Improving the role of language model in statistical machine translation (Indonesian-Javanese)

Herry Sujaini
Departement of Informatics, Tanjungpura University, Idonesia

Article Info

Article history:

Received May 4, 2019
Revised Oct 30, 2019
Accepted Nov 15, 2019

Keywords:

Indonesian-Javanese
language model
Statistical machine translation

ABSTRACT

The statistical machine translation (SMT) is widely used by researchers and practitioners in recent years. SMT works with quality that is determined by several important factors, two of which are language and translation model. Research on improving the translation model has been done quite a lot, but the problem of optimizing the language model for use on machine translators has not received much attention. On translator machines, language models usually use trigram models as standard. In this paper, we conducted experiments with four strategies to analyze the role of the language model used in the Indonesian-Javanese translation machine and show improvement compared to the baseline system with the standard language model. The results of this research indicate that the use of 3-gram language models is highly recommended in SMT.

Corresponding Author:

Herry Sujaini,
Department of Informatics,
Tanjungpura University,
Jl. Prof.Dr.H. Hadari Nawawi, Pontianak 78124, Indonesia.
Email: hs@unttan.ac.id

1. INTRODUCTION

Statistical machine translation (SMT) or known as statistical-based machine translation, is a paradigm of machine translation where the interpretation is created dependent on a statistical model which parameters come from bilingual corpus (parallel corpus) analysis. Corpus is a collection or sample of written or oral text in the form of data that can be read by using a set of machines and can be noted in the form of various linguistic information forms [1]. A quality corpus greatly influences the outcome of a statistical or neural-based translator machine. Many previous researchers have experimented with improving the quality of the corpus [2-6].

The best hypothesis for each input of sentence f is the goal of bilingual corpus analysis by:

$$\bar{e}^* = argmax_{\bar{e}} P(\bar{e}|\bar{f}) = argmax_{\bar{e}} P(\bar{f}|\bar{e})P(\bar{e}) \tag{1}$$

$P(\bar{e}|\bar{f})$ is a translation model that expresses the probability of the relationship between the source language and the target language. Language models that determine the probability of strings in the target language are denoted by $P(\bar{e})$, normally uses the standard word of trigram model from:

$$P(e_1, \dots, e_l) \approx \prod_{i=3}^{l} P(e_i|e_{i-1}, e_{i-2}) \tag{2}$$

which $\bar{e}$ = e1 ,..., el. In the trigram model form, each word is predicted based on the previous two-word history.

Although machine translator models have continued to develop in recent years, statistical machine translation (SMT) continues to grow rapidly, with more and more proposed new translation models being practiced in various languages [7-9]. Most of the work in SMT concentrates on developing better translation models. Little exertion has been made to maximize the role of language modeling for machine translation. The purpose of this research was to improve the role of language modeling, which in turn will improve the accuracy of the translation results of an SMT.

Figure 1 shows the general statistical language machine architecture. The decoder functions as a translator machine whose job it is to translate sentences from one language to another. The results of the work of the decoder can differ from one another. These results are influenced by the models used, namely the translation model (TM) and the language model (LM) as the main model, and the feature model (FM) besides. TM is generated through the training process of a parallel corpus, while the training process of a monolingual corpus from the target language generated LM. FM is usually used as an effort to improve the accuracy of machine translators by adding linguistic features such as Part of Speech (PoS) [10-12]. The generated PoS can be done with a supervised or unsupervised approach [13]. The main system functions as a translator machine to produce the target language from sentence input in the source language called the decoder. As shown in Figure 1, the parallel corpus is the primary source for building an SMT, while the monolingual corpus can use sentences that are on the target side of the parallel corpus. The size of the monolingual corpus can be enlarged by adding other sentences in the same language, even though it does not have a pair in the parallel corpus.

