

Lexical Chain dan Word Sense Disambiguation Untuk Peringkasan Artikel Berbahasa Indonesia

Lexical Chain and Word Sense Disambiguation
to Summarize Article in Bahasa Indonesia

Dika Muhammad Fazar¹, Nelly Indriani Widiastuti²

^{1,2}Program Studi Teknik Informatika, Unikom, Bandung

e-mail: *¹dika.fazar@gmail.com, ²nelly.indriani@email.unikom.ac.id

Abstrak

Text Summarization adalah sebuah proses untuk menghasilkan ringkasan suatu dokumen dengan tidak menghilangkan informasi utama dari artikel. Ada beberapa metode untuk melakukan peringkasan, seperti metode rantai leksikal atau lexical chain yang memiliki kinerja yang baik untuk dokumen peringkasan dokumen tunggal. Meskipun demikian penggunaan metode lexical chain ini masih memiliki kelemahan yaitu tidak dapat mengidentifikasi kata yang berambigu dalam pembentukan lexical chain. Dalam penelitian ini, untuk memperbaiki kekurangan tersebut metode lexical chain dilengkapi dengan word sense disambiguation. Metode lexical chain dengan word sense disambiguation adalah metode pengembangan metode dari lexical chain. Dalam penelitian ini, setiap kata diperiksa ambiguitasnya. Pemeriksaan ini dilakukan untuk mendapatkan akurasi dalam menentukan lexical chain yang sesuai dengan konteks kalimat. Perbedaan mendasar dari metode lexical chain dengan word sense disambiguation dibandingkan dengan lexical chain, yaitu metode lexical chain dengan pengecekan word sense disambiguation sebelum menentukan rantai ambiguitas leksikal sedangkan metode asli lexical chain ambiguitas tidak diperiksa. Metode lexical chain dengan word sense disambiguation diterapkan pada peringkasan artikel kesehatan. Dari hasil penelitian ini disimpulkan bahwa metode lexical chain dengan word sense disambiguation dapat memperbaiki kelemahan dalam metode sebelumnya karena metode lexical chain dengan word sense disambiguation dapat mengidentifikasi ambiguitas dan menghasilkan akurasi ringkasan yang lebih baik.

Kata kunci— Peringkasan teks, lexical chain, word sense disambiguation

Abstract

Text Summarization is a process to produce a summary of an article, but still, have a picture or in other words, do not eliminate the main information from the article. There are several methods for doing summarization text, such as Lexical chain method that has good performance for summarization single text document, but the use of this Lexical chain method still has a disadvantage that it can not identify the disambiguation word in a formation of a lexical chain. Therefore, to rectify these shortcomings Lexical chain method comes with word sense disambiguation. Lexical chain methods with word sense disambiguation is a method of development of the original method of a lexical chain. In this method, each candidate word that is used to build a chain of the lexical ambiguity of meaning will be checked first. This check is done to obtain the accuracy in determining the appropriate lexical relations with the sentence context. The fundamental difference of method lexical chain with word sense disambiguation compared to the original lexical chain that if the method of lexical chain with word sense disambiguation checking before determining the chain of lexical ambiguity whereas original method lexical ambiguity chain is not checked. Lexical chain method with word sense disambiguation implemented against summarization health articles. From the result of this

study concluded that the method of lexical chain with word sense disambiguation can improve the weakness in previous method because the method of lexical chain with Word sense disambiguation can identify meaning that ambiguity and produce summaries better accuracy.

Keywords— *Text Summarization, lexical chain method, word Sense disambiguation*

1. PENDAHULUAN

Peringkasan dokumen teks secara umum adalah suatu proses untuk menghasilkan ringkasan dari suatu artikel, meskipun demikian hasil ringkasan tetap memiliki gambaran atau dengan kata lain tidak menghilangkan inti informasi dari artikel tersebut. Cara pembaca dalam mendapatkan informasi dari sebuah bacaan teks tidak selalu sama, sehingga dapat mengakibatkan kesalahpahaman terhadap informasi yang diperoleh dari hasil membaca. Hal ini terjadi karena beberapa faktor, diantaranya usia, cara berfikir setiap orang yang tidak sama, cara membaca yang tidak efektif karena adanya gangguan dari lingkungan, sehingga mengakibatkan pembaca kesulitan untuk mendapatkan inti dari bacaan teks tersebut[1].

