

Optimalisasi Fitur Pencarian Pada E-Catalog Menggunakan *Query Expansion* dan Algoritma TF-IDF

Optimization Of Search Features In E-Catalog Using Query Expansion and TF IDF Algorithm

Selvia Ferdiana Kusuma¹, Tessy Badriyah², Prasetyo Wibowo³,
Rosiyah Faradisa⁴, Solichul Huda⁵

^{1,2,3,4}Teknik Informatika dan Komputer/Teknik Informatika,
Politeknik Elektronika Negeri Surabaya

⁵Teknik Informatika, Universitas Dian Nuswantoro

E-mail: ¹selvia@pens.ac.id, ²tessy@pens.ac.id, ³prasetyo@pens.ac.id,

⁴rosiyah@pens.ac.id, ⁵ solichul.huda@dsn.dinus.ac.id

Abstrak

E-catalog adalah sebuah katalog elektronik dapat digunakan untuk mempublikasikan dan mempromosikan sebuah produk atau layanan secara online kepada berbagai pihak. Pengguna dapat memperoleh informasi terkait produk yang diinginkan melalui fitur pencarian pada e-catalog tersebut. Oleh sebab itu fitur pencarian memiliki peran yang signifikan untuk menunjang performa e-catalog. Proses pencarian produk pada sebuah e-catalog dilakukan berdasarkan kesamaan kata yang dimasukkan oleh pengguna dengan judul produk yang tersedia. Biasanya semua judul yang mengandung kata kunci yang dimasukkan pengguna akan ditampilkan tanpa adanya pengurutan kedekatan hasil pencarian. Hal tersebut tentunya menyulitkan pengguna untuk menemukan produk yang sesuai dengan keinginan. Oleh sebab itu penelitian ini bertujuan untuk melakukan optimalisasi fitur pencarian produk yang ada pada e-catalog menggunakan *query expansion* dan algoritma TF-IDF. Berdasarkan hasil uji coba yang telah dilakukan, terbukti bahwa penambahan *query expansion* dan algoritma TF-IDF dapat mengoptimalkan kinerja fitur pencarian pada e-catalog tersebut. Hasil pencarian produk dapat terurut berdasarkan kedekatan kata kunci yang dimasukkan oleh pengguna dengan judul produk yang diinginkan oleh pengguna.

Kata kunci: e-catalog, fitur pencarian, query expansion, TF-IDF

Abstract

E-Catalog is an electronic catalog that can be used to publish and promote a product or service online to various parties. Users can obtain information related to the desired product through the search feature on the e-catalog. Therefore, the search feature has a significant role to support the performance of the e-catalog. The process of searching for products in an e-catalog is carried out based on the similarity of words entered by users with the available product titles. Usually, all titles containing keywords entered by users will be displayed without ranking the similarity of search results. This certainly makes it difficult for users to find products that match with their need. Therefore, this study aims to optimize the product search feature in the e-catalog using the query expansion and TF-IDF algorithm. Based on the results of experiments that have been carried out, it is proven that query expansion and the TF-IDF algorithm can optimize the performance of the search feature on the e-catalog. Product search results can be ranked based on the similarity of keywords entered by users with the available product titles.