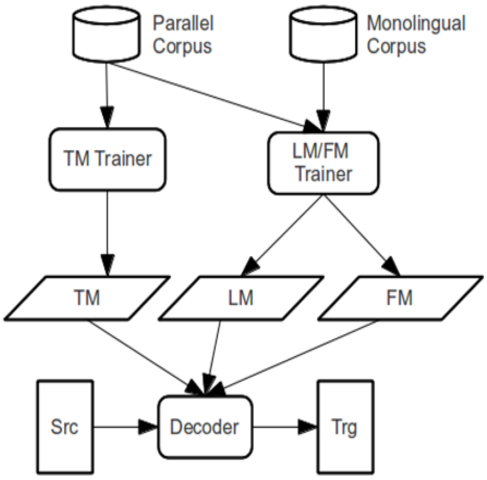


Figure 1. SMT architecture

Several studies have been conducted to improve the role of linguistic models in various languages and methods. Yu *et al.* [14] explained language models that triggered new topics by calculating contexts and topics and estimate n-gram probabilities under a given topic and adjust language model scores based on the distribution of different topics online. The resulting translation proved to improve the hypothesis considered best by the first stage of the system. Zhang *et al.* [15] have researched by improving the coding of automated veterinary diagnoses through large-scale language modeling. The algorithm proposed by them addresses important challenges in veterinary medicine and training in unsupervised learning for clinical language development. Mohaghegh [16] reported improved accuracy by enhancing the role of the language model in the English-Persian translator machine. Monz [17] reported improved accuracy by enhancing the role of the language model in Arabic- and Chinese-to-English translator machines. Banerjee *et al.* [18] reported their research to improve the language model by learning from speech recognition mistakes in a listening reading tutor. Sujaini *et al.* [19] reported the results of their research to improve the accuracy of machine translators by using the part of speech features. The results of this study can increase accuracy by 6.45% when compared to machine translators without using part of speech. Jaya and Gupta [20] proposed a better quality SMT that was improved by 2 points in the English to Hindi system and 2.93 points in the Hin-Eng system. These results were obtained as they explored the corpus augmentation approach for the English and Hindi Two-Way SMT.

The language model is designed to obtain the occurrence probability of words (or token). If W1 = (w1,..., wL) shows the string token of L on fixed vocabulary. The n-gram language model provides probability to w_1^L according to:

$$P(w_1^L) = \prod_{i=1}^{L} P(w_i | w_1^{i-1}) \approx \prod_{i=1}^{L} \hat{P}(w_i | w_{i-n+1}^{i-1}) \tag{3}$$

where the approximation reflects Markov's assumption that only n-1 token that the newest relevant in predicting the next word.

For each w substring, for example, f(w) shows the frequency of substring occurrence in the specified target language sequence, usually very long, which called training data. The maximum-likelihood probability for n-gram is given by its relative frequency.

$$r(w_i | w_{i-n+1}^{i-1}) = \frac{f(w_{i-n+1}^{i})}{f(w_{i-n+1}^{i-1})} \tag{4}$$

In principle, the predictive accuracy of the language model can be enhanced by increasing the order of n-grams [21]. However, under certain conditions, this can reduce the accuracy of translation when using complex data, especially if there are many errors in corpus data. This study discusses the best accuracy that can be produced by machine translators by conducting experiments on the use of the order of n-grams on LM.

2. RESEARCH METHOD

2.1. Experimental stages

The training data is a parallel corpus of Indonesia – Kromo Javanese language taken from folklore manually translated as many as 5108 sentences. In the process of training, 4500 pairs of sentences and 608 pairs of sentences were used for the testing process. The experimental stages conducted can be seen in Figure 2. Corpus preparation (preprocessing) was conducted by performing the process of cleaning, tokenizing, and lowercasing to the parallel corpus that has been prepared. The language model used in the baseline system was the trigram model of Javanese language trained by using toolkit SRILM [22], while parallel corpus that was ready to use was then trained to obtain word alignments, phrase table, language model, and model combination weights. The baseline used in this research was trained by using standard tools, namely GIZA++ [23], to train the word alignment and Moses for phrase-based coding. Moses is a tool that is an implementation of Statistic Machine Translation. Moses is used to training a statistical model of translated text from the source language to the target language. In translating the language, Moses requires a corpus in two languages, source language and target language. Moses is released under the license pf LGPL (Lesser General Public License) and is available as source code and binary for Windows and Linux. Its development is supported by the EuroMatrix project, with the funding by European Commission [24]. The decoder, as a translator machine, was set following the experimental strategy conducted, which was by changing the language model variables used. For each setting, testing was performed by inputting 608 sentences that had been prepared previously. The tests were performed using the BLEU automatic evaluation method [25].

