Nama: Rifa’atus Shalihah

Mata Ujian/ kelas: Sistem Temu Balik Informasi

Hari tanggal ujian: 14 Desember 2022

1. Jelaskan secara detil metode kueri dan similarity search pada proyek pengembangan stbi kelompok saya (Kelompok 5)

Metode yang kami gunakan yaitu metode vector space model. Sebelum metode masuk pada metode vector space model, Tahapan awal yang dilakukan adalah proses preprocessing dimana dokumen collection dan kueri pencarian akan mengalami text preprocessing.

* Tahapan preprocessing yang dilakukan oleh proyek pengembangan kami yaitu case folding, tokenizing, filtering dan stemming. Case folding yaitu merubah semua karakter huruf pada sebuah kalimat menjadi huruf kecil dan menghilangkan karakter yang dianggap tidak valid dan tidak diperlukan infor- masinya seperti angka, tanda baca, dan Uniform Resources Locator (URL). Kemudian akan dilanjutkan dengan tokenizing yaitu memotong sebuah kalimat berdasarkan tiap kata yang menyusunnya. Lalu selanjutnya Stemming yaitu merubah berbagai kata berimbuhan menjadi kata dasarnya dengan menggunakan library Sastrawi.
* Setelah proses preprocessing, Tahapan selanjutnya yaitu pembuatan inverted index. Inverted index merupakan konversi dari dokumen asli yang mengandung sekumpulan kata ke dalam daftar kata (dictionary) yang berasosiasi dengan dokumen terkait dimana kata-kata tersebut muncul (postings). Dictionary adalah daftar kata yang diperoleh dari hasil pengindeksan dokumen. Setelah proses pembuatan inverted index selesai, inverted index tersebut akan disimpan pada basis data dengan informasi-informasi yang dibutuhkan.
* Selanjutnya adalah proses similarity antara dokumen koleksi dan kueri yang diinputkan. Pada proyek ini menggunakan metode vector space model. Proses perhitungan vector space model (VSM) melalui tahapan perhitungan term frequency (tf), inverse document frequency (idf), term frequency inverse document frequency (tf– idf), jarak query dan dokumen, pengukuran similaritas query document (inner product), dan pengukuran Cosine Similarity (menghitung nilai cosine sudut antara dua vektor).
* Menghitung bobot dokumen dan kueri menggunakan Tf-idf

Term frequency–inverse document frequency atau biasa sering disebut tf – idf adalah metode pembobotan kata dengan menghitung term (t) yang terdapat pada setiap dokumen (d), banyaknya term pada setiap dokumen ditunjukkan dengan notasi Term Frequency (tf). Kehadiran term total di semua dokumen dikumpulkan dalam notasi (df). Inverse document frequency (idf) adalah jumlah dokumen yang mengandung sebuah term didasarkan pada seluruh dokumen yang ada pada data set. Terdapat beberapa cara atau metode dalam melakukan pembobotan kata pada metode tf – idf yaitu melalui skema pembobotan query dan dokuman. Formula yang digunakan pada term frequency dihitung dengan menggunakan persamaan dibawah ini:

Dimana, N adalah total jumlah dokumen dalam koleksi, dfi = jumlah dokumen yang berisi term j, dan idfi = frekuensi kemunculan suatu istilah dalam dokumen i Dengan demikian rumus untuk tf – idf dapat dilihat pada Persamaan berikut.

Wij = tfij \* idfi

Setelah bobot (W) masing-masing dokumen diketahui, maka dilakukan proses sorting atau pengurutan dimana semakin besar nilai W, semakin besar tingkat similaritas dokumen tersebut terhadap kata kunci, demikian seba- liknya.

* Menghitung Jarak Query dengan Dokumen

Perhitungan jarak dokumen dan jarak query dilakukan setelah perhitungan bobot. Jika, Q adalah vektor query dan D adalah vektor dokumen, merupakan 2 buah vektor dalam ruang berdimensi-n, dan θ adalah sudut yang dibentuk oleh 2 vektor tersebut. Perhitungan jarak query dan dokumen menggunakan Persamaan berikut.

