Building of Informatics, Technology and Science (BITS)
Volume 4, No 2, September 2022 Page: 588−594
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v4i2.1983

POS Tagger Improvisation using HMM with the Addition of Foreign Word Labels on Telkom University News

Winkie Setyono*, Donni Richasdy, Mahendra Dwifebri Purbolaksono

School of Computing, Informatics Engineering, Telkom University, Bandung, Indonesia
Email: 1,*winkieset@student.telkomuniversity.ac.id, 2donnir@telkomuniversity.ac.id, 3mahendradp@telkomuniversity.ac.id
Email Penulis Korespondensi: winkieset@telkomuniversity.ac.id
Submitted:**26/07/2022**; Accepted:**23/08/2022**; Published**: 30/09/2022**

Abstract−News is a medium of daily information usually obtained by the public. The news consists of a lot of information in it and is composed of sentence structures. Each language is unique with its own sentence structure, like Indonesian and other foreign languages. But nowadays, many media mix Indonesian with foreign languages, making the sentence structure different from Bahasa Indonesia. To classify these words, Part Of Speech Tagging needed to determine the class of words composed of sentences by learning from the Corpus of each language. The language structure can determine the results of tagging from the POS Tagger. If there are words that are not in the Corpus, it can reduce the accuracy of the POS Tagger. With the new sentence structure, POS Tagger requires a larger Corpus to learn, but the current corpus doesn't cover it yet. We conducted to enhance the research results by adding data to Corpus with a different sentence structure from the Indonesian Language Corpus using sentences from online media. Added about 242 sentences with 7,043 tokens on Corpus focused on Foreign Word tags, which total 3819 tags. After some testing and scenarios, the results of the accuracy of POS Tagger show an accuracy of 94.7% using the Hidden Markov Model method with the F1-Score tag FW 78%.

Keywords: POS Tagger; Sentences; Foreign Word; Corpus

1. INTRODUCTION

News is a medium often found every day in getting information in the community about something happening. That makes the general public interested in knowing information consisting of these facts or opinions [1]. News articles contain many crucial details such as titles, names, places, objects, and subjects through sentence arrangement. The sentence structure made consists of many word classes that are used to compose the whole news article. But now, many online media and everyday language contain many foreign words. This sometimes makes language acculturation so that the use of words in Indonesian has a different set of words than usual.

In Indonesian news, it is not uncommon to find words not included in the Indonesian language, such as using foreign words to compose sentences in the news. The function of the language in the broad scope of communication is carried out by Indonesian and several foreign languages ranging from intercultural and regional exchanges [2].

Part of Speech Tagging or POS Tagging, which can also be labelled according to word tag, in Natural Language Processing (NLP) is a process where words will be labelled as word tag in a given sentence according to the tagset [3]. From the tagset, POS Tagging learns from each set of words from each given language. The implementation of POS Tagging has many methods that can be used, an example is the Hidden Markov Model method. This method is a statistical model of a system imagined by a Markov process using unknown parameters, selecting hidden state parameters from parameters that can be seen and observed state. As in POS Tagging, the word class tag is invisible, but the word can be seen [4].

Each region or region has its own language structure and is unique to each region. For example, Indonesian's sentence structure is different from English's. Research on POS Tagging in Indonesian has been carried out as in the research: POS Tagging Indonesian with the Viterbi Algorithm method [5], where the algorithm looks at and determines the best path that can be called the Viterbi path. Indonesian POS Tagging Using the Viterbi Algorithm conducted by N. Sabloak et al. showed an average accuracy of 93.23%, measured using the 10-fold cross-validation measurement method with data from the dictionary for the dataset is 16,290 words with manual word input testing data [5].

The research was conducted using another method, namely comparing the POS Tagging method for Indonesian by Ahmad Z et al. in 2017. In this study, Unigram showed an accuracy of 88.37% [3]. POS Tagging is also applied to the Malayalam Language by Sindhya K Nambiar, et al. using the Hidden Markov Model [4], which can show improvisation results by being 85% of the 13 sentence tags for performance.

