INTENSIF, Vol.5 No.1 February 2021
ISSN: 2580-409X (Print) / 2549-6824 (Online)
DOI: https://doi.org/10.29407/intensif.v5i1.14670

Translation of the Lampung Language Text Dialect of Nyo into the Indonesian Language with DMT and SMT Approach

Penerjemahan Teks Bahasa Lampung Dialek Nyo ke Bahasa Indonesia dengan Pendekatan DMT dan SMT

Received:
12 August 2020
Accepted:
28 December 2020
Published:
1 February 2021

[1]*Zaenal Abidin, [2]Permata, [3]Farida Ariyani
[1]Sistem Informasi, Universitas Teknokrat Indonesia
[2]Informatika, Universitas Teknokrat Indonesia
[3]Magister Pendidikan Bahasa dan Kebudayaan Lampung, Universitas Lampung
[1,2,3]Lampung, Indonesia
E-mail: [1]zabin@teknokrat.ac.id, [2]permata@teknokrat.ac.id, [3]farida.ariyani@fkip.unila.ac.id
*Corresponding Author

Abstract—Research on the translation of Lampung language text dialect of Nyo into Indonesian is done with two approaches, namely Direct Machine Translation (DMT) and Statistical Machine Translation (SMT). This research experiment was conducted as a preliminary effort in helping students immigrants in the province of Lampung, translating the Lampung language dialect of Nyo through prototypes or models was built. In the DMT approach, the dictionary is used as the primary tool. In contrast, in SMT, the parallel corpus of Lampung Nyo and Indonesian language is used to make language models and translation models using Moses Decoder. The result of text translation accuracy with the DMT approach is 39.32%, and for the SMT approach is 59.85%. Both approaches use Bilingual Evaluation Understudy (BLEU) assessment.

Keyword—Lampung language dialect of Nyo, Direct machine translation, Statistical machine translation, Bilingual Evaluation Understudy

Abstrak—*Penelitian penerjemahan teks bahasa Lampung dialek Nyo ke bahasa Indonesia dilakukan dengan dua pendekatan yaitu Direct Machine Translation (DMT) dan Statistical Machine Translation (SMT). Eksperimen penelitian ini dilakukan sebagai upaya awal dalam membantu siswa/i para pendatang di provinsi Lampung dalam menerjemahkan bahasa Lampung dialek Nyo melalui purwarupa atau model yang dibangun. Pada pendekatan DMT digunakan kamus sebagai alat bantu utama sedangkan pada SMT digunakan korpus paralel bahasa Lampung Nyo dan bahasa Indonesia sebagai bahan dalam membuat model bahasa dan model translasi menggunakan Moses Decoder. Hasil akurasi penerjemahan teks dengan pendekatan DMT yaitu 39.32% dan untuk pendekatan SMT yaitu 59.85 %. Kedua pendekatan menggunakan penilaian Bilingual Evalution Understudy (BLEU).*

Kata Kunci—*Bahasa Lampung dialek Nyo, Direct machine translation, Statistical machine translation, Bilingual Evaluation Understudy*

ISSN: 2580-409X (Print) / 2549-6824 (Online)
DOI: https://doi.org/10.29407/intensif.v5i1.14670

I. INTRODUCTION

Lampung Province is a province located at the entrance gate to the island of Sumatra. Lampung Province has a wealth of culture, one of which is the Lampung language and Lampung script. In general, in Lampung province, there are two main dialects, namely the fire dialect and the Nyo dialect. The Lampung provincial government has great concern for the Lampung language. The provincial government continues to make various efforts to preserve and maintain the Lampung language. The Government of Lampung, through Governor Regulation number 39 of 2014 concerning Lampung Language and Script Subjects, stipulates that the Lampung language is a mandatory local content at the primary to senior secondary education unit levels and is supported by the availability of textbooks ranging from elementary, junior high and high school, along with the Lampung language dictionary. The Lampung language, both the fire dialect and the Nyo dialect, is used by the people of Lampung to communicate daily both in the family environment and at formal events. The Lampung language belongs to the Austronesian class in the Polynesian Malay language family. The two main dialects are dialect A (api) and dialect O (Nyo), which refers to the word 'Apa' [1].

For immigrants who send their children to school in Lampung province, one of the subjects learned at the SD, SMP, and SMK / SMA is the Lampung language. Parents of immigrants indeed find it challenging to help their children learn the Lampung language because it is not the local language itself. In response to this, academics at the Technocrat University of Indonesia and the University of Lampung are trying to find a solution. Through this research, it is expected to try to provide an initial solution to solve this problem by making a prototype with the Python programming language or the Lampung language translator model, especially for the Nyo dialect. There are two approaches to build this solution, namely Direct Machine Translation (DMT) and Statistical Machine Translation (SMT). In this study, only careful observations were made of the Lampung language, Nyo dialect.