Keywords: e-catalog, query expansion, search features, TF-IDF

1. PENDAHULUAN

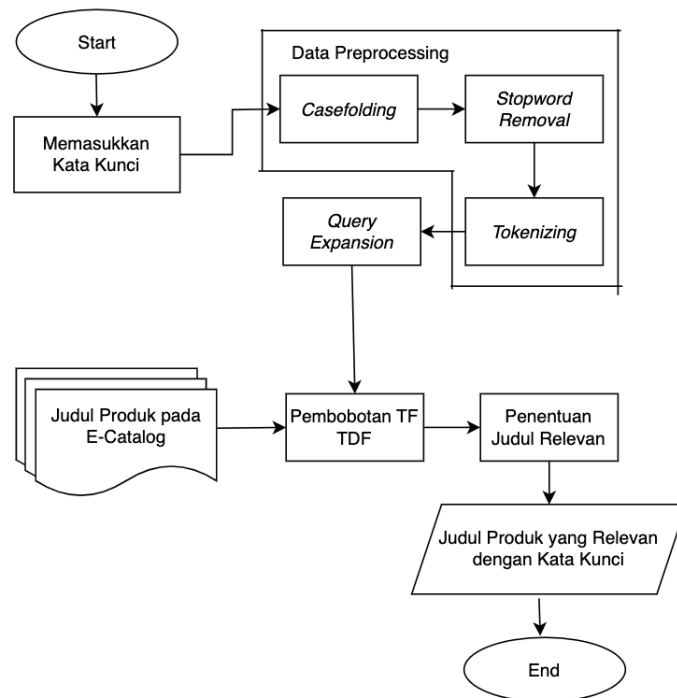
E-catalog adalah sebuah katalog elektronik dapat digunakan untuk mempublikasikan dan mempromosikan sebuah produk atau layanan secara *online* kepada berbagai pihak. Berdasarkan studi literatur yang telah dilakukan, diketahui bahwa tren penggunaan e-catalog sebagai media publikasi maupun promosi semakin meningkat dari tahun ke tahun. Pada tahun 2017 ada pembuatan e-catalog untuk CV Eklesia [1], kemudian di tahun 2018 ada pembuatan e-catalog untuk promosi produk batik [2], di tahun 2019 ada pembuatan e-catalog untuk produk kopi sapit [3], di tahun 2020 ada pembuatan e-catalog untuk produk di TPK-Koja [4], kemudian di tahun 2021 ada pembuatan e-catalog untuk mempromosikan produk khas dari Jombang [5] dan e-catalog untuk mempromosikan tempat wisata di Mamminasata [6]. Puncaknya adalah di tahun 2022 yang terjadi peningkatan jumlah pembuatan produk e-catalog yang digunakan untuk mempromosikan hp [7], perhiasan [8] dan UMKM [9] [10] [11]. Berdasarkan penelitian-penelitian tersebut, diketahui bahwa e-catalog efektif digunakan sebagai media publikasi maupun promosi dalam berbagai aspek.

Salah satu hal yang dapat menunjang performa e-catalog adalah adanya fitur pencarian. Fitur pencarian dapat mempermudah pengguna untuk mencari produk yang diinginkan. Namun, saat ini yang terjadi adalah banyak website yang menggunakan fitur pencarian yang belum optimal. Fitur pencarian yang ada melakukan proses pencarian produk/jasa berdasarkan adanya kesamaan kata kunci yang dimasukkan oleh pengguna dengan setiap kata yang ada di judul produk. Hal tersebut membawa ketidakefisienan karena tentunya akan ada banyak produk yang berhasil di filter jika hanya menggunakan adanya kesamaan kata kunci pada judul yang ada. Oleh sebab itu perlu adanya optimalisasi fitur pencarian yang ada agar pencarian tersebut menghasilkan produk yang spesifik dengan kata kunci yang dimasukkan oleh pengguna.

Oleh sebab itu pada penelitian ini menggabungkan *query expansion* dan algoritma *Term Frequency — Inverse Document Frequency* (TF-IDF) untuk meningkatkan akurasi proses pencarian produk pada fitur *search* disebuah e-catalog. Pemilihan metode tersebut juga didasari dari penelitian terdahulu yang juga memanfaatkan TF-IDF sebagai metode perengkingan pada suatu proses pencarian namun pada studi kasus yang berbeda yakni pencarian judul skripsi [12] dan pencarian dokumen arsip [13]. Berbeda dari penelitian sebelumnya yang hanya menggunakan TF-IDF, pada penelitian ini penulis juga mengoptimalkan kinerja algoritma TF-IDF yang telah ada dengan menggunakan *query expansion*. Penggunaan *query expansion* menghasilkan kata kunci baru yang berasal dari kedekatan kata dari kata kunci yang dimasukkan oleh pengguna [14][15]. Kata kunci baru tersebut digunakan untuk menunjang kinerja algoritma TF-IDF dalam melakukan perengkingan.

2. METODE PENELITIAN

Metoda penelitian yang dilakukan pada proses optimalisasi fitur pencarian menggunakan *query expansion* dan algoritma TF-IDF ini memiliki 7 tahapan. Tahapan proses optimalisasi hasil pencarian fitur *search* dijelaskan pada Gambar 1. Kata kunci yang dimasukkan oleh pengguna akan masuk pada fase preproses. Pada fase tersebut terdapat proses *case folding*, *stopword removal*, dan *tokenizing*. Kemudian hasil preproses akan masuk pada proses *query expansion*. Hasil kata kunci dari proses *query expansion* dan kata kunci yang aslinya kemudian diberi pembobotan sesuai dengan kedekatannya dengan judul produk pada e-catalog. Proses pembobotan menggunakan algoritma TF-IDF. Hasil dari proses tersebut kemudian dijadikan dasar untuk proses penentuan judul yang relevan dengan kata kunci. Hasil akhir dari proses ini adalah judul-judul produk yang paling relevan dengan kata kunci yang dimasukkan oleh pengguna. Detail dari setiap proses dijelaskan pada subbab 2.1 sampai dengan sub bab 2.5



