

Building Synsets for Indonesian WordNet using ROCK (Robust Clustering Using Links) Algorithm

Mubarq Iqbal^{[1]*}, Moch. Arif Bijaksana^[2], Widi Astuti^[3]

Bachelor of Informatics Engineering, Faculty of Informatics^{[1], [2], [3]}

Telkom University

Bandung, Indonesia

mubarqiqball@student.telkomuniversity.ac.id^[1], arifbijaksana@telkomuniversity.ac.id^[2], widiwdu@telkomuniversity.ac.id^[3]

Abstract-- On the development of Indonesian WordNet, the synonym set is an important part that represents the similarity of meaning between words. Synonym sets are built using the Indonesian Thesaurus as the lexical database. After going through the extraction process from the Indonesian Thesaurus, we will get a synonym set that has a similarity or word sense between words. In general, the difference between WordNet and the dictionary is their main focus, in which the dictionary usually focuses on just one word, while in WordNet the focus is on the meaning of words and connectedness with other words. Explained in previous research, the constructions of synonym sets were done using several approaches, which is clustering to produce synonym sets and WSD (Word Sense Disambiguation). In this article, the approach used to produce synonym sets is the ROCK (Robust Clustering Using Links) algorithm, which uses similarity and link values. The resulting synonym sets will then be used for lexical database development. Therefore, the main focus of this article is to produce synonym sets through the clustering process and calculate their accuracy, using the F-Measure method involving the gold standard for performance calculation and evaluation.

Keywords: ROCK Clustering, Synonym Sets, WordNet

I. INTRODUCTION

WordNet, also known as Princeton WordNet [1], was developed by a lexicographer who produced it based on lexical data. Princeton WordNet is made manually and requires a lot of resources with language experts and time that produces high-quality WordNet [1]. Based on other research [2], WordNet contains 155.000 nouns (nouns), adjectives (adjectives), adverbs (adverbs), and verbs (verbs), then these words are grouped according to their meaning into synonym sets (synsets) or a collection of synonyms that have the same meaning. In its development, WordNet has made into several languages, one of them is Dutch WordNet [3], Russian WordNet [4], and Korean WordNet [5], which was built using available lexical resources. In WordNet, there is a structure that contains word information, word classes, and resolutions of all the word sets in the discussion which then become a single, interconnected entity. Pronouns are different from WordNet because they are synchronous or set of a synonym that has the same meaning. Synchronization is a basic concept that supports semantic relations in a lexical database. In 2010, synsets for Indonesian WordNet were built using a language dictionary developed at the Sepuluh Nopember Institute of Technology using hierarchical clustering techniques. Based on the results of the research conducted, a stronger and more suitable grouping

technique is needed for the construction of synsets for WordNet Indonesian. In this article, the ROCK (Robust Clustering Using Links) clustering technique will be used. This technique uses a similarity measure called link [6]. This technique is a developed hierarchical grouping technique and is felt to be very appropriate to the needs in the construction of synonym sets that prioritize the similarity of meaning in its development. Based on previous research, the use of ROCK clustering is highly recommended in the development of Indonesian synsets, because this clustering technique is very supportive to do the classification process itself using category attribute [6] [7]. These advantages will be utilized to build the system so it can produce better synsets because it is an important component in WordNet development. Therefore, the construction of synonym sets using ROCK clustering in this article is expected to produce a system that can provide better performance or accuracy than those previous research in producing synsets for Indonesian WordNet.

II. LITERATURE REVIEW

A. WordNet

WordNet is a large-scale electronic dictionary for English that was created in 1986 at Princeton University, where it developed continuously by George A. Miller. He was inspired by experiments in artificial intelligence that try to understand human semantic memory. WordNet, also known as Princeton WordNet [8], was developed by lexicographers and the results are made into a lexical database. Princeton WordNet is created manually and requires a lot of resources to have a high quality such as linguists and time [2]. The difference between WordNet and the language dictionary is the dictionary focuses on words while WordNet focuses on the meaning of the word. WordNet consists of the following classes of words such as nouns, verbs, adjectives, and adverbs.

