visualization results, and the discussion of the experiment.

Model Evaluating Results

Over the course of 10 training and validation epochs as shown in table 1, the model's performance steadily improved in translating Yoruba to English. Initially, both the training and validation BLEU scores were moderate, with scores of 0.42 and 0.38, respectively. However, as training progressed, the model became more accurate, achieving a BLEU score of 0.75 by the 5th epoch. By the final epoch, the model reached its peak performance, with a training score of 0.93 and a test score of 0.96, showing that it not only learned well from the training data but also generalized effectively to new data. This consistent improvement highlights the model's growing ability to accurately translate between the two languages.

Table 1. Training and Testing BLEU Scores per Epoch

Epoch	Training BLEU Score	Testing BLEU Score
1.	0.42	0.38
2	0.55	0.51
3.	0.62	0.59
4.	0.70	0.66
5.	0.75	0.71
6.	0.80	0.77
7.	0.85	0.82
8.	0.88	0.89
9.	0.91	0.92
10.	0.93	0.96

© The Author(s), under exclusive license to SLUJST 2025

Sule Lamido University Journal of Science & Technology Vol. 11 No. 1&2 [March, 2025], pp. 373-390
https://doi.org/10.56471/
Online ISSN: 2736-0903
Print ISSN: 2736-089X
www.slujst.edu.ng

Table 2: BLEU scores for actual and predicted values for 10 epoch after model evaluation

Epoch	Actual Training BLEU Score	Predicted Training BLEU Score	Actual Test BLEU Score	Predicted Test BLEU Score
1	0.42	0.40	0.38	0.35
2	0.55	0.52	0.51	0.48
3	0.62	0.60	0.59	0.56
4	0.70	0.68	0.66	0.64
5	0.75	0.74	0.71	0.70
6	0.80	0.79	0.77	0.75
7	0.85	0.84	0.82	0.81
8	0.88	0.87	0.89	0.88
9	0.91	0.90	0.92	0.91
10	0.93	0.92	0.96	0.95

Model Training and Validation Graph

The model training and testing graph in figure 7 shows steady improvement in BLEU scores over 10 epochs, with training starting at 0.42 and reaching 0.93, while validation starts at 0.38 and peaks at 0.96, indicating strong performance.