Gambar 1. Tahapan optimalisasi fitur pencarian

2.1 Preproses

Preproses pada penelitian ini bertujuan untuk menyiapkan data yang akan diproses agar dapat dipergunakan pada proses selanjutnya. Preproses yang dilakukan meliputi *casefolding*, *stopword removal*, dan tokenisasi. Contoh kalimat yang dimasukkan oleh pengguna sebagai kata kunci adalah "Sistem informasi antrian pada klinik ". Kalimat tersebut kemudian masuk pada tahap preproses. Tahapan preproses yang pertama adalah *case folding*. *Case folding* adalah proses mengubah semua huruf dalam teks menjadi huruf kecil (*lowercase*). Tujuan dari case folding adalah untuk mengatasi perbedaan huruf kapital dan huruf kecil dalam teks yang dapat menyebabkan kesalahan pemrosesan atau pencocokan kata. Contoh kalimat yang telah mengalami *case folding* adalah "sistem informasi antrian pada klinik". Setelah melalui tahap *case folding* kemudian masuk pada tahap *stopword removal*. *Stopword removal* adalah proses menghilangkan kata-kata umum yang sering muncul dalam teks namun tidak memiliki nilai informasi yang signifikan. Tujuan utama dari *stopword removal* adalah untuk membersihkan teks dari kata-kata yang tidak memberikan nilai informasi yang berguna, sehingga meningkatkan efisiensi dan efektivitas analisis teks. Contoh kalimat yang telah mengalami *stopword removal* adalah "sistem informasi antrian klinik". Kata "pada" di kalimat tersebut dihilangkan karena tergolong dalam kata-kata yang tidak memiliki informasi yang significant terkait konten kalimat. Hasil proses *stopword removal* kemudian masuk ke proses tokenisasi. Tokenisasi adalah proses membagi teks menjadi unit-unit yang lebih kecil, yang disebut token. Token dapat berupa kata, frasa, simbol, atau bahkan karakter, tergantung pada tingkat granularitas yang diinginkan dalam analisis teks. Tujuan dari tokenisasi adalah untuk memecah teks menjadi unit-unit yang lebih terstruktur agar dapat diolah lebih lanjut dalam analisis teks. Dalam bahasa alami, teks biasanya tidak berbentuk struktur yang terorganisir, dan tokenisasi membantu dalam memperoleh unit-unit artikulasi yang lebih bermakna. Hasil tokenisasi yang terbentuk dari kalimat yang dijadikan masukan adalah ("sistem", "informasi", "antrian", "klinik". Hasil tokenisasi tersebut kemudian masuk ke tahap *query expansion* yang dijelaskan lebih detail pada sub bab selanjutnya.

2.2 Query Expansion

Query expansion adalah teknik yang digunakan dalam sistem pencarian informasi untuk meningkatkan relevansi dan kualitas hasil pencarian. Tujuan utama dari query expansion adalah memperluas atau memperkaya query yang diajukan oleh pengguna dengan kata kunci atau frasa tambahan yang relevan. Dalam query expansion, kata-kata atau frasa tambahan ditambahkan ke query asli untuk memperluas cakupan pencarian dan mencakup kemungkinan variasi kata kunci yang relevan dengan topik yang dicari. Hal ini dapat membantu mengatasi masalah ambiguitas, sinonim, polisemi, dan variasi istilah yang dapat terjadi dalam pencarian. Teknik *query expansion* yang dipergunakan pada penelitian ini adalah pemanfaatan tesaurus. Tesaurus disini berisi hubungan sinonim, antonim, atau hubungan semantik lainnya dapat membantu dalam memperkaya query dengan kata-kata yang serumpun atau relevan. Contoh tesaurus yang dipergunakan pada penelitian ini ditunjukkan pada Tabel 1. Contoh hasil *query expansion* dari kalimat "sistem informasi pasien klinik" adalah "aplikasi pasien klinik" dan "sistem informasi pasien rumah sakit". Kedua kalimat tersebut dianggap kalimat yang sama dengan kata kunci yang dimasukkan oleh pengguna. Selanjutnya kalimat hasil *query expansion* masuk pada proses pembobotan TF-IDF.

