

Evaluasi Sentence Extraction pada Peringkasan Dokumen Otomatis

Kania Evita Dewi*, Nelly Indriani Widiastuti, Ednawati Rainarli

Program Studi Teknik Informatika
Fakultas Teknik dan Ilmu Komputer
Universitas Komputer
Jl. Dipati Ukur 112-118, Bandung
kania.evita.dewi@email.unikom.ac.id

Abstrak—Salah satu proses yang dilakukan dalam peringkasan dokumen adalah proses ekstraksi ciri. Berdasarkan beberapa penelitian, metode Tf-Idf adalah metode yang sering digunakan dalam proses ekstraksi ciri karena menghasilkan akurasi yang baik. Akan tetapi, ada metode lain yaitu metode dengan skema LGN yang disampaikan oleh Erica dan diklaim lebih baik dari pada Tf-Idf. Oleh karena itu dalam penelitian ini dilakukan pengujian secara statistik terhadap penggunaan metode *term weighting* Tf-Idf dan LGN. Hasil pengujian menunjukkan bahwa tidak ada perbedaan yang signifikan antara akurasi hasil peringkasan menggunakan metode Tf-Idf dengan LGN. Berdasarkan hal ini maka dalam peringkasan teks otomatis metode Tf-Idf ataupun LGN dapat digunakan sebagai metode ekstraksi ciri.

Kata kunci—*term weighting*; peringkasan; *tf-idf*; uji hipotesis.

I. PENDAHULUAN

Menurut kamus Oxford English Dictionary¹, meringkas otomatis adalah membuat versi singkat dari suatu dokumen dengan menggunakan komputer. Hasil dari proses ini adalah poin-poin penting dari teks asli. Karen Spark Jones menyampaikan bahwa meringkas dokumen teks adalah proses kompresi yang melibatkan hilangnya informasi. Proses peringkasan dilakukan secara otomatis dengan algoritma atau sistem komputer[1].

Untuk memilih kalimat yang akan dijadikan ringkasan, dapat dilakukan dengan cara *sentence extraction* atau dengan cara mengekstraksi terlebih dahulu kalimat dari teks yang akan diringkaskan. *Sentence extraction* atau ekstraksi kalimat mulai dilakukan oleh Luhn tahun 1958 yang menyatakan bahwa pentingnya suatu kalimat didasarkan pada frekuensi kata [2]. Metode ekstraksi kalimat yang tepat akan memberikan pengaruh yang kuat terhadap kinerja sistem. Sampai saat ini, hal tersebut masih banyak dilakukan pada penelitian-penelitian lain, seperti pada penelitian dalam sistem *information retrieval*, *document clustering*, *summarization*, *text mining*, dan pemrosesan bahasa alami lainnya [3] [4][5].

Pendekatan ekstraksi kalimat yang paling populer hingga saat ini adalah metode Tf-Idf[6]. Ide dasar dari Tf-Idf adalah mengurangi bobot kata-kata yang sering terjadi dengan cara

membandingkan frekuensi proporsionalnya dalam kumpulan dokumen. Properti ini telah membuat Tf-Idf menjadi salah satu istilah yang digunakan secara universal dalam peringkasan ekstraktif [7][8][9][10][11]. Metode ekstraksi kalimat ini termasuk ke dalam salah satu skema LGN. Erica Christholm dan Tamara Kolda melakukan penelitian yang membandingkan beberapa kombinasi LGN dalam sistem temu kembali. Hasil penelitian ini menunjukkan bahwa kombinasi LGN menghasilkan akurasi yang lebih baik dari pada Tf-Idf [12]. Hasil yang berbeda ditunjukkan pada beberapa penelitian lain. Dalam penelitian Man Lan dkk. Tf-Idf digunakan untuk kasus pengelompokan teks. Dari hasil penelitian tersebut diperoleh fakta Tf-Idf tidak secara konsisten menunjukkan hasil yang baik[13][14].

