

FINAL PROJECT: APPLICATION OF CFG IN SYNTACTIC PARSING

Dosen : Dr. Anak Agung Istri Ngurah Eka Karyawati, S.Si., M.Eng.



Disusun Oleh :

Wayan Farel Nickholas Sadewa (2208561051)

David Brave Moarota Zebua (2208561063)

I Made Treshnanda Mas (2208561089)

Desak Putu Tia Rusilia Wati (2208561143)

PROGRAM STUDI INFORMATIKA

FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM

UNIVERSITAS UDAYANA

2023

KATA PENGANTAR

Kami panjatkan puji syukur atas kehadiran Tuhan yang Maha Esa yang telah melimpahkan rahmat, dan hidayahnya kepada kami Penulis, sehingga kami dapat menyelesaikan laporan mengenai "Final Project: Application of CFG in Syntactic Parsing" dari mata kuliah Teori Bahasa dan Otomata ini dengan sedemikian baik dan tepat Waktu. Laporan ini kami susun dengan semampunya dan tentunya mendapatkan segala bantuan dari berbagai sumber sehingga Laporan ini dapat dibuat dengan baik dan lancar.

Dengan demikian kami menyampaikan banyak terimakasih kepada semua pihak yang sudah berkontribusi pada pembuatan Laporan ini. Terlepas dari semua hal itu, kami menyadari bahwa masih banyak kekurangan dari segi susunan kalimat maupun Tata Bahasa. Oleh karena itu dengan senang hati kami menerima segala saran dan kritik dari pembaca agar Laporan ini dapat menjadi lebih baik. Akhir kata kami harap Laporan ini bermanfaat dan bisa menjadi bahan literasi yang baik.

Jimbaran, 25 Desember 2023

Penulis

DAFTAR ISI

DAFTAR ISI.....	3
BAB I.....	4
PENDAHULUAN.....	4
1.1 Latar Belakang.....	4
1.2 Rumusan Masalah.....	5
1.3 Tujuan.....	5
1.4 Batasan.....	5
BAB II.....	7
METODE CFG.....	7
2.1 Desain CFG.....	7
2.1.1 Set Of Non Terminals.....	7
2.1.2 Set Of Terminals.....	8
2.1.3 Set Of Production Rules.....	10
2.1.4 Start Symbol.....	13
2.2 CFG Rule Conversion to CNF.....	14
2.3 CYK Algorithm.....	16
BAB III.....	18
ANALISIS DAN DESAIN.....	18
3.1 Desain Eksperimen:.....	18
3.2 Desain Aplikasi.....	18
3.3 Source Code.....	23
BAB IV.....	24
HASIL DAN PEMBAHASAN.....	24
4.1 Hasil dan Pembahasan.....	24
BAB V.....	30
PENUTUP.....	30
DAFTAR PUSTAKA.....	31

BAB I

PENDAHULUAN

1.1 Latar Belakang

Sebelumnya sering kali kita menghadapi permasalahan yang dimana menentukan kalimat bahasa baku indonesia sesuai dengan tata bahasa indonesia yang ada merupakan hal yang sulit dihindari. Dalam mengatasi hal tersebut, Context-Free grammars digunakan untuk membantu menentukan bahasa alami dan oleh karena itu sangat cocok untuk mendefinisikan bahasa komputer (Carroll, J., & Long, D. (1989). Latar belakang penelitian ini dikaitkan dengan kebutuhan untuk memahami dan menganalisis struktur kalimat dalam bahasa secara otomatis.

Algoritma CYK menjadi pilihan utama dalam penelitian ini karena keunggulannya dalam mengurai struktur kalimat yang terdefinisi oleh CFG, sebuah konsep linguistik yang mendasar dalam memodelkan tata bahasa dan struktur sintaksis. Kombinasi antara teori bahasa formal dan kecerdasan komputasional menjadi pemicu utama penggunaan algoritma CYK dalam konteks ini.

Penelitian tentang parsing CFG dengan algoritma CYK menjadi relevan karena aplikasinya yang luas dalam berbagai bidang, termasuk pemrosesan bahasa alami, kompilasi kode, dan analisis teks. Algoritma ini memiliki kemampuan untuk mengurai struktur kalimat secara efisien, mengidentifikasi hubungan antara kata-kata dalam kalimat, serta memahami struktur sintaksisnya. Kegunaan praktisnya terlihat dalam pengembangan teknologi terkait pemrosesan bahasa alami, terjemahan otomatis, serta dalam pengolahan dan analisis teks yang kompleks.

Oleh Karena itu, untuk mempermudah menentukan sebuah kalimat sudah baku dan sesuai dengan tata bahasa indonesia atau tidak diperlukan aplikasi berbasis web dengan metode table filling yang menggunakan Context-Free Grammar dengan algoritma CYK (Cocke-Younger-Kasami) yang akan dikembangkan pada penelitian ini.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang ada pada penelitian ini, dapat disimpulkan rumusan masalah pada penelitian ini yaitu:

- 1). Bagaimana cara mengimplementasikan metode CFG dalam parsing kalimat bahasa baku indonesia?
- 2). Bagaimana cara mengubah CFG ke CNF (Chomsky Normal Form) hingga dapat meningkatkan efisiensi algoritma CYK?
- 3). Bagaimana cara mengembangkan aplikasi berbasis web streamlit dalam parsing kalimat bahasa baku indonesia?

1.3 Tujuan

Adapun tujuan utama dari penelitian ini yaitu:

- 1). Mengetahui bagaimana cara mengimplementasikan metode CFG pada penggunaannya dalam parsing kalimat bahasa baku indonesia
- 2). Mengetahui bagaimana cara mengubah CFG menjadi CNF dalam parsing kalimat bahasa baku indonesia
- 3). Mengetahui bagaimana cara mengembangkan aplikasi berbasis web streamlit sehingga dapat digunakan dalam parsing kalimat bahasa baku indonesia

1.4 Batasan

- 1). Jika kalimat yang diinput oleh user mengandung kata yang tidak berada dalam set of terminals pada CFG maka sistem akan mengatakan bahwa kalimat tersebut tidak valid meskipun aslinya kalimat tersebut adalah valid.

BAB II

METODE CFG

2.1 Desain CFG

Tata Bahasa Bebas Konteks (CFG) didefinisikan sebagai 4 tuple $G = (V, \Sigma, R, S)$. V adalah himpunan elemen berhingga yang dikenal sebagai variabel atau sering disebut dengan non-terminal. Σ adalah himpunan elemen berhingga yang dikenal sebagai terminal yang merupakan karakter pembentuk isi sebenarnya dari kalimat, dapat mencakup kata-kata atau huruf-huruf. R adalah sekumpulan elemen yang dikenal sebagai aturan dalam penyusunan sebuah kalimat. Dan S adalah non terminal khusus yang mewakili string awal yang akan dihasilkan oleh tata bahasa.

2.1.1 Set Of Non Terminals

Set Non Terminal (V) dalam tata bahasa bebas konteks (CFG) merupakan sekumpulan variabel yang memiliki peran utama dalam membentuk sebuah kalimat. Setiap variabel berfungsi sebagai "rumah" di mana berbagai bagian kalimat seperti subjek, predikat, objek, pelengkap, keterangan, dan lain-lain dapat ditempatkan. Variabel ini menciptakan kerangka struktural yang diperlukan penulis untuk menentukan sebuah kalimat sudah baku dan sesuai dengan tata bahasa indonesia, mempertimbangkan bagaimana setiap elemen saling berhubungan dan berperan. Berikut adalah penjelasan rinci tentang set Non Terminal, yang terdiri dari berbagai variabel yang menentukan cara kalimat dibangun :

- a) **K**: Kalimat
- b) **S**: Subjek
- c) **P**: Predikat
- d) **O**: Objek
- e) **Pel**: Pelengkap
- f) **Ket**: Keterangan
- g) **NP**: Noun Phrase

- h) **VP**: Verb Phrase
- i) **AdjP**: Adjectival Phrase
- j) **PP**: Prepositional Phrase
- k) **NumP**: Numeral Phrase
- l) **Verb**: kata kerja (verba)
- m) **Noun**: kata benda (nomina)
- n) **Adj**: kata sifat (Adjectival)
- o) **Adv**: kata keterangan (Adverb: sangat, agak, jarang, sudah, akan, sedang, ...)
- p) **Num**: numeralia
- q) **Prep**: Preposisi (pada, dengan, di, ke, untuk, ...)
- r) **PropNoun**: Proper Noun (nama-nama: Agung, Kali Unda, Denpasar, New York,...)
- s) **Pronoun**: kata ganti (ini, itu, saya, dia,..)