Penelitian *automatic text summarization* telah dilakukan beberapa kali dengan menerapkan berbagai metode. Ada beberapa metode untuk melakukan peringkasan teks salah satunya dengan menggunakan metode *Lexical chain*[2]. Metode *Lexical chain* memiliki performansi yang baik untuk peringkasan teks *single document* karena adanya pembentukan *lexical chain* pada setiap kalimat sehingga bobot kalimat untuk diekstraksi ditentukan oleh rantai terkuat tidak tergantung pada judul. Meskipun demikian penggunaan metode *Lexical chain* ini masih memiliki kelemahan, yaitu tidak bisa mengidentifikasi kata yang berambigu dalam pembentukan *lexical chain*. Dimana *lexical chain* tersebut memiliki peran penting dalam proses selanjutnya yaitu menentukan bobot kalimat sebelum menghasilkan ringkasan[3]. Berdasarkan hal tersebut, jika terjadi ambiguitas dalam pembentukan *lexical chain* akan sangat berpengaruh kepada pembobotan kalimat dan hasil ringkasan yang dihasilkan.

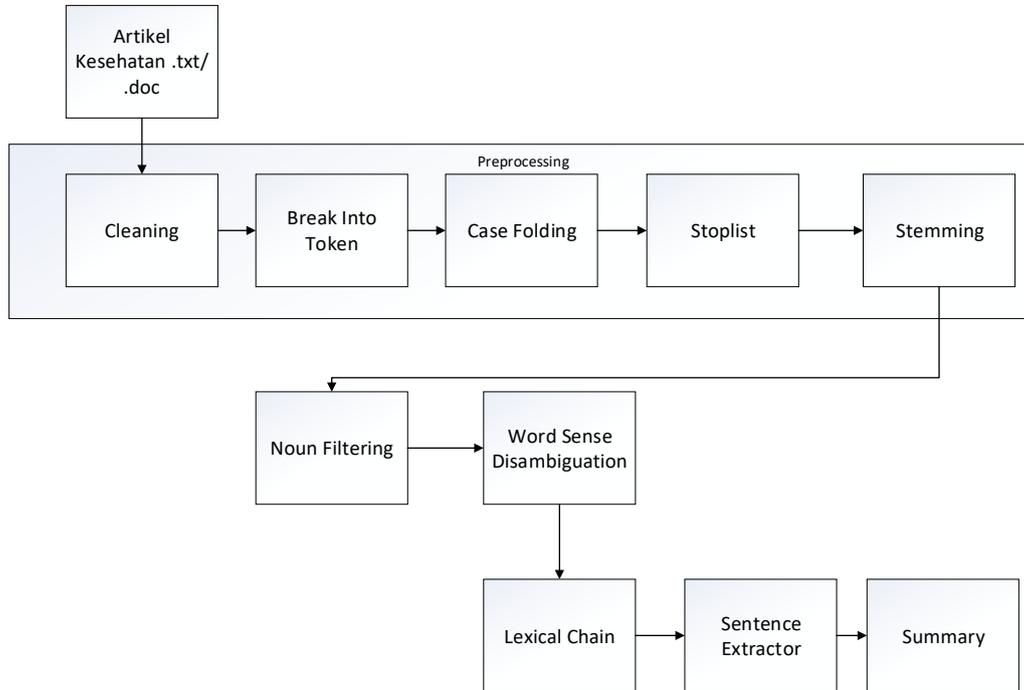
Berdasarkan uraian tersebut, maka dalam penelitian ini akan menghilangkan ambiguitas dengan menggunakan metode *lexical chain* dan *word sense disambiguation*. Metode dalam penelitian ini mengidentifikasi makna kata sebelum pembentukan *lexical chain* dan untuk mengetahui pengaruhnya terhadap hasil ringkasan. Metode *Lexical chain* adalah salah satu metode untuk meringkas teks yang membentuk *lexical chain* berdasarkan hubungan semantik antar kata dalam teks[2]. Sedangkan *Word sense disambiguation* adalah suatu proses mengidentifikasi makna kata yang digunakan dalam kalimat tertentu ketika satu kata memiliki sejumlah makna yang berbeda[4]. Algoritma *Word sense disambiguation* yang akan digunakan adalah algoritma *Lesk*. Algoritma *Lesk* adalah algoritma yang bekerja untuk mengetahui makna dari kata yang memiliki makna ganda atau lebih dari satu dalam kalimat dengan memperhatikan kata yang ada disekitarnya[5].

