

# Deteksi Struktur Kalimat Tunggal Bahasa Indonesia Menggunakan

## *POS Tagging dan Algoritma LALR Parser*

Diajukan untuk Menyusun Skripsi  
di Jurusan Teknik Informatika Fakultas Ilmu Komputer UNSRI



Oleh :

GIAN AGNAR MOSLEM  
NIM : 09021181520028

**Jurusan Teknik Informatika  
FAKULTAS ILMU KOMPUTER UNIVERSITAS SRIWIJAYA  
2020**

## LEMBAR PENGESAHAN TUGAS AKHIR

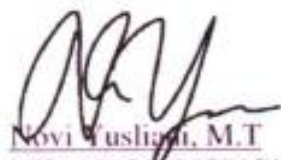
Deteksi Struktur Kalimat Tunggal Bahasa Indonesia Menggunakan POS Tagging  
dan Algoritma LALR Parser

Oleh:

GIAN AGNAR MOSLEM  
NIM : 09021181520028

Palembang, Januari 2021

Pembimbing,



Novi Yulianti, M.T  
NIP. 198211082012122001

Mengetahui,

Ketua Jurusan Teknik Informatika,



Alvi Syahrini Utami, M.Kom.  
NIP. 197812222006042003

## TANDA LULUS UJIAN SIDANG TUGAS AKHIR

Pada hari Jumat tanggal 29 Desember 2020 telah dilaksanakan ujian sidang tugas akhir oleh Jurusan Teknik Informatika Fakultas Ilmu Komputer Universitas Sriwijaya.

Nama : Gian Agnar Moslem  
NIM : 09021181520028  
Judul : Deteksi Struktur Kalimat Tunggal Bahasa Indonesia  
Menggunakan *POS Tagging* dan Algoritma *LALR Parser*

### 1. Pembimbing I

Novi Yusliani, M.T  
NIP. 198211082012122001



### 2. Penguji I

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




### 3. Penguji II

Kanda Januar Miraswan, S.Kom., M.T  
NIP. 199001092019031012



Mengetahui,  
Ketua Jurusan Teknik Informatika



  
Alvi Syahrini Utami, M.Kom  
NIP. 197812222006042003

## HALAMAN PERNYATAAN

Yang bertanda tangan di bawah ini :

Nama : Gian Agnar Moslem  
NIM : 09021181520028  
Program Studi : Teknik Informatika  
Judul Skripsi : Deteksi Struktur Kalimat Tunggal Bahasa Indonesia  
Menggunakan *POS Tagging* dan Algoritma *LALR Parser*

Hasil Pengecekan Software *iThenticate/Turnitin* : 20%

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 siapapun.

Palembang, Januari 2021



Gian Agnar Moslem  
NIM. 09021181520028

## **MOTTO DAN PERSEMBAHAN**

### **MOTTO:**

"You might not have any desire to live, but don't show that to everyone else struggling to survive. We should live more decently as human beings."

### **Kupersembahkan Karya Tulis ini kepada:**

- **Keluargaku tercinta**
- **Orang-orang tersayangku**
- **Sahabat-sahabatku**
- **Fakultas Ilmu Komputer,  
Universitas Sriwijaya**

# **Detection of Indonesian Sentence Structure Using POS Tagging and LALR Parser Algorithm**

**By :**  
**Gian Agnar Moslem**  
**09021181520028**

## **ABSTRACT**

The presentation of languages in written form can save knowledge from generation to generation. Every writing of a paper such a book, should be written in a correct grammar according to the rules of Indonesian Language, because it will affect the perfection of the process of delivering ideas or messages. This research was conducted to help examine the Indonesian sentence structure, with the aim of helping to examine the sentence structure of the Indonesian language using the POS Tagging and LALR Parser. POS Tagging is a technique used to identify the class of words that are in a sentence, such a nouns, verbs, adjectives and others. LALR Parsrer is one of the methods used in examinig sentence structure. By the 150 sentences tested on the system, 124 sentences show the correct sentence structure. Whereas 26 sentences show the wrong sentence structure. Detection failure caused by several factors such as sentences structure not in accordance with the rules of Indonesian language and the words contained in sentence are not available on the corpus.