Oleh karena itu, dalam penelitian ini akan dievaluasi kinerja Tf-Idf dengan skema LGN lain secara statistik. Penelitian ini akan membuktikan secara signifikan bahwa penggunaan Tf-Idf akan menghasilkan kualitas ringkasan yang lebih baik daripada LGN [15]. Pengujian dilakukan dengan cara mengukur kualitas ringkasan yang dihasilkan. Adapun, kualitas ringkasan akan diuji dari nilai akurasi.

II. METODE

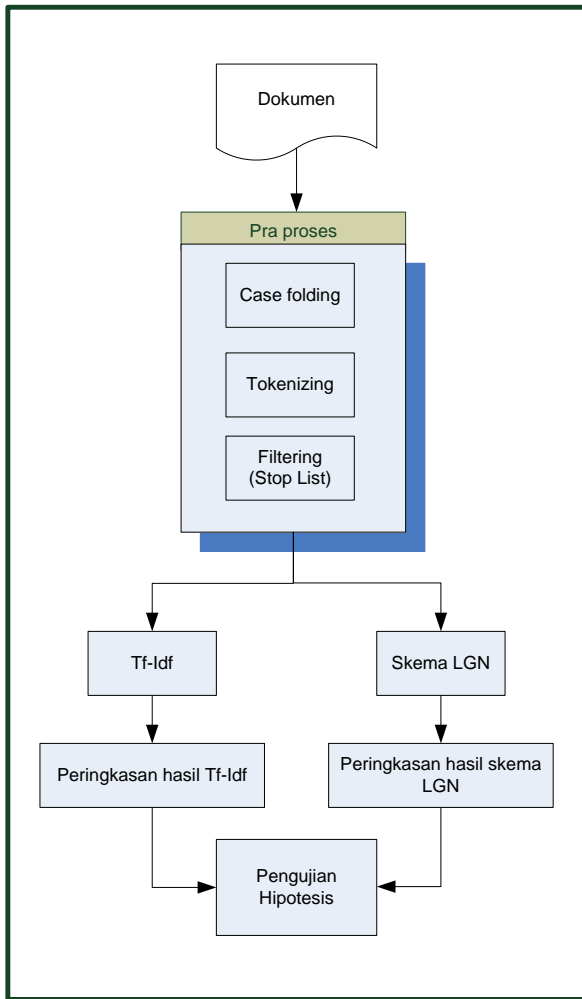
Untuk memberikan gambaran utuh tentang pengujian hipotesis yang dilakukan, berikut pada Gambar 1 adalah gambaran metode penelitian yang dilakukan. Secara garis besar ada empat tahapan yang dilakukan yaitu: pra proses, ekstraksi fitur, peringkasan otomatis dan pengujian secara statistik.

Tahapan pra proses yang digunakan adalah *tokenizing*, *case folding* dan *filtering*. Data yang telah melalui tahapan pra proses akan diekstraksi dengan menggunakan Tf-Idf dan LGN. Setelah dilakukan ekstraksi fitur, diperoleh nilai bobot tiap kalimat. Nilai bobot yang diperoleh lalu diranking dari yang terbesar hingga terkecil. Kalimat yang akan menjadi kandidat ringkasan adalah kalimat yang memiliki nilai bobot terbesar. Dalam penelitian ini digunakan kompresi ringkasan sebesar 50%. Hasil ringkasan dikomparasi dengan hasil ringkasan yang dilakukan oleh ahli bahasa untuk kemudian dihitung nilai akurasinya. Nilai akurasi dari masing-masing penggunaan metode Tf-Idf dan LGN inilah yang akan diuji dengan menggunakan uji statistik.

¹ <http://www.oed.com>.

Dokumen yang digunakan adalah bagian latar belakang skripsi mahasiswa teknik informatika. Dokumen tersebut sudah dalam format txt. Setiap dokumen akan melalui tahapan dari pra proses. Dalam penelitian ini, selain *stop word*, angka dan tanda baca akan diabaikan.