In addition, it is also carried out by Ryan A, et al. on regional languages in Indonesia, such as POS Tagging, which is applied to Javanese [6] and Balinese languages[7] by using the same method, namely the Hidden Markov Model method. The journal POS Tagging for the Javanese language used the Hidden Markov Model method. From the journal, it shows an accuracy of 92.6% by using the corpus tagset based on the "Javanese Vocabulary Book" dictionary. The dataset used comes from Javanese news on solopos.com/jagad-jawa which contains news in Javanese. Balinese Languages shows an average accuracy of model is 68,56% using 10 K-Fold Cross Validation [7].

The lack of foreign tag data and the size of the training data in the Corpus used in this study provide recommendations by adding sentences containing Indonesian and foreign languages to the Corpus, which is focused on increasing the size and the number of FW tags. The improvisation is expected to affect better results on POS Tagger by using data from Indonesian news portals using the Hidden Markov Model method. POS Tagger Model will be

Building of Informatics, Technology and Science (BITS)
Volume 4, No 2, September 2022 Page: 588−594
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v4i2.1983

implemented to Dataset Telkom University News. Furthermore, the results of the Telkom University news data tagging from the Original Corpus will be compared with the Modified Corpus.

2. METHODOLOGICAL RESEARCH

2.1 Stages Of Research

In this study, design and test the system built for POS Tagger using the Hidden Markov Model method. Overall, this has a system workflow, as shown in Figure 1.

