

***POS – TAGGING BAHASA INDONESIA BERBASISKAN ATURAN  
DAN LEKSIKON UNTUK KELAS KATA BENDA DAN ADJEKTIVA***

Skripsi



oleh

**YOSEP ADI KRISTANTO**

**71130058**

**PROGRAM STUDI TEKNIK INFORMATIKA FAKULTAS TEKNOLOGI INFORMASI  
UNIVERSITAS KRISTEN DUTA WACANA**

2017

**POS – TAGGING BAHASA INDONESIA BERBASISKAN ATURAN  
DAN LEKSIKON UNTUK KELAS KATA BENDA DAN ADJEKTIVA**

Skripsi



Diajukan kepada Program Studi Teknik Informatika Fakultas Teknologi Informasi

Universitas Kristen Duta Wacana

Sebagai Salah Satu Syarat dalam Memperoleh Gelar

Sarjana Komputer

Disusun oleh

**YOSEP ADI KRISTANTO**

**71130058**

PROGRAM STUDI TEKNIK INFORMATIKA FAKULTAS TEKNOLOGI INFORMASI

UNIVERSITAS KRISTEN DUTA WACANA

2017

## PERNYATAAN KEASLIAN SKRIPSI

Saya menyatakan dengan sesungguhnya bahwa skripsi dengan judul:

### **POS - TAGGING BAHASA INDONESIA BERBASISKAN ATURAN DAN LEKSIKON UNTUK KELAS KATA BENDA DAN ADJEKTIVA**

yang saya kerjakan untuk melengkapi sebagian persyaratan menjadi Sarjana Komputer pada pendidikan Sarjana Program Studi Teknik Informatika Fakultas Teknologi Informasi Universitas Kristen Duta Wacana, bukan merupakan tiruan atau duplikasi dari skripsi keserjanaan di lingkungan Universitas Kristen Duta Wacana maupun di Perguruan Tinggi atau instansi manapun, kecuali bagian yang sumber informasinya dicantumkan sebagaimana mestinya.

Jika dikemudian hari didapati bahwa hasil skripsi ini adalah hasil plagiasi atau tiruan dari skripsi lain, saya bersedia dikenai sanksi yakni pencabutan gelar keserjanaan saya.

Yogyakarta, 23 Oktober 2017



YOSEP ADI KRISTANTO  
71130058

## HALAMAN PERSETUJUAN

Judul Skripsi : POS - TAGGING BAHASA INDONESIA  
BERBASISKAN ATURAN DAN LEKSIKON  
UNTUK KELAS KATA BENDA DAN ADJEKTIVA

Nama Mahasiswa : YOSEP ADI KRISTANTO

N I M : 71130058

Matakuliah : Skripsi (Tugas Akhir)

Kode : TIW276

Semester : Gasal

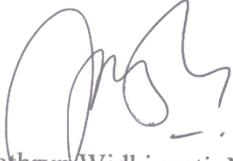
Tahun Akademik : 2017/2018

Telah diperiksa dan disetujui di  
Yogyakarta,  
Pada tanggal 23 Oktober 2017

Dosen Pembimbing I

  
Lucia Dwi Krisnawati, Dr.

Dosen Pembimbing II

  
Kathryn Widhiyanti, M.Cs.

## HALAMAN PENGESAHAN

### POS - TAGGING BAHASA INDONESIA BERBASISKAN ATURAN DAN LEKSIKON UNTUK KELAS KATA BENDA DAN ADJEKTIVA

Oleh: YOSEP ADI KRISTANTO / 71130058

Dipertahankan di depan Dewan Penguji Skripsi  
Program Studi Teknik Informatika Fakultas Teknologi Informasi  
Universitas Kristen Duta Wacana - Yogyakarta  
Dan dinyatakan diterima untuk memenuhi salah satu syarat memperoleh gelar  
Sarjana Komputer  
pada tanggal 18 Oktober 2017

Yogyakarta, 23 Oktober 2017  
Mengesahkan,

Dewan Penguji:

1. Lucia Dwi Krisnawati, Dr.
2. Kathryn Widhiyanti, M.Cs.
3. Antonius Rachmat C., S.Kom.,M.Cs.
4. Aditya Wikan Mahastama, S.Kom., M.Cs.

