

EKSTRAKSI KATA KUNCI PADA BAHASA INDONESIA MENGGUNAKAN METODE YAKE

Diajukan Sebagai Syarat Untuk Menyelesaikan
Pendidikan Program Strata-1 Pada
Jurusan Teknik Informatika



Oleh :

Gerald Plakasa
NIM : 09021381924100

**Jurusan Teknik Informatika
FAKULTAS ILMU KOMPUTER UNIVERSITAS SRIWIJAYA
2022**

LEMBAR PENGESAHAN SKRIPSI

EKSTRAKSI KATA KUNCI PADA BAHASA INDONESIA MENGGUNAKAN METODE YAKE

Oleh :

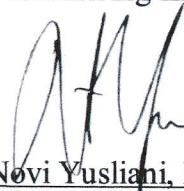
Gerald Plakasa
NIM : 09021381924100

Indralaya, 22 November 2022

Pembimbing I


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

Pembimbing II,


Novi Yusliani, M.T.
NIP. 198211082012122001

Mengetahui,



TANDA LULUS UJIAN KOMPREHENSIF

Pada hari Rabu tanggal 21 Desember 2022 Telah dilaksanakan ujian komprehensif skripsi oleh Jurusan Teknik Informatika Fakultas Ilmu Komputer Universitas Sriwijaya

Nama : Gerald Plakasa
NIM : 09021381924100
Judul : Ekstraksi Kata Kunci Pada Bahasa Indonesia Menggunakan Metode Yake

dan dinyatakan LULUS.

1. Ketua Pengaji

Mastura Diana Marieska, S.T., M.T.
NIP. 198603212018032001

2. Pengaji I

Alvi Syahrini Utami, M.Kom.
NIP. 197812222006042003

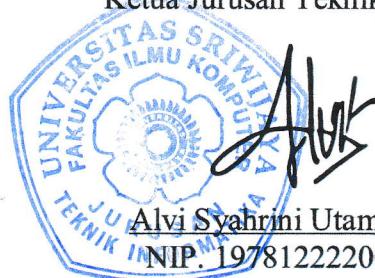
3. Pembimbing I

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

4. Pembimbing II

Novi Yusliani, M.T.
NIP. 198211082012122001

Mengetahui,
Ketua Jurusan Teknik Informatika



Alvi Syahrini Utami, M.Kom.
NIP. 197812222006042003

HALAMAN PERNYATAAN

Yang bertanda tangan di bawah ini:

Nama : Gerald Plakasa
NIM : 09021381924100
Program Studi : Teknik Informatika
Judul Skripsi : Ekstraksi Kata Kunci Pada Bahasa Indonesia
Menggunakan Metode Yake

Hasil pengecekan *Software iThenticate/Turnitin*: 7%

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 dari pihak mana pun.



Indralaya, 23 November 2022



Gerald Plakasa
NIM. 09021381924100

MOTTO DAN PERSEMBAHAN

Motto:

“Tidak ada rahasia untuk sukses. Ini adalah hasil dari persiapan, kerja keras, dan belajar dari kegagalan”

(Colin Powell)

Kupersembahkan Karya Tulis ini kepada :

- Allah SWT
- Kedua orang tua dan saudara saya
- Fakultas Ilmu Komputer
- Universitas Sriwijaya

ABSTRACT

Publication of current research results continues to grow. When a publication is accepted for submission to the reviewer, the publication must be appropriate to the field of knowledge possessed by the reviewer. Keyword Extraction can be one of the solutions to retrieve the most relevant words based on titles and abstracts from scientific publications. One of the Statistical Based methods for keyword extraction is Yake (Yet Another Keyword Extraction). This research will extract keywords in Indonesian language using the Yake method. The dataset used to conduct this research is 100 scientific publications from the websites jtiik, jatisi, and jepin, where the topic is limited to Computer Science. Based on the research conducted, the configuration of the Levenshtein Distance parameter has an influence on the resulting keywords. The evaluation results of this study obtained an accuracy value of 97.05% and an f-score value of 54.1% with the Levenshtein distance parameter < 2.