Tabel 1 Tesaurus

| Kata | Padanan Kata |
|-------------|--------------------------|
| Aplikasi | Sistem, sistem informasi |
| Klinik | Rumah sakit |
| Klasifikasi | Kategori |

2.3 Pembobotan Menggunakan Algoritma TF-IDF

Algoritma TF-IDF (Term Frequency-Inverse Document Frequency) adalah metode yang digunakan dalam pemrosesan bahasa alami dan sistem pencarian informasi untuk mengukur pentingnya sebuah kata dalam suatu kalimat dalam konteks koleksi kalimat yang lebih besar. Algoritma ini bertujuan untuk mengidentifikasi kata-kata kunci yang paling relevan dalam kalimat. Perhitungan pada algoritma TF-IDF ditunjukkan pada persamaan 1.

$$W_{i,j} = \frac{n_{i,j}}{\sum_{j=1}^p n_{i,j}} \log_2 \frac{D}{d_j} \quad (1)$$

Dimana:

$W_{i,j}$ = pembobotan tf-idf untuk term ke-j pada kalimat ke-i

$n_{i,j}$ = jumlah kemunculan term ke-j pada kalimat ke-i

p = banyaknya term yang terbentuk

$\sum_{j=1}^p n_{i,j}$ = jumlah kemunculan seluruh term pada kalimat ke-i

D = jumlah keseluruhan kalimat

d_j = banyaknya kalimat yang mengandung term ke-j.

Gambar 2 merupakan contoh perhitungan manual kesamaan judul produk pada e-catalog dengan kata kunci yang dimasukkan pengguna. Kata kunci yang digunakan adalah “sistem informasi pasien klinik” sedangkan judul produk yang tertera pada e-catalog dicontohkan ada tiga yaitu “sistem informasi apotek”, “sistem informasi rumah sakit di tulungagung” dan “aplikasi rekam medis pasien klinik“. Proses yang dilakukan pada perhitungan tersebut mulai dari tokenisasi kata pada kata kunci dan judul produk. Kemudian menghitung frekwensi setiap kata di setiap judul (tf), menghitung frekwensi kata di seluruh kalimat (df), menghitung frekwensi kemunculan kata terhadap keseluruhan judul produk (d/df), menghitung inverse dari frekwensi dokumen (idf), lalu dilanjutkan menghitung (tf-idf) dari setiap judul produk. Hasil akhir kemiripan kata kunci dan judul didapatkan dengan cara menjumlah nilai yang mirip dengan kata kunci. Pada Gambar 2 dicontohnya untuk menjumlahkan nilai di kolom (tf-idf) dari setiap dokumen namun hanya yang berwarna coklat muda.

| kata kunci | sistem informasi pasien klinik | | | | | | | | | | |
|----------------|---|----|----|----|----|------|-------------|-------------|-------------|-------------|--|
| produk 1 (p1) | sistem informasi apotek | | | | | | | | | | |
| produk 2 (p2) | sistem informasi rumah sakit di tulungagung | | | | | | | | | | |
| produk 3 (p3) | aplikasi rekam medis pasien klinik | | | | | | | | | | |
| jumlah dokumen | 3 | | | | | | | | | | |
| token | kk | tf | | | df | d/df | idf | tf-idf | | | |
| | | p1 | p2 | p3 | | | | p1 | p2 | p3 | |
| sistem | 1 | 1 | 1 | 0 | 2 | 1,5 | 0,176091259 | 0,058697086 | 0,029348543 | 0 | |
| informasi | 1 | 1 | 1 | 0 | 2 | 1,5 | 0,176091259 | 0,058697086 | 0,029348543 | 0 | |
| apotek | 0 | 1 | 0 | 0 | 1 | 3 | 0,477121255 | 0,159040418 | 0 | 0 | |
| aplikasi | 0 | 0 | 0 | 1 | 1 | 3 | 0,477121255 | 0 | 0 | 0,095424251 | |
| rumah | 0 | 0 | 1 | 0 | 1 | 3 | 0,477121255 | 0 | 0,079520209 | 0 | |
| sakit | 0 | 0 | 1 | 0 | 1 | 3 | 0,477121255 | 0 | 0,079520209 | 0 | |
| di | 0 | 0 | 1 | 0 | 1 | 3 | 0,477121255 | 0 | 0,079520209 | 0 | |
| tulungagung | 0 | 0 | 1 | 0 | 1 | 3 | 0,477121255 | 0 | 0,079520209 | 0 | |
| rekam | 0 | 0 | 0 | 1 | 1 | 3 | 0,477121255 | 0 | 0 | 0,095424251 | |
| medis | 0 | 0 | 0 | 1 | 1 | 3 | 0,477121255 | 0 | 0 | 0,095424251 | |
| pasien | 1 | 0 | 0 | 1 | 1 | 3 | 0,477121255 | 0 | 0 | 0,095424251 | |
| klinik | 1 | 0 | 0 | 1 | 1 | 3 | 0,477121255 | 0 | 0 | 0,095424251 | |
| hasil ranking | | | | | | | | 0,117394173 | 0,058697086 | 0,190848502 | |

