

## PEMERINGKATAN PENCARIAN PADA BUKU PEDOMAN AKADEMIK FILKOM UB MENUJU MERDEKA BELAJAR DAN FREE E-BOOK PEMBELAJARAN SEBAGAI PROTOTYPE LOCAL SMART MICRO SEARCH ENGINE MENGGUNAKAN ALGORITMA PAGERANK DAN TF-IDF

Imam Cholissodin<sup>\*1</sup>, Akhmad Sa'rony<sup>2</sup>, Rona Salsabila<sup>3</sup>, Ilham Firmansyah<sup>4</sup>, Guedho Augnifico Mahardika<sup>5</sup>, Andreas Pardede<sup>6</sup>, Zaien Bin Umar Alaydrus<sup>7</sup>

<sup>1,2,3,4,5</sup>Fakultas Ilmu Komputer Universitas Brawijaya Malang

Email: <sup>1</sup>imamcs@ub.ac.id, <sup>2</sup>aron97729@gmail.com, <sup>3</sup>ronasalsabila1391@gmail.com, <sup>4</sup>firman.ilham05@gmail.com, <sup>5</sup>edho08@gmail.com, <sup>6</sup>tyrantme1@gmail.com, <sup>7</sup>zenalaydrus054@gmail.com

\*Penulis Korespondensi

(Naskah masuk: 20 November 2020, diterima untuk diterbitkan: 19 Oktober 2021)

### Abstrak

Buku Pedoman Akademik FILKOM Universitas Brawijaya merupakan suatu kebutuhan informasi akademik yang cukup penting, dan juga buku penunjang pembelajaran seperti Free e-Book bagi para mahasiswa. Untuk memperoleh informasi yang relevan terhadap query yang diberikan seringkali belum sesuai dengan kebutuhan pencarian pengguna. Pengguna harus menguasai secara keseluruhan untuk mengetahui dokumen mana yang paling sesuai, dan proses ini akan memakan waktu yang banyak. Sistem ini mampu memberikan rekomendasi dokumen sesuai dengan hasil perhitungan pemeringkatan teks. Proses pemeringkatan teks dapat diselesaikan dengan algoritma PageRank, di mana dokumen yang memiliki bobot pemeringkatan terkecil, memiliki kata terbanyak pada dokumen tersebut. Algoritma ini telah dibuktikan mampu memberikan feedback dokumen yang relevan melalui dua tahap pengujian. Evaluasi yang dilakukan terhadap dua buah pengujian menghasilkan rata-rata nilai recall tertinggi yaitu 80.6% pada data ke-1, dan data ke-2 didapatkan korelasi terbaik antara precision, recall dan f-measure sebesar 0,98, 0,99, 0,99.

**Kata kunci:** *algoritma PageRank, tf-idf, pedoman akademik filkom ub, free e-book, local smart micro search engine*

## SEARCH RANKING IN ACADEMIC GUIDELINES OF FILKOM UB TOWARDS FREEDOM TO LEARN AND FREE E-BOOK LEARNING AS LOCAL SMART MICRO SEARCH ENGINE PROTOTYPE USING PAGERANK AND TF-IDF ALGORITHM

### Abstract

The Brawijaya University FILKOM Academic Handbook is an important academic information need, as well as learning support books such as Free e-Books for students. To obtain information that is relevant to the query given is often not in accordance with the wishes of the user. Users must master the whole to find out which documents are most suitable, which is where the process will take a lot of time. This system is able to provide document recommendations in accordance with the results of the text ranking calculation. The process of ranking the text can be solved by the PageRank algorithm, where documents that have the smallest ranking weight, have the most words in the document. This algorithm has been proven to be able to provide feedback on relevant documents through two stages of testing. The evaluation conducted on the two tests resulted in the highest average recall value of 80.6% on the 1st dataset, and 2nd dataset the best correlation was obtained between precision, recall and f-measure of 0.98, 0.99, 0.99.