Setelah melalui tahap pra proses, token atau *term* digunakan untuk memperoleh nilai *Term Weighting* baik Tf-Idf maupun skema LGN lainnya. Dalam hal ini skema LGN yang akan digunakan sebagai pembanding adalah skema SQRT - IGFS – COSN. Skema ini digunakan karena dianggap paling baik dalam penelitian [12]. Metode penelitian dapat dilihat pada Gambar 1.



Gambar 1. Metode Penelitian

A. Term Frequency-Inverse document frequency

Term frequency-Inverse document frequency atau lebih dikenal sebagai Tf-Idf adalah ekstraksi kalimat dengan cara memberikan nilai atau bobot pada kata. Pembobotan kata yang dimaksudkan bertujuan untuk mencerminkan betapa pentingnya sebuah kata dalam sebuah kumpulan dokumen atau korpus[16]. Metode ini sering digunakan sebagai faktor pembobotan dalam *information retrieval*, *text mining*, dan *user modeling*. Nilai Tf-Idf meningkat secara proporsional berdasarkan berapa kali sebuah kata muncul dalam dokumen.

Hal tersebut membantu menyesuaikan fakta bahwa secara umum beberapa kata lebih sering muncul. Namun hal ini sering diimbangi oleh frekuensi kata dalam *corpus*. Saat ini, Tf-Idf adalah salah satu skema pembobotan yang paling populer. Menurut literatur survei yang dilakukan Joran Beel dkk. 83% sistem rekomendasi berbasis teks dalam domain perpustakaan digital menggunakan Tf-Idf[17].

Persamaan Tf-Idf yang digunakan dapat dilihat pada *persamaan* (1) dan (2) sebagai berikut:

$$Tf = f_{t,d} \quad (1)$$

yang mana *Tf* adalah variabel *Term frequency*, $f_{t,d}$ adalah nilai frekuensi kemunculan kata *t* dalam dokumen *d*. Perhitungan *Inverse Document Frequency* (*IDF*) ditunjukkan pada *persamaan* (2)

$$IDF = \log\left(\frac{N}{n_i}\right) \quad (2)$$

dengan variabel *N* adalah jumlah dokumen yang digunakan dan n_i adalah jumlah dokumen di mana kata *i* muncul. Berdasarkan kedua *persamaan* tersebut maka Tf-Idf dinyatakan sebagai $Tf \times Idf$ [18]. Gambar 2. adalah proses perhitungan Tf-Idf dalam sistem peringkasan otomatis. Keluaran dari hasil Tf-Idf dapat dianggap sebagai bobot dari setiap kata dalam setiap kalimat.

Setiap token kalimat atau berupa kata dalam suatu dokumen akan ditentukan jumlah setiap term dalam setiap kalimat. Contohnya, jumlah kalimat dalam suatu latar belakang skripsi adalah 8. Jumlah kata “algoritma” yang muncul dalam kalimat merupakan Tf secara berturut-turut adalah {2,0,1,1,2,2,1,0}. Nilai *N* = 8, nilai n_i = 6. Berdasarkan hal tersebut, maka nilai global (*Idf*) = $\log(8/6) = 0,124939$ dan setiap kalimat memiliki nilai Tf-Idf berturut-turut {0,249877473; 0; 0,124938737; 0,124938737; 0,249877473; 0,249877473; 0; 0,124938737}.

B. Local Global Normalization

Local, Global and Normalization atau LGN adalah metode pembobotan yang mempertimbangkan faktor lokal, faktor global dan faktor normalisasi untuk dokumen[12]. Faktor lokal untuk menyatakan berapa banyak tiap kata muncul dalam sebuah dokumen, faktor global adalah fungsi untuk menghitung berapa kali tiap kata muncul dalam seluruh dokumen, dan faktor normalisasi adalah kompensasi terhadap panjang dokumen yang berbeda-beda [12].

Setiap elemen dalam LGN baik lokal, global maupun normalisasi memiliki beberapa skema. Dalam penelitian ini, untuk memperoleh faktor/bobot lokal digunakan *persamaan* (3), bobot global yang digunakan adalah *persamaan* (4) dan bobot normalisasi digunakan *persamaan* (5). Ketiga *persamaan* tersebut menggunakan skema SQRT -IGFS - COSN[12].