The way to translate the Lampung language text Nyo dialect can be done by using a dictionary. In this way, it will be tiring both for parents or students because they repeatedly see the words that need to be searched in the Lampung language dictionary. Research on translating the Lampung language text in the Nyo dialect has never been carried out on a dictionary basis. The approach to building machine translation can be made in three approaches, namely (1) direct approach or DMT using a dictionary, (2) a rule-based approach or Rule-based Machine Translation (RBMT) using a series of rules in the language, and (3) a data-driven approach that uses a parallel corpus [2].

ISSN: 2580-409X (Print) / 2549-6824 (Online)
DOI: https://doi.org/10.29407/intensif.v5i1.14670

In the DMT research for the Lampung language, the Nyo dialect, the main component needed is a bi-dictionary Indonesian Lampung language. The construction of a translating machine with a rule-based approach will require rules for analyzing sentences in the source language, rules for transforming the representation of the source language analysis results, and rules for generating sentences in the destination language. The construction of a translation machine using a data-driven or parallel corpus-based approach requires sentence pairs between the original language and the destination language [2]. Research on the translation of the Lampung language in the Api dialect has been carried out, using a parallel corpus in the form of 3000 Lampung language sentence pairs and their translation in Indonesian, using the Neural Machine Translation (NMT) method without the Attention mechanism [3] and the Neural Machine Translation (NMT) method with the Attention mechanism [4]. Statistical Machine Translation (SMT) research using a parallel corpus in the form of 3000 sentence pairs in the Lampung Dialek Api language and their translation in Indonesian has been carried out [5]. In this study, the dictionary is used as a database to build the DMT and a parallel corpus of the Lampung language Nyo dialect. Its translation in Indonesian is used to build the SMT model. Meanwhile, the Lampung language research from the speech research aspect was carried out for the first time [6].

The studies related to DMT developed are research conducted in India [7], where this research is based on a dictionary from Kannada to Telugu. The study was conducted in Sri Lanka, where the study was based on a dictionary from Pali to Sinhala [8]. Research conducted in Indonesia is based on a dictionary from Indonesian to Balinese using Android [9]. The research results are in the form of applications that can be installed on an Android smartphone. Research conducted in Indonesia is based on a dictionary from Indonesian to Javanese using mobile [10]. As for SMT, research on machine translation in Indonesia has been carried out by researchers including translation of Javanese and Indonesian with phrase-based SMT [11]. Research on translation of Sundanese into Indonesian using phrase-based SMT and utilizing the part of speech (PoS) Tag [12], Indonesian-Dayak Taman translation research with root word markings and affixes was carried out at the University of Tanjungpura [13], investigative research on the role of language models in the Indonesian-Dayak Kanayatn SMT research [14].

Research on the effect of corpus quantity on Bugis language SMT research Wajo into Indonesian [15], research on various models of translating Indonesian into Japanese has also been carried out [16], research on the measurement of translation results produced by machine translators using the Bilingual Evalution Understudy (BLEU) score calculation [17], observing the morphological aspects of language Lampung has been carried out by Lampung language researchers [1], while the references are few SMT research taken explicitly from the SINTA 2 accredited journal, namely efforts to improve the accuracy of the machine translator statistics in

ISSN: 2580-409X (Print) / 2549-6824 (Online)
DOI: https://doi.org/10.29407/intensif.v5i1.14670

Javanese to Indonesian with a lexical model probability improvement approach [18], experiments using Pivot Language SMT from English to Malay Sambas [19], influence research Dictionary lookup method on corpus cleaning on the accuracy of Indonesian-Malay Pontianak SMTs [20], observing the effect of increasing the accuracy of Indonesian-Minang SMT using the EWSB algorithm [21], and comparative research on the accuracy value of the smooting algorithm at SMT Indonesia - Melayu Sambas with the IRSLTM Language model toolkit [22].

By making a prototype application for translating the Lampung language, Nyo dialect, which maintains the Lampung language dictionary and the SMT model in Lampung, the Nyo dialect is expected to be the first way to help immigrant students in translating the Lampung language Nyo dialect. The dictionary acts as a database in making a prototype for translating the Lampung language Nyo dialect. The prototype was made using the Python programming language as a programming language that is reliable in processing data in the form of text and is open source for its use. Meanwhile, in building the SMT model, the parallel corpus of the Lampung language Nyo dialect and its translation in Indonesian plays a role as a raw material in making translation models and language models in SMT. This study aims to make a DMT prototype and an SMT model that can translate paragraph texts in Lampung dialect language that can translate paragraph texts in Lampung dialect of Nyo and analyze the translation results with Bilingual Evaluation Understudy (BLEU).

II. RESEARCH METHOD

A. The Subject / Material Studied

The subjects/materials in this study were the Lampung language dictionary Nyo dialect and the parallel corpus Lampung language Nyo dialect to Indonesian. The dictionary used is the Lampung language dictionary by Herman, S.Pd.I. The parallel corpus of the Lampung language in Nyo dialect and its translation in Indonesian was built through manual typing through the notepad application media. The subject/material researched can be accessed in detail at the link https://bit.ly/37YBNDe.

B. Design / Design made

The research stages used in the DMT research in the Lampung language, Nyo dialect to Indonesian, are illu image below.

