

**PERANGKAT LUNAK SISTEM PENENTUAN KATA DASAR
SUATU KATA DALAM SUATU KALIMAT
SECARA OTOMATIS**

Tacbir Hendro Pudjiantoro

**Sekolah Tinggi Manajemen Informatika dan Komputer
MARDIRA INDONESIA
Jalan Soekarno-Hatta No. 11 Bandung**

Abstrak

Bahasa merupakan alat komunikasi untuk menyampaikan gagasan atau ide dari komunikator ke komunikan yang bisa saling dimengerti. Banyak aneka cara dalam menyampaikan gagasan kepada komunikan : bisa dengan bahasa lisan (*oral*) bahasa tulis (*written*), dan bahkan dengan bahasa gerak tubuh (*gesture*). Bahasa Indonesia merupakan salah satu dari ribuan bahasa dan merupakan bahasa Nasional yang dibangun berdasarkan atas keempat dasar ini, sehingga penggunaan bahasa Indonesia haruslah mengacu pada aturan yang baik dan benar.

Kata-kata kunci: kata dasar

1. PENDAHULUAN

Bahasa merupakan alat komunikasi untuk menyampaikan gagasan atau ide dari komunikator ke komunikan yang bisa saling dimengerti. Banyak aneka cara dalam menyampaikan gagasan kepada komunikan: bisa dengan bahasa lisan (*oral*), bahasa tulis (*written*), dan bahkan dengan bahasa gerak tubuh (*gesture*).

Di dunia ini terdapat ribuan bahasa dan mempunyai susunan (*structure*) kalimat yang berbeda pula. Akan tetapi, dari keanekaragaman bahasa tersebut mempunyai ciri yang sama, yaitu kesemestaan bahasa (*language universals*) yang mencakup :

1. Maksud atau fungsi bahasa adalah sebagai sarana komunikasi.
2. Media utama bahasa adalah bunyi ujaran (*vocal sound*).
3. Semua bahasa memiliki leksikon atau kosa kata yang mengandung makna.
4. Semua bahasa memiliki tata bahasa atau grammer (Elson & pickett ; 1962 : 1).

Bahasa Indonesia merupakan salah satu dari ribuan bahasa dan merupakan bahasa Nasional yang dibangun berdasarkan atas keempat dasar ini, sehingga penggunaan bahasa Indonesia haruslah mengacu pada aturan yang baik dan benar.

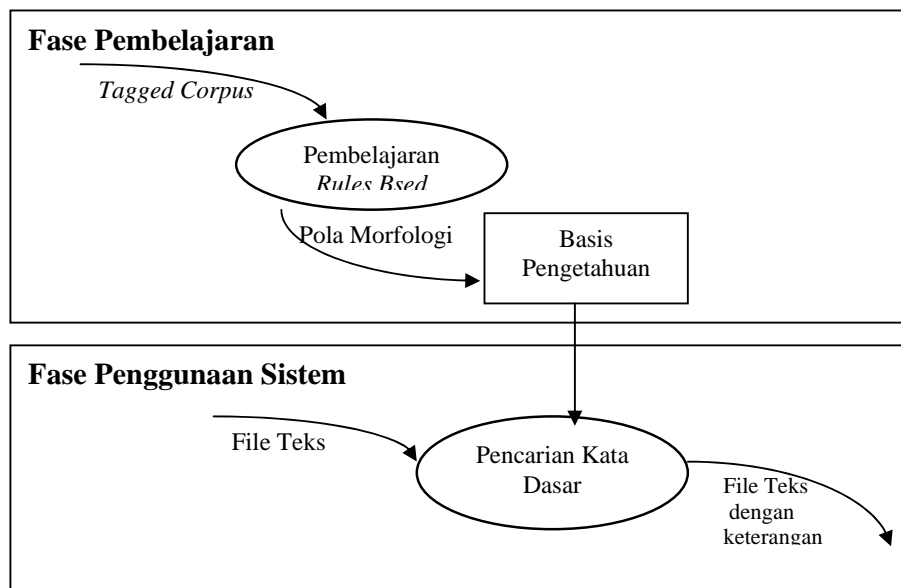
Secara umum bahasa terdiri dari lisan dan tulisan yang merupakan kumpulan dari kata-kata yang telah disusun menjadi sebuah kalimat. Setiap kata yang digunakan tersebut

masing-masing mempunyai kata dasar, dalam tata bahasa Indonesia sebuah kata dasar akan mengalami penambahan baik awalan, sisipan dan akhiran yang tentunya dapat memberikan penekanan arti atau bahkan mengandung arti yang berbeda. Untuk mempelajari dan memahami sebuah kata dasar yang terkandung dalam sebuah kata atau kalimat tentunya memerlukan suatu rujukan aturan dalam bahasa Indonesia, dan aturan itu dikenal dalam ilmu bahasa sebagai ilmu **morfologi** atau ilmu yang mempelajari perubahan bunyi dan bentuk dari morfem pembentuk sebuah kata.

Untuk mempermudah dan membantu pengajaran morfologi bagi guru, khususnya guru bahasa Indonesia yang selama ini masih menggunakan metode tatap muka dan menggunakan buku panduan yang telah ditetapkan, nampaknya perlu dibangun pengajaran morfologi dengan program komputer (*computer based*). Untuk itu sistem berbasis pengetahuan bisa dijadikan solusi pemecahan masalahnya. Sistem berbasis pengetahuan itu sendiri berarti sebuah program komputer yang mengambil pemikiran seorang pakar atau sebuah aturan yang ada untuk membantu memecahkan suatu permasalahan.

2. RUMUSAN MASALAH

Arsitektur Sistem Pencarian Kata Dasar dalam Bahasa Indonesia menggunakan metode pembelajaran *Rules Based Supervised* adalah seperti pada gambar 1.



Gambar 1. Arsitektur Sistem Pencarian Kata Dasar dalam Bahasa Indonesia menggunakan metode pembelajaran *Rules Based Supervised*

3. LANDASAN TEORI

3.1 Supervised Learning

Supervised Learning (Pembelajaran terselia) adalah sebuah metoda pembelajaran yang diawasi. Pengawasannya di sini dilakukan pada sebuah data *Corpus* yang telah ditandai (*tag*) pada setiap katanya. Tanda (*tag*) dalam kata tersebut adalah kata dasar dari kata tersebut.

contoh:

Pola : [awalan // kata dasar // akhiran]

[memper jual belikan]

Pola : [awalan // kata dasar]

[men dapat]

Pola : [kata dasar // akhiran]

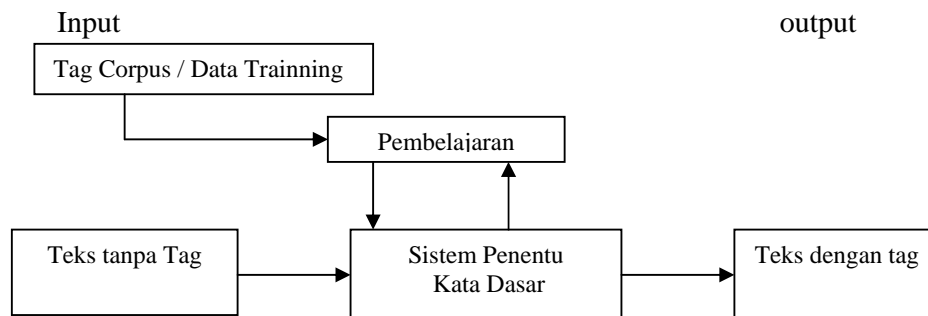
[sandara an]

Pemberian tanda (*tag*) dilakukan secara manual sebelum data *Corpus* dijadikan *data training*. *Supervised Learning* adalah metoda pembelajaran yang diberikan contohnya dan pada setiap contoh tersebut diperlihatkan apa yang akan ditampilkan.