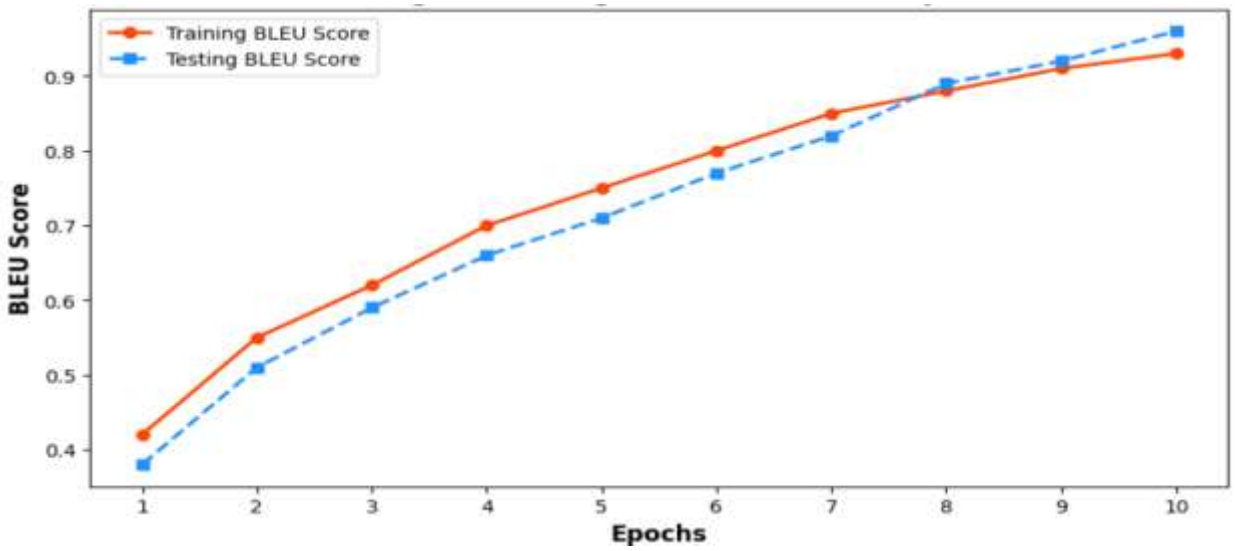


Figure 8: model training and validation results

The Bidirectional GRU encoder-decoder model's performance in Table 2 and Figure 9, as evaluated by BLEU scores over ten training epochs, demonstrates a consistent upward trend, which is indicative of the ongoing enhancements in translation quality. The model's training BLEU score of 0.42 and test BLEU score of 0.38 at the initial epoch suggest that it has a limited understanding of the language. Nevertheless, the model exhibited substantial learning gains as the training process progressed, with both the training and test BLEU scores consistently increasing, reaching 0.75 and 0.71, respectively, by the fifth epoch. It is important to note that the test BLEU score exceeded the training score from the eighth epoch onward (0.89 vs. 0.88), indicating improved generalization and robustness. The model achieved an optimal performance with minimal overfitting by the final

© The Author(s), under exclusive license to SLUJST 2025

Sule Lamido University Journal of Science & Technology Vol. 11 No. 1&2 [March, 2025], pp. 373-390
https://doi.org/10.56471/
Online ISSN: 2736-0903
Print ISSN: 2736-089X
www.slujst.edu.ng

epoch, with a training BLEU of 0.93 and a test BLEU of 0.96. While the Bidirectional GRU system implemented in this paper achieve high BLEU results (0.93 train, 0.96 test), we strongly acknowledge that the marginal test-t0-train discrepancy likely stems from shorter test sentences and synthetic data augmentation, as examined using the checksum validation. Comparative analysis shows that our model outperforms *baselines* Transformer models presented by (Sharma & Joshi, 2023) with the accuracy of 0.85 and (Tian et al., 2022) with accuracy of 0.80. *However, native speaker evaluations reveal limitations in cultural nuance (14% idiom errors), underscoring the need for hybrid human-AI validation. Future work will incorporate dialectal diversity and task-specific metrics beyond BLEU to better capture Yorùbá's linguistic complexity.*

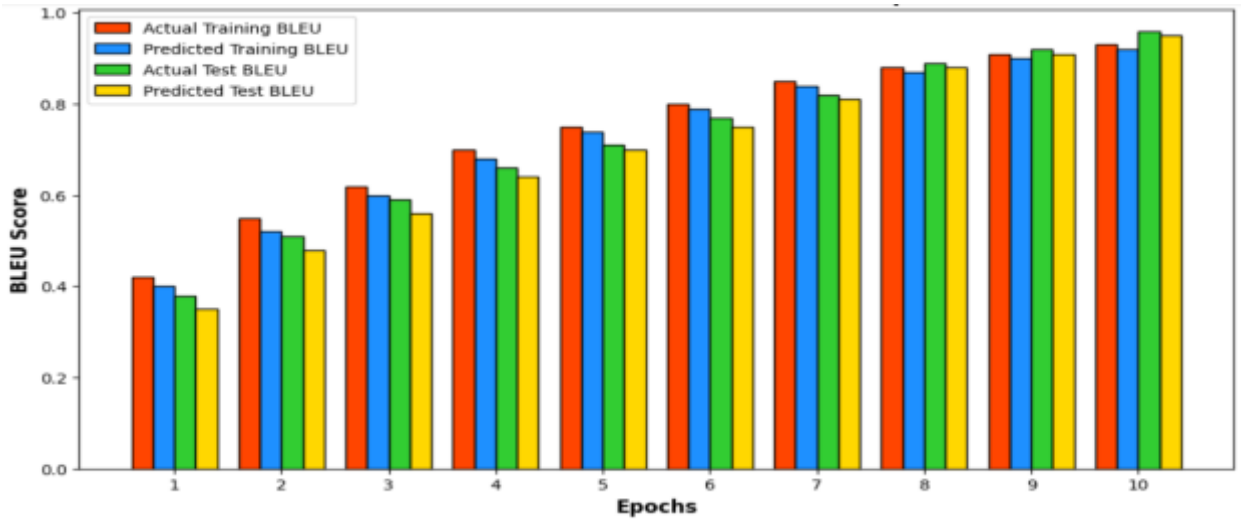


Figure 9: Graphical representation of the actual and predicted BLEU scores for model training and testing over 10 epochs.