Gambar 2. Perhitungan manual TF-IDF tanpa *query expansion*

Gambar 3 merupakan perhitungan TF-IDF dari kata kunci yang berasal dari *query expansion*. Cara perhitungan TF-IDF yang dilakukan tetap sama dari perhitungan sebelumnya namun hasilnya berubah menyesuaikan kata kunci yang digunakan. Hasil peringkat kemiripan judul pada Gambar 3 didapatkan dari penjumlahan nilai (tf-idf) pada token "aplikasi rumah sakit pasien".

| kata kunci | sistem informasi pasien klinik | | | | | | | | | | |
|-----------------|---|----|----|----|----|------|------------|------------|------------|------------|--|
| query expansion | aplikasi pasien rumah sakit | | | | | | | | | | |
| produk 1 (p1) | sistem informasi apotek | | | | | | | | | | |
| produk 2 (p2) | sistem informasi rumah sakit di tulungagung | | | | | | | | | | |
| produk 3 (p3) | aplikasi rekam medis pasien klinik | | | | | | | | | | |
| jumlah dokumen | 3 | | | | | | | | | | |
| token | qe | tf | | | df | d/df | idf | tf-idf | | | |
| | | p1 | p2 | p3 | | | | p1 | p2 | p3 | |
| sistem | 0 | 1 | 1 | 0 | 2 | 1,5 | 0,17609126 | 0,05869709 | 0,02934854 | 0 | |
| informasi | 0 | 1 | 1 | 0 | 2 | 1,5 | 0,17609126 | 0,05869709 | 0,02934854 | 0 | |
| apotek | 0 | 1 | 0 | 0 | 1 | 3 | 0,47712125 | 0,15904042 | 0 | 0 | |
| aplikasi | 1 | 0 | 0 | 1 | 1 | 3 | 0,47712125 | 0 | 0 | 0,09542425 | |
| rumah | 1 | 0 | 1 | 0 | 1 | 3 | 0,47712125 | 0 | 0,07952021 | 0 | |
| sakit | 1 | 0 | 1 | 0 | 1 | 3 | 0,47712125 | 0 | 0,07952021 | 0 | |
| di | 0 | 0 | 1 | 0 | 1 | 3 | 0,47712125 | 0 | 0,07952021 | 0 | |
| tulungagung | 0 | 0 | 1 | 0 | 1 | 3 | 0,47712125 | 0 | 0,07952021 | 0 | |
| rekam | 0 | 0 | 0 | 1 | 1 | 3 | 0,47712125 | 0 | 0 | 0,09542425 | |
| medis | 0 | 0 | 0 | 1 | 1 | 3 | 0,47712125 | 0 | 0 | 0,09542425 | |
| pasien | 1 | 0 | 0 | 1 | 1 | 3 | 0,47712125 | 0 | 0 | 0,09542425 | |
| klinik | 0 | 0 | 0 | 1 | 1 | 3 | 0,47712125 | 0 | 0 | 0,09542425 | |
| hasil ranking | | | | | | | | 0 | 0,15904042 | 0,1908485 | |

Gambar 3. Perhitungan manual TF-IDF dengan *query expansion*

2.4 Penentuan Judul Relevan

Berdasarkan perhitungan TF-IDF yang telah dilakukan pada Gambar 2 dan Gambar 3, diketahui bahwa judul produk yang paling relevan dengan kata kunci adalah produk ke 3. Judul produk yang memiliki nilai TF-IDF paling tinggi dianggap sebagai judul produk yang paling relevan dengan kata kunci yang dimasukkan pengguna. Pada contoh tersebut, hasil perhitungan TF-IDF menggunakan kata kunci (Gambar 2) dan menggunakan hasil *query expansion* (Gambar