2.1.2 Set Of Terminals

Set Terminal (Σ) dalam tata bahasa bebas konteks (CFG) merupakan kumpulan simbol-simbol terminal yang mencakup kata-kata asli dan elemen-elemen spesifik dalam kalimat. Simbol-simbol ini digunakan untuk mengisi peran yang telah ditentukan oleh variabel non-terminal. Simbol-simbol terminal ini merangkum berbagai kata dan frasa yang digunakan untuk membangun kalimat. Setiap simbol terminal mewakili kata atau frasa dalam bahasa Indonesia. Berikut adalah penjelasan rinci yang penulis jelaskan tentang set Terminal:

$\Sigma = \{ \text{sekolah , keluarga , telur , kucing , anjing , penjelasan , hewan , siang , kantor , pak , kuburan , rumah , semangat , pengunjung , kak , ayah , ibu , istri , suami , om , komputer , tas , roti , sungai , bulan , gelas , mobil , motor , kue , boneka , peralatan , anak , tempat , waktu , konsultasi , sepanjang , hormat , antara , sekali , SD , maraton , pasar , belanja , pagi , tadi , belakang , suara , pintu , sore , bawah , tidur , matematika , pelaut , kapal , toko , taman , safari , pensil , tengah , meja , laut , murid , muridnya , yayasan , uang , penjual , permen , SD , sekolah , dasar , jalan , lapangan , istana , presiden , keadaan ,$

teman , temannya , kelas , kode , kode-kode , harga , juara , proyektor , menara , ketampanannya , pahala , pahalanya , kejahatan , dunia , kemunafikan , kemunafikannya , rumahnya , foto , ponsel , ponselnya , lemari , nilai , nilainya , nasi , peraturan , kampus , teras , keseharian , kesehariannya , gunung , sayur , sayur-sayuran , paman , surat , pos , makan , makanan , jumlah , sejumlah , buah , tangan , kota , hingga , perjalanan , kebun , binatang , kemarin , malam , sepak , bola , letak , diletakkan , piring , kertas , pakaian , bencana , korban , mata , kuliah , dokter , resep , obat , pasien , ujian , akhir , dosen , tugas , kerja , kerjakan , mahasiswa , buku , dipelajari , wisata , kebersamaan , kunci , kebahagiaan , kerjasama , pemain , kemenangan , pelatih , petugas , keamanan , patroli , guna , produsen , produk , permintaan , konsumen , penulis , motivasi , pembaca , soal , ujian , diri , sendok , minggu , januari , selama , perusahaan , anak-anak , jam , pukul , hari , pantai , sesaat , hujan , laporan , pertemuan , informasi , sejarah , kemerdekaan , perihai , rencana , perkuliahan , depan , kebijakan , data , hasil , survei , perkembangan , teknologi , paham , keputusan , panitia , baju , katun , kayu , jati , lantai , mengendarai , sepeda , pulau , kanak-kanak , kejauhan , hidangan , atas , kesalahannya , kebaikan , kematian , anjingku , tanaman , kesayangannya , kesayangan , hitungan , detik , pintu , kamar , kamarnya , taruhan , rapornya , sedih , kekuatan , gawang , semester , syal , tahun , sebuah , jaket , antar-pemain , percaya , sepatu , permintaan , konsumen , dua , satu , delapan , belas , setiap , lima , semua , 2023 , banyak , Bersama , saya , kamu , kalian , dia , mereka , aku , ini , itu , ia , kami , tersebut , Kresna , Arista , Puja , Farel , Ilham , Jawa , China , Papua , Tabanan , Yulius , Anton , Rusdi , Ani , Lorraine , Jordan , Gerry , Rafael , David , Farel , Rafael , Anita , Yunus , Chan , Doni , Jokowi , Budi , Marita , Doni , Gery , Nanda , Risol , Renon , Eka , Marita , ninja , Han , Radja , cecep , Agung , Adi , Yogyakarta , Samarinda , Bandung , Tata , Bahasa , Otomata , Agus , Aldi , Rina , Andre , Rio , Tia , Indonesia , Rani , Desi , Putu , Shyta , Krisnawan , Wahyu , Calvin , Marcelino , Kavin , Okta , Sanglah , Lia , Badung , Surabaya , akan , sangat , sedang , baru , sudah , saja , tidak , belum , sekali , tentu , harus , sering , agak , lebih , telah , bersama , ingin , hati-hati , tersedu-sedu , jarang , penting , keras , mandiri , kecil , baru , hormat , sendiri , tampan , takut , gembira , kesal , kecewa , sabar , lelah , indah , ganteng , lambat , cerdas , luas , sempit , bersih , kotor , singkat , tentu , maraton , aneh , besar , nyenyak , dalam , lokal , giat , lebih , layak , mudah , banyak , sama , penuh , kuat ,

hangat , cekatan , dekat , kejauhan , lahap , serius , teliti , santai , malas, suka , pergi , datang , merokok , belajar , bermain , mengamati , membeli , memberikan , menyumbang , memberi , bernyanyi , terlihat , merasa , berniat , berpuasa , terkejut , percaya , terletak , mendengar , menyimpan , memasak , merasa , menunggu , terletak , termenung , tersembunyi , menuju , sampai , berperang , mereka , tidur , verb , lari , bekerja , jalan , jalan-jalan , pulang , makan , lantik , mandi , bergegas , memukul , terkesima , menang , berlatih , lebih , goreng , daki , mendaki , berjualan , membawa , sebut , tersebut , bersepeda , berjalan , berada , tunjuk , menunjuk , tempel , dinding , sumbang , menyumbangkan , pakai , adalah , menyewa , wisata , antar , menabung , melakukan , menjaga , membuat , memenuhi , menulis , memotivasi , mengerjakan , tekad , berolahraga , meninggalkan , dikumpulkan , berakhir , mengadakan , mengenai , mengetahui , menyusun , berkaitan , dibuat , dibangun , tersenyum , mengambil , berlayar , berlari , terburu-buru , melihat , memakan , menegur , menangis , memindahkan , menendang , kalah , menganyam , tergantung , berbelanja , paham , dikerjakan , ditempel , di , ke , dari , pada , untuk , dengan , oleh , sampai , sejak , antara , sebelum , akan , kepada , agar , hingga , bagi , buat , antar , guna , setelah , dimulai , mengenai , tentang , atas , karena}.

2.1.3 Set Of Production Rules

Set Of Production Rules (R) dalam tata bahasa bebas konteks (CFG) menggambarkan langkah-langkah untuk membentuk kalimat berdasarkan variabel-variabel non-terminal yang telah didefinisikan sebelumnya. Berikut adalah kumpulan aturan produksi yang penulis terapkan:

- a. Kalimat dapat terbentuk dengan kombinasi subjek (S), predikat (P), objek (O), pelengkap (Pel), dan keterangan (Ket).

$$K \rightarrow S P \mid S P O \mid S P Pel \mid S P Ket \mid S P O Pel \mid S P O Ket \mid S P Pel Ket \mid S P O Pel Ket$$

- b. Subjek (S) dapat merupakan frasa kata benda (NP).

$$S \rightarrow NP$$

- c. Predikat (P) dapat terdiri dari frasa preposisional (PP), frasa kata kerja (VP), frasa kata sifat (AdjP), dan juga frasa kata benda (NP).

$P \rightarrow PP \mid VP \mid AdjP \mid NP$

- d. Objek (O) dapat merupakan frasa kata benda (NP).

$O \rightarrow NP$

- e. Pelengkap (Pel) dapat terdiri dari frasa kata benda (NP), frasa kata kerja (VP), atau frasa kata sifat (AdjP), dan juga frasa preposisional (PP)

$Pel \rightarrow NP \mid VP \mid AdjP \mid PP$

- f. Keterangan (Ket) dapat terdiri dari frasa preposisional (PP), frasa kata benda (NP), frasa numeralia (NumP), atau frasa kata sifat (AdjP).

$Ket \rightarrow PP \mid NP \mid NumP \mid AdjP$

- g. Frasa kata benda (NP) dapat terbentuk dari nama khas (PropNoun), kata ganti (Pronoun), kata benda (Noun), kombinasi NP dan frasa kata sifat (AdjP), kombinasi NumP dan NP, kombinasi NP dan kata benda (Noun), kombinasi NP dan nama khas (PropNoun), atau kombinasi NP dan kata ganti (Pronoun), beserta kombinasi NP kata numerik

$NP \rightarrow PropNoun \mid Pronoun \mid Noun \mid NP \ AdjP \mid NumP \ NP \mid NP \ Noun \mid NP \ PropNoun \mid NP \ Pronoun \mid NP \ NumP$

- h. Frasa kata kerja (VP) dapat terdiri dari kata kerja (Verb) atau kombinasi frasa kata keterangan (AdvP) dan frasa kata kerja (VP).

