

## SISTEM KELAS KATA BERIMBUHAN MENGGUNAKAN ALGORITMA PORTER STEMMER SEBAGAI PEMBELAJARAN BAHASA INDONESIA

Muhamad Rosidin<sup>(1)</sup>, Abdul Fadlil<sup>(2)</sup>, Anton Yudhana<sup>(3)</sup>

<sup>1</sup>Teknik Informatika, Universitas Ahmad Dahlan

<sup>2,3</sup>Teknik Elektro, Universitas Ahmad Dahlan

e-mail : idin.ilkom.unpak@gmail.com<sup>(1)</sup>, fadlil@mti.uad.ac.id<sup>(2)</sup>, eyudhana@ee.uad.ac.id<sup>(3)</sup>

### Abstract

*Class of word are categories of forms, meanings and functions in the Indonesian grammar system. Determining the basic words of the word affix is still difficult. The word affix class system is a solution to determine the basic words at same time with the class of words that are affixed so that they can be used as learning Indonesian language. The word affix class system was developed using the PHP programming language. Porter algorithms are algorithms that are used to change the word affix into a root word. In this study, the word affix detected by system and separated by infiltration section, then system would process types of affixes juxtaposed with the root words to form a class of words. This can facilitate Indonesian language, basic words and word classes learning. The output is in the form of a table of steps in stemming process and determining affix words class.*

**Keywords:** Word Class, Porter, Porter Stemmer, Imbunan, Stemming Algorithm

### Abstrak

Kelas kata adalah kategori bentuk, makna dan fungsi dalam sistem bahasa Indonesia. Penentuan kata dasar dari kata berimbuhan masih sulit dilakukan. Sistem kelas kata berimbuhan adalah solusi untuk menentukan kata dasar sekaligus dengan kelas kata berimbuhan sehingga dapat digunakan sebagai media pembelajaran bahasa Indonesia. Pengembangan sistem kelas kata berimbuhan berbasis *web* dan proses codingnya dengan bahasa pemrograman PHP. *Porter algorithm* merupakan algoritma *stemming* yang digunakan untuk mengubah kata berimbuhan menjadi kata dasar. Pada penelitian ini, kata berimbuhan dideteksi oleh sistem dan dipisahkan per bagian imbuhan, kemudian sistem akan memroses jenis imbuhan yang disandingkan dengan kata dasar akan membentuk sebuah kelas kata. Hal ini dapat memudahkan pembelajaran Bahasa Indonesia mengenai kata berimbuhan, kata dasar dan kelas kata. *Stemming* dan perumusan formula sangat mempengaruhi penentuan kelas kata, karena bahasa Indonesia memiliki imbuhan yang kompleks dan memerlukan algoritma *stemming* yang memiliki kemampuan tinggi dalam menentukan *root word*. Terdapat sejumlah *stemming algorithm* yang telah dikembangkan, pada penelitian ini *stemming porter algorithm* yang dipakai dalam penentuan kata dasar. Output yang dihasilkan berupa *table* langkah-langkah dalam proses *stemming* dan penentuan kelas kata berimbuhan.

**Kata Kunci :** Kelas Kata, Algoritma Porter, Porter Stemmer, Imbunan, Stemming

## 1. PENDAHULUAN

Penggunaan Bahasa Indonesia dalam kehidupan sehari-hari sangatlah kompleks. Setiap orang memiliki kebebasan dalam penggunaannya. Selain penggunaan kata yang tepat, ada bagian kata yang harus disesuaikan yaitu imbuhan. Imbuhan dapat dipergunakan pada setiap kata dan mempunyai teknik penggabungan masing-masing. (Rosidin, Fadlil, & Yudhana, 2018). Terdapat 35 jenis imbuhan resmi bahasa Indonesia yang tercatat dalam KBBI (Kamus Besar Bahasa Indonesia). Imbuhan tersebut dapat digolongkan kedalam beberapa bagian kata antara lain : awalan, sisipan dan akhiran. Dalam pemecahan imbuhan menjadi *root word* tidaklah mudah. Terdapat kata dasar yang huruf awalnya mengalami perubahan ketika diberi suatu imbuhan. Aturan ini yang dapat menyulitkan dalam penggunaan kata yang tepat. Selain imbuhan, adapula yang dikenal dengan kelas kata. Dalam penentuan kelas kata dapat dilakukan dengan melihat imbuhan yang disisipkan pada kata dasar. Kelas kata yang terdapat dalam bahasa Indonesia berjumlah delapan macam dan mempunyai arti di masing-masing kelasnya. Dalam penggunaan imbuhan terdapat hubungan antara *root word* dengan penentuan kelas kata