B. Synsets (Synonym Sets)

Synsets or “synonym sets” are the main units used in WordNet. Synsets are set of one or more words that have the same meaning, or commonly called synonyms [9]. Each set member or set that can replace the use of the word. Every word that can replace another word in the same context cannot be from a different class of words, because some words can belong to more than one different word class [5] [10]. For example, the word “satu” in the Indonesian Thesaurus exists in both noun and numeric class.

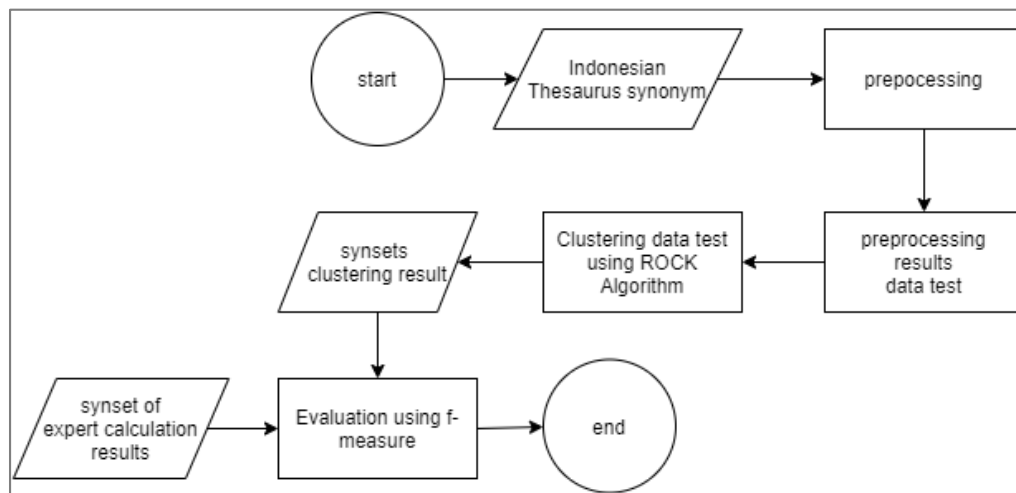


Fig. 1. System overview

C. Indonesian Thesaurus

Thesaurus is originated from the Greek word, which means “treasures”, and got developed on its meaning as “the book of information source”. A thesaurus is a dictionary of words that has interrelated meanings [11] [12]. Thesaurus consists of synonym and antonym relations. In the Indonesian Thesaurus, there were 48.484 items of Indonesian words. Thesaurus is different from a dictionary. While a dictionary can be used to search for information about the meaning of words, a thesaurus can be used to search for words that will be used to express user ideas. For example, if somebody wants to find another word for the word “aba-aba” in Indonesian Thesaurus, he can search for it on “aba-aba” which means “sign” in English. The purpose of using an Indonesian Thesaurus as a test data is because thesaurus is a large dictionary which has been used in several previous research [7] [13], also, the information contained in the thesaurus has been recognized by lexicographers and is an easy source to get and is provided by the Indonesian Language Center.

D. ROCK (Robust Clustering Using Links)

Clustering is a process of grouping data into a cluster or group so it contains objects that have a large similarity but have large dissimilarities with objects in other clusters [14] Or clustering is grouping data items into small groups so that each group has an essential equation [15]. Many clustering algorithms are intended to process numerical data, one of it is the Hierarchical Clustering algorithm that groups objects by creating a hierarchy where objects that have a large similarity will be placed in an adjacent hierarchy while objects that have large dissimilarities in the far apart hierarchy. However, problems arise when the algorithm is applied to data that has attribute values that are boolean or categorical [16]. The ROCK (Robust Clustering Using Links) clustering method uses a measure of similarity called “links” in forming clusters, unlike traditional clustering techniques such as Hierarchical Clustering techniques that use distance values [6]. Often, the results of the clustering process group object that do not have the same items and have a small

similarity value. To handle the problem of categorical data, in this article the ROCK clustering method will be used to cluster the data by grouping the data that has the most links or the same number of items with its neighbors, with the parameter number of clusters (k) and threshold value [16].