INTENSIF, Vol.5 No.1 February 2021
ISSN: 2580-409X (Print) / 2549-6824 (Online)
DOI: https://doi.org/10.29407/intensif.v5i1.14670

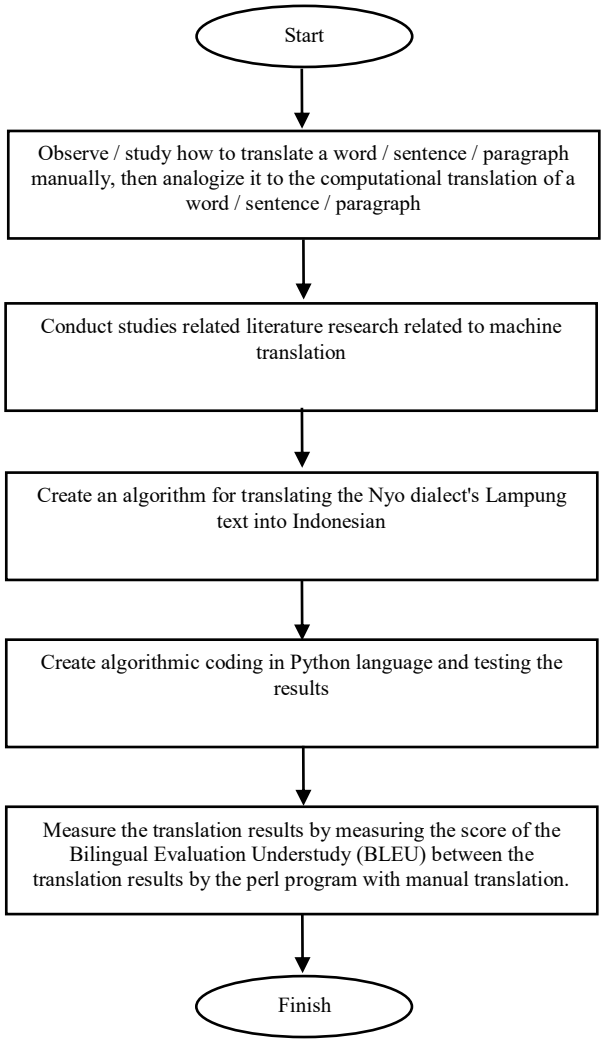

Figure 1. STAGES IN DEVELOPING DMT IN LAMPUNG LANGUAGE

The DMT algorithm for translating Lampung Nyo into Indonesian is made as steps to translate a word/sentence/paragraph from Lampung Nyo into Indonesian:

1. The DMT algorithm for translating Lampung Nyo into Indonesian is made as steps for translating 1. Enter a word/sentence/paragraph into the machine translator.
2. The machine translator reads a word/sentence/paragraph: (1) converts all words in a word/sentence/paragraph into lowercase letters, (2) separates all words and symbols in a word/sentence/paragraph This is done by tokenizing the paragraph based on spaces (3) Putting all the words and symbols obtained in a List of words and symbols.
3. The machine translator will take, one by one, the contents of the word and symbol list to be matched sequentially in the database provided.
4. If the contents of the word and Symbol list are found, then the results are placed in the *HASIL* or result list. If not found, then the word or symbol will be displayed again

ISSN: 2580-409X (Print) / 2549-6824 (Online)
DOI: https://doi.org/10.29407/intensif.v5i1.14670

according to the contents of the word and Symbol list and also placed in the *HASIL* or RESULTS list in the order it came.

5. Reorder the *HASIL* or RESULTS list in a new variable named OUTPUT OF *KATA* or WORD / *KALIMAT* or SENTENCE / *PARAGRAF* or PARAGRAPH

The research stages used in the research of SMT Lampung language Nyo dialect to Indonesian are shown in Figure 2 below. The implementation of the syntax used by SMT in Moses Decoder can be seen at this link https://bit.ly/37YBNDe.

C. Variables measured in translation

The research of the algorithm is for translating the text of a word/sentence/paragraph in Lampung from the Nyo dialect into Indonesian. The variables that want to be observed are the translation results by the application prototype made using the Python programming language and the translation results by the model generated from the Moses Decoder.

The SMT pre-processing phase in the Moses Decoder consists of sentence alignment, tokenization, cleaning, lowercase filtering, and actual case. Sentence alignment aligns the parallel corpus of the Nyo dialect with Indonesian as its translation. Tokenization is needed to provide spacing between words, including spacing between words and existing punctuation marks, while lowercase is a process to uniform the letters' case. In this proper casing process, each beginning of each sentence is converted to the most likely place. Cleaning is the process of limiting sentence length. Cleaning also functions to remove misaligned sentences. The next phase is the training phase. It is in this phase that the language model and model translation are carried out. Language model using software, in this study used KenLM, which has been integrated into Moses.