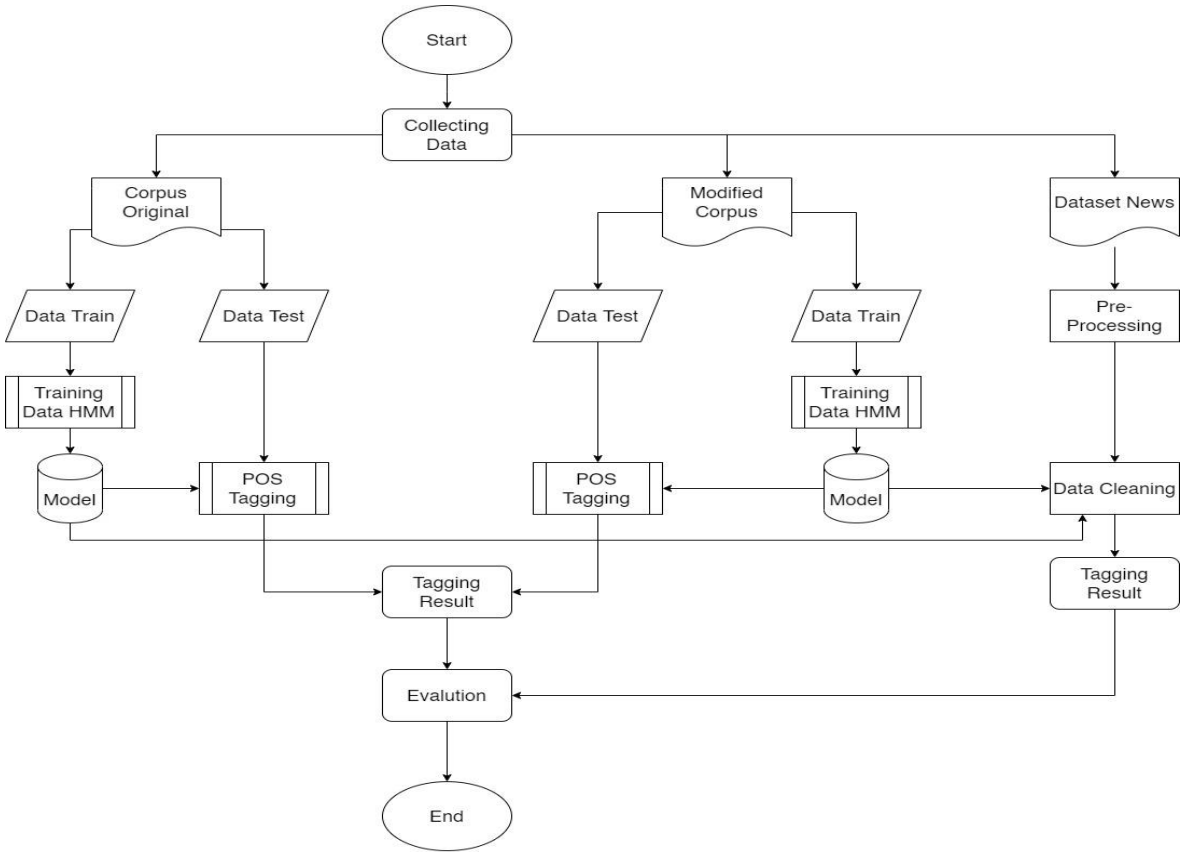

Figure 1. System Design

2.2 Dataset Corpus

Dataset The corpus dataset is the dictionary for the training data used in this study. The Corpus contains Indonesian sentences consisting of 10,000 sentences built with 256,683 tokens [8] with TSV file format.

Table 1. Tagset Corpus

Tag	Deskripsi
CC	Coordinating conjunction
CD	Cardinal number
OD	Ordinal Number
DT	Determiner
FW	Foreign Word
IN	Preposition
JJ	Adjective
MD	Modal
NEG	Negation
NN	Noun
NNP	Proper Noun
NND	Classifier
PR	Demonstrative pronoun
PRP	Personal pronoun
RB	Adverb
RP	Particle

Building of Informatics, Technology and Science (BITS)
Volume 4, No 2, September 2022 Page: 588−594
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v4i2.1983

SC	Subordinating conjunction
SYM	Symbol
UH	Interjection
VB	Verb
WH	Question
X	Unknown
Z	Punction

2.2 Dataset Corpus Modified

Dataset Modified Corpus is the dictionary for the training data used in this study. The Corpus contains Indonesian sentences adding 242 sentences, and over 7000 tokens were tagged manually. Size more than the Original Corpus, focusing on adding the FW label. Data is obtained from some articles on the website hypebeast.com/id with a mix of the words Bahasa and English. Tagset still has the same like on Table 1.

2.3 Dataset News

The dataset used in this study was obtained from scraping on the Detik.com news portal. News dataset related to Telkom University with json file format. The dataset includes 6 columns such as:

Table 2. Dataset News

Title	Category	Author	Date	Article	Scrape Time
Tentang Telkom University, Universitas Swasta ...	detikEdu	Novia Aisyah	Selasa, 29 Jun 2021 16:46 WIB	Tentang Telkom University, Universitas Swasta ...	2021-07-14 18:57:08
Universitas Terbaik di Jawa Timur Versi Webome...	detikEdu	Novia Aisyah	Selasa, 29 Jun 2021 15:44 WIB	Universitas Terbaik di Jawa Timur Versi Webome...	2021-07-14 18:57:37

2.4 Hidden Markov Model

Hidden Markov Model can help to find the probability of certain words and to predict the probability of remaining words in the sequence [9]. There are three processes Initialization which means getting the number of word labels, Transition of label search after the label is checked, and Emission of the number of words from the label from the training data. The formula (1) of the Hidden Markov Model is [10]:

$$P\ (word|tag) * P(tag|tag\ before\ \boldsymbol{n}) \quad (1)$$

2.5 Viterbi Algorithm

The algorithm Viterbi determines the possible Viterbi path sequences generated in the HMM circuit [11]. Formula Viterbi can be seen on formula number (2)

$$v_t(j) = max_{i=1}^{n}\ v_{t-1}(i)a_{ij}b_j(o_t) \quad (2)$$

Description:
$v_{t(j)}$ = probability of HMM state
$v_{t-1}(i)$ = probability of Viterbi path
a_{ij} = probability transition q_i to state q_j
$b_j(o_t)$ = probability emission

2.6 Unigram

Unigram tagger is a word tagging process by looking at one word at a time. It considers one word at a time and assigns each word to the most common tag [12]. The formula can be seen in formula number (3)

$$P\ (w_i|t_i) = freq(w_i|t_i)/freq(w_i) \quad (3)$$

Description:
w_i = word in index
v_i = tag in index

2.7 Pre-Processing Data

The process of preparing data that is used to become data that can be retrieved for subsequent processing. In this study, the preprocessing stage used for the Telkom University news dataset is tokenization. Tokenization is separating the words in the data. [7]

