International Journal of Artificial Intelegence Research
Vol. 9, No. 1, June 2025
ISSN 2579-7298

LSTM Model Using Adam's Optimizer for Indonesian – Bugis Bidirectional Translation System

Erliana Fajarwati [a,1,], Aji Prasetya Wibawa [a,2,*]

[a] Universitas Negeri Malang, Jl. Semarang No. 5, Malang 65145, Indonesia
[1] erliana.fajarwati.2105356@students.um.ac.id*; [2] aji.prasetya.ft@um.ac.id
* corresponding author

ARTICLE INFO

Article history
Received
Revised
Accepted

Keywords
Bugis
Indonesian
LSTM
Machine Translation
NLP

ABSTRACT

This research aims to develop a machine translation of Bugis to Indonesian and vice versa to preserve the Bugis language. This research utilizes a recent dataset of 30,000 Bugis-Indonesian sentence pairs from the online Bible. This research performs scraping to compile the corpus, followed by manual and automatic pre-processing. The method chosen is Neural Machine Translation (NMT), while Long Short-Term Memory (LSTM) is used for training and testing models. The Bilingual Evaluation Understudy (BLEU) score evaluates the model's performance to measure the accuracy of translation at various epochs. In addition, this study also compared the use of Adam's optimizer with a non-optimizer. The results showed that Adam's optimizer significantly improved the model's performance. At epoch 2000, the model achieved the highest BLEU score of 0.996261, indicating highly accurate translation quality. In contrast, the model without the optimizer showed lower performance. Other results also found that the translation from Bugis to Indonesian was more accurate than from Indonesian to Bugis. This is due to the more balanced word count difference in the Bugis to Indonesian translation, which makes it easier for the model to match words. In conclusion, using NMT with Adam optimizer effectively improves the accuracy of two-way translation from Bugis-Indonesian.

1. Introduction

A great deal of complexity is involved in translating local languages into Indonesian, especially when dealing with languages like Bugis that have complex grammatical structures. The Bugis language has a more complex vocabulary and syntax than Indonesian, making translation more difficult than in other regional languages. Translation from regional languages to Indonesian has been the subject of numerous studies; some examples are Javanese-Indonesian [1], [2], Sundanese-Indonesian [3], [4], [5], Madurese-Indonesian [6], [7], and Lampung-Indonesian [8], [9], [10]. In addition, it is noted that other language translations, for instance, English, have been comprehensively conducted from Indonesian [11], [12]. However, translating Bugis into Indonesian and Indonesian into Bugis is an area that still lacks the attention of researchers.

Still, as a part of her study on Bugis-Indonesian translation, one prominent example in Bugis-Indonesian translation shifts the focus into an application of Statistical Machine Translation (SMT) to translate Bugis Interlinear Literature into Indonesian where this study reports an accuracy of 16.342% [13] wherein fundamental differences between the target and source languages also complicate such structure. Similarly, studies that employed the SMT method in interlanguage from Bugis to Indonesian realized significantly positive translation results above 81.5% [14]. Although

SMT is beneficial for some purposes, it has drawbacks in translating long sentences, colloquial and rare expressions, and even the particulars of Bugis syntactical and morphological structures.

This study addresses these shortcomings by resorting to the NMT approach since it will provide the necessary assistance to achieve better translations specific to the Bugis-Indonesian language pair. NMT, on the other hand, can resolve long-range language dependencies and thus accomplish the realistic prediction of phrases with higher accuracy and fluency. Thanks to such an encoder-decoder structure, each sentence can be encoded into a respective context vector, which helps merge long and complex sentences in NMT. Hence, the adoption of NMT ought to significantly improve the translation of local languages by ensuring results that are more accurate and relevant within context than what was possible with earlier techniques.

Aside from enhancing translation accuracy, the NMT approach is also hoped to go a long way in preserving the Bugis language. There is already a decline in the use of Bugis, particularly among the youth who prefer Indonesian or other dominant languages in their social interactions. This decline is aggravated by sociological changes and urbanization, which create an ordinary sense among the youth that Bugis has less practical use in day-to-day activities. This research intends to preserve the Bugis language amidst globalization by designing a complete system of Bugis-Indonesian translations and vice versa.

2. Method

In this study's context, transferring meanings from Bugis to Indonesian and Indonesian to Bugis employs the LSTM-based Neural machine translation (NMT) approach. The evaluation criteria applied is the BLEU matrix. All the process steps are represented in Figure 1.

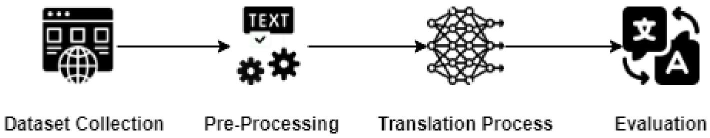


Fig. 1. Indonesian-Bugis Translation Process and Vice Versa

As can be seen in Figure 1, the Bugis to Indonesian and Indonesian to Bugis translation processes commence with Bugis and Indonesian compilation of datasets through scraping, followed by the pre- processing stage done automatically using the Python software application, which is then followed by Neural Machine Translation (NMT). Such translation method is used for Bugis to Indonesian and the reverse translation in the hope of getting a better translation. Finally, the model's performance is validated using the BLEU metric, which shows the degree of adequacy of the translation of target languages, such as Bugis and Indonesian.

