

KEYPHRASE EXTRACTION MENGGUNAKAN TEXTRANK
UNTUK TEXT BERBAHASA INDONESIA

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



Oleh :

Fadel Muhammad
NIM : 09021381924153

Jurusan Teknik Informatika
FAKULTAS ILMU KOMPUTER UNIVERSITAS SRIWIJAYA
2023

LEMBAR PENGESAHAN SKRIPSI

KEYPHRASE EXTRACTION MENGGUNAKAN *TEXTRANK* UNTUK *TEXT* BERBAHASA INDONESIA

Oleh :

Fadel Muhammad
NIM : 09021381924153

Palembang, 27 Desember 2022

Pembimbing I

Novi Yusliani, S.Kom., M.T.
NIP. 198211082012122001

Pembimbing II,

M Naufal Rachmatullah M.T.
NIP. 199212012022031008

Mengetahui,

Ketua Jurusan Teknik Informatika

Alvi Syahrini Utami, M.Kom.
NIP. 19781222200642003

HALAMAN PERNYATAAN

Yang bertanda tangan di bawah ini :

Nama : Fadel Muhammad
NIM : 09021381924153
Program Studi : Teknik Informatika
Judul Skripsi : *Keyphrase Extraction Menggunakan TextRank untuk
Text Berbahasa Indonesia*

Hasil Pengecekan Software *iThenticate/Turnitin* : 16%

Menyatakan bahwa laporan skripsi saya merupakan hasil karya sendiri dan bukan hasil penjiplakan/plagiat. Apabila ditemukan unsur penjiplakan/plagiat dalam laporan skripsi 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 siapapun.

Palembang, 27 Desember 2022



Fadel Muhammad
NIM. 09021381924153

TANDA LULUS UJIAN KOMPREHENSIF

Pada hari Selasa tanggal 03 Januari 2023 Telah dilaksanakan ujian komprehensif skripsi oleh Jurusan Teknik Informatika Fakultas Ilmu Komputer Universitas Sriwijaya.

Nama : Fadel Muhammad

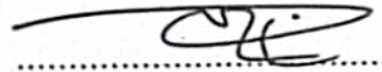
NIM : 09021381924153

Judul : *Keyphrase Extraction Menggunakan TextRank untuk Text Berbahasa Indonesia*

Dan dinyatakan LULUS.


1. Ketua Penguji

Osvari Arsalan, M.T.
NIP. 198603212018032001



2. Penguji 1

Dr. Abdiansah, S.Kom., M.CS.
NIP. 197812222006042003



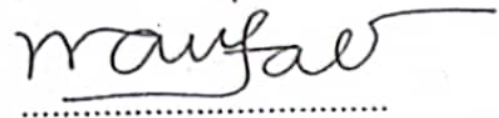
3. Pembimbing 1

Novi Yusliani, M.T.
NIP. 198211082012122001

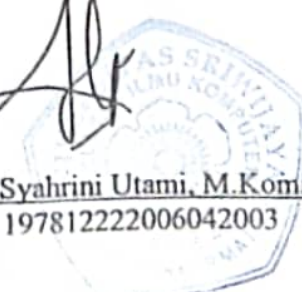



4. Pembimbing 2

Muhammad Naufal Rachmatullah, M.T.
NIP. 199212012022031008



Mengetahui,
Ketua Jurusan Teknik Informatika



Alvi Syahrini Utami, M.Kom.
NIP. 197812222006042003

MOTO DAN PERSEMBAHAN

”Sesulit apapun kehidupan ini, tetaplah hidup hari ini, bukan berhayal untuk hidup di masa depan” - Fadel muhammad

Skripsi ini kupersembahkan kepada:

- Allah SWT
- Orang Tua
- Kakak
- Dosen Pembimbing Akademik dan Skripsi
- Universitas Sriwijaya
- Teman – Teman

ABSTRACT

Keywords are commonly used as a form of summary from scientific publications. But in determining keywords, it requires expertise in the related field and a long amount of time because you have to read and understand the entire contents of scientific publications. Keyphrase Extraction can be a solution to get relevant keywords in a short time based on titles and abstracts from scientific publications. TextRank method is used to extract keywords. This research will perform keyword extraction using the TextRank method for Indonesian text. The evaluation results of this study showed an accuracy value of 95.53% and an f1-score of 59.32% with a threshold configuration of 80% and using all keyword candidates.