Dengan |Q| adalah jarak query dan Wiq adalah bobot query dokumen ke-i

* Menghitung Nilai cosine similarity

Melalui vector space model dan tf-idf weighting akan didapatkan representasi nilai numerik dokumen sehingga dapat dihitung kedekatan antar dokumen. Semakin dekat dua vektor didalam suatu vector space model, maka semakin mirip dua dokumen yang diwakili oleh dua vektor tersebut. Kemiripan antar dokumen dapat dihitung menggunakan suatu fungsi ukuran kemiripan (similarity measure). Ukuran ini memungkinkan perankingan dokumen sesuai dengan kemiripannya atau relevansinya terhadap query.

Pengukuran cosine similarity (menghitung nilai cosine sudut antara dua vektor) menggunakan Persamaan berikut.

1. Metode evaluasi yang kelompok kami (Kelompok 5) gunakan untuk mengevaluasi prototype STBI beserta hasil evaluasi yang diperoleh.

Metode evaluasi yang kami gunakan yaitu precision and recall. Precision adalah rasio dokumen relevan yang berhasil ditemu kembalikan dari seluruh dokumen. Recall adalah rasio antara dokumen yang relevan yang berhasil ditemu kembalikan dari seluruh dokumen relevan yang ada didalam sistem. Rumuas precision and recall dapat dilihat pada persamaan berikut.

Selain precision and recall kami juga melihat runtime dari sistem pencarian ini. skenario dari evaluasi kami yaitu dilakukan 10 kali percobaan untuk setiap query dengan jumlah query yang berbeda. Evaluasi ini berfungsi untuk melihat berapa nilai precision and recall serta waktu yang diperlukan untuk retrieve dokumen.

Untuk hasil evaluasi dari kuery 1 kata dapat dilihat pada tabel dibawah ini.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| No | Query | Runtime | Precision | Recall |
| 1 | musik | 0.5921080 | 100 | 37.5 |
| 2 | huruf | 0.5898170 | 100 | 16.666666 |
| 3 | hiragana | 0.5847039 | 100 | 40 |
| 4 | masak | 0.5654931 | 0 | 0 |
| 5 | nasi | 0.5742459 | 100 | 50 |
| 6 | manusia | 0.6029548 | 75 | 15 |
| 7 | angka | 0.9094681 | 100 | 8.1779053 |
| 8 | ginjal | 0.5639269 | 100 | 50 |
| 9 | tabel | 0.9389278 | 100 | 9.433962 |
| 10 | rambut | 0.5675659 | 100 | 50 |
| Rata-rata | | 0.648 | 87.5 | 27.67 |

Dari hasil pengujian precision dan recall yang dilakukan pada query yang terdiri dari 1 kata memiliki nilai rata-rata precision sebesar 87.5 dan nilai rata-rata recall sebesar 27.67. Untuk nilai rata-rata runtime sebesar 0.648

Adapun hasil pengujian precision dan recall yang dilakukan pada query yang terdiri dari 2 kata dapat dilihat pada Tabel berikut.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| No | Query | Runtime | Precision | Recall |
| 1 | Musik jawa | 1.0879840 | 75 | 33.33 |
| 2 | Penyakit menular | 1.0182321 | 0 | 0 |
| 3 | Belajar matematika | 1.0522079 | 0 | 0 |
| 4 | Lambung manusia | 1.1111891 | 0 | 50 |
| 5 | Ikan laut | 1.0558810 | 100 | 50 |
| 6 | Lambang Indonesia | 1.0504291 | 50 | 50 |
| 7 | Batu ginjal | 1.0574929 | 50 | 50 |
| 8 | Bahasa jepang | 1.0798890 | 100 | 40 |
| 9 | Burung Garuda | 1.0464317 | 100 | 50 |
| 10 | Pesawat terbang | 1.1254491 | 100 | 50 |
| Rata-rata | | 1.068 | 57.5 | 37.33 |

Dari hasil pengujian precision dan recall yang dilakukan pada query yang terdiri dari 2 kata memiliki nilai rata-rata precision sebesar 57.5 dan nilai rata-rata recall sebesar 37.33. Untuk nilai rata-rata runtime sebesar 1.068