$VP \rightarrow Verb \mid AdvP \ VP$

- i. Frasa preposisional (PP) dapat terdiri dari preposisi (Prep) dan frasa kata benda (NP), frasa kata keterangan (AdvP), frasa kata sifat (AdjP), atau frasa kata kerja (VP), kombinasi kata keterangan (VP) dan juga Preposisi (Prep), beserta Preposisi (Prep) frasa kata kerja (VP) dan juga frasa kata benda (NP)

$PP \rightarrow Prep \ NP \mid Prep \ AdvP \mid Prep \ AdjP \mid Prep \ VP \mid Adv \ Prep \mid Prep \ VP \ NP$

- j. Frasa kata sifat (AdjP) dapat terdiri dari kata sifat (Adj) atau kombinasi kata keterangan (Adv) dan frasa kata sifat (AdjP).

$AdjP \rightarrow Adj \mid Adv \ AdjP$

- k. Frasa kata keterangan (AdvP) dapat terdiri dari kata keterangan (Adv) atau kombinasi kata keterangan (Adv) dan frasa kata keterangan (AdvP).

AdvP → Adv | Adv AdvP

- l. Frasa numeralia (NumP) dapat terdiri dari numeralia (Num) atau kombinasi numeralia (Num) dan frasa numeralia (NumP).

NumP → Num | Num NumP

- m. **Noun** → sekolah | keluarga | telur | kucing | anjing | penjelasan | hewan | siang | kantor | pak | kuburan | rumah | semangat | pengunjung | kak | ayah | ibu | istri | suami | om | komputer | tas | roti | sungai | bulan | gelas | mobil | motor | kue | boneka | peralatan | anak | tempat | waktu | konsultasi | sepanjang | hormat | antara | sekali | SD | maraton | pasar | belanja | pagi | tadi | belakang | suara | pintu | sore | bawah | tidur | matematika | pelaut | kapal | toko | taman | safari | pensil | tengah | meja | laut | murid | muridnya | yayasan | uang | penjual | permen | SD | sekolah | dasar | jalan | lapangan | istana | presiden | keadaan | teman | temannya | kelas | kode | kode-kode | harga | juara | proyektor | menara | ketampanannya | pahala | pahalanya | kejahatan | dunia | kemunafikan | kemunafikannya | rumahnya | foto | ponsel | ponselnya | lemari | nilai | nilainya | nasi | peraturan | kampus | teras | keseharian | kesehariannya | gunung | sayur | sayur-sayuran | paman | surat | pos | makan | makanan | jumlah | sejumlah | buah | tangan | kota | hingga | perjalanan | kebun | binatang | kemarin | malam | sepak | bola | letak | diletakkan | piring | kertas | pakaian | bencana | korban | mata | kuliah | dokter | resep | obat | pasien | ujian | akhir | dosen | tugas | kerja | kerjakan | mahasiswa | buku | dipelajari | wisata | kebersamaan | kunci | kebahagiaan | kerjasama | pemain | kemenangan | pelatih | petugas | keamanan | patroli | guna | produsen | produk | permintaan | konsumen | penulis | motivasi | pembaca | soal | ujian | diri | sendok | minggu | januari | selama | perusahaan | anak-anak | jam | pukul | hari | pantai | sesaat | hujan | laporan | pertemuan | informasi | sejarah | kemerdekaan | perihal | rencana | perkuliahan | depan | kebijakan | data | hasil | survei | perkembangan | teknologi | paham | keputusan | panitia | baju | katun | kayu | jati | lantai | mengendarai | sepeda | pulau | kanak-kanak | kejauhan | hidangan | atas | kesalahannya | kebaikan | kematian | anjingku | tanaman | kesayangannya | kesayangan | hitungan | detik | pintu | kamar | kamarnya | taruhan | rapornya | sedih | kekuatan | gawang | semester | syal | tahun | sebuah | jaket | antar-pemain | percaya | sepatu | permintaan | konsumen
- n. **Pronoun** → saya | kamu | kalian | dia | mereka | aku | ini | itu | ia | kami | tersebut
- o. **PropNoun** → Kresna | Arista | Puja | Farel | Ilham | Jawa | China | Papua | Tabanan | Yulius | Anton | Rusdi | Ani | Lorraine | Jordan | Gerry | Rafael | David | Farel | Rafael | Anita | Yunus | Chan | Doni | Jokowi | Budi | Marita | Doni | Gery |

Nanda | Risol | Renon | Eka | Marita | ninja | Han | Radja | cecep | Agung | Adi | Yogyakarta | Samarinda | Bandung | Tata | Bahasa | Otomata | Agus | Aldi | Rina | Andre | Rio | Tia | Indonesia | Rani | Desi | Putu | Shyta | Krisnawan | Wahyu | Calvin | Marcelino | Kavin | Okta | Sanglah | Lia | Badung | Surabaya

- p. **Adv** → akan | sangat | sedang | baru | sudah | saja | tidak | belum | sekali | tentu | harus | sering | agak | lebih | telah | bersama | ingin | hati-hati | tersedu-sedu | jarang
- q. **Adj** → penting | keras | mandiri | kecil | baru | hormat | sendiri | tampan | takut | gembira | kesal | kecewa | sabar | lelah | indah | ganteng | lambat | cerdas | luas | sempit | bersih | kotor | singkat | tentu | maraton | aneh | besar | nyenyak | dalam | lokal | giat | lebih | layak | mudah | banyak | sama | penuh | kuat | hangat | cekatan | dekat | kejauhan | lahap | serius | teliti | santai | malas
- r. **Verb** -> suka | pergi | datang | merokok | belajar | bermain | mengamati | membeli | memberikan | menyumbang | memberi | bernyanyi | terlihat | merasa | berniat | berpuasa | terkejut | percaya | terletak | mendengar | menyimpan | memasak | merasa | menunggu | terletak | termenung | tersembunyi | menuju | sampai | berperang | mereka | tidur | verb | lari | bekerja | jalan | jalan-jalan | pulang | makan | lantik | mandi | bergegas | memukul | terkesima | menang | berlatih | lebih | goreng | daki | mendaki | berjualan | membawa | sebut | tersebut | bersepeda | berjalan | berada | tunjuk | menunjuk | tempel | dinding | sumbang | menyumbangkan | pakai | adalah | menyewa | wisata | antar | menabung | melakukan | menjaga | membuat | memenuhi | menulis | memotivasi | mengerjakan | tekad | berolahraga | meninggalkan | dikumpulkan | berakhir | mengadakan | mengenai | mengetahui | menyusun | berkaitan | dibuat | dibangun | tersenyum | mengambil | berlayar | berlari | terburu-buru | melihat | memakan | menegur | menangis | memindahkan | menendang | kalah | menganyam | tergantung | berbelanja | paham | dikerjakan | ditempel
- s. **Prep** -> di | ke | dari | pada | untuk | dengan | oleh | sampai | sejak | antara | sebelum | akan | kepada | agar | hingga | bagi | buat | antar | guna | setelah | dimulai | mengenai | tentang | atas | karena

2.1.4 Start Symbol

Start symbol digunakan sebagai titik awal dalam proses derivasi atau pembentukan suatu kalimat berdasarkan aturan produksi yang telah ditentukan sebelumnya. Dalam konteks tata bahasa bebas konteks (CFG) ini, start symbol yang penulis gunakan adalah $S=\{K\}$

2.2 CFG Rule Conversion to CNF

Konversi aturan produksi dalam Tata Bahasa Bebas Konteks (CFG) ke dalam bentuk Normal Form Chomsky (CNF) adalah proses penting yang melibatkan restrukturisasi aturan produksi sehingga memenuhi persyaratan CNF. Dalam CNF, setiap aturan produksi memiliki dua non-terminal atau satu terminal di sisi kanan. Konsep ini mempermudah proses analisis sintaktik dan membawa kejelasan dalam pemahaman struktur tata bahasa.