**Keywords:** Sentence Structure, POS Tagging, LALR Parser

# **Deteksi Struktur Kalimat Tunggal Bahasa Indonesia Menggunakan POS Tagging dan Algoritma LALR Parser**

**Oleh :**  
**Gian Agnar Moslem**  
**NIM: 09021181520028**

## **ABSTRAK**

Penyajian bahasa dalam bentuk tulisan dapat menyimpan pengetahuan dari generasi ke generasi. Setiap penulisan karya tulis seperti buku, hendaknya ditulis dengan tata bahasa yang benar sesuai kaidah Bahasa Indonesia, karena akan berpengaruh pada kesempurnaan proses penyampaian ide atau pesan. Penelitian ini dilakukan untuk membantu mendeteksi struktur kalimat Bahasa Indonesia dengan menggunakan *POS Tagging* dan *LALR Parser*. *POS Tagging* merupakan suatu teknik yang digunakan untuk mengidentifikasi kelas kata yang ada dalam suatu kalimat, seperti kata benda, kata kerja, kata sifat dan lain-lain. *LALR Parser* merupakan salah satu cara yang digunakan dalam pendeteksian struktur kalimat. Dari 150 kalimat yang diuji pada sistem, 124 kalimat menunjukkan struktur kalimat yang benar. Sedangkan 26 kalimat menunjukkan struktur kalimat yang salah. Kegagalan deteksi disebabkan oleh beberapa faktor seperti struktur kalimat tidak sesuai dengan kaidah bahasa Indonesia dan kata yang terdapat pada kalimat tidak tersedia maupun sesuai pada korpus

**Kata Kunci :** Struktur Kalimat, *POS Tagging*, *LALR Parser*

## **KATA PENGANTAR**

Puji syukur kepada Allah SWT atas berkat dan rahmat-Nya yang telah melimpahkan rahmat dan karunianya kepada Penulis sehingga mampu menyelesaikan Tugas Akhir ini dengan baik. Tugas akhir ini disusun untuk memenuhi salah satu syarat untuk menyelesaikan Pendidikan program Strata-1 pada Fakultas Ilmu Komputer Program Studi Teknik Informatika di Universitas Sriwijaya.

Dalam menyelesaikan Tugas Akhir ini, banyak pihak yang telah memberikan bantuan dan dukungan baik secara langsung maupun tidak langsung. Oleh karena itu, penulis ingin menyampaikan rasa terima kasih kepada:

1. Orang tuaku, Solihin dan Yuna Sariyanti yang jasanya tiada tara dan tak akan dapat tergantikan. Kedua adikku Radithya Maulana Moslem dan Muhammad Fachri Alwafi Moslem, serta seluruh keluarga besarku yang selalu mendoakan, menghibur, serta memberikan dukungan baik secara moril maupun materil.
2. Civitas Akademika Fakultas Ilmu Komputer Universitas Sriwijaya yang telah memberikan ilmu, serta membimbing jalan penulis dalam menghadapi perkuliahan.
3. Pembimbing Tugas Akhir, Ibu Novi Yusliani, MT yang telah membimbing penulis dengan sabar dan semangat.



4. Penguji Tugas Akhir, bapak Dr. Abdiansah, S.Kom., M.Cs dan bapak Kanda Januar Miraswan, S.Kom., M.T yang senantiasa memberikan masukan kepada penulis mengenai tugas akhir dan pengalaman kerja.
5. Pembimbing Akademik, Ibu Novi Yusliani, MT. yang telah banyak memberikan masukan kepada penulis mengenai perkuliahan.
6. Albar, Halim, Tirta, Pribadi, Ayep, Kerel, Blen dan Egan yang telah menemani penulis untuk bermain Dota dan tidak mengerjakan tugas akhir skripsi.
7. Teman-teman perjuangan Tugas Akhir sekaligus nongkrongku Faiz, Ega, Taca yang penulis sendiri tidak tahu apa kontribusi mereka selama pengerjaan tugas akhir skripsi,

Penulis menyadari dalam penyusunan Tugas Akhir ini masih terdapat banyak kekurangan, oleh karena itu kritik dan saran yang membangun sangat diharapkan untuk kemajuan penelitian selanjutnya. Akhir kata semoga Tugas Akhir ini dapat digunakan sebaik-baiknya serta bermanfaat bagi kita semua.