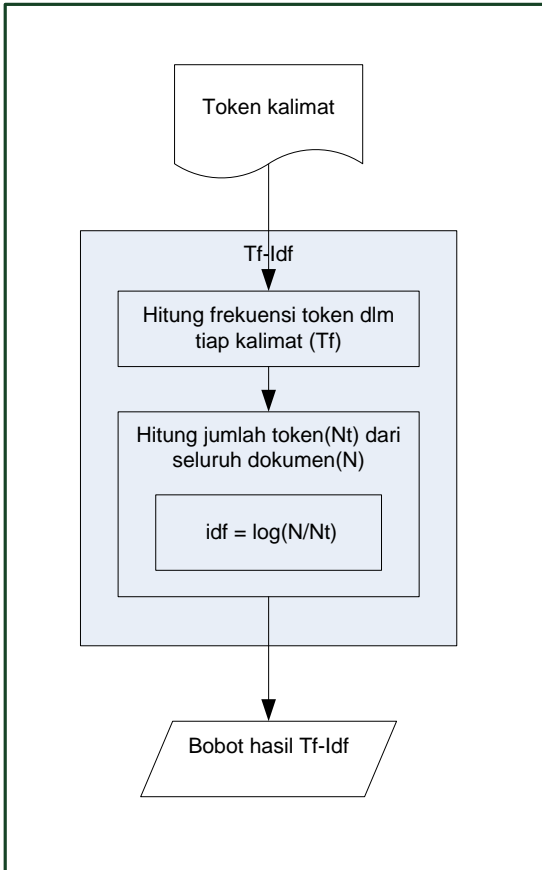
$$L_i = \begin{cases} \sqrt{F_{ij} - 0,5} + 1, & F_{ij} > 0 \\ 0, & F_{ij} = 0 \end{cases} \quad (3)$$

dimana L_i adalah nilai lokal dari kata ke-*i*. F_{ij} adalah frekuensi kata ke-*i* dalam dokumen *j*.

$$G_i = \sqrt{\frac{F_i}{n_i} - 0,9} \quad (4)$$

G_i adalah nilai global dari kata ke- i , F_i menyatakan frekuensi kata ke- i dalam seluruh dokumen. Bobot normalisasi menggunakan persamaan (5) dimana N_i adalah nilai normalisasi dari kata ke- i

$$N_i = \frac{1}{\sqrt{\sum(G_i L_{ij})^2}} \quad (5)$$



Gambar 2. Alur perhitungan Tf-Idf

Bobot normalisasi adalah hasil bagi satu dengan magnitude bobot vektor dokumen. Dengan cara ini maka memaksa besarnya bobot vektor dokumen menjadi satu. Hal ini memungkinkan untuk membandingkan sudut antara bobot vektor[12].

C. Uji Hipotesis

Uji hipotesis atau disebut juga konfirmasi analisis data adalah cara untuk menentukan apakah suatu hasil percobaan baik yang terkontrol maupun hasil yang tidak terkontrol (observasi) signifikan secara statistik. Dalam statistik hasil yang signifikan jika kejadian (hasil) tersebut dikatakan tidak mungkin disebabkan faktor kebetulan sesuai dengan batas probabilitas yang sudah ditentukan sebelumnya [15].

Uji hipotesis yang digunakan adalah pengujian hipotesis satu pihak dari perbedaan dua rata-rata. Sebelum pengujian hipotesis, dilakukan pengujian normalitas dan homogenitas

dari nilai akurasi metode Tf-IDF dan LGN. Untuk pengujian normalitas digunakan uji Kolmogorov-Smirnov, sedangkan pengujian homogenitas menggunakan uji Levene. Setelah pengujian tersebut selesai, dilakukan uji perbedaan antara rata-rata akurasi yang dihasilkan Tf-IDF dengan LGN. Hipotesis yang akan diuji adalah sebagai berikut:

$H_0 : \mu_1 = \mu_2$ artinya tidak terdapat perbedaan antara hasil peringkasan menggunakan Tf-idf dengan LGN.

