

PENGUKURAN KEMIRIPAN MAKNA KALIMAT DALAM BAHASA INDONESIA MENGGUNAKAN METODE PATH

Yessy Caterina ^{1)*}, Muhammad Ainul Yaqin ²⁾, Syahiduz Zaman ³⁾

Teknik Informatika Universitas Islam Negeri Maulana Malik Ibrahim Malang ^{1),2),3)}
15650090@student.uin-malang.ac.id ^{1)*}, yaqinov@ti.uin-malang.ac.id ²⁾, syahid@ti.uin-malang.ac.id ³⁾

Abstrak

Pengukuran kemiripan makna kalimat bertujuan untuk didapatkan nilai kemiripan antar satu kalimat dengan kalimat yang lain. Nilai kemiripan yang didapatkan tersebut kemudian dapat diimplementasikan untuk pengembangan sistem yang berbasis *matching sentence*, misalnya *search engine*. Pada beberapa penelitian yang dilakukan sebelumnya, membahas mengenai efektivitas algoritma semantik dalam mengukur kemiripan makna kalimat dalam bahasa Inggris. Sedangkan, penelitian ini mencari kemiripan makna antar kalimat dalam bahasa Indonesia. Dataset dalam pencarian dan pengukuran makna kalimat pada penelitian ini menggunakan *sinonimkata.com* yang berupa node atau percabangan. Pada perhitungan kemiripan makna kalimat menggunakan *WordNet*, pendekatan yang digunakan ialah *Wu Palmer*, *Lin*, *Path*, *Resnik*, dan *Hirst-St (HSO)*. Pada penelitian ini menggunakan pendekatan *path* karena paling sesuai untuk menghitung jumlah node atau relasi yang terhubung antar node lain dalam *sinonimkata.com*. Pengukuran ini dilakukan dengan 5 eksperimen yakni, berdasarkan susunan kalimat kata kerja – kata benda, kalimat aktif – kalimat pasif (makna sama), 2 kalimat aktif, 2 kalimat pasif, dan kalimat aktif – kalimat pasif (makna berbeda). Menghitung nilai kemiripan kata diurai dengan kriteria kelas kata kerja, dan kata benda kemudian dihitung berdasarkan kontekstual menggunakan pendekatan *path* yang kemiripan katanya dicari menggunakan *sinonimkata.com*. Dari proses perhitungan kelima eksperimen tersebut, dapat dihasilkan kemiripan kalimat dalam bahasa Indonesia yang memiliki tingkat kemiripan yang tinggi bernilai 0,875 pada eksperimen kriteria kalimat dengan susunan kata kerja – kata benda.

Kata kunci: kemiripan makna kalimat, *sinonimkata.com*, *path*

Abstract

[Measurement of the sentence similarity in Indonesian using the path method] Measurement of the sentence similarity aims to obtain the value of similarity between one sentence and another sentence. The similarity value obtained can be implemented for the development of a based system on matching sentence, for example search engines. In several previous studies, discussed the effectiveness of semantic algorithms in measuring the sentence similarity meanings in English. Meanwhile, this research looks for similarities of meaning of one sentence to another in Indonesian. The dataset in the search and measurement of the sentence similarity in this study using *sinonimkata.com* in the form of nodes or branches. In the calculation of the sentence similarity meanings using *WordNet*, the approach used is *Wu Palmer*, *Lin*, *Path*, *Resnik*, and *Hirst-St (HSO)*. This study uses a *path* approach because it is best suited to calculate the number of nodes or relationships connected between other nodes in *sinonimkata.com*. This measurement was done with 5 experiments, based on the composition of verb sentences – nouns, active sentences – passive sentences (same meaning), 2 active sentences, 2 passive sentences, and active sentences – passive sentences (different meanings). Calculating the likeness value of a word is parsed with the criteria of a verb class, and the noun is then calculated based on contextual using a *path* approach whose similarity is searched using *sinonimkata.com*. From the calculation process of the five experiments, it can be produced a sentences similarity in Indonesian that have a high level of similarity worth 0.875 in the experiment of sentence criteria with the arrangement of verbs – nouns.