E. Gold Standard

Gold Standard used to find out how big is the correlation between the score issued by the machine or system to the relevance of the word being tested. The gold standard value is obtained from a set of human opinions. This value is used as a reference measurement of similarity between words. In this article, the gold standard will be used to validating the synonym sets is performed by lexical experts (lexicographers). The validation will be held very carefully so that it can be used as a comparison for the results of the system as a measure of accuracy.

F. F-Measure

F-Measure is a popular performance metric, especially for tasks with unbalanced test data [17]. The F-Measure method involves precision and recall. For the calculation of recall (R) and precision (P), can be seen in (1) and (2) [18] [19]. Human intuition is very important in determining the gold standard and in determining the threshold value for grouping the words based on the value of similarity obtained. The F-Measure method calculates multiple propositions multiplied by the results of the precision and the recall divided by the sum of both of them, can be seen in (3).

$$\text{Precision} = \frac{\text{the right number of prediction}}{\text{number of prediction}} \quad (1)$$

$$\text{Recall} = \frac{\text{the right number of prediction}}{\text{exact number}} \quad (2)$$

$$\text{F-Measure} = \frac{2(\text{precision} * \text{recall})}{\text{precision} + \text{recall}} \quad (3)$$

III. RESEARCH METHODS

A. System Overview

In this article, the system that will be built is a system that able to do a clustering process on a synonym set that has passed the previous extraction process. The clustering algorithm used in this study is the ROCK (Robust Clustering Using Links) algorithm. F-measure will be used as the testing method and gold standard as the validation with the synonym sets performed by lexical experts (lexicographers). The overview of the system that will be built can be seen on Fig. 1.

At the beginning of the system, words are chosen randomly from the Indonesian Thesaurus. Then the synonym set obtained from the Indonesian Thesaurus will be processed at the preprocessing stage to remove unusable character from the data test, then the synonym set will be grouped based on the results of the clustering process to get new synonym sets, the two synonym sets that have been grouped through the clustering process will become a new matrix. Then the synonym set from the results of the clustering process will be compared with the results from the linguist expert and evaluated using the F-Measure to calculate the performance of the clustering process.

B. Data Test

The data test will be built in the form of a dataset derived from the extraction process from the Indonesian Thesaurus, as many as 86 Indonesian words will be selected from the Indonesian Thesaurus. In the process of identifying prospective synsets, only pairs of words that have a commutative relationship in the thesaurus are considered valid. Words in a thesaurus that do not have a commutative relationship with other words, whether that word has an item or not, then both words are considered to have no synonym or invalid relationship. Words that do not have a pair in the thesaurus will also be considered valid synonyms. After the selection process, the set of words that have a valid synonym relationship will be store in the data test. Here are the words that will be used in the data test in this article.

TABLE 1. DATA TEST WORDS

Word			
aborsi	eka	handuk	kabung
abrasi	eks	harap	kafilah
acuh	asap	harfiah	lamaran
bahu	gas	ialah	madukara
bajul	gelar	idealis	maharani
bantahan	eksponen	jahiliah	mambang
cabuk	faksi	arwah	mampir
cadar	fana	agar	nagari
cahar	fiber	kembang	nahas
dam	fiksi	tenda	nasrani
dandang	gagu	jalin	oral
darat	gajih	jambe	orasi
edan	galat	kabin	binatang
tiup	simpan	langgar	gosok
tulis	lompat	sulam	ayam
Tikar	tukar	Ahad	tembak
pacul	tendang	padas	abah
pahala	ceker	qadim	minggu
Aduk	pekan	ampun	pukul

Word			
bentak	potong	Bisik	sumpah
bingkai	buat	tolong	
umpat	usap	tusuk	

For example, the words “agraria”, “Pertanahan” and “pertanian” will be used. In the synonyms of the word “agraria”, there are the words “Pertanahan” and “pertanian”. This means that “agraria” has a synonymous relationship with the word “Pertanahan” and the word “pertanian”. In the word “Pertanahan”, there is the word “agraria”. So the word “Pertanahan” has a synonymous relationship with the word “agraria”. Synonym relations are considered valid for the words “agraria” and “Pertanahan” which are commutative relationships.