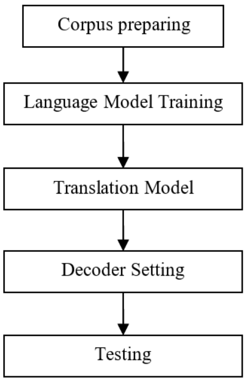


Figure 2. Experimental stages

2.2. Experiment strategy

To see the role of the language model used in Indonesian-Java SMT, in this experiment, four strategies were performed.

a. The language model was trained from 4500 sentences of the parallel corpus target; in other words, the test reference sentence was not included in the training. Furthermore, this first strategy was tested for 3 to 7-gram models.
b. The language model was trained from 4500 sentences of parallel corpus target plus 608 reference sentences; in other words, the test reference sentence was included in the training. Furthermore, this first strategy was tested for 3 to 7-gram models.
c. The language model was trained from 608 reference sentences and 3892 sentences of parallel corpus target, then added with 100 sentences of the remaining parallel corpus targets for each experiment. Six experiments were added; therefore, the corpus used in each experiment was 4600, 4700, 4800, 4900, 5000, and 5100 sentences.
d. The language model was trained from 4500 target sentences, then added 100 reference sentences for each experiment. Six experiments were added; therefore, the corpus used in each experiment was 4600, 4700, 4800, 4900, 5000, and 5100 sentences.

2.3. Result and discussion

The training data is in the form of the Indonesian-Java parallel corpus as shown in Figure 3. The left column is a collection of sentences in the Indonesian language, while the right column is a sentence cluster in Javanese, where each line is a translation of the corresponding sentence.

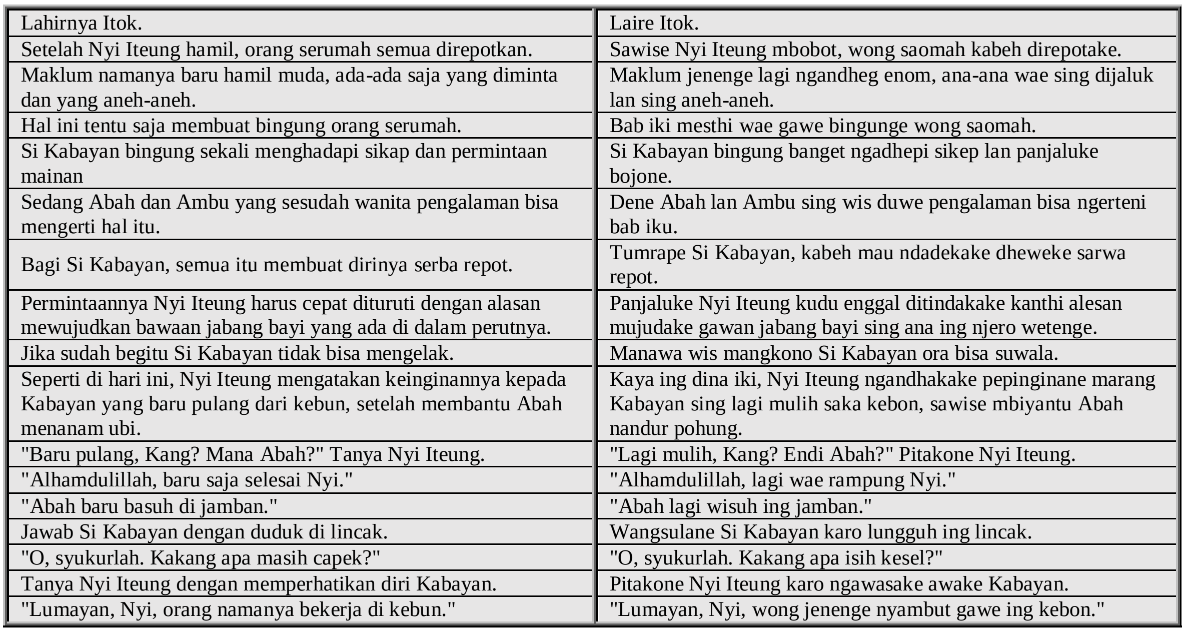
Examples of sentences that have been passed the process of cleaning, tokenizing, and lowercasing are:

"Baru pulang, Kang? Mana Abah?" Tanya Nyi Iteung

Result :

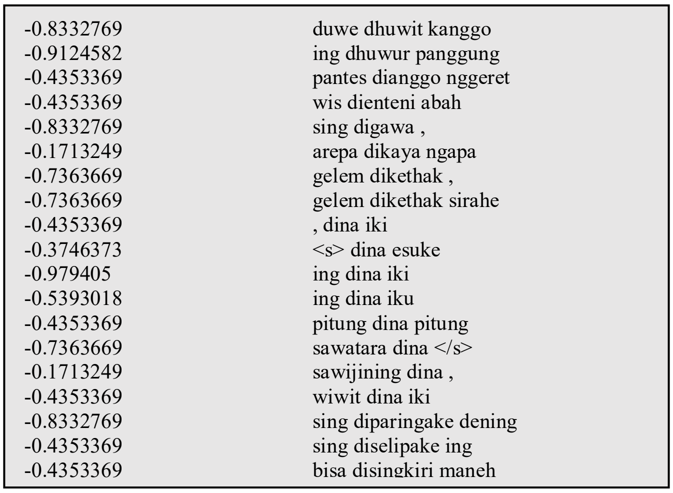
" lagi mulih , kang? endi abah? " pitakone nyi iteung

The language model was generated from the training process conducted on the target language of the parallel corpus, i.e., Java language. As a baseline, a trigram (3-gram) training was conducted, then training was also conducted to produce a comparison machine with 4-gram, 5-gram, 6-gram, and 7-gram models. The example of the 3-gram language model can be found in Figure 4, while the example of the 7-gram language model can be found in Figure 5. For instance, Figure 4. says that the probability of the first word in a sentence being "dina esuke" is 10-0.3746373 = 0.422, the probability gave the pair word "sawatara dina" that the next thing that happens is that the sentence ends 10-0.7363669 = 0.183, and so forth. From the training results, the number of token pairs with their probabilities for each n-gram are: 1-gram=5598, 2-gram=26350, 3-gram=3924, 4-gram=1768, 5-gram=646, 6-gram=181 and 7-gram=53.



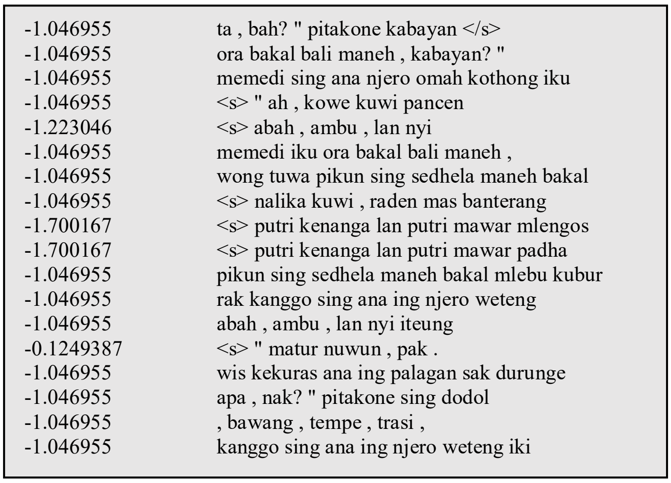
Lahirnya Itok.	Laire Itok.
Setelah Nyi Iteung hamil, orang serumah semua direpotkan.	Sawise Nyi Iteung mbobot, wong saomah kabeh direpotake.
Maklum namanya baru hamil muda, ada-ada saja yang diminta dan yang aneh-aneh.	Maklum jenenge lagi ngandheg enom, ana-ana wae sing dijaluk lan sing aneh-aneh.
Hal ini tentu saja membuat bingung orang serumah.	Bab iki mesthi wae gawe bingunge wong saomah.
Si Kabayan bingung sekali menghadapi sikap dan permintaan mainan	Si Kabayan bingung banget ngadhepi sikep lan panjaluke bojone.
Sedang Abah dan Ambu yang sesudah wanita pengalaman bisa mengerti hal itu.	Dene Abah lan Ambu sing wis duwe pengalaman bisa ngerteni bab iku.
Bagi Si Kabayan, semua itu membuat dirinya serba repot.	Tumrape Si Kabayan, kabeh mau ndadekake dheweke sarwa repot.
Permintaannya Nyi Iteung harus cepat dituruti dengan alasan mewujudkan bawaan jabang bayi yang ada di dalam perutnya.	Panjaluke Nyi Iteung kudu enggal ditindakake kanthi alesan mujudake gawan jabang bayi sing ana ing njero wetenge.
Jika sudah begitu Si Kabayan tidak bisa mengelak.	Manawa wis mangkono Si Kabayan ora bisa suwala.
Seperti di hari ini, Nyi Iteung mengatakan keinginannya kepada Kabayan yang baru pulang dari kebun, setelah membantu Abah menanam ubi.	Kaya ing dina iki, Nyi Iteung ngandhakake pepinginane marang Kabayan sing lagi mulih saka kebon, sawise mbiyantu Abah nandur pohung.
"Baru pulang, Kang? Mana Abah?" Tanya Nyi Iteung.	"Lagi mulih, Kang? Endi Abah?" Pitakone Nyi Iteung.
"Alhamdulillah, baru saja selesai Nyi."	"Alhamdulillah, lagi wae rampung Nyi."
"Abah baru basuh di jamban."	"Abah lagi wisuh ing jamban."
Jawab Si Kabayan dengan duduk di lincak.	Wangsulane Si Kabayan karo lungguh ing lincak.
"O, syukurlah. Kakang apa masih capek?"	"O, syukurlah. Kakang apa isih kesel?"
Tanya Nyi Iteung dengan memperhatikan diri Kabayan.	Pitakone Nyi Iteung karo ngawasake awake Kabayan.
"Lumayan, Nyi, orang namanya bekerja di kebun."	"Lumayan, Nyi, wong jenenge nyambut gawe ing kebon."

