

# IMPLEMENTASI METODE TERM RELATEDNESS TO QUERY UNTUK PENCARIAN BUAH DAN SAYURAN LOKAL JEMBER

Barki I. Huda<sup>1</sup>, Wiwik Suharso<sup>2</sup>

Program Studi Teknik Informatika, Fakultas Teknik,  
Universitas Muhammadiyah Jember

Email : [1barkiilmanhuda@gmail.com](mailto:barkiilmanhuda@gmail.com), [2wiwiksuharso@unmuhjember.ac.id](mailto:wiwiksuharso@unmuhjember.ac.id)

## ABSTRAK

Pengguna memiliki pengetahuan dan kemampuan yang berbeda dalam mengekspresikan kata kunci pencarian. Kualitas *query* menjadi salah satu faktor yang menentukan akurasi informasi yang dihasilkan sistem sehingga pengguna membutuhkan informasi *suggestion*. Penelitian ini mengajukan metode *term relatedness to query* (TRQ), sebuah metode untuk mengoptimasi pembobotan pada *query expansion*. Pembobotan TRQ memberikan nilai pada *term* berdasarkan hubungannya dengan *query*. Evaluasi hasil uji coba menunjukkan bahwa metode TRQ dapat melakukan temu kembali dokumen dengan hasil rata-rata kinerja *recall* terbaik dari skenario 1, 2, 3 dan 4 sebesar 100%. Hasil pengujian rata-rata kinerja *precision* terbaik dari skenario 1, 2, 3 dan 4 sebesar 7%. Dan hasil pengujian rata-rata kinerja *accuracy* terbaik dari skenario 1, 2, 3 dan 4 sebesar 87%. Metode TRQ untuk optimasi pembobotan pada *query* terbukti memberikan pengaruh untuk meningkatkan relevansi pencarian dokumen.

**Kata Kunci** : TRQ, pencarian, *term relatedness to query*, ensiklopedia, ekspansi *query*, *suggestion*.

## ABSTRACT

*Users have different knowledge and abilities in expressing search keywords. The quality of the query is one of the factors that determines the accuracy of information generated by the system so that users need suggestion information. This research proposes the term relatedness to query (TRQ) method, a method for optimizing the weighting of query expansion. TRQ weighting assigns values to terms based on their relationship to the query. Evaluation of trial results shows that the TRQ method can retrieve documents with the best average recall performance results from scenarios 1, 2, 3 and 4 of 100%. The best average performance test results from scenarios 1, 2, 3 and 4 are 7%. And the results of testing the best average accuracy performance of scenarios 1, 2, 3 and 4 by 87%. The TRQ method for optimizing weighting on queries has been proven to have an effect on increasing the relevance of document search.*

**Keywords** : TRQ, searching, *term relatedness to query*, encyclopedia, *query expansion*, *suggestion*.

## 1. PENDAHULUAN

Ensiklopedia sudah di kenal oleh kalangan pelajar sebagai media untuk mendapatkan informasi tentang topik tertentu yang di inginkan. Kebanyakan produk ensiklopedia di pasaran dalam bentuk buku, majalah, atlas dan kartu. Produk ensiklopedia fisik tersebut telah digunakan dalam proses pembelajaran siswa-siswi di sekolah. Akan tetapi produk ensiklopedia fisik memiliki keterbatasan dalam kemudahan akses dan kecepatan penyebaran informasi serta bersifat statis. Ensiklopedia buah dan sayuran lokal Jember berisi

48 jenis buah lokal dan 51 jenis sayuran lokal yang tersebar di Kabupaten Jember (Suharso, 2018).

Jumlah informasi yang tersedia secara elektronik meningkat secara dramatis. Dalam melakukan pencarian informasi membutuhkan metode untuk mengidentifikasi dokumen yang relevan terhadap *query user* (Boston, 2014). Pengguna memiliki pengetahuan dan kemampuan yang berbeda dalam mengekspresikan kata kunci pencarian (*query*). *Query* yang berbeda akan menghasilkan informasi yang berbeda sehingga kualitas *query* menjadi salah satu faktor yang

menentukan akurasi informasi yang dihasilkan sistem. Sebagian besar pengguna terutama masyarakat umum memiliki peluang untuk memberikan *query* yang tidak sesuai atau tidak lengkap sehingga pengguna membutuhkan informasi *suggestion*.