Blok Diagram Model Supervised

Inputan dalam sistem penentu kata dasar adalah sebuah teks tanpa *tag*, proses pencarian kata dasar yang pertama kali adalah membandingkan kata tersebut dengan kata-kata dasar yang ada di *tag corpus*. Selanjutnya dilakukan perbandingan dengan aturan morfologi (perubahan bentuk kata diakibatkan bertemunya sebuah morfem dengan morfem lain).

Dalam bentuk diagram dapat dilihat pada gambar 3.



Gambar 2. Blok Diagram Model Supervised

Konsep Pembelajaran

Dalam metoda pembelajaran induktif, pengajar menyediakan contoh untuk yang belajar, dan yang belajar mencari aturan umum untuk dijadikan sebagai aturan induksi secara umum. Secara umum, konsep pembelajaran dari contoh dapat dibuat secara simbolik seperti:

$U \rightarrow$ himpunan Universal dari objek

$C \rightarrow$ subset dari objek pada U

Pembelajaran konsep C berarti belajar untuk memahami objek dalam C . Objek diterangkan dalam pernyataan atributnya, sementara atribut mempunyai jumlah nilai yang mungkin. Tugas induksi adalah untuk mengembangkan sebuah klasifikasi aturan yang dapat digunakan untuk mendeterminasikan kelas dari setiap objek pada nilai dari atribut. Semua metoda pembelajaran induksi mempunyai tujuan yang sama yaitu untuk menginduksi aturan klasifikasi dari contoh. Sistem ini sering membangkitkan pengetahuan yang dapat diperlihatkan dalam bentuk pohon keputusan dimana dapat memberikan penyelesaian masalah yang sulit pada prakteknya.

Kata Dasar

Kata dasar adalah satuan terkecil yang menjadi asal atau permulaan sesuatu kata kompleks. Contohnya kata *bersandaran*, yang terbentuk dari kata dasar sandar memperoleh afiks-*an* menjadi sandaran, dan selanjutnya memperoleh afiks *ber-* menjadi bersandar.

Dasar kata adalah satuan, baik tunggal atau kompleks, yang menjadi dasar pembentukan bagi satuan yang lebih besar atau lebih kompleks. Kita ambil contoh kata *bersandaran* tadi, yang terbentuk dari dasar kata *sandar* dengan afiks *ber-*; seterusnya kata *sandaran* terbentuk dari dasar kata sandar dengan afiks-*an*. Kata *berkelanjutan* terbentuk dari dasar kata *kelanjutan* terbentuk dari kata dasar *lanjut* dengan afiks *ke-an*.

Perlu disadari dan dipahami benar-benar bahwa kata dasar, selalu berupa bentuk tunggal, sedangkan dasar kata mungkin berupa bentuk tunggal dan mungkin pula bentuk kompleks.

Contoh kata dasar :

| | | |
|-------|------|------------|
| dapat | pada | mendapat |
| mau | pada | kemauan |
| ajar | pada | pengajaran |

Contoh dasar kata :

- Berbentuk tunggal

| | | |
|-------|------|-----------|
| lamar | pada | melamar |
| buka | pada | terbuka |
| ajar | pada | pelajaran |

- Berbentuk kompleks

| | | |
|-------------|------|-----------------|
| terbelakang | pada | keterbelakangan |
| terbaca | pada | keterbacaan |
| berada | pada | keberadaan |

Pengertian Morfologi

Morfologi adalah ilmu yang mempelajari seluk-beluk bentuk kata serta fungsi perubahan-perubahan bentuk kata itu, baik fungsi gramatik maupun fungsi semantik (Ramlan, 1983 :16-17).

Dalam morfologi kita tidak menelaah bunyi tunggal beserta varian-variannya saja, tetapi justru menelaah bunyi-bunyi rangka beserta varian-ariannya (Heatherington;1980:47).

Atau lebih jelas lagi: "Morfologi mempelajari perubahan-perubahan fonem yang timbul sebagai akibat pertemuan morfem dengan morfem lain". (Ramlan:1983:73). [2], Berbicara mengenai proses morfologi dalam bahasa Indonesia, maka terdapat tiga hal yang penting yaitu :

- a) Proses perubahan fonem
- b) Proses penambahan fonem
- c) Proses penanggalan/fonem.

a) Proses Perubahan Fonem

Apabila kita menyinggung perubahan fonem dalam bidang proses morfofonemik dalam bahasa indonesia, maka dalam dua hal yang perlu mendapat perhatian, yaitu:

- a.1) Perubahan fonem/N/
- a.2) Perubahan fonem/R/

sudah barang tentu perubahan-perubahan fonem tersebut tergantung dari kata dasar yang mengikutinya.

a.1) Perubahan Fonem /N/

Adapun kaidah-kaidah perubahan fonem /N/ yaitu :

- (1) Fonem /N/ pada morfem {meN-} dan morfem {peN} berubah menjadi fonem

/m/ kalau kata dasar yang mengikutinya berawal dengan /b,f,p/ misalnya :

| | | | | | | | |
|-----|---|-------|---------|-----|---|--------|-----------|
| meN | + | bawa | membawa | peN | + | bela | pembela |
| meN | + | foto | memfoto | peN | + | fitnah | pemfitnah |
| meN | + | pakai | memakai | peN | + | pecah | pemecah |

- (2) Fonem /N/ pada {meN-} dan {peN-} berubah menjadi fonem /n/ kalau kata dasar yang mengikutinya berawal dengan fonem /d,s,t,/. Perlu kita catat fonem /s/ hanya khusus bagi sejumlah kata dasar yang berasal dari bahasa asing. Misalnya :

| | | | | | | | |
|-----|---|--------|-----------|-----|---|--------|-----------|
| meN | + | daki | mendaki | peN | + | debat | pendebat |
| meN | + | survei | mensurvei | peN | + | suplai | pensuplai |
| meN | + | tahan | menahan | peN | + | tembak | penembak |