Proses konversi ke CNF melibatkan beberapa langkah penting. Pertama, setiap produksi yang memiliki lebih dari dua simbol non-terminal di sisi kanan dipecah menjadi beberapa produksi yang lebih kecil, sehingga setiap produksi hanya memiliki dua simbol non-terminal atau satu terminal di sisi kanan. Langkah ini memastikan bahwa setiap langkah derivasi hanya menggabungkan dua non-terminal atau memperkenalkan satu terminal. Berikut adalah hasil konversi aturan produksi CFG ke CNF:

- a. $A \rightarrow S P$
- b. $B \rightarrow \text{Prep VP}$
- c. $C \rightarrow A \text{ Pel}$
- d. $D \rightarrow A O$
- e. $E \rightarrow D \text{ Pel}$
- f. $K \rightarrow A \text{ Ket} \mid C \text{ Ket} \mid E \text{ Ket} \mid A O \mid D \text{ Ket} \mid D \text{ Pel} \mid A \text{ Pel} \mid S P$
- g. $S \rightarrow NP$
- h. $P \rightarrow VP \mid PP \mid NP \mid \text{AdjP}$
- i. $O \rightarrow NP$
- j. $\text{Pel} \rightarrow VP \mid PP \mid NP \mid \text{AdjP}$
- k. $\text{Ket} \rightarrow \text{AdjP} \mid PP \mid NP \mid \text{NumP}$
- l. $NP \rightarrow \text{PropNoun} \mid NP \text{ NumP} \mid \text{Noun} \mid NP \text{ Pronoun} \mid NP \text{ Noun} \mid \text{NumP NP} \mid \text{Pronoun} \mid NP \text{ PropNoun} \mid NP \text{ AdjP}$
- m. $VP \rightarrow \text{Verb} \mid \text{AdvP VP}$
- n. $PP \rightarrow \text{Prep NP} \mid \text{Prep AdvP} \mid B NP \mid \text{Prep AdjP} \mid \text{Prep VP} \mid \text{Adv Prep}$
- o. $\text{AdjP} \rightarrow \text{Adj} \mid \text{Adv AdjP}$

p. **AdvP** → **Adv** | **Adv AdvP**

q. **NumP** → **Num** | **Num NumP**

r. **Noun** → sekolah | keluarga | telur | kucing | anjing | penjelasan | hewan | siang | kantor | pak | kuburan | rumah | semangat | pengunjung | kak | ayah | ibu | istri | suami | om | komputer | tas | roti | sungai | bulan | gelas | mobil | motor | kue | boneka | peralatan | anak | tempat | waktu | konsultasi | sepanjang | hormat | antara | sekali | SD | maraton | pasar | belanja | pagi | tadi | belakang | suara | pintu | sore | bawah | tidur | matematika | pelaut | kapal | toko | taman | safari | pensil | tengah | meja | laut | murid | muridnya | yayasan | uang | penjual | permen | SD | sekolah | dasar | jalan | lapangan | istana | presiden | keadaan | teman | temannya | kelas | kode | kode-kode | harga | juara | proyektor | menara | ketampanannya | pahala | pahalanya | kejahatan | dunia | kemunafikan | kemunafikannya | rumahnya | foto | ponsel | ponselnya | lemari | nilai | nilainya | nasi | peraturan | kampus | teras | keseharian | kesehariannya | gunung | sayur | sayur-sayuran | paman | surat | pos | makan | makanan | jumlah | sejumlah | buah | tangan | kota | hingga | perjalanan | kebun | binatang | kemarin | malam | sepak | bola | letak | diletakkan | piring | kertas | pakaian | bencana | korban | mata | kuliah | dokter | resep | obat | pasien | ujian | akhir | dosen | tugas | kerja | kerjakan | mahasiswa | buku | dipelajari | wisata | kebersamaan | kunci | kebahagiaan | kerjasama | pemain | kemenangan | pelatih | petugas | keamanan | patroli | guna | produsen | produk | permintaan | konsumen | penulis | motivasi | pembaca | soal | ujian | diri | sendok | minggu | januari | selama | perusahaan | anak-anak | jam | pukul | hari | pantai | sesaat | hujan | laporan | pertemuan | informasi | sejarah | kemerdekaan | perihai | rencana | perkuliahan | depan | kebijakan | data | hasil | survei | perkembangan | teknologi | paham | keputusan | panitia | baju | katun | kayu | jati | lantai | mengendarai | sepeda | pulau | kanak-kanak | kejauhan | hidangan | atas | kesalahannya | kebaikan | kematian | anjingku | tanaman | kesayangannya | kesayangan | hitungan | detik | pintu | kamar | kamarnya | taruhan | rapornya | sedih | kekuatan | gawang | semester | syal | tahun | sebuah | jaket | antar-pemain | percaya | sepatu | permintaan | konsumen

s. **Pronoun** → saya | kamu | kalian | dia | mereka | aku | ini | itu | ia | kami | tersebut

t. **PropNoun** → Kresna | Arista | Puja | Farel | Ilham | Jawa | China | Papua | Tabanan | Yulius | Anton | Rusdi | Ani | Lorraine | Jordan | Gerry | Rafael | David | Farel | Rafael | Anita | Yunus | Chan | Doni | Jokowi | Budi | Marita | Doni | Gery | Nanda | Risol | Renon | Eka | Marita | ninja | Han | Radja | cecep | Agung | Adi | Yogyakarta | Samarinda | Bandung | Tata | Bahasa | Otomata | Agus | Aldi | Rina | Andre | Rio | Tia | Indonesia | Rani | Desi | Putu | Shyta | Krisnawan | Wahyu | Calvin | Marcelino | Kavin | Okta | Sanglah | Lia | Badung | Surabaya

u. **Adv** → akan | sangat | sedang | baru | sudah | saja | tidak | belum | sekali | tentu | harus | sering | agak | lebih | telah | bersama | ingin | hati-hati | tersedu-sedu | jarang

- v. **Adj** → penting | keras | mandiri | kecil | baru | hormat | sendiri | tampan | takut | gembira | kesal | kecewa | sabar | lelah | indah | ganteng | lambat | cerdas | luas | sempit | bersih | kotor | singkat | tentu | maraton | aneh | besar | nyenyak | dalam | lokal | giat | lebih | layak | mudah | banyak | sama | penuh | kuat | hangat | cekatan | dekat | kejauhan | lahap | serius | teliti | santai | malas
- w. **Verb** -> suka | pergi | datang | merokok | belajar | bermain | mengamati | membeli | memberikan | menyumbang | memberi | bernyanyi | terlihat | merasa | berniat | berpuasa | terkejut | percaya | terletak | mendengar | menyimpan | memasak | merasa | menunggu | terletak | termenung | tersembunyi | menuju | sampai | berperang | mereka | tidur | verb | lari | bekerja | jalan | jalan-jalan | pulang | makan | lantik | mandi | bergegas | memukul | terkesima | menang | berlatih | lebih | goreng | daki | mendaki | berjualan | membawa | sebut | tersebut | bersepeda | berjalan | berada | tunjuk | menunjuk | tempel | dinding | sumbang | menyumbangkan | pakai | adalah | menyewa | wisata | antar | menabung | melakukan | menjaga | membuat | memenuhi | menulis | memotivasi | mengerjakan | tekad | berolahraga | meninggalkan | dikumpulkan | berakhir | mengadakan | mengenai | mengetahui | menyusun | berkaitan | dibuat | dibangun | tersenyum | mengambil | berlayar | berlari | terburu-buru | melihat | memakan | menegur | menangis | memindahkan | menendang | kalah | menganyam | tergantung | berbelanja | paham | dikerjakan | ditempel
- x. **Prep** -> di | ke | dari | pada | untuk | dengan | oleh | sampai | sejak | antara | sebelum | akan | kepada | agar | hingga | bagi | buat | antar | guna | setelah | dimulai | mengenai | tentang | atas | karena

2.3 CYK Algorithm

Algoritma Cocke-Younger-Kasami (CYK) adalah algoritma parsing yang efisien digunakan untuk menentukan apakah suatu kalimat dapat dihasilkan oleh tata bahasa bebas konteks (CFG) tertentu. Algoritma ini memanfaatkan teknik pemrograman dinamis dan tabel multidimensi untuk memeriksa kemungkinan struktur tata bahasa dari suatu kalimat.

Langkah pertama dalam penerapan CYK adalah memastikan aturan produksi CFG telah dikonversi ke dalam bentuk Normal Form Chomsky (CNF). Oleh karena itu, aturan produksi yang telah diubah menjadi CNF akan digunakan sebagai dasar dalam proses parsing dengan CYK. Selanjutnya inisialisasi tabel dalam penerapan CYK Algorithm, sebuah tabel berukuran $n \times n$ digunakan, dengan n adalah panjang kalimat yang akan dianalisis. Setiap sel di dalam tabel

tersebut akan berfungsi untuk menampung himpunan variabel non-terminal yang mungkin menghasilkan kata pada posisi tertentu dalam kalimat. Dalam pengisian tabel dilakukan secara iteratif, dimulai dari segmen dua kata hingga seluruh kalimat. Algoritma ini mempertimbangkan aturan produksi dalam bentuk Normal Form Chomsky (CNF) untuk menentukan apakah kombinasi non-terminal pada setiap segmen dapat menghasilkan struktur tata bahasa yang valid. Lalu memeriksa apakah simbol awal (start symbol) K terkandung di sel tabel yang sesuai dengan panjang kalimat. Jika simbol K ditemukan di sel tabel tersebut, hal ini menunjukkan bahwa kalimat dapat diterima berdasarkan tata bahasa Bebas Konteks (CFG) yang diberikan.