Key word : Keyword Extraction, Yake, Levenshtein Distance

ABSTRAK

Publikasi hasil penelitian saat ini terus berkembang. Ketika publikasi diterima untuk disampaikan kepada kepada *reviewer*, publikasi haruslah tepat dengan bidang ilmu yang dimiliki oleh *reviewer* tersebut. Ekstraksi Kata Kunci dapat menjadi salah satu solusi mengambil kata yang paling relevan berdasarkan judul dan abstrak dari publikasi ilmiah. Salah satu metode *Statistical Based* untuk Ekstraksi Kata Kunci adalah Yake (*Yet Another Keyword Extraction*). Penelitian ini akan melakukan Ekstraksi Kata Kunci pada Bahasa Indonesia menggunakan metode Yake. *Dataset* yang digunakan untuk melakukan penelitian ini adalah 100 publikasi ilmiah dari website *jtiik*, *jatisi*, dan *jepin* yang mana topiknya terbatas pada Ilmu Komputer. Berdasarkan penelitian yang dilakukan, konfigurasi parameter *Levenshtein Distance* memiliki pengaruh terhadap kata kunci yang dihasilkan. Hasil evaluasi dari penelitian ini mendapat nilai akurasi sebesar 97,05% dan nilai *f-score* sebesar 54,1% dengan parameter jarak *Levensthein* < 2.

Kata Kunci : Ekstraksi Kata Kunci, Yake, *Levensthein Distance*

KATA PENGANTAR

Puji syukur kepada Allah SWT atas rahmat dan nikmat Nya yang lebih diberikan kepada penulis sehingga dapat menyelesaikan skripsi ini dengan baik. Skripsi ini disusun sebagai salah satu syarat menyelesaikan Pendidikan program Strata-1 di Fakultas Ilmu Komputer Universitas Sriwijaya.

Dalam menyelesaikan skripsi ini, penulis menerima bantuan, bimbingan dan dukungan dari banyak pihak, baik secara langsung maupun tidak langsung. Oleh karena itu, penulis ingin menyampaikan terima kasih kepada:

1. Allah SWT atas rahmat dan nikmat-Nya penulis dapat menyelesaikan skripsi ini dengan baik.
2. Kedua orang tua, serta keluarga besar yang telah mendoakan, memberi semangat, memotivasi, dan nasihat untuk menyelesaikan skripsi ini.
3. Bapak Jaidan Jauhari, M.T. selaku Dosen Fakultas Ilmu Komputer Universitas Sriwijaya
4. Ibu Alvi Syahrini Utami, M.Kom. selaku Ketua Jurusan Teknik Informatika Universitas Sriwijaya
5. Bapak Dr. Abdiansah, S.Kom, M.Cs. selaku Dosen Pembimbing I dan Ibu Novi Yusliani, M.T. selaku dosen Pembimbing II yang telah membimbing, memberikan motivasi serta arahan kepada penulis dalam proses penggerjaan skripsi.
6. Bapak Alm. Drs. Megah Mulya, M.T. selaku Pembimbing akademik dari Semester 1 sampai 3 dan Bapak Muhammad Qurhanul Rizqie, S.Kom., M.T.

selaku Pembimbing akademik dari semester 4 sampai menyelesaikan perkuliahan di Universitas Sriwijaya.

7. Seluruh dosen program studi serta admin Jurusan Teknik Informatika Fakultas Ilmu Komputer Universitas Sriwijaya.
8. Teman – Teman penulis yang telah memberikan saran, motivasi, dan semangat selama mengerjakan skripsi ini.
9. Pihak – pihak lain yang tidak dapat penulis sebutkan satu-persatu.

Penulis menyadari bahwa dalam penyusunan skripsi ini masih banyak sekali kekurangan dikarenakan kurangnya pengalaman dan pengetahuan penulis. Oleh karena itu penulis mengharapkan saran dan kritik yang membangun guna kemajuan penelitian selanjutnya. Semoga tugas akhir ini dapat bermanfaat. Terima kasih.