Adapun hasil pengujian precision dan recall yang dilakukan pada query yang terdiri dari 3 kata dapat dilihat pada Tabel berikut.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| No | Query | Runtime | Precision | Recall |
| 1 | Struktur virus covid | 1.5175569 | 33.33333333 | 50 |
| 2 | Budaya wayang jawa | 1.5313150 | 100 | 41.66667 |
| 3 | Alat music gamelan | 1.8257658 | 100 | 55.555556 |
| 4 | Nasi arab kebuli | 1.6307568 | 0 | 50 |
| 5 | Perkalian matematika sederhana | 1.5602569 | 83.33333 | 83.33333 |
| 6 | Transportasi udara pesawat | 1.5167338 | 100 | 50 |
| 7 | Pembuluh darah manusia | 1.5598120 | 83.33333333 | 27.77778 |
| 8 | Tumbuhan yang mengandung vitamin | 1.5462398 | 100 | 100 |
| 9 | Bahasa jepang susah | 1.5550308 | 0 | 0 |
| 10 | Anatomi mata manusia | 1.7289500 | 3.03030303 | 50 |
| Rata-rata | | 1.597 | 60.30 | 50.83 |

Dari hasil pengujian precision dan recall yang dilakukan pada query yang terdiri dari 2 kata memiliki nilai rata-rata precision sebesar 60.30 dan nilai rata-rata recall sebesar 50.83. Untuk nilai rata-rata runtime sebesar 1.597.

1. Jelaskan perbedaan antara sistem rekomendasi dan sistem tanya jawab. Menurut ada proyek stbi anda masuk kategori mana?

Sistem rekomendasi adalah sistem yang menggunakan teknologi pembelajaran mesin untuk menganalisis preferensi dan perilaku pengguna, dan memberikan rekomendasi produk atau layanan yang sesuai dengan kebutuhan dan minat pengguna. Sistem ini biasanya digunakan dalam e-commerce, media sosial, atau aplikasi streaming video untuk memberikan rekomendasi konten kepada pengguna.

Sedangkan sistem tanya jawab adalah sistem yang mampu menjawab pertanyaan-pertanyaan yang diajukan oleh pengguna, dengan menggunakan teknologi natural language processing dan knowledge base yang terkait. Sistem ini biasanya digunakan dalam aplikasi chatbot atau virtual assistant untuk menjawab pertanyaan-pertanyaan dari pengguna secara otomatis.

Perbedaan utama antara kedua sistem ini adalah bahwa sistem rekomendasi memberikan rekomendasi produk atau layanan berdasarkan preferensi dan perilaku pengguna, sedangkan sistem tanya jawab menjawab pertanyaan pengguna dengan menggunakan basis pengetahuan yang terkait.

Proyek stbi kami (kelompok 5) tidak masuk dikeduanya. Proyek kami adalah mesin pencari atau search engine. Dimana search engine sangat bergantung pada kueri dengan melihat similaritas antara kueri dan dokumen yang ada pada koleksi.

1. A. Rangkuman dari Proyek STBI

**Introduction**

Banyaknya data yang tersedia membuat para ahli dan peneliti melakukan pengembangan pada bidang teknologi informasi. Pembuatan dan pengembangan sistem pencarian merupakan salah satu yang banyak diteliti dikarenakan Sistem pencarian sangat memudahkan untuk mengetahui berbagai macam informasi yang dibutuhkan. Sistem Temu Kembali Informasi (Information Retrieval) adalah proses menemukan materi yang biasanya dokumen dari data yang tidak terstruktur yang memenuhi kebutuhan informasi dimana Informasi memi- liki banyak jenis seperti teks (termasuk data numerik), audio, video, dan multimedia lainnya. dari sebuah kumpulan koleksi yang tersimpan pada komputer [2]. Salah satu teknologi sis- tem temu kembali yang secara terus menerus digunakan dan berkembang adalah content base image retrieval (CBIR) yang membahas mengenai proses temu kembali gambar berdasarkan pada content yang ada pada gambar tersebut [3].

Dalam sistem temu kembali informasi, diperlukan perankingan hasil pencarian dokumen sesuai dengan tingkat rel- evansinya dengan kata kunci yang dimasukkan. Vector Space Model (VSM) merupakan salah satu model dalam sistem temu kembali informasi yang mampu melakukan perankingan terhadap dokumen – dokumen yang relevan terhadap kata kunci yang dimasukkan. Vector Space Model (VSM) adalah metode untuk mengetahui tingkat kesamaan (similarity) term dengan cara pembobotan term. Vector Space Model (VSM) merepresentasikan suatu dokumen dan query dalam sebuah bentuk vektor yang memiliki jarak (magnitude) dan arah (direction).

**Methods**

Metode menggunakan vector space model dengan tahapan perhitungan adalah sebagai berikut.