Keywords: sentence similarity, *sinonimkata.com*, *path*

1. PENDAHULUAN

Kalimat adalah kumpulan dari beberapa kata yang tersusun untuk mengungkapkan suatu konsep pikiran dan perasaan. Suatu kalimat terdiri atas

beberapa unsur pembentuk kalimat yang terdiri atas unsur subjek dan predikat. Subjek dalam sebuah kalimat dapat berupa nama orang, hewan, sapaan, benda, dan lain-lain. Sedangkan predikat ialah yang

menandai apa yang telah diucapkan ataupun dituliskan oleh pihak pertama. Predikat dapat dikategorikan kata kerja (*verb*), kata benda (*noun*), kata sifat (*adjectiva*), kata depan (*preposisi*), dan kelas lainnya. Dalam satu kalimat dapat bersifat aktif yang berarti subjek aktif melakukan suatu pekerjaan, dan bersifat pasif yang berarti subjek dikenai suatu pekerjaan.

Dalam perbendaharaan bahasa terdapat kumpulan berbagai kata yang memiliki kemiripan makna. Kemiripan makna yaitu kata yang berbeda namun memiliki arti yang sama atau mirip. Membandingkan kemiripan antar kata dapat dilakukan dengan menentukan kata yang berada dalam kelas yang sama. Gabungan dari kemiripan antar kata dapat membentuk kemiripan makna kalimat. Tujuan dari pengukuran kemiripan makna kalimat ini untuk mendapatkan nilai kemiripan makna antar satu kalimat dengan kalimat yang lain. Manfaat dari nilai kemiripan yang didapatkan tersebut, dapat diimplementasikan untuk pengembangan sistem yang berbasis matching sentence, misalnya pada search engine.

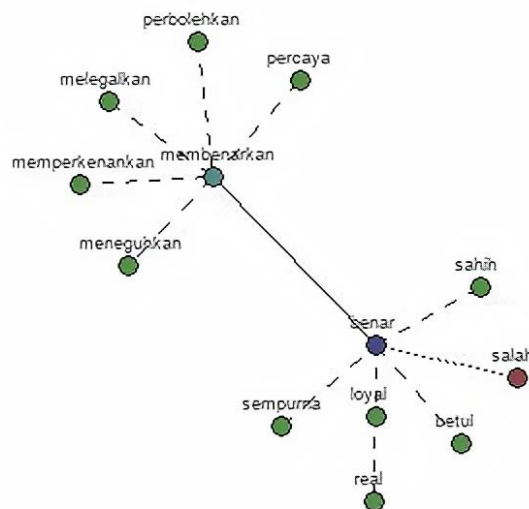
Perhitungan kemiripan makna antar kalimat dapat dilakukan dengan berbagai cara, yaitu metode Wu Palmer (WUP), Jiang-Conrath (JCN), Lin, Path, Resnik, dan Hirst-St (HSO) [1]. Perhitungan kemiripan tersebut menggunakan WordNet. WordNet merupakan basis pengetahuan dalam bentuk *lexical database* yang menyimpan arti kata dan saling berhubungan secara konseptual [2]. Kata benda, kata kerja, dan kata sifat diorganisasikan ke dalam *synonym sets* (*synset*) [3]. Dalam pendekatan *Wu Palmer* mampu mengukur relevansi antar dokumen ataupun antar term [4] [5]. *Lin* dinilai dari tiga asumsi, yaitu dua makna yang mirip akan sama, dua makna yang sama maka sedikit kemiripannya, dan kesamaan maksimum terjadi ketika dua makna yang identik [6]. *Path* pada prinsipnya dirancang untuk bekerja dengan struktur hierarki dan varian dari metode *distance* [7].

Penelitian ini berbeda dengan beberapa penelitian terdahulu, sebagai contoh pada penelitian yang dilakukan oleh Husni Thamrin dan Atiqa Sabardilla yang membahas mengenai efektivitas algoritma semantik dengan keterkaitan kata dalam mengukur kemiripan teks bahasa Indonesia [8]. Sedangkan pada penelitian ini mencari kemiripan makna kalimat satu dengan kalimat lain dalam bahasa Indonesia.