berimbuhan. Dalam penentuan kelas kata pemilihan imbuhan juga memiliki pengaruh penting. Imbuhan yang digunakan dalam sebuah *root word* akan dapat dikelompokkan dalam salah satu kelas kata. Seperti diketahui bahwa tahapan yang paling penting dalam temu kembali informasi adalah *stemming process*. (Afuan, 2013). Implementasi algoritma *stemming* sangat luas contohnya dalam mendeteksi plagiarisme. Sebagaimana diketahui fenomena plagiarisme dalam dunia akademik Indonesia sudah lama terjadi, sehingga lembaga tinggi dicemari oleh tindak plagiarisme. (Sunardi, Yudhana, & Mukaromah, 2019). Berdasarkan uraian diatas akan dibuat suatu sistem kelas kata berimbuhan dengan implementasi *porter stemmer algorithm*.

Penelitian ini mempermudah dalam proses *stemming* kata berimbuhan sehingga dapat membantu dalam ketepatan penggunaan kata yang sesuai dengan format EYD, membantu dalam penentuan kelas kata berimbuhan dari tahap-tahap proses *porter stemmer*. Kemudian dapat memberi pembelajaran tentang keterkaitan antara kata berimbuhan, kelas kata, dan kata dasar. Harapannya dapat digunakan sebagai pembelajaran bahasa Indonesia.

## 2. TINJAUAN PUSTAKA

Temu kembali informasi adalah tahapan dalam mengidentifikasi, memanggil (*retrieve*) suatu dokumen dari suatu *directory (file)* sebagai timbal balik atas permintaan informasi. (Agnew, 2017). Bagian terpenting dalam tahapan temu kembali informasi adalah proses *stemming*. Salah satu penelitian yang menerapkan metode *stemming* adalah evaluasi efektivitas dan efisiensi dari algoritma dilakukan dengan menghitung yang berasal waktu proses dan presisi. Kesimpulannya adalah Nazief & Adriani memiliki ketepatan yang baik untuk tujuan linguistik dari *porter*. (Agusta, 2009).

*Stemming* adalah proses penghapusan imbuhan dari suatu kata, kalimat, dokumen atau query menjadi *root word*. (Novitasari, 2016). Terdapat perbedaan antara proses algoritma *porter stemmer* pada teks berbahasa Indonesia dengan teks yang berbahasa Inggris. (Pramudita, 2014). *Stemming* adalah proses penentuan *root word* dari kata berimbuhan. Nazief-Adriani merupakan salah satu algoritma *stemming* yang sering digunakan. (Yudhana, Djayali, & Sunardi, 2017). Penggunaan *stemming* dalam pengolahan dokumen elektronik telah digunakan secara luas. *Stemming* diimplementasikan dalam sejumlah fokus keilmuan antara lain : cek *similarity*, pendeteksi ejaan, sistem penerjemah, *document clustering*, klasifikasi dokumen dan lain-lain. (Bagus, Wiguna, & Hantono, 2013).