Figure 3. Indonesian-Java parallel corpus



```
-0.8332769          duwe dhuwit kanggo
-0.9124582          ing dhuwur panggung
-0.4353369          pantes dianggo nggeret
-0.4353369          wis dienteni abah
-0.8332769          sing digawa ,
-0.1713249          arepa dikaya ngapa
-0.7363669          gelem dikethak ,
-0.7363669          gelem dikethak sirahe
-0.4353369          , dina iki
-0.3746373          <s> dina esuke
-0.979405           ing dina iki
-0.5393018          ing dina iku
-0.4353369          pitung dina pitung
-0.7363669          sawatara dina </s>
-0.1713249          sawijining dina ,
-0.4353369          wiwit dina iki
-0.8332769          sing diparingake dening
-0.4353369          sing diselipake ing
-0.4353369          bisa disingkiri maneh
```

Figure 4. 3-gram language model



```
-1.046955       ta , bah? " pitakone kabayan </s>
-1.046955       ora bakal bali maneh , kabayan? "
-1.046955       memedi sing ana njero omah kothong iku
-1.046955       <s> " ah , kowe kuwi pancen
-1.223046       <s> abah , ambu , lan nyi
-1.046955       memedi iku ora bakal bali maneh ,
-1.046955       wong tuwa pikun sing sedhela maneh bakal
-1.046955       <s> nalika kuwi , raden mas banterang
-1.700167       <s> putri kenanga lan putri mawar mlengos
-1.700167       <s> putri kenanga lan putri mawar padha
-1.046955       pikun sing sedhela maneh bakal mlebu kubur
-1.046955       rak kanggo sing ana ing njero weteng
-1.046955       abah , ambu , lan nyi iteung
-0.1249387      <s> " matur nuwun , pak .
-1.046955       wis kekuras ana ing palagan sak durunge
-1.046955       apa , nak? " pitakone sing dodol
-1.046955       , bawang , tempe , trasi ,
-1.046955       kanggo sing ana ing njero weteng iki
```

Figure 5. 7-gram language model

The first experimental strategy was using 4500 sentences parallel corpus, and 4500 sentences monolingual corpus of, the results of the experiments produced can be found in Table 1. Machine 1.3 means using strategy 1 with 3-gram; machine 1.4 means using strategy 1 with 4-gram, and so on. The experiment's results show that with the addition of n-grams in the monolingual corpus taken from the parallel corpus, it does not show a significant increase in accuracy (represented by BLEU value). The highest value on the 4-gram to 7-gram model can only increase the accuracy of ((32.46-32.42) /32.42) * 100% = 0.12%.