Pada paper ini kemiripan kata dapat dicari dengan menggunakan dataset yang disediakan oleh sinonimkata.com, yakni kamus tesaurus bahasa Indonesia dan bahasa Inggris yang terdiri dari sinonim, antonim, maupun arti dari suatu kata. Bentuk dari pencarian kata berupa percabangan atau node yang memiliki kemiripan maupun lawan kata. Contoh hasil pencarian kata dalam sinonimkata.com.

Gambar 1 merupakan hasil pencarian kemiripan kata yang membentuk percabangan. Dari hasilnya akan muncul beberapa garis percabangan kata yang terdiri dari kata dasar, dan kata yang memiliki makna hampir sama ditandai dengan garis dan node hijau. Selain itu,

terdapat node merah yang berarti antonim atau lawan katanya.



Gambar 1. Hasil pencarian kata dalam sinonimkata.com

Pada penelitian yang dilakukan oleh Gandhis Ulta dan M. Ainul Yaqin, mengukur kemiripan kontekstual menggunakan *Word Similarity for Java* (WS4J) dengan pendekatan *wu palmer*, *lin*, dan *path*. Diketahui nilai word similarity dan nilai bobot kriterianya menggunakan AHP untuk menentukan nilai *sentence similarity* [9]. Dalam penelitian ini mencari kemiripan kata dalam Bahasa Indonesia menggunakan sinonimkata.com. Pendekatan yang digunakan dalam menghitung nilai kemiripan kata, yaitu pendekatan *Path* karena yang paling sesuai digunakan untuk menghitung relasi atau jumlah node yang menghubungkan antar node dalam taksonomi [10]. Kemudian menghitung nilai kemiripan makna antar kalimat dengan *sentence similarity* dengan nilai pembobotan kata pada kalimat yang telah ditentukan sesuai dengan kriteria kelas.

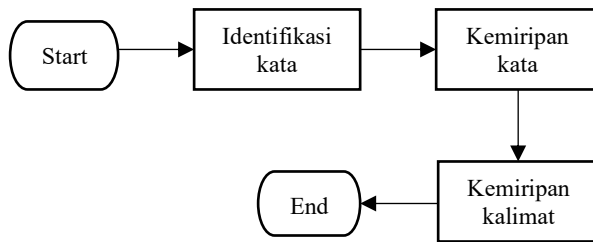
2. BAHAN DAN METODE

Bahan yang digunakan dalam mencari hasil kemiripan kalimat ialah, sinonimkata.com. Dalam sinonimkata.com hanya dapat mendeteksi kata yang bersifat aktif atau kata kerja, jika kata yang dicari merupakan kata pasif, maka tidak dapat terdeteksi.

Dalam penelitian ini menggunakan 5 eksperimen dalam mengukur kemiripan makna kalimat, dengan susunan kalimat yang terdiri dari:
 Eksperimen 1 : kata kerja (mirip) dan kata benda.
 Eksperimen 2 : 2 kalimat aktif (makna mirip)
 Eksperimen 3 : kalimat aktif dan kalimat pasif (makna mirip)
 Eksperimen 4 : 2 kalimat pasif (makna mirip)
 Eksperimen 5 : kalimat aktif dan kalimat pasif (makna berbeda)

Dari eksperimen tersebut, masing-masing dicari kemiripannya dalam sinonimkata.com.

Alur proses penelitian mulai dari identifikasi kata sampai menghitung kemiripan kalimat seperti pada gambar dibawah ini:



Gambar 2. Prosedur penelitian

Gambar 2 merupakan alur penelitian pengukuran kemiripan kalimat sebagai tujuan penelitian ini. Proses identifikasi kata diperlukan dalam penentuan kriteria sesuai kelasnya dan pencarian dalam sinonimkata.com. Proses perhitungan kemiripan kata dapat dihitung setelah proses sebelumnya telah didapatkan kata yang mirip. Setelah proses perhitungan kemiripan kata, dapat dihitung kemiripan antar kalimat yang menjadi percobaan dalam paper ini.

Metode yang digunakan dalam penelitian ini yakni, metode *Path* karena metode kemiripan makna kalimat yang paling sesuai dengan perpaduan dalam perhitungan node atau relasi dalam sinonimkata.com.