Dalam penelitian ini akan diimplementasikan sebuah metode yang nantinya digunakan untuk memberikan informasi *suggestion*, sehingga banyak membantu pengguna dalam memberikan kemudahan untuk mencari informasi kata yang dibutuhkan.

Penelitian ini bertujuan untuk mengimplementasikan sistem pencarian informasi untuk mengatasi permasalahan yang dialami pengguna dalam melakukan pencarian, maka diperlukan suatu metode agar hasil pencarian dapat maksimal. Terdapat beberapa metode yang dapat diimplementasikan untuk memberikan kata saran yang paling mendekati. Metode yang diajukan ini disebut dengan TRQ (*Term Relatedness to Query*). Pembobotan TRQ memberikan nilai pada *term* berdasarkan hubungannya dengan *query*. Proses pada metode TRQ yaitu dengan menggabungkan *lexical word frequency* (lwf) dengan *invers document frequency* (idf).

### 1.1. Rumusan Masalah

Dalam penelitian ini rumusan masalah yang diberikan adalah :

1. Bagaimana hasil implementasi metode *Term Relatedness to Query* (TRQ) dalam sistem pencarian informasi untuk menghasilkan *suggestion* informasi buah dan sayuran lokal Jember secara akurat dan spesifik sesuai *query* pengguna?
2. Berapa tingkat *precision*, *recall* dan *accuracy* dalam mengukur kinerja sistem?

### 1.2. Batasan Masalah

Batasan-batasan dalam penelitian ini sebagai berikut.

1. Dataset penelitian menggunakan data buah dan sayuran lokal Jember dari buku "Buah dan Sayur Lokal di Kabupaten Jember".
2. Jumlah data buah dan sayuran lokal Jember sebanyak 99 jenis (Sawitri, 2017).
3. Parameter dataset yang digunakan adalah nama, deskripsi, manfaat buah dan sayuran

dan *keyword* yang berada dalam database ensiklopedia.

4. Hasil pencarian dari rekomendasi *query* tidak berbasis frasa.
5. Untuk mengukur kinerja sistem menggunakan *precision*, *recall* dan *accuracy*.

### 1.3. Tujuan Penelitian

Tujuan penelitian ini adalah untuk :

1. Mengetahui hasil implementasi metode TRQ (*Term Relatedness to Query*) dalam sistem pencarian informasi untuk menghasilkan *suggestion* informasi buah dan sayuran lokal Jember secara akurat dan spesifik sesuai *query* pengguna.
2. Mengetahui tingkat *precision*, *recall* dan *accuracy* dalam mengukur kinerja sistem.

### 1.4. Manfaat Penelitian

1. Sebagai referensi bagi kalangan umum khususnya pelajar agar dapat menggunakan produk Ensiklopedia untuk membantu dalam mendapatkan informasi tentang topik buah dan sayuran lokal Jember dengan lebih mudah.
2. Memudahkan pengguna dalam mencari informasi yang diinginkan.
3. Sistem Ensiklopedia dapat memberikan rekomendasi *Query* berdasarkan *query* yang diinginkan pengguna dengan metode TRQ.

## 2. TINJAUAN PUSTAKA

### 2.1. Data Buah dan Sayuran

Sayuran dan buah merupakan sumber serat pangan yang sangat mudah ditemukan dalam bahan makanan. Akhir-akhir ini, karena pola konsumsi pangan di Indonesia menyebabkan kekurangan konsumsi sayuran dan buah-buahan pada hampir seluruh provinsi di Indonesia (Santoso, 2011).

Sawitri (2017) menyatakan bahwa Buah dan Sayuran lokal Jember merupakan merupakan bahan pangan utama dalam kehidupan sehari-hari. Jenis buah dan sayuran ini memiliki kurang lebih 99 macam jenis yang keanekaragamannya sangat bervariasi. Keanekaragaman warna pada buah bukanlah sekedar pembeda jenis antar buah yang satu dengan yang lainnya. Data Buah dan Sayur ini

sudah banyak di sajikan berupa Buku Bacaan, Majalah, Atlas dan lain-lain.