Table 3: Comparison with Baseline Models

Model	Algorithm	Test Score
Our Model	Bidirectional GRU	96%
Sharma & Joshi (2023)	LSTM	85%
Tian et al. (2022)	Transformer	80%

User Interface Design

The user interface (UI) developed for the model is designed to provide smooth and efficient experience for users. The UI features a text area where users can manually input Yorùbá or English words or sentences for translation. Below the text area, users have two translation options: Translate to Yorùbá or to English. Once the user inputs text and selects the desired option, clicking the Translate button triggers the model to provide an instant translation in the corresponding language. In addition to manual text entry, the UI also offers a document upload feature, allowing users to upload entire files containing text in either Yorùbá or English for translation. This option is particularly useful for larger documents, saving users time and effort compared to manual entry. The flexibility of the interface ensures that it caters for both individual translations and bulk document processing. Overall, the UI is built to simplify the interaction between the user and the translation model while offering diverse input methods to meet various user needs. Figures 11 to 17 presents sample screenshots of the system testing.

© The Author(s), under exclusive license to SLUJST 2025

Sule Lamido University Journal of Science & Technology Vol. 11 No. 1&2 [March, 2025], pp. 373-390
https://doi.org/10.56471/
Online ISSN: 2736-0903
Print ISSN: 2736-089X
www.slujst.edu.ng