- (3) Fonem /N/ pada {meN-} dan {peN-} berubah menjadi /n/ atau /ny/ apabila kata dasar yang mengikutinya berawal dengan /c, j, s, sy/ misalnya :

| | | | | | | | |
|-----|---|---------|-----------|-----|---|------------|--------------|
| meN | + | cabut | mencabut | peN | + | cari | pen cari |
| meN | + | jual | menjual | peN | + | jepit | penjepit |
| meN | + | sayat | penyayat | peN | + | suruh | penyuruh |
| meN | + | syukuri | menyukuri | meN | + | syariatkan | menyariatkan |

- (4) Fonem /N/ pada {meN-} dan {peN-} berubah menjadi /ng/ apabila kata dasar yang mengikutinya berfonem awal /g, h, k, kh/ dan vokal. Misalnya :

| | | | | | | | |
|-----|---|----------|--------------|-----|---|--------|------------|
| meN | + | ganti | mengganti | peN | + | gubah | penggubah |
| meN | + | harap | mengharap | peN | + | halang | penghalang |
| meN | + | kait | mengait | peN | + | karang | pengarang |
| meN | + | khianati | mengkhianati | peN | + | khayal | pengkhayal |
| meN | + | eja | mengeja | peN | + | adu | pengadu |

a.2) Perubahan Fonem / R/

Fonem /r/ pada morfem {ber} dan morfem {per} berubah menjadi fonem /l/ sebagai akibat pertemuan morfem tersebut dengan kata dasar yang berupa morfem {ajar}.

Perlu dicatat bahwa proses perubahan fonem /r/ ini tidak produktif dalam bahasa Indonesia.

Contoh:

Ber- + ajar belajar Per- + ajar pelajar

b) Proses Penambahan Fonem

Proses penambahan fonem biasanya terjadi pada kata dasar yang bersuku kata tunggal. Hal inipun sangat terbatas dan terjadi sebagai akibat pertemuan kata dasar yang bersuku tunggal dengan morfem {meN-} dan {peN-}. Fonem tambahan itu ialah /nge/, Misalnya:

| | | | | | | | |
|-----|---|-----|----------|-----|---|-----|----------|
| meN | + | bom | mengebom | peN | + | las | pengelas |
| meN | + | cat | mengecat | peN | + | bor | pengebor |

c) Proses Penanggalan Fonem

Dalam proses penanggalan fonem ini terjadi tiga hal yang perlu untuk dibicarakan, yaitu :

1. Proses penanggalan fonem /N/ pada kata dasar yang berawalan /l,r,y,w,n/

Contoh :

| | | | | | | | |
|-----|---|----------|------------|-----|---|-----------|-------------|
| meN | + | lawan | melawan | peN | + | lawak | pelawak |
| meN | + | rapatkan | merapatkan | peN | + | rasa | perasa |
| meN | + | yakinkan | meyakinkan | | | | |
| meN | + | wajibkan | mewajibkan | peN | + | wawancara | pewawancara |
| meN | + | nasihati | menasihati | peN | + | minum | peminum |

2. Proses penanggalan fone /r/ pada morfem-morfem {ber, per, ter} pada kata dasar yang berawalan dengan fonem /e'/ dan kata dasar yang suku pertamanya berakhiran /er/

Contoh :

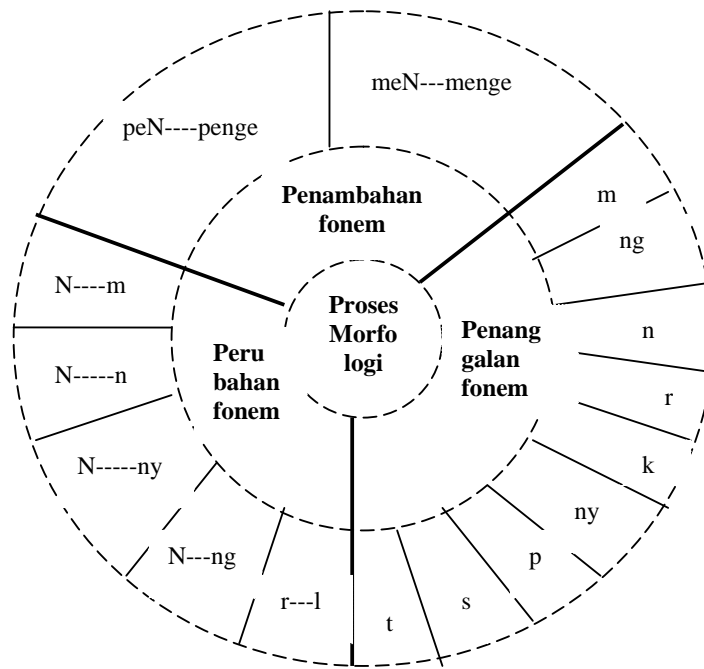
| | | | | | | | |
|-----|---|-------|---------|-----|---|---------|-----------|
| Ber | + | racun | beracun | ber | + | serbuk | berserbuk |
| Per | + | rindu | perindu | ter | + | perdaya | teperdaya |

3. Proses penghilangan fonem-fonem /k, p, t, s/ sebagai akibat pertemuan dengan morfem-morfem {meN, peN} dengan kata dasar yang bermula dengan fonem-fonem tersebut.

Contoh :

| | | | |
|----------------|------------|--------------|----------|
| meN + kacau | mengacau | peN + kagum | pengagum |
| meN + perintah | memerintah | peN + pesan | pemesan |
| meN + tukar | menukar | peN + tumpas | penumpas |
| meN + susun | menyusun | peN + sapu | penyapu |

Untuk memperjelas aturan morfologi di atas dapat dilihat dalam gambar 10.



Gambar 3. Proses morfologi

Dasar pengetahuan dari sistem berbasis pengetahuan akan dituangkan dalam kajian bagaimana kaidah-kaidah morfologi dalam bahasa Indonesia bisa dituangkan dalam komputer, sebagai media dalam proses transformasi ilmu pengetahuan ilmu morfologi. Kajian dari dasar pengetahuan tersebut akan dikenalkan pada tulisan ini.

Sebagai basis pengetahuan dalam pencarian kata dasar maka akan direpresentasikan dalam regular expression, *finite automata* yang terkumpul dalam metode *scanning* dan

penentuan CFG (*Contex Free Grammer*), serta pohon urai yang terkumpul dalam metode parsing.

Pertama *regular expression* yang bertugas membaca karakter input dan memproduksi token yang diperlukan. Token-token tersebut biasanya dipakai menjadi kata kunci. Lalu token yang dijadikan sebagai kata kunci akan dikirim kepengurai (*parsing*). Dalam penyusunan *regular expression* aturan didalamnya, diantaranya adalah :

1. Simbol (|) menyatakan pilihan (atau)
2. Simbol (U) menyatakan gabungan dari satu notasi dengan notasi yang lainnya
3. Simbol (*) atau (+) menyatakan statement berantai atau berulang
4. Simbol () menyatakan statement pembuka dan penutup

Setelah token-token yang diperlukan dibentuk, maka langkah selanjutnya adalah dengan membuat *finite automata* untuk masing-masing kemungkinan yang terjadi pada program. Dalam hal ini adalah pencarian kata dasar dalam bahasa Indonesia.