## 2.2. Web Ensiklopedia

Mauludin (2012) menyatakan bahwa Web Ensiklopedia yang berbasis Teknologi Informasi (TI) merupakan tempat untuk mencari informasi sesuai kebutuhan sehari-hari, sehingga pengunjung web tersebut dapat mengendalikan situs tersebut sesuai fitur-fitur yang di sediakan. Beberapa situs website yang meyediakan layanan yang mengandung Ensiklopedia sebagai berikut :

<http://www.wikipedia.org> merupakan situs yang mengandung layanan Ensiklopedia. Ensiklopedia bebas berbahasa Indonesia dan di bangun oleh para sukarelawan, serta gratis. Salah satu keunggulannya adalah di sediakan fasilitas pencarian, seperti situs *search engine*.

### Kelemahan :

1. Karena platform itu memungkinkan siapapun dapat terlibat dalam penulisan/pengeditan.
2. Kurangnya kesimpulan yang mendetail.
3. Kepengarangan/sumber informasi kurang di kenal.

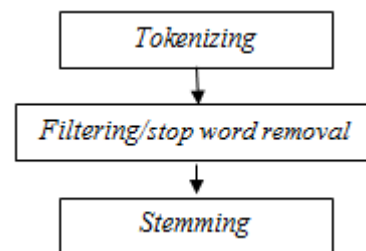
## 2.3. Text Mining

*Text mining* didefinisikan sebagai proses penemuan kembali relasi dan fakta yang terkubur didalam teks dan tidak harus baru. Dalam penelitian ini *text mining* meliputi *tokenizing*, *stoplist/wordlist/Filt*, *stemming*. Dalam penelitian ketiga operasi *text mining* tersebut sebagai tahap persiapan untuk menyaring dan mengurangi jumlah informasi atau term-term yang tidak memiliki relevansi dalam pengukuran derajat kemiripan suatu dokumen. Proses selanjutnya adalah pembobotan kata dan pengukuran kemiripan dokumen. (Salim. 2018)

### 2.3.1. Preprocessing

*Pre-processing* dalam proses klarifikasi dokumen digunakan untuk membangun sebuah *index* dari koleksi dokumen *index* dari koleksi dokumen. *Index* adalah himpunan *term* yang menunjukkan isi atau topik yang dikandung oleh dokumen. Pembuatan *inverted index* harus melibatkan konsep *linguistic processing* yang bertujuan meng-ekstrak term-term penting dari dokumen yang dipresentasikan sebagai *bag-of-*

*word*. Tahap *Pre-processing* yang digunakan antara lain :



**Gambar 2.1.** Tahapan Preprocessing

#### 1. *Tokenizing*

Pemisahan rangkaian term (*tokenization*).

*Tokenization* adalah tugas memisahkan deretan kata didalam kalimat, paragraf atau halaman menjadi token atau potongan kata tunggal atau *termmedword*. Tahapan ini juga menghilangkan karakter-karakter tertentu seperti tanda baca dan mengubah semua token ke bentuk huruf kecil (*lower case*).

#### 2. *Filtering/stop word removal*

Tahap filtering adalah tahap pengambilan kata-kata penting dari hasil *tokenizing* menggunakan algoritma *stopword removal*, *Stopword removal* adalah proses penyaringan (*filtering*) terhadap kata-kata yang tidak layak untuk dijadikan sebagai pembeda atau kata kunci sehingga kata-kata tersebut dapat dihilangkan dari dokumen, seperti, kata sambung, kata depan, kata ganti, kata sifat, dan lain sebagainya.

#### 3. *Stemming*

Kata-kata yang muncul didalam dokumen sering mempunyai banyak varian morfologik. Oleh karena itu setelah menjalani proses *tokenizing* dan *stopword removal*, kata-kata yang tersisa menjalani proses stemming. Stemming bertujuan untuk mengubah atau mengembalikan kata mejadi bentuk kata dasarnya (*root word*) dengan menghilangkan imbuhan. Pada penelitian ini menggunakan algoritma Nazief & Adriani yang disesuaikan atau dikembangkan dalam Bahasa Indonesia.