BAB III

ANALISIS DAN DESAIN

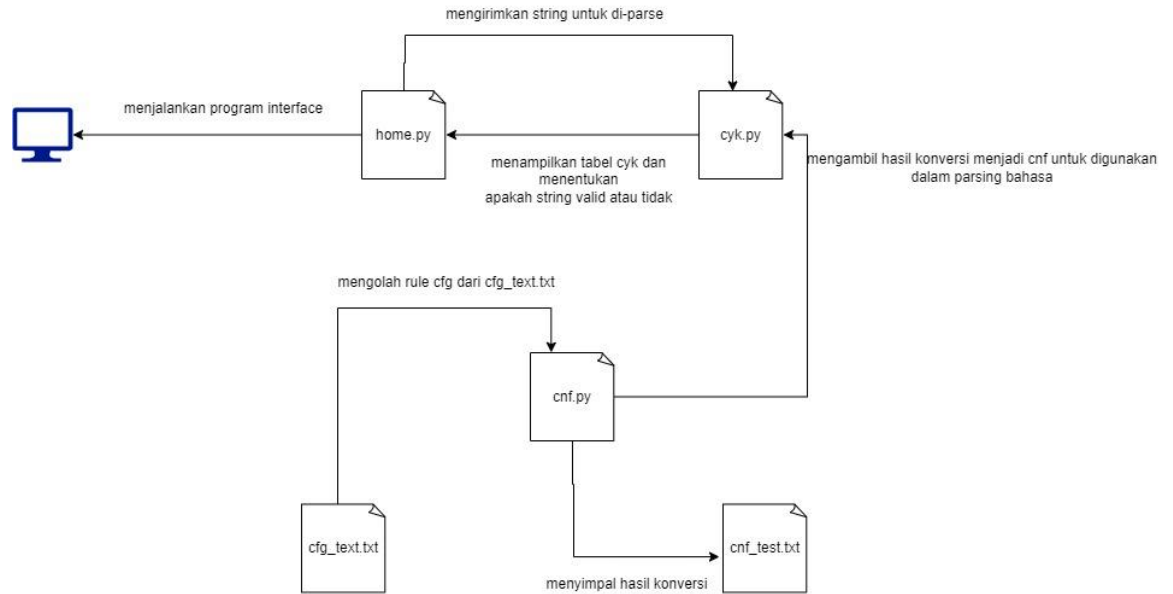
3.1 Desain Eksperimen:

Berdasarkan dari analisa yang dilakukan pada Bab I laporan ini, maka dapat diartikan bahwa dibutuhkan suatu aplikasi yang dapat digunakan untuk melakukan *parsing* kalimat bahasa Indonesia menggunakan algoritma Cocke-Younger-Kasami (CYK) dan menggunakan bentuk Context-Free Grammar (CFG) yang telah diubah menjadi bentuk Chomsky Normal Form (CNF).

Karena aplikasi ini menggunakan algoritma CYK yang berbasis dari konsep CFG, maka artinya terdapat batasan yang akan dimiliki oleh aplikasi ini yaitu apabila pada kalimat yang ingin di-*parse* terdapat suatu kata yang tidak dimasukkan dalam *rule* CFG maka otomatis aplikasi ini akan mengatakan bahwa kalimat tersebut tidak diterima dalam bentuk kalimat baku walaupun semisalnya kalimat tersebut aslinya baku.

3.2 Desain Aplikasi

- a. Arsitektur Aplikasi

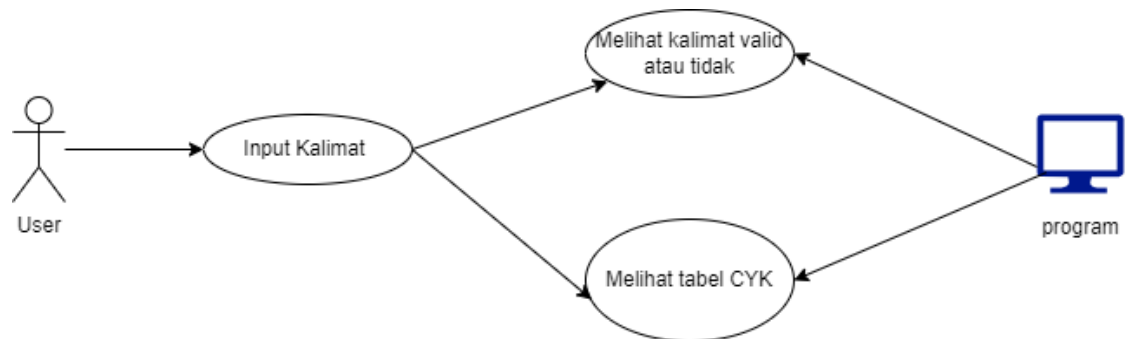


Gambar 3.1 Arsitektur Aplikasi

Berdasarkan gambar arsitektur aplikasi yang telah dilampirkan di atas `home.py` bertugas untuk menjalankan User Interface dari aplikasi ini, Streamlit digunakan dalam aplikasi ini untuk membuatnya dapat diakses oleh user dalam bentuk Web App. `home.py` memberikan input string dari user kepada `cyk.py` untuk diproses menggunakan algoritma CYK, setelah diproses `cyk.py` akan memberikan suatu tabel kepada `home.py` untuk ditentukan apakah kalimat tersebut valid atau tidak dan menampilkan tabel tersebut pada user.

`Cnf.py` bertugas untuk melakukan perubahan bentuk CFG yang telah disimpan dalam file `cfg_text.py` kedalam bentuk CNF yang kemudian akan ditulis pada suatu file yang bernama `cnf_test.py`. Hasil konversi ke bentuk CNF akan digunakan oleh `cyk.py` untuk melakukan algoritma CYK.

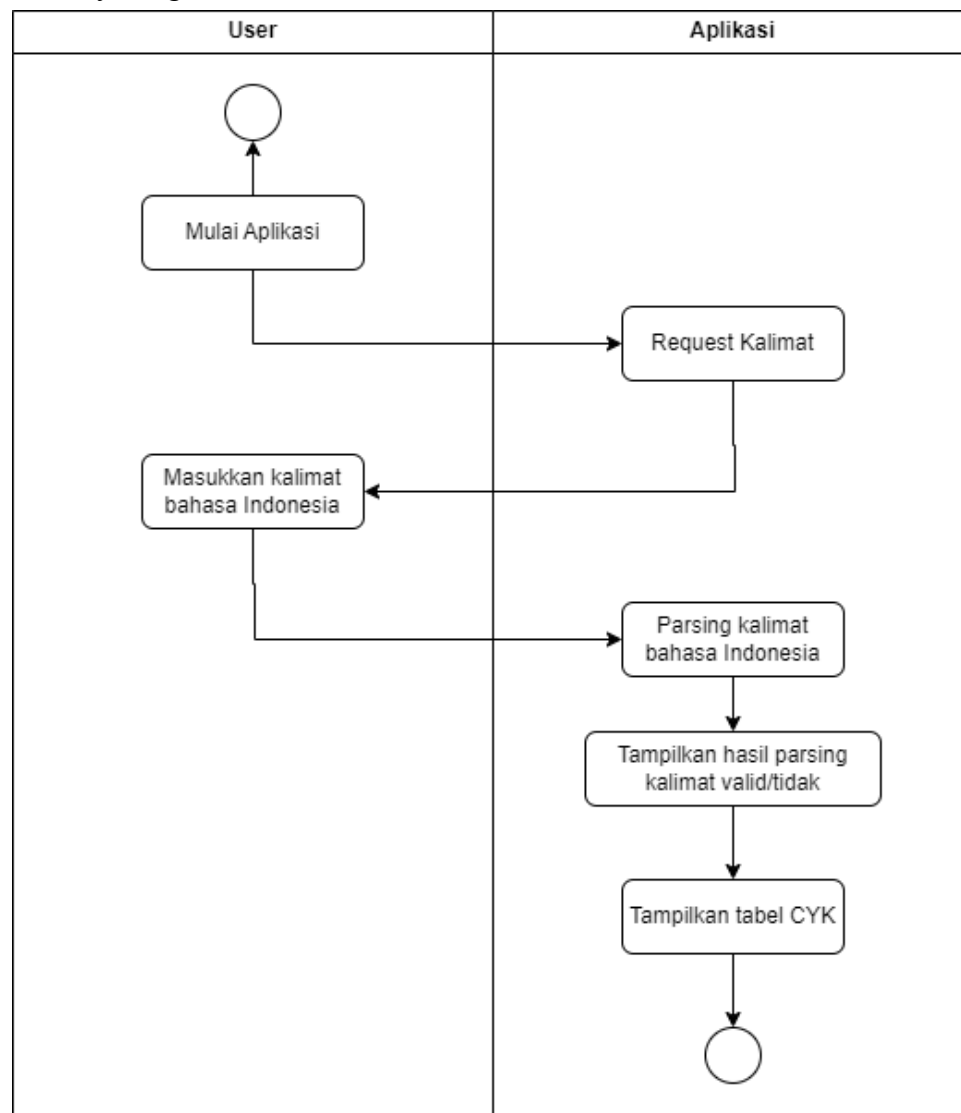
b. Use Case Diagram



Gambar 3.2 Use Case Diagram

Gambar di atas merupakan Use Case Diagram dari aplikasi ini di mana user dapat memberikan suatu input kalimat lalu melihat apakah kalimat tersebut diterima dalam bentuk pola kalimat baku bahasa Indonesia atau tidak dan juga melihat tabel CYK untuk proses *parsing* kalimat tersebut yang telah disediakan oleh program.

