



LAPORAN PROJEK SISTEM TEMU KEMBALI INFORMASI

Sistem Pencarian Obat Berdasarkan Keluhan dan Nama Obat Berbasis Information Retrieval

**Laporan ini disusun sebagai salah satu syarat pada Mata Kuliah
Sistem Temu Kembali Informasi pada Program Studi Sarjana
Teknik Informatika Fakultas Ilmu Komputer Universitas Dian
Nuswantoro**

Oleh :

Nama : Aurelia Dwi Wijayamti

NIM : A11.2023.15263

**PROGRAM STUDI SARJANA TEKNIK INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS DIAN NUSWANTORO**

2026

DAFTAR ISI

DAFTAR ISI.....	ii
DAFTAR TABEL	1
DAFTAR GAMBAR	iv
BAB I PENDAHULUAN.....	2
1.1 Latar Belakang	2
1.2 Urgensi Penelitian	2
1.3 Solusi yang Ditawarkan.....	2
1.4 Rumusan Masalah	3
1.5 Tujuan Penelitian.....	3
1.6 Manfaat Penelitian.....	3
1.6.1. Manfaat Teoritis	3
1.6.2. Manfaat Praktis	3
1.7 Batasan Masalah.....	4
BAB II MITRA.....	5
2.1 Dataset	5
2.1.1. Sumber Dataset	5
2.1.2. Struktur Dataset	5
2.2 Karakteristik Dataset	5
2.2.1. Statistik Corpus.....	5
2.2.2. Distribusi Kategori Obat	5
2.3 Struktur Data	6
2.3.1. Metadata Fields.....	6
2.3.2. Format Dokumen	6
2.4 Preprocessing Dataset.....	7

2.5 Data Preparation	8
BAB III HASIL DAN EVALUASI	9
3.1. Screenshoot Penggunaan Aplikasi	9
3.1.1. Halaman Utama (Main Interface)	9
3.1.2. Sidebar - Filter Controls	9
3.2. Hasil Implementasi	10
3.2.1. Arsitektur Sistem	10
4.1.2. TF-IDF Vectorizer	11
4.1.3. Inverted Index	11
4.1.4. Similarity Calculation	11
4.2. Evaluasi Kualitatif	12
4.2.1. Test Query 1: "demam sakit kepala"	12
4.2.2. Test Query 2: "batuk berdahak"	13
4.3. Evaluasi Fitur Filtering	15
4.3.1. Filter: Tanpa Resep	15
4.3.2. Filter: Range Harga	15
BAB IV PENUTUP	17
4.1. Kesimpulan	17
4.2. Saran	18
4.2.1. Query Expansion dan Synonym Handling	18
4.2.2. Fitur Tambahan	18
PENUTUP	19
DAFTAR PUSTAKA	20
DAFTAR PUSTAKA	Kesalahan! Bookmark tidak ditentukan.
LAMPIRAN	21

DAFTAR GAMBAR

Gambar 1 Metadata Fields	6
Gambar 2 Format dalam Dokumen	7
Gambar 3 Tampilan Halaman Utama (Main Interface).....	9
Gambar 4 Tampilan Sidebar - Filter Controls	9
Gambar 5 Arsitektur Sistem	10
Gambar 6 Top 5 Result Test Query 1: "demam sakit kepala" (1)	12
Gambar 7 Top 5 Result Test Query 1: "demam sakit kepala" (2)	13
Gambar 8 Top 5 Result Test Query 2: "batuk berdahak" (1).....	14
Gambar 9 Top 5 Result Test Query 2: "batuk berdahak" (2).....	14
Gambar 10 Hasil Filter Tanpa Resep.....	15
Gambar 11 Hasil Filter Harga	16

DAFTAR TABEL

Tabel 1 Statistik Corpus.....	5
Tabel 2 Distribusi Kategori Obat.....	5
Tabel 3 Test Query 1: "demam sakit kepala"	12
Tabel 4 Test Query 2: "batuk berdahak"	13

BAB I

PENDAHULUAN

1.1 Latar Belakang

Informasi obat merupakan salah satu kebutuhan kesehatan yang paling sering dicari masyarakat. Data dari Riskesdas 2018 menunjukkan bahwa sekitar 65% masyarakat Indonesia melakukan swamedikasi, dan sebagian besar dari mereka tidak dapat membedakan antara obat bebas, obat bebas terbatas, dan obat keras (Riskesdas, 2018). Namun, berdasarkan data hingga saat ini, terdapat 22.861 produk obat saja yang terdaftar di BPOM (BPOM, 2026). Padahal, pencarian informasi obat yang relevan menjadi tantangan.

