

PENCARIAN INFORMASI BERITA DETIKNEWS
MENGUNAKAN OKAPI BM25

*Diajukan Sebagai Syarat Untuk Menyelesaikan
Pendidikan Program Strata-1 Pada
Jurusan Teknik Informatika Fakultas Ilmu Komputer UNSRI*



Oleh :

Euis Annisa Rachman
NIM : 09021282025057

Jurusan Teknik Informatika
FAKULTAS ILMU KOMPUTER UNIVERSITAS SRIWIJAYA
2024

LEMBAR PENGESAHAN SKRIPSI

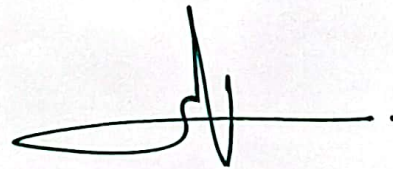
PENCARIAN INFORMASI BERITA DETIKNEWS MENGGUNAKAN
OKAPI BM25

Oleh:

Euis Annisa Rachman
NIM: 09021282025057

Indralaya, 31 Mei 2024

Pembimbing I



Dr. Abdiansah, S.Kom., M.Cs.
NIP. 198410012009121005

Mengetahui,
Ketua Jurusan Teknik Informatika



Dr. M. Fachrurrozi, S.Si., M.T.
NIP. 198005222008121002

TANDA LULUS UJIAN SIDANG SKRIPSI

Pada hari Senin tanggal 31 Mei 2024 telah dilaksanakan ujian komprehensif skripsi oleh Jurusan Teknik Informatika Fakultas Ilmu Komputer Universitas Sriwijaya

Nama : Euis Annisa Rachman
NIM : 09021282025057
Judul : Pencarian Informasi Berita DetikNews Menggunakan Okapi BM25

Dan dinyatakan **LULUS**.

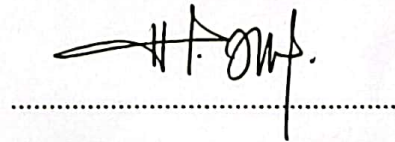
1. Ketua Penguji

Osvari Arsalan, M.T.
NIP. 198806282018031001



2. Penguji

Annisa Darmawahyuni, M.Kom.
NIP. 199006302023212044



3. Pembimbing

Dr. Abdiansah, S.Kom., M.Cs
NIP. 198410012009121005



Mengetahui,

Ketua Jurusan Teknik Informatika



M. Faehurrozi, S.Si., M.T.
NIP. 198005222008121002

HALAMAN PERNYATAAN

Yang bertanda tangan di bawah ini:

Nama : Euis Annisa Rachman

NIM : 09021282025057

Jurusan : Teknik Informatika

Judul Skripsi : Pencarian Informasi Berita DetikNews Menggunakan Okapi BM25

Hasil pengecekan Software iThenticate/Turnitin : 18%

Menyatakan bahwa laporan proyek saya merupakan hasil karya sendiri dan bukan hasil penjiplakan/plagiat. Apabila ditemukan unsur penjiplakan/plagiat dalam laporan proyek ini, maka saya bersedia menerima sanksi akademik dari Universitas Sriwijaya sesuai dengan ketentuan yang berlaku.

Demikian pernyataan ini saya buat dengan sebenarnya dan tidak ada paksaan oleh siapa pun.



Indralaya, 31 Mei 2024



Euis Annisa Rachman
NIM. 09021282025052

MOTTO DAN PERSEMBAHAN

“life can be heavy, especially if you try to carry it all at once, part of growing up and moving into new chapters of your life is about catch or release. what I mean by that is, knowing what things to keep and what things to release. you can't carry all things”
-Taylor Swift-

Kupersembahkan karya tulis ini kepada:

- Allah Subhanahu Wa Ta 'ala dan Rasulullah SAW
- Keluarga saya
- Dosen Pengampu, Pembimbing Akademik dan Skripsi Saya
- Universitas Sriwijaya
- Sahabat dan teman-teman seperjuangan
- Sastrawan Indonesia