Keywords: Keyphrase Extraction, TextRank, Indonesian Text

ABSTRAK

Kata kunci biasa digunakan sebagai bentuk kesimpulan dari publikasi ilmiah. Namun dalam menentukan kata kunci, dibutuhkan keahlian terhadap bidang ilmu dan waktu yang lama karena harus membaca dan memahami keseluruhan isi publikasi ilmiah. *Keyphrase Extraction* dapat menjadi solusi untuk mendapatkan kata kunci yang relevan dalam waktu singkat berdasarkan judul dan abstrak dari publikasi ilmiah. metode *TextRank* digunakan untuk mendapatkan kata kunci. Penelitian ini akan melakukan ekstraksi kata kunci menggunakan metode *TextRank* untuk *text* berbahasa indonesia. Hasil evaluasi dari penelitian ini mendapat nilai akurasi sebesar 95.53% dan nilai *f1-score* sebesar 59.32% dengan konfigurasi *Threshold* sebesar 80% dan menggunakan seluruh kandidat kata kunci.

Kata Kunci: *Keyphrase Extraction*, *TextRank*, *Text* Berbahasa Indonesia

KATA PENGANTAR

Puji syukur penulis ucapkan kepada Allah SWT, karena berkat rahmat dan karunianya, penulis dapat menyelesaikan tugas akhir dengan judul “*Keyphrase Extraction Menggunakan TextRank untuk Text Berbahasa Indonesia*” dengan baik. Penelitian ini disusun sebagai syarat dalam menyelesaikan studi Strata-I program studi Teknik Informatika, Fakultas Ilmu Komputer Universitas Sriwijaya. Pada kesempatan ini penulis mengucapkan banyak terima kasih khususnya kepada:

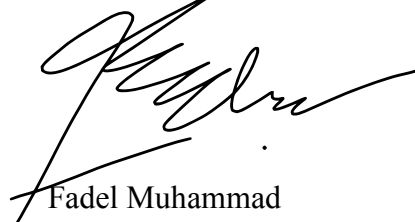
1. Allah SWT atas berkah, rahmat, dan karunianya.
2. Orang Tua saya, mama dan papa atas segala dukungannya.
3. Kakak” saya, cece, aak dan kak ade atas dukungannya.
4. Bapak Jaidan Jauhari, M.T. selaku Dekan Fakultas Ilmu Komputer Universitas Sriwijaya.
5. Ibu Alvi Syahrini Utami, M.Kom. selaku Ketua Jurusan Teknik Informatika Fakultas Ilmu Komputer Universitas Sriwijaya.
6. Ibu Novi Yusliani S.Kom., M.T. selaku Dosen Pembimbing I yang telah membantu memberikan arahan dan bimbingan.
7. Bapak Muhammad Naufal Rachmatullah S.Kom., M.T. selaku Dosen Pembimbing II yang telah memberikan arahan dan bimbingan.
8. Bapak Qurhanul Rizkie S.Kom., M.T. selaku Dosen Pembimbing Akademik atas segala dukungannya selama masa perkuliahan.
9. Seluruh Dosen Jurusan Teknik Informatika dan Dosen Fakultas ilmu Komputer yang telah membagikan ilmu dan pengetahuan selama masa perkuliahan.