Fungsi *finite automata* adalah untuk menggambarkan kemungkinan-kemungkinan yang bisa terjadi. Digambarkan seperti bentuk jaringan kerja yang mempunyai *node*. Dimana node yang satu dengan node yang lainnya bisa saling berkaitan.

Setelah tahap scanning selesai, maka dilanjutkan ke proses parsing yang terdiri dari pembentukan CFG (*Contex Free Grammer*) dimana aturan-aturan proses morfologi didefinisikan di sini. Khususnya yang berkaitan dengan sistem pencarian kata dasar dalam bahasa Indonesia. Tahap terakhir adalah dengan membentuk pengurai dari keseluruhannya dalam bentuk pohon urai.

Regular Expression

Bagian pertama dari dasar pengetahuan ini adalah bagaimana kaidah-kaidah morfologi dalam tata bahasa Indonesia dituangkan dalam komputer. Pada bagian ini akan dipelajari bagaimana perubahan-perubahan morfem dan fonem dalam proses nasalisasi dapat diterima dalam kaidah komputer. Dalam *regular expression* akan dibentuk token-token yang akan digunakan pada implementasi dalam kaidah morfologi dalam komputer. Adapun token-token tersebut adalah sebagai berikut :

Hrp = (a | b | c | d | e | f | g | h | i | j | k | l | m | n | o | p | q | r | s | t | u | v | w | x | y | z | A | | C | D |
E | F | G | H | I | J | K | L | M | N | O | P | Q | R | S | T | U | V | W | X | Y | Z)

K = (b | c | d | f | g | h | j | k | l | m | n | p | q | r | s | t | v | w | x | y | z)
 V = (a | i | e | u | o)
 Afk = ((me | ber | per | di | ke | ter | se | pe) | (an | kan | nya | i) | (el | er | em))
 Afkser = (is | ik | isasi | al | i | iah | wiah | ni)
 Sp = (_)
 Tk = eol (.)
 Kt = ((k | v) * eol)
 Kl = ((k | v) * | _ | not(eol) | k | v | eol |) *
 MeN = (mem | men | meny | meng | me | menge)
 Mem = (b | f | p)
 Men = (d | s | t)
 Meny = (s)
 Meng = (g | h | k | x | a | i | e | u | o)
 Me = (l | r | w | y | n)
 Menge = (sk1) / (kata dasar bersuku kata tunggal)
 PeN = (pem | pen | peny | peng | pe | penge)
 Pem = (b | f | p)
 Pen = (d | s | t)
 Peny = (s)
 Peng = (g | h | k | x | a | i | e | u | o)
 Pe = (l | r | w | y | n)
 Penge = (sk1) / (kata dasar bersuku kata tunggal)
 Ber = (be | bel | ber)
 Be = (r | sk1=r) / (kata dasar yang suku pertamanya berakhiran er)
 Bel = (kt=ajar)
 Ber = ((a | b | c | d | e | f | g | h | i | j | k | l | m | n | o | p | q | s | t | u | v | w | x |
 y | z) | (not kt=ajar))
 Per = (pe | pel | per)
 Pe = (r | sk1=r) / (kata dasar yang suku pertamanya berakhiran er)
 Pel = (kt=ajar)
 Per = ((a | b | c | d | e | f | g | h | i | j | k | l | m | n | o | p | q | s | t | u | v | w | x |
 y | z) | (not kt=ajar))

Ter = (te | ter)
 Te = (r | sk1=r) / (kata dasar yang suku pertamanya berakhiran er)
 Ter = ((a | b | c | d | e | f | g | h | i | j | k | l | m | n | o | p | q | s | t | u | v | w | x |
 y | z) | (sk1=r)) / (kata dasar yang suku pertamanya tidak berakhiran er)
 Di = (di)
 Se = (se)
 Pe = (pe)

Keterangan token :

Hrf = merupakan hurup-hurup abjad dalam bahasa Indonesia yang mana apabila dirangkai akan membentuk sebuah kata.
 K = kumpulan hurup-hurup konsonan dalam bahasa Indonesia
 V = kumpulan hurup-hurup vokal dalam bahasa Indonesia
 Afk = kumpulan imbuhan yang digabung dengan kata dasar dalam bahasa Indonesia, imbuhan itu terdiri atas awalan akhiran dan sisipan.
 Afkser = kumpulan imbuhan serapan yang digabung dengan kata dasar yang berasal dari bahasa asing
 Sp = token yang digunakan untuk menandakan karakter spasi
 Tk = karakter titik menandakan akhir dari sebuah kalimat atau bisa disebut juga dengan token eol (end of line)
 Kt = token yang digunakan untuk menandakan sebuah kata dalam sebuah kalimat
 Kl = token yang digunakan untuk sebuah kalimat
 meN = token yang menandakan terjadinya proses nasalisasi dari awalan me berubah menjadi : mem, men, meny, meng, me, menge. (Proses nasalisasi diinisialkan dengan tanda N).
 Mem = merupakan token yang menampung hurup-hurup awal dari kata dasar yang mana apabila bertemu dengan awalan meN maka berubah bunyi dan penulisannya menjadi mem.
 Men = merupakan token yang menampung hurup-hurup awal dari kata dasar yang mana apabila bertemu dengan awalan meN maka berubah bunyi dan penulisannya menjadi men.

- Meny = merupakan token yang menampung hurup-hurup awal dari kata dasar yang mana apabila bertemu dengan awalan meN maka berubah bunyi dan penulisannya menjadi meny.
- Meng = merupakan token yang menampung hurup-hurup awal dari kata dasar yang mana apabila bertemu dengan awalan meN maka berubah bunyi dan penulisannya menjadi meng.
- Me = merupakan token yang menampung hurup-hurup awal dari kata dasar yang mana apabila bertemu dengan awalan meN maka berubah bunyi dan penulisannya menjadi me.
- Menge = merupakan token yang menampung hurup-hurup awal dari kata dasar yang mana apabila bertemu dengan awalan meN maka berubah bunyi dan penulisannya menjadi menge.
- PeN = token yang menandakan terjadinya proses nasalisasi dari awalan pe berubah menjadi : pem, pen, peny, peng, pe, penge.
- Pem = merupakan token yang menampung hurup-hurup awal dari kata dasar yang mana apabila bertemu dengan awalan peN maka berubah bunyi dan penulisannya menjadi pem.
- Pen = merupakan token yang menampung hurup-hurup awal dari kata dasar yang mana apabila bertemu dengan awalan peN maka berubah bunyi dan penulisannya menjadi pen.
- peny = merupakan token yang menampung hurup-hurup awal dari kata dasar yang mana apabila bertemu dengan awalan peN maka berubah bunyi dan penulisannya menjadi peny.
- peng = merupakan token yang menampung hurup-hurup awal dari kata dasar yang mana apabila bertemu dengan awalan peN maka berubah bunyi dan penulisannya menjadi peng.
- pe = merupakan token yang menampung hurup-hurup awal dari kata dasar yang mana apabila bertemu dengan awalan peN maka berubah bunyi dan penulisannya menjadi pe.
- Penge = merupakan token yang menampung hurup-hurup awal dari kata dasar yang mana apabila bertemu dengan awalan peN maka berubah bunyi dan penulisannya menjadi penge.