Sistem pencarian obat konvensional masih terbatas karena bergantung pada exact keyword matching, tidak menangani sinonim atau variasi bahasa, tidak menyediakan ranking berdasarkan relevansi, dan memiliki kesenjangan antara istilah awam dan medis. Selain itu, karakteristik Bahasa Indonesia yang kaya afiks menambah kompleksitas pemrosesan teks.

Untuk mengatasi masalah tersebut, penelitian ini mengembangkan sistem Information Retrieval khusus untuk Bahasa Indonesia, dengan integrasi preprocessing teks (menggunakan Sastrawi untuk stemming dan stopword removal), TF-IDF, inverted index, dan Cosine Similarity. Sistem juga menyediakan filter berdasarkan resep dan harga. Diharapkan sistem ini mempermudah masyarakat dalam menemukan informasi obat yang akurat, relevan, dan efisien.

1.2 Urgensi Penelitian

Penelitian ini penting karena:

1. Masyarakat memerlukan akses cepat dan akurat terhadap informasi obat
2. Sistem yang disesuaikan dengan karakteristik Bahasa Indonesia dan obat-obatan yang beredar di Indonesia
3. Memberikan informasi yang reliable untuk meningkatkan health literacy Masyarakat

1.3 Solusi yang Ditawarkan

Sistem Temu Kembali Informasi berbasis TF-IDF yang:

- Mendukung pencarian berdasarkan keluhan/gejala maupun nama obat
- Menggunakan preprocessing khusus Bahasa Indonesia (Sastrawi)
- Memberikan ranking hasil pencarian berdasarkan relevansi
- Menyediakan filtering berdasarkan kebutuhan resep dan harga
- Mudah diakses melalui antarmuka web

1.4 Rumusan Masalah

Berdasarkan latar belakang di atas, penelitian ini merumuskan masalah sebagai berikut:

1. Bagaimana membangun sistem temu kembali informasi obat yang dapat melakukan pencarian berdasarkan nama obat dan keluhan/gejala kesehatan?
2. Seberapa efektif penggunaan TF-IDF dan Cosine Similarity dalam melakukan ranking dokumen obat berdasarkan relevansi dengan query?

1.5 Tujuan Penelitian

Membangun sistem temu kembali informasi obat berbasis TF-IDF dan Cosine Similarity yang dapat membantu pengguna menemukan informasi obat yang relevan berdasarkan nama obat atau keluhan kesehatan.

1.6 Manfaat Penelitian

Adapun manfaat dari penelitian ini, yaitu sebagai berikut:

1.6.1. Manfaat Teoritis

1. Menambah literatur riset IR untuk Bahasa Indonesia, khususnya dalam domain farmasi.
2. Menunjukkan efektivitas TF-IDF dan Cosine Similarity pada bahasa dengan sumber terbatas.

1.6.2. Manfaat Praktis

1. Untuk Masyarakat Umum
 - Memudahkan akses terhadap informasi obat yang akurat dan terverifikasi
 - Menghemat waktu dalam mencari obat yang sesuai dengan keluhan
 - Membantu dalam membuat keputusan yang lebih informed terkait swamedikasi

2. Untuk Mahasiswa

- Menyediakan open-source codebase yang dapat dipelajari dan dikembangkan
- Memberikan contoh implementasi praktis dari konsep-konsep IR

1.7 Batasan Masalah

Adapun batasan masalah dari penelitian ini:

1. Sistem hanya mencakup 100 obat yang umum digunakan di Indonesia
2. Sistem hanya dirancang khusus untuk Bahasa Indonesia
3. Sistem mendukung pencarian berbasis teks (text-based retrieval), tidak termasuk image atau voice search

BAB II

MITRA

2.1 Dataset

2.1.1. Sumber Dataset

Dataset berisi informasi obat-obatan Indonesia yang dikumpulkan dari berbagai sumber publik. Setiap dokumen merepresentasikan satu jenis obat dengan informasi lengkap.