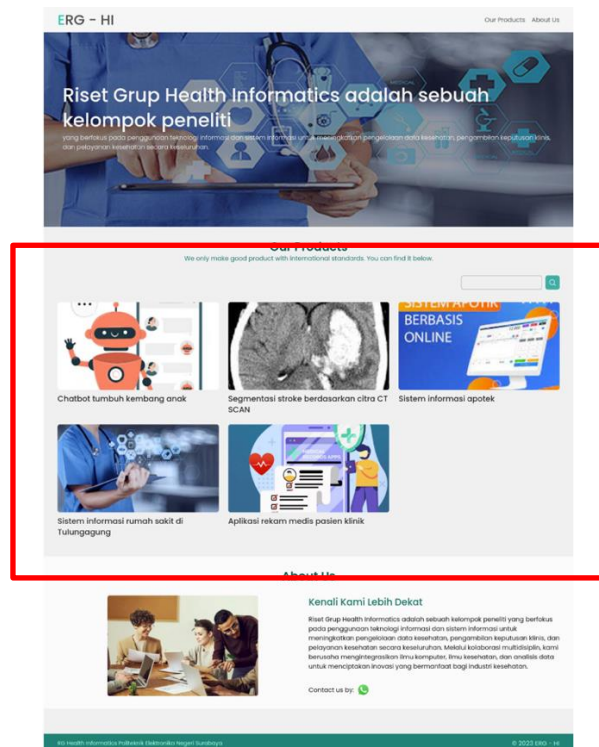
3) menunjukkan kesamaan hasil akhir namun memiliki perbedaan urutan kemiripan judul. Apabila terjadi yang demikian maka urutan didasarkan pada hasil penjumlahan nilai TF-IDF di setiap judul produk. Namun jika ada judul produk yang bernilai 0 maka akan langsung diabaikan karena dianggap tidak memiliki kesamaan sama sekali dengan kata kunci yang dimasukkan oleh pengguna.

2.5 Evaluasi

Proses evaluasi diperlukan untuk menilai apakah sistem yang dikembangkan sudah berjalan seperti apa yang telah direncanakan. Selain itu proses evaluasi juga diperlukan untuk menilai keberhasilan metode yang diusulkan dalam menyelesaikan sebuah permasalahan yang diangkat. Evaluasi yang dilakukan pada penelitian ini yaitu dengan cara menyamakan hasil pencarian judul produk berdasarkan perhitungan manual dan hasil pencarian judul produk yang dilakukan otomatis oleh sistem. Selain itu penulis juga melakukan evaluasi terhadap efektifitas penggunaan *query expansion* dan pembobotan TF-IDF untuk menunjang fitur pencarian pada sebuah e-catalog. Efektifitas penggunaan *query expansion* dan pembobotan TF-IDF dapat terlihat jika pencarian produk bisa menghasilkan produk yang spesifik dan relevan dengan kata kunci yang dimasukkan oleh pengguna.

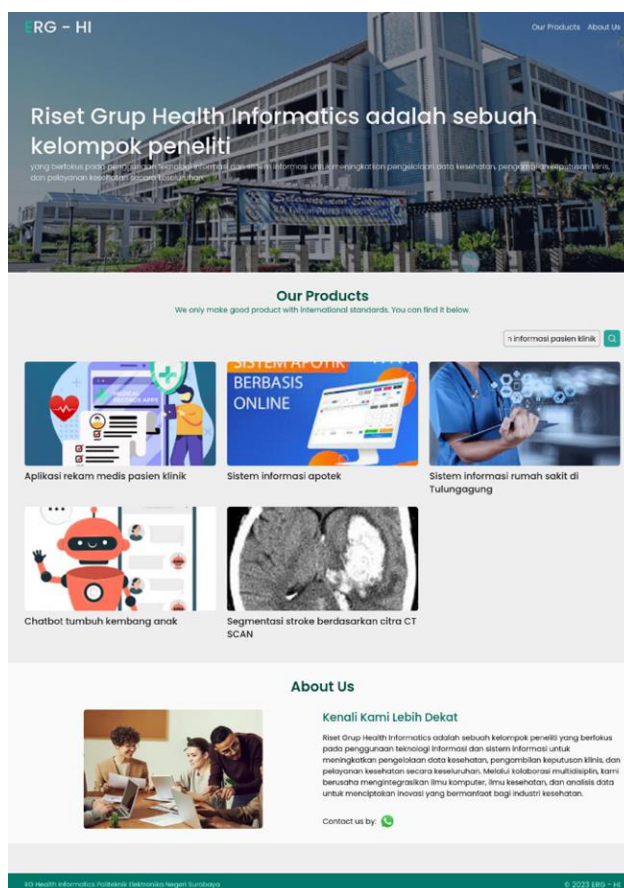
3. HASIL DAN PEMBAHASAN

Penelitian ini menghasilkan sebuah e-catalog yang mengoptimalkan fitur pencarian menggunakan *query expansion* dan algoritma TF-IDF. Seluruh tahapan pengembangan fitur pencarian yang dilengkapi dengan *query expansion* dan TF-IDF telah berhasil diimplementasikan. Hasil implementasi tersebut kemudian dievaluasi untuk menilai keberhasilan metode yang diusulkan untuk mengoptimalkan fitur pencarian. Proses evaluasi yang dilakukan pada penelitian ini terdiri dari 2 bagian yaitu evaluasi terhadap kesesuaian perhitungan sistem dengan perhitungan manual dan evaluasi terhadap efektifitas penggunaan *query expansion* dan pembobotan TF-IDF untuk menunjang fitur pencarian pada sebuah e-catalog. Gambar 3 menunjukkan produk-produk yang ada pada e-catalog.