- Ber = token yang menandakan terjadinya proses nasalisasi dari awalan ber berubah menjadi : be, bel, ber.
- Be = merupakan token yang menampung hurup-hurup awal dari kata dasar dan kata dasar yang suku kata pertamanya berahir ‘r’ apabila bertemu dengan awalan ber maka berubah bunyi dan penulisannya menjadi be.
- Bel = merupakan token yang menampung khusus kata dasar “ajar” yang mana apabila bertemu dengan awalan ber maka berubah bunyi dan penulisannya menjadi bel.
- Ber = merupakan token yang menampung hurup-hurup awal dari kata dasar dan bukan kata dasar “ajar”, yang mana apabila bertemu dengan awalan ber maka tidak berubah bunyi dan penulisannya, tetap menjadi ber.
- Per = token yang menandakan terjadinya proses nasalisasi dari awalan per berubah menjadi : pe, pel, per.
- pe = merupakan token yang menampung hurup-hurup awal dari kata dasar dan kata dasar yang suku kata pertamanya berahir ‘r’ apabila bertemu dengan awalan per maka berubah bunyi dan penulisannya menjadi pe.
- Pel = merupakan token yang menampung khusus kata dasar “ajar” yang mana apabila bertemu dengan awalan per maka berubah bunyi dan penulisannya menjadi pel.
- Per = merupakan token yang menampung hurup-hurup awal dari kata dasar dan bukan kata dasar “ajar”, yang mana apabila bertemu dengan awalan per maka tidak berubah bunyi dan penulisannya, tetap menjadi per.
- Ter = token yang menandakan terjadinya proses nasalisasi dari awalan ter berubah menjadi te dan ter.
- Te = merupakan token yang menampung hurup-hurup awal dari kata dasar dan kata dasar yang suku kata pertamanya berahir ‘r’ apabila bertemu dengan awalan ter maka berubah bunyi dan penulisannya menjadi te.
- Ter = merupakan token yang menampung hurup-hurup awal dari kata dasar dan kata dasar yang suku kata pertamanya tidak berahir ‘r’ apabila bertemu dengan awalan ter maka tidak berubah bunyi dan penulisannya tetap menjadi ter.
- Di, se, ke = masing-masing berfungsi sebagai awalan yang tidak mengalami perubahan bunyi dan penulisannya dan menampung semua kata dasar

Parser

Pada kajian ini akan dibahas bagaimana sebuah kalimat dalam bahasa Indonesia dapat diuraikan menjadi sebuah pohon urai, yang menerangkan rangkaian dari huruf-huruf menjadi sebuah string dan terdapat kaidah morfologi/kata dasar didalamnya. Dalam sebuah kalimat akan terdapat sebuah aturan morfologi atau bahkan lebih.

Dalam kajian *regular expression* token yang membentuk sebuah kata adalah terdiri dari token hrf (huruf-huruf dalam bahasa Indonesia) yaitu terdiri dari vokal dan konsonan, sedang aturan morfologi terjadi apabila bertemu antara afik me, pe, ber, per, ter dengan kata dasar yang memenuhi aturan morfologi maka terjadi perubahan bunyi dan penulisanya.

Parser sendiri terbentuk berdasarkan aturan bahasa yang membentuknya dengan tujuan membentuk pohon urai untuk masing-masing tata bahasa yang terbentuk didalamnya. Pada kajian morfologi disini, tata bahasa yang akan terbentuk berkaitan dengan aturan-aturan morfologi bahasa Indonesia yang dibahas. Namun sebelum pembentukan parser yang akan menggunakan metoda *top down parser* yaitu membentuk pohon urai yang dimulai dari atas ke bawah, terlebih dahulu dibentuk CFG (*Context Free Grammar*) sebagai landasan teori untuk memperkuat tata bahasa aturan morfologi.

Komponen dari *Context Free Grammar* akan dibentuk oleh empat komponen yaitu (V_t , V_n , S , P) dimana , komponen itu diambil dari teori chomski, dimana :

- V_t : Himpunan lambang terminal
- V_n : Himpunan lambang non terminal
- P : Himpunan daftar produksi
- S : Simbol awal

Komponen CFG (*Context Free Grammar*)

1. S (Simbol awal)

S = program

2. V_t (Himpunan lambang terminal)

{Morfem mem, morfem men, morfem meny, morfem meng, morfem me, morfem menge, morfem pem, morfem pen, morfem peny, morfem peng, morfem pe, morfem penge, morfem ber, morfem bel, morfem be, morfem per, morfem pel, morfem pe, morfem ter, morfem te}

3. Vn (himpunan lambang nonterminal)

{ hrf, k, v, kt, kl, sp, tk, mem, men, meny, meng, me, menge, pem, pen, peny, peng, pe, penge, ber, bel, be, per, pel, pe, ter, te, di, se }

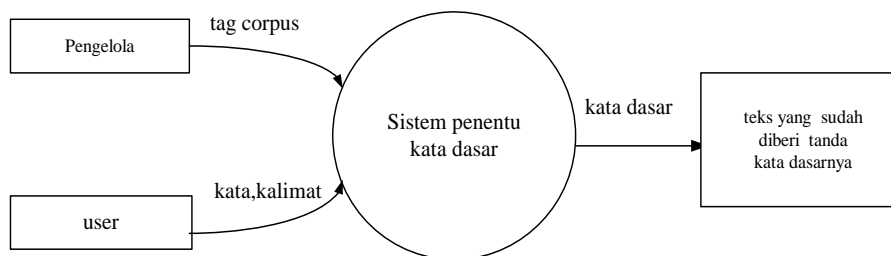
4. P (Daftar produksi)