### 2.3.2. Pembobotan TF-IDF

Dalam penelitian ini, digunakan kombinasi pembobotan lokal dan pembobotan global yaitu *Term Frequency – Inverse Dokument Frequency* (TF-IDF).

- a. *Term Frequency* (TF) adalah jumlah kemunculan atau frekuensi ( $f$ ) suatu *term* ( $t$ ) didalam sebuah dokumen  $ij$ .

$$tf_{ij} = f_{ij}$$

**Keterangan :**

**tf** : jumlah kemunculan suatu *term*

**f** : kemunculan *term*.

TF merupakan pendekatan dalam pembobotan lokal yang paling banyak digunakan. Faktor ini menyatakan banyaknya kemunculan suatu *term* dalam sebuah dokumen, berarti semakin penting *term* tersebut.

- b. *Inverse Dokument Frequency* (IDF) adalah logaritma dari jumlah dokumen keseluruhan ( $n$ ) dibagi dengan jumlah dokumen yang memuat *term*  $i$  ( $df_i$ ).

$$idf_i = \log\left(\frac{n}{df_i}\right)$$

**Keterangan :**

**idf** : bobot dari jumlah keseluruhan dokumen

**n** : jumlah dokumen keseluruhan

**df** : dokumen yang memuat *term*.

Logaritma digunakan untuk meredam efek relatif terhadap TF. IDF juga merupakan pembobotan global yang paling banyak digunakan. IDF didasari oleh aspek *term* yang muncul pada banyak dokumen tidak lebih penting daripada *term* yang muncul pada sedikit dokumen.

- c. Pembobotan dengan TF-IDF untuk sebuah *term*  $i$  dan dokumen  $j$  didapatkan dari hasil perkalian nilai  $tf_{ij}$  dan nilai  $idf_i$ .

$$w_{ij} = tf_{ij} \cdot idf_i$$

**Keterangan :**

**w** : bobot *term* dalam sebuah dokumen

**tf** : jumlah kemunculan suatu *term*

**idf** : bobot dari jumlah keseluruhan dokumen.

## 2.4. Term Relatedness to Query

Xu dan Croft (2000) memperkenalkan metode *query expansion* dengan gabungan metode lokal (*relevance feedback*) dan metode global, yang disebut *local context analysis*. Pada metode yang mereka ajukan, pemilihan *term* ekspansi ditingkatkan dengan mempertimbangkan konsep dalam dokumen peringkat atas (*top-rank*) yang

sering memuat banyak *term query* di seluruh koleksi. Dibandingkan dengan *relevance feedback* klasik, kandidat *term* ekspansi lebih relevan dengan *query*, karena *term* tersebut telah diteliti sering muncul pada dokumen.

Pembobotan TRQ memberikan nilai lebih pada *term* yang berhubungan dengan *query*. Rumusan dari TRQ memadukan antara perhitungan IDF (*invers document frequency*) dengan LWF (*lexical word frequency*), dimana formula lwf menggunakan **Persamaan 4** dan TRQ dihitung menggunakan **Persamaan 5**.

### 2.4.1. Lexical Word

*Lexical Word* merupakan penggalan kecil (*small fragment*) dari dokumen yang memuat *term query*. Pada metode ini, ketika pengguna memasukkan *query*, sistem temu kembali informasi mengembalikan set awal dokumen hasil pencarian kemudian meminta pengguna untuk menilai apakah beberapa dokumen relevan atau tidak. Setelah itu, sistem merumuskan *query* berdasarkan penilaian pengguna, dan mengembalikan satu set hasil baru dengan persamaan lwf :

$$lwf_{ti} = \frac{1}{1 + \log\left(\frac{K}{K_i}\right)} \quad (4)$$

Dimana notasi K pada persamaan 4 adalah jumlah *query keyword* dan  $K_i$  adalah jumlah *query keyword* yang berhubungan dengan *lexical word*  $i$ .