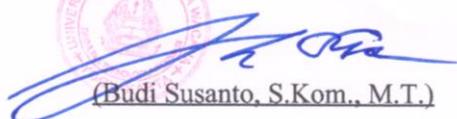
*Lucia*

*Kathryn*

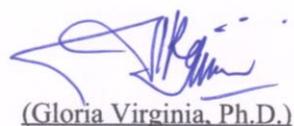
*Antonius*

*Aditya*

Dekan

  
(Budi Susanto, S.Kom., M.T.)

Ketua Program Studi

  
(Gloria Virginia, Ph.D.)

## UCAPAN TERIMA KASIH

Dalam menyelesaikan pembuatan program dan penyusunan laporan Tugas Akhir ini, penulis telah menerima banyak bantuan yang berupa bimbingan, saran dan masukan dari berbagai pihak baik secara langsung maupun secara tidak langsung. Untuk itu, dengan segala kerendahan hati, pada kesempatan ini penulis menyampaikan ucapan terima kasih kepada:

1. **Ibu Lucia Dwi Krisnawati, Dr.** dan **Ibu Kathryn Widhiyanti, M.Cs.** selaku dosen pembimbing I dan dosen pembimbing II yang telah memberikan nasehat, saran, dan masukan serta mengoreksi kesalahan-kesalahan yang terjadi dalam proses pengerjaan Tugas Akhir.
2. **Bapak Antonius Rachmat C., S.Kom., M.Cs.** selaku koordinator Tugas Akhir.
3. Tim Himpunan Sastra Sanata Dharma (*HSSD*) dan Tim Pendidikan Bahasa Sastra Indonesia (*PBSI*) yang telah bekerjasama dengan penulis dalam proses pengujian data.
4. Keluarga terkasih yang selalu memberikan semangat, dorongan dan mendoakan agar dapat menyelesaikan Tugas Akhir Bapak Yusuf M. Kristanto, Ibu Lydia Anijati, adik laki-laki Joash Ananda Kristanto.
5. Kekasih yang tersayang yang selalu setia memberikan dukungan dan semangat dalam suka maupun duka Emylia Intan L.
6. Sahabatku yang telah memberikan bantuan doa, dukungan, semangat yang selalu menjadi tempat berkeluh kesah selama proses pengerjaan Tugas Akhir ini Alvin Tanjung D., Zefanya Widya P., Amsal Maestro, Yonatan Adi N., Yohanes, Yudhit Anindita dan sahabat lainnya yang tidak dapat disebutkan satu persatu.
7. Teman-teman jurusan Teknik Informatika dan Sistem Informasi UKDW angkatan 2013 yang senantiasa ada untuk memberikan dukungan dan bersama-sama berjuang untuk menyelesaikan Tugas Akhir.

8. Terakhir, penulis hendak menyapa setiap nama yang tidak dapat penulis cantumkan satu per satu, terima kasih atas doa yang senantiasa mengalir tanpa sepengetahuan penulis.

Dalam penyusunan Tugas Akhir ini, tentunya penulis masih memiliki banyak kekurangan pada topik Tugas Akhir ini dan masih terdapat banyak kekurangan dalam penulisan Tugas Akhir ini.

Oleh karena itu, penulis sangat menghargai dan menerima jika ada berbagai masukan dari para pembaca baik berupa kritik maupun saran yang sifatnya membangun demi penyempurnaan penulisan-penulisan Tugas Akhir di masa yang akan datang. Akhir kata penulis ingin meminta maaf apabila terdapat kesalahan dalam penyusunan laporan maupun yang pernah penulis lakukan sewaktu pelaksanaan Tugas Akhir ini.