ABSTRACT

In the process of searching for information, internet users generally rely on search engines. The internet plays a significant role in society, particularly in the dissemination of news through online media. News serves as a source of information about events or opinions that are important and interesting enough to be published online. Information retrieval is the science of studying methods to retrieve stored information from various sources. This study implements scraping to collect news data from the DetikNews website and uses the Okapi BM25 method for news information retrieval. Okapi BM25 is a probabilistic model that effectively ranks terms within documents for relevant searches. Based on the research conducted, the Okapi BM25 algorithm has a relevance percentage of 48.88%, which is higher compared to the DetikNews search engine with a relevance percentage of 23.33%.

Keywords: Scraping, Information Retrieval, Search Engine, Okapi BM25

ABSTRAK

Pada proses pencarian sebuah informasi, pengguna internet umumnya menggunakan bantuan mesin pencari. Internet memiliki kontribusi yang besar di dalam masyarakat, terutama dalam proses penyebaran berita dengan menggunakan internet atau media *online*. Berita merupakan sebuah sumber informasi mengenai suatu peristiwa atau pendapat yang memiliki nilai penting dan menarik untuk dipublikasikan melalui media online. Pencarian informasi adalah ilmu yang mempelajari metode untuk mendapatkan kembali informasi yang tersimpan dari berbagai sumber. Pada penelitian ini menerapkan *scraping* untuk mengumpulkan data berita pada situs berita DetikNews dan menggunakan metode Okapi BM25 untuk pencarian informasi berita. Okapi BM25 adalah model probablistik yang secara efektif memberi peringkat istilah dalam dokumen untuk pencarian yang relevan. Berdasarkan penelitian yang dilakukan, algoritma Okapi BM25 memiliki persentase relevansi dengan nilai 48.88% lebih tinggi dibandingkan dengan mesin pencari DetikNews dengan nilai persentase relevansi sebesar 23.33%.

Kata Kunci: *Scraping, Information Retrieval, Pencarian, Okapi BM25*

KATA PENGANTAR

Puji syukur penulis panjatkan kehadiran Allah SWT, Tuhan Semesta Alam atas berkat, rahmat, rahim dan karunia-Nya yang telah diberikan kepada penulis sehingga Tugas Akhir berjudul “Pencarian Informasi Berita DetikNews Menggunakan Okapi BM25” dapat disusun dengan baik sebagai syarat dalam menyelesaikan studi Strata-1 program studi Teknik Informatika, Fakultas Ilmu Komputer Universitas Sriwijaya. Pada kesempatan ini penulis ingin mengucapkan ucapan terima kasih kepada semua pihak yang memberikan dukungan, motivasi dan bimbingan selama penyusunan Tugas Akhir dan penelitian. Secara khusus ucapan terima kasih ini ditujukan kepada:

1. Allah Subhanallahu Wa Ta’ala atas segala berkah, rahmat, rahim dan karunia-Nya.
2. Keluarga terutama kedua orang tua tercinta, kakak Adit, Adek Nayla, Nenek yang selalu memberikan dukungan positif dan doa kepada penulis. Terima kasih juga karena telah memberikan kepercayaan ini tanpa tekanan kepada penulis.
3. Prof. Dr. Erwin, S.Si., M.Si. selaku Dekan Fakultas Ilmu Komputer Universitas Sriwijaya
4. Bapak Dr. M. Fachrurrozi, S.Si., M.T. selaku Ketua Jurusan Teknik Informatika Fakultas Ilmu Komputer Universitas Sriwijaya.
5. Bapak Dr. Abdiansah, S.Kom., M.Cs. selaku Dosen Pembimbing saya yang senantiasa membagikan ilmu, membimbing, memberi arahan dan memotivasi saya untuk menyelesaikan Tugas Akhir ini dengan baik.
6. Ibu Desty Rodiah, S.Kom., M.T. selaku Dosen Pembimbing Akademik yang senantiasa memberikan saran, arahan dan bimbingan kepada saya selama masa perkuliahan.
7. Seluruh Dosen Fakultas Ilmu Komputer, khususnya Dosen Jurusan Teknik Informatika, yang telah banyak memberikan bekal, arahan, dan saran selama masa perkuliahan.