Gambar 3 Produk pada e-catalog

Berdasarkan percobaan yang telah dilaksanakan, diketahui bahwa sistem yang dibangun telah dapat melakukan filterisasi produk sesuai perhitungan TF-IDF manual seperti yang ditunjukkan pada sub bab 2.3. Hal tersebut membuktikan bahwa sistem yang dibangun telah mengimplementasikan algoritma TF-IDF dengan benar. Hasil pencarian pada sistem ditunjukkan pada Gambar 4. Gambar 4 menunjukkan hasil pencarian dari kata kunci “Sistem Informasi pasien klinik”. Hasil pencarian menunjukkan bahwa produk berjudul “Aplikasi rekam medis pasien” berada di urutan pertama hasil pencarian. Hasil tersebut menunjukkan adanya kesamaan hasil pencarian dengan hasil perhitungan manual yang telah dilakukan sebelumnya.



Gambar 4 Hasil pencarian menggunakan optimasi "fitur pencarian"

Hasil efektifitas penggunaan *query expansion* dan pembobotan TF-IDF menunjukkan hasil yang positif. Proses pencarian menghasilkan produk yang lebih spesifik dibandingkan tanpa menggunakan *query expansion* dan pembobotan TF-IDF. Hal tersebut juga terbukti dari hasil percobaan pada Gambar 5. Gambar 5 merupakan contoh hasil pencarian tanpa penggunaan *query expansion* dan TF-IDF. Kata kunci yang digunakan sebagai *input* adalah "sistem informasi pasien klinik". Pencarian yang dilakukan menghasilkan judul yang sangat bervariasi. Sistem mengambil semua judul yang memiliki kata atau kombinasi kata yang sama dengan kata kunci. Pada contoh di Gambar 5 semua judul yang memiliki kata "sistem informasi" akan ditampilkan walaupun secara khusus tidak berelasi dengan kata kunci yang dimasukkan oleh pengguna. Misalkan "sistem informasi kafe", "sistem informasi koperasi", dan "sistem informasi antrian kasir". Hal tersebut tentunya tidak efisien dan akan menyulitkan jika ada banyak produk yang memiliki kata kunci sama. Penggunaan *query expansion* dan TF-IDF dapat meminimalisir ketidakefisienan tersebut dengan cara menghitung kesamaan/kemiripan kata kunci dengan judul produk yang ada dan hanya menampilkan beberapa judul produk yang memiliki tingkat kemiripan yang tinggi dengan kata kunci yang dimasukkan oleh pengguna.

```
input: sistem informasi pasien klinik

output:
sistem informasi kafe
sistem informasi koperasi
sistem informasi rumah sakit
sistem informasi antrian kasir
aplikasi rekam medis pasien klinik
sistem informasi klinik
```

Gambar 5. Hasil pencarian tanpa penggunaan *query expansion* dan TF-IDF

4. KESIMPULAN DAN SARAN

E-catalog adalah sebuah katalog elektronik yang dapat digunakan untuk mempublikasikan dan mempromosikan sebuah produk atau layanan secara online kepada berbagai pihak. Pengguna dapat memperoleh informasi terkait produk yang diinginkan melalui fitur pencarian pada e-catalog. Oleh sebab itu fitur pencarian memiliki peran penting dalam sebuah e-catalog. Penelitian ini mencoba untuk mengoptimalkan fitur pencarian pada sebuah e-catalog menggunakan *query expansion* dan algoritma TF-IDF. *Query expansion* digunakan sebagai metode untuk memperluas cakupan kata kunci yang dimasukkan oleh pengguna sedangkan algoritma TF-IDF digunakan sebagai metode untuk melakukan pengurutan dan penentuan judul produk yang paling relevan dengan kata kunci yang dimasukkan oleh pengguna. Berdasarkan eksperimen yang telah dilakukan terbukti bahwa penggunaan *query expansion* dan algoritma TF-IDF dapat memperbaiki kinerja pencarian dengan menghasilkan hasil pencarian yang lebih spesifik dan terurut sesuai kedekatan kata yang dimasukkan oleh pengguna dengan judul yang ada pada setiap produk e-catalog. Jika dibandingkan dengan penelitian sebelumnya yang belum menggunakan optimasi fitur pencarian, penggunaan *query expansion* dan TF-IDF dinilai lebih efektif untuk menghasilkan hasil pencarian yang lebih relevan. Hasil penelitian ini dapat digunakan oleh seluruh pengguna e-catalog / seluruh pengembang website untuk mengoptimalkan fitur pencarian pada website yang telah dikembangkan.