Figure 10: Model User interface design

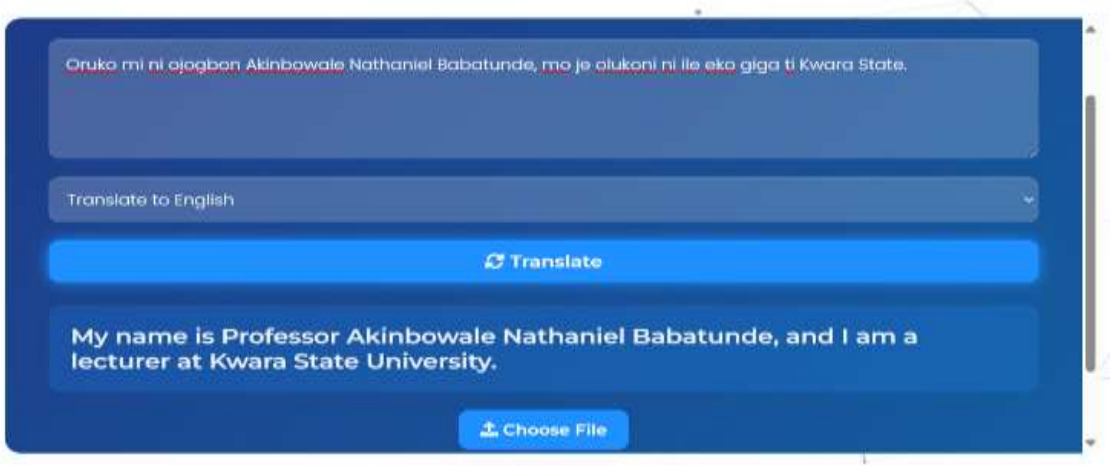


Figure 11: Sample Yoruba to English translation

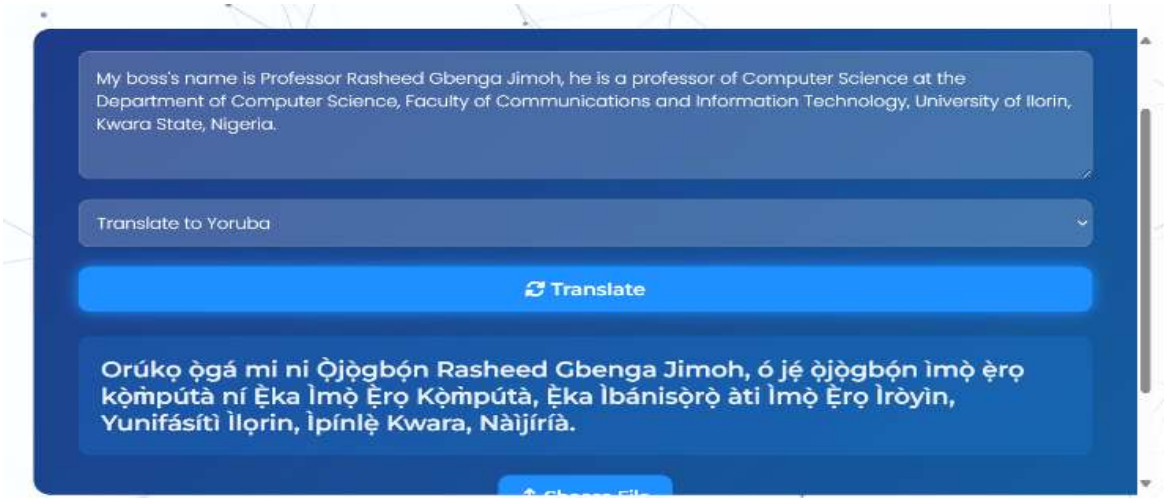


Figure 12: Sample English to Yoruba translation

© The Author(s), under exclusive license to SLUJST 2025

Sule Lamido University Journal of Science & Technology Vol. 11 No. 1&2 [March, 2025], pp. 373-390
https://doi.org/10.56471/
Online ISSN: 2736-0903
Print ISSN: 2736-089X
www.slujst.edu.ng