8. Seluruh Staf Administrasi dan Pegawai yang telah membantu dalam urusan administrasi.
9. Sahabat saya Arya Yunanta yang telah membantu dan memberikan dukungan dari awal hingga akhir dalam proses pengerjaan Tugas Akhir ini.
10. Teman – temanku kelas REG C, Tiba-tiba epic (Diva, Enot, Gandi, Dzaki, Nabil, Adit), Adhela, Putmar, kak Caca, dan Liza yang telah membantu dan memberikan dukungan secara langsung maupun tidak langsung dalam proses pengerjaan Tugas Akhir ini.
11. Terima kasih untuk diri sendiri, karena telah bertahan sampai dititik ini. Terima kasih telah begitu sabar serta tidak menyerah meskipun keadaan begitu sulit. *Long story short, I survived.*

Penulis secara penuh menyadari akan kekurangan dalam penyusunan Tugas Akhir ini. Kekurangan ini semata – mata adalah karena keterbatasan pengetahuan juga pengalaman penulis. Oleh karena itu, kritik dan saran yang membangun sangat diharapkan untuk menyempurnakan Tugas akhir ini sehingga dapat membawa manfaat lebih banyak dan luas. Akhir kata, semoga Tugas Akhir ini bermanfaat bagi kita semua.

Indralaya, 31 Mei 2024
Penulis,



Euis Annisa Rachman

DAFTAR ISI

LEMBAR PENGESAHAN SKRIPSI.....	i
TANDA LULUS UJIAN SIDANG SKRIPSI.....	ii
HALAMAN PERNYATAAN	iii
MOTTO DAN PERSEMBAHAN	iv
ABSTRACT	v
ABSTRAK	vi
DAFTAR ISI	ix
DAFTAR TABEL.....	xii
DAFTAR GAMBAR	xiii
BAB 1 PENDAHULUAN	I-1
1.1 Pendahuluan.....	I-1
1.2 Latar Belakang Masalah	I-1
1.3 Rumusan Masalah.....	I-4
1.5 Manfaat Penelitian	I-5
1.6 Batasan Masalah	I-5
1.7 Sistematika Penulisan	I-5
1.8 Kesimpulan	I-7
BAB II KAJIAN LITERATUR.....	II-1
2.1 Pendahuluan.....	II-1
2.2 Landasan Teori.....	II-1
2.2.1 Document Object Model	II-1
2.2.2 Web Scraping	II-2
2.2.3 Search Engine.....	II-4
2.2.4 Okapi BM25	II-5
2.2.5 ChatGPT 4.0.....	II-8
2.2.6 Rational Unified Process.....	II-8
2.3 Penelitian Lain yang relevan.....	II-10
BAB III METODOLOGI PENELITIAN.....	III-1
3.1 Pendahuluan.....	III-1
3.2 Pengumpulan Data.....	III-2

3.2.1	Jenis dan Sumber Data.....	III-2
3.2.2	Metode Pengumpulan Data.....	III-3
3.3	Tahapan Penelitian.....	III-3
3.3.1	Kerangka Kerja.....	III-6
3.3.2	Kriteria Pengujian.....	III-7
3.3.3	Format Data Pengujian.....	III-7
3.3.4	Menentukan Alat Bantu Penelitian.....	III-9
3.3.5	Pengujian Penelitian.....	III-9
3.3.6	Menganalisis Hasil Pengujian Membuat Kesimpulan.....	III-9
3.4	Metode Pengembangan Perangkat Lunak.....	III-11
3.4.1	Fase Inception.....	III-11
3.4.2	Fase Elaboration.....	III-11
3.4.3	Fase Construction.....	III-11
3.4.4	Fase Transition.....	III-12
3.5	Kesimpulan.....	III-12
BAB IV PENGEMBANGAN PERANGKAT LUNAK.....		IV-1
4.1	Pendahuluan.....	IV-1
4.2	Fase Insepsi.....	IV-1
4.2.1	Pemodelan Bisnis.....	IV-1
4.2.2	Kebutuhan Sistem.....	IV-2
4.2.3	Analisis dan Desain.....	IV-2
4.3	Fase Eleborasi.....	IV-8
4.3.1	Pemodelan Bisnis.....	IV-9
4.3.2	Kebutuhan sistem.....	IV-11
4.3.3	Diagram.....	IV-11
4.4	Fase Konstruksi.....	IV-13
4.4.1	Kebutuhan.....	IV-14
4.4.2	Implementasi.....	IV-14
4.5	Fase Transisi.....	IV-16
4.5.1	Pemodelan Sistem.....	IV-16
4.5.2	Kebutuhan.....	IV-16
4.5.3	Analisis dan Perancangan.....	IV-17