**Keywords:** *PageRank algorithm, tf-idf, academic guidelines of filkom ub, free e-book, local smart micro search engine*

### 1. PENDAHULUAN

Pemeringkatan *text retrieval* adalah suatu proses dalam Sistem Temu Kembali Informasi di

mana dilakukan proses pemeringkatan terhadap hasil-hasil *retrieval* yang dikembalikan oleh sistem, sehingga nantinya dapat memudahkan pengguna untuk mendapatkan hasil yang paling relevan. Salah

satu algoritma pemeringkatan yang populer yaitu algoritma PageRank. Algoritma PageRank (Brin dan Page, 1998) adalah komponen penting dari Google Web yang merupakan mesin pengambilan informasi (IR) Google dan memicu penelitian dan penerapan analisis tautan ke Web. Hal ini memungkinkan perhitungan kepentingan priori untuk halaman-halaman Web. Perhitungan tersebut dilakukan secara luring dan bersifat *independent* terhadap search *query*. Nilai PageRank membentuk vektor Eigen yang sesuai dengan nilai Eigen 1 dari matriks Web. Pada saat *query* time, hasil perhitungan digabungkan dengan skor spesifik dari *query* yang digunakan untuk menentukan peringkat halaman Web dengan mengambil nilai tertinggi (Mihalcea, 2004).

Pada era modern saat ini, mayoritas lini kehidupan menjadi serba digital. Mulai dari kegiatan transaksi, pertukaran pesan hingga dokumen-dokumen semuanya semakin mudah dilakukan dan diakses karena perkembangan era digital yang sangat pesat. Hal ini tentu semakin memudahkan manusia dalam menjalani aktivitas sehari-hari. Saat ini, terdapat banyak sekali dokumen digital yang dengan mudah dapat didapatkan dalam internet. Seiring berkembangnya waktu semuanya menjadi semakin mudah. Manusia dapat memperoleh bacaan tanpa harus memiliki buku fisik. Sejak beberapa dekade terakhir, sistem temu kembali informasi dikembangkan dengan sangat pesat. Hal ini dikarenakan adanya sistem temu kembali informasi dianggap sangat memudahkan dalam memperoleh informasi dari dokumen digital. Proses perolehan informasi baik secara umum maupun spesifik dapat dilakukan dengan lebih mudah dan efisien, dan lagi proses tersebut dapat dilakukan pada banyak dokumen dalam satu waktu, sehingga hal ini sangat bermanfaat. Dalam sebuah dokumen, terdapat banyak sekali informasi yang tersimpan didalamnya. Tentu menjadi hal yang sulit bagi pembaca untuk memperoleh informasi secara spesifik dari dokumen tersebut. Apabila proses pencarian dilakukan secara manual maka tentu akan memakan waktu yang tidak sebentar. Oleh karena itu, diperlukan adanya sebuah sistem temu kembali informasi yang dapat membantu pengguna untuk memperoleh informasi relevan dari halaman-halaman spesifik yang mengandung informasi yang dicari oleh pengguna.

Pada penelitian ini, dibuat sistem temu kembali informasi untuk menemukan daftar halaman-halaman dari buku pedoman akademik FILKOM UB dan Free e-Book yang mengandung informasi spesifik yang dibutuhkan oleh pengguna. Algoritma yang akan digunakan dalam mengembangkan sistem ini, yaitu algoritma PageRank yang digunakan Google dalam mesin pencariannya (Chung, 2014), dengan basis ekstraksi tf-idf dari suatu dokumen tunggal. Nantinya, pengguna akan memasukkan kata kunci dari informasi yang ingin dicari pada dokumen tersebut, kemudian sistem akan mulai mencari halaman-halaman yang mengandung informasi yang

dicari oleh pengguna tersebut. Kemudian, setelah halaman-halaman yang sesuai telah didapatkan, kemudian akan dilakukan proses pemeringkatan berdasarkan tingkat kesesuaiannya, sehingga nantinya hal ini akan semakin memudahkan pengguna dalam mendapatkan halaman dokumen yang paling relevan terhadap informasi yang dicari oleh pengguna. Berdasarkan paper yang menggunakan pemeringkatan dokumen bahasa Arab berbasis Latent Semantic Indexing (Wahib, dkk, 2014), digunakan pula basis ekstraksi tf-idf dalam melakukan pemeringkatan. Hasilnya, didapatkan nilai f-measure 41%, 51% dan 60% pada ambang cosine similarity 0,3, 0,4, dan 0,5. Waktu komputasi rata-rata yang dihasilkan yaitu berada di kisaran 2 menit 8 detik. Sehingga dapat disimpulkan bahwa penggunaan algoritma pemeringkatan dengan basis ekstraksi tf-idf efektif digunakan.