Yogyakarta, 29 September 2017

Penulis

## KATA PENGANTAR

Puji Syukur kepada Tuhan Yang Esa atas berkat dan karunianya selama pengerjaan Tugas Akhir sehingga penulis dapat membuat serta menyelesaikan skripsi berjudul “*POS – Tagging Bahasa Indonesia Berbasiskan Aturan dan Leksikon untuk Kelas Kata Benda dan Adjektiva*” dengan lancar.

Tugas Akhir merupakan salah satu syarat wajib dalam memperoleh gelar Sarjana Komputer dalam Program Studi Teknik Informatika Universitas Kristen Duta Wacana. Penulisan Laporan Tugas Akhir ini bertujuan untuk memberikan laporan tentang penelitian yang telah dilakukan sehingga dapat bermanfaat dan menjadi sumber referensi untuk pengembangan selanjutnya.

Penulis menyadari bahwa penelitian dan laporan Tugas Akhir ini belum sempurna dalam berbagai hal. Oleh karena itu, penulis mengharapkan kritik dan saran yang membangun dari semua pihak agar penelitian dan laporan ini menjadi lebih baik. Terima kasih.

Yogyakarta, 29 September 2017

Penulis

## INTISARI

Kelas kata adalah golongan kata dalam satuan bahasa berdasarkan kategori bentuk, fungsi, dan makna dalam sistem gramatikal. *Part of Speech Tagging (POS Tagging)* adalah suatu proses yang memberikan label kelas kata secara otomatis pada suatu kata dalam kalimat. Penelitian ini melakukan implementasi pelabelan kelas kata benda dan kelas kata adjektiva. Penelitian ini adalah penelitian tahap awal atau dasar dan dapat berguna untuk penelitian *information retrieval*, *language generator*, *information extraction*, *summarization* serta *machinetranslation*. Sistem ini dibangun menggunakan metode berbasis aturan bahasa (*grammar*), untuk bahasa Indonesia. Penggunaan metode tersebut adalah berguna untuk dapat merumuskan pelabelan kelas kata benda dan kata adjektiva untuk kata yang akan diolah oleh sistem *POS Tagging* ini. Tujuan dari penelitian ini adalah untuk dapat merumuskan aturan *tagging* yang benar pada sistem *POS Tagging* untuk kata dalam bahasa Indonesia khususnya pada kata benda + adjektiva. Pengujian dilakukan menggunakan *precision*, *recall*. Hasil yang didapat untuk *precision*, *recall*, dari 35 artikel dokumen yakni 96%.

Kata Kunci: *Part of Speech Tagging*, Gramatikal, Pelabelan, Kelas Kata, Metode *Rule Based*

## DAFTAR ISI

HALAMAN JUDUL .....	
PERNYATAAN KEASLIAN SKRIPSI.....	iii
HALAMAN PERSETUJUAN.....	iv
HALAMAN PENGESAHAN .....	v
UCAPAN TERIMA KASIH.....	v
KATA PENGANTAR .....	viii
INTISARI .....	ix
DAFTAR ISI.....	x
DAFTAR TABEL.....	xi
DAFTAR GAMBAR .....	xii
DAFTAR LAMPIRAN.....	xivv
<b>BAB 1</b> .....	1
1.1. Latar Belakang Masalah .....	1
1.2. Perumusan Masalah .....	2
1.3. Batasan Masalah .....	2
1.4. Tujuan Penelitian .....	3
1.5. Metode Penelitian .....	3
1.6. Sistematika Penulisan .....	5
<b>BAB 2</b> .....	6
2.1. Tinjauan Pustaka .....	6
2.2. Landasan Teori.....	7
2.2.1. Kelas Kata.....	7
2.2.2. Metode .....	13
2.2.3. Tagset.....	14
2.2.4. Leksikon .....	15
2.2.5. Normalisasi .....	17
2.2.6. Tokenisasi .....	17
<b>BAB 3</b> .....	18