$H_1 : \mu_1 > \mu_2$ artinya hasil peringkasan menggunakan tf-idf lebih baik daripada hasil peringkasan LGN.

Dalam uji perbedaan rata-rata terdapat tiga kemungkinan yang terjadi. Tabel 1. menunjukkan kemungkinan pengujian hasil uji perbedaan rata-rata. Jika data berdistribusi normal dan homogen maka digunakan uji-t. Jika data berdistribusi normal dan tidak homogen maka digunakan uji-t' dan jika data tidak berdistribusi normal maka digunakan uji non parametrik yaitu Mann-Whitney U. Seluruh perhitungan pengujian tersebut dilakukan dengan menggunakan program komputer SPSS versi 12.0.

TABEL 1. UJI HIPOTESIS

Pengujian	Kemungkinan hasil		
	1	2	3
Normalitas	√	√	x
Homogenitas	√	x	x
Uji Hipotesis	t	t'	Mann-Whitney U

Untuk pengujian normalitas digunakan uji Kolmogorov Smirnov satu sampel. Adapun pengujian Kolmogorov-Smirnov dapat dilihat pada persamaan (6) sebagai berikut:

H_0 : Data sampel mengikuti distribusi populasi yang diuji

H_1 : Data sampel tidak mengikuti distribusi populasi yang diuji

Statistik ujinya sebagai berikut:

$$D = \max|F_0(X) - S_N(X)| \quad (6)$$

dimana $F_0(X)$ adalah fungsi distribusi frekuensi kumulatif yang akan diuji (pada kasus ini adalah fungsi distribusi kumulatif normal). $S_N(X)$ adalah fungsi distribusi frekuensi kumulatif yang diobservasi dari suatu sampel random dengan N observasi. X adalah sembarang skor yang mungkin, $S_N(X) = \frac{k}{N}$, dimana k adalah banyak observasi yang sama atau kurang dari X [19].

Kriteria pengujian: H_0 ditolak jika $D > Kolmogorov_{tabel}$

Pengujian homogenitas dilakukan dengan menggunakan uji Levene. Adapun pengujian Levene adalah sebagai berikut:

$$H_0: \sigma_1^2 = \sigma_2^2 = \dots = \sigma_k^2$$

H_1 : paling sedikit ada satu σ_i^2 yang tidak sama

Statistik uji dapat dilihat pada persamaan (7) sebagai berikut:

$$W = \frac{(N-k) \sum_{i=1}^k N_i (\bar{Z}_i - \bar{Z}_..)^2}{k-1 \sum_{i=1}^k \sum_{j=1}^{N_i} (Z_{ij} - \bar{Z}_i)^2} \quad (7)$$

Dimana N adalah jumlah observasi, k adalah banyak kelompok, $Z_{ij} = |Y_{ij} - \bar{Y}_i|$, dimana \bar{Y}_i Adalah rata-rata dari kelompok ke- i . \bar{Z}_i adalah rata-rata kelompok dari Z_i , $\bar{Z}_..$ adalah rata-rata menyeluruh dari Z_{ij} [20].

Kriteria pengujian: H_0 ditolak jika $W > F_{(\alpha; k-1, n-k)}$

Seperti dijelaskan pada Tabel 1. maka jika data berdistribusi normal dan homogen, maka pengujian hipotesis perbedaan dua rata-rata menggunakan uji-t, berikut adalah statistik ujinya dapat dilihat pada persamaan (8) dan (9):

$$t = \frac{\bar{x}_1 - \bar{x}_2}{s \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \quad (8)$$

dengan

$$s^2 = \frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1 + n_2 - 2} \quad (9)$$

Dengan taraf nyata α , maka kriteria pengujian adalah: terima H_0 jika $-t_{1-\frac{1}{2}\alpha} < t < t_{1-\frac{1}{2}\alpha}$ dimana $t_{1-\frac{1}{2}\alpha}$ didapat dari daftar student dengan $dk = n_1 + n_2 - 2$ peluang $1 - \frac{1}{2}\alpha$. Dalam hal lainnya H_0 ditolak.