10. Seluruh Staf Administrasi dan Pegawai yang telah membantu dalam urusan administrasi.

11. Seluruh teman-temanku di Universitas Sriwijaya.

Penulis menyadari bahwa masih terdapat kekurangan terhadap penyusunan tugas akhir ini. Oleh karena itu, penulis sangat berlapang dada apabila terdapat masukan atau saran yang membantu dalam menyempurnakan tugas akhir ini. Agar bisa bermanfaat bagi kita semua.

Palembang, 27 Desember 2022



Fadel Muhammad

DAFTAR ISI

	Halaman
LEMBAR PENGESAHAN SKRIPSI	ii
TANDA LULUS UJIAN KOMPREHENSIF	iii
HALAMAN PERNYATAAN	iv
MOTO DAN PERSEMBAHAN	v
ABSTRACT	vi
ABSTRAK	vii
KATA PENGANTAR	viii
DAFTAR ISI	x
DAFTAR TABEL	xiii
DAFTAR GAMBAR	xv
DAFTAR LAMPIRAN	xvi
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-3
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>POS Tag</i>	II-2
2.2.3 <i>TextRank</i>	II-3
2.2.4 <i>Confusion Matrix</i>	II-6

2.3	Penelitian Lain yang relevan	II-8
2.4	Kesimpulan.....	II-9
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-3
3.3.4	Menentukan Format Data Pengujian.....	III-4
3.3.5	Menentukan Alat Bantu Penelitian	III-5
3.3.6	Melakukan Pengujian Penelitian.....	III-5
3.3.7	Melakukan Analisis dan Menarik Kesimpulan	III-6
3.4	Kesimpulan.....	III-6
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-1
4.2.3	Analisis dan Desain.....	IV-2
4.2.3.1	Analisis Kebutuhan Perangkat Lunak.....	IV-2
4.2.3.2	Analisis Data.....	IV-2
4.2.3.3	Analisis <i>Text Pre-processing</i>	IV-2
4.2.3.4	Analisis <i>Word Weighting</i>	IV-7
4.2.3.5	Analisis <i>Rank Keyphrase</i>	IV-8
4.2.3.6	Analisis <i>Keyphrase List</i>	IV-9
4.2.3.7	Analisis <i>Confusion Matrix</i>	IV-9
4.2.3.8	Desain Perangkat Lunak	IV-10
4.3	Fase Elaborasi.....	IV-13

4.3.1	Pemodelan Bisnis	IV-13
4.3.1.1	Perancangan Data	IV-13
4.3.1.2	Perancangan <i>Interface</i>	IV-14
4.3.2	Kebutuhan Sistem	IV-14
4.3.3	Analisis dan Perancangan	IV-15
4.3.3.1	Diagram <i>Activity</i>	IV-15
4.3.3.2	Diagram <i>Sequence</i>	IV-15
4.4	Fase Konstruksi	IV-16
4.4.1	Kebutuhan Sistem	IV-16
4.4.2	Implementasi	IV-17
4.4.2.1	Implementasi Kelas	IV-17
4.4.2.2	Implementasi <i>Interface</i>	IV-18
4.5	Fase Transisi	IV-18
4.5.1	Pemodelan Bisnis	IV-18
4.5.2	Rencana Pengujian	IV-19
4.5.3	Implementasi	IV-19
4.6	Kesimpulan	IV-20
BAB V HASIL DAN ANALISIS		V-1
5.1	Pendahuluan	V-1
5.2	Hasil Penelitian	V-1
5.2.1	Konfigurasi Percobaan	V-1
5.2.2	Data Hasil Pengujian Menggunakan <i>Top Ranking</i>	V-4
5.2.3	Data Hasil Pengujian Menggunakan <i>Cosine Similarity</i>	V-9
5.3	Analisis Hasil Penelitian	V-13
BAB VI KESIMPULAN DAN SARAN		VI-1
6.1	Pendahuluan	VI-1
6.2	Kesimpulan	VI-1
6.3	Saran	VI-1
DAFTAR PUSTAKA		xiv
DAFTAR LAMPIRAN		xvi