INTENSIF, Vol.5 No.1 February 2021
ISSN: 2580-409X (Print) / 2549-6824 (Online)
DOI: https://doi.org/10.29407/intensif.v5i1.14670

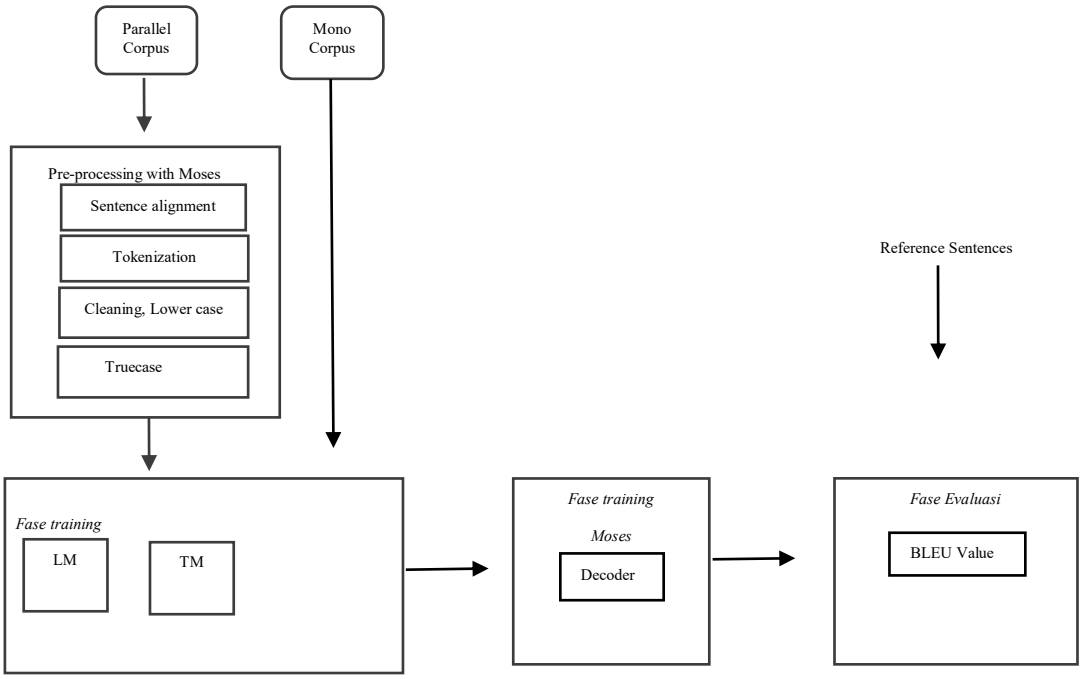

Figure 2. ARCHITECTURE IN SMT BUILDING

D. Data collection techniques for translation trials

The research on the algorithm is for translating the text of a word/sentence/paragraph from Lampung from the Nyo dialect to Indonesian. The data collection technique scenario for testing DMT and SMT was through a random selection of sentences in the Lampung language Nyo dialect, which had been translated by native speakers of the Nyo dialect. Details of the test sentences are provided at the following link https://bit.ly/37YBNDe.

E. The analysis assesses the translation results

Evaluation of translation results is done by comparing the translated sentences with the reference sentences using the Bilingual Evaluation Understudy (BLEU) application available on the Moses Decoder. BLEU is an algorithm aimed to evaluate the quality of the translation results that have been translated by a machine from a source language to a destination language. BLEU measures the modified statistical precision score between the translation results automatically and the reference translation using a constant called the brevity penalty (BP) [17].

INTENSIF, Vol.5 No.1 February 2021
ISSN: 2580-409X (Print) / 2549-6824 (Online)
DOI: https://doi.org/10.29407/intensif.v5i1.14670

III. RESULT AND DISCUSSION

Information on the accuracy of the results obtained from translating the Lampung language Nyo dialect into Indonesian using the DMT and SMT approaches is given in table 1.

Table 1. BLEU RESULTS ACCURACY VALUE IN DMT AND SMT

Translation Results	BLEU Value (%)	
	DMT	***SMT***
Lampung language Nyo dialect to Indonesian	39.32	59.85

Table 1 shows that the translation results from the Lampung language Nyo dialect to Indonesian. In which the SMT approach gives higher translation accuracy results than the DMT approach in the test material used in the form of 25 single sentences in the Nyo dialect Lampung language because the SMT approach can learn from training data available via parallel corpus and mono corpus are used.

A. DMT Testing Result

Translation of words or sentences, or paragraphs from the Lampung language Nyo dialect can be done through a prototype that has been made using Python 2.7 in the form of a console. The prototype test for the Lampung language translator application - Indonesian was carried out by using more than one single sentence in the Lampung language, Nyo dialect. A list of twenty-five test sentences is given at the https://bit.ly/37YBNDe.

The results of translation, as shown in Figure 3 below, from the prototype application show that the application can translate, as stated in the application's database. If there are words that are not in the database, it shows that the application will give results in the initial/original word. The accuracy obtained through the DMT approach in testing the 25 test sentences obtained a value of 39.32%, meaning that the DMT application is only able to translate words that are already in its database. The use of the dictionary is considered insufficient due to the limitations of the words contained in the Lampung language dictionary by Herman, S.Pd.I.