- Morfem mem → apabila me → bertemu kata dasar dengan hurup awal ; b ,f , p maka berubah menjadi mem.
- Morfem men → apabila me → bertemu kata dasar dengan hurup awal : d , s , t maka berubah menjadi men.
- Morfem meny → apabila me → bertemu kata dasar dengan hurup awal : s maka berubah menjadi meny.
- Morfem meng → apabila me → bertemu kata dasar dengan hurup awal : g, h, k, x, a, e, i, o, u maka berubah menjadi meng.
- Morfem me → apabila me → bertemu kata dasar dengan hurup awal : l, r, w, y, n maka tidak berubah tetap me.
- Morfem menge → apabila me → bertemu dengan kata dasar yang bersuku kata tunggal maka berubah menjadi menge.
- Morfem pem → apabila pe → bertemu kata dasar dengan hurup awal ; b ,f , p maka berubah menjadi pem.
- Morfem pen → apabila pe → bertemu kata dasar dengan hurup awal : d , s , t maka berubah menjadi pen.
- Morfem peny → apabila pe → bertemu kata dasar dengan hurup awal : s maka berubah menjadi peny.
- Morfem peng → apabila pe → bertemu kata dasar dengan hurup awal : g, h, k, x, a, e, i, o, u maka berubah menjadi peng.
- Morfem pe → apabila pe → bertemu kata dasar dengan hurup awal : l, r, w, y, n maka tidak berubah tetap pe.
- Morfem penge → apabila pe → bertemu dengan kata dasar yang bersuku kata tunggal maka berubah menjadi penge.
- Morfem be → apabila ber → bertemu kata dasar dengan hurup awal : r dan kata dasar yang bersuku kata pertamanya berakhir r maka berubah menjadi be.
- Morfem bel → apabila ber → bertemu dengan kata dasar “ajar” maka berubah menjadi bel.
- Morfem ber → apabila ber → bertemu kata dasar dengan hurup awal : selain r dan kata dasar yang bersuku kata pertamanya tidak berakhir r serta bukan kata dasar “ajar” maka berubah menjadi ber.

- Morfem pe → apabila per → bertemu kata dasar dengan hurup awal : r dan kata dasar yang bersuku kata pertamanya berakhir r maka berubah menjadi pe.
- Morfem pel → apabila per → bertemu dengan kata dasar “ajar” maka berubah menjadi pel.
- Morfem per → apabila per → bertemu kata dasar dengan hurup awal : selain r dan kata dasar yang bersuku kata pertamanya tidak berakhir r serta bukan kata dasar “ajar” maka berubah menjadi per.
- Morfem te → apabila ter → bertemu kata dasar dengan hurup awal : r dan kata dasar yang bersuku kata pertamanya berakhir r maka berubah menjadi te.
- Morfem ter → apabila ter → bertemu kata dasar dengan hurup awal : selain r dan kata dasar yang bersuku kata pertamanya tidak berakhir r maka berubah menjadi ter.
- Morfem se → apabila se → bertemu kata dasar dengan hurup awal : a, b, c, d, e, f, g, h, i, j, k, l, m, n, o, p, q, r, s, t, u, v, w, x, y, z maka tidak berubah tetap se.
- Morfem di → apabila di → bertemu kata dasar dengan hurup awal : a, b, c, d, e, f, g, h, i, j, k, l, m, n, o, p, q, r, s, t, u, v, w, x, y, z. maka tidak berubah tetap di.
- Morfem ke → apabila ke → bertemu kata dasar dengan hurup awal : a, b, c, d, e, f, g, h, i, j, k, l, m, n, o, p, q, r, s, t, u, v, w, x, y, z. maka tidak berubah tetap ke.

4. Analisis Sistem

4.1. Context Diagram



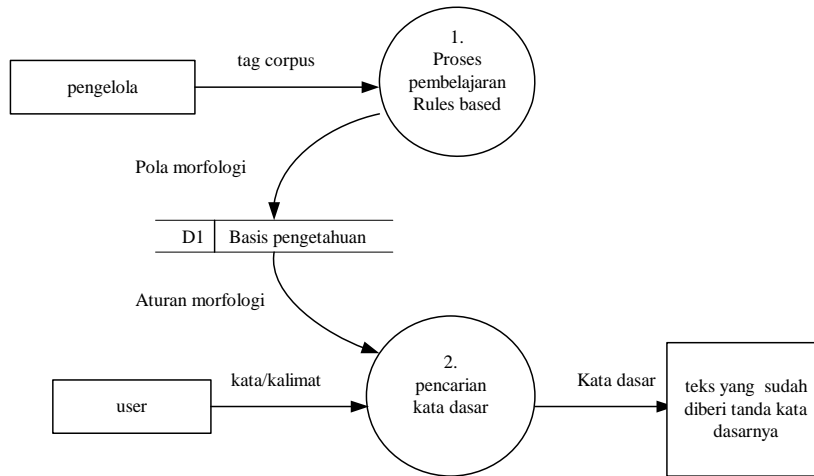
Gambar 3.46 Context Diagram

Keterangan :

Pengelola adalah pembuat sistem yang menyediakan kamus dalam bentuk tag corpus atau file teks yang sudah diberi tanda. Tag corpus ini akan dijadikan dasar pembelajaran sistem penentu kata dasar. Kemudian user memberikan input berupa kata atau kalimat yang

diketik dalam bahasa Indonesia kemudian system akan memproses kata dan kalimat tersebut dan menghasilkan file data kata dasar.

4.2. DFD Level 1 Sistem Penentu Kata Dasar



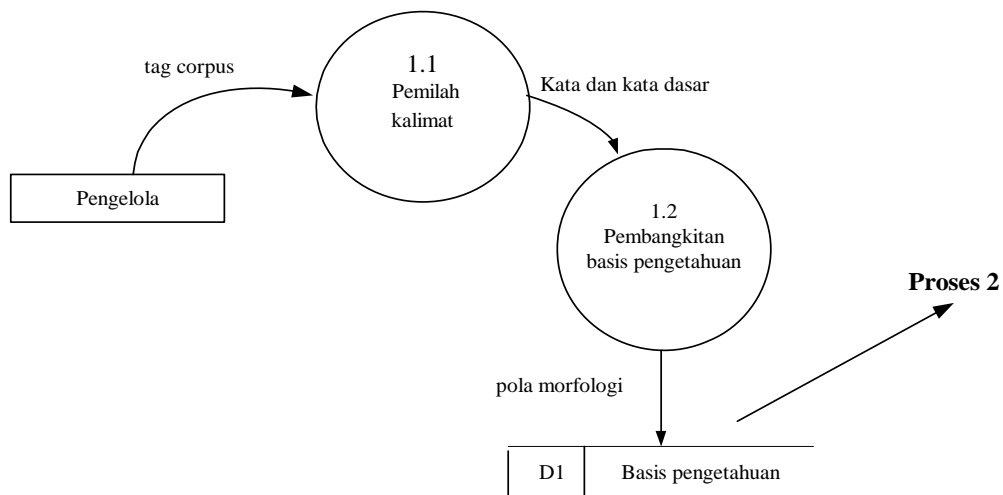
Gambar 4 DFD level 1 Sistem Informasi penentu kata dasar

Keterangan:

Dalam sistem penentu kata dasar ini terdiri dari dua proses :

- Proses pertama yaitu proses pembelajaran yang menentukan untuk mendapatkan aturan-aturan morfologi yang kemudian dihimpun dalam suatu basis pengetahuan. Proses ini membangun aturan-aturan perubahan atau pembentukan sebuah kata dalam bahasa Indonesia dengan merujuk pada sebuah kamus yang sebelumnya telah dibuat dalam bentuk tag corpus atau file teks yang sudah diberi tanda oleh pengelola.
- Proses kedua melakukan pencarian kata dasar, yang merespon inputan dari user berupa kata atau kalimat yang diketik kemudian diproses dengan cara membandingkan aturan pembentukan dan perubahan-perubahan kata yang sudah tersimpan dalam basis pengetahuan. Selanjutnya hasil pencarian ini akan ditampilkan dalam sebuah file teks yang berisi kata atau kalimat yang telah diberi penggalan kata-kata yang disertai dengan kata dasarnya.