Jika maka jika data berdistribusi normal dan tidak homogen, maka pengujian hipotesis perbedaan dua rata-rata menggunakan uji-t', berikut adalah statistik ujinya pada persamaan (10):

$$t' = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \quad (10)$$

Dengan taraf nyata α , maka kriteria pengujian adalah terima H_0 jika (lihat persamaan 11):

$$-\frac{w_1 t_1 + w_2 t_2}{w_1 + w_2} < t' < \frac{w_1 t_1 + w_2 t_2}{w_1 + w_2} \quad (11)$$

Dengan: $w_i = \frac{s_i^2}{n_i}$ dan $t_i = t_{(1-\frac{1}{2}\alpha), (n_i-1)}$ dengan $i = 1, 2$.

Dalam hal lainnya H_0 ditolak[21]. Jika maka jika data tidak berdistribusi normal maka pengujian hipotesis perbedaan dua rata-rata menggunakan uji Mann withney U, berikut adalah statistik ujinya, lihat persamaan 12.

$$U_1 = n_1 n_2 + \frac{n_2(n_2+1)}{2} - \sum R_2 \quad (12)$$

Dan $U_2 = n_1 n_2 - U_1$

Dimana n_i adalah jumlah sampel ke - i , $\sum R_2$ adalah jumlah ranking ke-2, lihat persamaan 13.

$$Z = \frac{U - \frac{n_1 n_2}{2}}{\sqrt{\frac{n_1 n_2 (n_1 + n_2 + 1)}{12}}} \quad (13)$$

Dimana U dapat dipilih antara U_1 atau U_2 , hasil penggunaan salah satunya akan menghasilkan negatif dari yang lainnya.

Dengan taraf nyata α , maka kriteria pengujian adalah: terima H_0 jika $-z_{1/2(1-\alpha)} < z < z_{1/2(1-\alpha)}$ dimana $z_{1/2(1-\alpha)}$ didapat dari daftar normal baku dengan peluang $1/2(1-\alpha)$. Dalam hal lainnya H_0 ditolak [22].

III. HASIL DAN DISKUSI

Dalam penelitian ini dokumen yang digunakan adalah latar belakang skripsi mahasiswa teknik informatika. Deskripsi data yang digunakan dapat dilihat pada Tabel 2.

TABEL 2. DESKRIPSI DATASET

	N	Min	Max	Mean	Standar Deviasi
Tfidf	100	0,31	0,93	0,6104	0,12734
LGN	100	0,25	0,86	0,6109	0,12931
Valid N (listwise)	100				

Dapat dilihat dari tabel 2 bahwa rata-rata dan standar deviasi akurasi untuk tf-idf adalah 0,6107 atau 61,07% dan 0,1273 sedangkan untuk LGN diperoleh 0,6109 atau 61,09% dan 0,1293. Jika dilihat dari hasil deskripsi data yaitu rata-rata dan standar deviasi, TFIDF lebih baik dibandingkan LGN dikarenakan standar deviasi TFIDF lebih kecil dibandingkan LGN. Hasil ini akan diperkuat dengan pengujian hipotesis, tetapi data set akan diuji normalitas terlebih dahulu. Pengujian pertama dilakukan terhadap data hasil akurasi tf-idf dengan hipotesis sebagai berikut:

H_0 : Data sampel berasal dari populasi berdistribusi normal

H_1 : Data sampel berasal dari populasi yang tidak berdistribusi normal

Kriteria uji: "signifikansi" $> 0,05$ maka H_0 diterima

Pengujian dilakukan dengan menggunakan bantuan *software* IBM SPSS versi 23 dengan output pada Tabel 3.