INTENSIF, Vol.5 No.1 February 2021
ISSN: 2580-409X (Print) / 2549-6824 (Online)
DOI: https://doi.org/10.29407/intensif.v5i1.14670

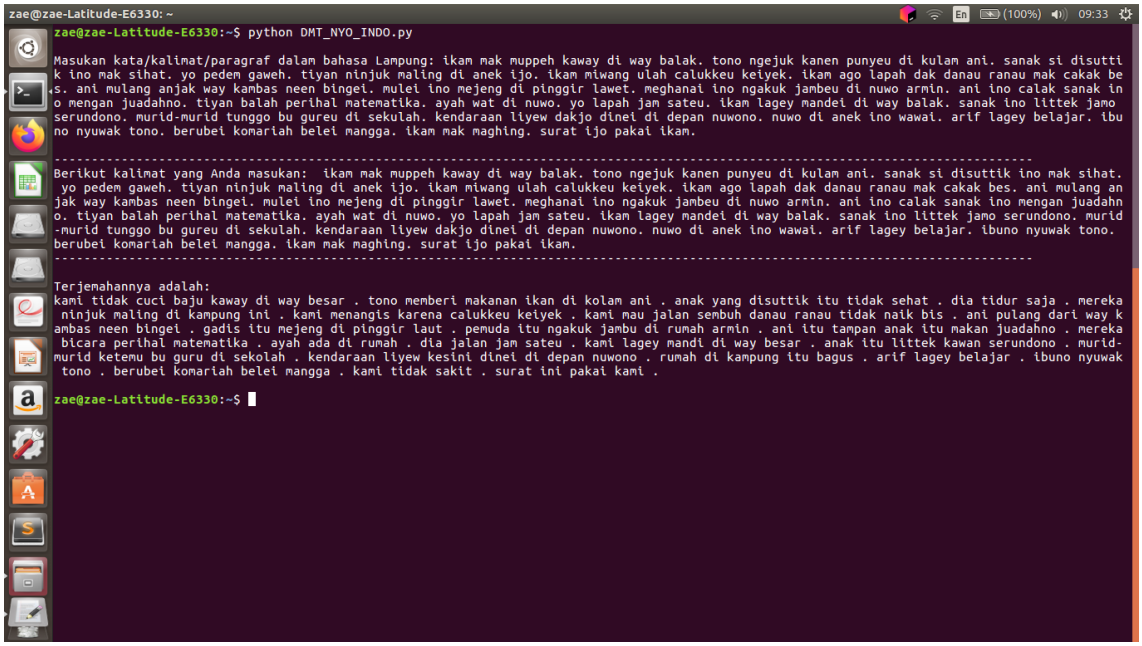

Figure 3. RESULTS OF PURWARUPA TRANSLATING LAMPUNG WITH NYO DIALECT

Figure 3 shows the results of the translation that can be done by the translator application prototype. The application is only able to translate words from the Lampung language Nyo dialect that have been found in the application database. The application only displays the input in the application as for symbols other than words and other words that are not in the database. It is proven that the dictionary used in this study still lacks words that should be present.

B. SMT TESTING RESULT

The implementation of SMT on the Moses Decoder, for the translation experiment from the Lampung language Nyo dialect to Indonesian, can be seen in detail at the link https://bit.ly/37YBNDe. Various steps have been taken to produce a language model and a translation model that can be used to translate the Lampung language, Nyo dialect. The training data used in this study were a number of the parallel corpus as many as 4057 sentence pairs from the Lampung language Nyo dialect to Indonesian and 13759 mono corpora Indonesian. After the data training was carried out, the testing was carried out using twenty-five test sentences in the Nyo Lampung dialect made by speakers of the Nyo dialect Lampung language. Twenty-five test sentences are given in full on the https://bit.ly/37YBNDe.

The results of testing sentences in the Lampung language, Nyo dialect with SMT, get an accuracy value of 59.85%, as shown in Table 1 above. The main factor for translating sentences from the Lampung language in Nyo dialect into Indonesian is the compatibility of the trigram or bigram or unigram data in the test sentence with the trigram or bigram or unigram data from the results of the training data. The test results in the test sentence show the BLEU number of

ISSN: 2580-409X (Print) / 2549-6824 (Online)
DOI: https://doi.org/10.29407/intensif.v5i1.14670

59.85%. In general, the facts found from the translation of the Lampung language in the Nyo dialect are as follows:

1. SMT can give the same results as the translation results by speakers of the Lampung language Nyo dialect.
2. SMT can provide results that have the same meaning or have different meanings, such as the results of translation by speakers of the Lampung language in Nyo dialect.
3. SMT gives results with different word structures as translated by speakers of the Lampung language, Nyo dialect.
4. SMT gives results that there are still words in the Lampung language Nyo dialect which are not translated into Indonesian.

Table 2. EXAMPLE OF SMT TRANSLATION RESULTS ACCORDING TO THE FIRST FACT

Sentences in Lampung Language	Reference Meaning in Indonesian	Results of SMT Lampung in Nyo dialect
nuwo di anek ino wawai	rumah di kampung itu bagus	rumah di kampung itu bagus

5. Table 2 above shows the SMT results being able to translate the Lampung language sentence according to the reference sentence well. SMT produces translations that match the meaning of reference sentences in Indonesian. Of the twenty test sentences used, 14 test sentences were obtained whose SMT translation gave the same results as the reference sentences given.