2.1 Datasets

Data for this study was gathered from the Online Bible Site, which combines more than 30,000 Bugis and Indonesian sentence pairs. A web scraping method using the Web Scraper extension, an automatic data extraction tool for website pages, was employed to build this bibliography. Such a method makes it possible to obtain large amounts of data within a short time and cost-effectively, ensuring the dataset has all the required variations for the translation model to be trained. In Table 1, an illustrative example of a dataset that has been integrated consisting of Bugis and Indonesian languages from the chapter verses 1-5 output is demonstrated.

Table 1. Example Dataset

Bugis	Indonesian
“Makkedai PUWANGNGE lao ri Musa,”	“Kemudian, TUHAN berkata kepada Musa, firman-Nya,”
"Patettongngi Kéma-Ku ri tanggala séddi ulengséddi.”	“Pada hari pertama dalam bulan pertama, kamu harus mendirikan Tenda Suci, tenda pertemuan itu.”
“Puttama’i ri lalenna Petti Assijancingngé iya mallise’é Seppulo Parénta sibawa pasanni kaingpattongkoé ri yolona.”	“Kamu harus menempatkan tabut kesaksian itu di sanadan tudungilah tabut itu dengan tirai.”
“Taroi méjangngé sibawa passakke’na. Puttama towi ajé lampué sibawa pasangngi lampunna.”	“Kamu harus membawa masuk meja itu dan mengatur segala sesuatu yang harus diatur di atasnya, dan kamu juga

	harus membawa masuk kaki dian itu beserta pelita-pelitanya.”
“Palénne’i mézba ulaweng onrong mattunu dupaé ri yolona Petti Assijancingngé, namugattungngi kaing paddenringngé ri sumpanna Kémaé.”	“Kamu harus menempatkan mazbah emas untuk dupa itudi depan tabut kesaksian, lalu pasanglah tirai di pintu ke Tenda Suci.”

2.2 Pre-Processing

Dataset collection has been completed, and the next step is pre-processing, which aims to clean and prepare text data so that machine learning models can more easily process it. There are two methods used: manually and automatically. Manual pre-processing is done because there are differences in some verses in Bugis. Some verses in Bugis reference a previous verse due to the sentence length within the verse. Table 2 shows a difference in the scraping results for verse 2 of Surah Raja-Raja 1, where the Indonesian translation is not merged like the Bugis translation. In this research, there are two ways to do it. The first approach is to parse the Bugis translation if a verse contains more than three sentences, while the second combines them into a single verse if it contains fewer than three sentences.

Table 2. Comparison of Scraping Results

Indonesian	**Bugis**
2. “Inilah para pembesarnya: Azarya bin Zadok menjadiimam;”	2. “Iyanaé pajaba-pajaba tanré iya nakkaé Salomo: Imang-ngimang: Zadok, Azarya ana’ Zadok, Abyatar; Jurutulisi’na wanuwaé: Elihoréf sibawa Ahia, iyanaritu ana’-ana’na Sisia; Bendaharana wanuwaé: Yosafat ana’ Ahilud; Pallima tentara: Bénaya, ana’ Yoyada; Jennanna sining bupatié: Azarya, ana’ Natan; Pappangajana arungngé: Imang Zabud, ana’ Natan; Kapala ruma tanggana saorajaé: Ahisar; Kapalana pajjama rodié: Adoniram, ana’ Abda.”
3. “Elihoref dan Ahia, anak-anak Sisa menjadi paniteranegara; Yosafat bin Ahilud menjadi bendahara negara;”	3(4:2)
4. “Benaya bin Yoyada menjadi panglima; Zadok dan Abyatar menjadi imam.”	4(4:2)
Indonesian	Bugis
5. “**Azarya bin Natan mengawasi para kepala daerah; Zabut bin Natan, seorang imam, menjadi sahabat raja;**”	5(4:2)
6. “**Ahisar menjadi kepala istana; Adoniram bin Abda menjadi kepala rodi.**”	6(4:2)

Furthermore, automatic pre-processing using Python where the text is cleaned from irrelevant characters such as Tokenization, which is the process of breaking the text into small units, then Lower Casing converts uppercase letters to lowercase letters, and finally, Punctuation Removal which is the process of removing punctuation marks. This process is essential to ensure that the data used in training the model does not contain errors that can affect the translation results. An example can be seen in Table 3.

Table 3. Automatic Pre-Processing

Step	**Example Sentence**
Raw Sentence	”Engkalingai, umma’-Ku, maélo-Ka mabbicara, Israél,maélo-Ka mabbéré asabbing lao ri iko; Iyya’naé Allataala, Allataalamu.”
Tokenization	[”Engkalingai,” ”umma’-Ku,” ” maélo-Ka” ” mabbicara,”” Israél,” ” maélo-Ka” ” mabbéré” ”asabbing” ” lao” ” ri””iko;” ”Iyya’naé” ” Allataala,” ” Allataalamu.”]
Lowercasing	[”engkalingai,” ”umma’-ku,” ” maélo-ka” ” mabbicara,”” israél,” ” maélo-ka” ” mabbéré” ”asabbing” ” lao” ” ri””iko;” ”iyya’naé” ” allataala,” ” allataalamu.”]