2.8 Evaluation

Building of Informatics, Technology and Science (BITS)
Volume 4, No 2, September 2022 Page: 588−594
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v4i2.1983

Evaluation in POS Tagger this time is measured in percent units, starting from the accuracy of the tagger to compare the results of the Modified Corpus with the Original Corpus. TP is True Positive,FP is False Positive,TN is True Negative,FN is False Negative

2.8.1 Accuracy

Accuracy is an accurate calculation model that is formed to classify data correctly [13]. Formula (4) is a formula for accuracy.

$$accuracy = \frac{Correct\ Tag}{Total\ Word\ Tag} \tag{4}$$

2.8.2 Precision

Precision is the calculation of the accuracy between the data and the prediction results from the model. Equation (5) is the calculation formula for Precision [14].

$$precision = \frac{TP}{TP\ + FP} \tag{5}$$

2.8.3 Recall

Recall is a calculation model in determining the return of an information [15].

$$recall = \frac{TP}{TP\ + FN} \tag{6}$$

2.8.4 F1-Score

F1-Score is a performance matrix considering Recall and Precision results [15].

$$F1\ Score = 2 \times \frac{precision\ \times recall}{precision\ + recall} \tag{7}$$

3. RESULT AND DISCUSSION

In the evaluation phase of this research, there are 4 test scenarios to evaluate the system that has been built. Scenario 1 is a baseline test. Scenario 2 is a comparison test of each K value with K-Fold Cross Validation by testing the Corpus on each K value and taking the average value. Scenario 3 tests the results of the Testing accuracy by comparing methods other than the Hidden Markov Model, namely with Unigram. The fourth scenario is to provide tagging data that does not yet have a tag from the Telkom University news dataset that has been processed using the Hidden Markov Model formed from each Corpus.

3.1 Split Data

The first scenario is to determine the baseline or initial data that will be used for testing the next scenario. Testing this scenario is carried out on Corpus data between the Modified Corpus and the Original Corpus. The results of the first scenario can be seen in the table that has been created.

Table 3. Baseline Original Corpus

Data Ratio	Accuracy(Training)	Accuracy(Predict)	Precision(FW)	Recall(FW)	F1-Score(FW)
80:20	97.35	94.7	59	83	69
75:25	97.28	94.36	57	8	67
70:30	97.24	94.3	57	8	67

Table 4. Baseline Modified Corpus

Data Ratio	Accuracy(Training)	Accuracy(Predict)	Precision(FW)	Recall(FW)	F1-Score(FW)
80:20	97.23	94.7	72	85	78
75:25	97.22	94.58	71	86	78
70:30	97.22	94.4	7	86	77

From the results of testing scenario 1, the data ratio of 80:20 gets the best results compared to other data ratios. The results accuracy of 97.35% from the Original Corpus, while for Modified Corpus, it shows 97.23% for accuracy training. But, the Precision, Recall, and F1-Score of the Modified Corpus are greater than the Original Corpus. With that, the test results of scenario one will be used for testing the next scenario, which is split data with a ratio of 80:20

3.2 K-Fold Cross Validation

Building of Informatics, Technology and Science (BITS)
Volume 4, No 2, September 2022 Page: 588−594
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v4i2.1983

In scenario 2, we will test the effect of K-Fold on the performance of each Corpus, which will find the average evaluation result of each K.

Table 5. Result K-Fold Original Corpus

Fold	Accuracy(Avg)	Precision(Avg)	Recall(Avg)	F1-Score(Avg)
K = 3	81.66	86.73	81.66	82.94
K = 5	91.54	92.84	91.54	91.90
K = 10	92.91	93.82	92.91	93.17

Table 6. Result K-Fold Modified Corpus

Fold	Accuracy(Avg)	Precision(Avg)	Recall(Avg)	F1-Score(Avg)
K = 3	82.26	86.30	82.26	82.91
K = 5	91.66	92.46	91.67	91.73
K = 10	92.62	93.05	92.62	92.62

The results of testing scenario 2 prove that the value of K = 10 has the best average compared to the other average K values in each tested Corpus. Here, the results of the Original Corpus accuracy after 10 folds show better results than the Modus Corpus because the results from the Precision, Recall, and F1-Score calculations are the results of the whole Corpus, not from the FW tag, which is focused on this research.