DAFTAR PUSTAKA

- [1] Hendra, "Perancangan Sistem Informasi Produk Katalog Berbasis Web," *J. Ilm. Core IT*, vol. 9, no. 1, pp. 17–23, 2017.
- [2] D. M. A. Jatmika, "Use of Technology Appropriate to Batik 's Small Medium Enterprises Up Class Through E-Catalog," *IJEBA*, vol. 8, no. 3, pp. 61–65, 2000.
- [3] D. Yulistina and B. D. D. Arianti, "E-Katalog Sebagai Sistem Informasi Pemasaran Kopi Sapit Berbasis Web," *EDIMATIC J. Pendidik. Informatika*, vol. 3, no. 2, pp. 45–52, 2019, doi: 10.29408/edumatic.v3i2.1766.
- [4] E. Y. T. Safuan; Budiandru; Arief, Rudi Kurniawan; Adesta, "Web Based E-Catalog Implementation at TPK-KOJA : Case Study of Stock Inventory Division," *Int. J. Adv. Sci. Technol.*, vol. 29, no. 7, pp. 481–488, 2020.
- [5] N. Khafidhoh and V. F. S. Ningrum, "Design and Build an E-Catalog Application for Typical Jombang Products," in *Multidisiplin - International Conference 2021*, 2021, pp. 79–84.
- [6] E. S. Rahman and A. NFH, "PERANCANGAN E-KATALOG OBJEK WISATA MAMMINASATA MAMMINASATA TOURISM OBJECT E-CATALOG DESIGN," *Media Elektr.*, vol. 18, no. 2, pp. 43–48, 2021.
- [7] J. M. Butarbutar, R. Nisa, and S. Amriza, "Perancangan Sistem Informasi E-Catalogue Berbasis Website Menggunakan Metode Waterfall," *JSON*, vol. 3, no. 4, pp. 438–449,

- 2022, doi: 10.30865/json.v3i4.4165.
- [8] R. Wati, “Perancangan E-Katalog Berbasis Web Pada RR Collection Sampit Sebagai Media Branding Menggunakan Aplikasi Figma,” *EJECTS-E-Journal Comput. Technol. Informations Syst.*, vol. 02, no. 02, pp. 60–65, 2022.
 - [9] S. Mariam and H. A. Ramli, “Pengenalan Digital Marketing E-Katalog bagi UMKM Binaan Jakpreneur,” *J. Komunitas J. Pegabdian Kpd. Masy.*, vol. 5, no. 1, pp. 74–83, 2022.
 - [10] A. B. Santoso, I. Permana, E. Zusrony, and M. U. Dewi, “Implementasi Aplikasi digitalisasi Produk UMKM dengan E- Katalog dan GIS secara terpadu untuk Pemetaan dan optimasi penjualan,” *Jurnala Ilm. Elektron. dan Komput.*, vol. 15, no. 2, pp. 383–392, 2022.
 - [11] F. A. Diani, Fitri; Lubis, “Analisis Implementasi E-Katalog Terhadap Perkembangan UMKM di Kota Medan dalam Mendukung Kemajuan Ekonomi Syariah,” *JIEI*, vol. 8, no. 02, pp. 1970–1981, 2022.
 - [12] A. P. P. U. Fitroh, Meiga Ayu Ariyanti Nur; Wibawa, “Metode term frequency - invers document frequency pada mekanisme pencarian judul skripsi,” *TEKNO*, vol. 28, no. 2, pp. 177–190, 2018.
 - [13] D. Asmarajati, “ANALISIS PERBANDINGAN ALGORITMA TF-IDF DENGAN SQL QUERY UNTUK KASUS PENCARIAN PADA SISTEM INFORMASI DOKUMENTASI,” *J. Device*, vol. 10, no. 1, pp. 1–8, 2020.
 - [14] Maskur, N. P. D. N, and N. Hayatin, “Ekspansi Query Berbasis Semantik Pada Online Public Access Catalog,” *REPOSITOR*, vol. 1, no. 2, pp. 87–94, 2019.
 - [15] F. P. Wardani and B. Rahayudi, “Query Expansion Pada Sistem Temu Kembali Informasi Dokumen Jurnal Berbahasa Indonesia Menggunakan Metode BM25,” vol. 3, no. 3, pp. 2619–2625, 2019.