### 2.4.2. Pembobotan TRQ

Pembobotan pada *term* didapatkan dengan memadukan antara perhitungan *idf* dengan perhitungan lwf dengan persamaan TRQ :

$$TRQ_{ti} = \alpha \times lwf_{ti} + (1 - \alpha) \times idf_t \quad (5)$$

$$\alpha \in [0, 1]$$

Konstanta  $\alpha$  digunakan untuk mengatur keberimbangan bobot lwf dan idf. Berdasarkan eksperimen yang dilakukan oleh Saneifar (2014) hasil optimal ketika  $\alpha = 0,25$ .

### 2.4.3. Confusion Matrix

Raharjo (2013) menyatakan bahwa dalam “dunia” pengenalan pola (*pattern recognition*) dan temu kembali informasi (*information retrieval*), *precision* dan *recall* adalah dua perhitungan kinerja dari sistem atau metode yang digunakan. *Precision* adalah tingkat ketepatan antara informasi yang

diminta oleh pengguna dengan jawaban yang diberikan oleh sistem. *Recall* adalah tingkat keberhasilan sistem dalam menemukan kembali sebuah informasi. Sedangkan *accuracy* didefinisikan sebagai tingkat kedekatan antara nilai prediksi dengan nilai aktual.

Secara umum *precision*, *recall* dan *accuracy* dapat dirumuskan sebagai berikut :

Gambar 2.2. *precision*, *recall* dan *accuracy*

		Nilai sebenarnya	
		TRUE	FALSE
Nilai prediksi	TRUE	TP (True Positive) <i>Correct result</i>	FP (False Positive) <i>Unexpected result</i>
	FALSE	FN (False Negative) <i>Missing result</i>	TN (True Negative) <i>Correct absence of result</i>

Rumus untuk mengukur *precision*, *recall* dan *accuracy* sebagai berikut:

$$precision = \frac{TP}{TP + FP}$$

$$recall = \frac{TP}{TP + FN}$$

$$accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

Keterangan :

TP : Jumlah data yang terambil oleh sistem dan di anggap relevan.

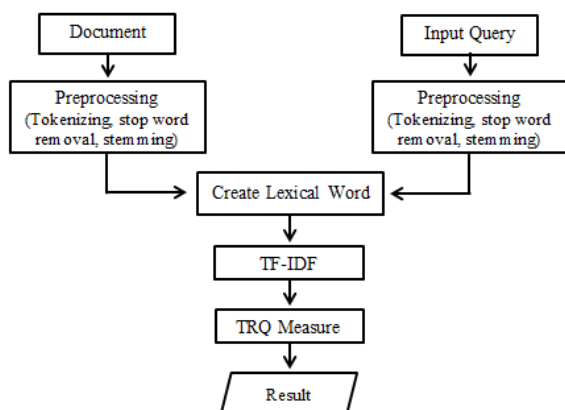
FP : Jumlah data yang terambil oleh sistem dan di anggap tidak relevan.

FN : Jumlah data yang tidak terambil oleh sistem dan di anggap relevan.

TN : Jumlah data yang tidak terambil oleh sistem dan di anggap tidak relevan.

### 3. METODOLOGI PENELITIAN

#### 3.1 Rancangan Sistem



Gambar 3.1. Rancangan Sistem

Tahapan yang ditunjukkan pada rancangan diatas, adalah ketika dilakukan input *query* awal

yang dimasukkan oleh pengguna kemudian dilakukan *preprocessing*. *Preprocessing query* dan dokumen dimulai dengan pemotongan teks dokumen menjadi kumpulan kata (*tokenizing*), menghilangkan kata yang tidak representatif (*stopword removal*), dan menjadikan kata berimbuhan menjadi akar kata (*stemming*). Pada tahap *preprocessing* dokumen, yang di proses adalah *keywordnya*. Kemudian dilakukan pembentukan *Lexical Word*, *lexical word* sendiri merupakan penggalan kecil *term* dari dokumen yang memuat *term query*. Tujuan pembentukan *lexical word* untuk menentukan nilai dari suatu *term* dalam dokumen. Selanjutnya dilakukan TF-IDF *query* dan dokumen yang mengandung *term* dari *query* yang terkumpul dalam *lexical word*. Hasil dari tahap sebelumnya kemudian diperluas lagi menggunakan metode TRQ dengan cara memberikan nilai lebih pada *term* yang berhubungan dengan *query*.