Palembang, Januari 2020

Penulis

Gian Agnar Moslem

## DAFTAR ISI

Halaman

HALAMAN JUDUL .....	i
LEMBAR PENGESAHAN SKRIPSI .....	ii
TANDA LULUS UJIAN SIDANG SKRIPSI .....	iii
HALAMAN PERNYATAAN .....	iv
MOTTO DAN PERSEMBAHAN .....	v
ABSTRACT .....	vi
ABSTRAK .....	vii
KATA PENGANTAR .....	viii
DAFTAR ISI .....	x
DAFTAR TABEL .....	xiii
DAFTAR GAMBAR .....	xiv
DAFTAR LAMPIRAN .....	xvii
BAB I PENDAHULUAN	
1.1 Pendahuluan .....	I-1
1.2 Latar Belakang .....	I-1
1.3 Rumusan Masalah .....	I-4
1.4 Tujuan Penelitian .....	I-5
1.5 Manfaat Penelitian .....	I-5
1.6 Batasan Masalah .....	I-5
1.7 Sistematika Penulisan .....	I-6
1.8 Kesimpulan .....	I-6
BAB II KAJIAN LITERATUR	
2.1 Pendahuluan .....	II-1
2.2 Landasan Teori .....	II-1
2.2.1 Struktur Kalimat Bahasa Indonesia .....	II-1
2.2.2 <i>Tokenizing</i> .....	II-3
2.2.3 <i>POS Tagging</i> .....	II-3
2.2.4 <i>LALR Parser</i> .....	II-5
2.3 Penelitian Lain Yang Relevan .....	II-9
2.3.1 <i>Extraction of Link Context using Tag Tree</i> <i>and LALR Parsing</i> .....	II-9
2.3.2 Analisis perbandingan Algoritma <i>LCP (Left-Corner-Parsing)</i> dan Algoritma <i>CYK (Cocke-Younger-Kasami)</i> untuk Memeriksa Pola Kalimat Baku Bahasa Indonesia .....	II-9
2.4 Kesimpulan .....	II-10
BAB III METODOLOGI PENELITIAN	
3.1 Pendahuluan .....	III-1
3.2 Metode Pengumpulan Data .....	III-1
3.2.1 Jenis dan Sumber Data .....	III-1

3.2.2	Metode Pengumpulan Data .....	III-2
3.3	Tahapan Penelitian.....	III-2
3.3.1	Menetapkan Kerangka Kerja .....	III-2
3.3.2	Menetapkan Kriteria Pengujian .....	III-16
3.3.3	Menetapkan Format Data Pengujian .....	III-17
3.3.4	Menentukan Alat yang Digunakan dalam Pelaksanaan Penelitian .....	III-17
3.3.5	Melakukan Pengujian Penelitian .....	III-18
3.3.6	Melakukan Analisis Hasil Pengujian dan Membuat Kesimpulan Penelitian .....	III-18
3.4	Metode Pengembangan Perangkat Lunak .....	III-19
3.4.2	Fase Insepsi .....	III-19
3.4.3	Fase Elaborasi .....	III-20
3.4.4	Fase Konstruksi .....	III-20
3.4.5	Fase Transisi.....	III-21
3.5	Manajemen Proyek Penelitian .....	III-21
3.6	Kesimpulan .....	III-22
<b>BAB IV PENGEMBANGAN PERANGKAT LUNAK</b>		
4.1	Pendahuluan.....	IV-1
4.2	Fase Insepsi.....	IV-1
4.2.1	Pemodelan Bisnis .....	IV-1
4.2.2	Kebutuhan Sistem.....	IV-3
4.2.2.1	Fungsi Praproses Data.....	IV-3
4.2.2.2	Fungsi Evaluasi Pendeteksian Struktur Kalimat... ..	IV-5
4.2.3	Analisis dan Desain .....	IV-6
4.3	Fase Elaborasi .....	IV-7
4.3.1	Pemodelan Bisnis.....	IV-7
4.3.2	Kebutuhan Sistem.....	IV-10
4.3.3	Analisis Desain .....	IV-10
4.4	Fase Konstruksi.....	IV-12
4.4.1	Pemodelan Bisnis.....	IV-12
4.4.2	Kebutuhan Sistem.....	IV-15
4.5	Fase Transisi .....	IV-17
4.5.1	Pemodelan Bisnis.....	IV-17
4.5.2	Kebutuhan.....	IV-17
4.5.3	Pengujian .....	IV-18
4.5.3.1	Rencana Pengujian .....	IV-18
4.5.3.2	Pengujian Perangkat Lunak .....	IV-19
4.6	Kesimpulan .....	IV-24
<b>BAB V HASIL DAN ANALISIS PENELITIAN</b>		
5.1	Pendahuluan.....	V-1
5.2	Data Hasil Penelitian .....	V-1
5.2.1	Pengujian Tahap Pertama .....	V-1
5.2.2	Pengujian Tahap Kedua .....	V-5
5.3	Kesimpulan .....	V-6
<b>BAB VI KESIMPULAN DAN SARAN</b>		