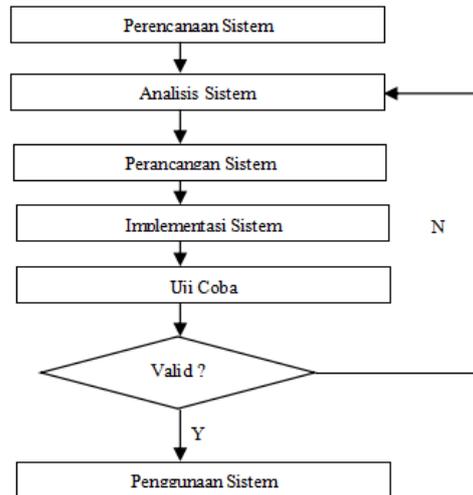
Algoritma *porter stemmer* merupakan proses penentuan kata dasar melalui penghapusan afiks atau imbuhan, kata yang ambigu dan membingungkan dapat terjadi karena aturan morfologi yang tidak konsisten dalam tata bahasa Indonesia. (Nurul Justina Mahardianingroem, 2018). Algoritma atau tahapan dalam *Porter Stemmer* berbahasa Indonesia adalah penghilangan partikel dalam sebuah kata contohnya: -lah pada kata adalah, -kah pada kata kamukah, menghilangkan possessive pronoun atau kata ganti yang terkandung pada sebuah kata, contohnya -mu pada kata bukumu, -ku, -nya, menghilangkan awalan satu jika awalan satu tidak terdeteksi, maka lanjut ke tahap menghapus awalan kedua, dan jika ada maka lanjut ke tahap menghapus akhiran dan tahap akhir menghilangkan awalan ke-2 lalu dilanjutkan pada langkah ke penghapusan suffix dan menghapus akhiran, jika tidak ditemukan maka kata tersebut dianggap sebagai kata dasar (*root word*). (Indriyono, Utami, & Sunyoto, 2017).

Kelas kata merupakan perangkat kata yang memiliki sintaksis sama. (Yesi Nur Cahyati, 2015). Kelas kata merupakan golongan kata yang memiliki persamaan dalam perilaku formalnya; klasifikasi atas kata benda (nomina), kata kerja (verba), dan kata sifat (adjektiva). (Ismayasari, Pastika, & Putra, 2018). Kata turunan atau berimbuhan merupakan kata disisipkan atau disandingkan sebuah imbuhan. (Yosephine & Prabowo, 2017). Kata berimbuhan dapat didefinisikan sebagai kata dasar yang mendapat penambahan berupa imbuhan di bagian awal, akhir, sisipan atau gabungan yang sering disebut morfem terikat. (Wibowo, 2016)

---

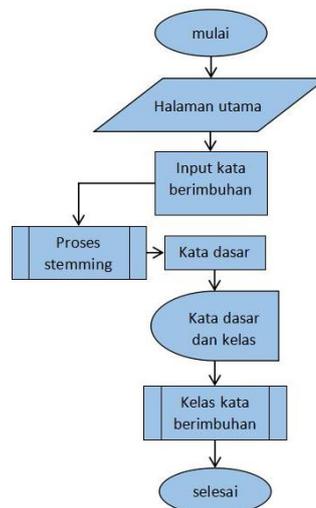
### 3. METODE PENELITIAN

Metode penelitian yang diterapkan pada Sistem Kelas Kata Berimbuhan Menggunakan Algoritma *Porter Stemmer* sebagai pembelajaran Bahasa Indonesia yaitu *System Development Life Cycle* (SDLC) atau metode Siklus Hidup Pengembangan Sistem. Tahap yang dilakukan secara umum dalam pengembangan sistem berdasarkan SDLC terdiri dari 5 tahap: Tahap Perencanaan, Tahap Analisis Sistem, Tahap Perancangan, Tahap Implementasi dan Tahap Uji Coba, pengembangan sistem SDLC dapat diamati pada gambar 1 sebagai berikut:



**Gambar 1. Siklus Hidup Pengembangan Sistem (System Development Life Cycle )**

Tahap Analisis Sistem dilakukan dalam penelitian ini sebagai acuan dalam penyusunan sistem yang akan dibangun. Dalam pengembangan sistem, aliran data pada Sistem penentuan kelas kata berimbuhan pada gambar 2 sebagai berikut:



**Gambar 2. Proses Alur Sistem**

Tahap implementasi terdapat beberapa hal yang harus dilakukan antara lain, pembangunan *database* menggunakan *MySQL*, desain *form-form* dan *source code* dilakukan dengan memakai aplikasi *sublime text 3* sebagai halaman editor dengan bahasa pemrograman *PHP* dan pembuatan *database* dibangun menggunakan *MySQL*. Tahap uji coba menggunakan *browser Mozilla Firefox*, dan *local server* yaitu *XAMPP*.