The second experimental strategy was using 4500 sentences parallel corpus, and 5108 sentences monolingual corpus, the results of the experiments produced can be found in Table 2. Machine 2.3 means using strategy 2 with 3-gram; machine 2.4 means using strategy 2 with 4-gram, and so on. The experiment's results show that with the addition of n-grams in the monolingual corpus taken from the parallel corpus, it does not show an increase in accuracy (represented by BLEU value), even lower than the 3-gram baseline. The highest score remains on the 3-gram model, which is 40.79.

Table 1. Result from strategy 1

Machine	n-gram	BLEU Score (%)
1.3	3-gram	32.42
1.4	4-gram	32.46
1.5	5-gram	32.46
1.6	6-gram	32.46
1.7	7-gram	32.46

Table 2. Result from strategy 2

Machine	n-gram	BLEU Score (%)
2.3	3-gram	40.79
2.4	4-gram	40.69
2.5	5-gram	40.69
2.6	6-gram	40.67
2.7	7-gram	40.71

The third strategy experiment was using a parallel corpus of 4500 sentences and a monolingual corpus of 4500 to 5100 sentences, 4500 baseline sentences consist of 608 reference sentences and 3892 sentences of parallel corpus targets. The experiment's result produced can be found in Table 3 indicates that the addition of monolingual corpus quantities taken from the parallel corpus does not show significant increases at accuracy. Machine 3.45 means using strategy 3 with 4500 sentences; machine 3.46 means using strategy 3 with 4600 sentences, and so on. The highest value on a 3.51 can only increase the accuracy of ((40.81-40.59) / 40.59) * 100% = 0.54%.

The fourth strategy experiment was using a parallel corpus of 4500 sentences and a monolingual corpus of 4500 to 5100 sentences, and the whole 4500 baseline sentences were taken from the parallel corpus target sentence. The results of the experiments produced can be found in Table 4 indicates that the addition of monolingual corpus quantities taken from the reference sentence shows a significant increase at accuracy. Machine 4.45 means using strategy 4 with 4500 sentences; machine 4.46 means using strategy 4 with 4600 sentences, and so on. The highest value on a 4.51 machine with 5100 sentences in monolingual corpus can increase accuracy by ((40.63-32.42) / 32.42) * 100% = 25.32%.

Table 3. Result from strategy 3

Machine	Monolingual corpus	BLEU Score (%)
3.45	4500	40.59
3.46	4600	40.46
3.47	4700	40.20
3.48	4800	40.46
3.49	4900	40.64
3.50	5000	40.63
3.51	5100	40.81

Table 4. Result from strategy 4

Machine	Monolingual corpus	BLEU Score (%)
4.45	4500	32.42
4.46	4600	35.48
4.47	4700	37.19
4.48	4800	37.82
4.49	4900	38.55
4.50	5000	39.52
4.51	5100	40.63

Experiments conducted on strategies 1 and 2 show that the use of n-gram model from 3-gram to 7-gram does not affect the accuracy of the Indonesian-Java translator machine with a parallel corpus of 4500 sentences. This is due to the small number of sentences used in the corpus. The small quantity of corpus sentences results in no variation in the probability of each pair of tokens, as seen in the 7-gram language model; thus, for SMT using the small corpus, it is best to keep using the 3-gram language model.

The monolingual corpus quantity addition experiments used for gradual language model training, the results are demonstrated by strategies 3 and 4. From the experimental results, it is found that the best results are obtained by increasing the quantity of the monolingual corpus outside the parallel corpus in strategy 4, in other words, the monolingual corpus taken from the parallel corpus target language, then added with another sentence beyond the existing sentence in the parallel corpus. The increase of the BLEU score of each machine to the baseline can be seen in Figure. 6. The experiment's results on strategy 4 show a significant increase for each addition of 100 sentences to the monolingual corpus, as seen in Table 5. From the results of this study, it can be concluded that the role of the language model is quite important in anticipating the sentences to be translated on the SMT, especially when the phrase in the sentence is not contained in the translation model. This will certainly be more apparent on SMT with small resources because the possibility of a sentence to be translated does not exist in the translation model is certainly very large compared to SMT with large resources.