The format used in the test data uses square brackets as a sign that the synonym set is valid and has a bidirectional relationship. If there is a synonym set that does not have a two-way relationship but has a two-way relationship with other words, it can be used as a candidate for a synonym set and use square brackets that are adjusted to all words that have a two-way relationship. for example ['bandrek', 'serbat'] in both words has a two-way relationship. While ['benalu', 'parasit', 'pasilan'] and ['benalu', 'parasit', 'sakat'] become two synonym sets because there is one word that is not connected or does not have a two-way relationship to the main word but has a two-way relationship with other words.

C. Clustering

ROCK (Robust Clustering using link) is a clustering analysis algorithm that develops from a hierarchical agglomerative clustering method to classify categorical data. This algorithm uses a similarity measure called links in the process of grouping, unlike traditional grouping techniques such as hierarchical clustering techniques that use distance values [6]. This algorithm naturally uses similarity values in clustering. Below are the steps of the algorithm [6] [20].

- 1) Calculating the similarity of a measure, the similarity value between the pair of object p_i with object p_j is calculated by the following formula.

$$\text{Sim}(p_i, p_j) = \frac{|p_i \cap p_j|}{|p_i \cup p_j|} \tag{4}$$

$\text{Sim}(p_i, p_j)$ is similarity values between p_i and p_j , $|p_i \cap p_j|$ is number of the same word between p_i and p_j , and $|p_i \cup p_j|$ is the number of words contained in p_i and p_j .

- 2) Determining neighbors, objects p_i and p_j are defined as neighbors if $\text{Sim}(p_i, p_j) \geq \theta$. Threshold (θ) is a parameter determined by researchers that can be used to control how close the relationship between objects. The value of θ that can be used is $0 < \theta < 1$. By default, the threshold value (θ) used is 0.5.
- 3) Count the number of links between groups with the following formula.

$$\text{link}(p_i, p_j) = |T_{p_i} \cap T_{p_j}| \tag{5}$$

Link(p_i, p_j) is number of neighbors p_i and p_j , T_{p_i} is the number of p_i neighbors, and T_{p_j} is the number of p_j neighbors.

- 4) Calculate the value of goodness measure between groups if the link $\neq 0$, with the following formula.

$$g(p_i, p_j) = \frac{\text{link}(p_i, p_j)}{a-b-c} \quad (6)$$

$$a = (n_i + n_j)^{(1+2(\theta))} \quad (7)$$

$$b = (n_i)^{(1+2(\theta))} \quad (8)$$

$$c = (n_j)^{(1+2(\theta))} \quad (9)$$

$$f(\theta) = \frac{1-\theta}{1+\theta} \quad (10)$$

From the formulas above, $g(p_i, p_j)$ is the value of good measure when p_i and p_j is the number of member of the group p_i , and n_j is the number of member of the group p_j .

- 5) Merge pairs that have the largest goodness measure value into groups, then add links between groups that are grouped and update the new goodness measure values.
6) Do step 5 to form and groups or until there are no more links between groups.

D. Calculation of the Evaluation

Evaluation of calculations in this article uses the F-Measure which aims to measure the accuracy of the results of clustering that has been done. As explained in the literature review, F-Measure uses precision and recall to do calculations.

Precision is taken from the calculation of the correct number of words in the synsets that have been generated by the system compared to the gold standard resulting from manual calculations by humans, divided by the number of words in the synsets produced.