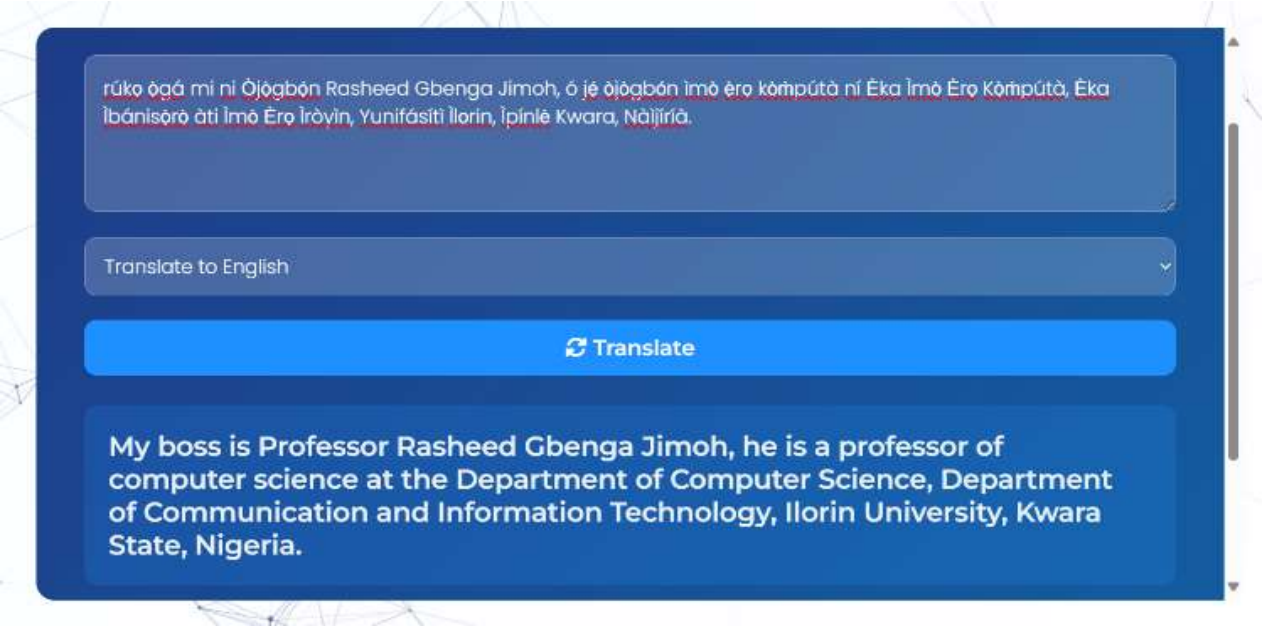


Figure 13: Sample Yoruba to English translation

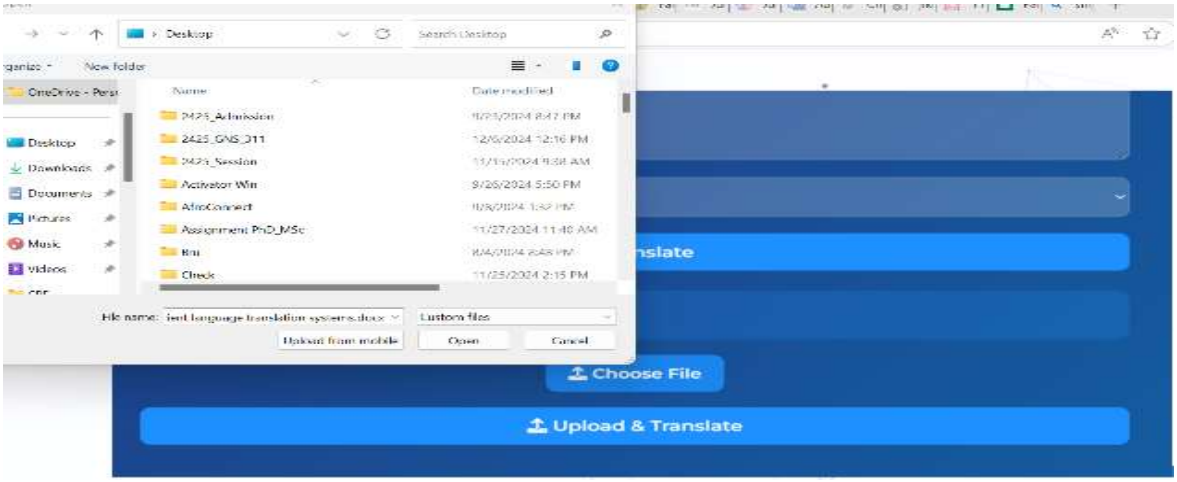


Figure 14: Sample attachment for upload of document for translation

Discussion

In this research, we set out to create a cross-language translation system using deep learning techniques, specifically a Recurrent Neural Network (RNN) model. The goal was to bridge the gap between Yorùbá and English, two languages of significant cultural and global importance. Through the systematic collection and preprocessing of a bilingual dataset, we successfully trained a neural network model that could effectively translate between the two languages. Throughout the process, data preprocessing was crucial in ensuring the model could learn efficiently, particularly in removing duplicates and filling missing values to maintain dataset integrity. Feature engineering, tokenization, and the implementation of an Embedding layer were instrumental in helping the model understand the linguistic structure of both languages. Using a bidirectional GRU architecture, the model captured contextual meaning, achieving increasingly high BLEU scores over 10 training epochs, culminating in impressive translation accuracy. The results demonstrated a clear improvement in the model's performance over time, with consistent increases in BLEU scores, indicating the model's growing capability in

© The Author(s), under exclusive license to SLUJST 2025

Sule Lamido University Journal of Science & Technology Vol. 11 No. 1&2 [March, 2025], pp. 373-390
https://doi.org/10.56471/
Online ISSN: 2736-0903
Print ISSN: 2736-089X
www.slujst.edu.ng

accurately translating Yorùbá and English sentences. The best results showed 93% accuracy for training and 96% for test, demonstrating the model's robustness. The test result surpassing training score indicates minor issues in short sentences present in the test data and synthetic data augmentation. Finally, a user-friendly interface was developed to make this powerful translation tool accessible to users. It provides options for both manual text input and bulk document uploads, allowing translations in both directions Yorùbá to English and vice versa. However, feedback received from the Yoruba native speakers shows model limitations in cultural nuance (14% idiom errors). This setback will be addressed in the future. This research not only contributes to the field of natural language processing but also provides a practical solution for breaking down language barriers and promoting cross-cultural communication between Yorùbá and English speakers.