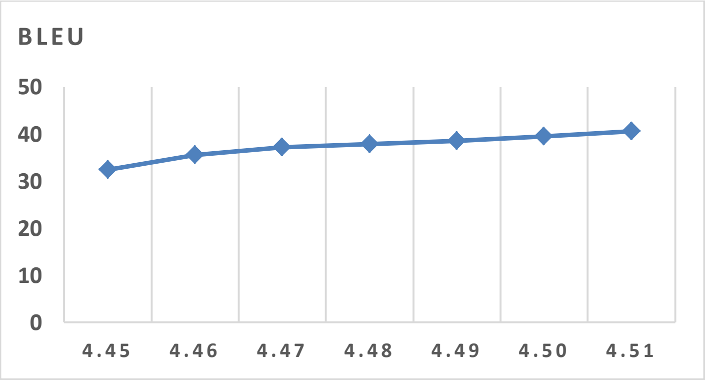


Figure 6. Increasing of BLEU scores

Table 5. Increasing accuracy on strategy 4

Machine	Monolingual corpus	Increasing (%)
4.45	32.42	
4.46	35.48	9.44
4.47	37.19	4.82
4.48	37.82	1.69
4.49	38.55	1.93
4.50	39.52	2.52
4.51	40.63	2.81

3. CONCLUSION

The utilization of the n-gram model from 3-gram to 7-gram does not affect the accuracy of the Indonesian-Java translator machine. It is recommended that SMT using a small corpus should keep using a 3-gram language model. The best result for improving the language model role is to use the the parallel corpus target language as the monolingual corpus, then added as much as possible with other sentences beyond the existing sentence in the parallel corpus.

REFERENCES

[1] M. Volk, “Parallel Corpora, Terminology Extraction und Machine Translation,” *In: 16. DTTSymposion. Terminologie und Text(e), Mannheim*, 22 - 24 March 2018, 3-14. 2018.

[2] E. Yıldız, A.C. Tantu˘g, and B. Diri. “The effect of parallel corpus quality vs size in English-to-Turkish SMT,” In *Proceedings of the Sixth International Conference on Web services and Semantic Technology (WeST 2014)*, 2014.

[3] A. Imankulova, T. Sato, M. Komachi, “Improving Low-Resource Neural Machine Translation with Filtered Pseudo-parallel Corpus,” In *Proceedings of the 4th Workshop on Asian Translation (WAT2017)*, Taipei, 2017

[4] K.K. Arora and S.S. Agrawa, “Pre-Processing of English-Hindi Corpus for Statistical Machine Translation,” *Computación y Sistemas*, Vol. 21, No. 4, 2017.

[5] H. Tran, Y. Guo, P. Jian, S. Shi, and H. Huang, “Improving Parallel Corpus Quality for Chinese-Vietnamese Statistical Machine Translation,” *Journal of Beijing Institute of Technology*, Vol. 27, No. 1, 2018.

[6] M.G. Asparilla, H. Sujaini, and R.D. Nyoto, “Corpus Quality Improvement to Improve the Quality of Statistical Translator Machines (Case Study of Indonesian Language to Java Krama),” *Jurnal Linguistik Komputasional,* Vol. 1, No. 2, 2018.

[7] J. Su, H. Wu, H. Wang, Y. Chen, X. Shi, H. Dong, and Q. Liu, “Translation Model Adaptation for Statistical Machine Translation with Monolingual Topic Information,” in *Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers),* Jeju Island, 2012.

[8] G. Neubig, T. Watanabe, "Optimization for Statistical Machine Translation: A Survey," *Computational Linguistics*, Vol. 42, No. 1, 2016.

[9] K.N. Dew, A.M.Turner, Y.K. Choi, A. Bosold, and K. Kirchhoffe, "Development of Machine Translation Technology for Assisting Health Communication: A Systematic Review," *Journal of Biomedical Informatics*, Vol. 85, 2018

[10] P.J. Antony and K.P. Soman, “Kernel Based Part of Speech Tagger for Kannda,” in *International Conference on Machine Learning and Cybernetics*, ICMLC 2010, Qingdao, Shandong, 2010.