TABEL 3. HASIL UJI KOLMOGOROV-SMIRNOV

		Tf-Idf
N		100
Normal Parameter	Mean	0,6104
	Standar deviation	0,12734
Most extreme differences	Absolute	0,110
	Positive	0,056
	Negative	-0,110
Test Statistic		0,110
Asymp. Sig (2-tailed)		0,005

Berdasarkan Tabel 3, hasil uji Kolmogorov-Smirnov, nilai signifikansi yang diperoleh yaitu 0,005 untuk metode Tf-Idf maka berdasarkan kriteria uji H_0 ditolak, sehingga dapat disimpulkan bahwa data akurasi hasil dari metode Tf-Idf berasal dari populasi yang tidak berdistribusi normal.

Uji normalitas juga dilakukan terhadap data set hasil akurasi LGN, berikut adalah hipotesisnya:

H_0 : Data sampel berasal dari populasi berdistribusi normal

H_1 : Data sampel berasal dari populasi yang tidak berdistribusi normal

Kriteria uji: "signifikansi">0,05 maka H_0 diterima

TABEL 4. HASIL UJI KOLMOGOROV-SMIRNOV

		LGN
N		100
Normal Parameter	Mean	0,6109
	Standar deviation	0,12931
Most extreme differences	Absolute	0,079
	Positive	0,54
	Negative	-0,079
Test Statistic		0,079
Asymp. Sig. (2-tailed)		0,131

Berdasarkan Tabel 4, hasil uji Kolmogorov-Smirnov, nilai signifikansi yang diperoleh yaitu 0,131 untuk metode LGN maka berdasarkan kriteria uji H_0 diterima. Sehingga dapat disimpulkan bahwa data akurasi hasil dari metode LGN berasal dari populasi yang berdistribusi normal. Setelah dilakukan uji normalitas, pada Tabel 3 dan Tabel 4 menunjukkan normalitas yang berbeda antara hasil akurasi Tf-Idf dan LGN. Berdasarkan Tabel 1, maka pengujian hipotesis perlu dilakukan dengan menggunakan uji hipotesis Mann-Whitney U.

TABEL 5. HASIL UJI HIPOTESIS MANN-WHITNEY U

	Tf-Idf
Mann-Whitney U	4985.500
Wilcoxon W	10035.500
Z	-0,035
Asymp. Sig. (2-tailed)	0,972

Pada Tabel 5. Ditunjukkan hasil pengujian menggunakan uji hipotesis Mann-Whitney U yang menunjukkan signifikansi lebih besar dari 0,05. Hal tersebut dapat disimpulkan bahwa H_0 diterima atau peringkasan hasil Tf-Idf sama dengan hasil peringkasan LGN.

IV. KESIMPULAN

Sejumlah penelitian dalam peringkasan dokumen menyatakan bahwa peringkasan menggunakan ekstraksi kalimat Tf-Idf lebih baik. Akan tetapi penelitian lain pun menunjukkan hal yang bertolak belakang. Dalam penelitian ini hipotesis bahwa hasil peringkasan Tf-Idf lebih baik daripada hasil peringkasan menggunakan skema LGN diuji menggunakan metode uji hipotesis non-parametrik Mann-Whitney U. Hasil pengujian menunjukkan bahwa hasil peringkasan Tf-Idf tidak secara signifikan menunjukkan keunggulan dibandingkan dengan hasil peringkasan skema LGN.

UCAPAN TERIMAKASIH

Penelitian ini merupakan bagian dari penelitian dosen pemula (PDP) tahun 2016 dengan judul Klasifikasi Dokumen Skripsi Menggunakan *Support Vector Machine*. Ucapan terima kasih kepada Dikti khususnya DIPA Direktorat Penelitian Pengabdian kepada Masyarakat Kementerian Pendidikan dan Kebudayaan yang memberikan dukungan materil atas terlaksananya penelitian ini dan kepada Balai Bahasa Jawa Barat atas bantuannya dalam memvalidasi hasil pengujian ringkasan yang telah dilakukan.