Whereas recall is taken from the calculation of the number of correct words in the synsets the system has been generated compared to the results of manual calculations divided by the number of words in synsets that have been calculated manually by humans or the gold standard.

Then the precision and recall values that have been obtained will be calculated using the formula in equation (3) to get the accuracy value of the system that has been built, as can seen from the formula the accuracy score is obtained from twice the value of precision times recall divided by recall plus precision.

IV. RESULT AND DISCUSSION

A. Testing Scenarios

The testing method used in the evaluation process of this study is the F-Measure and gold standard methods as a comparison. The method will compare the synonym sets produced by the program from the clustering process with the synonym sets produced by the results of validation by experts

or the gold standard. In addition to using the F-Measure, an experiment will be held by changing the threshold value (θ) from a range of 0.1 to 0.9. The experiment does not reach the value 1.0 because there will be a divider with a value of 0, causing the program to stop in the middle of the clustering process. The experiment was held to find the threshold (θ) with what value can produce the correct synonym sets and have better performance.

B. Testing Results

The results of tests that have been done by conducting a clustering process using the scenarios that has been discussed in the previous section. The results can be seen in the following table 2.

TABLE 2. TESTING RESULTS

Threshold	Maximum Goodness Measure	Number of loops	Number of synsets
0.1	0.08	7	59
0.2	0.14	7	61
0.3	0.21	4	75
0.4	0.34	2	90
0.5	0.54	2	93
0.6	0.85	2	98
0.7	1.41	1	104
0.8	2.57	1	104
0.9	6.14	1	104

Based on table 2, it can be seen that the greater the threshold value, the fewer the number of loops. This is caused by the threshold value which is the minimum equation between the two synsets to obtain neighbors, the greater the threshold value, the fewer synsets can be seen through the search process. So as discussed before, looping will only stop compiling when there are no more links between synchronizations. The smaller the threshold, the more small synsets that will be generated, because the words that have obtained relatively small ones are combined by the clustering process. After getting the test results in table 2, we get some of the best threshold values that can be seen in table 3 below.

TABLE 3. TESTING RESULTS USING F-MEASURE

Result of testing	Precision	Recall	F-Measure
Before clustering	92.11%	102.94%	97.22%
Threshold 0.4	86.54%	88.24%	87.38%
Threshold 0.5	87.74%	91.18%	89.42%
Threshold 0.6	89.91%	96.08%	92.89%

After passing the clustering process the accuracy calculation process will be carried out using the f-measure which is the system evaluation. Accuracy calculations will be performed using the gold standard synsets produced by experts. The testing analysis process will use threshold values of 0.4, 0.5, and 0.6 as the object of analysis because they are the three threshold values that produce the largest recall value that does not exceed one hundred percent and are the best threshold value in generating synsets. This analysis process is carried out in order to find some facts or information from experiments conducted using the ROCK algorithm.

C. Testing Results Analysis

In the Rock algorithm the threshold value is something that is very influential during the process of clustering, one of which is to calculate the value of the goodness measure used to determine whether two synsets will be grouped or not. Threshold value is also very influential in the cessation of the clustering process, which is affecting the number of links between synset because the clustering process will only stop when there are no more links or neighbors in the test data that are processed through clustering.

Based on the results of experiments conducted, it was found that the threshold values of 0.4, 0.5, and 0.6 can produce the best recall value that does not exceed one hundred percent and produce the most synsets, this accuracy value is better when compared with previous research using various methods [21] [22]. A threshold value of 0.4 produces a total of 90 synsets with an F-Measure of 87.38, a precision level of 86.54, and a recall value of 88.24. The threshold value of 0.5 produces 93 synsets with an F-Measure of 89.42, a precision level of 87.74, and a recall value of 91.18. Finally, the threshold value 0.6 produces 98 synsets with an F-Measure value of 92.89, a precision value of 89.91, and a recall value of 96.08. the greater the threshold value, the greater the value of the goodness measure produced, this can be seen in table 2. This shows why the threshold value greatly influences the synsets generated by the ROCK (Robust Clustering Using Links) process clustering algorithm.