Indralaya, 22 November 2022

Penulis

Gerald Plakasa

DAFTAR ISI

LEMBAR PENGESAHAN	ii
TANDA LULUS UJIAN KOMPREHENSIF.....	iii
HALAMAN PERNYATAAN.....	iv
MOTTO DAN PERSEMPAHAN	v
ABSTRACT.....	vi
ABSTRAK.....	vii
KATA PENGANTAR	viii
DAFTAR ISI.....	x
DAFTAR TABEL	xiv
DAFTAR GAMBAR.....	xvi
DAFTAR LAMPIRAN.....	xvii
BAB I PENDAHULUAN.....	I-1
1.1 Pendahuluan	I-1
1.2 Latar Belakang Masalah	I-1
1.3 Rumusan Masalah.....	I-2
1.4 Tujuan Penelitian	I-3
1.5 Manfaat Penelitian	I-3
1.6 Batasan Masalah	I-3
1.7 Sistematika Penulisan	I-4
1.8 Kesimpulan	I-5
BAB II KAJIAN LITERATUR.....	II-1
2.1 Pendahuluan	II-1
2.2 Landasan Teori	II-1
2.2.1 <i>Keyphrase Extraction</i>	II-1
2.2.2 <i>Yake</i>	II-2
2.2.2.1 <i>Text Pre-processing</i>	II-3
2.2.2.2 <i>Feature Extraction</i>	II-3

2.2.2.3 <i>Individual Term Weighting</i>	II-6
2.2.2.4 <i>Candidate Keyword List Generation</i>	II-7
2.2.3 <i>Levenshtein Distance</i>	II-7
2.2.4 <i>Confusion Matrix</i>	II-9
2.2.5 <i>Rational Unified Process</i>	II-10
2.3 Penelitian Lain yang Relevan	II-11
2.4 Kesimpulan	II-12
 BAB III METODOLOGI PENELITIAN	III-1
3.1 Pendahuluan	III-1
3.2 Pengumpulan Data	III-1
3.2.1 Jenis dan Sumber Data.....	III-1
3.2.2 Metode Pengumpulan Data	III-1
3.3 Tahapan Penelitian.....	III-2
3.3.1 Mengumpulkan Data	III-2
3.3.2 Menentukan Kerangka Kerja Penelitian.....	III-3
3.3.3 Menentukan Kriteria Pengujian.....	III-4
3.3.4 Menentukan Format Data Pengujian.....	III-4
3.3.5 Menentukan Alat Bantu Penelitian	III-6
3.3.6 Melakukan Pengujian Penelitian	III-6
3.3.7 Melakukan Analisis dan Menarik Kesimpulan Penelitian	III-7
3.4 Kesimpulan	III-7
 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-3
4.2.3.1 Analisis Kebutuhan Perangkat Lunak	IV-3
4.2.3.2 Analisis Data	IV-3

4.2.3.3 Analisis <i>Text Pre-processing</i>	IV-5
4.2.3.4 Analisis <i>Term Frequency</i>	IV-8
4.2.3.5 Analisis <i>Feature Extraction</i>	IV-9
4.2.3.6 Analisis <i>Individual Term Weighting</i>	IV-15
4.2.3.7 Analisis <i>Candidate Keyword List Generation</i>	IV-16
4.2.3.8 Analisis <i>Levensthein Distance</i>	IV-17
4.2.3.9 Analisis <i>Confusion Matrix</i>	IV-19
4.2.3.10 Desain Perangkat Lunak.....	IV-21
4.3 Fase Elaborasi	IV-24
4.3.1 Pemodelan Bisnis	IV-24
4.3.1.1 Perancangan Data	IV-25
4.3.1.2 Perancangan Antarmuka	IV-25
4.3.2 Kebutuhan Sistem	IV-28
4.3.3 Analisis dan Perancangan	IV-28
4.3.3.1 Diagram Aktivitas.....	IV-29
4.3.3.2 Diagram <i>Sequence</i>	IV-30
4.4 Fase Konstruksi	IV-32
4.4.1 Kebutuhan Sistem	IV-33
4.4.2 Implementasi.....	IV-34
4.4.2.1 Implementasi Kelas	IV-34
4.4.2.2 Implementasi <i>Interface</i>	IV-34
4.5 Fase Transisi.....	IV-37
4.5.1Pemodelan Bisnis	IV-37
4.5.2 Rencana Pengujian	IV-38
4.5.3 Implementasi.....	IV-38
4.6 Kesimpulan	IV-39
 BAB V HASIL DAN ANALISIS	V-1
5.1 Pendahuluan	V-1
5.2 Hasil Penelitian.....	V-1
5.2.1 Konfigurasi Percobaan Parameter	V-1