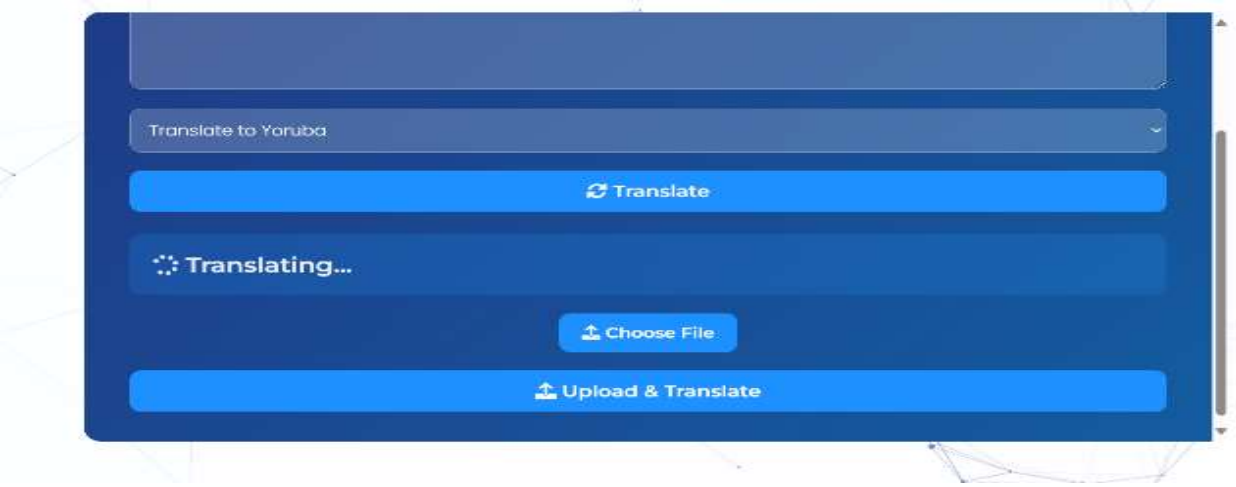


Figure 15: Sample translating document

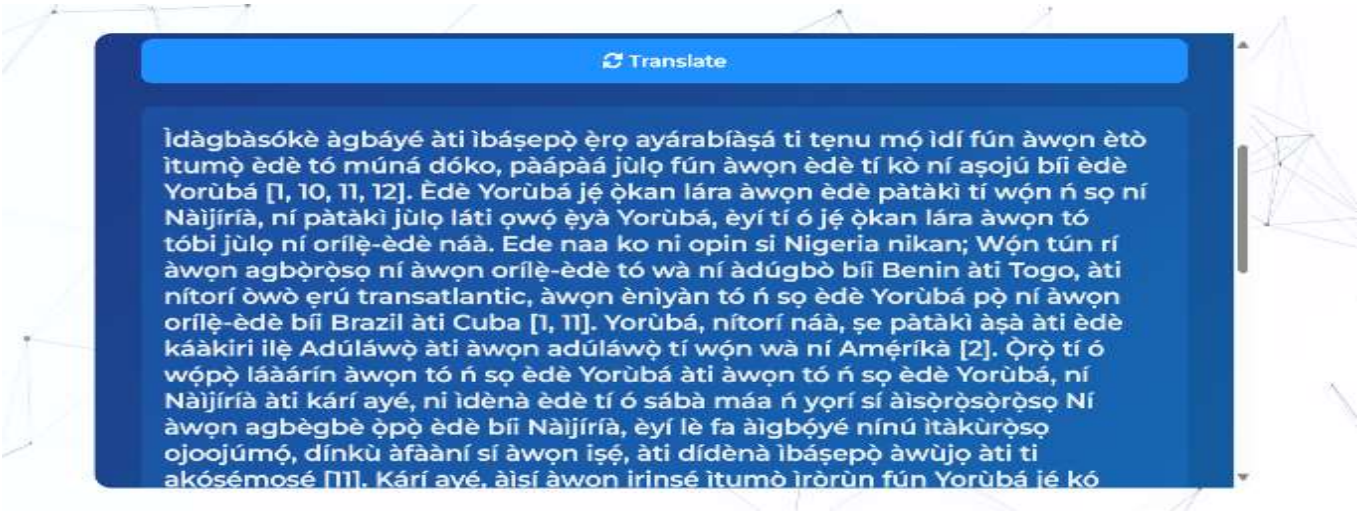


Figure 16: Sample translated document

5. Conclusion and future work

This paper advances Yorùbá-English machine translation by applying Bidirectional GRU achitecture that achives state-of-the-art scores (0.96), outperforming Transformer and LSTM baseline models in prior studies in tonal accuracy and training efficiency. By integrating HCI principles, diacritic-aware preprocessing and handling Yorùbá rich morphology and addressing tonal challenges, we brige critical gaps in low-resource NLP, offering a

© The Author(s), under exclusive license to SLUJST 2025

Online ISSN: 2736-0903
Print ISSN: 2736-089X
www.slujst.edu.ng

deployable tool for real-world use. We therefore suggest future improvements for this translation system in four key areas: intergrating support for different Yorùbá dialects to better understand regional variations. Likewise, incorperating voice input so users can speak rather than typing would improve HCI of the system. Improving how the model handles Yorùbá idioms and proverbs by working with language expert and training the system with more data to boost overal translation qaulity and accuracy.