Based on the results obtained, it is known that the greater the threshold value, the greater the F-Measure, precision, and recall value generated by the program. The accuracy calculation value discussed earlier is due to the greater threshold value, the merging between two or more synsets when going through the clustering process, especially when comparing similarity values with threshold values to determine whether these two synsets are neighbors or not experience a reduction in the amount due to the threshold value as determining neighbors or not getting bigger. As discussed earlier, this number of neighbors will affect the number of links defined as the number of neighbors of a synset.

From the number of synsets produced by the three previous threshold values, it is known that for a threshold value of 0.3 resulting in 90 synsets there are 18 synsets from the test data that are combined and produce nine new synsets with four synsets of which are not appropriate and five synsets are following validation synsets, the threshold value is 0.5 produces 93 synsets, 14 synsets from the test data are combined and produce seven new synsets with four synsets including those that do not match and three synsets according to the validation synsets, and a threshold value of 0.6 produces 98 synsets, there are eight synsets from the test data that are combined and produce four new synsets with three synsets which are not appropriate and one synset are following validation synsets.

Synsets that are not compatible with validation synsets are caused by many similarities in the words of this synsets even though the main words are different, for example, the

word "asap" as the main word with the contents of the synsets ['asap', 'gas'] and "gas" as the main word with the contents of the synsets ['gas', 'asap']. These two synsets are then combined into one by the program because they have the same two words as members and are considered identical. From the results previously written, it can be seen that the threshold value of 0.4 produces the newest synsets, and with the most number of synsets that are following the validation synsets.

The number of loops produced is also calculated a little with a fairly large accuracy value, with all three threshold values both having a loop as much as two times during the clustering process. When compared with previous research [21] [22], the number of loops counted quite a lot namely 13 times loops [21] and 15 times loops [22] of the best results made by the results of the research process which has been done. When compared to previous research [22] which only processed 50 synsets of test data with 114 synsets of test data in this study, the less time needed for the ROCK algorithm to complete the process. So the algorithm can be used to process large amounts of data with less time to produce synsets.

V. CONCLUSION

Based on the results of tests conducted in the previous chapter, it can be concluded that to produce synonym sets better the clustering process is done with a threshold range of 0.1 to 0.9. It is known that the best threshold value is 0.4 because it has the most number of synsets that match the most validation synsets data compared to another. At a threshold value of 0.4, the total number of synonym sets generated is 90 synonym sets out of a total of 114 synonym sets of test data, from the evaluation that had been done resulting in an accuracy value of 87.38% using the F-Measure. This accuracy value is better when compared with previous research.

Based on the number of loops performed during the clustering process the ROCK algorithm, does not have more loops than the other algorithm used in the previous research, so the algorithm can be used to process large amounts of data with less time to produce synsets. So after this research is done it can be said that the ROCK (Robust Clustering Using Links) algorithm is better than the algorithms used previously to produce synsets for Indonesian wordnet.

REFERENCES

- [1] G. A. Miller, R. Beckwith dan C. Fellbaum, "Introduction to WordNet: An on-line lexical database," *International Journal of lexicography*, vol. 3, no. 4, pp. 235-244, 1990.
- [2] C. Fellbaum, "WordNet," *The encyclopedia of applied linguistics*, no. Wiley Online Library, 1998.
- [3] M. Postma, E. v. Miltenburg, R. Segers, A. Schoen dan P. Vossen, "Open Dutch WordNet," dalam *Proceedings of the Eight Global Wordnet Conference*, Bucharest, 2016.
- [4] N. Loukachevitch, G. Lashevich, A. Gerasimova, V. Ivanov dan B. Dobrov, "Creating Russian WordNet by Conversion," dalam *Computational Linguistics and Intellectual Technologies: papers from the Annual conference "Dialogue"*, 2016, p. 22.
- [5] D. M. Bikel, "Automatic WordNet mapping using word sense