5.2.2 Hasil Pengujian	V-6
5.3 Analisis Hasil Penelitian	V-9
5.4 Kesimpulan	V-19
BAB VI KESIMPULAN DAN SARAN	VI-1
6.1 Pendahuluan	VI-1
6.2 Kesimpulan	VI-1
6.3 Saran	VI-1
DAFTAR PUSTAKA.....	xviii
DAFTAR LAMPIRAN.....	xx

DAFTAR TABEL

Tabel II-1. Perbedaan antara Keyword dan Keyphrase Extraction	II-2
Tabel III-1. Hasil kinerja seluruh <i>dataset</i>	III-5
Tabel III-2. Hasil Analisis Ekstraksi Kata Kunci.....	III-7
Tabel IV-1. Kebutuhan Fungsional	IV-2
Tabel IV-2. Kebutuhan Non-Fungsional	IV-2
Tabel IV-3. Contoh Data Judul	IV-5
Tabel IV-4. Contoh Data Abstrak	IV-5
Tabel IV-5. Hasil Penggabungan Teks Judul dan Teks Abstrak	IV-6
Tabel IV-6. Hasil Tokenisasi	IV-7
Tabel IV-7. Hasil Penghapusan Karakter Spesial	IV-8
Tabel IV-8. Hasil <i>Term Frequency</i>	IV-9
Tabel IV-9. Hasil <i>Word Casing</i>	IV-10
Tabel IV-10. Hasil <i>Word Position</i>	IV-11
Tabel IV-11. Hasil <i>Word Frequency</i>	IV-12
Tabel IV-12. Hasil <i>Word Relatedness to Context</i>	IV-13
Tabel IV-13. Hasil <i>Word DifSentence</i>	IV-14
Tabel IV-14. Hasil <i>Individual Term Weighting</i>	IV-15
Tabel IV-15. Hasil <i>Candidate Keyword List Generation</i>	IV-17
Tabel IV-16. Hasil <i>Levenshtein Distance</i>	IV-18
Tabel IV-17. Hasil <i>Cosine Similarity</i> sampel jurnal	IV-19
Tabel IV-18. Hasil <i>Confusion Matrix</i> sampel jurnal	IV-21
Tabel IV-19. Definisi <i>Actor</i>	IV-22
Tabel IV-20. Definisi <i>Use Case</i>	IV-22
Tabel IV-21. Skenario <i>Use Case</i> Evaluasi.....	IV-22
Tabel IV-22. Skenario <i>Use Case</i> Demonstrasi.....	IV-23
Tabel IV-23. Perancangan Data	IV-25
Tabel IV-24. Keterangan Implementasi Kelas	IV-34
Tabel IV-25. Rencana Pengujian <i>Use Case</i> Evaluasi	IV-38
Tabel IV-26. Rencana Pengujian <i>Use Case</i> Demonstrasi	IV-38
Tabel IV-27. Pengujian <i>Use Case</i> Evaluasi	IV-39

Tabel IV-28. Pengujian <i>Use Case</i> Demonstrasi	IV-39
Tabel V-1. Hasil Pengujian <i>Tuning</i> Top 3 Kata Kunci	V-2
Tabel V-2. Hasil Pengujian <i>Tuning</i> Top 5 Kata Kunci	V-2
Tabel V-3. Hasil Pengujian <i>Tuning</i> Top 10 Kata Kunci	V-3
Tabel V-4. Hasil Pengujian <i>Tuning</i> Top 50 Kata Kunci	V-4
Tabel V-5. Hasil Pengujian <i>Tuning</i> Top 100 Kata Kunci	V-4
Tabel V-6. Hasil Pengujian <i>Confusion Matrix</i> seluruh Kata Kunci (<i>all</i>) dengan jarak <i>Levensthein</i> < 2.....	V-8
Tabel V-7. Hasil Pengujian <i>Confusion Matrix</i> seluruh Kata Kunci (<i>all</i>) dengan jarak <i>Levensthein</i> < 4.....	V-8
Tabel V-8. Hasil Evaluasi Kata kunci 3 Sampel Jurnal	V-10
Tabel V-9. <i>Confusion Matrix</i> hasil akhir	V-14
Tabel V-10. <i>Confusion Matrix</i> tanpa Evaluasi <i>Similarity</i>	V-17