Punctuation Removal	[”engkalingai,” ”ummaku,” ” maéloka” ” mabbicara,” ”israél” ” maéloka” ” mabbéré” ”asabbing” ” lao” ” ri” ”iko” ”iyyanaé” ” allataala” ” allataalamu”]

2.3 Translation Process

The machine-assisted translation process, neural machine translation, employs artificial intelligence to translate text from one language to another. NMT is end-to-end as it does not rely on many rule-based models like statistical machine translation (SMT) to perform its tasks because a neural network handles all tasks required for translation. In NMT, the source text is first passed through an encoder, which captures its meaning and translates it into a vector representation. Then, the decoder takes this representation and represents it as text written in the target language. NMT is very interesting due to its effectiveness in addressing the weaknesses of statistical models, especially with respect to long dependencies and even more intricate grammatical structures [16].

Neural Machine Translation performed by Long Short Term Memory (LSTM) neural network is one of the automatic translation methods that is quite popular because of its long-term solid memory capability [17]. The language model incorporates an encoder-decoder architecture that uses LSTM networks to allow the model to remember key concepts from the beginning and use them later in the sentence while generating individual words. LSTM for language translating NMT systems has proven helpful as it aids in representing semantic context, thereby improving translation quality. Transformative works were published at this junction, including works combining LSTM features with attention mechanism and transformer model which began gaining momentum in NMT in 2019 [18]. Therefore, parameter settings are done to obtain a more reliable picture of the results, as shown in Table 4.

Table 4. Parameter Setting

Number	Parameter	Value
1	Activation function	Softmax
Number	Parameter	Value
2	Epoch	100, 500, 1000, 2000
3	Batch Size	64
4	Dropout	0.2
5	Verbose	2
6	Optimizer	Non - Optimizer, Adam

2.4 Evaluation

The translation evaluation technique utilizes the Bilingual Evaluation Understudy (BLEU) approach, which is ranked among machine translation's most popular assessment tools. BLEU works by matching translated words with n-grams of the reference translation text. In addition to focusing on individual word (unigram) overlaps, BLEU also quantifies the degree of coverage of larger n-grams, such as bigrams, trigrams and four-grams (BLEU-4).

$$BP_{BLEU} = \{1, if\ c > r\ e^{1\frac{r}{c}}, if\ c \le r\} \quad (1)$$

$$P_n = \frac{\Sigma C \in corpus\ n-gram \in C\ \Sigma count\ clip^{(n-gram)}}{\Sigma C \in corpus\ n-gram \in C\ \Sigma count_{(n-gram)}} \quad (2)$$

$$BLEU = BP_{BLEU}\ e^{\Sigma_{n-1}^{N} W_n} \log \log p_n \quad (3)$$

A good BLEU score can only be achieved if the length of the translated sentence is comparable with that of the reference sentence. A more critical factor relates to the words: the translated sentence must have the same structural formations of words and their sequence as the reference [19]. This makes it possible for BLEU to impose a more significant disadvantage on word positioning in the context. This cold check leads to higher expectations in translatology. Equation (1) defines the BLEU formula to correlate image scores with the idea that scores express the proportion of the narrative segment's length. BP is the brevity penalty that assists in avoiding a bias which favors short translations as the quality of evaluation provides [20]. In this context, C represents the total number of words in the automated translation result, r represents the count of reference words, Pn refers to the adjusted precision score, Wn equals 1/ N with the standard value of N for BLEU being 4, and pn is the ratio of n-grams that match the reference translation to the total number of n-grams translated

3. Results and Discussion

This chapter will discuss the automatic translation model's performance from Bugis into Indonesian using the Bilingual Evaluation Understudy (BLEU) metric. The evaluation is done by comparing the BLEU scores obtained using the model through different epochs and optimization, specifically Adam and non-optimization. Each training scenario was subjected to five trials, and the average BLEU scores from these tests are presented in the table to ensure reliable results. Tables 5 and 6 provide the results for the Bugis – Indonesian and Indonesian – Bugis translations, respectively. Such tables are essential for evaluating the model's performance while simultaneously understanding the performance when subject to various factors. These outcomes are crucial in improving the translation model's efficiency and dependability.

Table 5. Bugis - Indonesian BLEU Results

Epoch	**Optimizer**	**BLEU 1**	**BLEU 2**	**BLEU 3**	**BLEU 4**
100	Non - Optimizer	0.1315684	0.0413288	0.0324946	0.0139508
500		0.6640464	0.5208988	0.4477474	0.3259126
Epoch	Optimizer	BLEU 1	BLEU 2	BLEU 3	BLEU 4
1000		0.9560474	0.9296268	0.9108834	0.8724642
1500		0.9940302	0.9911026	0.9894104	0.9847982
2000		0.9957000	0.9941394	0.9935520	0.9911304
100	Adam	0.1572318	0.0648332	0.0515580	0.0228378
500		0.7901396	0.6770774	0.6141508	0.5027650
1000		0.9851690	0.9771286	0.9712722	0.9583208
1500		0.9904526	0.9854116	0.9824768	0.9752026
2000		0.9962110	0.9946584	0.9941034	0.9917902

According to Table 5, the model gains the highest value of the best BLEU score at epoch 2000, with the highest BLEU score being BLEU 1, which is 0.9962110 for the model that uses the Adam optimizer. On the other hand, the lowest BLEU value is found at epoch 100, with a BLEU 4 value of 0.0139508 without using the optimizer. Table 5 also shows that the value of BLEU 1 is always higher than BLEU 2, BLEU 3, and BLEU 4 at each epoch and optimizer combination. This is reasonable since BLEU 1 only assesses single-word similarity between the translation and the reference. In contrast, higher BLEUs consider longer n-gram sequences, which are more difficult to match perfectly. Figure 2 compares the epoch graphs of the best model, i.e. epoch 2000, when using the Adam optimizer and without the optimizer.

