

Nama : Suryani Ayu Dewanti

NIM : A11.2023.15018

Kelompok : A11.4703

1. Definisi STKI dan Perbedaan dengan Database Retrieval

Sistem Temu Kembali Informasi (STKI) adalah sistem yang dirancang untuk menemukan dan mengambil informasi yang relevan berdasarkan kebutuhan pengguna dari koleksi besar dokumen, terutama teks tak terstruktur. Berbeda dengan sistem basis data tradisional yang bersifat terstruktur dan menggunakan pencocokan eksak, STKI berfokus pada pencarian semantik yang bersifat toleran terhadap variasi bahasa seperti sinonim dan ambiguitas.

Database retrieval bekerja dengan query yang pasti (misalnya SQL), di mana hasilnya berupa data yang memenuhi kondisi logika tertentu secara presisi. Sebaliknya, STKI mengembalikan dokumen berdasarkan tingkat relevansi. Hal ini membuat STKI lebih fleksibel untuk menangani konteks pencarian manusia, sementara database retrieval lebih cocok untuk kebutuhan administratif dan transaksi data yang presisi.

2. Peran Index dan Ranking dalam STKI

Index berfungsi sebagai struktur data yang mempercepat pencarian dokumen tanpa harus memindai seluruh koleksi. Inverted index adalah bentuk paling umum, yang menyimpan hubungan antara setiap term dengan daftar dokumen tempat term tersebut muncul.

Ranking menentukan urutan hasil pencarian berdasarkan tingkat relevansi terhadap query pengguna. Dalam model Boolean, hasil bersifat biner (relevan atau tidak), sedangkan dalam model berbasis vektor seperti Vector Space Model (VSM), sistem menghitung skor relevansi menggunakan ukuran seperti cosine similarity, sehingga hasil dapat diurutkan dari yang paling relevan hingga yang paling tidak.

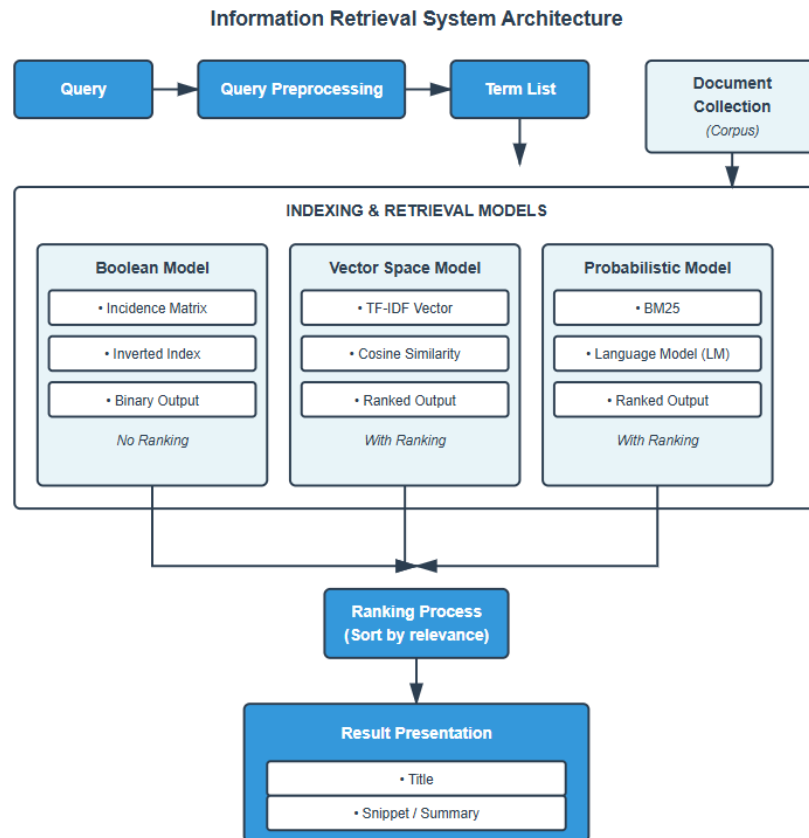
3. Garis Besar Arsitektur Search Engine Klasik

Arsitektur klasik sistem temu kembali informasi umumnya terdiri dari tiga komponen utama:

- 1) Indexing pipeline, yang bertanggung jawab untuk preprocessing dan pembuatan indeks.
- 2) Query processing & retrieval, yang menangani pemrosesan query dan pencocokan dengan indeks.
- 3) Presentation layer, yang menampilkan hasil pencarian dalam bentuk daftar dokumen atau halaman hasil pencarian (SERP).

Alur kerja dimulai dari pengguna yang mengajukan query. Sistem melakukan preprocessing seperti tokenisasi, case folding, stopword removal, dan stemming. Kemudian query dibandingkan dengan indeks untuk menemukan dokumen yang relevan. Setelah itu, sistem menghitung skor relevansi, mengurutkannya (ranking), dan menampilkan hasil akhir kepada pengguna.

4. Sketsa Arsitektur Retrieval (Boolean & Vector Space Model)



5. Demo Singkat Arsitektur IR (Mini Search Engine)

1) Boolean (AND, OR, & NOT)

Mini Search Engine (UTS)

Masukkan query:

Skema pencarian

Boolean

Top-k dokumen

4

Hasil Pencarian

	0	1	2
0	1	itb.txt	informasi umum institut teknologi bandung itb nama singkatan nama resmi institut...
1	2	ui.txt	informasi umum universitas indonesia ui nama singkatan nama resmi universitas in...

