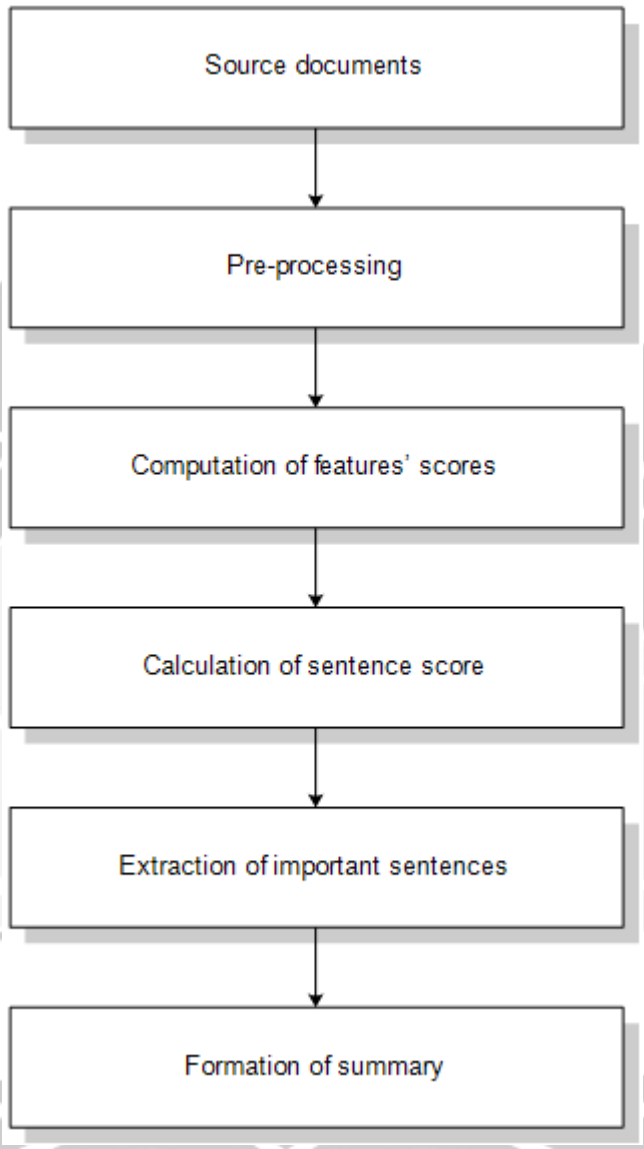
BAB III

LANDASAN TEORI

3.1 Peringkasan Teks Secara Otomatis

Sering kali kita memerlukan ringkasan dari sebuah dokumen untuk dapat memahami dengan cepat isi dari bacaan tersebut. Memahami isi bacaan melalui ringkasan teks memerlukan waktu yang lebih singkat dibandingkan membaca seluruh isi dokumen, sehingga ringkasan teks menjadi sangat penting. Selain dapat menghemat waktu, pembaca juga dapat menghindari pembacaan teks yang tidak relevan dengan informasi yang diharapkan oleh pembaca, terutama ketika sangat banyak informasi tersedia di internet.

Peringkasan teks adalah proses pemampatan teks sumber ke dalam versi lebih pendek namun tetap mempertahankan informasi pokok yang terkandung didalamnya (Alami, Meknassi, & Rais, 2015). Defenisi lain menurut (Fattah & Ren, 2009) adalah proses otomatis untuk menciptakan versi singkat dari sebuah teks atau dokumen dengan memilih bagian yang paling penting dan menghasilkan ringkasan yang relevan. Maka dapat dikatakan peringkasan teks secara otomatis (*automatic text summarization*) adalah proses menciptakan versi yang lebih singkat dari sebuah teks dengan memanfaatkan sistem yang dijalankan pada komputer, dimana pada hasil ringkasan tetap terkandung pikiran pokok dari teks sumber. Gambar 1 dibawah ini menunjukan alur peringkasan teks secara otomatis menurut (Gupta & Gambhir, 2016).



Gambar 1 Diagram alur peringkasan teks secara otomatis

Gambar 1 di atas menampilkan diagram blok sistem peringkasan teks otomatis berdasarkan pendekatan statistik. Pertama, preprocessing dokumen sumber dilakukan di mana teknik linguistik diterapkan yang mencakup segmentasi kalimat, penghapusan stop-words, penghapusan tanda baca, stemming, dan lain-lain. Proses segmentasi berhubungan dengan membagi teks menjadi kalimat. Kemudian, eliminasi stop-words dilakukan. Kata-kata yang sering terjadi

dalam teks namun tidak ada kontribusi dalam memilih kalimat penting, misalnya preposisi, artikel, kata ganti, dll disebut sebagai kata berhenti. Mereka dianggap sebagai istilah bising dalam teks. Jadi, pengangkatan kata-kata tersebut akan sangat membantu sebelum menjalankan pemrosesan alami. Kemudian, stemming dilakukan. Stemming adalah proses mengurangi kata-kata dengan akar yang sama atau kembali pada bentuk yang sama, sehingga menghilangkan akhiran variabel. Beberapa algoritma stemming populer dan efisien adalah Porter dan Lovins. Kemudian beberapa fitur dipilih yang akan membantu dalam ekstraksi kalimat penting. Fitur ini mungkin bersifat statistik atau linguistik atau kombinasi keduanya. Untuk setiap kalimat, semua nilai fitur yang dipilih dihitung dan kemudian ditambahkan bersama untuk mendapatkan skor setiap kalimat. Kalimat yang memiliki skor paling baik kemudian dipilih untuk membentuk ringkasan sambil menyimpan kalimat asli dalam dokumen. Panjang ringkasan tergantung pada tingkat kompresi yang diinginkan (Gupta & Gambhir, 2016).