## 2. DASAR TEORI

### 2.1. Tentang Dataset

Pada penelitian ini, data yang digunakan ialah buku “Pedoman Akademik Fakultas Ilmu Komputer (FILKOM) Universitas Brawijaya (UB) Menuju Merdeka Belajar” (Kemdikbud, 2020)(Filkom, 2020), yaitu berupa teks. Data yang digunakan ialah 90 halaman dari 218 halaman, di mana data tersebut diambil secara acak dan berurutan dari halaman 12 hingga halaman 102. Dan data ke-2 yaitu dokumen e-book yang diunduh secara gratis dari <https://www.deeplearningbook.org>. E-Book yang di-*download* merupakan e-Book tentang Deep Learning yang diterbitkan oleh MIT Press dan dapat diunduh secara gratis dengan format HTML/WEB. Dataset kemudian dirubah menjadi berformat TXT yang memungkinkan kemudahan pembacaan oleh program. Dokumen berisi 800 halaman yang akan digunakan secara keseluruhan.

### 2.2. Temu Kembali Informasi

Temu Kembali Informasi (Information Retrieval) adalah proses pencarian informasi yang biasanya berupa dokumen dari suatu kondisi atau kumpulan dokumen yang tidak terstruktur untuk memenuhi kebutuhan informasi dari koleksi yang besar (Manning, Raghavan, & Schütze, 2009). Dokumen yang tidak terstruktur tersebut biasanya dalam bentuk teks akan tetapi dewasa ini juga termasuk data tidak terstruktur lainnya seperti media berupa citra, video audio dan lainnya.

Term Frequency (TF) merupakan cara penghitungan suatu frekuensi kemunculan term dalam sebuah dokumen (Cholissodin, Setiawan, 2013). Semakin tinggi nilai frekuensi dari suatu term maka nilai pembobotannya akan semakin tinggi. Terdapat beberapa cara dalam menghitung nilai TF yaitu TF biner, TF murni dan TF logaritmik. TF logaritmik banyak digunakan karena dapat mengatasi masalah kurangnya dominasi dan terlalu

dominasi dokumen dengan TF besar pada TF biner dan TF murni. Persamaan 1 adalah rumus menghitung TF logaritmik (Thomas, 2013).

$$TF = \begin{cases} 1 + \log_{10} f_t, & f_t > 0 \\ 0 & \end{cases} \quad (1)$$

Inverse Document Frequency (IDF) adalah invers dari TF, di mana nilai distribusi dari term pada kumpulan dokumen atau korpus (Thomas, 2013). Nilai IDF suatu term akan semakin kecil apabila term akan semakin kecil apabila term tersebut tersebar pada banyak dokumen pada corpus. Berikut adalah rumus untuk menghitung IDF.

tf-idf adalah metode pembobotan yang di mana salah satu caranya adalah mengkalikan TF dengan. Pembobotan tf-idf banyak digunakan karena TF-IDF memberikan nilai pembobotan yang kecil pada term yang tidak dapat memisahkan dokumen dan term yang sedikit ditemukan pada dokumen. tf-idf dapat dihitung menggunakan rumus berikut.

$$W_{ij} = TF_{ij} IDF_i \quad (2)$$

### 2.3. PageRank

Algoritma PageRank digunakan pada web mining untuk menampilkan hasil di search engine (Z. Hao et. al., 2015)(J. Berkhout, 2016)(F. Ali, 2016)(Usha & Dr.N.Nagadeepa, 2018). Algoritma PageRank terbagi menjadi dua jenis yaitu Simplified Algorithm dan Random Surfer Model, yang di mana keduanya terdiri atas dua cara penyelesaian, yaitu dengan metode iteratif dan representasi matriks. Penelitian ini menggunakan Random Surfer Model dengan cara representasi matriks, hal tersebut dapat diselesaikan dengan Persamaan (3).

$$PR_{t+1} = (M^T)^{(t+1)} . PR_0 \quad (3)$$

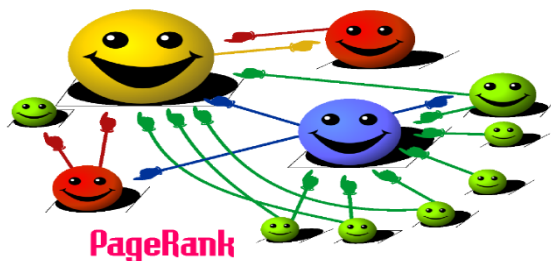
Keterangan:

$PR_{t+1}$  = Nilai PageRank

$(M^T)^{(t+1)}$  = Transposisi dari matriks yang dihitung

$PR_0$  = Faktor damping (bernilai antara 0 dan 1)

Secara sederhana, PageRank adalah cara mengukur pentingnya halaman situs web. Ilustrasi pada Gambar 1 menggambarkan prinsip dasar algoritma PageRank. Ukuran setiap wajah sebanding dengan tingkatan seberapa pentingnya suatu halaman.