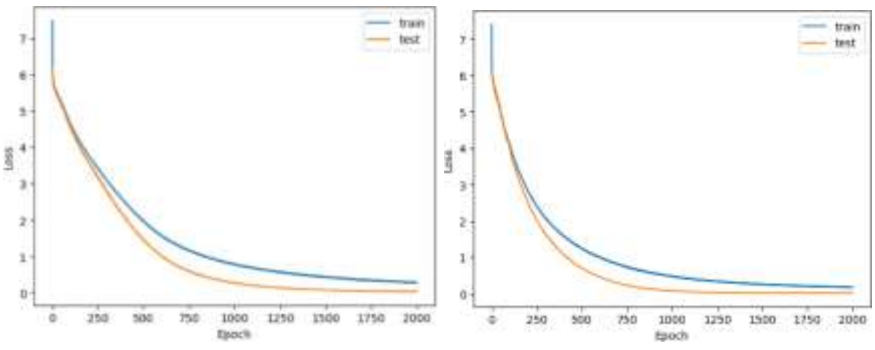


Fig. 2. Epoch graph (left) Non-optimizer, (right) Adam

Figure 2 shows two graphs comparing the model's performance at epoch 2000, both with and without the Adam optimizer. These two graphs show the best results achieved by the model, where the model achieves the most accurate translation performance at that epoch. Figure 2a shows the model's performance at epoch 2000 without using the optimizer. This graph shows that the model performs best; however, convergence occurs slowly or with more significant fluctuations before reaching stability. This is due to the lack of an optimization mechanism that usually helps the model reach the optimal point faster in the training process.

In contrast, Figure 2b shows the model's performance at epoch 2000 and the use of Adam's optimizer helps the model reach convergence faster and with less fluctuation in translation quality. Adam's optimizer works by dynamically adjusting the weight updates in the model based on the calculated gradient, allowing the model to learn the translation patterns more efficiently and stably. At epoch 2000, the graph shows that the model is already in an optimal state with no significant difference in translation quality between neighbouring epochs. For example, translations can be seen in Figure 3 to Figure 5.

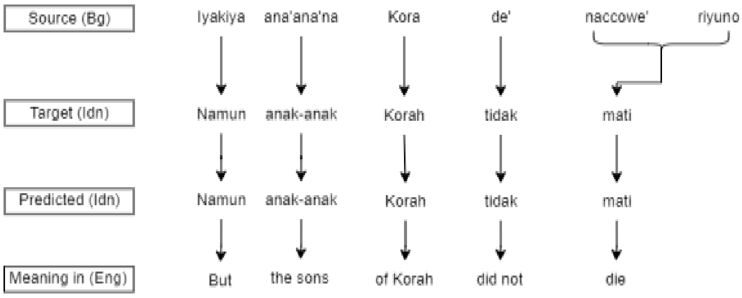


Fig. 3. Ideal Bugis - Indonesian translation result

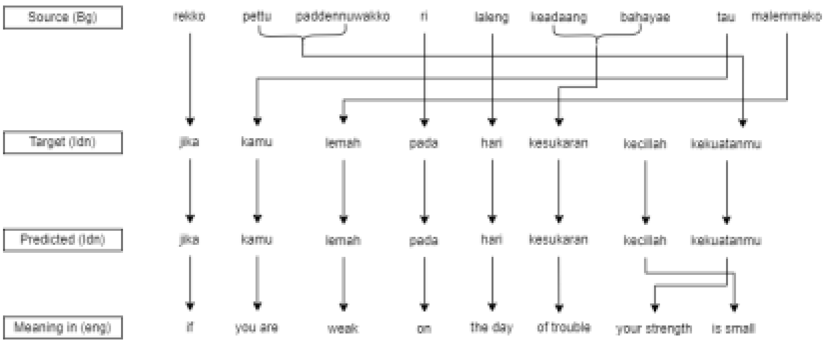


Fig. 4. Bugis - Indonesian translation results that are not ideal in word order

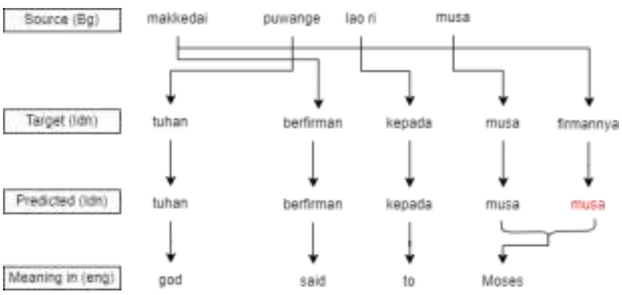


Fig. 5. Bugis - Indonesian translation results that are not ideal between target and predicted

In Figure 3, we can see an ideal example of the word-by-word translation process. Each word in the source Bugis language has its translation pair following the context and original meaning, except for the sentence narrower riyuno, which is translated into only one word, namely mati. Then, the target and the predicted show harmonious translation results, with the exact and accurate translation. This shows that the algorithm or translation model works well, providing reliable results that align with expectations. In contrast to the example sentences in

In Figure 4, we can see examples that are not ideal in the translation process, caused by the difference in word structure between Bugis and Indonesian. One example is in the sentence "rekko pettu paddennuwakko ri laleng keadaang dangere tau malemmako" If translated word by word, the sentence means "if your hope box is in danger of weak people". However, if the sentence is translated as a whole, the result is "if you are weak on the day of adversity, let down your strength", which is contextually correct and follows the target and prediction.