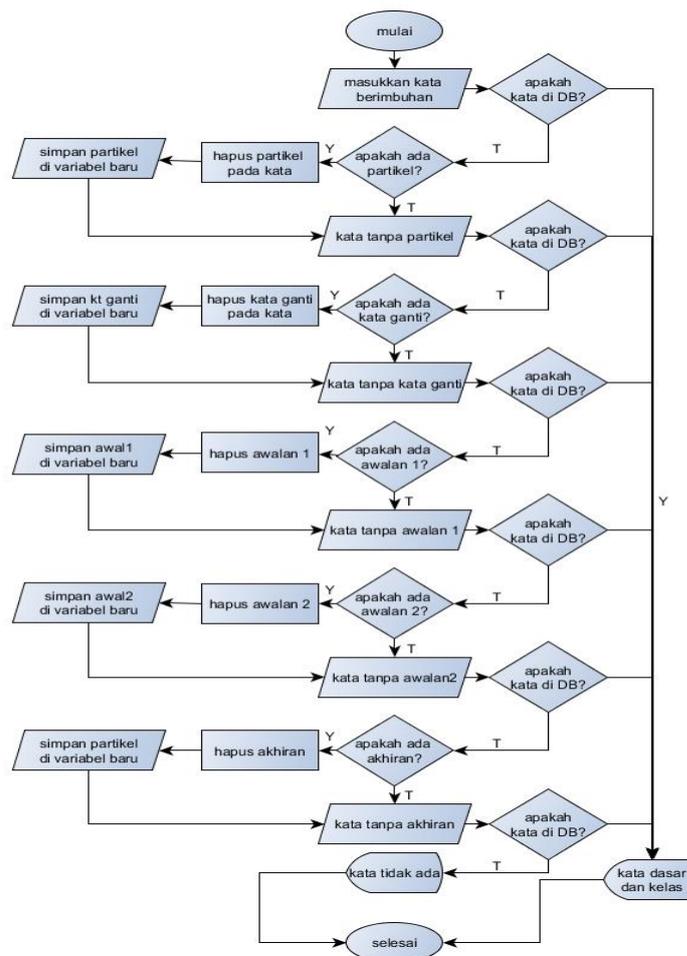
Tahapan pengujian sistem adalah tahap untuk dapat menyimpulkan sebuah sistem mampu menjalankan fungsinya. Tahap ini terdiri dari tiga uji coba yakni, pengujian secara struktur, pengujian secara fungsi dan pengujian validasi.

Pada pengujian struktural dilakukan dengan mengamati *form* pada sistem harus sesuai dengan rancangan yang ada, dengan membandingkan rancangan awal dan hasil akhirnya. Dalam mengetahui ketepatan fungsi dari sistem kelas kata harus dilakukan pengujian fungsional sistem, pengujian dilakukan dengan memeriksa fungsi setiap *button-button* yang ada pada *form* dan hasil penyimpanan pada database. Hasil akurasi sistem adalah hal yang harus diketahui pada uji coba validasi. Keberhasilan pada proses pengujian adalah dengan melihat keterkaitan sistem dalam mengeluarkan hasil dengan perancangan manual.