2. METODE PENELITIAN

Sistem dalam penelitian ini bekerja dengan cara memasukan artikel dengan format .txt atau .doc. Teks artikel tersebut akan diolah dengan berbagai proses yang akan menghasilkan ringkasan artikel. Setiap kandidat kata yang digunakan untuk membangun rantai leksikal akan dilakukan pengecekan ambiguitas makna terlebih dahulu. Pengecekan ini dilakukan untuk memperoleh keakuratan dalam menentukan hubungan leksikal yang sesuai dengan konteks kalimat[6][7].

Proses-proses yang terjadi dalam sistem ini yang memiliki 9 fungsi. Alur proses untuk sistem peringkasan artikel dapat dilihat pada Gambar 1. Dokumen yang digunakan dalam sistem ini adalah artikel kesehatan. Artikel yang masuk ke dalam sistem akan diproses tahap *preprocessing* terlebih dahulu sebelum masuk ke proses utama. Tahap *preprocessing* terdiri dari

proses *break into token*, *stoplist* dan *stemming*. Hasil dari *preprocessing* tersebut kemudian masuk ke proses utama dalam sistem yaitu *noun filtering*, *word sense disambiguation*, *lexical chain* dan terakhir adalah *sentence extractor* untuk menghasilkan *output* berupa ringkasan artikel(*summary*).



Gambar 1. Alur Proses

Deskripsi proses yang terjadi dalam tiap tahap adalah sebagai berikut:

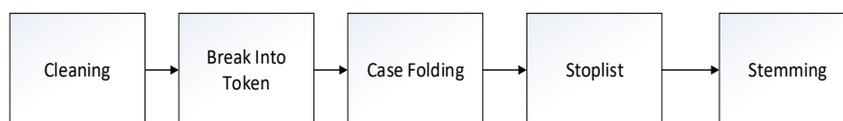
A. *Preprocessing*

Preprocessing adalah tahap dimana sebelum masuk ke proses utama dalam sistem. *Preprocessing* dilakukan terlebih dahulu agar menghasilkan informasi yang siap digunakan pada proses selanjutnya. Pada sistem ini artikel dipenggal menjadi kumpulan kalimat. Setiap kalimat dalam suatu artikel akan dilakukan *preprocessing* yaitu:

- 1) *Cleaning* : Tahap pembersihan karakter-karakter yang tidak diperlukan dalam proses selanjutnya. Dalam penelitian ini karakter yang dihilangkan adalah karakter seperti angka, tanda tanya, tanda seru dan lain-lain.
- 2) *Break into token* : untuk memudahkan pengidentifikasian kata yang ambigu, maka setiap kalimat yang menyusun artikel tersebut dipecah sehingga menjadi kumpulan kata.
- 3) *Case folding* : Pada tahap *case folding* ini setiap token yang memiliki huruf kapital akan diubah menjadi lowercase atau huruf kecil.
- 4) *Stoplist* : dalam proses ini sejumlah kata dihilangkan karena dianggap tidak memiliki makna penting dalam dokumen seperti "adalah", "yang", "dari", dan sebagainya. Hasil pembuangan *stoplist* disebut dengan *non-stopword tokens*. Hal ini diperlukan untuk mengurangi jumlah kata yang harus diproses.

- 5) Stemming : mengubah token atau kata yang berimbuhan diubah menjadi kata dasar pembentuknya. Misalnya, kata "kelangsungan" diubah menjadi kata "langsung", kata "merawat" diubah menjadi kata "rawat". Dalam penelitian ini metode stemming yang digunakan adalah metode Nazif dan Adriani[8].

Secara umum tahapan *preprocessing* dapat dilihat pada Gambar 2 berikut ini:



Gambar 2. *Preprocessing*

Setiap tahap dalam preprocessing ini merupakan tahap yang berkelanjutan. Artikel yang akan diringas, akan melalui proses ini setahap demi setahap sebelum memasuki proses selanjutnya.