Table 3. RESULTS OF SMT TRANSLATION ACCORDING TO SECOND FACT

Sentences in Lampung Language	Reference Meaning in Indonesian	Results of SMT Lampung in Nyo dialect
tiyan balah perihal matematika	mereka **bicara tentang** matematika	mereka **berkata perihal** matematika
sanak si disuttik ino mak sihat	anak yang disuntik itu **tidak** sehat	anak yang disuntik itu **tetap** sehat

6. Table 3 above shows that SMT can translate Nyo's Lampung language sentences even though it gives different results but has the same meaning. SMT can translate Nyo's Lampung language sentences even though it gives different results. The word "bicara tentang" has the same meaning as "berkata perihal," while the word "tidak" has a different meaning from the word "tetap." The results of the SMT that produce the second fact are 3 out of 25 test sentences.

INTENSIF, Vol.5 No.1 February 2021
ISSN: 2580-409X (Print) / 2549-6824 (Online)
DOI: https://doi.org/10.29407/intensif.v5i1.14670

Table 4. RESULTS OF SMT TRANSLATION ACCORDING TO THIRD FACTS

Sentences in Lampung Language	Reference Meaning in Indonesian	Results of SMT Lampung in Nyo dialect
tiyan ninjuk maling di anek ijo	mereka menangkap maling di dusun ini	mereka menangkap di dusun maling ini

7. Table 4 above shows that SMT can translate sentences in the Lampung Nyo language even though it provides a different word structure from the reference sentence to give a different meaning. The results of the SMT that produce like the third fact are 3 out of 25 test sentences.

Table 5. RESULTS OF SMT TRANSLATION ACCORDING TO FOURTH FACT

Sentences in Lampung Language	Reference Meaning in Indonesian	Results of SMT Lampung in Nyo dialect
sanak ino littek jamo serundono	anak itu bertengkar dengan temannya	anak itu bermusuhan sama serundono

8. Table 5 above shows an example of the translation results of SMT in the Lampung language, Nyo dialect, according to the fourth fact. SMT failed in translating the word 'serundono' because the parallel corpus did not exist. Another thing that is implied from the above example is that SMT gives different translation results on 'littek.' The SMT results that produce like the third fact are five sentences out of 25 test sentences. The complete results on testing the translation of the sentence text from the Lampung language Nyo dialect to Indonesian are given in table 6.

Table 6. RESULTS OF THE TRANSLATION OF DMT AND SMT IN 25 TEST SENTENCES

Sentences in Lampung Language	Reference Meaning in Indonesian	Results of DMT Lampung dialect of Nyo	Results of SMT Lampung in Nyo dialect
ikam mak muppeh kaway di way balak	saya tidak mencuci baju di sungai	kami tidak mencuci kaway di way besar	saya tidak kaway mencuci kain di sungai
tono ngejuk kanen punyeu di kulam ani	tono memberi makanan ikan di kolam ani	tono memberi makanan ikan di kolam ani	tono memberi makanan ikan di kolam ani
sanak si disuttik ino mak sihat	anak yang disuntik itu tidak sehat	anak yang disuttik itu tidak sehat	anak yang disuntik itu tetap sehat
yo pedem gaweh	dia tidur saja	dia tidur saja	dia tidur saja
tiyan ninjuk maling di anek ijo	mereka menangkap maling di dusun ini	mereka ninjuk maling di kampung ini	mereka menangkap di dusun maling ini

ISSN: 2580-409X (Print) / 2549-6824 (Online)
DOI: https://doi.org/10.29407/intensif.v5i1.14670

Table 6. RESULTS OF THE TRANSLATION OF DMT AND SMT IN 25 TEST SENTENCES [CONTINUE]