#### 3.2. Skenario Pengujian

Skenario pengujian dilakukan dengan membuat 3 test-case *query* pengguna yang berbeda baik jumlah kata dan jenis data yang diinginkan. Diberikan contoh data perhitungan dalam menghasilkan informasi yang dibutuhkan berdasarkan *query* pengguna sebagai berikut :

##### 3.2.1 Preprocessing Query dan Dokumen.

Q : “BUAH ATAU SAYUR YANG DAPAT MENCEGAH”

- Tokenizing : “buah atau sayur yang dapat mencegah”
- Filtering/Stopword removal: “buah sayur dapat mencegah”
- Stemming : “buah sayur dapat cegah”

##### 3.2.2 Pembentukan Lexical Word

Pembentukan *lexical word* dari *query* dan dokumen digunakan untuk menentukan nilai dari suatu *term*. Sehingga dengan *lexical word* dapat diidentifikasi kumpulan *term* yang muncul di sekitar suatu *term* dalam dokumen.

Tabel 3.1. Pembentukan Lexical Word

No	Term	Q	D1	D2	D3
1	buah	1	4	3	0
2	sayur	1	0	0	3
3	dapat	1	0	0	0
4	cegah	1	1	1	1



### 3.2.3 Perhitungan TF-IDF

Pada tahap ini, dilakukan perhitungan untuk tiap *term* yang berada dalam kumpulan *lexical word*.

Perhitungan IDF menggunakan persamaan :

$$idf_i = \log\left(\frac{n}{df_i}\right) + 1$$

$$idf_i = \log\left(\frac{4}{3}\right) + 1 = 1,12494$$

dimana  $n$  bernilai 4 yang merupakan jumlah dari keseluruhan dokumen dan *query* yang terlibat, sedangkan  $df_i$  bernilai 3 merupakan jumlah dokumen yang mengandung *term* tersebut. Didapatkan hasil :

**Tabel 3.2.** Perhitungan TF-IDF

No	Term	Q	TF				IDF	TF-IDF			
			D1	D2	D3	DF		Q	D1	D2	D3
1	buah	1	4	3	0	3	1,12494	1,12494	4,49975	3,37482	0
2	sayur	1	0	0	3	2	1,3010	1,3010	0,0000	0,0000	3,90309
3	dapat	1	0	0	0	1	1,6021	1,6021	0	0	0
4	cegah	1	1	1	1	4	1	1	1	1	1

#### Keterangan :

**Q:** Jumlah kemunculan *term* dalam query

**D1:** Jumlah kemunculan *term* dalam dokumen 1

**D2:** Jumlah kemunculan *term* dalam dokumen 2

**D3:** Jumlah kemunculan *term* dalam dokumen 3

**DF:** Jumlah keseluruhan dokumen yang memuat *term*

**IDF** : bobot dari jumlah keseluruhan dokumen dibagi jumlah dokumen yang memuat *term*.

### 3.2.4 Pembobotan TRQ

Pembobotan TRQ memberikan nilai lebih pada *term* yang berhubungan dengan *query*. Rumusan dari TRQ dengan memadukan antara perhitungan IDF dan LWF. Maka sebelum memasuki tahap pembobotan TRQ, dilakukan perhitungan LWF untuk setiap *term* dalam kumpulan *lexical word*. Perhitungan LWF pada tabel 3.3 menggunakan persamaan :

$$lwf_{ti} = \frac{1}{1 + \log\left(\frac{K}{K_i}\right)}$$

$$lwf_{ti} = \frac{1}{1 + \log\left(\frac{4}{1}\right)} = 0,6242$$

dimana  $K$  bernilai 4 yang merupakan jumlah dari *query* dan  $K_i$  bernilai 1 adalah jumlah *query* yang berhubungan dengan *lexical word*  $i$ . Didapatkan hasil :

**Tabel 3.3.** Perhitungan LWF

No	Term	LWF			
		Q	D1	D2	D3
1	buah	0,6242	1,0000	0,8889	1,0000
2	sayur	0,6242	1,0000	1,0000	0,8889
3	dapat	0,6242	1,0000	1,0000	1,0000
4	cegah	0,6242	0,6242	0,6242	0,6242