4.5.4	Implementasi	IV-518
4.6	Kesimpulan	IV-19
BAB V HASIL DAN ANALISIS PENELITIAN		V-1
5.1	Pendahuluan.....	V-1
5.2	Hasil Program	V-1
5.3	Hasil Pengujian	V-2
5.3.1	Konfigurasi Pengujian.....	V-2
5.3.2	Pengujian Dengan Kata Kunci Normal.....	V-3
5.3.2	Pengujian Dengan Kata Kunci Tanpa Spasi.....	V-5
5.3.3	Pengujian Dengan Kata Kunci Terbalik.....	V-6
5.3.4	Pengujian Dengan Posisi Kata Kunci Terbalik dan Tanpa Spasi ...	V-8
5.3.5	Pengujian Dengan Kata Kunci Kurang Ketik	V-8
5.3.6	Pengujian Dengan Kata Kunci Kelebihan Ketik.....	V-10
5.3.7	Pengujian Dengan Kata Kunci Kiasan.....	V-12
5.3.8	Pengujian Dengan Kata Kunci Slangword.....	V-14
5.3.9	Pengujian Dengan Kata Kunci Multi Bahasa.....	V-16
5.4	Analisis Hasil Penelitian	V-18
5.5	Kesimpulan	V-21
BAB VI KESIMPULAN DAN SARAN		VI-1
6.1	Pendahuluan.....	VI-1
6.2	Kesimpulan	VI-1
6.3	Saran	VI-2
DAFTAR PUSTAKA		xv
LAMPIRAN		Error! Bookmark not defined.

DAFTAR TABEL

Tabel III- 1. Contoh Dataset Berita.....	III-2
Tabel III- 2. Rancangan Tabel Pengujian Pencarian Berita	III-8
Tabel III- 3. Hasil Analisis Berita Relevan	III-10
Tabel IV- 1. Kebutuhan Fungsional.....	IV-2
Tabel IV- 2. Kebutuhan Non Fungsional.....	IV-2
Tabel IV- 3. Contoh data teks berita.....	IV-4
Tabel IV- 4. Hasil Stop Word dan Stemming MSS data teks berita.	IV-5
Tabel IV- 5. Hasil Tokenisasi data teks artikel.	IV-5
Tabel IV- 6. Definisi Aktor	IV-7
Tabel IV- 7. Definisi Use case.....	IV-7
Tabel IV- 8. Skenario Pencarian informasi berita	IV-7
Tabel IV- 9. Implementasi Kelas	IV-15
Tabel IV- 10. Rencana Pengujian Melakukan Pencarian Informasi Berita	IV-17
Tabel IV- 11. Pengujian Use Case Proses Pencarian Informasi Berita.....	IV-51
Tabel V- 1. Hasil Pengujian Dengan Kata Kunci Normal.....	V-4
Tabel V- 2. Hasil Pencarian Dengan Kata Kunci Tanpa Spasi	V-5
Tabel V- 3. Hasil Pencarian Dengan Kata Kunci Terbalik	V-6
Tabel V- 4. Hasil Pencarian Posisi Kata Kunci Terbalik dan Tanpa Spasi	V-8
Tabel V- 5. Hasil Pencarian Dengan Kata Kunci Kurang Ketik.....	V-9
Tabel V- 6. Hasil Pencarian Kata Kunci Kelebihan Ketik.....	V-10
Tabel V- 7. Hasil Pencarian Dengan Kata Kunci Kiasan	V-12
Tabel V- 8. Hasil Pencarian Kata Kunci Slangword.....	V-14
Tabel V- 9. Hasil Pencarian Kata Kunci Multi Bahasa	V-16
Tabel V- 10. Persentase Relevansi.....	V-19