* Menghitung bobot dokumen dan kueri menggunakan Tf-idf

Term frequency–inverse document frequency atau biasa sering disebut tf – idf adalah metode pembobotan kata dengan menghitung term (t) yang terdapat pada setiap dokumen (d), banyaknya term pada setiap dokumen ditunjukkan dengan notasi Term Frequency (tf). Kehadiran term total di semua dokumen dikumpulkan dalam notasi (df). Inverse document frequency (idf) adalah jumlah dokumen yang mengandung sebuah term didasarkan pada seluruh dokumen yang ada pada data set. Terdapat beberapa cara atau metode dalam melakukan pembobotan kata pada metode tf – idf yaitu melalui skema pembobotan query dan dokuman. Formula yang digunakan pada term frequency dihitung dengan menggunakan persamaan dibawah ini:

Dimana, N adalah total jumlah dokumen dalam koleksi, dfi = jumlah dokumen yang berisi term j, dan idfi = frekuensi kemunculan suatu istilah dalam dokumen i Dengan demikian rumus untuk tf – idf dapat dilihat pada Persamaan berikut.

Wij = tfij \* idfi

Setelah bobot (W) masing-masing dokumen diketahui, maka dilakukan proses sorting atau pengurutan dimana semakin besar nilai W, semakin besar tingkat similaritas dokumen tersebut terhadap kata kunci, demikian seba- liknya.

* Menghitung Jarak Query dengan Dokumen

Perhitungan jarak dokumen dan jarak query dilakukan setelah perhitungan bobot. Jika, Q adalah vektor query dan D adalah vektor dokumen, merupakan 2 buah vektor dalam ruang berdimensi-n, dan θ adalah sudut yang dibentuk oleh 2 vektor tersebut. Perhitungan jarak query dan dokumen menggunakan Persamaan berikut.

Dengan |Q| adalah jarak query dan Wiq adalah bobot query dokumen ke-i

* Menghitung Nilai cosine similarity

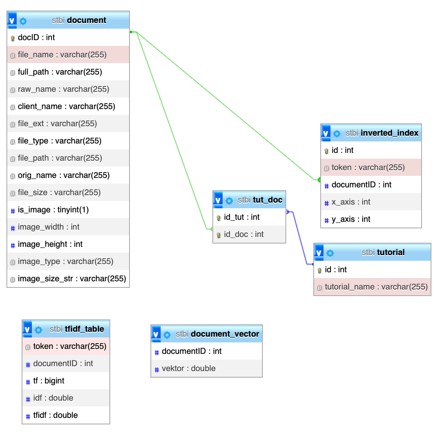
Melalui vector space model dan tf-idf weighting akan didapatkan representasi nilai numerik dokumen sehingga dapat dihitung kedekatan antar dokumen. Semakin dekat dua vektor didalam suatu vector space model, maka semakin mirip dua dokumen yang diwakili oleh dua vektor tersebut. Kemiripan antar dokumen dapat dihitung menggunakan suatu fungsi ukuran kemiripan (similarity measure). Ukuran ini memungkinkan perankingan dokumen sesuai dengan kemiripannya atau relevansinya terhadap query.

Pengukuran cosine similarity (menghitung nilai cosine sudut antara dua vektor) menggunakan Persamaan berikut.

**Result and Discussion**

*Database*

Tampilan dari *database* yang telah dibuat bisa diliat pada gambar 6.

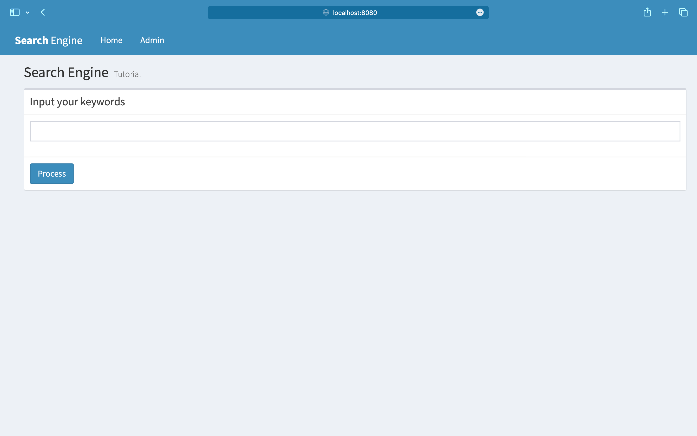


Gambar Rancangan *Database*

*Database* dalam membangun sistem temu balik informasi menggunakan databse mySql dengan 6 tabel yang saling berelasi.