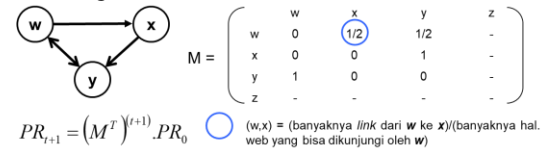


Gambar 1. Ilustrasi Algoritma PageRank

### 3. IMPLEMENTASI DAN HASIL

Pada implementasi Random Surfer Model dengan cara representasi matriks, berikut diberikan ilustrasi dalam bentuk graf, misal merepresentasikan link-link antar 3 halaman website, yaitu w, x, dan y. (z telah di-*remove*). Berikut detail langkah-langkahnya.

- Skema graf.



Misal nilai faktor damping (*teleport*) = 0,85 menggunakan “*power method*”, misal dengan iterMax = 6, di mana  $PR_0 = teleport * ones(nNodeAktif, 1)$ ,  $nNodeAktif = 3$ .

- Proses iterasi.

ke-1 dari {0,1,2,3,4,5}.

$$PR_1 = (M^T)^1 . PR_0 = \begin{pmatrix} 0 & 0 & 1 \\ 1/2 & 0 & 0 \\ 1/2 & 1 & 0 \end{pmatrix} * \begin{pmatrix} 0,85 \\ 0,85 \\ 0,85 \end{pmatrix} = \begin{pmatrix} 0,8500 \\ 0,4250 \\ 1,2750 \end{pmatrix}$$

ke-2 dari {0,1,2,3,4,5}.

$$PR_2 = (M^T)^2 . PR_0 = \begin{pmatrix} 0 & 0 & 1 \\ 1/2 & 0 & 0 \\ 1/2 & 1 & 0 \end{pmatrix}^2 * \begin{pmatrix} 0,85 \\ 0,85 \\ 0,85 \end{pmatrix} = \begin{pmatrix} 1,2750 \\ 0,4250 \\ 0,8500 \end{pmatrix}$$

ke-3 dari {0,1,2,3,4,5}.

$$PR_3 = (M^T)^3 . PR_0 = \begin{pmatrix} 0 & 0 & 1 \\ 1/2 & 0 & 0 \\ 1/2 & 1 & 0 \end{pmatrix}^3 * \begin{pmatrix} 0,85 \\ 0,85 \\ 0,85 \end{pmatrix} = \begin{pmatrix} 0,8500 \\ 0,6375 \\ 1,0625 \end{pmatrix}$$

ke-4 dari {0,1,2,3,4,5}.

$$PR_4 = (M^T)^4 . PR_0 = \begin{pmatrix} 0 & 0 & 1 \\ 1/2 & 0 & 0 \\ 1/2 & 1 & 0 \end{pmatrix}^4 * \begin{pmatrix} 0,85 \\ 0,85 \\ 0,85 \end{pmatrix} = \begin{pmatrix} 1,0625 \\ 0,4250 \\ 1,0625 \end{pmatrix}$$

ke-5 dari {0,1,2,3,4,5}.

$$PR_5 = (M^T)^5 . PR_0 = \begin{pmatrix} 0 & 0 & 1 \\ 1/2 & 0 & 0 \\ 1/2 & 1 & 0 \end{pmatrix}^5 * \begin{pmatrix} 0,85 \\ 0,85 \\ 0,85 \end{pmatrix} = \begin{pmatrix} 1,0625 \\ 0,5313 \\ 0,9563 \end{pmatrix} \Rightarrow \begin{matrix} \text{Hasil PageRank} \\ \begin{pmatrix} w \\ x \\ y \end{pmatrix} \rightarrow \begin{pmatrix} 3 \\ 1 \\ 2 \end{pmatrix} \end{matrix}$$