#### 4. HASIL DAN PEMBAHASAN

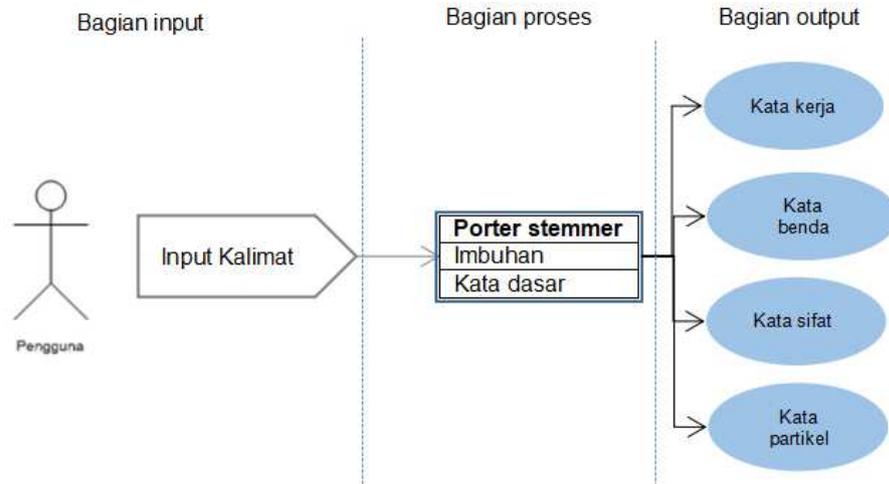
Secara garis besar ada dua tahapan *flowchart* sistem. Pertama, pemrosesan input kata berimbuhan atau sekumpulan kalimat menjadi sejumlah kata dasar dengan algoritma *porter stemmer*. Kedua, sistem akan memroses hasil pada tahap pertama menjadi kelas kata berimbuhan.

Kata dasar atau *root word* otomatis akan memberikan output berupa kelas kata. Aliran proses pada *stemming* kata dasar menggunakan *porter stemmer* pada gambar 3 sebagai berikut.



Gambar 3. Flowchart Porter Stemmer

Penentuan kelas dari kata berimbuhan bisa dilakukan dengan proses *stemming* terlebih dahulu. Berikut proses yang di usulkan pada gambar 4.



Gambar 4. Proses yang diusulkan

#### Praproses Data

Masukan yang berupa kalimat akan dipecah menjadi kata kemudian dilakukan proses *stemming*. Kata atau kalimat yang di *input* harus dilakukan pencocokan pada kamus data, hal ini dilakukan untuk mencegah kesalahan sistem dalam mendeteksi imbuhan contohnya kata “berat”, ketika diproses maka sistem akan membuang kata “ber” sebagai imbuhan dan mengeluarkan kata “at” sebagai kata dasar.

#### Hapus Partikel dan Imbuhan

Beberapa contoh kata partikel yang akan dideteksi sistem antara lain; apapun, perlembar adalah dan lain-lain. Setelah itu sistem akan mendeteksi kata ganti, contohnya; miliknya, bukumu, dan uangku. Selanjutnya sistem akan mendeteksi jumlah awalan pada kata, jika ada dua imbuhan, maka akan menghilangkan dua awalan tersebut, contohnya; memperberat, mempercantik, dan memperkaya. Kemudian sistem akan menghilangkan kata awalan contohnya; memakan, dimakan, menghapus, menjadi, permudah dan lain-lain. Pada tahap akhir sistem akan mendeteksi keberadaan akhiran dan menghilangkannya, contoh kata yang dilengkap akhiran antara lain; ajari, baikan, jalani, cintai dan lain-lain.

#### Penentuan Kelas Kata

Penentuan kelas kata akan dilakukan dengan *query* yang mengadopsi aturan dari penentuan kelas kata. Aturan-aturan penentuan kelas kata berimbuhan pada tabel 1 sebagai berikut.

Tabel 1. Aturan penentuan kelas kata berimbuhan

Imbuhan			Kelas Kata
Awalan	Akhiran	Gabungan	
pe-	-an, -nya	pe-an, ke-an, per-an	Benda
me-, ber-, di-, ter-, pe-	-kan, -i	-	Kerja
-	-	Se-nya	Sifat
-per	-kah, -lah, -pun	-	Partikel

Proses *stemming* menggunakan *porterstemmer* diharapkan kata yang diproses sesuai dan penentuan kelas kata berimbuhan mendapatkan hasil yang benar seperti pada tabel 2 sebagai berikut.