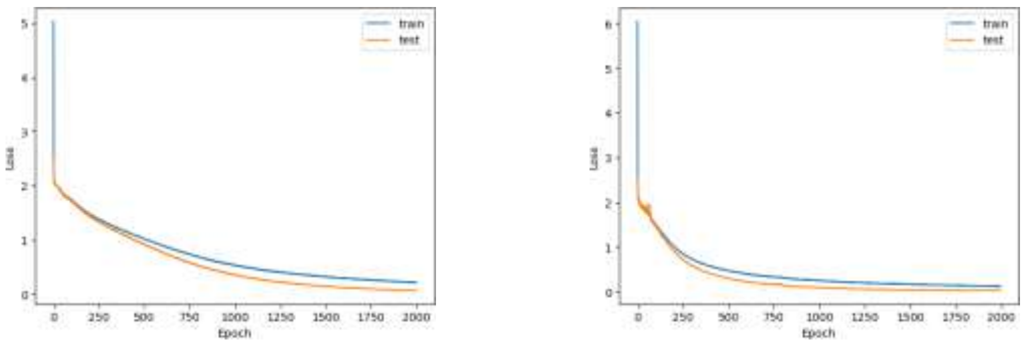
In contrast to Figure 5, the example sentence "makkedai puwange lao ri musa," when translated word-for-word, becomes "said the god went at musa", while the overall translation is "his god said to musa". The overall translation is similar to the predetermined target, and it is considered a correct translation from Bugis to Indonesian. However, it was found that there was a slight error between the target and predicted. This shows that although the model has been well-trained, there is room for improvement to capture more accurate meaning in complex translations.

Table 6. Indonesian – Bugis BLEU Results

Epoch	Optimizer	BLEU 1	BLEU 2	BLEU 3	BLEU 4
100	Non - Optimizer	0.0994370	0.0342514	0.0276688	0.0105202
500		0.3076380	0.1746502	0.1355458	0.0733792
1000		0.7645586	0.6488190	0.5863852	0.4729458
1500		0.9091258	0.8601788	0.8314568	0.7687746
2000		0.9613458	0.9391604	0.9269750	0.8988296
100	Adam	0.1023402	0.0373482	0.0289044	0.0106406
500		0.7246588	0.5921174	0.5246050	0.4057016
1000		0.9188474	0.8735880	0.8466356	0.7898370
1500		0.9682544	0.9526952	0.9436190	0.9236994
2000		0.9833640	0.9766722	0.9731812	0.9629886

Table 6 shows that the highest BLEU value is achieved at epoch 2000, with a peak value at BLEU 1 of 0.9833640 in the model using the Adam optimizer. In contrast, the lowest BLEU value occurs at epoch 100, where BLEU 4 only reaches 0.0105202 without using the optimizer. This shows that as epochs increase, the model can increasingly produce translations close to the reference translation, mainly when supported by the optimizer. In addition, Table 6 also indicates that the value of BLEU 1 is consistently higher than BLEU 2, BLEU 3, and BLEU 4 in every combination of epoch and optimizer. The difference in BLEU values between the Adam optimizer and without the optimizer is not too significant, with a difference of only about 0.04 in some scenarios. However, Adam's optimizer consistently resulted in slightly higher BLEU scores than without the optimizer. This suggests that Adam's optimizer helps the model achieve better translation results, although the gains may not be significant in every case.

Adam's optimizer contributed significantly to improving the model's performance, resulting in higher BLEU scores than without the optimizer, especially at BLEU 1 and BLEU 2. This suggests that using Adam's optimizer helps the model achieve more accurate and stable translation results. The effect of the optimizer is most pronounced at more straightforward n-gram levels, such as BLEU 1, where the model more easily matches individual words between the translation and the reference. Figure 6 shows a graphical comparison of the model's performance at the best epoch, epoch 2000, both when using Adam's optimizer and without it, providing a clear visualization of the effectiveness of using the optimizer in improving the quality of the translation produced by the model.



(a) **(b)**

Fig. 6. Epoch graph (a) Non-optimizer, (b) Adam

Figure 6 presents two graphs comparing the model's performance at epoch 2000, both with and without the Adam optimizer. Both graphs illustrate the best performance achieved by the model, where it achieved the highest accuracy in translation at epoch 2000. Figure 6a shows that although the model eventually achieves optimal performance in translation, the convergence process is slower, which means that stable and accurate translation performance is achieved with more difficulty.

In contrast, in Figure 6b, which shows the use of Adam's optimizer, the model can achieve convergence faster and with less fluctuation. Adam's optimizer uses a more adaptive and intelligent weight adjustment approach, allowing the model to learn more effectively from the training data. At epoch 2000, the graph shows that the model is already in an optimal state with a more consistent performance improvement from previous epochs. This confirms that Adam's optimizer speeds up the training process and helps the model achieve its best performance in less time and with higher stability. For an illustration of the translation of the model results, see Figure 7 to Figure 8.