6.1 Pendahuluan.....	VI-1
6.2 Kesimpulan .....	VI-1
6.3 Saran .....	VI-1

DAFTAR PUSTAKA .....	xvi
----------------------	-----

## DAFTAR TABEL

	Halaman
II-1. Tabel Pola Kalimat Dasar .....	II-7
II-2. Tabel Label Kata .....	II-7
II-3. Tabel Aturan Produksi .....	II-8
II-4. Tabel <i>Action Table</i> .....	II-8
II-5. Tabel <i>Goto Table</i> .....	II-8
III-1. Tabel Label Kata untuk Subjek .....	III-4
III-2. Tabel Label Kata untuk Prdikat .....	III-5
III-3. Tabel Label Kata untuk Objek .....	III-6
III-4. Tabel Label Kata untuk Pelengkap .....	III-7
III-5. Tabel Label Kata untuk Keterangan .....	III-8
III-6. Rancangan Tabel Hasil Pendeteksian Struktur Kalimat .....	III-17
III-7. Rancangan Tabel Hasil Pendeteksian Struktur Kalimat Tunggal .....	III-19
IV-1. Definisi Aktor <i>Use Case</i> .....	IV-2
IV-2. Definisi <i>Use Case</i> .....	IV-3
IV-3. Skenario <i>Use Case</i> Melakukan Praproses Data Teks .....	IV-4
IV-4. Skenario <i>Use Case</i> Mengevaluasi Hasil Pendeteksian Menggunakan LALR Parser .....	IV-6
IV-5. Tabel Implementasi Kelas .....	IV-15
IV-6. Rencana Pengujian <i>Use Case</i> Melakukan Praproses Data Teks .....	IV-18
IV-7. Rencana Pengujian <i>Use Case</i> Mengevaluasi Hasil Pendeteksian Struktur Kalimat dengan LALR Parser .....	IV-19
IV-8. Pengujian Fitur Praproses Data Teks .....	IV-21
IV-9. Pengujian Fitur Evaluasi Hasil Pendeteksian Menggunakan LALR Parser .....	IV-22
V-1. Rancangan Struktur Kalimat dan Contoh Kalimat Berdasarkan Buku Alwi .....	V-1
V-2. Hasil <i>POS Tagging</i> dan Hasi Penguraian Kalimat .....	V-3
V-3. Hasil Evaluasi Presentasi Penguraian Kalimat Tunggal pada Media Cetak Kompas .....	V-5