[11] M. Mohaghegh, A. Sarrafxadeh, and T. Moir, "Improved Language Modeling for English-Persian Statistical Machine Translation," in *Proceedings of SSST-4, Fourth Workshop on Syntax and Structure in Statistical Translation, COLING 2010*, Beijing, 2010.

[12] J. Sangeetha, S. Jothilakshmi, and R.N.D. Kumar, "An Efficient Machine Translation System for English to Indian Languages Using Hybrid Mechanism," *International Journal of Engineering and Technology (IJET)*, Vol. 6, No. 4, 2014.

[13] H. Sujaini, Kuspriyanto, A.A. Arman, and A. Purwarianti, “Extended Word Similarity Based Clustering on Unsupervised PoS Induction to Improve English-Indonesian Statistical Machine Translation,” in *16th ORIENTAL COCOSDA/CASLRE-2013*, Gurgaon, India, 2013.

[14] H. Yu, J. Su, Y. Lv, and Q. Liu, “A Topic-Triggered Language Model for Statistical Machine Translation,” in *Proceedings of the Sixth International Joint Conference on Natural Language Processing,* Nagoya, 2013.

[15] Y. Zhang, A. Nie, A. Zehnder, L. Rodney, and J. Zou, “*VetTag: improving automated veterinary diagnosis coding via large-scale language modeling,*” Digital Medicine, 2019.

[16] M. Mohaghegh, A. Sarrafzadeh, and T. Moir, “Improved Language Modeling for English-Persian Statistical Machine,” in *SSST-4, Fourth Workshop on Syntax and Structure in Statistical Translation*, Beijing, 2010.

[17] C. Monz, “Statistical Machine Translation with Local Language Models,” in *Conference on Empirical Methods in Natural Language Processing*, Edinburgh, 2011.

[18] S. Banerjee, J. Mostow, J. Beck, and W. Tam, "Improving Language Models by Learning from Speech Recognition Errors in a Reading Tutor that Listens," in *Second International Conference on Applied Artificial Intelligence*, 2003.

[19] H. Sujaini, Kuspriyanto, A.A. Arman, and A. Purwarianti, “A Novel Part-of-Speech Set Developing Method for Statistical Machine Translation,” *TELKOMNIKA (Telecommunication Comput. Electron. Control.)*, vol. 12, no. 3, 2014.

[20] K. Jaya and D. Gupta, “Exploration of Corpus Augmentation Approach for English-Hindi Bidirectional Statistical Machine Translation System,” *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 6, no. 3, 2016.

[21] [z] C. Shaoul, C.F. Westbury, and R.H. Baayen,"*The Subjective Frequency of Word n-grams*," PSIHOLOGIJA, Vol. 46, No. 4, 2013.

[22] A. Stolcke, J. Zheng, W. Wang, and V. Abrash, “SRILM at sixteen: Update and outlook,” in *Automatic Speech Recognition and Understanding (ASRU)*, 2011 IEEE Workshop, Waikoloa, 2011.

[23] F. J. Och and H. Ney, “A Systematic Comparison of Various Statistical Alignment Models,” *Computational Linguistics*, vol. 1, no. 29, pp. 19-51, 2003.

[24] W. Xu and P. Koehn, “Extending Hiero Decoding in Moses with Cube Growing,” The Prague Bulletin of Mathematical Linguistics, 8(1), 2012

[25] K. Papineni, S. Roukos, T. Ward and W.-J. Zhu, “BLEU: A Method For Automatic Evaluation of Machine Translation,” in *Proceedings of the 40th Annual Meeting of the Association of Computational Linguistics (ACL), Pennsylvania*, 2002.

BIOGRAPHY OF AUTHOR

Herry Sujaini graduated from a bachelor's degree in the Electrical Engineering Department, University of Tanjungpura. He got his master and a doctoral degree from STEI, Bandung Institute of Technology. Since 1997, he has become a lecturer at Informatics Department, Engineering Faculty, University of Tanjungpura. Her research interest is on computational linguistics, mainly on machine translation and machine learning