B. *Noun filtering*

Dalam penelitian ini kata yang berambigu dibatasi hanya kata benda. Sebelum memasuki pendeteksian kata yang ambigu, dalam tahap ini dilakukan penyaringan kata benda (*noun*) dari kumpulan token hasil *preprocessing* yang kemudian dianggap sebagai input dalam membentuk lexical chain. Penyaringan kata benda dilakukan dengan cara pengecekan setiap kata dengan database kata benda sebagai acuan untuk mendapatkan kata benda.

C. *Word sense disambiguation*

Pada tahap ini dilakukan identifikasi makna kata dengan menggunakan algoritma *lesk*. Algoritma *lesk* bekerja dengan membandingkan definisi dari kata yang berambigu dengan definisi dari kata tetangganya berdasarkan definisi kamus. Pada tabel 1. adalah contoh hasil kata-kata hasil *noun filtering* yang memiliki makna ganda.

Tabel 1. Hasil Word Sense Disambiguation

Kata bermakna ganda	hati
Makna dari kamus	<ol style="list-style-type: none"> 1. organ tubuh yang berwarna kemerah-merahan di bagian kanan atas rongga perut gunanya untuk mengambil sari-sari makanan di dalam darah dan menghasilkan empedu 2. tempat segala perasaan batin dan tempat menyimpan pengertian

Setelah proses pengenalan kata berambigu, kata tersebut akan ditentukan bobotnya. Nilai didapat dari banyaknya kata pada makna kata dibandingkan dengan kata dalam makna kata pembanding. Setiap kata yang sama memiliki nilai 1. Sebagai contoh pada tabel 2. berikut akan membandingkan kata hati dengan organ.

Tabel 2. Contoh menentukan bobot makna kata

Kata Berambigu	Makna kata	Kata pembanding	Makna pembanding	bobot
Hati	[1]organ tubuh yang berwarna kemerah-merahan	organ	alat yang mempunyai tugas tertentu di dalam tubuh	5

Kata Berambigu	Makna kata	Kata pembandingan	Makna pembandingan	bobot	
di bagian kanan atas rongga perut gunanya untuk mengambil sari-sari makanan di dalam darah dan menghasilkan empedu			manusia		
		vital	organ tubuh yang sangat penting bagi kehidupan	3	
		tubuh	keseluruhan struktur organisme manusia atau binatang yang kelihatan dr bagian ujung kaki sampai ujung rambut	2	
		manusia	makhluk yang berakal budi mampu menguasai atau makhluk insan orang	1	
		peran	pemain sandiwara	0	
		hidup	masih terus ada bergerak dan bekerja sebagaimana mestinya seperti manusia binatang atau tumbuhan	1	
		orang	makhluk hidup yang dapat mengembangkan kemampuan bertindak independen atau manusia	1	
	[2] tempat segala perasaan batin dan tempat menyimpan pengertian		organ	alat yang mempunyai tugas tertentu di dalam tubuh manusia	0
			vital	organ tubuh yang sangat penting bagi kehidupan	0
			tubuh	keseluruhan struktur organisme manusia atau binatang yang kelihatan dr bagian ujung kaki sampai ujung rambut	0
			manusia	makhluk yang berakal budi mampu menguasai atau makhluk insan orang	0
			peran	pemain sandiwara	0
			hidup	masih terus ada bergerak dan bekerja sebagaimana mestinya seperti manusia binatang atau tumbuhan	1
		orang	makhluk hidup yang dapat mengembangkan kemampuan bertindak independen atau manusia	0	