3.2 Comparison Method of Testing Corpus

In scenario 3 testing, we will perform a test scenario by comparing the two methods to validate the original and modified Corpus. The technique in this corpus test compares the Hidden Markov Model method with the Unigram method, which in this test sees the results of precision accuracy up to F1-Score.

Table 7. Result Comparing Method with Original Corpus

Method	Accuracy(Predict)	Precision(FW)	Recall(FW)	F1-Score(FW)
Hidden Markov Model	94,7	59	83	69
Unigram	94.3	94	70	80

Table 8. Result Comparing Method with Original Corpus

Method	Accuracy(Predict)	Precision(FW)	Recall(FW)	F1-Score(FW)
Hidden Markov Model	94.7	72	85	78
Unigram	94.3	95	78	86

The results of scenario 3 testing show that each modified Corpus tagging model performs better but not too far from the Original Corpus. It can be seen from the accuracy that the F1-Score shows good results from the Modified Corpus of 52,342 tokens.

3.3 Implementation

In the last test, the model formed is tested to label words from the news dataset related to Telkom University, where the data is not contained in the Corpus so that it labels the real word. In this scenario, the tag results from fragments of sentences in the news dataset will be compared and evaluated for each tagging result.

Table 9. Example Test Sentence

Example Sentence
Times Higher Education (THE) merilis daftar THE Young University Rankings 2021. Data pemeringkatan ini merangkum daftar universitas terbaik dunia yang berusia kurang dari atau 50 tahun. Dari 475 universitas di 68 negara dan wilayah yang disaring THE, ada satu universitas asal Indonesia yang masuk daftar THE Young University Rankings 2021, yaitu Telkom University.

Table 10. Result Tagging Original Corpus

Result Tagging	Total Token	Tag Correct
('Times', 'NNP'), **('Higher', 'FW'), ('Education', 'FW'), ('(THE)', 'FW'), ('merilis', 'FW'), ('daftar', 'FW'), ('THE', 'FW'), ('Young', 'FW'), ('University', 'FW'), ('Rankings', 'FW'), ('2021.', 'FW'),** ('Data', 'NN'), ('pemeringkatan', 'NN'), ('ini', 'PR'), ('merangkum', 'VB'), ('daftar', 'NN'), ('universitas', 'NN'), ('terbaik', 'JJ'), ('dunia', 'NN'), ('yang', 'SC'), ('berusia', 'VB'), ('kurang', 'RB'), ('dari', 'IN'), ('atau', 'CC'), ('50', 'CD'), **('tahun.', 'CD')**, ('Dari', 'IN'), **('475', 'NN')**, ('universitas', 'NN'), ('di', 'IN'), ('68', 'CD'), ('negara', 'NN'), ('dan', 'CC'), ('wilayah', 'NN'), ('yang', 'SC'), **('disaring', 'NEG'),** ('THE,', 'MD'), ('ada', 'VB'), ('satu', 'CD'),	53	39

Building of Informatics, Technology and Science (BITS)
Volume 4, No 2, September 2022 Page: 588−594
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v4i2.1983

('universitas', 'NN'), ('asal', 'NN'), ('Indonesia', 'NNP'), ('yang', 'SC'), ('masuk', 'VB'), ('daftar', 'NN'), ('THE', 'NNP'), ('Young', 'NNP'), ('University', 'NNP'), ('Rankings', 'NNP'), **('2021,', 'Z')**, ('yaitu', 'SC'), ('Telkom', 'NNP'), ('University.', 'NNP')		