DAFTAR GAMBAR

Gambar II-1. Rational Unified Process (RUP) Architecture (Manalil, 2010)..	II-10
Gambar III- 1. Tahapan Penelitian.....	III-4
Gambar III- 2. Kerangka Kerja Penelitian.....	III-6
Gambar IV- 1. Use case Diagram.....	IV-6
Gambar IV- 2. Rancangan Antarmuka Pencarian Berita.....	IV-10
Gambar IV- 3. Rancangan Antarmuka Hasil Pencarian Berita	IV-10
Gambar IV- 4. Diagram Aktivitas Pencarian Berita.....	IV-12
Gambar IV- 5. Diagram Squence Pencarian Informasi Berita	IV-13
Gambar IV- 6. Diagram Kelas.....	IV-14
Gambar IV- 7. Desain Tampilan Pencarian	IV-15
Gambar IV- 8. Desain Tampilan Hasil Pencarian.....	IV-16
Gambar V- 1. Tampilan Fitur Pencarian Berita	V-2
Gambar V- 2. Grafik Persentase Relevansi	V-21

BAB I

PENDAHULUAN

1.1 Pendahuluan

Bab ini akan menjelaskan tentang pokok-pokok pikiran dalam penelitian yang mencakup latar belakang masalah, rumusan masalah, tujuan penelitian, manfaat penelitian, dan batasan masalah yang akan menjadi acuan dalam kajian penelitian ini.

1.2 Latar Belakang Masalah

Dalam era digital saat ini, internet memiliki kontribusi yang besar di dalam masyarakat. Salah satunya yaitu proses penyebaran berita dengan menggunakan internet atau media *online*. Berita merupakan sebuah sumber informasi mengenai suatu peristiwa atau pendapat yang memiliki nilai penting, menarik untuk dipublikasikan melalui media *online*. Salah satu media *online* yang terus melakukan pembaruan ialah Detik. Detik adalah salah satu situs yang menyajikan artikel dan berita pada media *online* di Indonesia. Detik memiliki *platform* bernama DetikNews yang secara khusus fokus pada berita-berita terkini. Penyebaran berita melalui media *online* menjadi sebuah berita terpercaya dalam mengetahui sebuah informasi, tetapi ada beberapa kekurangan pada media *online* berita khususnya pada pencarian.

Dalam proses pencarian sebuah informasi, pengguna internet umumnya menggunakan bantuan mesin pencari. Perkembangan ini meningkatkan penyimpanan dokumen digital dengan bertambahnya jumlah artikel yang

diterbitkan dan bertambahnya ukuran data membuat pengguna semakin sulit menemukan artikel yang relevan, sementara pengguna ingin menerima informasi dengan cepat dan akurat (Eldira *et al.*, 2011).

Pencarian informasi (*information retrieval*) adalah ilmu yang mempelajari metode untuk mendapatkan kembali informasi yang tersimpan dari berbagai sumber atau koleksi informasi yang relevan dan dibutuhkan berdasarkan *keyword* pada *query* yang diberikan (Pardede *et al.*, 2018). Penerapan *information retrieval* diharapkan membuat pencarian informasi lebih mudah dan akurat. Keberhasilan sistem pencarian informasi sangat tergantung pada efektivitas fungsi pencarian yang dapat mencocokkan proporsional antara kata kunci dan dokumen (Russell & Norvig, 2010). Dibantu dengan metode Okapi BM25, untuk menentukan bagaimana meningkatkan kemampuan pencarian informasi berita dengan mengimplementasikan algoritma Okapi BM25. Okapi BM 25 adalah model probabilistik yang didasarkan pada teori probabilitas. Model ini secara efektif memberi peringkat istilah dalam dokumen untuk pencarian yang relevan. Okapi BM25 memiliki unsur standar TF-IDF (*term frequency-inverse document frequency*) yang merupakan ukuran statistik yang menggambarkan pentingnya suatu istilah terhadap sebuah dokumen dalam sebuah kumpulan atau korpus. Model ini menggunakan TF-IDF (*term frequency-inverse document frequency*) untuk pembobotan istilah dan normalisasi panjang dari suatu dokumen yang diberikan (Garcia, 2016).