## DAFTAR GAMBAR

	Halaman
III-1. Gambar LALR Parser.....	III-10
III-2. Gambar Arsitektur Umum.....	III-14
III-3. Gambar Flowchart Sistem Pendeteksian Struktur Kalimat.....	III-15
III-4. Gambar Tahapan Pengujian Penelitian .....	III-18
III-5. Gambar Penjadwalan untuk Tahap Menentukan Ruang Lingkup dan Unit Penelitian .....	III-29
III-6. Gambar Penjadwalan untuk Tahap Menentukan Dasar Teori yang Berkaitan dengan Penelitian .....	III-30
III-7. Gambar Penjadwalan untuk Tahap Menentukan Kriteria Pengujian ...	III-30
III-8. Gambar Penjadwalan untuk Tahap Menentukan Alat yang Digunakan untuk Pelaksanaan Penelitian Fase Insepsi.....	III-31
III-9. Gambar Penjadwalan untuk Tahap Menentukan Alat yang Digunakan untuk Pelaksanaan Penelitian Fase Elaborasi .....	III-31
III-10. Gambar Penjadwalan untuk Tahap Menentukan Alat yang Digunakan untuk Pelaksanaan Penelitian Fase Konstruksi.....	III-32
III-11. Gambar Penjadwalan untuk Tahap Menentukan Alat yang Digunakan untuk Pelaksanaan Penelitian Fase Transisi .....	III-33
III-12. Gambar Penjadwalan untuk Tahap Melakukan Pengujian Penelitian	III-33
III-13. Gambar Penjadwalan untuk Tahap Analisa Hasil Pengujian Penelitian dan Membuat Kesimpulan .....	III-34
IV-1. Diagram <i>Use Case</i> .....	IV-2
IV-2. <i>Sequence Diagram</i> Melakukan Praproses Dokumen.....	IV-8
IV-3. <i>Sequence Diagram</i> Melakukan Evaluasi Hasil Pendeteksian dengan LALR Parser .....	IV-9
IV-4. Rancangan Antar Muka Perangkat Lunak .....	IV-10
IV-5. <i>Activity Diagram</i> Melakukan Praproses Dokumen.....	IV-11
IV-6. <i>Activity Diagram</i> Evaluasi Hasil Pendeteksian Menggunakan LALR Parser .....	IV-12
IV-7. Diagram Kelas.....	IV-14

# **BAB 1**

## **PENDAHULUAN**

### **1.1 Pendahuluan**

Pada bab ini membahas mengenai latar belakang masalah, rumusan masalah, tujuan dan manfaat penelitian, batasan masalah, sistematika penulisan dan kesimpulan. Pada bab ini juga memberikan penjelasan umum mengenai keseluruhan penelitian.

### **1.2 Latar Belakang Masalah**

Sistem kalimat Bahasa Indonesia tersusun menurut struktur kalimatnya, tidak tersusun secara acak atau sembarangan contohnya jenis kalimat berupa ambigu, idiom dan peribahasa. Setiap penulisan karya ilmiah, skripsi, surat kabar, kalimat paragraf, bahkan kalimat tunggal hendaknya ditulis dengan kalimat yang efektif. Harus memperhatikan struktur kalimat dan maknanya karena berpengaruh dalam proses penyampaian dan menerima pesan. Bidang teknologi *Natural Language Processing (NLP)* atau yang biasa disebut dengan pemrosesan bahasa alami dapat digunakan untuk mempermudah pemeriksaan struktur kalimat yang benar (Arif, 2018).

Pemrosesan bahasa alami merupakan suatu topik yang banyak diperbincangkan oleh para ahli komputer. Sebagian besar para ahli ini memiliki pertanyaan yang hampir sama, yaitu seberapa jauh teknologi komputer yang

mampu menerjemahkan atau menginterpretasikan bahasa manusia atau bahasa alami ini ke dalam bahasa komputer (Sager, 1981).

Penelitian sebelumnya melakukan penelitian tentang penguraian kalimat dengan pendekatan *linguistic string analysis* menggunakan aturan *Backus Naus Form (BNF)*. Dari 65 kalimat yang diproses, 56 kalimat berhasil diuraikan. Namun pada penelitian tersebut masih membutuhkan tambahan aturan kalimat atau tata bahasa untuk menangani masalah tata bahasa yang semakin luas (Iqbal *et al.*, 2007).