Table 11. Result Tagging Modified Corpus

Result Tagging	Total Token	Tag Correct
('Times', 'NNP'), **('Higher', 'Z')**, **('Education', 'SC')**, **('(THE)', 'MD'),** ('merilis', 'VB'), ('daftar', 'NN'), ('THE', 'NNP'), ('Young', 'NNP'), ('University', 'NNP'), **('Rankings', 'Z')**, **('2021.', 'CC')**, ('Data', 'NN'), ('pemeringkatan', 'NN'), ('ini', 'PR'), ('merangkum', 'VB'), ('daftar', 'NN'), ('universitas', 'NN'), ('terbaik', 'JJ'), ('dunia', 'NN'), ('yang', 'SC'), ('berusia', 'VB'), ('kurang', 'RB'), ('dari', 'IN'), ('atau', 'CC'), ('50', 'CD'), **('tahun.', 'CD')**, ('Dari', 'IN'), **('475', 'NN')**, ('universitas', 'NN'), ('di', 'IN'), ('68', 'CD'), ('negara', 'NN'), ('dan', 'CC'), ('wilayah', 'NN'), ('yang', 'SC'), **('disaring', 'NEG')**, **('THE,', 'MD')**, ('ada', 'VB'), ('satu', 'CD'), ('universitas', 'NN'), ('asal', 'IN'), ('Indonesia', 'NNP'), ('yang', 'SC'), ('masuk', 'VB'), ('daftar', 'NN'), ('THE', 'NNP'), ('Young', 'NNP'), ('University', 'NNP'), ('Rankings', 'NNP'), **('2021,', 'Z')**, ('yaitu', 'SC'), ('Telkom', 'NNP'), ('University.', 'NNP')	53	43

After the model is formed and the news dataset is ready to be processed, the model is tried for the implementation of sentences that are not in the Corpus 10 times. The results of each Corpus can be seen in tables 12 and 13.

Table 12. Result Tagging Telkom University News Original Corpus

No. News	Total Token	Total Correct	Total Incorrect	Accuracy
1	34	27	7	79.41
2	73	60	13	82.19
3	37	30	7	81.08
4	49	40	9	81.63
5	24	21	3	87.5
6	32	29	3	90.62
7	44	36	8	81.81
8	42	30	12	.71.42
9	53	39	14	73.58
10	38	27	11	71.05
			Average	80.03

Table 13. Result Tagging Telkom University News Original Corpus

No. News	Total Token	Total Correct	Total Incorrect	Accuracy
1	34	29	5	85.29
2	73	61	12	83.56
3	37	31	6	83.78
4	49	42	7	85.71
5	24	21	3	87.5
6	32	29	3	90.62
7	44	36	8	81.81
8	42	30	12	.71.42
9	53	43	10	81.13
10	38	28	10	73.68
			Average	82.45

In testing with Telkom University news, the model formed by each Corpus shows different results. The Original Corpus results have many tagging errors, but the results are still not in place compared to the Modified Corpus, which has tag accuracy on every word. This is because the effects of Precision, recall, and F1-Score are better than Corpus Original even though the accuracy from the Corpus is lower so that the accuracy of 10 news data is obtained, the average accuracy of Corpus Modification is 82.45 while for Corpus Original 80.03.

4. CONCLUSION

Based on the research and discussion results, the Corpus can show good performance after modifications are made by adding sentences that contain many acculturation languages found in online news media. In measuring the evaluation results, it can be seen that the difference is not too significant. Still, when looking at one of the "FW" tags which is the main focus of this study, it can show better Precision, recall, and F1-score results than the unmodified Corpus. To

Building of Informatics, Technology and Science (BITS)
Volume 4, No 2, September 2022 Page: 588−594
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v4i2.1983

validate the model made, a K-Fold Cross Validation was carried out where from the experiment K = 10 showed the best results from each model, and the Modified Corpus had better results than the Original Corpus. This effect can be seen when given news data that does not yet have a tag. Corpus modifications can place FW labels according to the word class contained in the text of the news sentence. From these results indicate that the model can perform tagging well. Suggestions for further research need to add a larger dataset to improve accuracy by considering all the tags in the dataset and taking source data from many media such as novels, books, social media, or other developing media. Other methods for implementing POS Tagger can also affect further research to find the best method for POS Tagger Indonesia.

REFERENCES

[1] A. Y. Rofiqi, "Clustering Berita Olahraga Berbahasa Indonesia Menggunakan Metode K-Medoid Bersyarat," *J. Simantec*, vol. 6, no. 1, pp. 25–32, 2017.

[2] Badan Pengembangan dan Pembinaan Bahasa, *Badan pengembangan dan Pembinaan Bahasa Kementerian pendidikan dan kebudayaan*. 2017.