4.3. DFD Level 2 Proses 1 (Pembelajaran Rules Based)

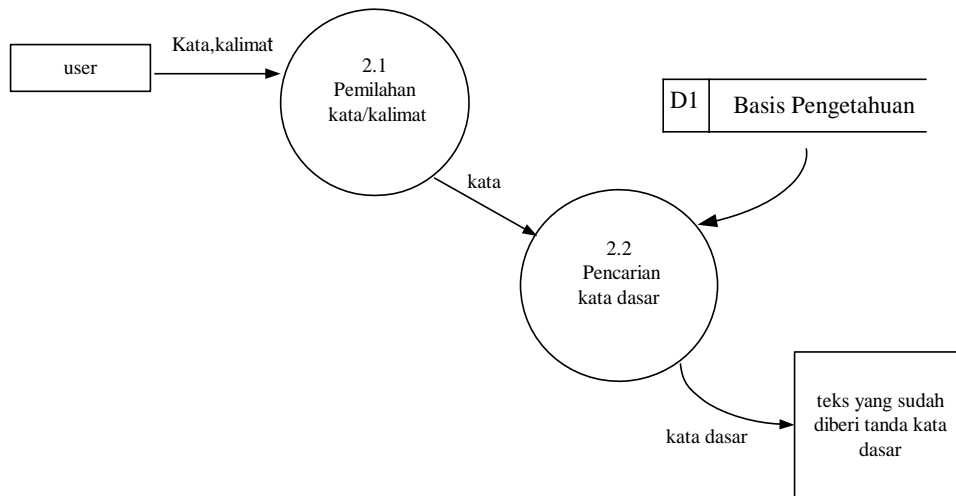


Gambar 5. DFD level 2 Proses pembelajaran Rules based

Keterangan :

Pengelola adalah pembuat sistem yang menyediakan kamus dalam bentuk tag corpus atau file teks yang sudah diberi tanda. Kemudian tag corpus ini dipilah dalam proses pemilah kata yang selanjutnya menghasilkan kata dan kata dasar yang sudah dipisahkan dengan tanda pemenggalan. Kata dan kata dasar ini dijadikan dasar pembentukan pembangkitan basis pengetahuan yang berisi aturan-aturan morfologi, yang akan dibuat setelah nantinya digabung dan dibandingkan dengan aturan rules based morfologi dalam bahasa Indonesia. Dari semua aturan pembentukan kata yang terdiri atas beberapa morfem akan dianalisa dan disimpan dalam knowledge base (basis pengetahuan). Selanjutnya basis pengetahuan ini akan diambil oleh proses 2 (Pencarian kata dasar).

4.4. DFD Level 2 Proses 2 (Pencarian Kata Dasar)



Gambar 6. DFD level 2 Proses pencarian kata dasar

Keterangan :

Kata atau kalimat yang diinputkan akan dipilah menjadi penggalan kata-kata yang membentuk sebuah kalimat, proses pemilahan ini dilakukan untuk mempermudah dalam upaya pemeriksaan antara kata asal dengan kata dasarnya. Setelah terbentuk data kata kemudian dilakukan proses pencarian kata dasar. Pada proses pencarian kata dasar ini dilakukan proses membandingkan data kata yang sudah terbentuk dengan aturan-aturan morfologi yang tersimpan di basis pengetahuan. Selanjutnya hasil pencarian aturan morfologi ini akan ditampilkan dalam sebuah teks kata yang sudah diberi tanda kata dasar.

5. Perancangan Algoritma

Setelah melalui tahap analisa sistem yang telah dibahas pada Bab III dan sebelum dibuatkan rancangan dari tampilan input dan output program maka selanjutnya dibuat perancangan algoritma. Perancangan algoritma ini akan dijadikan dasar pada saat pembuatan. Perancangan algoritma yang dibahas adalah algoritma pada proses yang terdapat dalam kamus data pada Bab III.

5.1 Pemilahan Kalimat

Pada proses pemilahan kalimat sesuai yang terdapat dalam aliran data ada dua macam. Kedua proses tersebut sama-sama memilah kalimat namun dalam konteks yang berbeda. Perbedaan proses tersebut dapat dilihat pada penjelasan algoritma dibawah ini :

Nama Proses : Pemilahan kalimat dari tag corpus

Algoritma :

Ambil kata sampai tanda //, masukan ke dalam sebuah variabel kata

Ambil 2 kata sampai tanda //, masukan ke dalam sebuah variabel kata dasar

Ulangi kedua proses tersebut sampai ketemu tanda titik

Nama Proses : Pemilahan kata dari file teks / kalimat

Algoritma :

Ambil kata sampai spasi, masukan ke dalam sebuah variabel

5.2 Penentuan Kata Dasar

Proses penentuan kata dasar berfungsi untuk mendapatkan hasil yang diinginkan. Penentuan kata dasar berdasarkan aturan pertama yang bekerja berdasarkan penandaan kata sesuai dengan kata yang ditemukannya, setelah itu baru dirubah sesuai dengan konteksnya.

Adapun algoritma proses tersebut adalah sebagai berikut :

Nama Proses : Penentuan kata dasar

Algoritma :

Ambil 1 kata sampai tanda spasi, simpan divariabel 1

Ambil 1 kata sesudahnya, simpan divariabel 2

Cari di basis pengetahuan yang sesuai dengan variabel yang telah disimpan

Ulangi kedua proses di atas sampai ketemu tanda titik

Jalankan prosedur pembangkitan basis pengetahuan

Pilih yang paling besar kemunculannya

Tampilkan hasilnya.