DAFTAR GAMBAR

Gambar II-1. Arsitektur Sistem Yake.....	II-2
Gambar II-2. Perbandingan <i>String</i> menggunakan GLD	II-8
Gambar II-3. Tabel <i>Confusion Matrix</i>	II-10
Gambar II-4. Arsitektur RUP	II-10
Gambar III-1. Rincian Kegiatan Penelitian	III-2
Gambar III-2. Kerangka Kerja Penelitian.....	III-3
Gambar IV-1. Diagram <i>Use Case</i>	IV-21
Gambar IV-2. Rancangan Tampilan <i>Home</i>	IV-26
Gambar IV-3. Rancangan Tampilan Evaluasi.....	IV-26
Gambar IV-4. Rancangan Tampilan Detail Evaluasi.....	IV-27
Gambar IV-5. Rancangan Tampilan Demo	IV-27
Gambar IV-6. Rancangan Tampilan Hasil Demo.....	IV-28
Gambar IV-7. Diagram Aktivitas Evaluasi.....	IV-29
Gambar IV-8. Diagram Aktivitas Demonstrasi.....	IV-30
Gambar IV-9. Diagram <i>Sequence</i> Evaluasi	IV-31
Gambar IV-10. Diagram <i>Sequence</i> Demonstrasi.....	IV-32
Gambar IV-11. Diagram Kelas	IV-33
Gambar IV-12. Implementasi Tampilan <i>Home</i>	IV-35
Gambar IV-13. Implementasi Tampilan Evaluasi	IV-35
Gambar IV-14. Implementasi Tampilan Detail Evaluasi	IV-36
Gambar IV-15. Implementasi Tampilan Demo	IV-36
Gambar IV-16. Implementasi Tampilan Hasil Demo	IV-37
Gambar V-1. <i>Tuning</i> jarak <i>Levenshtein Distance</i>	V-5
Gambar V-2. Nilai <i>F-Score</i> jarak <i>Levenshtein</i> < 2 dan < 4	V-6
Gambar V-3. Nilai <i>Accuracy</i> jarak <i>Levenshtein</i> < 2 dan < 4	V-7
Gambar V-4. Perbandingan Jarak <i>Levenshtein</i>	V-9

DAFTAR LAMPIRAN

Lampiran 1. Kode Program xx

BAB I

PENDAHULUAN

1.1 Pendahuluan

Pada bab ini akan membahas latar belakang masalah, rumusan masalah, tujuan penelitian, manfaat penelitian dan batasan masalah. Bab ini akan berisi penjelasan umum mengenai keseluruhan penelitian.

1.2 Latar Belakang Masalah

Publikasi hasil penelitian biasanya dilakukan oleh para peneliti, dosen, dan juga mahasiswa yang sudah melakukan kegiatan penelitian. Publikasi hasil penelitian sedang ditingkatkan di Indonesia agar ter indeks secara nasional, tidak hanya internasional (Arini at el, 2019). Berdasarkan pernyataan tersebut terlihat bahwa publikasi di Indonesia sedang berkembang. Banyaknya publikasi berbahasa Indonesia akan memerlukan banyak *reviewer* dengan berbagai bidang keahlian yang dimiliki. Ketika publikasi diterima untuk di sampaikan ke *reviewer*, publikasi haruslah tepat dengan bidang ilmu yang dimiliki (Arini at el, 2019).

Kata Kunci atau *Keyword* merupakan kata yang mewakili dari sebuah dokumen teks, karena kata kunci dapat terdiri dari satu atau lebih susunan kata yang dipilih secara khusus dari sebuah teks, oleh karena itu kata kunci biasanya di ambil dari bagian penting dari teks tersebut (Perdana, 2018). Ekstraksi Kata Kunci atau *Keyword Extraction* biasa juga orang menyebutnya *Keyphrase Extraction* diartikan

proses otomatis untuk pengambilan *term* dari sebuah dokumen teks (Arini et al, 2019).