c. Activity Diagram



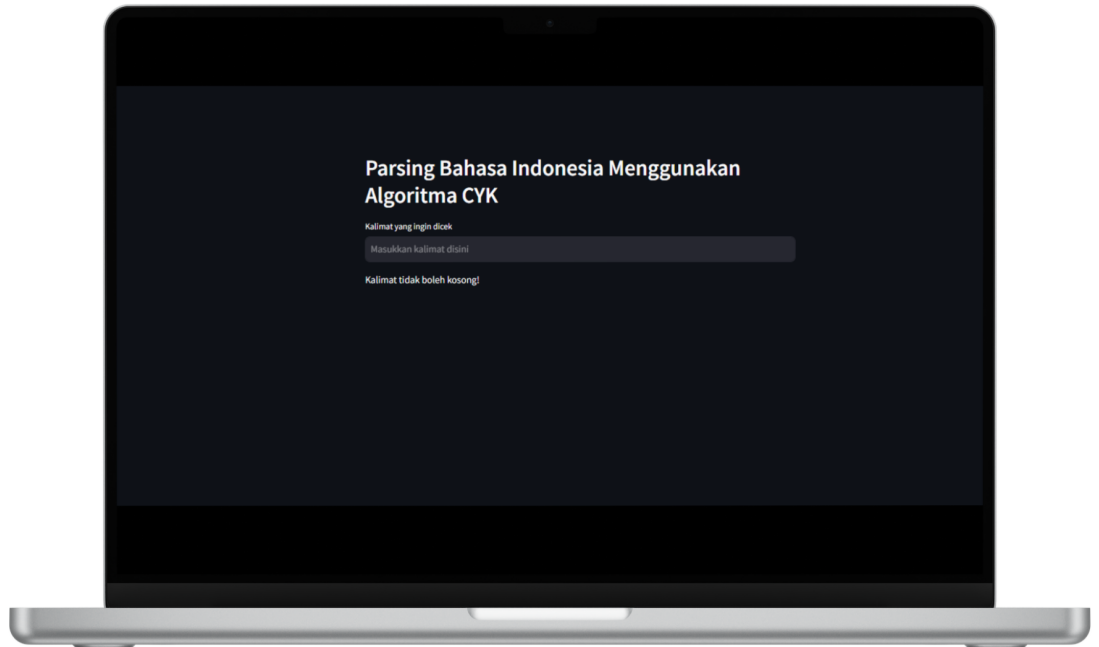
Gambar 3.3 Activity Diagram

Program dimulai dari sisi user di mana user akan memulainya aplikasi parsing kalimat bahasa Indonesia lalu aplikasi akan meminta sebuah input kalimat dari user, user kemudian memberikan kalimat input untuk aplikasi lalu kalimat input tersebut akan dilakukan proses *parsing*. Setelah selesai *parsing* akan ditunjukkan apakah kalimat input yang diberikan oleh user merupakan kalimat

yang diterima bentuknya dalam pola kalimat bahasa Indonesia atau tidak dan juga menunjukkan tabel CYK yang terbentuk dari kalimat tersebut,

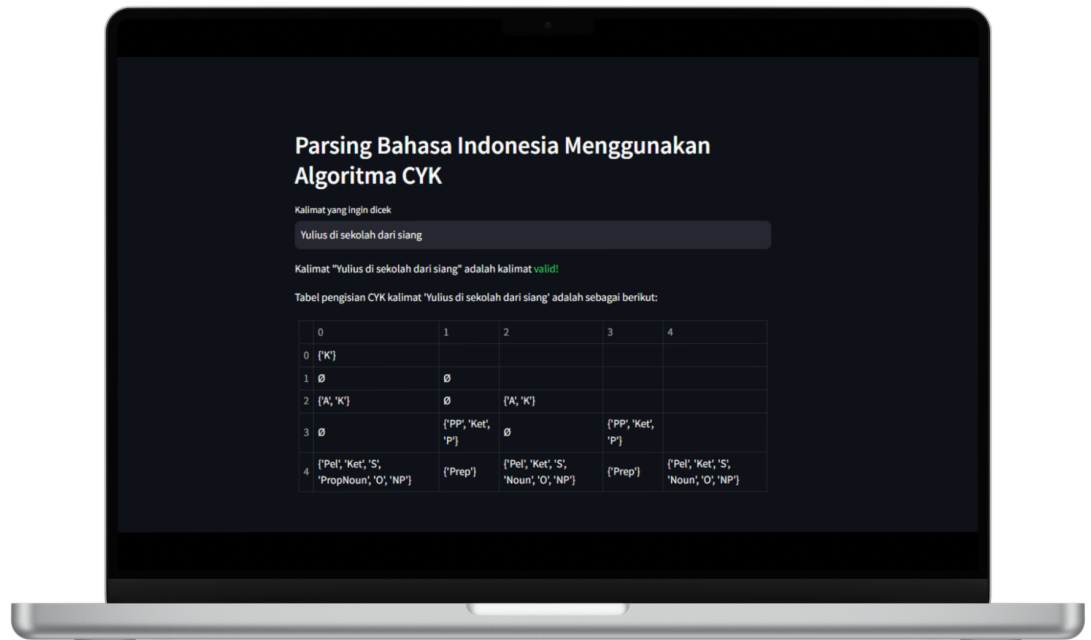
d. Mockup Aplikasi

Berikut adalah mockup dari aplikasi *parsing* kalimat bahasa Indonesia yang didesain menggunakan *library* Streamlit Python



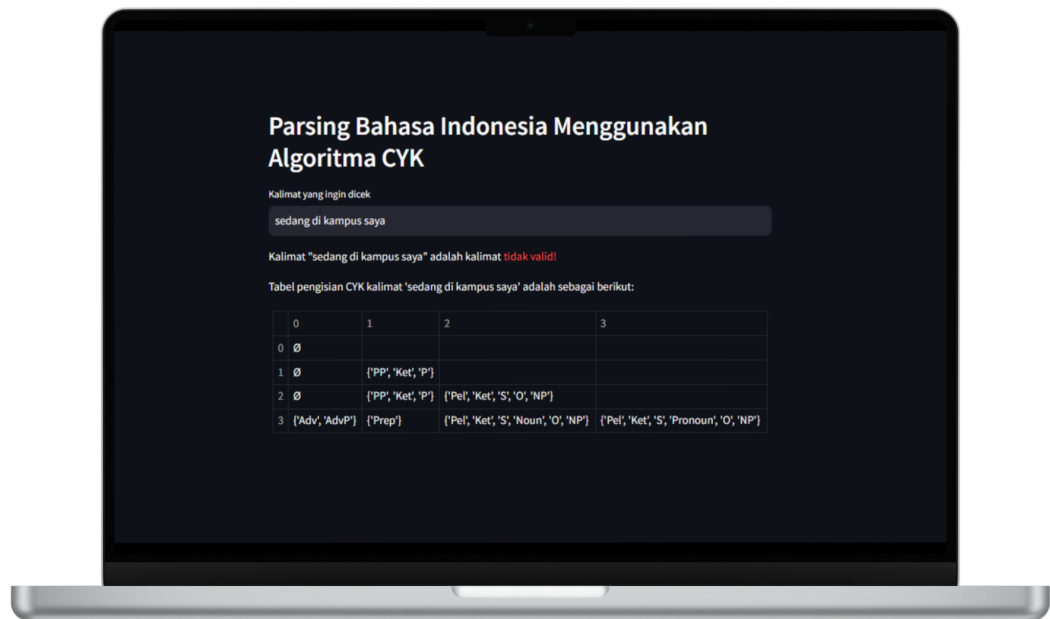
Gambar 3.4 Mockup Main Page

Gambar di atas adalah *mockup* dari laman yang ditampilkan pertama kali saat memulai aplikasi. Dapat terlihat judul dari aplikasi tersebut dan sebuah *text field* untuk meminta input berupa kalimat dari user.