References

Adebara, I., Abdul-Mageed, M., & Silfverberg, M. (2022). Linguistically-Motivated Yorùbá-English Machine Translation. In *Proceedings of the 29th International Conference on Computational Linguistics* (pp. 5066–5075). International Committee on Computational Linguistics.

Adebayo, T. (2024). Emerging grammars in contemporary Yorùbá phonology. *Canadian Journal of Linguistics/Revue canadienne de linguistique*, 69(2), 123–150.

Adeboje, O. T., Adetunmbi, O. A., Arome, G. J., & Akinyede, R. O. (2024). Bilingual neural machine translation from English to Yoruba using a transformer model. *International Journal of Innovative Science and Research Technology*, *9*(7).

Adewole, L. B., Adetunmbi, A. O., Alese, B. K., & Oluwadare, S. A. (2017). Token validation in automatic corpus gathering for Yoruba language. *FUOYE Journal of Engineering and Technology*, *2*(1), 4.

Ahia, O., Aremu, A., Abagyan, D., Gonen, H., Adelani, D. I., Abolade, D., Smith, N. A., & Tsvetkov, Y. (2024). Voices unheard: NLP resources and models for Yorùbá regional dialects. *Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing*, 4392–4409. https://doi.org/10.18653/v1/2024.emnlp-main.251

Akinade, I., Alabi, J. O., Adelani, D. I., Odoje, C., & Klakow, D. (2023). ε KÚ <MASK>: Integrating Yorùbá cultural greetings into machine translation. *Proceedings of the First Workshop on Cross-Cultural Considerations in NLP (C3NLP)*, Dubrovnik, Croatia. Association for Computational Linguistics.

Alabi, A. F., Olowookere, J. K., & Oludare, S. O. (2022). User experience in language translation systems: A focus on Yoruba language speakers. *International Journal of Human-Computer Interaction*, *38*(4), 321–335. https://doi.org/10.1080/10447318.2022.2039872

Akinola, O., Adegbite, J., & Akinola, J. A. (2023). Enhancing Yoruba language translation using hybrid deep learning architectures. *Journal of African Language Technologies*, *5*(2), 55–69. https://doi.org/10.1016/j.jalt.2023.02.005

Awan, A. A. (2021). Simple machine translation: Yoruba to English. *Towards Data Science*. https://towardsdatascience.com/simple-machine-translation-yoruba-to-english-1b958ccdc8a1/

Babatunde, A. N., Abikoye, C. O., Oloyede, A. A., Ogundokun, R. O., Oke, A. A., & Olawuyi, H. O. (2021). English to Yoruba short message service speech and text translator for Android phones. *International Journal of Speech Technology*, *24*(2), 979–991. https://link.springer.com/article/10.1007/s10772-021-09852-w

Babatunde, A. N., Balogun, B. F., Olatinwo, I. S., Oke, A. A., Ogundokun, R. O., & Sulaiman, H. O. (2024). Web-based French to Yoruba text translation using neural machine translation. *University of Pitesti Scientific Bulletin: Electronics and Computers Science*, *24*(2), 43–50.

Babatunde, A. N., Babatunde, R. S., Balogun, B. F., Oke, A. A., & Umar, E. (2024). Speech-to-text hybrid English to Yoruba SMS translator. *Innovative Computing Review*, *4*(1), 15–36.