DAFTAR TABEL

Tabel II-1. Tabel <i>Confusion Matrix</i>	II-6
Tabel III-1. Tabel Hasil Evaluasi.....	III-4
Tabel III-2. Tabel Hasil Analisis <i>Candidate Keyphrase</i>	III-6
Tabel III-3. Tabel Hasil Analisis dengan <i>Cosine Similarity</i>	III-6
Tabel IV-1. Kebutuhan Fungsional.....	IV-1
Tabel IV-2. Kebutuhan Non-Fungsional.....	IV-2
Tabel IV-3. Contoh Data Judul	IV-3
Tabel IV-4. Contoh Data Abstrak	IV-3
Tabel IV-5. Hasil Penggabungan Data Judul dan Abstrak	IV-3
Tabel IV-6. Hasil <i>Tokenizing</i>	IV-4
Tabel IV-7. Hasil <i>Stopwords Filtering</i>	IV-5
Tabel IV-8. Hasil <i>POS Tagging</i>	IV-5
Tabel IV-9. Hasil <i>Extract Phrase</i>	IV-6
Tabel IV-10. Hasil <i>Word Weighting</i>	IV-7
Tabel IV-11. Hasil Pemeringkatan Kandidat Kata Kunci.....	IV-8
Tabel IV-12. Hasil <i>Keyphrase</i>	IV-9
Tabel IV-13. Hasil Evaluasi	IV-9
Tabel IV-14. Hasil <i>Confusion Matrix</i> data sampel	IV-10
Tabel IV-15. Definisi Actor	IV-11
Tabel IV-16. Definisi <i>Use Case</i>	IV-11
Tabel IV-17. Skenario <i>Use Case Extract Keyphrase</i>	IV-12
Tabel IV-18. Implementasi Kelas	IV-17
Tabel IV-19. Rencana Pengujian <i>Use Case</i> Demonstrasi	IV-19
Tabel IV-20. Pengujian <i>Use Case</i> Demonstrasi.....	IV-19
Tabel V-1. Tabel Data Pengujian dan <i>Candidate Keyphrase</i>	V-2
Tabel V-2. Tabel <i>Selected Keyphrase</i> untuk Top 5.....	V-4
Tabel V-3. Tabel <i>Selected Keyphrase</i> untuk Top 10.....	V-4
Tabel V-4. Tabel <i>Selected Keyphrase</i> untuk Top 50.....	V-4
Tabel V-5. Tabel <i>Selected Keyphrase</i> untuk Top 100.....	V-5

Tabel V-6. Tabel Hasil Evaluasi untuk Top 5	V-6
Tabel V-7. Tabel Hasil Evaluasi untuk Top 10	V-7
Tabel V-8. Tabel Hasil Evaluasi untuk Top 50	V-7
Tabel V-9. Tabel Hasil Evaluasi untuk Top 100	V-8
Tabel V-10. Rata-rata hasil pengukuran tiap ketentuan pengambilan	V-8
Tabel V-11. Tabel <i>Selected Keyphrase</i> untuk <i>Threshold 40%</i>	V-9
Tabel V-12. Tabel <i>Selected Keyphrase</i> untuk <i>Threshold 60%</i>	V-10
Tabel V-13. Tabel <i>Selected Keyphrase</i> untuk <i>Threshold 80%</i>	V-10
Tabel V-14. Tabel Hasil Evaluasi untuk <i>Threshold 40%</i>	V-11
Tabel V-15. Tabel Hasil Evaluasi untuk <i>Threshold 60%</i>	V-11
Tabel V-16. Tabel Hasil Evaluasi untuk <i>Threshold 80%</i>	V-12
Tabel V-17. Rata-rata hasil pengukuran tiap konfigurasi <i>Threshold</i>	V-12