3.1. Perancangan Sistem .....	18
3.1.1. Perancangan Masukan .....	19
3.1.2. Perancangan Proses .....	20
3.1.3. Perancangan Keluaran ( <i>Output</i> ) .....	21
3.1.4. Perancangan Leksikon .....	22
3.1.5. Perancangan Antarmuka .....	23
3.2. Skenario Pengujian .....	25
3.2.1. Pelabelan / Anotasi .....	25
3.2.2. Perancangan Tagset .....	26
3.2.3. Ukuran yang dipakai .....	26
<b>BAB 4</b> .....	29
4.1. Implementasi Sistem .....	29
4.2. Alur Penelitian .....	55
4.3. Analisis Sistem .....	57
<b>BAB 5</b> .....	64
5.1. Kesimpulan .....	64
5.2. Saran .....	65
DAFTAR PUSTAKA .....	66
SOURCE CODE .....	Lampiran A
KARTU KONSULTASI .....	Lampiran B
FORM REVISI .....	Lampiran C
LAMPIRAN	

## DAFTAR TABEL

Tabel 2.1 .....	14
Tabel 2.2 .....	16
Tabel 3.1 .....	26
Tabel 4.1 .....	53
Tabel 4.2 .....	55
Tabel 4.3 .....	58
Tabel 4.4 .....	59

©UKDW

## DAFTAR GAMBAR

Gambar 1.1 Ilustrasi proses.....	4
Gambar 3.1 Rancangan halaman utama beserta menu utama.....	19
Gambar 3.2 Rancangan tampilan masukan kalimat.....	20
Gambar 3.3 Diagram alur sistem.....	21
Gambar 3.4 Rancangan tampilan keluaran ( <i>output</i> ) .....	22
Gambar 3.5 Rancangan tampilan halaman depan .....	23
Gambar 3.6 Rancangan tampilan halaman pelabelan.....	24
Gambar 3.7 Rancangan tampilan halaman tentang .....	24
Gambar 3.8 Rancangan tampilan halaman penelitian .....	25
Gambar 4.1 Tampilan awal sistem.....	30
Gambar 4.2 Tampilan pelabelan.....	30
Gambar 4.3 Tampilan hasil setelah proses pelabelan.....	31
Gambar 4.4 Tampilan tentang sistem.....	31
Gambar 4.5 Tampilan penelitian sistem.....	32
Gambar 4.6 Tampilan update jumlah data leksikon.....	32
Gambar 4.7 Flowchart normalisasi.....	33
Gambar 4.8 Flowchart tokenisasi .....	34
Gambar 4.9 Flowchart proses aturan pe-...-an .....	35
Gambar 4.10 Flowchart proses aturan ke-...-an .....	36
Gambar 4.11 Flowchart proses aturan ...-asi .....	37
Gambar 4.12 Flowchart proses aturan ...-itas .....	38
Gambar 4.13 Flowchart proses aturan ...-asme .....	39
Gambar 4.14 Flowchart proses aturan ...-nya .....	40
Gambar 4.15 Flowchart proses aturan berulang atau reduplikasi.....	41
Gambar 4.16 Flowchart proses aturan pe-...-an .....	42
Gambar 4.17 Flowchart proses aturan ke-...-an .....	43
Gambar 4.18 Flowchart proses aturan ...-asi .....	44
Gambar 4.19 Flowchart proses aturan ...-itas .....	45
Gambar 4.20 Flowchart proses aturan ...-asme .....	46