5.3 Pembangkitan Basis Pengetahuan

Proses pembangkitan pengetahuan berfungsi untuk mendapatkan basis pengetahuan dari data yang ada. Proses ini dibangkitkan dari kata yang dikirimkan untuk kata dasarnya

dan kemudian dilakukan proses analisa terhadap kata tersebut. Penjelasan algoritma dari proses ini adalah sebagai berikut :

Nama Proses : Pembangkitan basis pengetahuan

Algoritma :

Mem → if char1 = 'm' and

char2 = 'e' and

char3 = 'm' then 'me' U 'b','f','p'

Men → if char 1 = 'm' and

char2 = 'e' and

char3 = 'n' then 'me' U 'd','s','t'

Meny → if char1 = 'm' and

char2 = 'e' and

char3 = 'n' and

char4 = 'y' then 'me' U 's'

Meng → if char1 = 'm' and

char2 = 'e' and

char3 = 'n' and

char4 = 'g' then 'me' U 'g','h','k','x','a','i','u','o','e'

Menge → if char1 = 'm' and

char2 = 'e' and

char3 = 'n' and

char4 = 'g' and

char3 = 'e' then 'me' U 'sk1'

Me → if char1 = 'm' and

char2 = 'e' then 'me' U 'l','r','w','y','n'

Pem → if char1 = 'p' and

char2 = 'e' and

char3 = 'm' then 'pe' U 'b','f','p'

Pen → if char 1 = 'p' and

char2 = 'e' and

char3 = 'n' then 'pe' U 'd','s','t'

Peny → if char1 = 'p' and

char2 = 'e' and
char3 = 'n' and
char4 = 'y' then 'pe' U 's'

Peng → if char1 = 'p' and
char2 = 'e' and
char3 = 'n' and
char4 = 'g' then 'pe' U 'g','h','k','x','a','i','u','o','e'

Penge → if char1 = 'p' and
char2 = 'e' and
char3 = 'n' and
char4 = 'g' and
char3 = 'e' then 'pe' U 'sk1'

Pe → if char1 = 'p' and
char2 = 'e' then 'pe' U 'l','r','w','y','n'

Per → if char 1 = 'p' and
char2 = 'e' and
char3 = 'r' then 'per' U 'r', 'not ajar'

Pel → if char 1 = 'b' and
char2 = 'e' and
char3 = 'l' then per U 'ajar'

Pe → if char1 = 'p' and
char2 = 'e' then 'per' U 'r'

Ber → if char 1 = 'b' and
char2 = 'e' and
char3 = 'r' then 'ber' U 'r', 'not ajar'

Bel → if char 1 = 'b' and
char2 = 'e' and
char3 = 'l' then 'ber' U 'ajar'

Be → if char1 = 'b' and
char2 = 'e' then 'ber' U 'r'

Ter → if char 1 = 't' and
char2 = 'e' and

char3 = 'r' then 'ter' U 'r', 'sk1not r'

Te → if char1 = 't' and

char2 = 'e' then 'ter' U 'r', 'sk1= r'

Se → if char1 = 's' and

char2 = 'e' then 'se' U semua kata dasar

Di → if char1 = 'd' and

char2 = 'i' then 'di' U semua kata dasar

Ke → if char1 = 'k' and

char2 = 'e' then 'ke' U semua kata dasar

5.4. Rancangan Input / Output

Pada tahap ini akan dirancang tampilan input dan output program sistem penentu kata dasar. Rancangan input merupakan rancangan untuk inputan data yang dibuat, sedangkan rancangan output adalah rancangan tampilan keluaran setelah program dijalankan. Adapun perancangan input/output tersebut adalah sebagai berikut :



Gambar 7. Rancangan input /output

Keterangan :

1. Kotak (masukan kata/kalimat) rancangan tempat user memasukan inputan terhadap sistem berupa teks dalam bahasa Indonesia yang ditik.
2. Tombol tag corpus adalah rancangan yang akan memproses pengisian dan penambahan tag corpus yang selanjutnya akan disimpan pada basis pengetahuan.
3. Tombol mulai cari adalah rancangan yang akan memproses inputan teks kata/kalimat yang telah ditik oleh user dan hasilnya akan terlihat pada tampilan kotak hasil (file teks) dan file kata dasar serta kotak pembelajaran morfologi
4. Tombol buka adalah rancangan yang akan digunakan untuk membuka file
5. Tombol baru adalah rancangan yang akan digunakan untuk menambahkan inputan baru yang akan ditik atau diisi oleh user.
6. Tombol simpan adalah rancangan untuk menyimpan semua inputan dan output yang telah dimasukan dan ditampilkan oleh sistem setelah melalui proses sebelumnya.
7. Tombol hapus adalah rancangan untuk menghapus semua inputan dan output yang telah dimasukan dan ditampilkan oleh system setelah melalui proses sebelumnya. Sehingga hasilnya menjadi kosong kembali.
8. Tombol keluar adalah rancangan untuk keluar dari sistem
9. Kotak file kata dasar adalah rancangan output untuk menampilkan hasil proses pencarian kata dasar. Yang ditampilkan adalah semua kata dasar dari kata/kalimat (teks yang diinputkan).
10. Kotak hasil (file teks) adalah rancangan output untuk menampilkan kata beserta kata dasarnya yang ditandai dengan penggalan kata antara kata dan kata dasarnya.
11. Kotak pembelajaran morfologi adalah rancangan untuk menampilkan proses pembentukan kata yang tercakup dalam aturan morfologi

6. Kesimpulan

Berdasarkan hasil penelitian yang didapat dari analisa sistem penentuan kata dasar serta perancangan yang ada, maka penulis dapat menarik beberapa kesimpulan sebagai berikut :

1. Menentukan sebuah kata dasar perlu dikaji dan diteliti dari proses pembentukan kata itu sendiri, sebab sebuah kata dasar akan mengalami penambahan baik awalan, sisipan

dan akhiran yang tentunya dapat memberikan penekanan arti atau bahkan mengandung arti yang berbeda.

2. Morfologi adalah ilmu yang mempelajari seluk-beluk bentuk kata serta fungsi perubahan-perubahan kata itu, baik fungsi gramatik maupun fungsi semantik
3. Sistem berbasis pengetahuan penentu kata dasar membantu dalam mempelajari penerapan aturan-aturan yang terdapat dalam kaidah morfologi
4. Mengimplementasikan metode rules based supervised dalam sistem berbasis pengetahuan dengan basis pengetahuan morfologi. Salah satu metode yang digunakan dengan aplikasi membandingkan kondisi (fakta) dengan aturan produksi ilmu morfologi yang akhirnya menghasilkan penentuan kata dasar.
5. Dalam implementasinya tidak terlalu memberikan kesulitan kepada user sehingga sistem berbasis pengetahuan penentuan kata dasar akan mudah dioperasikan.

7. Daftar Pustaka

- [1] Brill, Eric, *Unsupervised Learning of Disambiguation Rules for Part of Speech Tagging*, Departemen of Computer Science, Johns Hopkins University.1995.
- [2] Jurafsky, Daniel & Martin, James H., *Speech and Language Processing, An Introduction to Natural Language Processing Computational Linguistics and Speech Recognition*, Prentice Hall, Upper Saddle River, New Jersey,2000
- [3] Tofani, M. Abi, & Nugroho, G.S., *Sari Kata Bahasa Indonesia Lengkap*, Kartika, Agustus 1995, Surabaya.