2.1. Identifikasi Jenis Kata

Prosedur pertama yang dilakukan ialah identifikasi kata dengan menentukan sesuai kriteria kelasnya baik itu kelas kata kerja atau kelas kata benda. Penentuan kata kerja dan kata benda dicari sesuai dalam sinonimkata.com untuk mendapatkan kemiripan kata. Contoh kalimat yang digunakan ialah:

Kalimat 1: melengkapi data barang

Kalimat 2: melengkapi informasi produk

Pada contoh kalimat 1 dan kalimat 2 sama-sama terdiri dari tiga susunan kata. Dari susunan kata tersebut masing-masing dicari kelasnya yang termasuk dalam kata benda ataupun kata kerja.

Tabel 1. Hasil kelas kata pada sinonimkata.com

Kata	Jenis kata
Melengkapi	Kata kerja (v)
Data	Kata benda (n)
Barang	Kata benda (n)
Informasi	Kata benda (n)
Produk	Kata benda (n)

Tabel 1 menunjukkan hasil pencarian jenis kelas kata pada sinonimkata.com. Kemudian, masing-masing susunan kata dalam kalimat 1 dibandingkan kalimat 2 dan dicari kemiripan katanya, jika dua kata memiliki kelas yang sama dapat dicari kemiripannya, jika jenis kelas berbeda maka tidak memiliki kemiripan.

2.2. Metode Path

Metode *Path* merupakan salah satu perhitungan kemiripan antar kata yang pada prinsipnya dirancang untuk bekerja dengan struktur hierarki. Persamaannya sebagai berikut [6]:

$$sim(C1, C2) = 2 \times Max(C1, C2) - SP \quad (1)$$

$$path = \frac{1}{path_length}$$

Persamaan (1) untuk menghitung kemiripan kata menggunakan metode *path*. *Max* adalah panjang jalur maksimum antara C1 dan C2 dalam taksonomi, dan *SP* adalah jalur pendek yang menghubungkan (jumlah minimum tautan) antara C1 dan C2.

Contoh perhitungan metode path:

Kalimat 1: melengkapi data barang

Kalimat 2: melengkapi informasi produk

Dalam pencarian kata 'data' dan 'informasi' masing-masing memiliki jumlah jalur sama, yakni 2.

$$sim(C1, C2) = 2 \times Max(C1, C2) - SP$$

$$= 2 \times 2 - 2$$

$$= 2$$

$$path = \frac{1}{path_length}$$

$$= \frac{1}{2}$$

$$= 0,5$$

2.3. Kemiripan Kalimat

Sentence similarity merupakan pengukuran kemiripan *semantic* berdasarkan makna kata. Dalam perhitungan persamaan *contextual*, menggunakan algoritma *Path* untuk menghitung persamaan kata pada kalimat, kemudian menentukan bobot nilai setiap struktur pembentuk kalimat, yaitu *verb*, *adjective*, dan *noun* menggunakan algoritma AHP yang telah ditentukan pada penelitian yang dilakukan oleh Gandhis Ulta dan Muhammad Ainul Yaqin [9]. Hasil yang didapat kemudian digunakan untuk menentukan persamaan kalimat pada persamaan *contextual*.

$$sentence\ sim = (N_{max} \times B_N) + (V_{max} \times B_V) \quad (2)$$

Persamaan *sentence similarity* (2) merupakan persamaan untuk menghitung nilai kemiripan kalimat. N_{max} merupakan nilai maksimal dari kriteria *Noun*. V_{max} merupakan nilai maksimal dari kriteria *Verb*. Sedangkan untuk B_N merupakan bobot nilai dari kriteria *Noun*, dan B_V merupakan bobot nilai dari kriteria *Verb*. Contoh perhitungan *sentence similarity*:

$$sentence\ sim = (N_{max} \times B_N) + (V_{max} \times B_V)$$

$$= (0,5 \times 0,25) + (1 \times 0,75)$$

$$= 0,875$$

3. HASIL DAN PEMBAHASAN

3.1 Identifikasi Jenis Kata

Mengidentifikasi kata dengan kriteria kelas kata kerja dan kata benda sesuai Kamus Besar Bahasa Indonesia. Kemudian, dicari kemiripan kata yang menunjukkan kelas yang sama, yakni kata kerja dengan kata kerja, dan kata benda dengan kata benda, menggunakan sinonimkata.com.