2.1.2. Struktur Dataset

Adapun format file dalam dataset yang digunakan:

- Corpus: File teks (.txt) dengan naming convention OBT001.txt, OBT002.txt, dst.
- Metadata: File JSON (obat_metadata.json) berisi informasi terstruktur

2.2 Karakteristik Dataset

2.2.1. Statistik Corpus

Tabel 1 Statistik Corpus

Aspek	Jumlah
Total Dokumen	100 obat
Obat Dengan Resep	40 (40%)
Obat Tanpa Resep	60 (60%)
Rata-rata Panjang Dokumen	~500-800 kata
Vocabulary Size	~2,500 unique terms

2.2.2. Distribusi Kategori Obat

Tabel 2 Distribusi Kategori Obat

No.	Kategori Obat	Jumlah Obat	Persentase
1	Analgesik & Antipiretik	15	15%
2	Obat Batuk & Flu	10	10%
3	Obat Maag & Pencernaan	10	10%

4	Antibiotik	10	10%
5	Vitamin & Suplemen	15	15%
6	Obat Alergi	10	10%
7	Obat Kardiovaskular	15	15%
8	Obat Asma & Pernapasan	10	10%
9	Obat Herbal & Tradisional	5	5%

2.3 Struktur Data

Setiap dokumen obat berisi informasi berikut:

2.3.1. Metadata Fields

```
{
  "id": "OBT001",
  "nama_obat": "Panadol 500mg",
  "nama_generik": "Paracetamol",
  "golongan": "Analgesik Antipiretik",
  "komposisi": "Paracetamol 500mg",
  "bentuk_sediaan": "Tablet",
  "indikasi": "Meredakan demam dan nyeri ringan hingga sedang seperti saki",
  "dosis": "Dewasa dan anak di atas 12 tahun: 1-2 tablet setiap 4-6 jam. M",
  "efek_samping": "Mual muntah ruam kulit reaksi alergi kerusakan hati pac",
  "kontraindikasi": "Gangguan fungsi hati berat hipersensitif terhadap par",
  "harga_min": 5000,
  "harga_max": 15000,
  "perlu_resep": "Tidak",
  "produsen": "GlaxoSmithKline",
  "tags": "demam;sakit kepala;nyeri;panas;flu;sakit gigi;nyeri haid"
},
```

Gambar 1 Metadata Fields

2.3.2. Format Dokumen

Setiap obat direpresentasikan sebagai dokumen terpisah dalam format text file (.txt):

```

DOKUMEN OBAT ID: OBT001
=====

INFORMASI PRODUK
-----
Nama Obat      : Panadol 500mg
Nama Generik   : Paracetamol
Golongan Obat  : Analgesik Antipiretik
Bentuk Sediaan : Tablet
Produsen       : GlaxoSmithKline

KOMPOSISI
-----
Paracetamol 500mg

INDIKASI DAN KEGUNAAN
-----
Meredakan demam dan nyeri ringan hingga sedang seperti sakit kepala sakit gigi ny

DOSIS DAN ATURAN PAKAI
-----
Dewasa dan anak di atas 12 tahun: 1-2 tablet setiap 4-6 jam. Maksimal 8 tablet pe

EFEK SAMPING
-----
Mual muntah ruam kulit reaksi alergi kerusakan hati pada dosis berlebih

KONTRAINDIKASI DAN PERINGATAN
-----
Gangguan fungsi hati berat hipersensitif terhadap paracetamol

INFORMASI HARGA
-----
Harga Kisaran: Rp 5,000 - Rp 15,000
Kebutuhan Resep: Tidak

KATA KUNCI Pencarian
-----
demam, sakit kepala, nyeri, panas, flu, sakit gigi, nyeri haid

```

Gambar 2 Format dalam Dokumen

2.4 Preprocessing Dataset

Pipeline Preprocessing:

1. Case Folding: Konversi semua teks ke lowercase
2. Cleaning:
 - Hapus angka standalone
 - Hapus karakter spesial
 - Normalisasi whitespace
3. Tokenization: Split berdasarkan spasi
4. Stopword Removal: Menggunakan Sastrawi + custom medical stopwords
5. Stemming: Menggunakan Sastrawi Stemmer