Gambar 3,5 Mockup Kalimat Diterima

Gambar di atas menunjukkan tampilan dari aplikasi bila kalimat bahasa Indonesia yang dimasukkan oleh user diterima dalam pola kalimat bahasa Indonesia. Dapat terlihat aplikasi tersebut menampilkan ke layar tulisan “Kalimat {kalimat user} adalah kalimat valid!” dan di bawahnya terdapat tabel CYK yang digunakan dalam proses *parsing* kalimat tersebut.



Gambar 3.6 Mockup Kalimat Ditolak

Gambar di atas menunjukkan tampilan dari aplikasi bila kalimat bahasa Indonesia yang dimasukkan oleh user ditolak dalam pola kalimat bahasa Indonesia. Dapat terlihat aplikasi tersebut menampilkan ke layar tulisan “Kalimat {kalimat user} adalah kalimat tidak valid!” dan di bawahnya terdapat tabel CYK yang digunakan dalam proses parsing kalimat tersebut. Terlihat perbedaan dari kalimat yang valid dan tidak adalah kalimat valid memiliki “K” dalam tabel paling kiri atas dengan kalimat tidak valid tidak memiliki “K”.

3.3 Source Code

Source Code dari aplikasi ini dapat diakses di link sebagai berikut:
<https://github.com/Tresnanda/FP-TBO-IndonesianParsing>

BAB IV

HASIL DAN PEMBAHASAN

4.1 Hasil dan Pembahasan

Aplikasi *Parsing* Bahasa Indonesia yang telah dibuat selanjutnya akan dilakukan pengujian terhadap performanya dalam menentukan apakah suatu kalimat itu dapat diterima atau tidak sesuai dengan aturan pola kalimat Bahasa Indonesia. Pengujian untuk aplikasi ini akan dilakukan dengan total 100 kalimat yang sudah diketahui pola kalimatnya sesuai dengan aturan pola kalimat Bahasa Indonesia, terdapat sebanyak 81 kalimat yang benar bentuknya dan 19 yang tidak benar bentuknya, adapun kalimat tersebut adalah sebagai berikut:

Diterima
Ilham mendaki di gunung Agung Adi berjualan sayur-sayuran di pasar Cecep membeli sepatu di Surabaya Paman sedang ke Yogyakarta Kamu harus membawa surat ini ke kantor pos Aku membeli makanan ini dari rumah makan tersebut Paman memberikan kami sejumlah buah tangan dari kota Samarinda Saya bersepeda hingga sekolah Dia tidur dalam perjalanan hingga kebun binatang Kami telah sampai Bandung kemarin malam Mereka berjalan sampai lapangan sepak bola Gelas itu diletakkan antara dua piring Rumah Ilham berada antara Rumah temannya Dia menunjuk pada pintu itu Kertas itu ditempel pada dinding Mereka menyumbangkan pakaian layak pakai bagi korban bencana Mata kuliah Tata Bahasa Otomata adalah sangat mudah Dokter memberikan resep obat untuk pasien Agus belajar untuk ujian akhir Dosen memberikan sangat banyak tugas untuk dikerjakan mahasiswa

Agus membeli buku untuk dipelajari bersama
Kami menyewa mobil untuk perjalanan wisata
Kebersamaan adalah kunci kebahagiaan buat kami
Kerjasama antar-pemain adalah kunci kemenangan buat pelatih
Agus menabung buat membeli buku
Petugas keamanan melakukan patroli guna menjaga keamanan
Produsen membuat produk baru guna memenuhi permintaan konsumen
Penulis menulis buku motivasi guna memotivasi pembaca
Aldi mengerjakan soal ujian tersebut dengan penuh percaya diri
Rina mengerjakan tugas kuliah dengan penuh semangat
Adi belajar dengan tekad kuat
Saya makan nasi goreng dengan sendok
Agus akan pergi pada hari Minggu
Paman akan datang di bulan Januari
Ayah akan pergi selama dua minggu
Andre bekerja di perusahaan itu
Anak-anak pulang sekolah setelah jam satu
Rio ke sekolah sebelum pukul delapan
Kelas hari ini dimulai hingga dua belas siang
Tia berolahraga setiap pagi
Mereka meninggalkan pantai sesaat sebelum hujan
Laporan itu dikumpulkan sebelum pukul lima sore
Ujian akan berakhir dalam waktu dua jam
Mereka akan mengadakan pertemuan pada hari Minggu
Buku ini memberikan informasi mengenai sejarah kemerdekaan Indonesia
Dosen memberikan informasi perihal rencana perkuliahan minggu depan
Saya ingin mengetahui tentang kebijakan baru perusahaan
Rani menyusun laporan perihal data hasil survei
Kebijakan baru ini sangat berkaitan dengan perkembangan teknologi baru
Desi tidak paham tentang keputusan panitia
Baju ini dibuat dari katun
Rumah itu dibangun dari kayu jati
Nanda belajar dengan giat
Farel tidur dengan sangat nyenyak di lantai
Puja tersenyum dengan hangat kepada Putu
David mengendarai sepeda motor dengan sangat hati-hati sekali
Shyta sedang mengambil permen dengan cekatan di lantai
Nanda berlayar ke pulau Jawa
Ilham berlari dengan terburu-buru ke taman kanak-kanak dekat rumahnya
Aku melihat Agus dari kejauhan
Krisnawan memakan semua hidangan itu dengan lahap
Nanda menegur Ilham atas kesalahannya untuk kebaikan bersama
Aku menangis dengan tersedu-sedu karena kematian anjingku
Wahyu memindahkan tanaman kesayangannya dalam hitungan detik ke dalam kamarnya
Nanda menendang pintu kamarnya dengan sangat kesal karena kalah taruhan

Calvin membawa rapornya dengan sedih
 Marcelino menendang bola itu dengan kekuatan penuh ke gawang
 Kavin mengerjakan soal ujian akhir semester dengan serius
 Okta menganyam syal indah itu dengan teliti
 Kresna sering pulang dengan sangat gembira ke Tabanan
 Farel jalan-jalan dengan santai di taman kota
 Yunus sering pulang ke Papua
 Anton jarang pergi ke rumah Kresna
 Anton lari maraton di Lapangan Renon
 Pak Jokowi di dalam Istana Presiden
 Nanda ke pulau Jawa pada tahun 2023
 Anton sangat malas mandi di rumah
 Lorraine sedang pergi ke pasar Sanglah
 Sebuah jaket tergantung di belakang pintu
 Lia sering berbelanja sangat banyak di Pasar Badung
 Istana Presiden Pak Jokowi

Tabel 4.1 Testing Data Kalimat Diterima

Tidak Diterima

Memberikan informasi perihal rencana perkuliahan minggu depan
 Tentang kebijakan baru perusahaan
 Laporan perihal data hasil survei
 Kebijakan baru ini dengan perkembangan teknologi baru
 Tidak paham tentang keputusan panitia
 Baju ini membuat
 Dibangun dari kayu jati
 Dengan giat belajar
 Dengan sangat nyenyak tidur di lantai
 Tersenyum dengan hangat kepada Putu
 Sepeda motor dengan sangat hati-hati sekali
 Dengan cekatan di lantai
 Berlayar ke pulau Jawa
 Menganyam syal indah itu dengan teliti
 Dengan sangat gembira ke Tabanan
 Jalan-jalan dengan santai di taman kota
 Sering pulang ke Papua
 Ke rumah Kresna