Gambar 4.21 Flowchart proses aturan ...-nya .....	47
Gambar 4.22 Flowchart proses aturan ...-al.....	48
Gambar 4.23 Flowchart proses aturan ...-er .....	49
Gambar 4.24 Flowchart proses aturan ...-if .....	50
Gambar 4.25 Leksikon kata benda.....	52
Gambar 4.26 Leksikon kata adjektiva .....	52
Gambar 4.27 Alur penelitian.....	56

©UKDW

## DAFTAR LAMPIRAN

SOURCE CODE.....	Lampiran A
KARTU KONSULTASI.....	Lampiran B
FORM REVISI .....	Lampiran C

©UKDW

# BAB 1

## PENDAHULUAN

### 1.1. Latar Belakang Masalah

Kata merupakan bagian dari kalimat yang tidak dapat dipisahkan. Kalimat adalah satuan bahasa terkecil, dalam wujud lisan atau tulisan, yang mengungkapkan pikiran yang utuh (Alwi, Dardjowidjojo, & Moeliono, 2003). Peranan dan penggunaan klasifikasi kata sangat diperlukan agar dapat diketahui kategori kata tersebut ke dalam kelas kata dan jenis – jenis kata. Dalam hal ini konteks penggunaan kata dalam bahasa Indonesia yang baik menjadi sangat penting. Saat ini masih terdapat banyak sekali penggunaan bahasa yang masih kurang digunakan dengan maksimal. Untuk itu kelas kata berperan penting penggunaannya dalam bahasa Indonesia. Dalam era teknologi saat ini pengolahan kata yang bekerja pada sebuah mesin sangat diperlukan, agar mesin dapat mengenali setiap bahasa yang ada. Pada permasalahan tersebut diperlukan sebuah proses yang menggunakan sebuah metode *rule based* sebagai aturannya (Jurafsky, 2000).

Sebuah sistem yang berguna bagi pengetahuan mesin tersebut hadir untuk pengolahan kata. Sistem tersebut adalah *Part of Speech Tagging (POS Tagging)* adalah suatu proses yang memberikan label kelas kata secara otomatis pada suatu kata dalam kalimat (Jurafsky, 2000). Nantinya hasil atau *output* yang didapatkan dari sistem *POS Tagging* tersebut akan berbentuk teks yang terlabeli kelas kata. Kumpulan teks – teks yang terlabeli tersebut harapannya dapat menjadi sebuah korpus yang sudah terlabeli dengan kelas kata.

*POS Tagging* hadir dengan memberi peran penting dalam proses pengolahan kata bahasa Indonesia yang baik agar dapat dikenali oleh mesin. Harapannya bermanfaat untuk berbagai hal seperti : *information retrieval*, *language generator*, *information extraction*, *summarization* serta *machine translation*. Sebagai contoh penerapan *POS Tagging* dalam sebuah kalimat adalah sebagai berikut : “Alex/NNP bermain/VB bola/NN dilapangan/VB”, maka

penjelasannya berdasarkan penamaan kelas kata adalah NNP (*propernoun*), VB (kata kerja), NN (kata benda) (M., Pragantha, & Purnamasari, 2012). Dalam studi kasus yang akan ditangani tersebut terlihat mudah, namun pada kenyataannya dalam melakukan pelabelan tersebut adalah mesin, maka dari itu dalam penanganannya sebelum melakukan pelabelan kelas kata, sistem perlu melakukan training terlebih dahulu. *Training* dalam sistem tersebut diperoleh dari penerapan *POS Tagging* dengan metode berbasis aturan (*rule based*). Selanjutnya setelah dihasilkan kata yang sudah terlabeli tersebut, dimasukkan ke dalam leksikon. Leksikon merupakan daftar kata yang disusun seperti kamus, tetapi dengan penjelasan yang singkat dan praktis (Hardiyanto, 2008).

Penulis dalam penelitian *POS Tagging* ini menggunakan metode berbasis aturan (*rule based*), karena untuk melabeli kelas kata berbahasa Indonesia yang tepat diperlukan aturan dalam pelabelan tersebut, sehingga akan menghasilkan sebuah korpus (Jurafsky, 2000). Dalam penelitian ini, kata yang akan dianalisa adalah kata benda + adjektiva.