Web Scraping adalah sebuah teknik dalam pengambilan informasi dari halaman web dalam bahasa *markup* seperti DOM-HTML (*Document Object Model*)

yang merupakan suatu antarmuka pemrograman aplikasi (API) atau model penulisan standar yang digunakan untuk dokumen XML (*Extensible Markup Language*) atau HTML (*HyperText Markup Language*). *Web scraping* sebelumnya dikenal sebagai *screen scraping*, yaitu praktik mengekstraksi data dari representasi visual yang terjadi pada masa-masa awal komputasi sekitar tahun 1960-1980-an dan biasanya mengarah ke "terminal" berbasis teks sederhana (Broucke & Baesens, 2018). Teknik ini melibatkan sedikit manusia dan merupakan agen penggalian untuk mengunduh, mengurai dan mengorganisir data dari web secara otomatis. Data otomatis dalam *web scraping* mengacu pada data yang dikumpulkan secara otomatis oleh program komputer sehingga kelebihanannya dalam efisiensi waktu lebih cepat, data yang dihasilkan lebih akurat, dapat mengumpulkan data dalam jumlah besar, dan menghasilkan data yang konsisten. Dibandingkan dengan mengambil data secara manual, pengumpulan data manual dapat memakan waktu dan rentan terhadap kesalahan manusia.

Teknik *Web Scraping* ini telah diimplementasikan pada beberapa penelitian dan metode BM25 dikembangkan oleh beberapa peneliti. Sebagai contoh, (Yondra *et al.*, 2022) melakukan penelitian tentang pembangunan aplikasi *web scraping* yang menggunakan teknik pemrosesan paralel dan multithreading untuk mengumpulkan data dari beberapa situs web *e-commerce* di Indonesia. Hasil pengujian menunjukkan bahwa proses *scraping* paralel dengan multithreading dapat mempercepat akuisisi produk dari beberapa situs secara bersamaan (Yondra *et al.*, 2022). Speriosu dan Tashiro (2017) dalam penelitian mereka menyimpulkan bahwa model Okapi BM25 menunjukkan kinerja yang lebih unggul dibandingkan

dengan Algoritma *Language Modeling*. Penelitian sebelumnya yang dilakukan oleh Sari dan Adriani (2014) telah membuktikan bahwa BM25 dapat diidentifikasi dengan baik sebagai fitur untuk proses pemeringkatan dan pada penelitian (Bintana & Agustian, 2012) tentang Sistem Temu Kembali Informasi menggunakan metode Okapi BM25 menyimpulkan beberapa kata kunci yang diujikan performa sistem dinilai baik dengan *precision* 74% - 92% dan *recall* 100%. Dengan model Okapi BM25 dapat digunakan untuk situs-situs pencarian lokal organisasi, seperti perpustakaan digital, arsip dokumen hukum, teknik, dan sebagainya (Bintana & Agustian, 2012). Jadi, peneliti akan melakukan penelitian pencarian informasi berita yang mana akan menggunakan metode Okapi BM25 dengan data otomatis scraping pada situs DetikNews.

1.3 Rumusan Masalah

Berdasarkan latar belakang yang telah dijelaskan, rumusan masalah dalam penelitian ini adalah sebagai berikut:

1. Bagaimana penerapan metode Okapi BM25 untuk pencarian informasi berita Detiknews?
2. Bagaimana kinerja Okapi BM25 dalam melakukan pencarian informasi berita DetikNews?