Terdapat beberapa algoritma untuk *Keyphrase Extraction*, algoritma tersebut berdasarkan pendekatan *Statistic*, *Graph*, *Machine Learning* dan sebagainya. Salah Satu algoritma yaitu TextRank yang melakukan ekstraksi kata kunci berdasarkan pengaplikasian pada graf (Kurniawan, 2021). Terdapat juga algoritma TF.IDF, SingleRank, RAKE, dan Yake. Berdasarkan hasil penelitian terbukti bahwa Yake Mencapai hasil yang lebih baik dari pada keempat algoritma lainnya (Campos, Ricardo, et al, 2018). Yake (*Yet Another Keyword Extraction*) adalah *Unsupervised Term Extraction Method* yang tidak bergantung pada bahasa (Yunmar, et al. 2020). Jadi, Peneliti akan melakukan penelitian Ekstraksi Kata Kunci pada bahasa Indonesia yang mana akan menggunakan metode Yake dengan data publikasi ilmiah.

1.3 Rumusan Masalah

Berdasarkan permasalahan yang telah dijelaskan pada latar belakang, rumusan masalah penelitian ini adalah

1. Bagaimana mengembangkan Ekstraksi Kata Kunci pada Bahasa Indonesia menggunakan metode Yake?
2. Bagaimana kinerja Ekstraksi Kata Kunci pada dokumen publikasi Bahasa Indonesia menggunakan metode Yake dibandingkan dengan kata kunci dari penulis?

1.4 Tujuan Penelitian

Tujuan penelitian ini adalah

1. Menghasilkan sistem yang dapat melakukan Ekstraksi Kata Kunci pada Bahasa Indonesia menggunakan metode Yake.
2. Mengetahui kinerja Ekstraksi Kata Kunci pada dokumen publikasi Bahasa Indonesia menggunakan metode Yake dibandingkan dengan kata kunci dari penulis.

1.5 Manfaat Penelitian

Manfaat yang didapat dari penelitian ini adalah

1. Sistem dapat melakukan Ekstraksi Kata Kunci pada Bahasa Indonesia menggunakan metode Yake.
2. Hasil penelitian dapat dijadikan sebagai rujukan penelitian terkait

1.6 Batasan Masalah

1. Data yang digunakan adalah 100 data Publikasi Ilmiah dari *website jtiik*¹, *jatisi*², *jepin*³.
2. Data Publikasi Ilmiah terbatas pada topik Ilmu Komputer.

¹ <https://jtiik.ub.ac.id/index.php/jtiik/index>

² <https://jurnal.mdp.ac.id/index.php/jatisi/issue/archive>

³ <https://jurnal.untan.ac.id/index.php/jepin/index>

1.7 Sistematika Penulisan

Sistematika penulisan tugas akhir mengikuti standar penulisan tugas akhir Fakultas Ilmu Komputer Universitas Sriwijaya yaitu sebagai berikut:

BAB I. PENDAHULUAN

Pada bab ini membahas latar belakang, rumusan masalah, tujuan dan manfaat penelitian, batasan masalah serta sistematika penelitian yang akan dijadikan sebagai pokok pikiran penelitian ini.

BAB II. KAJIAN LITERATUR

Pada bab ini membahas landasan teori yang digunakan dalam penelitian, seperti definisi *Keyphrase Extraction* dan metode Yake, serta beberapa literatur yang relevan dengan penelitian ini.

BAB III. METODOLOGI PENELITIAN

Pada bab ini membahas proses yang akan dilaksanakan selama penelitian, Seperti pengumpulan data, analisis data dan perancangan perangkat lunak. Setiap tahap akan dijelaskan berdasarkan kerangka kerja yang dibuat.

BAB IV. PENGEMBANGAN PERANGKAT LUNAK

Pada bab ini membahas analisis dan rancangan perangkat lunak yang akan dikembangkan. Diawali dari analisis kebutuhan, perancangan dan konstruksi perangkat lunak, dan diakhiri dengan pengujian untuk memastikan sistem yang dikembangkan sudah sesuai dengan rancangan dan kebutuhan penelitian.