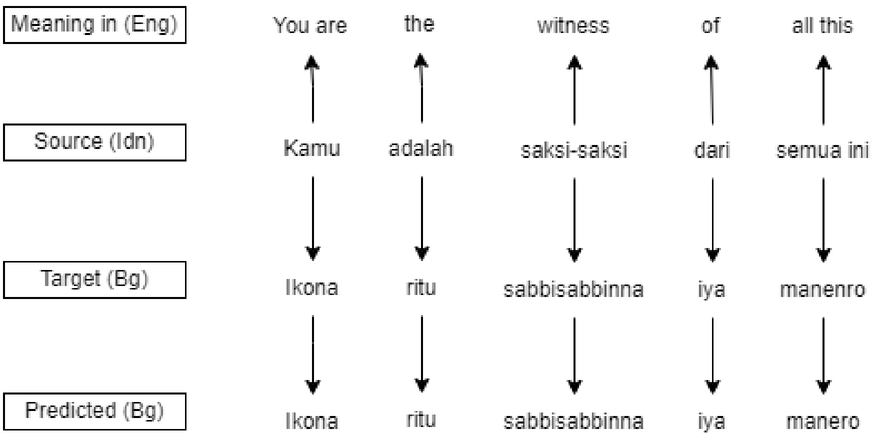


Fig. 7. Ideal Indonesian - Bugis translation result

Fig. 8. Indonesian - Bugis translation results that are not ideal in word order

In Figure 7, we can see an idealized example of the word-by-word translation process. Each Indonesian word has an appropriate Bugis translation pair that matches the original context and meaning. This result shows that the translation model works effectively, where both the target and predicted show perfect alignment, resulting in accurate and consistent translations.

In contrast, Figure 8 shows an example of a less-than-ideal translation caused by the difference in word count and word structure between Indonesian and Bugis. The number of words in Indonesian is only nine, while in Bugis, it is 24. The sentence "They devoted part of the spoils of war to repair the temple of God", when translated, means "Sibagiyang polé ri ammusurengngéro napakéi mennang untu' patettongngi Bolana PUWANGNGE". However, the target and predicted in the Bugis language Bible become "iyaro sining pabbere iyanaritu sibagiyang barang rappa iya nalolongengnge mennang ri laleng ammusurengnge sibawa iya nakkasuwiyangengnge mennang untu napake banna ri laleng bolana puwangnge". This is what affects the accuracy of the model in matching words.

The overall results of the existing translations (Indonesia - Bugis and Bugis - Indonesia) show that the utilization of the Neural Machine Translation (NMT) method with Long Short-Term Memory (LSTM) can result in a significant improvement in translation quality, as evidenced by the high BLEU scores in various data sets [21]. The BLEU scores show that BLEU 1 is higher than BLEU 2 to BLEU

4. This can occur because BLEU 1 only evaluates similarity at the individual word level without regard to the order or context of the words [22]. The model must only match one word at a time, which is statistically more straightforward than matching longer n-grams. As for BLEU 2 to BLEU 4, they provide a stricter assessment as they consider the order of the words in the translation, which is crucial in maintaining the sentence's overall meaning [23]. Besides, let us also consider the two additional aspects: the optimizer user and the epochs.

As evidenced in [24], however, Adam's optimizer is better at speeding up convergence or producing more stable convergence. Adam's optimizer is known to be able to change its learning quickly in response to the training gradient to improve the speed at which the model achieves more optimal results. Many studies conducted during the last five years show the same trend — the optimizer Adam outperforms the Stochastic Gradient Descent (SGD) optimizer in almost all models with vast and complex amounts of data [25].

The epoch counts varied and ranged from a few hundreds to about 2000 to test the model and confirm it was producing continuous and optimal results. It was found that the model achieved its best performance at epoch 2000. Epoch 2000 is considered the most optimal because the model has had enough time to learn complex language patterns and nuances, yet not too long to experience overfitting, where the model starts memorizing the training data instead of understanding the context in general. Studies from 2019 to 2024 show that choosing the correct number of epochs is crucial in model training, as too few epochs can lead to underfitting, while too many epochs can lead to overfitting [26].

On the other hand, the entire translation works better because it considers the context and structure of the sentence. This is supported by recent studies that focus on context understanding models - such models are critical in cross-language translation to improve accuracy and relevance in the target language [27].

This also stresses the limitations that arise during automatic translators' construction, particularly for languages whose structure is structurally distant [28]. Furthermore, there are still some gaps between the target and predicted results in the practice. This mismatch arises from translation models failing to account for several features and particular contexts of the language being translated [29]. Like most other languages, the Bugis language contains grammatical and idiomatic structures that are not easily captured in the models. However, subtleties such as sentence context, word order and some specific words tend to make the outcome of the translation models less desirable.

As students are taught, language and translation are not a mere stringing together of words, for every word in a translation must be qualified, thus achieving more accurate and quality results. Several scholars stress that successful translation converts sentences built in one particular language to sentences constructed in another language and can also convert the meaning by understanding the surrounding context and knowledge in a complex sentence form [30]. This also indicates that the translation algorithms and models require several modifications to address the issues regarding structural dissimilarities between different languages.

In addition, such a contextualized method makes it possible for the translation system to go beyond the constraints of literal translation by ensuring better output and adequate fluency in the output. This is crucial since some studies have pointed out a huge gap in the quality of translation from Bugis to Indonesian and Indonesian to Bugis. It was established that the translation from Bugis was much better than that from Indonesian. Even though the disparity is not exceedingly significant, it should be of great concern in designing translation systems in the future.