Penelitian ini menggunakan ekstraksi fitur tf-idf, di mana algoritma tf-idf ini berfungsi untuk menganalisis hubungan antara suatu term dengan sekumpulan dokumen, melalui bobot setiap term yang diperoleh dari hasil perhitungan tf-idf. Setelah itu, dilakukan perhitungan jarak antar dokumen menggunakan persamaan *Cosine similarity* dari hasil yang diperoleh pada perhitungan tf-idf. Setelah itu, penelitian ini menggunakan algoritma PageRank di mana algoritma ini berfungsi untuk mengambil keputusan ketika melakukan pemeringkatan halaman berdasarkan hasil voting dari suatu dokumen, hal ini dilakukan untuk mengetahui tingkat kepentingan suatu halaman berdasarkan keterkaitan antar halaman pada dokumen tersebut.

Pemeringkatan menggunakan algoritma PageRank dimulai dengan proses *case folding*, tokenisasi, kemudian melakukan stemming terhadap term yang telah ditokenisasi untuk memperoleh kata dasar dari setiap term. Setelah menemukan kata dasar dari hasil proses stemming, maka langkah selanjutnya ialah melakukan pembobotan kata menggunakan ekstraksi fitur tf-idf. Setelah menemukan nilai dari tf-idf, maka proses selanjutnya ialah melakukan perhitungan jarak menggunakan *Cosine similarity* (Nyein, 2011). *Cosine similarity* digunakan untuk mengetahui jarak dari setiap dokumen yang menunjukkan kemiripan dari suatu dokumen. Setelah mengetahui jarak, langkah selanjutnya ialah melakukan pemeringkatan dokumen dengan menggunakan algoritma PageRank berdasarkan nilai jarak yang telah diperoleh dari perhitungan *cos similarity*. Langkah terakhir ialah melakukan pengurutan data dari yang memiliki nilai PageRank terkecil. Proses dari pemeringkatan tersebut ditunjukkan dari potongan Kode Program 1.

```

1 def kosim_(tf_idf_):
2     pembilang_ = [ [0 for y_ in range \
3         (len(dokumen_)) ] for x_ in range \
4         (len(dokumen_)) ]
5
6     penyebut_ = [ [0 for y_ in range \
7         (len(dokumen_)) ] for x_ in range \
8         (len(dokumen_)) ]
9
10    for i_ in range (len(pembilang_)):
11        for j_ in range \
12            (len(penyebut_)):
13            for k_ in range \
14                (len(tf_idf_[i_])):
15                pembilang_[i_][j_] += \
16                    tf_idf_[i_][k_] \
17                    *tf_idf_[j_][k_]
18                penyebut_[i_][j_] += \
19                    pembilang_[i_][j_]**2
20
21    kosim_ = [ [0 for y_ in \
22        range(len(dokumen_)) ] for x_ in \
23        range(len(dokumen_)) ]
24
25    for i_ in range (len(kosim_)):
26        for j_ in range \
27            (len(kosim_[i_])):
28            if(penyebut_[i_][j_] == 0):
29                kosim_[i_][j_] = 0
30            else:
31                kosim_[i_][j_] = \
32                    pembilang_[i_][j_] \
33                    /math.sqrt( \
34                    penyebut_[i_][j_])
35    return kosim_
36
37 def pagerank_(kosim_):
38     mt_ = kosim_.transpose()
39     pr0_ = [1/len(dokumen_) for x_ in \
40         range(len(dokumen_)) ]
41     hasil_ = np.matmul(mt_, pr0_)
42     return hasil_
43
44

```

Kode Program 1. Pemeringkatan Teks

Penjelasan dari Kode Program 1:

- Baris 1-35 merupakan proses mencari nilai jarak dari setiap dokumen menggunakan persamaan *cos similarity*.
- Baris 37-42 merupakan proses pemeringkatan teks dengan menggunakan persamaan algoritma PageRank.

### 3.1. Pengujian Pencarian pada Buku Pedoman FILKOM UB

Hasil evaluasi pengujian pertama pada Tabel 1 dan pengujian kedua pada Tabel 2. Pengujian ini menggunakan 20 variasi data uji. Pengujian dilakukan dua kali, di mana setiap pengujianya menggunakan 10 *query* yang akan diuji. Pengujian dilakukan terhadap 90 dokumen buku pedoman FILKOM UB dari halaman 12 hingga 102.