Penelitian pemrosesan teks pada bahasa Indonesia telah dilakukan yaitu sistem yang menghasilkan simulasi dari pengurai Bahasa Indonesia yang mengadaptasi proses yang digunakan pada algoritma *Collins* yang biasa digunakan untuk pemrosesan pada bahasa Inggris. Pada penelitian tersebut hanya menggunakan sampel kalimat berjumlah 30, dan 28 kalimat berhasil diuraikan. Namun akan memerlukan pengembangan pada sistem untuk menguraikan kalimat dengan jumlah yang besar (Sibarani *et al.*, 2013).

Berdasarkan penelitian pemrosesan teks diatas, proses penguraian (parsing) kalimat adalah kunci utamanya. Untuk memproses suatu bahasa alami, proses penguraian kalimat adalah masalah fundamental antara mesin dan manusia (Hasan *et al.*, 2011). Sehingga pada penelitian ini penulis memperhatikan proses penguraian kalimat untuk mempermudah pemeriksaan struktur kalimatnya.

Proses penguraian kalimat dalam bahasa alami hampir sama dengan proses penguraian tata bahasa pemrograman dalam dunia komputer. Perbedaan yang mendasar pada keduanya adalah tata bahasa dunia komputer merupakan tata bahasa



yang bebas konteks (*context free grammar*), sedangkan tata bahasa pada manusia adalah tata bahasa yang peka terhadap konteks (*context sensitive*).

Pemrosesan bahasa alami merupakan suatu hal yang sangat kompleks. Sangat sulit untuk merancang dan membangun suatu aplikasi komputer yang dapat menganalisa dan memahami bahasa yang biasa dipakai oleh manusia dalam kehidupan sehari-hari. Pemrosesan bahasa alami memerlukan pengetahuan tentang kamus kata dan aturan tata bahasa (Rao *et al.*, 2010). Kamus kata diperlukan untuk memeriksa apakah kata-kata yang dimasukkan merupakan kata-kata yang benar. Aturan tata bahasa diperlukan untuk memeriksa apakah kalimat yang dimasukkan sesuai dengan kaidah tata bahasa dari yang digunakan. Beberapa pemrosesan bahasa memerlukan proses awal (*preprocessing*), diantaranya adalah perlu adanya proses pemberian jenis kata pada setiap kata masukkan yaitu *Part-Of-Speech Tagging*. *Part-Of-Speech Tagging* perlu dilakukan untuk membuat sebuah file korpus masukan untuk pengurai, dan setiap kata pada file tersebut telah diberi tag (Sukanto, 2009).

Penelitian Sebelumnya melakukan penelitian tentang mengekstraksi konteks link dari halaman web menggunakan Tag Tree dan LALR Parser (Gupta, 2013). Hasil ekstraksi sangat berguna untuk mengindeks dokumen. Penelitian berikutnya yaitu perancangan penganalisis struktur kalimat Bahasa Indonesia dengan menggunakan algoritma *Constraint-Based Formalism*. Pada penelitian tersebut dilakukan penguraian dengan aturan-aturan sintaks yang di definisikan (Sulastra, 2014). Penelitian lainnya yaitu analisis perbandingan algoritma LCP (*Left-Corner Parsing*) dan algoritma CYK (*Cocke-Younger-Kasami*) untuk memeriksa pola

kalimat baku Bahasa Indonesia (Susanti, 2016). Hasil pengujian pemeriksaan poa kalimat baku didapatkan tingkat akurasi dari algoritma CYK sebesar 65% dengan laju error 0.35 dan algoritma LCP sebesar 60% dengan laju error 0.40. Namun tingkat akurasi yang didapat belum cukup besar karena kamus POS Tag yang digunakan masih terdapat kekurangan yaitu masih banyak kata yang tidak dapat dikenali.