DAFTAR PUSTAKA

- [1] K. Spärck-Jones, "What Might Be In Summary?," *Inf. Retrieval '93*, Pp. 9–26., 1993.
- [2] H. P. Luhn, "The Automatic Creation Of Literature Abstracts," *Ibm*, Vol. 2, Pp. 159–165., 1958.
- [3] P. P. Klassen, "Calculating Llr Topic Signatures With Dependency Relations For Automatic Text Summarization," *Univ. Washingt.*, 2012.
- [4] S. Beliga And A. Me', "Toward Selectivity-Based Keyword Extraction For Croatian News," In *Surfacing The Deep And The Social Web*, 2014, Pp. 1–14.
- [5] R. Mihalcea And P. Tarau, "Textrank: Bringing Order Into Texts," In *Empirical Methods In Natural Language Processing*, 2004, Vol. 85.
- [6] E. Filatova And V. Hatzivassiloglou, "A Formal Model For Information Selection In Multi-Sentence Text Extraction," In *Proceedings Of The 20th International Conference On Computational Linguistics, (Ccl' 04)*, *Acm*, 2004, Pp. 397–403.
- [7] Y. J. Kumar, O. S. Goh, H. Basiron, N. H. Choon, And P. C. Suppiah, "A Review On Automatic Text Summarization Approaches," *J. Comput. Sci.*, Vol. 12, No. 4, 2016.
- [8] R. G. Easterling, *Performance Measures For Machine Learning*, Vol. 29, No. 3. 1987, Pp. 267.
- [9] M. Galley, "A Skip-Chain Conditional Random Field For Ranking Meeting Utterances By Importance", 2004.
- [10] E. Hovy, L. Gerber, U. Hermjakob, C. Lin, D. Ravichandran, And M. Rey, "Toward Semantics-Based Answer Pinpointing", March, 2001.
- [11] E. Hovy And C. Lin, "Automated Text Summarization In Summarist," 1999.
- [12] E. Chisholm And T. G. Kolda, "New Term Weighting Formulas For The Vector Space Method In Information Retrieval," *Comput. Sci. Math. Div.*, Pp. 1–16, 1999.
- [13] M. Lan, C. Lim, T. Senior, And J. Su, "Supervised And Traditional Term Weighting Methods For Automatic Text Categorization," *J. Ieee Pami*, Vol. 10, No. 10, Pp. 1–36, 2008.
- [14] V. Murthy,G, D. B. V. Vardhan, K. Sarangam, And P. V. Reddy, PaI, "A Comparative Study On Term Weighting Methods For Automated Telugu Text Categorization With," *Int. J. Data Min. Knowl. Manag. Process*, Vol. 3, No. 6, Pp. 95–105, 2013.
- [15] R. A. Fisher, "Statistical Methods For Research Workers," In *Landmark Writings In Western Mathematics 1640-1940*, 2005, Pp. 856–870.
- [16] J. D. Rajaraman, A.; Ullman, "Data Mining," In *Mining Of Massive Datasets*, Cambridge University Press, 2011, Pp. 1–17.
- [17] J. Beel, B. Gipp, S. Langer, C. Breiting, And C. Breiting, "Research-Paper Recommender Systems : A Literature Survey," *Int. J. Digit. Libr.*, June, 2015.
- [18] J. Manuel And T. Moreno, *Summarization*, 1st Ed. London; Hoboken: Iste; Willey, 2014.
- [19] S. Siegel, *Statistik Non Parametrik Untuk Ilmu-Ilmu Sosial*. Gramedia, 2011.
- [20] S. S. Uyanto, *Pedoman Analisis Data Dengan Spss*, 3rd Ed. Graha Ilmu, 2006.
- [21] Sudjana, *Statistik*, 6th Ed. Bandung: Tarsito Bandung, 2005.
- [22] T. Sriwidadi, "Penggunaan Uji Mann-Whitney pada Analisis Pengaruh Pelatihan Wiraniaga dalam Penjualan Produk Baru," *BINUS*, vol. 2, no. 2, pp. 751–762, 2011.