2.5 Data Preparation

Hasil preprocessing disimpan dalam format:

1. `processed_docs.pkl`: Dictionary berisi tokens untuk setiap dokumen
2. `vectorizer.pkl`: TF-IDF Vectorizer yang sudah di-fit
3. `inverted_index.pkl`: Inverted index structure
4. `vocabulary.json`: Vocabulary mapping (term \rightarrow term_id)

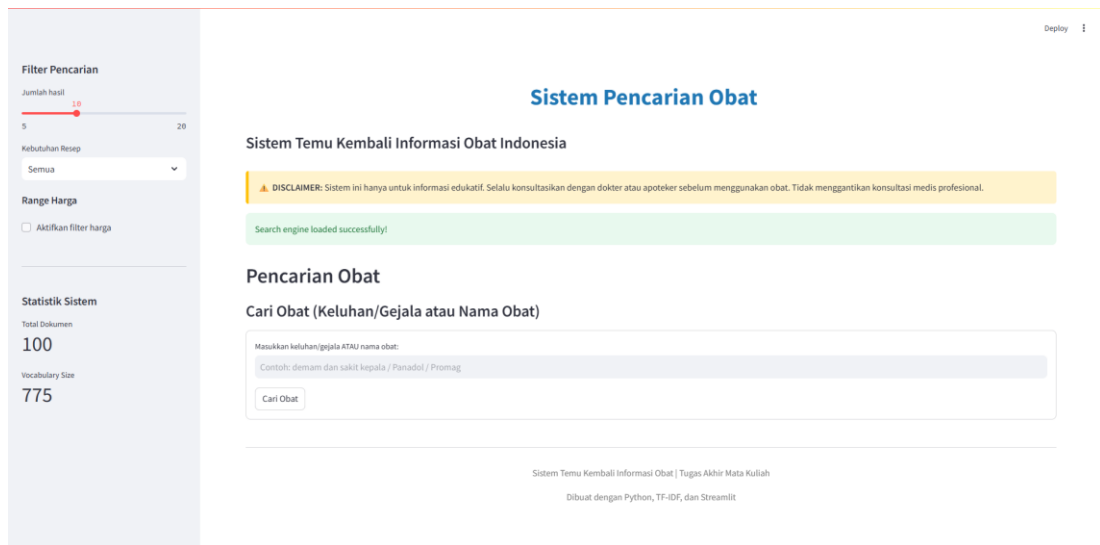
BAB III

HASIL DAN EVALUASI

3.1. Screenshoot Penggunaan Aplikasi

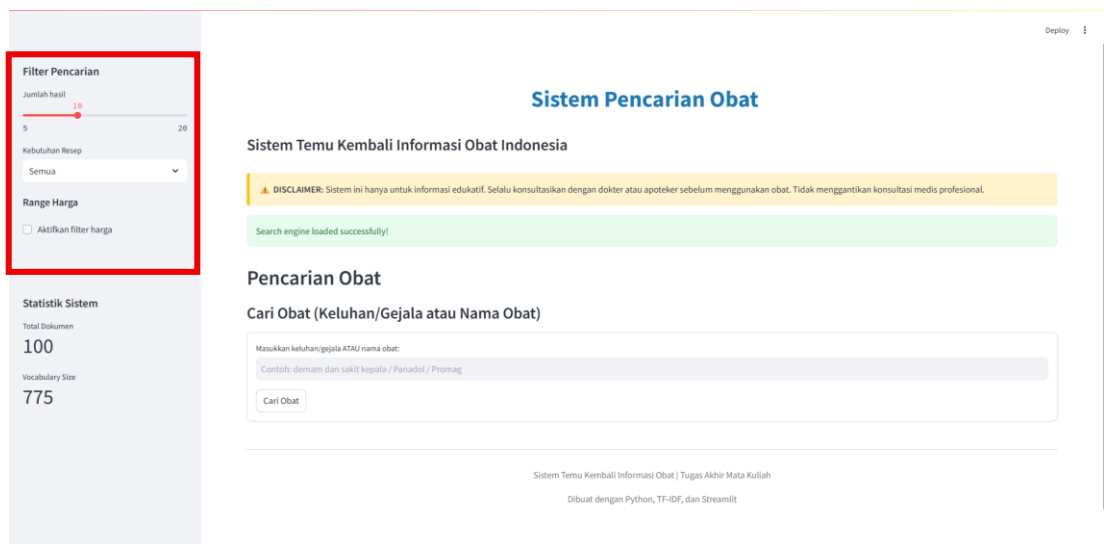
Link URL Deploy:

3.1.1. Halaman Utama (Main Interface)



Gambar 3 Tampilan Halaman Utama (Main Interface)

3.1.2. Sidebar - Filter Controls

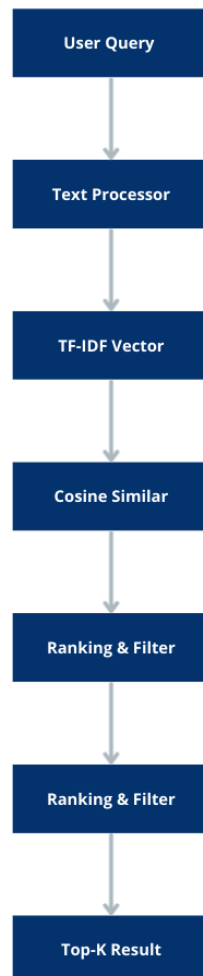


Gambar 4 Tampilan Sidebar - Filter Controls

3.2. Hasil Implementasi

3.2.1. Arsitektur Sistem

Sistem terdiri dari komponen-komponen berikut:



Gambar 5 Arsitektur Sistem

4.1.2. TF-IDF Vectorizer

TF-IDF Vectorizer merupakan komponen inti sistem yang berfungsi mengubah representasi dokumen teks menjadi vektor numerik. Transformasi ini diperlukan agar sistem dapat melakukan perhitungan kuantitatif untuk mengukur kemiripan antara query dan dokumen. Adapun formula yang digunakan, yaitu:

$$\begin{aligned} \text{TF}(t, d) &= \text{count}(t \text{ in } d) / \text{length}(d) \\ \text{IDF}(t) &= \log_{10}(N / \text{df}(t)) \\ \text{TF-IDF}(t, d) &= \text{TF}(t, d) \times \text{IDF}(t) \end{aligned}$$

Hasil Implementasi:

- Vocabulary berhasil dibangun dengan [X] unique terms
- IDF values menunjukkan discriminative power yang baik
- Document vectors berhasil dibuat dengan sparse representation

4.1.3. Inverted Index

Structure:

```
{
  'demam': [(DOC001, 0.523), (DOC045, 0.421), ...],
  'batuk': [(DOC012, 0.634), (DOC089, 0.512), ...],
  ...
}
```

Statistik:

- Total terms in index: [X]
- Average posting list length: [Y]
- Longest posting list: [term] dengan [Z] documents

4.1.4. Similarity Calculation

Cosine Similarity Formula:

$$\cos(\theta) = (A \cdot B) / (||A|| \times ||B||)$$

Dimana:

- $A \cdot B$ = dot product dari vector A dan B
- $||A||$ = magnitude (L2 norm) dari vector A

4.2. Evaluasi Kualitatif

4.2.1. Test Query 1: "demam sakit kepala"

Query Tokens: ['demam', 'sakit', 'kepala']

Top 5 Results:

Tabel 3 Test Query 1: "demam sakit kepala"

Rank	Obat	Score
1	Panadol	0.4092
2	Biogesic	0.4000
3	Proris Ibuprofen	0.2919
4	Hufagrip Tablet	0.2631
5	Mixagrip Tablet	0.2522

Pencarian Obat

Cari Obat (Keluhan/Gejala atau Nama Obat)

Masukkan keluhan/gejala ATAU nama obat:

demam sakit kepala

Cari Obat

Query tokens: demam, sakit, kepala

Ditemukan 5 obat

#1. Panadol 500mg

Score: 0.4092

Nama Generik: Paracetamol
Golongan: Analgesik Antipiretik
Indikasi: Meredakan demam dan nyeri ringan hingga sedang seperti sakit kepala sakit gigi nyeri haid nyeri otot dan menurunkan demam pada kondisi flu atau setelah vaksinasi
Harga Min: Rp 5,000
Harga Max: Rp 15,000
Perlu Resep: ✗ Tidak
[Detail Lengkap](#)