Dalam skripsi ini, algoritma yang digunakan dalam mendeteksi struktur kalimat tunggal adalah *LALR Parser*. *LALR Parser* memiliki tingkat keberhasilan yang lebih tinggi dari algoritma *LR* lainnya, yaitu mencapai 70% (Pratama, 2017). Namun pada penelitian tersebut belum memanfaatkan *POS Tagging* dalam menentukan jenis kata pada kalimat. Dengan menambahkan *POS Tagging* dipemrosesan awal, maka akan lebih memudahkan sistem dalam mengidentifikasi jenis kata pada kalimat masukan yang akan diproses dan diharapkan agar keberhasilan sistem dalam menguraikan kalimat bahasa Indonesia semakin baik. Berdasarkan permasalahan tersebut, penulis akan melakukan penelitian menggunakan *Part-Of-Speech Tagging* dan algoritma *LALR Parser* untuk menganalisis dan memeriksa struktur kalimat tunggal bahasa Indonesia apakah bisa mendapatkan tingkat keberhasilan dan keakuratan yang baik.

### **1.3 Rumusan Masalah**

Rumusan masalah dalam penelitian ini adalah :

1. Bagaimana cara kerja *LALR Parser* dengan menggunakan *POS Tagging* dalam pendeteksian struktur kalimat tunggal Bahasa Indonesia?

2. Bagaimana tingkat persentase keakuratan *LALR Parser* dengan menggunakan *POS Tagging* dalam pendeteksian struktur kalimat tunggal Bahasa Indonesia?

#### **1.4 Tujuan Penelitian**

Tujuan dari penelitian ini adalah mengetahui seberapa baik cara kerja dan tingkat keberhasilan maupun keakuratan dari algoritma *LALR Parser* dengan menggunakan *POS Tagging* dalam mendeteksi struktur kalimat tunggal Bahasa Indonesia..

#### **1.5 Manfaat Penelitian**

Manfaat penelitian ini adalah:

1. Mempermudah dalam memeriksa struktur kalimat tunggal Bahasa Indonesia dengan sistem komputer.
2. Mengetahui kemampuan *POS Tagging* dan *LALR Parser* dalam pemeriksaan struktur kalimat tunggal Bahasa Indonesia.

#### **1.6 Batasan Masalah**

Adapun batasan masalah pada penelitian ini adalah sebagai berikut:

1. Struktur kalimat yang digunakan pada penelitian ini adalah struktur kalimat yang sesuai dengan tata Bahasa Indonesia berdasarkan ejaan yang disempurnakan.
2. Tidak memeriksa jenis kalimat berupa ambigu, idiom dan peribahasa.
3. Kalimat yang digunakan berupa kalimat tunggal, tidak memeriksa kalimat majemuk setara maupun kalimat majemuk bertingkat.

4. Tidak memeriksa makna kalimat.

## **1.7 Sistematika Penulisan**

Sistematika penulisan ini adalah sebagai berikut:

### **BAB I. PENDAHULUAN**

Bab I menguraikan latar belakang masalah, rumusan masalah, tujuan dan manfaat penelitian, batasan masalah, sistematika penulisan penelitian dan kesimpulan.

### **BAB II. KAJIAN LITERATUR**

Dalam bab ini memaparkan mengenai pembahasan dan dasar-dasar teori yang terkait dengan penelitian serta bahan materi penyusunan penelitian tugas akhir.

### **BAB III. METODOLOGI PENELITIAN**

Pada bab ini membahas mengenai tahap-tahap yang akan dilakukan pada penelitian ini. Pada setiap tahapan penelitian akan dijelaskan secara rinci berdasarkan dengan kerangka kerja dan dilanjutkan dengan perancangan manajemen proyek pada pelaksanaan penelitian.

## **1.8 Kesimpulan**

Berdasarkan penjelasan sebelumnya, maka penelitian ini membangun sebuah sistem yang melakukan pemeriksaan pada struktur kalimat Bahasa Indonesia. Dengan menggunakan *POS Tagging* dan algoritma *LALR Parser* yang digunakan dalam penelitian ini, sehingga akan didapatkan tingkat akurasi dalam memeriksa struktur kalimat Bahasa Indonesia dalam penelitian ini.