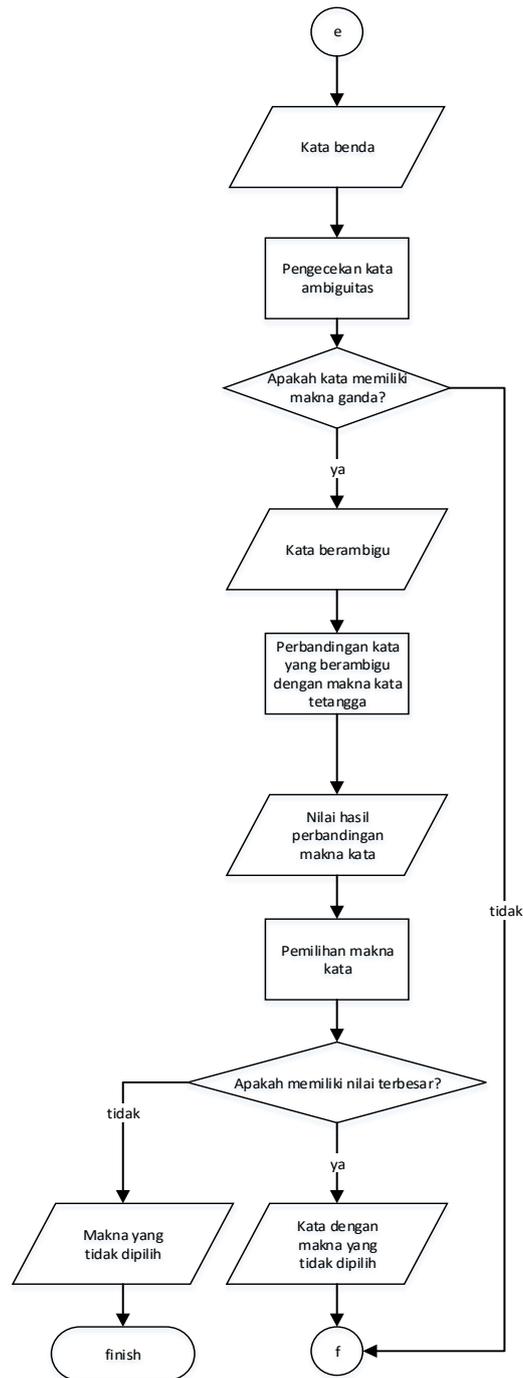
Bobot dihitung juga dengan kata berikutnya dalam satu kalimat. Hasil pembobotan makna kata yang bertetangga, dihitung berdasarkan kesamaan makna dengan kata yang ambigu dapat dilihat pada tabel 3. Kata yang terpilih adalah kata yang memiliki bobot terbesar.

Tabel 3. Hasil Pembobotan

Kata Berambigu	Makna kata	Total Bobot
Hati	[1] organ tubuh yang berwarna kemerah-merahan di	13

	bagian kanan atas rongga perut gunanya untuk mengambil sari-sari makanan di dalam darah dan menghasilkan empedu	
	[2] tempat segala perasaan batin dan tempat menyimpan pengertian	1

Secara lengkap proses *word sense disambiguation* dapat dilihat pada Gambar 4.



Gambar 4. Word Sense Disambiguation

E. Lexical chain

Pada tahap ini dilakukan pembentukan rantai yang menggambarkan hubungan semantik dari suatu kata dengan kata yang lain dan pemberian nilai. Dengan menggunakan rumus (1) diperoleh hasil perhitungan homogeneity indeks dan score chain bar[2], dapat dilihat pada tabel 4.

$$Score(chain) = length * homogeneityindex \tag{1}$$

$$\text{Homogeneityindex} = 1 - (\text{distinctmember}/\text{length})$$

Keterangan:

- Length*: merupakan jumlah total dari seluruh kemunculan seluruh anggota penyusun rantai dalam teks.
- Distinctmember*: banyaknya anggota yang berbeda yang menyusun rantai.
- Homogeneityindex*: nilai kemiripan dari anggota-anggota penyusun rantai.