Tabel 2. Contoh kalimat percobaan dalam penelitian

Eksperimen	Contoh kalimat
1	- Melengkapi data barang - Melengkapi informasi produk
2	- Kami mengonsumsi makanan bergizi - Kami menggunakan barang bermanfaat
3	- Dia mengonsumsi makanan bergizi - Makanan bergizi dikonsumsi olehnya
4	- Penerbangan dibatalkan lantaran kabut - Seminar tertunda akibat pandemi
5	- Aku menunggu pesannya sejak semalam - Mereka telah bertemu kemarin malam

Tabel 2 merupakan susunan 5 eksperimen yang akan dicari nilai kemiripannya dengan masing - masing kriteria yang berbeda.

Eksperimen 1 mencari hasil kemiripan makna kalimat dengan kriteria susunan kalimat berdasarkan kata kerja (mirip) dan kata benda. Contoh kalimat:

Kalimat 1: melengkapi data barang

Kalimat 2: melengkapi informasi produk

Dari kalimat tersebut dicari kemiripan kata sesuai kelas yang sama dan dilakukan pencarian kata yang mirip hingga maksimal lima kali proses. Jika tidak ditemukan sinonim yang sama, maka dapat dikatakan bahwa kedua kata tersebut tidak memiliki kemiripan.

Tabel 3. Hasil percabangan kalimat 1 dan 2

Kalimat 2	Kalimat 1		
	melengkapi (v)	data (n)	barang (n)
melengkapi (v)	melengkapi	-	-
informasi (n)	-	bahan	
produk (n)	-		komoditas

Tabel 3 merupakan hasil kemiripan kata pada kalimat 1 dan kalimat 2.



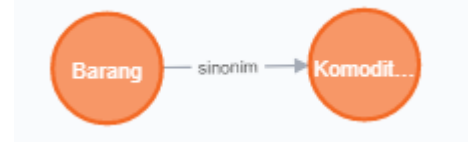
Gambar 3. Graf kata 'informasi'



Gambar 4. Graf kata 'data'



Gambar 5. Graf kata 'produk'



Gambar 6. Graf kata 'barang'

Gambar 3 merupakan hasil pencarian percabangan kata 'informasi' yang dipadukan dengan Gambar 4 hasil percabangan kata 'data' yang terdapat kemiripan kata, yakni 'bahan'. Gambar 5 merupakan hasil percabangan kata 'produk' yang dipadukan dengan Gambar 6 percabangan kata 'barang' yang memiliki kemiripan kata 'komoditas'.

Eksperimen 2 menggunakan susunan 2 kalimat aktif (makna mirip). Contoh kalimat:

Kalimat x: kami mengonsumsi makanan bergizi

Kalimat y: kami menggunakan barang bermanfaat

Pada kalimat x dan y merupakan susunan kalimat aktif, sedangkan kalimat z merupakan kalimat pasif.

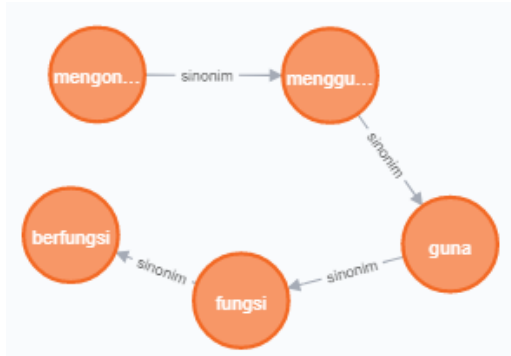
Tabel 4. Hasil percabangan kalimat x dan y

Kalimat y	Kalimat x		
	mengonsumsi	makanan	bergizi
menggunakan	menggunakan	-	-
barang	-	-	-
bermanfaat	berfungsi	-	-

Tabel 4 merupakan hasil kemiripan kata pada kalimat x dan kalimat y dengan sinonimkata.com.



Gambar 7. Graf kata ‘mengonsumsi’



Gambar 8. Graf kata ‘mengonsumsi’ kedua



Gambar 9. Graf kata ‘bermanfaat’

Gambar 7 merupakan hasil pencarian percabangan kata ‘mengonsumsi’ yang dipadukan dengan kata ‘menggunakan’. Gambar 8 merupakan hasil pencarian percabangan kata ‘mengonsumsi’ yang dipadukan dengan Gambar 9 percabangan kata ‘bermanfaat’ yang memiliki kemiripan, yakni kata ‘berfungsi’.