## DAFTAR PUSTAKA

- Abka, A.F. 2016. Evaluating the Use of Word Embeddings for Part-of Speech Tagging in Bahasa Indonesia, *International Confrence on Computer, Control, Informatics and its Applications*, pp. 209214
- Allen, J. 1994. Natural Language Understanding. *The Benjamin/Cumming Publishing Company, Inc.* Redwood City.
- Alwi, H, Dardjowidjojo, S., Lapowila, H. & Moeliono, A.M. 1998. *Tata Bahasa Baku Bahasa Indonesia*. Departemen Pendidikan dan Kebudayaan Republik Indonesia: Jakarta.
- Arif. 2018. Sistem Pemeriksaan Struktur Kalimat Pada Teks Bahasa Indonesia Dengan Menggunakan Part-Of-Speech Tagging dan Constraint-Based Formalism. Skripsi. Universitas Sumatera Utara.
- Bermudez, M.E. 1988. *A Unifying Model for Lookahead LR Parsing*. U.S.A: University of Florida.
- Felicia. 2001. *Peranan dan Fungsi Bahasa*. Jakarta: Rineka Cipta.
- Gupta, S. 2013. Extraction of Link Context using Tag Tree and LALR Parsing. *Proceedings of 2013 IEEE Conference on Information and Communication Technologies (ICT 2013), (ICT)*, 253–257.
- Hasan, K. M., Al-Mahmud, Mondal, A. & Saha, A. 2011. Recognizing Bangla Grammar Using Predictive Parser. *International Journal of Computer Science and Information Technology (IJCSIT)* 3 (6): 61-73.
- Iqbal, W.M., Darmantoro, D. & Puspandari, D. 2007. Sistem Penguraian Kalimat Berdasarkan Aturan Sintaks Bahasa Indonesia Dengan Pendekatan Linguistic String Analysis. Skripsi. Universitas Telkom
- Jaber, A. M. 1983. An Implementation of a Parsing Algorithm for LALR Grammars. Thesis. Lehigh University.
- Jurafsky, D. & Martin, J.H. 2000. *Speech and Language Processing – An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition*. Prentice-Hall, Inc.: New Jersey.
- Logothetis, George, Bermudez, Manuel E. 1989. *Relation-based Algorithms for Generating SLR (1) and LALR (1) Parsers*. U.S.A.: AT&T Bell Laboratories and University of Florida.

- Manning, C.D, Raghavan, P. & Schutze, H. 2008. *Introduction to Information Retrieval*. Cambridge: Cambridge University Press.
- Pratama, M.R., Kusumadewi, S. & Hidayat, T. 2017. Penerapan Algoritma *LALR Parser* dan *Context-Free Grammar* Untuk Struktur Kalimat Bahasa Indonesia. *Jurnal Teknologi Elektro* 8(1): 1-7.
- Rao, G., Agarwal, C., Chaudhry, S., Kulkarni, N. & Patil, S.H. 2010. Natural Language Query Processing Using Semantic Grammar. *International Journal on Computer Science and Engineering* 2(2): 219-223.
- Sager, N. 1981. Natural Language Information Processing : A Computer Grammar of English and its Aplications. *Addison-Wesley Publishing Company*. Massachusetts.
- Sibarani, E.M., Nadial, M., Panggabean, E. & Meryana S. 2013. A Study of Parsing Process on Natural Language Processing in Bahasa Indonesia. *16<sup>th</sup> International Conference on Computational Science and Engineering*, pp.309-316.
- Sukamto, R.A. 2009. Penguraian bahasa Indonesia dengan menggunakan pengurai collins.Tesis. Institut Teknologi Bandung.
- Sulastra, I.M.D.J. 2014. Perancangan Penganalisis Struktur Kalimat Bahasa Indonesia dengan Menggunakan *Constraint-Based Formalism*. *Lontar Komputer* 5(2):424-434.
- Susanti, S. 2016. Analisis Perbandingan Algoritma *LCP (Left-Corner-Parsing)* Dan Algoritma *CYK (Cocke-Younger-Kasami)* Untuk Memeriksa Pola Kalimat Baku Bahasa Indonesia. *Jurnal Ilmiah Komputer dan Informatika*. 1-8.
- Wicaksono, A.F. & Purwarianti, A. 2010. HMM Based Part-of-Speech Tagger for Bahasa Indonesia. *On Proceedings of 4th International MALINDO (Malay and Indonesian Language) Workshop*, Jakarta.