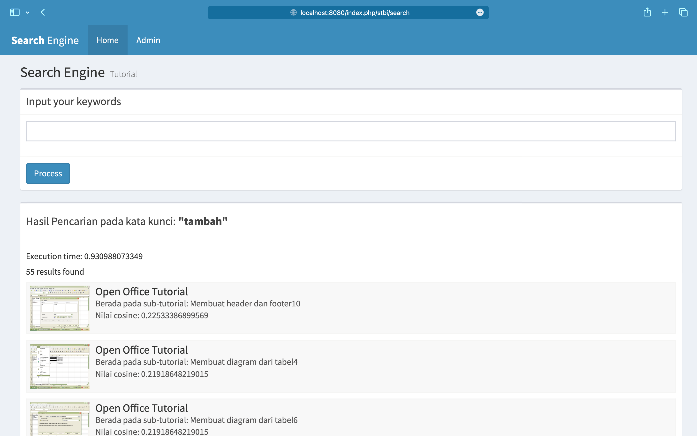
*Tampilan sistem pencarian*

Pada halaman utama menampilkan kotak pencarian untuk memasukkan *query* yang ingin dicari serta tombol cari.



Gambar Tampilan Awal Sistem Pencarian

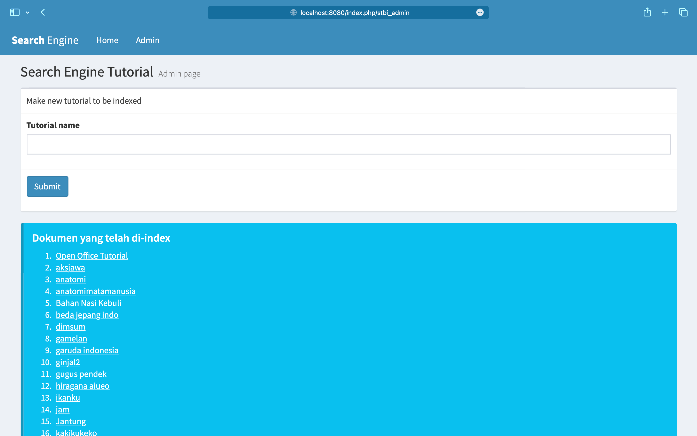
Setelah itu hasil pencarian akan menampilkan dokumen terkait dengan kata kunci yang dimasukkan.



Gambar Tampilan Hasil Pencarian

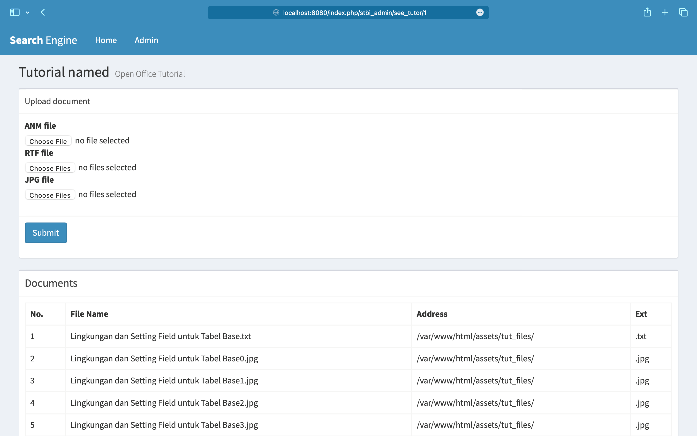
*Tampilan admin*

Pada halaman kelola data admin dapat melakukan tambah koleksi dokumen serta melihat data apa saja yang sudah ada di dalam sistem.



Gambar Halaman Utama Admin

Dokumen yang dapat dimasukkan pada koleksi bisa berupa rtf file, anm file, dan jpg.