## **1.2. Perumusan Masalah**

Berdasarkan pada masalah yang dijelaskan pada latar belakang tersebut, maka didapat perumusan masalah sebagai berikut :

1. Bagaimana mengimplementasi *Part of Speech Tagging* untuk bahasa Indonesia dengan menggunakan metode berbasis aturan (*rule based*)?
2. Bagaimana menyusun aturan – aturan *tagging* kelas kata untuk kata benda + adjektiva?

## **1.3. Batasan Masalah**

Dalam penelitian ini, penulis memberikan batasan masalah untuk sistem yang akan dibuat :

1. Bahasa yang digunakan dalam penelitian ini adalah bahasa Indonesia.
2. Inputan pada sistem ini hanya berbentuk teks.

3. Menggunakan data leksikon dari Kamus Besar Bahasa Indonesia.
4. Kata yang diinputkan tidak dapat menggunakan bahasa Inggris

#### **1.4. Tujuan Penelitian**

Melalui penelitian ini, harapannya tujuan dari penelitian ini adalah sebagai berikut :

1. Dapat merumuskan aturan *tagging* yang benar pada sistem *POS Tagging* untuk kata dalam bahasa Indonesia khususnya pada kata benda + adjektiva.
2. Dapat mengetahui kelas kata pada suatu dokumen berbahasa Indonesia, yang dihasilkan oleh sistem *POS Tagging* menggunakan metode *rule based*.

#### **1.5. Metode Penelitian**

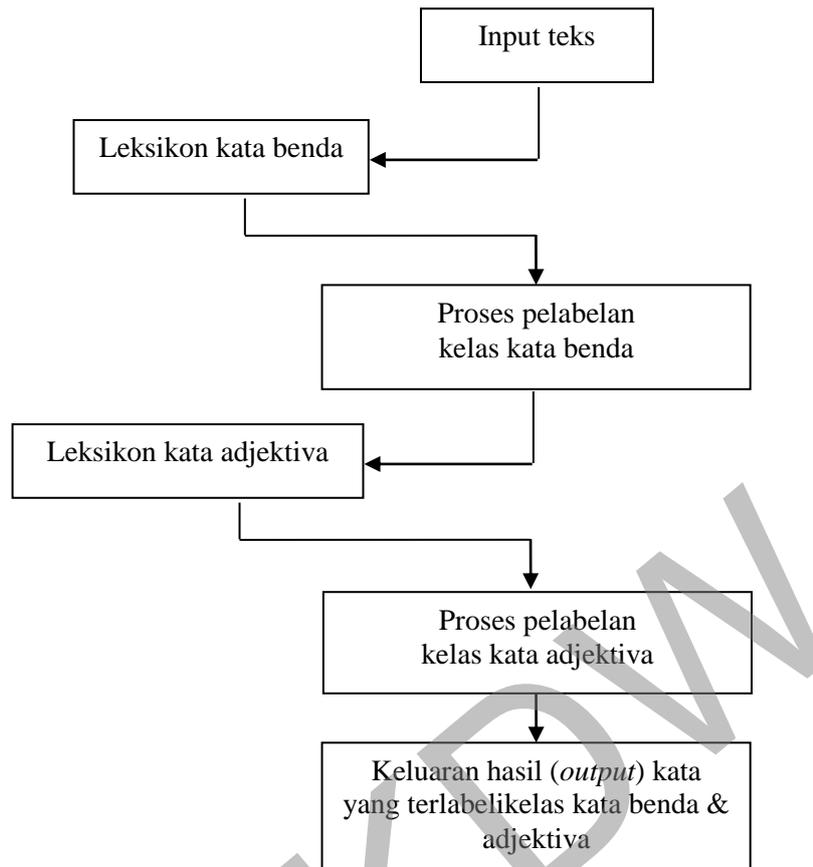
Dalam penelitian ini penulis menggunakan metode, perancangan, implementasi dan penelitian terhadap sistem yang dibuat. Berikut ini adalah penjelasannya :

##### **1. Studi Pustaka**

Studi Pustaka dilakukan dengan mempelajari teori – teori melalui buku, artikel, jurnal dan bahan lain yang mendukung yang berhubungan dengan penggunaan berbasis aturan (*rule based*).