Also, the translation from Bugis into Indonesian achieved greater precision since the total number of words in both languages is reasonably correlated. Consequently, it is easier for the model to pair words and expressions accurately. Seen otherwise, in the translation from Indonesian to Bugis, such a substantial discrepancy in the number of words has been documented. According to what can be seen in Figure 8, since Bugis has many words, it is also a burden for the model to maintain coherence or accuracy during translation.

The findings in this study show that the NMT method proved effective in translating Bugis to Indonesian and vice versa. The LSTM model helps in capturing long-term dependencies and language complexity better. Coupled with the combination of NMT, Adam's Optimizer, and the correct number of epochs, it can improve the quality of Bugis-Indonesian translation.

4. Conclusion

Based on the research results, it was found that the NMT approach with LSTM proved effective in Bugis-Indonesian translation and vice versa. The LSTM encoder-decoder model better translates Bugis to Indonesian (BLEU 0.9962110) than Indonesian to Bugis (BLEU 0.9833640). This happens because the translation from Bugis to Indonesian has a relatively balanced number of words. In contrast, the translation from Indonesian to Bugis shows a more significant difference in the number of words. Bugis has a broader and more complex language structure and word count.

In addition, Adam's optimizer is proven to significantly improve performance compared to models that do not use an optimizer. That has significant epochs, namely at epoch 2000, also improved model performance. However, there are still weaknesses in this research where there is still one Indonesian sentence with more than one translation in Bugis. For future studies, researchers could utilize filtered data to ensure that unique sentences are included in the dataset without repetitions.

Acknowledgment

We want to express our deepest gratitude to several parties who have been involved in helping with this research.

Declarations

Funding statement. No external sources funded this research. The researchers assumed full responsibility for all costs associated with the study, ensuring complete independence and objectivity in the research process.

Conflict of interest. The authors declare no conflict of interest.

Additional information. No additional information is available for this paper

Data and Software Availability Statements

Data and Software availability statements provide a statement about where data and software supporting the results reported in a published article can be found, including hyperlinks to publicly archived datasets and software analyzed and generated during the study/experiments.

References

[1] H. Sujaini, “Improving the role of language model in statistical machine translation (Indonesian-Javanese),” Int. J. Electr. Comput. Eng. IJECE, vol. 10, no. 2, p. 2102, Apr. 2020, doi: 10.11591/ijece.v10i2.pp2102-2109.

[2] S. H. Wijono, K. Azizah, and W. Jatmiko, “Canonical Segmentation For Javanese- Indonesian Neural Machine Translation,” no. 4.

[3] R. B. Primandhika, M. N. Munawar, and A. R. Saifullah, “Experiment on a Transformer Model Indonesian-to-Sundanese Neural Machine Translation with Sundanese Speech Level Evaluation:,” presented at the Thirteenth Conference on Applied Linguistics (CONAPLIN 2020), Bandung, Indonesia, 2021. doi: 10.2991/assehr.k.210427.069.

[4] B. D. Wijanarko, Y. Heryadi, D. F. Murad, C. Tho, and K. Hashimoto, “Recurrent Neural Network-based Models as Bahasa Indonesia-Sundanese Language Neural Machine Translator,” in 2023 International Conference on Computer Science, Information Technology and Engineering (ICCoSITE), Jakarta, Indonesia: IEEE, Feb. 2023, pp. 951–956. doi: 10.1109/ICCoSITE57641.2023.10127817.

[5] Y. Heryadi, B. D. Wijanarko, D. Fitria Murad, C. Tho, and K. Hashimoto, “Indonesian-Sundanese Language Machine Translation using Bidirectional Long Short-term Memory Model,” in 2023 International Conference on Computer Science, Information Technology and Engineering (ICCoSITE), Jakarta, Indonesia: IEEE, Feb. 2023, pp. 945–950. doi: 10.1109/ICCoSITE57641.2023.10127691.

[6] D. A. Sulistyo, “LSTM-Based Machine Translation for Madurese-Indonesian,” J. Appl. Data Sci., vol. 4, no. 3, pp. 189–199, Sep. 2023, doi: 10.47738/jads.v4i3.113.

[7] F. H. Rachman, Qudsiyah, and F. Solihin, “Finite State Automata Approach for Text to Speech Translation System in Indonesian-Madurese Language,” J. Phys. Conf. Ser., vol. 1569, no. 2, p. 022091, Jul. 2020, doi: 10.1088/1742-6596/1569/2/022091.

[8] Z. Abidin, P. Permata, and F. Ariyani, “Translation of the Lampung Language Text Dialect of Nyo into the Indonesian Language with DMT and SMT Approach,” INTENSIF J. Ilm. Penelit. Dan Penerapan Teknol. Sist. Inf., vol. 5, no. 1, pp. 58–71, Feb. 2021, doi: 10.29407/intensif.v5i1.14670.

[9] Z. Abidin, Permata, I. Ahmad, and Rusliyawati, “Effect of mono corpus quantity on statistical machine translation Indonesian – Lampung dialect of nyo,” J. Phys. Conf. Ser., vol. 1751, no. 1, p. 012036, Jan. 2021, doi: 10.1088/1742-6596/1751/1/012036.

[10] J. Fakhrurozi, Z. Abidin, and F. Ariyani, “Prototype of Application Machine Translation Lampung Language Sentence of Api Dialect Dictionary Based Using Python Programming,” 2019.