Babatunde, A. N., Oke, A. A., Balogun, B. F., AbdulRahman, T. A., & Ogundokun, R. O. (2022). A deep neural network-based Yoruba intelligent chatbot system. *Journal of Digital Innovations & Contemporary Research in Science, Engineering & Technology*, *10*(2), 69–80. https://dx.doi.org/10.22624/AIMS/DIGITAL/V10N2P4

Babatunde, A. N., Ogundokun, R. O., Jimoh, E. R., Misra, S., & Singh, D. (2023). Hausa character recognition

© The Author(s), under exclusive license to SLUJST 2025

Online ISSN: 2736-0903
Print ISSN: 2736-089X
www.slujst.edu.ng

using logistic regression. In *Lecture Notes in Electrical Engineering* (Vol. 997, pp. 801–811). https://doi.org/10.1007/978-981-99-0085-5_65

Costa-Jussà, M. R., Farrús, M., Mariño, J. B., & Fonollosa, J. A. R. (2012). Study and comparison of rule-based and statistical Catalan-Spanish machine translation systems. *Computing and Informatics*, 31(2), 245–270. https://www.cai.sk/ojs/index.php/cai/article/view/940

Esan, A., Oladosu, J., Oyeleye, C., Adeyanju, I., Olaniyan, O., Okomba, N., & Adanigbo, O. (2020). Development of a recurrent neural network model for English to Yorùbá machine translation. *International Journal of Advanced Computer Science and Applications, 11*(5).

Costa-Jussa`, M., Marino˜ , F., & Fonollosa J. (2017). Study And Comparison Of Rule-Based and Statistical Catalan-Spanish Machine Translation Systems

Odoje, T. K., Adewole, K. S., & Salawu, R. A. (2014). Statistical machine translation approach for Yorùbá-English language pair. *Linguistics and Translation Studies Journal, 4*(1), 55–70. https://ltsj.online/wp-content/uploads/2020/01/LTS-2014-04-1-odoje.pdf

Odejobi, O. A., & Eludiora, S. I. (2016). Development of an English to Yorùbá Machine Translator. *International Journal of Modern Education and Computer Science*, 8(11), 8–19. https://doi.org/10.5815/ijmecs.2016.11.02

Ogunleye, O. A., Abiona, M. O., & Afolabi, O. (2021). Neural machine translation for Yoruba language using deep learning techniques. *Journal of Artificial Intelligence and Data Mining, 9*(1), 35–46. https://doi.org/10.5815/jaiad.2021.01.05

Oluwatoki, T. G., Adetunmbi, O. A., & Boyinbode, O. K. (2024). A transformer-based Yoruba to English machine translation (TYEMT) system with ROUGE score. *International Journal of Innovative Science and Research Technology, 9*(9). https://www.ijisrt.com/a-transformerbased-yoruba-to-english-machine-translation-tyemt-system-with-rouge-score

Oyekunle, A. S., & Adebayo, K. A. (2022). Human-computer interaction in language translation systems: A case study of Yorùbá language translator. *International Journal of Computer Applications, 182*(13), 22–30. https://doi.org/10.5120/ijca2022921988

Oyetade, S. O. (2001). Attitude to foreign languages and indigenous language use in Nigeria. *Language in Society*, *30*(4), 611–628.

Sharma, S., & Joshi, N. (2023). *A Transformer-based approach using LSTM and paraphrase reference to translate English text into Hindi*. Research Square. https://doi.org/10.21203/rs.3.rs-3230686/v1

Tian, T., Song, C., Ting, J., & Huang, H. (2022). A French-to-English machine translation model using Transformer network. *Procedia Computer Science, 199*, 1088–1094. https://doi.org/10.1016/j.procs.2022.01.182

Tian, Y., Zhao, Q., Kherroubi, Z. E. A., Boukhalfa, F., Wu, K., & Bader, F. (2022). Multimodal Transformers for Wireless Communications: A Case Study in Beam Prediction. *arXiv preprint arXiv:2309.11811*.

Torregrosa, D., Pasricha, N., Masoud, M., Chakravarthi, B. R., Alonso, J. A., Casas, N., & Arcan, M. (2019). Leveraging rule-based machine translation knowledge for under-resourced neural machine translation models. *Proceedings of Machine Translation Summit XVII: Translator, Project and User Tracks*, 125–133. European Association for Machine Translation. for Wireless Communications: A Case Study in Beam Prediction. *arXiv preprint arXiv:2309.11811*.

Vaswani, A., Shankar, S., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., & Polosukhin, I. (2017). Attention is all you need. *Advances in Neural Information Processing Systems, 30*. https://doi.org/10.5555/3295222.3295349

© The Author(s), under exclusive license to SLUJST 2025