3.2 Menghitung Kemiripan Kata

Perhitungan kemiripan kata menggunakan pendekatan *path*. Dari pencarian kemiripan kata yang dilakukan sebelumnya, dibandingkan antara percabangan kedua kata tersebut dengan melihat jumlah jalur node awal sampai akhir.

a. Pada eksperimen 1 dengan kalimat berikut:

Kalimat 1: melengkapi data barang

Kalimat 2: melengkapi informasi produk

Perhitungan kemiripan kata kalimat 1 dan kalimat 2:

Dalam pencarian kata ‘data’ dan ‘informasi’ masing-masing memiliki jumlah jalur sama, yakni 2. Nilai kemiripan kata ‘data’ dan ‘informasi’:

$$\begin{aligned} sim(C1, C2) &= 2 \times Max(C1, C2) - SP \\ &= 2 \times 2 - 2 \\ &= 2 \end{aligned}$$

$$path = \frac{1}{path_length}$$

$$= \frac{1}{2}$$

$$= 0,5$$

Dalam pencarian kata ‘barang’ dan ‘produk’ masing-masing memiliki jumlah jalur yang sama pula, yaitu 2. Nilai kemiripan kata ‘barang’ dan ‘produk’:

$$\begin{aligned} sim(C1, C2) &= 2 \times Max(C1, C2) - SP \\ &= 2 \times 2 - 2 \\ &= 2 \end{aligned}$$

$$path = \frac{1}{path_length}$$

$$= \frac{1}{2}$$

$$= 0,5$$

Tabel 5. Nilai kemiripan kata pada kalimat 1 dan 2

Kalimat 2	Kalimat 1		
	melengkapi (v)	data (n)	barang (n)
melengkapi (v)	1	-	-
informasi (n)	-	0,5	
produk (n)	-		0,5

Tabel 5 hasil perhitungan kemiripan kata pada kalimat 1 dan kalimat 2 menggunakan *path* sesuai hasil pencarian percabangan pada proses sebelumnya.

b. Perhitungan kemiripan kata pada eksperimen 2:

Kalimat x: kami mengonsumsi makanan bergizi

Kalimat y: kami menggunakan barang bermanfaat

Pada pencarian kata ‘mengonsumsi’ dan ‘bermanfaat’, jalur maksimum bernilai 5 dan jalur terpendek bernilai 2.

$$\begin{aligned} sim(C1, C2) &= 2 \times Max(C1, C2) - SP \\ &= 2 \times 5 - 2 \\ &= 8 \end{aligned}$$

$$path = \frac{1}{path_length}$$

$$= \frac{1}{8}$$

$$= 0,125$$

Tabel 6. Nilai kemiripan kata pada kalimat x dan y

Kalimat y	Kalimat x		
	mengonsumsi	makanan	bergizi
menggunakan	0,5	-	-
barang	-	-	-
bermanfaat	0,125	-	-

Tabel 6 hasil perhitungan kemiripan kata pada kalimat x dan kalimat y menggunakan *path*.

3.3 Menghitung Kemiripan Kalimat

Setelah diperoleh nilai kemiripan kata pada ketiga contoh kalimat tersebut, kemudian dilakukan perhitungan kemiripan kalimatnya. Hasil perhitungan

kemiripan kalimat menggunakan dua jenis bobot, yaitu 0,25 untuk bobot kata benda, dan 0,75 untuk bobot kata kerja [9].

Hasil perhitungan kemiripan makna kalimat pada eksperimen 1:

kalimat 1: melengkapi data barang

kalimat 2: melengkapi informasi produk

$$\begin{aligned} \text{sentence sim} &= (N_{max} \times B_N) + (V_{max} \times B_V) \\ &= (0,5 \times 0,25) + (1 \times 0,75) \\ &= 0,875 \end{aligned}$$

Pada hasil eksperimen 1 perhitungan kemiripan makna kalimat 1 dan kalimat 2 diperoleh nilai 0,875 dengan kriteria sedang atau mirip.