Tabel 2. Hasil proses *stemming* dan penentuan kelas kata yang diharapkan

Stemming			Kelas Kata
Kata Imbuhan	Imbuhan	Kata Dasar	
pelari	Pe-	Lari	Benda
melihat	Me-	Lihat	kerja
sebagusnya	Se-nya	bagus	Sifat
perlembar	Per-	lembar	Partikel

Pada tabel di atas kelas kata yang dihasilkan itu merujuk pada kata berimbuhan. Untuk kata dasar, memiliki kelas kata tersendiri. Pada penentuan kelas kata dasar dapat dilakukan *query* langsung ke kamus kata dasar yang dibuat kedalam sebuah database.

### Pembuatan Antarmuka Pengguna

Halaman antarmuka pada awal sistem ini untuk memudahkan pengguna dengan tampilan *simple*. Terdapat sebuah logo "Klask" yang mendeskripsikan sebuah sistem kelas kata seperti terdapat pada gambar 5. terdapat *form* masukkan yang dibuat untuk mengetik kata berimbuhan ataupun berupa kalimat oleh pengguna, dan tiga tombol yaitu tombol proses, tombol hapus dan tombol bantuan.

## Klask

Proses Kata
Hapus
Bantuan

Gambar 5. Antarmuka halaman utama

Pada halaman hasil rancangan tampilan antarmuka adalah sebuah *form input* berisi kalimat yang sama pada saat diinput di halaman utama, dan hasil dari proses *stemming* beserta urutannya seperti terdapat pada gambar 6.

Proses Kata

Tabel Hasil Stemming dengan Algoritma Porter Bahasa Indonesia

Inputan	Hapus Partikel	Hapus Possesive Pronoun	Hapus Awalan 1	Hapus Awalan 2	Hapus Akhiran	Hasil Stemming	Kata Dasar	Kelas Kata Berimbuhan	Kelas Kata Dasar
saya	saya	saya	saya	saya	saya	saya	saya	Belum Tersedia	Pronomina
perorang	perorang	perorang	perorang	orang	orang	orang	orang	Partikel	Nomina
perdalam	perdalam	perdalam	perdalam	dalam	dalam	dalam	dalam	Verba / Kerja	Adjektiva
perlembar	perlembar	perlembar	perlembar	lembar	lembar	lembar	lembar	Partikel	Nomina
tercantik	tercantik	tercantik	tercantik	cantik	cantik	cantik	cantik	Adjektiva / Sifat	Adjektiva
memakan	memakan	memakan	makan	makan	makan	makan	makan	Verba / Kerja	Verba
pemakan	pemakan	pemakan	makan	makan	makan	makan	makan	Nomina / Benda	Verba

Gambar 6. Antarmuka halaman hasil dengan input kalimat

## KESIMPULAN

Sistem Kelas Kata Berimbuhan Menggunakan Algoritma *Porter Stemmer* Sebagai Pembelajaran Bahasa Indonesia dikembangkan berbasis *web*. Bahasa pemrograman yang digunakan adalah *PHP*. Pada pembuatan *sourcecode* dibagi menjadi dua tahap, pertama adalah penentuan kelas dan kata dasar dengan metode *stemming* dan kedua adalah menentukan kelas kata berimbuhan. Sistem ini dibangun dengan menerapkan algoritma Porter

Stemmer dalam proses stemming. Penyelesaian sistem ini dapat diuji dengan menggunakan *web client* seperti *google chrome*, *firefox interface* atau antar mukanya terdiri dari dua bagian yaitu *form* utama atau *home* dan *form* hasil. Sistem kelas kata dapat memroses inputan berupa kata imbuhan tunggal ataupun masukan berupa kalimat. Maksimal kalimat yang dapat dimasukan tidak ditentukan. Kinerja sistem lebih baik dalam memroses satu kata imbuhan.

Harapannya aplikasi sistem kelas kata dapat menambah pengetahuan dalam pemilihan kata yang sesuai baik untuk percakapan sehari-hari, untuk sebuah penulisan ataupun sebagai media pembelajaran bahasa Indonesia.