1.4 Tujuan Penelitian

Tujuan penelitian ini adalah sebagai berikut:

1. Menerapkan Okapi BM25 untuk pencarian informasi berita DetikNews

2. Mengetahui kinerja Okapi BM25 untuk pencarian informasi berita Detiknews.

1.5 Manfaat Penelitian

Manfaat dari penelitian ini adalah sebagai berikut:

1. Mempermudah pencarian informasi berita berdasarkan kata kunci yang dimasukkan.
2. Memberikan rekomendasi berita sesuai dengan kata kunci yang dimasukkan.
3. Menambah referensi informasi berita yang sejenis.

1.6 Batasan Masalah

Batasan masalah dalam penelitian ini sebagai berikut:

1. Data berita yang digunakan yaitu kumpulan berita pada DetikNews dengan rentang waktu 1 Januari – 27 April 2024, dengan jumlah 27.249 baris yang memiliki kolom yaitu berisi judul berita, deskripsi berita, dan tautan berita.

1.7 Sistematika Penulisan

Sistematika penulisan dalam penelitian ini adalah sebagai berikut:

BAB I. PENDAHULUAN

Bab ini akan menjelaskan tentang pokok-pokok pikiran dalam penelitian yang mencakup latar belakang masalah, rumusan masalah, tujuan penelitian,

manfaat penelitian, dan batasan masalah yang akan menjadi acuan dalam kajian penelitian ini.

BAB II. KAJIAN LITERATUR

Bab ini akan membahas tentang landasan teori yang digunakan dalam penelitian ini, yang diantaranya adalah *Document Object Model*, *Web Scraping*, *Search Engine*, Okapi *BM25*, *RUP* dan penelitian yang relevan.

BAB III. METODOLOGI PENELITIAN

Bab ini akan membahas tentang pengumpulan data, metode yang digunakan dalam mengumpulkan data dan kerangka kerja penelitian yang akan dibahas secara rinci.

BAB IV. PENGEMBANGAN PERANGKAT LUNAK

Bab ini akan membahas tentang perancangan pengembangan perangkat lunak. Diawali dengan analisis kebutuhan, perancangan dan konstruksi, serta diakhiri dengan tahap pengujian yang untuk memastikan kesesuaian sistem yang dikembangkan dengan rancangan dan kebutuhan penelitian.

BAB V. HASIL DAN ANALISIS PENELITIAN

Bab ini akan dibahas hasil pengujian berdasarkan langkah – langkah yang telah direncanakan sebelumnya. Analisis diberikan sebagai dasar kesimpulan yang akan diambil dari penelitian ini.

BAB VI. KESIMPULAN DAN SARAN

Bab ini akan dibahas tentang kesimpulan dari seluruh uraian yang telah disajikan pada bab sebelumnya serta saran yang diberikan berdasarkan penelitian yang telah dilakukan.

1.8 Kesimpulan

Bab ini telah dibahas mengenai latar belakang masalah, rumusan masalah, tujuan penelitian, manfaat penelitian, batasan masalah, dan sistematika penelitian yang akan dijadikan sebagai pokok pikiran peneliti