Sentences in Lampung Language	Reference Meaning in Indonesian	Results of DMT Lampung dialect of Nyo	Results of SMT Lampung in Nyo dialect
ikam miwang ulah calukkeu keiyek	hamba menangis karena kakiku terinjak	kami menangis karena calukkeu keiyek	aku menangisi keiyek karena kakiku
ikam ago lapah dak danau ranau mak cakak bes	kami mau pergi ke danau ranau tidak naik bus	kami mau jalan sembuh danau ranau tidak naik bus	saya mau pergi ke danau ranau tidak naik bus
ani mulang anjak way kambas neen bingei	ani pulang dari way kambas nanti malam	ani pulang dari way kambas neen bingei	ani pulang dari way kambas nanti malam
mulei ino mejeng di pinggir lawet	gadis itu duduk di pinggir laut	gadis itu mejeng di pinggir laut	gadis itu duduk dipinggir laut
meghanai ino ngakuk jambeu di nuwo armin	bujang itu mengambil jambu di rumah armin	pemuda itu ngakuk jambu di rumah armin	itu bujang mengambil jambu armin di rumah
ani ino calak	ani itu cantik	ani itu tampan	ani itu cantik
sanak ino mengan juadahno	anak itu makan kuenya	anak itu makan juadahno	anak itu makan juadahno
tiyan balah perihal matematika	mereka bicara tentang matematika	mereka bicara perihal matematika	mereka berkata perihal matematika
ayah wat di nuwo	ayah ada di rumah	ayah ada di rumah	ayah ada di rumah
yo lapah jam sateu	dia pergi jam satu	dia jalan jam satue	dia pergi satu jam
ikam lagey mandei di way balak	saya sedang mandi di sungai	saya lagey mandi di way besar	saya sedang mandi di sungai
sanak ino littek jamo serundono	anak itu bertengkar dengan temannya	anak itu littek kawan serundono	anak itu bermusuhan sama serundono
murid-murid tunggo bu gureu di sekulah	murid-murid bertemu bu guru di sekolah	murid-murid ketemu bu guru di sekolah	murid-murid bertemu bu guru di sekolah
kendaraan liyew dakjo dinei di depan nuwono	kendaraan berlalu lalang di depan rumahnya	kendaraan liyew kesini diney di depan nuwono	kendaraan liyew disana di depan rumahnya ke sini
nuwo di anek ino wawai	rumah di kampung itu bagus	rumah di kampung itu bagus	rumah di kampung itu bagus
arif lagey belajar	arif sedang belajar	arif lagey belajar	arif sedang belajar
ibuno nyuwak tono	ibunya memanggil tono	ibuno nyuwok tono	ibunya memanggil tono
berubei komariah belei mangga	kemarin komariah membeli mangga	berubei komariah belei mangga	kemarin komariah membeli mangga
ikam mak maghing	saya tidak sakit	kami tidak sakit	saya tidak sakit
surat ijo pakai ikam	surat ini untuk saya	surat ini pakai kami	saya surat ini buat

ISSN: 2580-409X (Print) / 2549-6824 (Online)
DOI: https://doi.org/10.29407/intensif.v5i1.14670

IV. CONCLUSION

In this experiment, it is proven that the translation of the Lampung language Nyo dialect into Indonesian can be done using the DMT and SMT approaches. The accuracy results can be seen from the BLEU value obtained, namely, the DMT approach 39.32% and the SMT approach 59.85%. The DMT approach is useful in translating words that are already in the database but cannot capture aspects of the meaning of a given test sentence. The SMT approach can learn from the training data provided and is also able to accommodate the meaning of the test sentence so that it can be said that the SMT approach gives better results than the DMT approach.

ACKNOWLEDGMENTS

The author would like to thank the Indonesian Technocrat University, which has provided financial support for this research through an internal grant for Capacity Building Research (PPK).

REFERENCES

[1] F. Ariyani, "Distribusi Verba Berfrefiks (N-) Pada Bahasa Lampung dalam Kitab Kuntara Raja Niti dan Buku Ajar. Ranah: Jurnal Kajian Bahasa 3," *Ranah J. Kaji. Bhs.*, vol. 3, no. 2, pp. 124–134, 2014, doi: https://doi.org/10.26499/rnh.v3i2.43.

[2] P. Bhattacharyya, *Machine Translation*. Boca Raton: Taylor & Francis Group, 2015.

[3] Z. Abidin, "Penerapan Neural Machine Translation untuk Eksperimen Penerjemahan secara Otomatis pada Bahasa Lampung – Indonesia," *Pros. Semin. Nas. Metod. Kuantitatif 2017*, no. 978, pp. 53–68, 2017.

[4] Z. Abidin, A. Sucipto, and A. Budiman, "Penerjemahan Kalimat Bahasa Lampung-Indonesia Dengan Pendekatan Neural Machine Translation Berbasis Attention Translation of Sentence Lampung-Indonesian Languages With Neural Machine Translation Attention Based," *J. Kelitbangan*, vol. 06, no. 02, pp. 191–206, 2018.

[5] P. Permata and Z. Abidin, "Statistical Machine Translation Pada Bahasa Lampung Dialek Api Ke Bahasa Indonesia," *Media Inform. Budidarma*, vol. 4, no. 3, pp. 519–528, 2020, doi: 10.30865/mib.v4i3.2116.

[6] S. Ningsih and S. Saniati, "Eksperimen Pengenalan Ucapan Aksara Lampung Dengan CMU Sphinx 4," *J. Teknoinfo*, vol. 12, no. 1, p. 33, 2018, doi: 10.33365/jti.v12i1.40.

[7] D. V Sindhu and B. M. Sagar, "Dictionary Based Machine Translation from Kannada to Telugu," *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 225, p. 012182, 2017, doi: 10.1088/1757-899x/225/1/012182.

[8] R. M. M. Shalini and B. Hettige, "Dictionary Based Machine Translation System for Pali to Sinhala," *Sri Lanka*, no. October 2017, p. 6, 2017.