BAB V. HASIL DAN ANALISIS PENELITIAN

Pada bab ini menyajikan hasil pengujian berdasarkan langkah – langkah yang telah direncanakan. Analisis diberikan sebagai dasar kesimpulan yang akan diambil dari penelitian ini.

BAB VI. KESIMPULAN DAN SARAN

Pada bab ini membahas kesimpulan yang diambil berdasarkan uraian dalam bab sebelumnya serta saran yang diberikan berdasarkan penelitian yang telah dilakukan.

1.8 Kesimpulan

Pada bab ini telah dijelaskan latar belakang, rumusan masalah, tujuan dan manfaat penelitian, batasan masalah serta sistematika penelitian yang akan dijadikan sebagai pokok pikiran peneliti.

DAFTAR PUSTAKA

- Arini, Ni Wayan Sri, Ida Bagus Putu Widja, and I. Komang Rinartha Yasa Negara. "Analisis Frekuensi Kata untuk Mengekstrak Kata Kunci dari Artikel Ilmiah Berbahasa Indonesia." *Jurnal Eksplora Informatika* 8.2 (2019): 80-84.
- Behara, Krishna NS, Ashish Bhaskar, and Edward Chung. "A novel approach for the structural comparison of origin-destination matrices: Levenshtein distance." *Transportation Research Part C: Emerging Technologies* 111 (2020): 513-530.
- Campos, Ricardo, et al. "Yake! collection-independent automatic keyword extractor." European Conference on Information Retrieval. Springer, Cham, 2018.
- Campos, Ricardo, et al. "YAKE! Keyword extraction from single documents using multiple local features." *Information Sciences* 509 (2020): 257-289.
- Denny, Matthew J., and Arthur Spirling. "Text preprocessing for unsupervised learning: Why it matters, when it misleads, and what to do about it." *Political Analysis* 26.2 (2018): 168-189.
- Firdausillah, Fahri, and Erika Devi Udayanti. "Keyphrase Extraction on Covid-19 Tweets Based on Doc2Vec and YAKE." *Journal of Applied Intelligent System* 6.1 (2021): 23-31.
- Hakim, Zaenal, and Robby Rizky. "Analisis Perancangan Sistem Informasi Pembuatan Paspor Di Kantor Imigrasi Bumi Serpong Damai Tangerang Banten Menggunakan Metode Rational Unified Process." vol 6 (2020): 103-112.
- IBM, 1998. Rational Unified Process Best Practices for Software Development Teams. United States: IBM Rational Software.

- Kurniawan, Ahmad. Aplikasi sistem ekstraksi kata kunci berbahasa indonesia menggunakan algoritma textrank studi kasus data wikipedia Indonesia. BS thesis. Fakultas Sains dan Teknologi UIN Syarif Hidayatullah Jakarta.
- Lamasigi, Zulfrianto Yusrin. "DCT Untuk Ekstraksi Fitur Berbasis GLCM Pada Identifikasi Batik Menggunakan K-NN." *Jambura Journal of Electrical and Electronics Engineering* 3.1 (2021): 1-6.
- Lubis, Andre Hasudungan, Ali Ikhwan, and Phak Len Eh Kan. "Combination of levenshtein distance and rabin-karp to improve the accuracy of document equivalence level." *International Journal of Engineering & Technology* 7.2.27 (2018): 17-21.
- Nasar, Zara, Syed Waqar Jaffry, and Muhammad Kamran Malik. "Textual keyword extraction and summarization: State-of-the-art." *Information Processing & Management* 56.6 (2019): 102088.
- Perdana, Novario Jaya. "IMPLEMENTASI ALGORITMA GOOGLE LATENT SEMANTIC DISTANCE UNTUK EKSTRAKSI RANGKAIAN KATA KUNCI ARTIKEL JURNAL ILMIAH." *Computatio: Journal of Computer Science and Information Systems* 2.2 (2018): 186-195.
- Yunmar, Rajif Agung, Andika Setiawan, and Hartanto Tantriawan. "The Combination of YAKE and Language Processing for Unsupervised Term Extraction Ontology Learning." *IOP Conference Series: Earth and Environmental Science*. Vol. 537. No. 1. IOP Publishing, 2020.