DAFTAR PUSTAKA

- Anwar, A. (2014). A Review of RUP (Rational Unified Process). In Ashraf Anwar International Journal of Software Engineering (IJSE) (Issue 5).
- Bintana, R.R., Agustian, S., 2012. Penerapan Model Okapi BM25 Pada Sistem Temu Kembali Informasi.
- Bonetti, L., 2021. Design and implementation of a real-world search engine based on Okapi BM25 and SentenceBERT.
- Caramancion, K.M., 2023. News Verifiers Showdown: A Comparative Performance Evaluation of ChatGPT 3.5, ChatGPT 4.0, Bing AI, and Bard in News Fact-Checking, in: 2023 IEEE Future Networks World Forum (FNWF). Presented at the 2023 IEEE Future Networks World Forum (FNWF), IEEE, Baltimore, MD, USA, pp. 1–6. <https://doi.org/10.1109/FNWF58287.2023.10520446>
- Eldira, H., Martiana, E. dan Rosyid, N, 2011. . Web Mining untuk Pencarian Dokumen Bahasa Inggris Menggunakan Hill Climbing Automatic Clustering. Eepis, 2(Pencarian Dokumen).
- F. Sudirjo, P. Diawati, Y. Riady, A. M. A. Ausat, and S. Suherlan, “The Role of ChatGPT in Enhancing the Information Search and Decision-Making Process of Travellers,” Jurnal Minfo Polgan, vol. 12, no. 2, pp. 500–507, 2023, doi: <https://doi.org/10.33395/jmp.v12i2.12443>.
- Frinta, K., 2019. Pencarian Berita Berbahasa Indonesia Menggunakan Metode BM25.
- Garcia, E., 2016. A Tutorial on the BM25F Model.
- Gunawan, R., Rahmatulloh, A., Darmawan, I., Firdaus, F., 2019. Comparison of Web Scraping Techniques : Regular Expression, HTML DOM and Xpath, in: Proceedings of the 2018 International Conference on Industrial Enterprise and System Engineering (IcoIESE 2018). Presented at the Proceedings of the 2018 International Conference on Industrial Enterprise and System Engineering (IcoIESE 2018), Atlantis Press, Yogyakarta, Indonesia. <https://doi.org/10.2991/icoiese-18.2019.50>
- Hakim, Z., Rizky, R., 2018. Analisis Perancangan Sistem Informasi Pembuatan Paspor Di Kantor Imigrasi Bumi Serpong Damai Tangerang Banten Menggunakan Metode Rational Unified Process 6.

- Jarmul, K., Lawson, R., 2017. Python Web Scraping - Second Edition, 2nd ed. ed. Packt Publishing, Birmingham.
- Josi, A., Abdillah, L.A., 2014. Penerapan Teknik Web Scraping Pada Mesin Pencari Artikel Ilmiah.
- Kruchten, P. (2000). The Rational Unified Process An Introduction Second Edition
- Manalil, J. (2010). Rational Unified Process. Computer, August.
- Manning, C.D., Raghavan, P., Schütze, H., 2008. Introduction to information retrieval. Cambridge University Press, New York.
- Pardede, J., Husada, M.G., Riansyah, R., 2018. Implementasi dan Perbandingan Metode Okapi BM25 dan PLSA Pada Aplikasi Information Retrieval 1-10.
- Riyaldi, T., 2009. Jenis-jenis Search Engine. hal.1-5.
- Russell, S., & Norvig, P. (2010). Artificial Intelligence: A Modern Approach, 3rd Edition. New Jersey: Prentice Hall.
- Rosario, B.Mario., Pratama, Y., & Fachruddin. (2017). Penerapan Web *Scraping* Pada Website Company Profile. Sekolah Tinggi Ilmu Komputer Dinamika Bangsa, 4(4), 37–43.
- Sahria Y. (2022) implementasi teknik web scraping pada Jurnal SINTA untuk analisis topik penelitian kesehatan Indonesia. University Research Colloquium. (pp 297–306). Yogyakarta.
- Sari, S. dan Adriani, M., 2014. Learning to Rank for Determining Relevant Document in Indonesian-English Cross Language Information Retrieval using BM25. hal.309-314.
- Speriosu, Michael, dan Tetsuya Tashiro. “Comparison of Okapi BM25 and Language Modeling Algorithms for NTCIR-6,” Justsystems Corporation, 14 September 2006. (Rev. 6 Desember 2007)
- Vanden Broucke, S., Baesens, B., 2018. Practical Web Scraping for Data Science. Apress, Berkeley, CA. <https://doi.org/10.1007/978-1-4842-3582-9>
- Whissell, J.S., Clarke, C.L.A., 2011. Improving document clustering using Okapi BM25 feature weighting. Inf Retrieval 14, 466–487. <https://doi.org/10.1007/s10791-011-9163-y>

Yondra, A.S., Triyanto, D., Bahri, S., 2022. Implementasi Web Scraping Untuk Mengumpulkan Informasi Produk Dari Situs E-commerce Dan Marketplace Dengan Teknik Pemrosesan Pararel 10.