Tabel 1. Hasil Pengujian ke-1 Pada Buku Pedoman FILKOM UB

Query	Precision	Recall	F-Measure
1	0,17	0,75	0,75
2	0,50	0,78	0,78
3	0,20	0,25	0,33
4	0,33	0,33	0,33
5	0,08	1,00	0,15
6	0,02	0,50	0,04
7	0,13	1,00	0,23
8	0,12	1,00	0,21
9	0,11	1,00	0,19
10	0,06	1,00	0,12
Rata-rata	0,17	0,76	0,24

Tabel 2. Hasil Pengujian ke-2 Pada Buku Pedoman FILKOM UB

Query	Precision	Recall	F-Measure
1	0,83	0,70	0,76
2	0,86	0,57	0,68
3	0,11	0,80	0,19
4	0,33	0,33	0,33
5	0,04	1,00	0,07
6	0,89	1,00	0,94
7	0,67	1,00	0,80
8	0,31	0,92	0,46
9	0,20	0,78	0,32
10	0,02	1,00	0,04
Rata-rata	0,42	0,81	0,46

### 3.2. Pengujian Pencarian Pada *Free e-Book* Penunjang Pembelajaran

Pengujian *precision*, *recall* dan *f-measure* dengan variabel *independent* persentase peringkasan menunjukkan hubungan korelasi kuat positif (nilai 0,98, 0,99, dan 0,99). Akan tetapi pada Tabel 3, kenaikan *precision*, *recall* dan *f-measure* pada persentase peringkasan 10% - 30% tidak terlalu tinggi dengan nilai berturut-turut 9%, 7%, dan 8%. Pengujian ini menunjukkan bahwa semakin panjang kueri pada sistem IR yang relevan dengan dokumen yang ingin di-*retrieve* maka semakin banyak dokumen relevan yang di-*retrieve* sistem IR. Akan tetapi, pengujian menguji panjang kueri yang relevan pada IR relatif terhadap jumlah dokumen pada cluster dokumen yang relevan dan tidak menguji panjang kueri absolut sehingga panjang

kueri yang dimasukkan mencari dokumen tiap cluster dokumen memiliki jumlah yang berbeda.

Tabel 3. Pengujian Persentase Peringkasan

Persentase Peringkasan	Precision	Recall	F-Measure
10%	0,58	0,44	0,36
30%	0,60	0,45	0,37
50%	0,64	0,48	0,38

Tabel 4. Pengujian Banyak Cluster

Banyak Cluster	Precision	Recall	F-Measure
4	0,67	0,40	0,50
6	0,60	0,37	0,47
8	0,58	0,36	0,46
10	0,55	0,35	0,43

Pengujian *precision*, *recall* dan *f-measure* dengan variabel *independent* jumlah cluster menunjukkan hubungan kuat negatif (nilai -0,97, -0,97 dan -0,97). Pada Tabel 4, variabel terikat *precision*, *recall* dan *f-measure* menunjukkan penurunan yang cukup signifikan dengan nilai berturut-turut -22%, -13%, dan -16%. Pengujian ini menunjukkan bahwa semakin tinggi standar relevan dokumen yang digunakan maka performa sistem IR dalam *re-trtrieve* dokumen relevan semakin buruk.

#### 4. KESIMPULAN DAN SARAN

Pemeringkatan *text retrieval* dalam pencarian isi halaman yang relevan pada dokumen Buku Pedoman Akademik FILKOM UB dan pada *Free e-Book* pembelajaran mampu memberikan *feedback* yang relevan berdasarkan hasil pemeringkatan teks menggunakan algoritma PageRank. Pengujian dengan ekstraksi fitur tf-idf pada kasus dokumen buku pedoman FILKOM UB menghasilkan nilai rata-rata tertinggi ketika menggunakan evaluasi *recall* pada pengujian pertama, yaitu sebesar 76,1% dengan nilai *precision* sebesar 17,3% dan nilai *f-measure* sebesar 23,9%. Sedangkan pada pengujian kedua, diperoleh nilai *precision*, *recall*, dan *f-measure* sebesar 42,4%, 80,9%, dan 45,8%. Nilai evaluasi *recall* yang tergolong baik membuktikan bahwa algoritma PageRank baik dalam melakukan pemeringkatan teks pada kasus pemeringkatan dokumen buku pedoman FILKOM UB.