- disambiguation,” dalam *2000 Joint SIGDAT Conference on Empirical Methods in Natural Language Processing and Very Large Corpora*, 2000, pp. 142--147.
- [6] S. Guha, R. Rastogi dan K. Shim, “ROCK: A Robust Clustering Algorithm for Categorical Attributes,” *Information systems*, vol. 25, no. Elsevier, pp. 345--366, 1999.
- [7] A. Saputra dan G. , “Building synsets for Indonesian Wordnet with monolingual lexical resources,” dalam *2010 International Conference on Asian Language Processing*, Harbin, IEEE, 2010, pp. 297--300.
- [8] G. A. Miller, “WordNet: a lexical database for English,” *Communications of the ACM*, vol. 38, no. 11, pp. 39--41, 1995.
- [9] C. Havasi, R. Speer dan J. Alonso, “ConceptNet 3: a flexible, multilingual semantic network for common sense knowledge,” dalam *Recent advances in natural language processing*, Citeseer, 2007, pp. 27--29.
- [10] N. Taghizadeh dan H. Faili, “Automatic Wordnet Development for Low-Resource Languages using Cross-Lingual WSD,” *Journal of Artificial Intelligence Research*, vol. 56, pp. 61--87, 2016.
- [11] G. “Akuisisi Gloss Berbasis Ekstraksi Sinonim Set Menggunakan Supervised Learning,” Sepuluh Nopember Institute of Technology, Surabaya, 2016.
- [12] E. Endarmoko, *Tesaurus Bahasa Indonesia Second edition*, Jakarta: Gramedia Pustaka Utama, 2016.
- [13] I. P. P. ananda, M. A. Bijaksana dan I. Asror, “Pembangunan Synsets untuk WordNet Bahasa Indonesia dengan Metode Komutatif,” *eProceedings of Engineering*, vol. 3, p. 5, 2018.
- [14] P. Berkhin, “A survey of clustering data mining techniques,” dalam *Grouping multidimensional data*, Springer, 2006, pp. 25--71.
- [15] H. Patil dan R. S. Thakur, “Document Clustering: A Summarized Survey,” dalam *Information Retrieval and Management: Concepts, Methodologies, Tools, and Applications*, IGI Global, 2018, pp. 47--64.
- [16] N. S. Belinda, I. R. Hg dan H. Yoza, “Penerapan Analisis Cluster Ensemble dengan Metode Rock Untuk Mengelompokkan Provinsi di Indonesia Berdasarkan Indikator Kesejahteraan Rakyat,” Andalas University, Padang, 2019.
- [17] Y. Nan, K. Chai, W. Lee dan H. Chieu, “Optimizing F-measure: A tale of two approaches,” *arXiv preprint arXiv:1206.4625*, 2012.
- [18] M. Arora, U. Kanjilal dan D. Varshney, “Evaluation of information retrieval: precision and recall,” *International Journal of Indian Culture and Business Management*, vol. 12, pp. 224--236, 2016.
- [19] D. Hand dan P. Christen, “A note on using the F-measure for evaluating data linkage algorithms,” *Pre-print NII6047, Isaac Newton Institute, Cambridge, UK*, 2016.
- [20] T. D. Latifah, C. Suhaeni dan R. Anisa, “Segmentasi Pelanggan Susu Formula Menggunakan Cluster Ensemble Berbasis Algoritme ROCK,” Bogor Agricultural University, Bogor, 2018.
- [21] D. P. Sari, M. A. Bijaksana dan T. Suhardijanto, “Membangun synset untuk wordnet bahasa indonesia menggunakan hierarchical clustering,” Telkom University, Bandung, 2019.
- [22] L. D. Anggaraini, M. A. Bijaksana dan D. Pusparani, “Analisis Pembangunan Word Sense pada WordNet Bahasa Indonesia Menggunakan Metode Hierarchical Clustering,” Telkom University, Bandung, 2019.