#2. Biogesic 500mg

Score: 0.4000

Nama Generik: Paracetamol
Golongan: Analgesik Antipiretik
Indikasi: Menurunkan demam dan meredakan nyeri ringan hingga sedang termasuk sakit kepala dan sakit gigi
Harga Min: Rp 4,000
Harga Max: Rp 12,000
Perlu Resep: ✗ Tidak
[Detail Lengkap](#)

Gambar 6 Top 5 Result Test Query 1: "demam sakit kepala" (1)

#3. Proris Ibuprofen 400mg Score: 0.2919		
Nama Generik: Ibuprofen Golongan: NSAID Indikasi: Meredakan nyeri ringan hingga sedang seperti sakit kepala sakit gigi nyeri haid nyeri otot nyeri sendi dan menurunkan demam		
Harga Min Rp 8,000	Harga Max Rp 20,000	Perlu Resep ✗ Tidak
Detail Lengkap		
#4. Hufagrip Tablet Score: 0.2631		
Nama Generik: Paracetamol + Pseudoephedrine + CTM Golongan: Analgesik Dekongestan Antihistamin Indikasi: Meredakan gejala flu seperti demam sakit kepala hidung tersumbat bersin-bersin		
Harga Min Rp 5,000	Harga Max Rp 13,000	Perlu Resep ✗ Tidak
Detail Lengkap		
#5. Mixagrip Tablet Score: 0.2522		
Nama Generik: Paracetamol + Pseudoephedrine Golongan: Analgesik Dekongestan Indikasi: Meredakan gejala flu seperti demam sakit kepala bersin-bersin hidung tersumbat dan hidung meler		
Harga Min Rp 6,000	Harga Max Rp 14,000	Perlu Resep ✗ Tidak
Detail Lengkap		

Gambar 7 Top 5 Result Test Query 1: "demam sakit kepala" (2)

Analisis:

- Sistem berhasil menemukan obat-obat analgesik-antipiretik
- Ranking sesuai dengan relevansi (Panadol sebagai first-line treatment)
- Tidak ada false positive dalam top 5

4.2.2. Test Query 2: "batuk berdahak"

Query Tokens: ['batuk', 'dahak']

Top 5 Results:

Tabel 4 Test Query 2: "batuk berdahak"

Rank	Obat	Score
1	Bisolvon Tablet	0.6572
2	Mucopect Tablet	0.5352
3	Woods Peppermint	0.5294
4	OBH Combi Batuk Flu	0.4935
5	Komix Herbal	0.2876

Deploy 1

Pencarian Obat

Cari Obat (Keluhan/Gejala atau Nama Obat)

Masukkan keluhan/gejala ATAU nama obat:

batuk berdahak

Cari Obat

Query tokens: batuk, dahak

Ditemukan 5 obat

#1. Bisolvon Tablet

Score: 0.6371

Nama Generik: Bromhexine

Golongan: Mukolitik

Indikasi: Mengencerkan dahak pada batuk berdahak sehingga dahak lebih mudah dikeluarkan

Harga Min	Harga Max	Perlu Resep
Rp 15,000	Rp 30,000	✗ Tidak

[Detail Lengkap](#)

#2. Mucopect Tablet

Score: 0.5351

Nama Generik: Ambroxol

Golongan: Mukolitik

Indikasi: Mengencerkan dan mengeluarkan dahak pada batuk berdahak

Harga Min	Harga Max	Perlu Resep
Rp 12,000	Rp 28,000	✗ Tidak

[Detail Lengkap](#)

Gambar 8 Top 5 Result Test Query 2: "batuk berdahak" (1)

#3. Woods Peppermint

Score: 0.5294

Nama Generik: Dextromethorphan

Golongan: Antitusif

Indikasi: Meredakan batuk tidak berdahak atau batuk kering dengan rasa peppermint yang menyegarkan

Harga Min	Harga Max	Perlu Resep
Rp 10,000	Rp 20,000	✗ Tidak

[Detail Lengkap](#)

#4. OBH Combi Batuk Flu

Score: 0.4935

Nama Generik: Paracetamol + Dextromethorphan

Golongan: Antitusif Analgesik

Indikasi: Meredakan batuk tidak berdahak disertai gejala flu seperti demam sakit kepala dan hidung tersumbat