Tabel 4.2 Testing Data Kalimat Tidak Diterima

Setelah melakukan pengujian menggunakan *dataset* kalimat-kalimat Bahasa Indonesia yang diterima dan tidak diterima, diperoleh hasil pengujiannya adalah sebagai berikut

```
Anak-anak pulang sekolah setelah jam satu Diterima
Rio ke sekolah sebelum pukul delapan Diterima
Kelas hari ini dimulai hingga dua belas siang Diterima
Tia berolahraga setiap pagi Diterima
Mereka meninggalkan pantai sesaat sebelum hujan Diterima
Laporan itu dikumpulkan sebelum pukul lima sore Diterima
Ujian akan berakhir dalam waktu dua jam Diterima
Mereka akan mengadakan pertemuan pada hari Minggu Diterima
Buku ini memberikan informasi mengenai sejarah kemerdekaan Indonesia Diterima
Dosen memberikan informasi perihal rencana perkuliahan minggu depan Diterima
Saya ingin mengetahui tentang kebijakan baru perusahaan Diterima
Rani menyusun laporan perihal data hasil survei Diterima
Kebijakan baru ini sangat berkaitan dengan perkembangan teknologi baru Diterima
Desi tidak paham tentang keputusan panitia Diterima
Baju ini dibuat dari katun Diterima
Rumah itu dibangun dari kayu jati Diterima
Nanda belajar dengan giat Diterima
Farel tidur dengan sangat nyenyak di lantai Diterima
Puja tersenyum dengan hangat kepada Putu Diterima
David mengendarai sepeda motor dengan sangat hati-hati sekali Diterima
Shyta sedang mengambil permen dengan cekatan di lantai Diterima
Nanda berlayar ke pulau Jawa Diterima
Ilham berlari dengan terburu-buru ke taman kanak-kanak dekat rumahnya Diterima
Aku melihat Agus dari kejauhan Diterima
Krisnawan memakan semua hidangan itu dengan lahap Diterima
Nanda menegur Ilham atas kesalahannya untuk kebaikan bersama Diterima
Aku menangis dengan tersedu-sedu karena kematian anjingku Diterima
Wahyu memindahkan tanaman kesayangannya dalam hitungan detik ke dalam kamarnya Di
Nanda menendang pintu kamarnya dengan sangat kesal karena kalah taruhan Diterima
Calvin membawa rapornya dengan sedih Diterima
Marcelino menendang bola itu dengan kekuatan penuh ke gawang Diterima
Kavin mengerjakan soal ujian akhir semester dengan serius Diterima
Okta menganyam syal indah itu dengan teliti Diterima
Kresna sering pulang dengan sangat gembira ke Tabanan Diterima
Farel jalan-jalan dengan santai di taman kota Diterima
Yunus sering pulang ke Papua Diterima
Anton jarang pergi ke rumah Kresna Diterima
Anton lari maraton di Lapangan Renon Diterima
Pak Jokowi di dalam Istana Presiden Diterima
Nanda ke pulau Jawa pada tahun 2023 Diterima
Anton sangat malas mandi di rumah Diterima
Lorraine sedang pergi ke pasar Sanglah Diterima
Sebuah jaket tergantung di belakang pintu Diterima
Lia sering berbelanja sangat banyak di Pasar Badung Diterima
-----
Kalimat ditolak:
Dosen memberikan sangat banyak tugas untuk dikerjakan mahasiswa Ditolak
Accept: 79
Reject: 1

(venv) C:\Users\Acer\Desktop\Wanda\TBO\Parse Bahasa>
```

Gambar 4.1 Hasil Pengujian Kalimat Diterima

Terlihat bahwa dari seluruh kalimat yang digunakan untuk pengujian kalimat diterima hanya 1 yang tidak diterima, hal tersebut karena pada pola kalimat tersebut

‘Dosen memberikan sangat banyak tugas untuk dikerjakan mahasiswa’ seharusnya adalah S P O Ket dimana Subjeknya adalah ‘Dosen’, Predikatnya adalah ‘memberikan’, Objeknya adalah ‘sangat banyak tugas’, Keterangan nya adalah ‘untuk dikerjakan mahasiswa’. Alasan kalimat tersebut ditolak adalah karena dalam rule CFG yang digunakan hanya terdapat O -> NP, tetapi tidak terdapat suatu aturan penyusunan yang memungkinkan penyusunan NP dengan Adj di depan NP (AdjP NP) sehingga algoritma CYK gagal mengenali ‘sangat banyak tugas’ sebagai Objek dan kalimatpun terdeteksi tidak baku

```
(venv) C:\Users\Acer\Desktop\Nanda\TBO\Parse Bahasa>py app.py
Kalimat diterima:
Laporan perihal data hasil survei Diterima
Kebijakan baru ini dengan perkembangan teknologi baru Diterima
Baju ini membuat Diterima
Sepeda motor dengan sangat hati-hati sekali Diterima
Istana Presiden Pak Jokowi Diterima
-----
Kalimat ditolak:
Memberikan informasi perihal rencana perkuliahan minggu depan Ditolak
Tentang kebijakan baru perusahaan Ditolak
Tidak paham tentang keputusan panitia Ditolak
Dibangun dari kayu jati Ditolak
Dengan giat belajar Ditolak
Dengan sangat nyenyak tidur di lantai Ditolak
Tersenyum dengan hangat kepada Putu Ditolak
Dengan cekatan di lantai Ditolak
Berlayar ke pulau Jawa Ditolak
Menganyam syal indah itu dengan teliti Ditolak
Dengan sangat gembira ke Tabanan Ditolak
Jalan-jalan dengan santai di taman kota Ditolak
Sering pulang ke Papua Ditolak
Ke rumah Kresna Ditolak
Lari maraton di Lapangan Renon Ditolak
Accept: 5
Reject: 15
```

Gambar 4.2 Hasil Pengujian Kalimat Tidak Diterima

Dapat dilihat bahwa terdapat 5 kalimat yang seharusnya tidak diterima tetapi diterima oleh sistem, hal ini dikarenakan oleh bagaimana penyusunan *rule* CFG yang digunakan oleh algoritma CYK sistem. Contohnya adalah seperti pada kalimat ‘Laporan perihal data hasil survei’, seluruh kata yang ada kalimat tersebut merupakan noun dan terdapat rule K -> S P, S -> NP, P -> NP, NP -> noun | NP Noun, sehingga kalimat

tersebut menggunakan NP yang tersusun oleh noun yang terdapat pada kalimat tersebut sebagai Predikat dan Subjek sehingga kalimat tersebut diterima.

Untuk kalimat 'Kebijakan baru ini dengan perkembangan teknologi baru', Subjeknya adalah 'Kebijakan' karena kata tersebut adalah noun sehingga dapat menjadi NP, Predikatnya adalah 'baru' karena kata tersebut adalah Adj dan AdjP dapat digunakan sebagai Predikat, Objeknya adalah 'ini' karena kata tersebut adalah pronoun yang dapat menjadi NP. Pelengkapnya adalah 'dengan perkembangan teknologi baru' dimana substring tersebut adalah PP dengan rule PP -> Prep NP. Sehingga sistem ini menganggap bahwa kalimat tersebut memiliki pola S P O Pel yang dapat diterima.

Kalimat 'Baju ini membuat' diterima oleh sistem karena rule yang memperbolehkan VP digunakan sebagai predikat dalam kalimat. Kalimat ini berpola S P dimana S adalah NP (Baju ini) dan P adalah VP (membuat) sehingga sistem menganggap bahwa kalimat ini baku.

Untuk kalimat 'Sepeda motor dengan sangat hati-hati sekali' dideteksi oleh sistem memiliki pola kalimat S P, hal tersebut karena aturan NP -> Noun | NP Noun memungkinkan 'Sepeda Motor' untuk menjadi Subjek karena rule dari subjek adalah S -> NP dan 'dengan sangat hati-hati sekali' memungkinkan dipakai sebagai Predikat karena Predikat memiliki aturan P -> Prep AdvP dan terdapat aturan AdvP -> Adv | Adv AdvP sehingga sistem juga mengira kalimat ini berbentuk baku.

BAB V

PENUTUP

5.1 Kesimpulan

Context Free Grammar (CFG) adalah sebuah formalisme tata bahasa yang berguna untuk analisis sintaksis pada Bahasa Indonesia baku. Untuk melakukan analisis sintaksis tersebut, dibutuhkan sekumpulan aturan produksi dari CFG. Dalam implementasinya dengan menggunakan bahasa pemrograman Python, algoritma Cocke–Younger–Kasami (CYK) dengan metode Table Filling dapat digunakan untuk melakukan analisis sintaksis pada Bahasa Indonesia baku, namun membutuhkan CFG yang telah diubah menjadi bentuk Normal Chomsky (CNF). Untuk memudahkan proses analisis, aplikasi berbasis web dikembangkan dengan menggunakan streamlit. Berdasarkan hasil dan pembahasan, kesimpulan dapat diambil bahwa dari 100 kalimat yang diuji, sistem berhasil menebak dengan benar sebanyak 94 kalimat. Dengan demikian, aplikasi ini dapat melakukan analisis sintaksis pada kalimat-kalimat Bahasa Indonesia baku dengan tingkat akurasi sebanyak 94% dan juga mampu menampilkan triangular table.

DAFTAR PUSTAKA

Carroll, J., & Long, D. (1989). Theory of finite automata with an introduction to formal languages.