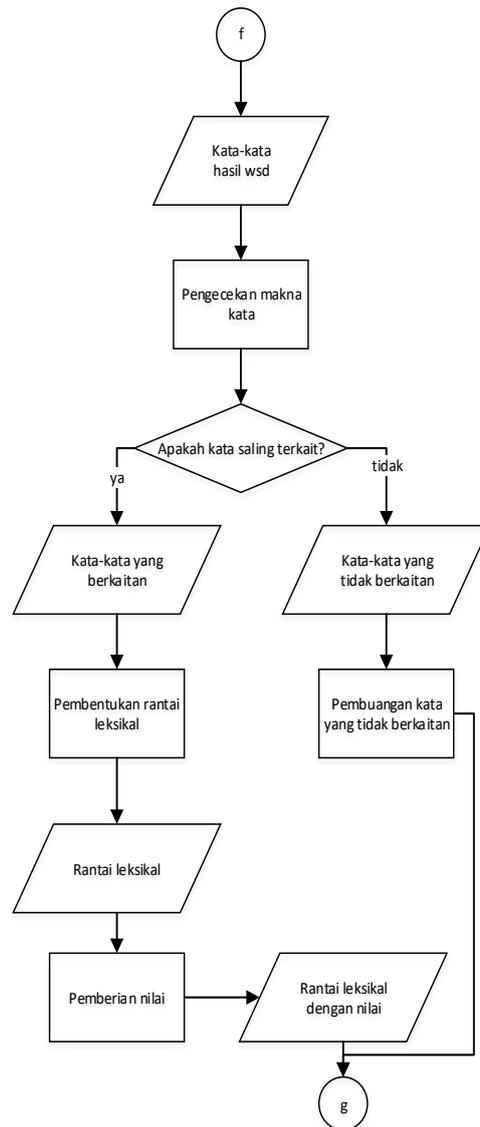
Perhitungan *Homogeneityindex* untuk digunakan untuk menunjukkan tingkat kemiripan anggota-anggota yang membentuk rantai. *Distinctmember* dibagi dengan *length* menyatakan tingkat perbedaan anggota penyusun rantai.

Pembentukan lexical chain didapat dari kandidat kata yang telah terpilih dari proses *noun filtering* serta telah didisambiguasi pada proses word sense disambiguation.

Tabel 4. Lexical Chain

No.	Lexical chain	Pemberian Nilai	Score
1	hati organ vital tubuh manusia kinerja terus hidup orang	homo=1-(14/23)=0.39 score= 22*0.39=8.97	8.97
2	organ kelenjar tubuh ikut bawah cara hati manusia	Homo=1-(18/20)=0.10 Score=20*0.10	2
3	vital hati	Homo=1-(7/7)= 0 Score=7*0=0	0
4	atur metabolisme saring racun masuk tubuh bentuk antibodi hati	Homo=1-(10/16)=0.38 Score=16*0.38=6.08	6.08
5	pola hidup jadi organ tubuh jaga hati	Homo=1-(11/19)=0.42 Score=19*0.42	7.98

Pembentukan lexical chain didapat dari kandidat kata yang telah terpilih dari proses *noun filtering* serta telah didisambiguasi pada proses word sense disambiguation. Proses *lexical chain* dapat dilihat pada Gambar 5.

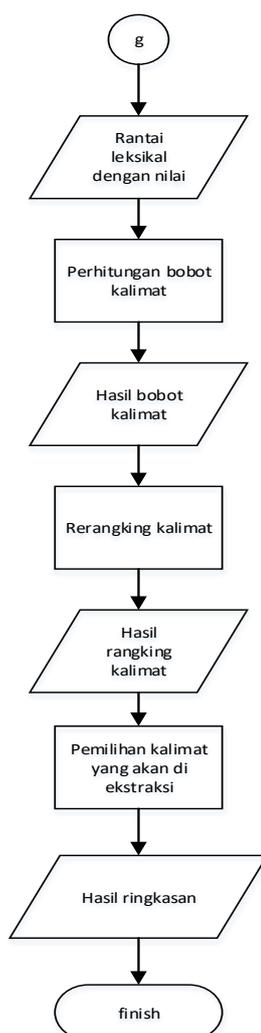


Gambar 5. Lexical Chain

F. Sentence Extractor

Dengan menggunakan hasil pada tahap lexical chain atau pada contoh tabel 4, setiap kalimat akan dihitung bobotnya menggunakan rumus (2). Setelah itu, kemudian dilakukan pemeringkatan secara *descending*. Hasil pemeringkatan diekstraksi pada proses *sentence extraction*. Tujuan dari *sentence extraction* adalah untuk menemukan kalimat yang akan di ekstrak sebagai penyusun hasil ringkasan. Proses *sentence extractor* dapat dilihat pada gambar 6.

$$score(sen_j) = \sum_{i=1}^m \left(\frac{score(ChainMember_i, Sen_j)}{S_i} \right) \quad (2)$$



Gambar 6. Sentence Extractor

Hasil perangkaan dalam tahap ini akan menentukan urutan kalimat dalam hasil peringkasan. Berikut adalah contoh hasil penghitungan sentence extractor pada tabel 5. Dalam tahap ini kalimat dipulihkan kembali sehingga siap ditampilkan sebagai ringkasan sebuah artikel.