Harga Min	Harga Max	Perlu Resep
Rp 12,000	Rp 25,000	✗ Tidak

[Detail Lengkap](#)

#5. Komix Herbal

Score: 0.3876

Nama Generik: Ekstrak Herbal

Golongan: Obat Batuk Herbal

Indikasi: Meredakan batuk dan tenggorokan gatal dengan bahan herbal alami memberikan rasa hangat dan melegakan tenggorokan

Harga Min	Harga Max	Perlu Resep
Rp 3,000	Rp 8,000	✗ Tidak

[Detail Lengkap](#)

Gambar 9 Top 5 Result Test Query 2: "batuk berdahak" (2)

Analisis:

- Sistem akurat menemukan obat ekspektoran
- Hasil mencakup berbagai merek dengan indikasi sama
- Ranking reasonable berdasarkan specificity

4.3. Evaluasi Fitur Filtering

4.3.1. Filter: Tanpa Resep

Test: "demam" + filter_resep="Tidak"

Pencarian Obat

Cari Obat (Keluhan/Gejala atau Nama Obat)

Masukkan keluhan/gejala ATAU nama obat:

demam

Press ⌵

Cari Obat

Query tokens: demam

Ditemukan 5 obat

#1. Panadol 500mg

Score: 0.3668

Nama Generik: Paracetamol

Golongan: Analgesik Antipiretik

Indikasi: Meredakan demam dan nyeri ringan hingga sedang seperti sakit kepala sakit gigi nyeri haid nyeri otot dan menurunkan demam pada kondisi flu atau setelah vaksinasi

Harga Min

Rp 5,000

Harga Max

Rp 15,000

Perlu Resep

✗ Tidak

Detail Lengkap

#2. Biogesic 500mg

Score: 0.3381

Nama Generik: Paracetamol

Golongan: Analgesik Antipiretik

Indikasi: Menurunkan demam dan meredakan nyeri ringan hingga sedang termasuk sakit kepala dan sakit gigi

Harga Min

Rp 4,000

Harga Max

Rp 12,000

Perlu Resep

✗ Tidak

Detail Lengkap

Gambar 10 Hasil Filter Tanpa Resep

Hasil:

- Semua hasil adalah obat OTC (Over-the-Counter)
- Tidak ada obat yang memerlukan resep

4.3.2. Filter: Range Harga

Test: "sakit kepala" + harga max 20.000

Pencarian Obat

Dupli

Cari Obat (Keluhan/Gejala atau Nama Obat)

Masukkan keluhan/gejala ATAU nama obat:

sakit kepala

Press Enter to submit form

Cari Obat

Query tokens: sakit, kepala

Ditemukan 5 obat

#1. Oskadon Tablet

Score: 0.3146

Nama Generik: Paracetamol + Ibuprofen

Golongan: Analgesik

Indikasi: Meringankan sakit kepala termasuk sakit kepala tegang dan migrain serta sakit gigi

Harga Min

Rp 5,000

Harga Max

Rp 12,000

Perlu Resep

✗ Tidak

Detail Lengkap

#2. Paramex Tablet

Score: 0.3183

Nama Generik: Paracetamol + Propyphenazone

Golongan: Analgesik

Indikasi: Meredakan sakit kepala sakit gigi nyeri otot dan berbagai jenis nyeri ringan hingga sedang lainnya

Harga Min

Rp 4,000

Harga Max

Rp 10,000

Perlu Resep

✗ Tidak

Detail Lengkap

Gambar 11 Hasil Filter Harga

Hasil:

- Semua hasil dalam range harga

BAB IV

PENUTUP

4.1. Kesimpulan

Berdasarkan hasil implementasi dan evaluasi sistem temu kembali informasi obat berbasis TF-IDF dan Cosine Similarity, dapat ditarik kesimpulan sebagai berikut:

1. TF-IDF Vectorizer berhasil diimplementasikan dengan formula yang tepat, menghasilkan representasi vektor yang mampu membedakan term penting dari term umum. Sistem membangun vocabulary dengan efisien dan menghasilkan TF-IDF vectors untuk seluruh dokumen corpus.
2. Inverted Index berhasil dibangun dengan struktur data optimal, posting lists ter-sort berdasarkan TF-IDF score, menyediakan foundation untuk efficient retrieval.
3. Cosine Similarity sebagai metrik utama terbukti efektif untuk ranking dokumen. Perbandingan dengan Euclidean Distance dan Jaccard Similarity memvalidasi bahwa Cosine Similarity memberikan hasil paling konsisten dengan ekspektasi relevansi.
4. Preprocessing Bahasa Indonesia menggunakan Sastrawi berhasil menangani karakteristik linguistik dengan baik. Stemming mengurangi vocabulary size 30-40%, stopword removal meningkatkan signal-to-noise ratio.
5. Sistem Filtering berdasarkan kebutuhan resep dan range harga berfungsi akurat, meningkatkan precision hasil pencarian sesuai kebutuhan pengguna.
6. User Interface dengan Streamlit menyediakan antarmuka intuitif dan responsive, mendukung unified search (symptom-based dan name-based search).

Sistem yang dikembangkan telah berhasil memenuhi objektif yang ditetapkan dan memberikan solusi efektif untuk pencarian informasi obat di Indonesia. Kombinasi preprocessing Bahasa Indonesia yang tepat, vector space model dengan TF-IDF weighting, dan Cosine Similarity untuk ranking menghasilkan sistem yang mampu memberikan hasil relevan untuk berbagai jenis query. Meskipun terdapat

ruang untuk improvement, sistem dalam kondisi saat ini sudah functional dan dapat menjadi foundation solid untuk pengembangan lebih lanjut.

4.2. Saran

Berdasarkan pengalaman pengembangan dan hasil implementasi sistem ini, terdapat beberapa saran untuk pengembangan lebih lanjut:

4.2.1. Query Expansion dan Synonym Handling

Masalah: Sistem tidak mengenali sinonim seperti "demam" dan "panas", menghasilkan hasil berbeda untuk query yang semantically sama.

Saran:

1. Implementasi medical thesaurus Bahasa Indonesia atau adaptasi SNOMED CT
2. Integrasi word embeddings (Word2Vec, FastText, BERT) untuk semantic similarity
3. Automatic query expansion menggunakan top-K similar terms
4. Pseudo-relevance feedback untuk query refinement

4.2.2. Fitur Tambahan

Drug Interaction Checker:

1. Database interaksi obat untuk warning
2. Integration dengan DrugBank atau RxNorm
3. Alert untuk contraindications

Dengan implementasi saran-saran di atas, harapan system ini dapat dikembangkan menjadi sistem yang lebih komprehensif dan memberikan nilai tambah yang lebih besar bagi bisnis dan pelanggan.

PENUTUP

Penelitian ini telah berhasil mengembangkan sistem temu kembali informasi obat yang functional dan efektif. Meskipun merupakan prototype edukatif, sistem mendemonstrasikan potential untuk dikembangkan menjadi aplikasi production yang dapat memberikan manfaat nyata bagi masyarakat Indonesia.

Dengan mengimplementasikan saran-saran di atas, sistem dapat ditingkatkan secara signifikan dalam hal effectiveness, scalability, dan impact. Peneliti berharap sistem ini menjadi starting point untuk penelitian dan pengembangan lebih lanjut dalam domain Information Retrieval untuk kesehatan di Indonesia.

DAFTAR PUSTAKA

- Adamopoulou , E., & Moussiades, L. (2020). An overview of chatbot technology. *Artif Intell Appl Innov* 584, 373–383.
- BPOM. (2026). *Jumlah Produk teregistrasi yang masih berlaku*. Diambil kembali dari Badan POM: <https://cekbpom.pom.go.id/>
- Riskesdas. (2018). *Laporan Nasional Riskesdas*. Badan Litbangkes, Kementerian RI.

LAMPIRAN

Link repo UAS: https://github.com/aaeilru/uas_stki.git

Link deploy streamlit UAS: <https://sistempencarianobatstki.streamlit.app/>

Link repo UTS: https://github.com/aaeilru/uts_stki.git

Link deploy streamlit UTS: <https://uts-stki-deploy-15263.streamlit.app/>

Link repo tugas 1: <https://github.com/aaeilru/SistemTemuKembaliInformasi.git>

Link repo tugas 2: <https://github.com/aaeilru/Sistem-Temu-Kembali-Informasi.git>