[9] K. T. C. Resmawan, I. K. R. Arthana, and I. M. G. Sunarya, "Pengembangan Aplikasi Kamus Dan Penerjemah Bahasa Indonesia–Bahasa Bali Menggunakan Metode Rule Based Berbasis Android," *KARMAPATI (Kumpulan Artik. Mhs. Pendidik. Tek. Inform.*, vol. 4, no. 2, pp. 70–81, 2015.

[10] N. Afifah, T. B. Santoso, and M. Yuliana, "Pembuatan Kamus Elektronik Kalimat Bahasa Indonesia dan Bahasa Jawa untuk Aplikasi Mobile Menggunakan Interpolation Search," *Semin. Proy. Akhir Jur. Tek. Telekomun. PENS-ITS*, pp. 1–7, 2010.

ISSN: 2580-409X (Print) / 2549-6824 (Online)
DOI: https://doi.org/10.29407/intensif.v5i1.14670

[11] R. Nugroho Aditya, T. Adji Bharata, and B. Hantono S, "Penerjemahan Bahasa Indonesia dan Bahasa Jawa Menggunakan Metode Statistik Berbasis Frasa," *Semin. Nas. Teknol. Inf. dan Komun.*, vol. 2015, no. Sentika, 2015.

[12] A. A. Suryani, D. H. Widyantoro, A. Purwarianti, and Y. Sudaryat, "Experiment on a phrase-based statistical machine translation using PoS Tag information for Sundanese into Indonesian," *2015 Int. Conf. Inf. Technol. Syst. Innov. ICITSI 2015 - Proc.*, 2016, doi: 10.1109/ICITSI.2015.7437678.

[13] Y. Jarob, H. Sujaini, and N. Safriadi, "Uji Akurasi Penerjemahan Bahasa Indonesia – Dayak Taman Dengan Penandaan Kata Dasar Dan Imbuhan," *J. Edukasi dan Penelit. Inform.*, vol. 2, no. 2, pp. 78–83, 2016, doi: 10.26418/jp.v2i2.16520.

[14] H. Sujaini, "Meningkatkan Peran Model Bahasa dalam Mesin Penerjemah Statistik (Studi Kasus Bahasa Indonesia-Dayak Kanayatn)," *Khazanah Inform. J. Ilmu Komput. dan Inform.*, vol. 3, no. 2, p. 51, 2017, doi: 10.23917/khif.v3i2.4398.

[15] T. Apriani, H. Sujaini, and N. Safriadi, "Pengaruh Kuantitas Korpus Terhadap Akurasi Mesin Penerjemah Statistik Bahasa Bugis Wajo Ke Bahasa Indonesia," *J. Sist. dan Teknol. Inf.*, vol. 1, no. 1, pp. 1–6, 2016.

[16] C. Adiputra, Krisna and Y. Arase, "Performance of Japanese-to-Indonesian Machine Translation on Different Models," *Proc. 23rd Annu. Meet. Linguist. Process. Soc.*, no. C, pp. 7–10, 2017.

[17] M. A. Sulaeman and A. Purwarianti, "Development of Indonesian-Japanese statistical machine translation using lemma translation and additional post-process," *Proc. - 5th Int. Conf. Electr. Eng. Informatics Bridg. Knowl. between Acad. Ind. Community, ICEEI 2015*, no. i, pp. 54–58, 2015, doi: 10.1109/ICEEI.2015.7352469.

[18] S. Mandira, H. Sujaini, and A. B. Putra, "Perbaikan Probabilitas Lexical Model Untuk Meningkatkan Akurasi Mesin Penerjemah Statistik," *J. Edukasi dan Penelit. Inform.*, vol. 2, no. 1, pp. 3–7, 2016, doi: 10.26418/jp.v2i1.13393.

[19] D. W. Ningtyas, H. Sujaini, and N. Safriadi, "Penggunaan Pivot Language pada Mesin Penerjemah Statistik Bahasa Inggris ke Bahasa Melayu Sambas," *J. Edukasi dan Penelit. Inform.*, vol. 4, no. 2, p. 173, 2018, doi: 10.26418/jp.v4i2.27414.

[20] M. D. Etsa, H. Sujaini, and N. Safriadi, "Pengaruh Metode Dictionary Lookup pada Cleaning Korpus Terhadap Akurasi Mesin Penerjemah Statistik Indonesia–Melayu Pontianak," *J. Edukasi dan Penelit. Inform.*, vol. 4, no. 1, p. 49, 2018, doi: 10.26418/jp.v4i1.24595.

[21] H. Priyatman, Hendro; Saleh, Muhammad ; Sujaini, "Analisis Akurasi Algoritma Extended Word Similarity Based Clustering (EWSB) pada Mesin," *J. Edukasi dan Penelit. Inform.*, vol. 6, no. 3, pp. 323–327, 2020.

[22] H. Sujaini and R. D. Nyoto, "Perbandingan Nilai Akurasi Algoritma Smoothing pada Mesin Penerjemah Statistik Bahasa Indonesia ke Bahasa Melayu Sambas dengan Language Model Toolkit IRSTLM," *J. Edukasi dan Penelit. Inform.*, vol. 6, no. 3, pp. 291–300, 2020.