Tabel 5. Hasil Sentence Extractor

Peringkat	Kalimat	Nilai sentence scoring	Kalimat ke
1	Hati merupakan organ vital dari tubuh manusia dengan peran yang sangat penting untuk melengkapi kinerja organ lain dalam meneruskan kelangsungan hidup setiap orang.	1.79	1
2	Pola hidup yang tidak sehat selalu menjadi momok bagi kesehatan organ tubuh, khususnya menjaga kesehatan hati.	1.60	5
3	Mengatur metabolisme, menyaring racun yang masuk ke tubuh, hingga membentuk antibodi merupakan sebagian dari fungsi hati.	1.22	4

4	Sebagai organ kelenjar terbesar dalam tubuh membuat kita mau tidak mau harus merawatnya dengan baik, berikut dibawah ini cara merawat kesehatan organ hati pada manusia.	0.40	2
5	Vitalnya fungsi hati mengharuskan Anda untuk senantiasa merawatnya.	0	3

3. HASIL DAN PEMBAHASAN

Pengujian akurasi dibagi menjadi dua yaitu pengujian akurasi dalam mengidentifikasi makna kata yang berambigu dan pengujian akurasi hasil peringkasan sistem yang menggunakan metode lexical chain dengan word sense disambiguation.

A. Pengujian Akurasi Identifikasi Makna Kata Berambigu

Pengujian akurasi dalam mengidentifikasi makna kata yang berambigu dilakukan terhadap 60 kata yang diambil dari artikel-artikel yang akan dilakukan peringkasan. Pengujian didapat setelah membandingkan hasil sistem dengan kamus yang menjadi acuan. Adapun tujuan pengujian adalah untuk dapat mengetahui akurasi dalam mengidentifikasi makna kata yang memiliki makna ganda atau ambigu. Hasil pengujian menunjukkan sebanyak 46 kata terdeteksi ambigu dari total 60 kata yang berambigu, sehingga akurasi pengujian dapat dilihat sebagai berikut:

$$\text{Akurasi} = \frac{46}{60} \times 100\% = 77\%$$

Berdasarkan hal tersebut maka akurasi yang didapat dalam mengidentifikasi makna kata yang berambigu terhadap 60 kata yang berambigu didapat akurasi sebesar 77%.

B. Akurasi Hasil peringkasan

Pengujian peringkasan artikel yang dilakukan dengan menggunakan 30 artikel. Berikut hasil pengujian ringkasan oleh manusia, dan sistem dengan mempertimbangkan ambiguitas dan tanpa mempertimbangkan ambiguitas pada tabel 7.

Tabel 7. Hasil Pengujian Ringkasan

Artikel Ke-	Pengujian	Dengan WSD	Tanpa WSD	Artikel Ke-	Pengujian	Dengan WSD	Tanpa WSD
1	<i>precision</i>	85%	77%	16	<i>precision</i>	75%	75%
	<i>recall</i>	44%	40%		<i>recall</i>	40%	40%
2	<i>precision</i>	67%	56%	17	<i>precision</i>	73%	64%
	<i>recall</i>	35%	29%		<i>recall</i>	36%	32%
3	<i>precision</i>	58%	33%	18	<i>precision</i>	62%	62%
	<i>recall</i>	29%	17%		<i>recall</i>	31%	31%
4	<i>precision</i>	59%	53%	19	<i>precision</i>	82%	73%
	<i>recall</i>	29%	26%		<i>recall</i>	41%	36%
5	<i>precision</i>	62%	62%	20	<i>precision</i>	55%	45%
	<i>recall</i>	32%	32%		<i>recall</i>	29%	24%
6	<i>precision</i>	57%	71%	21	<i>precision</i>	73%	64%
	<i>recall</i>	29%	36%		<i>recall</i>	38%	35%
7	<i>precision</i>	67%	60%	22	<i>precision</i>	70%	40%
	<i>recall</i>	33%	30%		<i>recall</i>	35%	20%
8	<i>precision</i>	58%	42%	23	<i>precision</i>	45%	55%
	<i>recall</i>	29%	21%		<i>recall</i>	24%	29%