Gambar Halaman Tambah Data

Untuk hasil evaluasi dari kuery 1 kata dapat dilihat pada tabel dibawah ini.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| No | Query | Runtime | Precision | Recall |
| 1 | musik | 0.5921080 | 100 | 37.5 |
| 2 | huruf | 0.5898170 | 100 | 16.666666 |
| 3 | hiragana | 0.5847039 | 100 | 40 |
| 4 | masak | 0.5654931 | 0 | 0 |
| 5 | nasi | 0.5742459 | 100 | 50 |
| 6 | manusia | 0.6029548 | 75 | 15 |
| 7 | angka | 0.9094681 | 100 | 8.1779053 |
| 8 | ginjal | 0.5639269 | 100 | 50 |
| 9 | tabel | 0.9389278 | 100 | 9.433962 |
| 10 | rambut | 0.5675659 | 100 | 50 |
| Rata-rata | | 0.648 | 87.5 | 27.67 |

Dari hasil pengujian precision dan recall yang dilakukan pada query yang terdiri dari 1 kata memiliki nilai rata-rata precision sebesar 87.5 dan nilai rata-rata recall sebesar 27.67. Untuk nilai rata-rata runtime sebesar 0.648

Adapun hasil pengujian precision dan recall yang dilakukan pada query yang terdiri dari 2 kata dapat dilihat pada Tabel berikut.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| No | Query | Runtime | Precision | Recall |
| 1 | Musik jawa | 1.0879840 | 75 | 33.33 |
| 2 | Penyakit menular | 1.0182321 | 0 | 0 |
| 3 | Belajar matematika | 1.0522079 | 0 | 0 |
| 4 | Lambung manusia | 1.1111891 | 0 | 50 |
| 5 | Ikan laut | 1.0558810 | 100 | 50 |
| 6 | Lambang Indonesia | 1.0504291 | 50 | 50 |
| 7 | Batu ginjal | 1.0574929 | 50 | 50 |
| 8 | Bahasa jepang | 1.0798890 | 100 | 40 |
| 9 | Burung Garuda | 1.0464317 | 100 | 50 |
| 10 | Pesawat terbang | 1.1254491 | 100 | 50 |
| Rata-rata | | 1.068 | 57.5 | 37.33 |

Dari hasil pengujian precision dan recall yang dilakukan pada query yang terdiri dari 2 kata memiliki nilai rata-rata precision sebesar 57.5 dan nilai rata-rata recall sebesar 37.33. Untuk nilai rata-rata runtime sebesar 1.068

Adapun hasil pengujian precision dan recall yang dilakukan pada query yang terdiri dari 3 kata dapat dilihat pada Tabel berikut.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| No | Query | Runtime | Precision | Recall |
| 1 | Struktur virus covid | 1.5175569 | 33.33333333 | 50 |
| 2 | Budaya wayang jawa | 1.5313150 | 100 | 41.66667 |
| 3 | Alat music gamelan | 1.8257658 | 100 | 55.555556 |
| 4 | Nasi arab kebuli | 1.6307568 | 0 | 50 |
| 5 | Perkalian matematika sederhana | 1.5602569 | 83.33333 | 83.33333 |
| 6 | Transportasi udara pesawat | 1.5167338 | 100 | 50 |
| 7 | Pembuluh darah manusia | 1.5598120 | 83.33333333 | 27.77778 |
| 8 | Tumbuhan yang mengandung vitamin | 1.5462398 | 100 | 100 |
| 9 | Bahasa jepang susah | 1.5550308 | 0 | 0 |
| 10 | Anatomi mata manusia | 1.7289500 | 3.03030303 | 50 |
| Rata-rata | | 1.597 | 60.30 | 50.83 |

Dari hasil pengujian precision dan recall yang dilakukan pada query yang terdiri dari 2 kata memiliki nilai rata-rata precision sebesar 60.30 dan nilai rata-rata recall sebesar 50.83. Untuk nilai rata-rata runtime sebesar 1.597.