## DAFTAR PUSTAKA

- Afuan, L. (2013). Stemming Dokumen Teks Bahasa Indonesia. *Jurnal Telematika*, 6(2), 34–40.
- Agnew, H. L. (2017). Our Natural Language. *Origins of the Czech National Renaissance*, (1982), 51–92. <https://doi.org/10.2307/j.ctt5hjshh.6>
- Agusta, L. (2009). Perbandingan Algoritma Stemming Porter dengan Algoritma Nazief & Andriani Untuk Stemming Dokumen Teks Bahasa Indonesia. *Konferensi Nasional Sistem Dan Informatika*, 196–201.
- Bagus, P., Wiguna, S., & Hantono, B. S. (2013). Peningkatan Algoritma Porter Stemmer Bahasa Indonesia berdasarkan Metode Morfologi dengan Mengaplikasikan 2 Tingkat Morfologi dan Aturan Kombinasi Awalan dan Akhiran, 2(2), 1–6. <https://doi.org/http://dx.doi.org/10.22146/jnteti.v2i2.49>
- Indriyono, B. V., Utami, E., & Sunyoto, A. (2017). Pemanfaatan Algoritma Porter Stemmer Untuk Bahasa Indonesia Dalam Proses Klasifikasi Jenis Buku. *World Journal of Surgical Oncology*, 15(1), 301–310. <https://doi.org/10.1186/s12957-017-1193-5>
- Ismayasari, R., Pastika, I. W., & Putra, A. P. (2018). Analisis Morfologi Kelas Kata Terbuka Pada Editorial Media Cetak, 17(2008), 379–384. <https://doi.org/10.1109/icuas.2018.8453477>
- Novitasari, D. (2016). Perbandingan Algoritma Stemming Porter Dengan Arifin, 1(2), 120–129.
- Nurul Justina Mahardianingroem, A. S. (2018). Penerapan Kamus Dasar Pada Algoritma Porter Untuk Mengurangi Kesalahan Stemming Bahasa Indonesia, 10(2), 103–112.
- Pramudita, H. R. (2014). Penerapan Algoritma Stemming Nazief & Andriani dan Similarity Pada Penerimaan Judul Thesis. *Jurnal Ilmiah DASI*, 15(04), 15–19.
- Rosidin, M., Fadlil, A., & Yudhana, A. (2018). Perancangan Sistem Penentuan Kelas Kata Berimbuhan Menggunakan Algoritma Porter Stemmer. *Seminar Nasional Informatika (SEMNASIF)*, 2018(November), 108–114.
- Sunardi, S., Yudhana, A., & Mukaromah, I. A. (2019). Implementasi Deteksi Plagiarisme Menggunakan Metode N-Gram Dan Jaccard Similarity Terhadap Algoritma Winnowing. *Transmisi*, 20(3), 105. <https://doi.org/10.14710/transmisi.20.3.105-110>
- Wibowo, J. (2016). Pada Kalimat Bahasa Indonesia Dengan Algoritma Stemming. *Jurnal Riset Komputer (JURIKOM)*, 3(5), 346–350.
- Yesi Nur Cahyati. (2015). Analisis Kelas Kata Dan Pola Kalimat Pada Tulisan 2 Cerita Pendek Siswa Kelas Viii C Smp Negeri 2 Sawit 3 Boyolali, 151, 10–17. <https://doi.org/10.1145/3132847.3132886>
- Yosephine, M., & Prabowo, Y. D. (2017). Pengembangan Aplikasi Pemeriksaan Kata Dasar dan Imbuhan pada Bahasa Indonesia. *Jurnal Sains Dan Teknologi*, 4, 1–11. Retrieved from <http://research.kalbis.ac.id/Research/Files/Article/Full/53O3ZHIP42P77K0UM3SPBM8N2.pdf>
- Yudhana, A., Djayali, A. D., & Sunardi. (2017). Sistem Deteksi Plagiarisme Dokumen Karya Ilmiah Dengan Algoritma Pencocokan Pola.