Hasil perhitungan kemiripan makna kalimat pada eksperimen 2:

Kalimat x: kami mengonsumsi makanan bergizi

Kalimat y: kami menggunakan barang bermanfaat

$$\begin{aligned} \text{sentence sim} &= (N_{max} \times B_N) + (V_{max} \times B_V) \\ &= (0 \times 0,25) + (0,5 \times 0,75) \\ &= 0,375 \end{aligned}$$

Pada hasil eksperimen 2 kalimat aktif dan kalimat pasif, perhitungan kemiripan makna kalimat x dan y diperoleh nilai dibawah 0,5 dengan kriteria kemiripannya rendah.

Tabel 7. Hasil kemiripan kalimat keseluruhan

Eksperimen	Nilai kemiripan kalimat
1	0,875
2	0,375
3	0,250
4	0
5	0,125

Tabel 7 merupakan hasil nilai kemiripan kalimat yang telah dihitung. Dari kelima hasil pencarian kemiripan makna kalimat, bahwa pada eksperimen 1 dengan kriteria antara kata kerja (yang sama) dan kata benda, memiliki tingkat kemiripan yang tinggi dibandingkan dengan eksperimen yang lainnya. Sesuai dengan prinsip sinonimkata.com yang hanya dapat mendeteksi pencarian kata aktif atau kata kerja, dan kata dasar. Maka, pada hasil eksperimen lain dengan kriteria kalimat aktif yang dikombinasikan dengan kalimat pasif memiliki tingkat kemiripan yang rendah.

4. KESIMPULAN

Berdasarkan penelitian yang telah dilakukan dapat disimpulkan bahwa hasil pencarian dan pengukuran kemiripan makna kalimat dalam bahasa Indonesia menggunakan dataset sinonimkata.com dengan metode *Path*, didapatkan nilai 0,875 pada eksperimen susunan kata kerja dan kata benda. Sehingga, kriteria kalimat dengan susunan kata kerja dan kata benda berpengaruh tinggi dibandingkan eksperimen kriteria susunan lainnya.

DAFTAR PUSTAKA

- [1] A. W. Prasetya, M. A. Yaqin dan S. Zaman, "Common Process Extraction pada Scalable Model Proses Bisnis," *Konferensi Nasional Sistem Informasi (KNSI) 2018*, 2018.
- [2] A. Ali, F. Alfayez dan H. Alquhayz, "Semantic Similarity Measures Between Words: A Brief Survey," *Sci. Int (Labore)*, vol. 30, pp. 907-914, 2018.
- [3] S. Christina, "Kinerja Cosine Similarity dan Semantic Similarity Dalam Pengidentifikasian Relevansi Nomor Halaman Pada Daftar Indeks Istilah," *Sentika 2014*, no. 2089-9813, 2014.
- [4] P. Kharismadita dan F. Rahutomo, "Implementasi Tokenizing Plus pada Sistem Pendeteksi Kemiripan Jurnal Skripsi," *Jurnal Informatika Polinema*, vol. 2, no. 1, p. 24, 2017.
- [5] D. Guessoum, M. Miraoui dan C. Tadj, "A Modification of Wu and Palmer Semantic Similarity Measure," pp. 42-46, 2016.
- [6] L. D, "An information-theoretic definition of similarity," *Proceeding of International, Canada*, 1998.
- [7] T. Slimani, "Description and Evaluation of Semantic Similarity Measure Approaches," *International Journal of Computer Applications*, vol. 80, pp. 25-33, 2013.
- [8] H. Thamrin dan A. Sabardilla, "Evektivitas Algoritma Semantik dengan Keterkaitan Kata dalam Mengukur Kemiripan Teks Bahasa Indonesia," *Khazanah Informatika*, vol. 1, p. 1, 2015.
- [9] G. U. Abriani dan M. A. Yaqin, "Implementasi Metode Semantic Similarity untuk Pengukuran Kemiripan Makna antar Kalimat," *Ilkomnika*, vol. 1, p. 2, 2019.
- [10] B. Mcinnes dan T. Pedersen, "Evaluating Measure of Semantic Similarity and Relatedness to Disambiguate Terms in Biomedical Text," *International Journal of Biomedical Informatics*, vol. 46, 2013.