**References**

1. A. Bahrudin, P. Permata, and J. Jupriyadi, Optimasi Arsip Penyimpanan Dokumen Foto Menggunakan Algoritma Kompresi Deflate (Studi Ka- susStudio Muezzart), J. Ilm. Infrastruktur Teknol. Inf., vol. 1, no. 2, pp. 14–18, 2020, doi: 10.33365/jiiti.v1i2.582.
2. Christioko, B. V., Daru, A. F, ”Sistem Temu Kembali Informasi Untuk Pencarian Judul Tugas Akhir Berbasis Kata Kunci”, Jurnal Pengemban- gan Rekayasa Dan Teknologi, 14(2), 41, 2019. Jurnal
3. Musrifah. (2018). Strategi Pengembangan Sistem Temu Kembali Infor- masi Berbasis Gambar (Content Based Image Retrieval System) di Per- pustakaan Perguruan Tinggi Kedokteran. JIPI (Jurnal Ilmu Perpustakaan Dan iNformasi), 3(1), 18–37
4. Sanjaya, F, ”Pemanfaatan Sistem Temu Kembali Informasi dalam Pen- carian Dokumen Menggunakan Metode Vector Space Model”, Journal of Information and Technology, 5(2), 147–153, 2018. Jurnal
5. D. Soyusiawaty and Y. Zakaria, “Book data content similarity detector with cosine similarity (case study on digilib.uad.ac.id),” Proceeding 2018 12th Int. Conf. Telecommun. Syst. Serv. Appl. TSSA 2018, 2018, doi: 10.1109/TSSA.2018.8708758.
6. P. P. Gokul, B. K. Akhil, and K. K. M. Shiva, “Sentence similarity detection in Malayalam language using cosine similarity,” RTEICT 2017 - 2nd IEEE Int. Conf. Recent Trends Electron. Inf. Commun. Technol. Proc., vol. 2018–Janua, pp. 221–225, 2017, doi: 10.1109/RTEICT.2017.8256590.
7. A. R. Lahitani, A. E. Permanasari, and N. A. Setiawan, “Cosine similarity to determine similarity measure: Study case in online essay assessment,” Proc. 2016 4th Int. Conf. Cyber IT Serv. Manag. CITSM 2016, 2016, doi: 10.1109/CITSM.2016.7577578
8. D. Kurniadi, S. F. C. Haviana, and A. Novianto, “Implementasi Algoritma Cosine Similarity pada sistem arsip dokumen di Universitas Islam Sultan Agung,” J. Transform., vol. 17, no. 2, p. 124, 2020, doi: 10.26623/transformatika.v17i2.1613.
9. P. dwi Nurfadila, A. P. Wibawa, I. A. E. Zaeni, and A. Nafalski, “Journal Classification Using Cosine Similarity Method on Title and Abstract with Frequency-Based Stopword Removal ,” Int. J. Artif. Intell. Res., vol. 3, no. 2, 2019, doi: 10.29099/ijair.v3i2.99.
10. M. Mhatre, D. Phondekar, P. Kadam, A. Chawathe, and K. Ghag, “Dimensionality Reduction for Sentiment Analysis using Pre-processing Techniques,” no. Iccmc, pp. 16–21, 2017.
11. Asian J. 2007. Effective Techniques for Indonesian Text Retrieval. PhD thesis School of Computer Science and Information Technology RMIT University Australia
12. Arifin, A.Z., I.P.A.K. Mahendra dan H.T. Ciptaningtyas. 2009. Enhanced Confix Stripping Stemmer and Ants Algorithm for Classifying News Document in Indonesian Language, Proceeding of International Conference on Information & Communication Technology and Systems (ICTS)
13. A. D. Tahitoe, D. Purwitasari. 2010. Implementasi Modifikasi Enhanced Confix Stripping Stemmer Untuk Bahasa Indonesia dengan Metode Corpus Based Stemming, Institut Teknologi Sepuluh Nopember (ITS) – Surabaya, 60111, Indonesia
14. D. K. Setiawan and H. Bunyamin, “Pemanfaatan Inverted Index Pada Proses Penelusuran Kesamaan Isi File Dokumen Pdf Tugas Akhir Mahasiswa,” SENTIKA, 2018
15. Wicaksosno, V. B., Saptono, R., dan Sihwi, S. W, ”Analisis Perbandin- gan Metode Vector Space Model dan Weighted Tree Similarity dengan Cosine Similarity pada kasus Pencarian Informasi Pedoman Pengobatan Dasar di Puskesmas”, Jurnal Teknologi Informasi ITSmart, 4(2), 73, 2016. Jurnal
16. C. D. Manning, P. Raghavan, and H. Schutze, An Introduction to Information Retrieval. 2009
17. N.P. Lestari, “Uji Recall and Precision Sistem Temu Kembali,” Univ. Airlangga, 2016.

B. Apakah proyek STBI yang kelompok kami buat merupakan multimedia information retrieval?

Proyek kami bukan termasuk multimedia information retrieval. Karena proyek kami hanya mengindeks label teks dari keterangan gambar yang ada pada data. Sehingga ketika dokumen di retrieve maka hasilnya keluar gambar tp informasi yg ada pada gambar tidak diolah melainka keterangan gambar yg berupa teks.