DAFTAR GAMBAR

Gambar II-1. Arsitektur Ekstraksi Kata Kunci (Pudota et al., 2010).....	II-2
Gambar II-2. Arsitektur Sistem <i>TextRank</i>	II-4
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-11
Gambar IV-2. Rancangan <i>Interface</i>	IV-14
Gambar IV-3. Diagram <i>Activity</i>	IV-15
Gambar IV-4. Diagram <i>Sequence</i>	IV-16
Gambar IV-5. <i>Class Diagram</i>	IV-17
Gambar IV-6. Implementasi Tampilan	IV-18
Gambar V-1. Diagram Evaluasi <i>Top Ranking</i>	V-9
Gambar V-2. Diagram Evaluasi <i>Cosine Similarity</i>	V-13
Gambar V-3. Perbandingan Kedua Konfigurasi Pengujian.....	V-14

DAFTAR LAMPIRAN

Lampiran 1. Kode Program.....	xvi
Lampiran 2. Hasil Evaluasi untuk Top 5.....	xxi
Lampiran 3. Tabel Hasil Evaluasi untuk Top 10.....	xxiv
Lampiran 4. Tabel Hasil Evaluasi untuk Top 50.....	xxvii
Lampiran 5. Tabel Hasil Evaluasi untuk Top 100.....	xxx
Lampiran 6. Tabel Hasil Evaluasi untuk <i>Threshold</i> 40%	xxxiii
Lampiran 7. Tabel Hasil Evaluasi untuk <i>Threshold</i> 60%	xxxvi
Lampiran 8. Tabel Hasil Evaluasi untuk <i>Threshold</i> 80%	xxxix

BAB I

PENDAHULUAN

1.1 Pendahuluan

Pada bab ini akan membahas tentang latar belakang masalah, rumusan masalah, tujuan penelitian, manfaat penelitian, dan batasan masalah. Pokok-pokok pikiran ini akan menjadi landasan pada bab selanjutnya.

1.2 Latar Belakang Masalah

Kata kunci biasanya digunakan sebagai bentuk kesimpulan dari publikasi ilmiah. Kata kunci dapat membantu dalam memahami isi dari publikasi ilmiah tanpa perlu membaca secara keseluruhan. Kata kunci dapat mewakili informasi yang terkandung dalam dokumen teks, seperti publikasi ilmiah. Penulis umumnya memberikan kata kunci yang dapat menggambarkan isi dari tulisannya (Asrori et al., 2020).

Kata kunci yang terdapat dalam publikasi ilmiah dibuat secara manual dengan membaca dan memahami keseluruhan isi untuk mendapatkan sebuah kesimpulan. Hal ini akan membutuhkan waktu apabila ingin membuat kata kunci pada publikasi ilmiah dalam jumlah banyak. Untuk alasan ini sistem ekstraksi kata kunci dibangun.

Dalam membangun sistem ekstraksi kata kunci, terdapat pendekatan yang bisa digunakan, antara lain pendekatan *Supervised* dan *Unsupervised*. Kelebihan pendekatan *Unsupervised* ialah metode tersebut tidak memerlukan pelatihan,

sehingga sistem ini dapat berjalan dengan baik di berbagai topik publikasi ilmiah (Mothe et al., 2020).

Dalam penelitian ini sistem ekstraksi kata kunci dikembangkan dengan algoritma *TextRank*. Kelebihan algoritma ini dibandingkan dengan algoritma lain ialah karena menggunakan pendekatan *Unsupervised* dan hanya mengandalkan kalimat yang terdapat dalam publikasi ilmiah, sehingga tidak diperlukannya pengetahuan mendalam akan sebuah bahasa (Abdurrohman, 2018).

TextRank menggunakan pendekatan *Unsupervised* sehingga tidak diperlukannya data pelatihan dan tidak ada masukan yang diperlukan untuk memproses data aktual. *TextRank* hanya didasarkan pada kata dan tidak memerlukan pengetahuan tata bahasa. Saat ini sudah banyak implementasi yang menggunakan *TextRank*, karena itu mudah bagi pengembang apabila ingin menggunakannya (Balcerzak et al., 2014).