Setelah ditemukan nilai dari LWF pada tabel diatas, selanjutnya dilakukan pembobotan TRQ menggunakan persamaan :

$$TRQ_{ti} = \alpha \times lwf_{ti} + (1 - \alpha) \times idf_t$$

$$TRQ_{ti} = 0,25 \times 0,6242 + (1 - 0,25) \times 1,12494 = 0,9998$$

dimana  $\alpha$  dengan nilai optimal 0,25,  $lwf_{ti}$  bernilai 0,6242 dan  $idf_i$  bernilai 1,12494. Didapatkan hasil :

**Tabel 3.4.** Pembobotan TRQ

No	Term	TRQ			
		Q	D1	D2	D3
1	buah	0,9998	1,0937	1,0659	1,0937
2	sayur	1,1318	1,2258	1,2258	1,1980
3	dapat	1,3576	1,4515	1,4515	1,4515
4	cegah	0,9060	0,9060	0,9060	0,9060
AVERAGE		1,0988	1,1693	1,1623	1,1623

Setelah didapatkan bobot tertinggi berdasarkan *query* awal yang dimasukkan ke sistem yaitu “Buah atau sayur yang dapat mencegah”, maka sistem akan menghasilkan beberapa saran pencarian. Yang mana saran rekomendasi *query* didapatkan dari *keyword* didalam database ensiklopedia.

Dari proses ekspansi masing-masing dokumen tersebut akan menghasilkan rekomendasi *query* yang berbeda-beda, rekomendasi dengan bobot tertinggi akan menjadi rekomendasi *query* teratas, sehingga rata-rata tertinggi pada tabel 3.4 yaitu D1, D2 dan D3. Hasil dari ekspansi *query* sebagai berikut :

**Tabel 3.5.** Contoh hasil ekspansi query dengan metode TRQ

D	Q	Suggestions/Rekomendasi
D1	Q1	Buah atau sayur yang dapat mencegah <b>osteoporosis</b>
D2	Q2	Buah atau sayur yang dapat mencegah <b>penuaan dini</b>
D3	Q3	Buah atau sayur yang dapat mencegah <b>diabetes</b>

### DAFTAR PUSTAKA

1. Boston, C., Fang, H., Carberry, S., Wu, H., & Liu, X. 2014. Wikimantic: Toward effective disambiguation and expansion of queries. *Data & Knowledge Engineering*, 90, 22-37.
2. Raharjo, B. 2013. *Perbedaan Recall, Precision dan Accuracy*.
3. Ludviani, R., Hayati, K. F. 2015. Optimasi Pembobotan pada Query Expansion dengan Term Relatedness to query – entropy based (TRQE).

4. Salim, A. 2018. Pencarian Link Informasi Pada Aplikasi Ensiklopedia Buah dan Sayuran Lokal Dengan Metode Cosine Similarity. Universitas Muhammadiyah Jember.
5. Santoso, A. 2011. Serat Pangan (Dietary Fiber) dan Manfaatnya Bagi Kesehatan.
6. Sawitri, K. 2017. Ensiklopedia Buah-buahan Lokal Jember Berbasis Potensi Alam Jember. p-ISSN 2527-7111; e-ISSN 2528-1615.
7. Sawitri, K., Herrianto, E., Suharso, W. 2018. Buah dan Sayur Lokal di Kabupaten Jember. Jember : Universitas Muhammadiyah Jember.
8. Suharso, W. 2018. Bisnis Buah dan Sayuran Lokal di Kabupaten Jember sebagai Dukungan Ketahanan Pangan, Indonesia
9. Wahyudi, D. 2017. Implementasi dan Analisis Algoritma Stemming Nazief & Adriani dan Porter pada Dokumen Berbahasa Indonesia. Jurnal Ilmiah Sinus 15 (2), 49-56.
10. Xu, J., & Croft, W. B. 2000. Improving the effectiveness of information retrieval with local context analysis. ACM Transactions on Information Systems (TOIS), 18(1), 79-112.