[11] D. Puspitaningrum, "A Study of English-Indonesian Neural Machine Translation with Attention (Seq2Seq, ConvSeq2Seq, RNN, and MHA): A Comparative Study of NMT on English-Indonesian," in 6th International Conference on Sustainable Information Engineering and Technology 2021, Malang Indonesia: ACM, Sep. 2021, pp. 271–280. doi: 10.1145/3479645.3479703.
[12] T. W. Guntara, A. F. Aji, and R. E. Prasojo, "Benchmarking Multidomain English-Indonesian Machine Translation".
[13] M. Mulyana, H. Sujaini, and H. S. Pratiwi, "Algortima Division of Phrases in Sentence to Improve the Accuracy of Statistical Translation Machine of Indonesian Language - Bugis Wajo Language," J. Sist. Dan Teknol. Inf. JUSTIN, vol. 6, no. 2, p. 39, Apr. 2018, doi: 10.26418/justin.v6i2.23984.
[14] T. Apriani, H. Sujaini, and N. Safriadi, "The Effect Of Corpus Quantity On The Accuracy Of The Wajo Bugis Language Statist Translation Machine To Indonesia," 2016.
[15] N. Hamka and N. Imansari, "The Preservation of Bugis Language in Sinjai City: A Sociolinguistic Perspective," vol. 7, no. 1.
[16] F. Stahlberg, "Neural Machine Translation: A Review," J. Artif. Intell. Res., vol. 69, pp. 343–418, Oct. 2020, doi: 10.1613/jair.1.12007.
[17] H. Zhang, H. Huang, and H. Han, "Attention-Based Convolution Skip Bidirectional Long Short-Term Memory Network for Speech Emotion Recognition," IEEE Access, vol. 9, pp. 5332–5342, 2021, doi: 10.1109/ACCESS.2020.3047395.
[18] S. J. Reddi, S. Kale, and S. Kumar, "On the Convergence of Adam and Beyond," 2019, arXiv. doi: 10.48550/ARXIV.1904.09237
[19] R. Bawden, B. Zhang, L. Yankovskaya, A. Tättar, and M. Post, "A Study in Improving BLEU Reference Coverage with Diverse Automatic Paraphrasing," 2020, arXiv. doi: 10.48550/ARXIV.2004.14989
[20] N. Mathur, T. Baldwin, and T. Cohn, "Tangled up in BLEU: Reevaluating the Evaluation of Automatic Machine Translation Evaluation Metrics," 2020, arXiv. doi: 10.48550/ARXIV.2006.06264.
[21] S.-K. Im and K.-H. Chan, "Neural Machine Translation with CARU-Embedding Layer and CARU-Gated Attention Layer," Mathematics, vol. 12, no. 7, p. 997, Mar. 2024, doi: 10.3390/math12070997.
[22] F. Beken Fikri, K. Oflazer, and B. Yanikoglu, "Semantic Similarity Based Evaluation for Abstractive News Summarization," in Proceedings of the 1st Workshop on Natural Language Generation, Evaluation, and Metrics (GEM 2021), Online: Association for Computational Linguistics, 2021, pp. 24– 33. doi: 10.18653/v1/2021.gem-1.3.
[23] T. Glushkova, C. Zerva, and A. F. T. Martins, "BLEU Meets COMET: Combining Lexical and Neural Metrics Towards Robust Machine Translation Evaluation," 2023, arXiv. doi: 10.48550/ARXIV.2305.19144.
[24] M. Agarwal, K. S. Gill, S. Malhotra, and S. Devliyal, "Analysing Digits with Sequential Convolutional Neural Networks and Adam Optimizer," in 2024 IEEE International Conference on Information Technology, Electronics and Intelligent Communication Systems (ICITEICS), Bangalore, India: IEEE, Jun. 2024, pp. 1–4. doi: 10.1109/ICITEICS61368.2024.10625453.
[25] M. Agarwal, K. S. Gill, S. Malhotra, and S. Devliyali, "Exploring Numerical Analysis with Sequential Convolutional Neural Networks Leveraging Adam Optimization Technique," in 2024 International Conference on Innovations and Challenges in Emerging Technologies (ICICET), Nagpur, India: IEEE, Jun. 2024, pp. 1–5. doi: 10.1109/ICICET59348.2024.10616317
[26] W. Wang, Z. Fan, and S. H. Ng, "Trajectory-Based Multi-Objective Hyperparameter Optimization for Model Retraining," 2024, arXiv. doi: 10.48550/ARXIV.2405.15303.
[27] R. Appicharla, B. Gain, S. Pal, A. Ekbal, and P. Bhattacharyya, "A Case Study on Context-Aware Neural Machine Translation with Multi-Task Learning," 2024, arXiv. doi: 10.48550/ARXIV.2407.03076.
[28] M. Abrishami, M. J. Rashti, and M. Naderan, "Machine Translation Using Improved Attention-based Transformer with Hybrid Input," in 2020 6th International Conference on Web Research (ICWR), Tehran, Iran: IEEE, Apr. 2020, pp. 52–57. doi: 10.1109/ICWR49608.2020.9122317.
[29] S. Kumar, A. Anastasopoulos, S. Wintner, and Y. Tsvetkov, "Machine Translation into Low-resource Language Varieties," in Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 2: Short Papers), Online: Association for Computational Linguistics, 2021, pp. 110–121. doi: 10.18653/v1/2021.acl-short.16.

[30] A. Popescu-Belis, "Context in Neural Machine Translation: A Review of Models and Evaluations," Jan. 25, 2019, arXiv: arXiv:1901.09115. Accessed: Sep. 27, 2024. [Online]. Available: http://arxiv.org/abs/1901.09115