[3] A. Z. Amrullah, R. Hartanto, and I. W. Mustika, "A comparison of different part-of-speech tagging technique for text in Bahasa Indonesia," *Proc. - 2017 7th Int. Annu. Eng. Semin. Ina. 2017*, 2017, doi: 10.1109/INAES.2017.8068538.

[4] S. K. Nambiar, A. Leons, S. Jose, and Arunsree, "POS Tagger for Malayalam using Hidden Markov Model," *Proc. 2nd Int. Conf. Smart Syst. Inven. Technol. ICSSIT 2019*, no. Icssit, pp. 957–960, 2019, doi: 10.1109/ICSSIT46314.2019.8987786.

[5] N. Sabloak, "Part-of-Speech (POS) Tagging Bahasa Indonesia Menggunakan Algoritma Viterbi," no. x, pp. 1–11, 2016.

[6] Ryan Armiditya Pratama, A. A. Suryani, and W. Maharani, "Part of Speech Tagging for Javanese Language with Hidden Markov Model," *J. Comput. Sci. Informatics Eng.*, vol. 4, no. 1, pp. 84–91, 2020, doi: 10.29303/jcosine.v4i1.346.

[7] I. G. M. H. Pradiptha and N. A. Sanjaya ER, "Building Balinese Part-of-Speech Tagger Using Hidden Markov Model (HMM)," *JELIKU (Jurnal Elektron. Ilmu Komput. Udayana)*, vol. 9, no. 2, p. 303, 2020, doi: 10.24843/jlk.2020.v09.i02.p18.

[8] A. Dinakaramani, F. Rashel, A. Luthfi, and R. Manurung, "Designing an Indonesian part of speech tagset and manually tagged Indonesian corpus," *Proc. Int. Conf. Asian Lang. Process. 2014, IALP 2014*, pp. 66–69, 2014, doi: 10.1109/IALP.2014.6973519.

[9] P. Alva and V. Hegde, "Hidden Markov model for POS tagging in word sense disambiguation," *2016 Int. Conf. Comput. Syst. Inf. Technol. Sustain. Solut. CSITSS 2016*, pp. 279–284, 2016, doi: 10.1109/CSITSS.2016.7779371.

[10] F. M. Hasan, N. Uzzaman, and M. Khan, "Comparison of different POS tagging techniques (n-gram, HMM and brill's tagger) for Bangla," *Adv. Innov. Syst. Comput. Sci. Softw. Eng.*, pp. 121–126, 2007, doi: 10.1007/978-1-4020-6264-3_23.

[11] F. Ramadhanti, Y. Wibisono, and R. A. Sukamto, "Analisis Morfologi untuk Menangani Out-of-Vocabulary Words pada Part-of-Speech Tagger Bahasa Indonesia Menggunakan Hidden Markov Model," *J. Linguist. Komputasional*, vol. 2, no. 1, p. 6, 2019, doi: 10.26418/jlk.v2i1.13.

[12] S. Rathod and S. Govilkar, "Survey of various POS tagging techniques for Indian regional languages," *Int. J. Comput. Sci. Inf. Technol.*, vol. 6, no. 3, pp. 2525–2529, 2015, [Online]. Available: www.ijcsit.com.

[13] D. E. Cahyani and M. J. Vindiyanto, "Indonesian part of speech tagging using hidden markov model - Ngram viterbi," *2019 4th Int. Conf. Inf. Technol. Inf. Syst. Electr. Eng. ICITISEE 2019*, pp. 353–358, 2019, doi: 10.1109/ICITISEE48480.2019.9003989.

[14] V. M. Patro and M. Ranjan Patra, "Augmenting Weighted Average with Confusion Matrix to Enhance Classification Accuracy," *Trans. Mach. Learn. Artif. Intell.*, vol. 2, no. 4, 2014, doi: 10.14738/tmlai.24.328.

[15] F. Rahmad, Y. Suryanto, and K. Ramli, "Performance Comparison of Anti-Spam Technology Using Confusion Matrix Classification," *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 879, no. 1, 2020, doi: 10.1088/1757-899X/879/1/012076.