Sedangkan pada *Free e-Book* didapatkan hubungan korelasi yang tinggi antara *precision*, *recall* dan *f-measure* yaitu 0,98, 0,99, 0,99 pada pengujian persentase peringkasan. Tetapi nilai-nilai tersebut masih belum begitu bagus. Kemudian pada pengujian jumlah cluster didapatkan nilai korelasi yang tinggi tetapi negatif dengan nilai -0,97, -0,97 dan -0,97. Hal ini dikarenakan jika membuat standar relevannya tinggi maka akan dapat berdampak kurang baik pada data yang akan ditemukan. Hal itu tampak dari hasil pengujian variabel independennya menunjukkan hubungan kuat negatif. Pada penelitian ini, *feedback* yang

dihasilkan oleh sistem masih terpacu pada kesesuaian term hasil dari proses *case folding*, tokenisasi, dan *stemming*. Sehingga perlu adanya peningkatan kemampuan dari sistem untuk dapat memahami dokumen secara semantik.

#### DAFTAR PUSTAKA

- BRIN, S., & PAGE, L., 1998. The Anatomy of a Large-Scale Hypertextual Web Search Engine. Diambil kembali dari The Anatomy of Search Engine: <http://infolab.stanford.edu/~backrub/google.html>
- CHOLISSODIN I., SETIAWAN B.D., 2013. Sentiment Analysis Dokumen E-Complaint Kampus Menggunakan Additive Selected Kernel SVM. Seminar Nasional Teknologi Informasi Dan Aplikasinya (SNATIA).
- CHUNG, F., 2014. A Brief Survey of PageRank Algorithms. Ieee Transactions On Network Science And Engineering, Vol. 1, No. 1.
- F. ALI, I. ULLAH AND S. KHUSRO, 2016. An Empirical Investigation of PageRank and Its Variants in Ranking Pages on the Web. International Conference on Frontiers of Information Technology (FIT), Islamabad, pp. 354-359, doi: 10.1109/FIT.2016.071.
- FILKOM, 2020. FILKOM UB Selenggarakan Konsolidasi Penyusunan Kurikulum Kampus Merdeka. Di ambil dari Web FILKOM: <https://filkom.ub.ac.id/page/read/news/filkom-ub-selenggarakan-konsolidasi-penyusunan-kurikulum-kampus-merdeka/07c25e5>
- J. BERKHOUT, 2016. Google's PageRank algorithm for ranking nodes in general networks. 13th International Workshop on Discrete Event Systems (WODES), Xi'an, pp. 153-158, doi: 10.1109/WODES.2016.7497841.
- KEMDIKBUD, 2020. Reformasi Pendidikan Nasional Melalui Merdeka Belajar. Di ambil dari Web kemdikbud: <https://www.kemdikbud.go.id/main/blog/2020/05/reformasi-pendidikan-nasional-melalui-merdeka-belajar>
- MIHALCEA, R., 2004. Graph-based Ranking Algorithms for Sentence Extraction, Applied to Text Summarization. Di ambil dari Association for Computational Linguistics: <https://www.aclweb.org/anthology/P04-3020>.
- NYEIN, S. S., 2011. Mining Contents in Web Page Using Cosine Similarity. [ieeexplore.ieee.org](http://ieeexplore.ieee.org).

- WAHIB, A., PASNUR P., SANTIKA, P. P., ARIFIN, A. Z., 2015. Perangkingan Dokumen Berbahasa Arab Menggunakan Latent Semantic Indexing. 6, p.83-91.
- THOMAS, R., 2013. Information Retrieval Models. London: Synthesis Lectures On Information Concepts, Retrieval, And Services .
- USHA, M., & DR. N. NAGADEEPA., 2018. Combined Two Phase Page Ranking Algorithm for Sequencing the Web Pages. Proceedings of the Second International Conference on Inventive Systems and Control (ICISC 2018) . India: IEEE Xplore Compliant.
- Z. HAO, P. QIUMEI, Z. HONG AND S. ZHIHAO, 2015. An Improved PageRank Algorithm Based on Web Content. 14th International Symposium on Distributed Computing and Applications for Business Engineering and Science (DCABES), Guiyang, pp. 284-287, doi: 10.1109/DCABES.2015.78.