Artikel Ke-	Pengujian	Dengan WSD	Tanpa WSD	Artikel Ke-	Pengujian	Dengan WSD	Tanpa WSD
9	<i>precision</i>	73%	55%	24	<i>precision</i>	89%	67%
	<i>recall</i>	36%	27%		<i>recall</i>	44%	33%
10	<i>precision</i>	62%	38%	25	<i>precision</i>	50%	40%
	<i>recall</i>	32%	20%		<i>recall</i>	26%	21%
11	<i>precision</i>	60%	50%	26	<i>precision</i>	70%	80%
	<i>recall</i>	30%	25%		<i>recall</i>	37%	42%
12	<i>precision</i>	73%	64%	27	<i>precision</i>	67%	56%
	<i>recall</i>	38%	33%		<i>recall</i>	35%	29%
13	<i>precision</i>	58%	50%	28	<i>precision</i>	67%	60%
	<i>recall</i>	33%	26%		<i>recall</i>	33%	30%
14	<i>precision</i>	60%	70%	29	<i>precision</i>	64%	64%
	<i>recall</i>	32%	37%		<i>recall</i>	32%	32%
15	<i>precision</i>	67%	44%	30	<i>precision</i>	67%	56%
	<i>recall</i>	33%	22%		<i>recall</i>	33%	28%

Pada pengujian menunjukkan bahwa nilai persentase *precision* dan *recall* terbesar didapat oleh metode *lexical chain* dengan *word sense disambiguation* masing-masing sebesar 89% dan 44% terjadi pada pengujian ke-24. Metode *lexical chain* dengan *word sense disambiguation* memiliki rata-rata *precision* sebesar 66% dan *recall* sebesar 34% sedangkan untuk metode *lexical chain* tanpa *word sense disambiguation* memiliki rata-rata *precision* sebesar 57% dan *recall* sebesar 29%.

4. KESIMPULAN

Berdasarkan hasil penelitian yang telah dilakukan dapat disimpulkan bahwa Metode *Lexical Chain* dengan *Word Sense Disambiguation* dapat mengidentifikasi kata yang memiliki ambiguitas atau makna ganda, dan dapat memilih makna dari kata yang memiliki ambiguitas dengan baik. Selain itu metode *lexical chain* dengan *word sense disambiguation* memiliki akurasi yang lebih baik daripada metode *lexical chain* dalam peringkasan artikel.

5. SARAN

Penelitian dalam *text summarization* atau peringkasan dokumen masih memiliki kendala dalam sumber daya kamus dan wordnet bahasa Indonesia yang masih terbatas. Penelitian dalam bidang bahasa akan jauh lebih baik jika sumber daya tersebut telah tersedia.

DAFTAR PUSTAKA

- [1] L. Suanmali, N. Salim, and M. S. Binwahlan, "Fuzzy Logic Based Method for Improving Text Summarization," *J. Comput. Sci.*, vol. 2, no. 1, p. 6, 2009.
- [2] R. Barzilay and M. Elhadad, "Using lexical chains for text summarization," *Proc. ACL Work. Intell. Scalable Text Summ.*, vol. 17, no. 48, pp. 10–17, 1997.
- [3] M. Augat and M. Ladlow, "CS65 : An NLTK Package for Lexical-Chain Based Word Sense Disambiguation," *Word J. Int. Linguist. Assoc.*, 2004.
- [4] J. Ekedahl, "Word sense disambiguation using WordNet and the Lesk algorithm," *Computing*, 2004.
- [5] S. Banerjee, "Adapting the Lesk Algorithm for Word Sense Disambiguation to

- WordNet,” Duluth, Minnesota, 2002.
- [6] Y. F. A. w. R. Yuni Susanti, M. A. Bijaksana, “ANALISIS PERBANDINGAN PERINGKASAN TEKS MENGGUNAKAN METODE LEXICAL CHAIN DAN METODE TF-IDF DENGAN EKSTRAKSI FRASE UTAMA,” Bandung, 2009.
- [7] Ș. D. Dumitrescu, A. Găinaru, and Ș. Trăușan-Matu, “A study on lexical chain identification and word sense disambiguation,” *UPB Sci. Bull. Ser. C Electr. Eng.*, vol. 73, no. 4, pp. 197–212, 2011.
- [8] D. Keke and D. P. Ryan, Chikita, Agus, “Sistem temu balik informasi,” Yogyakarta, 2012.