Penelitian ini akan membangun sistem *Keyphrase Extraction* menggunakan metode *TextRank* untuk *text* berbahasa indonesia.

1.3 Rumusan Masalah

Rumusan masalah dalam penelitian ini adalah sebagai berikut:

1. Bagaimana mengekstrak kata kunci dari publikasi ilmiah menggunakan metode *TextRank*?
2. Bagaimana kinerja metode *TextRank* dalam mengekstrak kata kunci?

1.4 Tujuan Penelitian

Tujuan penelitian ini adalah sebagai berikut:

1. Menghasilkan aplikasi yang dapat mengekstrak kata kunci pada publikasi ilmiah menggunakan metode *TextRank*.
2. Mengetahui kinerja dari metode *TextRank* dalam mengekstrak kata kunci.

1.5 Manfaat Penelitian

Manfaat penelitian ini adalah sebagai berikut:

1. Mempermudah dalam mendapatkan kata kunci dari publikasi ilmiah tanpa harus membaca dan memahami isi secara keseluruhan.
2. Dapat digunakan sebagai rujukan dalam penelitian berikutnya.

1.6 Batasan Masalah

Batasan masalah dalam penelitian ini adalah sebagai berikut:

1. Publikasi ilmiah yang digunakan adalah publikasi berbahasa Indonesia.

1.7 Sistematika Penulisan

Sistematika penulisan dalam penelitian ini adalah sebagai berikut:

BAB I. PENDAHULUAN

Pada bab ini akan membahas tentang latar belakang masalah, rumusan masalah, tujuan penelitian, manfaat penelitian, dan batasan masalah. Pokok-pokok pikiran ini akan menjadi landasan pada bab selanjutnya.

BAB II. KAJIAN LITERATUR

Pada bab ini akan membahas tentang landasan teori yang digunakan dalam penelitian, yang diantaranya adalah *Keyphrase Extraction*, *POS Tag*, *TextRank*, *Confusion Matrix* dan penelitian yang relevan.

BAB III. METODOLOGI PENELITIAN

Pada 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

Pada bab ini akan menjelaskan tentang pengembangan perangkat lunak menggunakan metode *Rational Unified Process* (RUP). Dimulai dari fase Insepsi, Elaborasi, Konstruksi, dan Transisi.

BAB V. HASIL DAN ANALISIS PENELITIAN

Pada bab ini akan ditampilkan hasil dari proses *Keyphrase Extraction* menggunakan *TextRank* dalam text berbahasa indonesia. Hasil evaluasi akan diberikan dengan dua konfigurasi, mengambil peringkat teratas berdasarkan dengan ketentuan pengambilan dan menggunakan *Similarity*. Serta hasil analisis berdasarkan hasil tersebut.

BAB VI. KESIMPULAN DAN SARAN

Pada bab ini akan membahas tentang kesimpulan dan saran berdasarkan penelitian yang telah dilakukan. Hasil penelitian ini dapat digunakan sebagai referensi dan acuan untuk penelitian selanjutnya.

1.8 Kesimpulan

Pada bab ini telah dibahas mengenai latar belakang masalah, rumusan masalah, tujuan penelitian, manfaat penelitian, Batasan masalah, dan sistematika penulisan.