Evaluasi

Precision: 0.4, Recall: 0.13, F1: 0.2, MAP@4: 0.13, nDCG@4: 0.55

Mini Search Engine (UTS)

Masukkan query:

Skema pencarian

Boolean

Top-k dokumen

5

Hasil Pencarian

	0	1	2
0	1	itb.txt	informasi umum institut teknologi bandung itb nama singkatan nama resmi institut...
1	2	ub.txt	informasi umum universitas brawijaya ub nama singkatan nama resmi universitas br...
2	3	udinus.tx t	informasi umum universitas dian nuswantoro udinus nama singkatan nama resmi univ...
3	4	ugm.txt	informasi umum universitas gadjah mada ugm nama singkatan nama resmi universitas...
4	5	ui.txt	informasi umum universitas indonesia ui nama singkatan nama resmi universitas in...

Evaluasi

Precision: 1.0, Recall: 0.33, F1: 0.5, MAP@5: 0.33, nDCG@5: 1.0

Mini Search Engine (UTS)

Masukkan query:

Skema pencarian

Boolean

Top-k dokumen

1 6 10

Hasil Pencarian

	0	1	2
0	1	stmik_bm_palu.txt	informasi umum stmik bina mulia palu nama singkatan nama resmi stmik bina mulia ...
1	2	ub.txt	informasi umum universitas brawijaya ub nama singkatan nama resmi universitas br...
2	3	udinus.txt	informasi umum universitas dian nuswantoro udinus nama singkatan nama resmi univ...
3	4	ugm.txt	informasi umum universitas gadjah mada ugm nama singkatan nama resmi universitas...
4	5	unand.txt	informasi umum universitas andalas unand nama singkatan nama resmi universitas a...
5	6	undip.txt	informasi umum universitas diponegoro undip nama singkatan nama resmi universitas...

Evaluasi

Precision: 0.0, Recall: 0.0, F1: 0, MAP@6: 0.0, nDCG@6: 0.0

Mini Search Engine (UTS)

Masukkan query:

Skema pencarian

Boolean

Top-k dokumen

10

Hasil Pencarian

	0	1	2
0	1	itb.txt	informasi umum institut teknologi bandung itb nama singkatan nama resmi institut...
1	2	stmik_bm_palu.txt	informasi umum stmik bina mulia palu nama singkatan nama resmi stmik bina mulia ...
2	3	ub.txt	informasi umum universitas brawijaya ub nama singkatan nama resmi universitas br...
3	4	udinus.txt	informasi umum universitas dian nuswantoro udinus nama singkatan nama resmi univ...
4	5	ugm.txt	informasi umum universitas gadjah mada ugm nama singkatan nama resmi universitas...
5	6	ui.txt	informasi umum universitas indonesia ui nama singkatan nama resmi universitas in...
6	7	undip.txt	informasi umum universitas diponegoro undip nama singkatan nama resmi universitas...
7	8	ung.txt	informasi umum universitas negeri gorontalo ung nama singkatan nama resmi univer...
8	9	unima.txt	informasi umum universitas negeri manado unima nama singkatan nama resmi univers...
9	0	unimal.txt	informasi umum universitas malikussaleh unimal nama singkatan nama resmi univers...

2) VSM TF-IDF vs VSM Sublinear

Mini Search Engine (UTS)

Masukkan query:

Skema pencarian

VSM Sublinear

Top-k dokumen

5

Hasil Pencarian

	0	1	2	3
0	1	unud.txt	0.136	informasi umum universitas udayana unud nama singkatan nama resmi universitas ud...
1	2	ung.txt	0.132	informasi umum universitas negeri gorontalo ung nama singkatan nama resmi univer...
2	3	unsri.txt	0.124	informasi umum universitas sriwijaya unsri nama singkatan nama resmi universitas...
3	4	undip.txt	0.121	informasi umum universitas diponegoro undip nama singkatan nama resmi universitas...
4	5	ub.txt	0.120	informasi umum universitas brawijaya ub nama singkatan nama resmi universitas br...

Evaluasi

Precision: 1.0, Recall: 0.36, F1: 0.53, MAP@5: 0.36, nDCG@5: 1.0

Mini Search Engine (UTS)

Masukkan query:

Skema pencarian

VSM TF-IDF

Top-k dokumen

5

Hasil Pencarian

	0	1	2	3
0	1	undip.txt	0.180	informasi umum universitas diponegoro undip nama singkatan nama resmi universita...
1	2	unud.txt	0.162	informasi umum universitas udayana unud nama singkatan nama resmi universitas ud...
2	3	unmubuton.txt	0.150	informasi umum universitas muhammadiyah buton nama singkatan nama resmi universi...
3	4	ub.txt	0.149	informasi umum universitas brawijaya ub nama singkatan nama resmi universitas br...
4	5	ung.txt	0.149	informasi umum universitas negeri gorontalo ung nama singkatan nama resmi univer...

Evaluasi

Precision: 1.0, Recall: 0.36, F1: 0.53, MAP@5: 0.36, nDCG@5: 1.0

5. Peta Materi ke RPS

Berdasarkan outline mata kuliah, hubungan antara soal-soal dan materi RPS dapat dijelaskan sebagai berikut:

No.	Materi RPS	Keterkaitan
1	Materi 1, 2.2, Search Engine Concept	Konsep dasar STKI, arsitektur, dan mapping RPS
2	Materi 2.2 (Preprocessing)	Tokenisasi, stopword removal, stemming, dan normalisasi
3	Materi 3.2 (Boolean Model)	Inverted index, incidence matrix, evaluasi query
4	Materi 4 (Evaluasi.pdf)	Precision, recall, dan metrik evaluasi sistem
5	Materi 5–6 (Word Embedding, Word2Vec)	TF-IDF, cosine similarity, dan semantic retrieval