##### **2. Perancangan Sistem**

Pada tahap ini merupakan tahap pembuatan program *Part of Speech Tagging* menggunakan berbasis aturan (*rule based*). Leksikon yang akan dibangun mengacu pada KBBI (Kamus Besar Bahasa Indonesia). Program ini akan dibuat menggunakan bahasa pemrograman PHP dengan tool sublime text 2.



Gambar 1.1 Ilustrasi Proses

### 3. Testing

Pada tahap ini dilakukan pengujian terhadap program *Part of Speech Tagging* dengan kata terbuka menggunakan berbasis aturan (*rule based*) apakah dapat memberi pelabelan atau pengkategorian kelas kata. Pengukuran testing program menggunakan *precision* dan *recall*.

### 4. Analisis Hasil Percobaan dan Evaluasi

Pada tahap ini kesimpulan dapat ditarik berdasarkan pada tingkat akurasi hasil dari penelitian *POS Tagging* tersebut.

## 1.6. Sistematika Penulisan

Sistematika penulisan terbagi menjadi lima bagian yang dilakukan dalam skripsi ini, Berikut merupakan rincian dari bab beserta sub bab yang akan dijelaskan dalam penelitian ini.

Bab 1 Pendahuluan, pada bab ini menjelaskan mengenai latar belakang masalah, rumusan masalah, batasan masalah, tujuan dilakukan penelitian, manfaat penelitian, metodologi, dan sistematika penulisan.

Bab 2 Landasan Teori, pada bab ini memuat sejarah pos tagging bahasa indonesia, tinjauan pustaka, dan menjelaskan mengenai aturan-aturan yang digunakan dalam sistem.

Bab 3 Perancangan Sistem, membahas mengenai perancangan sistem secara keseluruhan yang meliputi algoritma yang digunakan dalam sistem, serta rancangan output sistem.

Bab 4 Implementasi dan Analisis Sistem, bab ini penulis akan mengimplementasikan rancangan sistem dalam bentuk aplikasi. Penulis juga akan membuat sebuah strategi implementasi. Hasil pengolahan data akan tercantum dan dibahas pada bab ini.

Bab 5 Kesimpulan dan Saran, pada bab ini berisi kesimpulan dari rangkaian penelitian yang dikerjakan lengkap dengan kelebihan dan kekurangan sistem.

## BAB 5

### KESIMPULAN DAN SARAN

#### 5.1. Kesimpulan

Berdasarkan pada hasil penelitian yang dilakukan, maka dapat disimpulkan :

- a. Untuk pengujian artikel berita yang memiliki hasil tinggi didapatkan dari portal kompas.com yaitu *recall* sebesar 100% dan *precision* sebesar 100%. Sedangkan untuk hasil terendah artikel yang didapatkan yaitu *recall* sebesar 90% dan *precision* sebesar 92%.
- b. Presentase *recall* paling rendah adalah 87% dengan *precision* sebesar 92%. pada pengujian artikel yang diambil dari antaranews.com.
- c. Sistem juga masih belum akurat 100% dalam melakukan pelabelan kelas kata oleh karena leksikon yang harus dikembangkan lagi dan keterbatasan aturan perumusan dalam kelas kata benda dan kelas kata adjektiva yang memang hanya sedikit.
- d. Sistem dapat melakukan pelabelan kata benda genetif, seperti : -ku, -mu, -nya dengan pemisahan dengan kata benda dasar.
- e. Untuk kata benda yang memiliki suku kata belakang (-ku), program masih memiliki kelemahan dalam melakukan pelabelan karena menggunakan algoritma Regex (*Regular Expression*) yang dasarnya mengenali berdasarkan pada pola kata.
- f. Anotator dalam pengujian sistem masih melakukan banyak kesalahan dalam pelabelan manual oleh manusia.
- g. Penyebab anotator melakukan kesalahan salah satunya adalah karena anotator melakukan pelabelan berdasarkan pewarnaan untuk melabeli kelas kata.