DAFTAR PUSTAKA

- Abdurrohman. (2018). *Evaluasi Algoritma Textrank pada peringkasan teks berbahasa indonesia*. 4–16.
- Asrori, R. B., Setyawan, R., & Muljono, M. (2020). Performance analysis graph-based keyphrase extraction in Indonesia scientific paper. *Proceedings - 2020 International Seminar on Application for Technology of Information and Communication: IT Challenges for Sustainability, Scalability, and Security in the Age of Digital Disruption, ISemantic 2020*, 185–190. <https://doi.org/10.1109/iSemantic50169.2020.9234231>
- Bałcerzak, B., Jaworski, W., & Wierzbicki, A. (2014). Application of textrank algorithm for credibility assessment. *Proceedings - 2014 IEEE/WIC/ACM International Joint Conference on Web Intelligence and Intelligent Agent Technology - Workshops, WI-IAT 2014, 1*, 451–454. <https://doi.org/10.1109/WI-IAT.2014.70>
- Diana Permata Sari; Ayu Purwarianti. (2014). Ekstraksi Kata Kunci Otomatis Untuk Dokumen Bahasa Indonesia Studi Kasus: Artikel Jurnal Ilmiah Koleksi Pdi Lipi. *Baca: Jurnal Dokumentasi Dan Informasi*, 35(2), 139–147. <https://doi.org/http://dx.doi.org/10.14203/j.baca.v35i2.192>
- Dostal, M., & Ježek, K. (2011). Automatic keyphrase extraction based on NLP and statistical methods. *CEUR Workshop Proceedings*, 706, 140–145.
- Gunawan, B., Pratiwi, H. S., & Pratama, E. E. (2018). Sistem Analisis Sentimen pada Ulasan Produk Menggunakan Metode Naive Bayes. *Jurnal Edukasi Dan Penelitian Informatika (JEPIN)*, 4(2), 113. <https://doi.org/10.26418/jp.v4i2.27526>
- Informatika, J. T. (2019). *SKRIPSI Oleh : MUHAMMAD AUFA SHIDDIQ*.
- Kurniawan, A. (2021). Aplikasi sistem ekstraksi kata kunci berbahasa indonesia menggunakan algoritma textrank studi kasus data wikipedia Indonesia. *Repository.Uinjkt.Ac.Id*. <https://repository.uinjkt.ac.id/dspace/handle/123456789/56134%0Ahttps://repository.uinjkt.ac.id/dspace/bitstream/123456789/56134/1/AHMAD>

KURNIAWAN-FST.pdf

- Mothe, J., Ramiandrisoa, F., Rasolomanana, M., Mothe, J., Ramiandrisoa, F., & Rasolomanana, M. (2020). *Automatic keyphrase extraction using graph-based methods* To cite this version : HAL Id : hal-02640988.
- Pramudita, H. R., Utami, E., & Amborowati, A. (2016). Pengaruh Part of Speech Tagging Berbasis Aturan dan Distribusi Probabilitas Maximum Entropy untuk Bahasa Jawa Krama. *Jurnal Buana Informatika*, 7(4), 235–244. <https://doi.org/10.24002/jbi.v7i4.764>
- Pudota, N., Dattolo, A., Baruzzo, A., & Tasso, C. (2010). A new domain independent keyphrase extraction system. *Communications in Computer and Information Science*, 91 CCIS(September), 67–78. https://doi.org/10.1007/978-3-642-15850-6_8
- Rina, S. (2018). UNIVERSITAS SUMATERA UTARA Poliklinik UNIVERSITAS SUMATERA UTARA. *Jurnal Pembangunan Wilayah & Kota*, 1(3), 82–91.
- Santra, A. K., & Christy, C. J. (2012). Genetic Algorithm and Confusion Matrix for Document Clustering. *International Journal of Computer Science*, 9(1), 322–328. <http://ijcsi.org/papers/IJCSI-9-1-2-322-328.pdf>
- Wicaksono, A. F., & Purwarianti, A. (2010). HMM based part-of-speech tagger for Bahasa Indonesia. In Fourth International MALINDO Workshop, Jakarta. *Fourth International MALINDO ...*, June. https://www.researchgate.net/profile/Alfan-Farizki-Wicaksono/publication/209387036_HMM_Based_Part-of-Speech_Tagger_for_Bahasa_Indonesia/links/04d39ed44efdacf36e3af1a2/HMM-Based-Part-of-Speech-Tagger-for-Bahasa-Indonesia.pdf