- h. Pada anotator ke-2 dan ke-4 menghasilkan hasil keakurasian yang rendah. Hal tersebut dipengaruhi oleh faktor-faktor anotator yang menyebabkan rendahnya hasil yang diperoleh. Seperti tergesa-gesa dalam melakukan pelabelan.

## 5.2. Saran

Adapun saran yang diberikan untuk pengembangan sistem selanjutnya adalah :

- a. Dapat dikembangkan lagi khususnya untuk media pembelajaran kelas kata bahasa indonesia bagi para pelajar.
- b. Penggunaan tagset yang dapat dirumuskan khusus untuk bahasa indonesia, karena tagset yang saat ini digunakan masih menggunakan tagset berbahasa inggris dari penelitian sebelumnya.
- c. Pengembangan data leksikon untuk kelas kata benda dan kata adjektiva yang harus dikembangkan lagi yang dapat mengacu pada KBBI (Kamus Besar Bahasa Indonesia) yang dibangun dibawah pengembangan Badan Pengembangan dan Pembinaan Bahasa, Kementerian Pendidikan dan Kebudayaan Republik Indonesia, karena data leksikon yang terus diperbaharui akan semakin meningkatkan kinerja sistem.
- d. Untuk penelitian selanjutnya dapat memperbaiki pelabelan dengan menambahkan aturan-aturan baru yang dapat digunakan oleh sistem *POS Tagging*.
- e. Dapat mempercepat pencarian kata dalam leksikon secara efisien menggunakan dasar-dasar dan teknik untuk mempercepat pencarian kata, menggunakan *filter* abjad dalam pembangunan leksikon.

## DAFTAR PUSTAKA

- Alwi, H., Dardjowidjojo, S., & Moeliono, A. M. (2003). *Tata Bahasa Baku Bahasa Indonesia*. Jakarta: Balai Pustaka.
- Budiasih, Ni Nyoman, Wirayuda, Tjokorda A. B., Dayawati Retno N. (2009). *Analisis dan Implementasi Stemming Teks Berbahasa Indonesia dengan Menggunakan Porter Stemmer*.
- Jurafsky, D. S. (2000). *Speech and Language Processing "An Introduction to Natural Language Processing, Computational Linguistics and Speech Recognition"*. New Jersey: Prentice-Hall.
- Karibun H.S., M. (2005). *Penerapan Metode Unsupervised Learning Rule-Based*.
- Kridalaksana, H. (2005). *Kelas Kata Dalam Bahasa Indonesia*. Jakarta: Gramedia.
- M., V. C., Pragantha, J., & Purnamasari, E. (2012). Jurnal Teknik dan Ilmu Komputer. *Implementasi Brill Tagger untuk memberikan POS Tagging pada dokumen bahasa indonesia*.
- Pisceldo F., Adriano M., Manurung R. (2009). *Probabilistic Part Of Speech Tagging for Bahasa Indonesia*
- Utami, E., & Hartati, S. (2007). ResearchGate. *Pendekatan Metode Rule Based dalam Mengalihbahasakan Teks Bahasa Inggris ke Teks Bahasa Indonesia*, 42-53.
- Widhiyanti, K., & Harjoko, A. (2012). INFORMATIKA. *POS Tagging Bahasa Indonesia Dengan HMM dan Rule Based*.