3.1.1 Jenis Masukan

Berdasarkan jumlah dokumen masukan, *text summarization* dibedakan atas dua kategori yaitu *single-document* dimana dokumen sumber hanya terdiri dari satu dokumen sedangkan pada *multi-document summarization* merupakan peringkasan teks dengan dokumen sumber berjumlah lebih dari satu (Lloret & Palomar, 2012). Sedangkan model dalam peringkasan teks secara otomatis ada dua yaitu ringkasan yang umum (*generic summary*) adalah perwakilan dari teks asli yang mempresentasikan intisari dari sebuah teks asal. Dengan menggunakan pendekatan bottom up (*information retrieval*) dimana sebuah ringkasan

didasarkan pada spesifikasi kebutuhan informasi yang diberikan oleh pemakai (*query driven*) seperti topik atau query yang menggunakan pendekatan top down (*information extraction*).

3.1.2 Kriteria Peringkasan Teks

Ada dua kriteria dalam peringkasan teks secara otomatis yaitu peringkasan teks berdasarkan eksktraksi dan abstraksi. Teknik ekstraksi merupakan suatu teknik untuk menyalin bagian-bagian teks yang paling penting atau paling informatif dari teks sumber menjadi sebuah ringkasan, sedangkan teknik abstraksi adalah mengambil intisari dari teks sumber kemudian menghasilkan sebuah ringkasan dengan menciptakan kalimat baru yang merepresentasikan intisari teks sumber dalam bentuk berbeda (Gupta & Gambhir, 2016). Berdasarkan teori, hasil teknik peringkasan eksktraksi lebih baik dibandingkan dengan ringkasan abstraksi. Hal ini dikarenakan pada peringkasan abstraksi, seperti representasi semantik, inferens dan pembangun natural language relatif lebih sulit, dibandingkan pendekatan data driven, seperti ekstraksi kalimat. Oleh karena itu kebanyakan penelitian dilakukan menggunakan metode eksktraksi.

3.2 Stemming

Information Retrieval (IR) adalah suatu proses pemisahan dokumen-dokumen yang dianggap memiliki keterkaitan satu dengan yang lain dari sekumpulan dokumen yang ada. Salah satu proses awal yang terdapat dalam IR adalah Stemming. Stemming adalah proses mentransformasi kata-kata yang

terdapat dalam suatu dokumen ke kata-kata dasarnya (*root word*) dengan menggunakan beberapa aturan tertentu. Sebagai contoh, kata berdiri, pendirian, mendirikan, akan distem ke kata dasarnya yaitu "diri". Proses stemming pada teks bahasa Indonesia berbeda dengan stemming dalam teks bahasa Inggris. Pada teks berbahasa Inggris, proses yang diperlukan hanya proses menghilangkan kata akhiran atau sufiks. Sedangkan pada teks berbahasa Indonesia, selain akhiran, awalan, dan juga konfiks dihilangkan (Adriani, Asian, Nazief, Tahagohi, & Williams, 2007). Proses *stemming* dalam teks bahasa Indonesia lebih kompleks karena terdapat banyak sekali variasi imbuhan yang harus dihilangkan agar bisa mendapatkan kata dasar dari sebuah kata. Pada umumnya kata dasar dalam teks bahasa Indonesia terdiri dari kombinasi sebagai berikut:

Prefiks 1 + Prefiks 2 + Kata dasar + Sufiks 3 + Sufiks 2 + Sufiks 1

Salah satu algoritma stemming pada teks bahasa indonesia yang populer yaitu algoritma Nazief & Adriani yang dibuat oleh Bobby Nazief dan Mirna Adriani. Algoritma tersebeut memiliki tahapan sebagai berikut:

1. Langkah pertama adalah mencari kata yang akan diistem dalam kamus kata dasar. Jika ditemukan maka diasumsikan kata tersebut adalah sebuah kata dasar (*root word*) dan algoritma berhenti pada tahap ini.
2. Tahap selanjutnya adalah *Inflection Suffixes* ("-lah", "-kah", "-ku", "-mu", atau "-nya") dihilangkan. Jika kata tersebut berupa *particles* ("-lah", "-kah", "-tah" atau "-pun") maka langkah ini akan diulangi lagi untuk menghapus *Possesive Pronouns* ("-ku", "-mu", atau "-nya") jika ada.

3. Langkah berikutnya menghapus *Derivation Suffixes* (“-i”, “-an” atau “-kan”). Jika kata tersebut ditemukan dalam kamus, maka algoritma berhenti pada tahap ini. Jika tidak maka dilanjutkan ke langkah 3a
    a. Jika akhiran “-an” sudah dihapus dan huruf terakhir dari kata tersebut adalah “-k”, maka “-k” juga ikut dihapus. Jika kata tersebut ditemukan dalam kamus maka algoritma berhenti pada proses ini. Jika tidak ditemukan maka dilanjutkan pada langkah 3b.
    b. Akhiran yang dihapus (“-i”, “-an” atau “-kan”) dikembalikan ke kata asal, kemudian dilanjutkan ke langkah 4.
4. Menghapus *Derivation Prefix*. Jika pada langkah 3 ada akhiran yang dihapus maka lanjut ke langkah 4a, jika tidak maka lanjutkan ke langkah 4b.
    a. Lakukan pengecekan tabel kombinasi awalan-akhiran yang tidak diijinkan. Jika ditemukan maka algoritma berhenti pada tahap ini, jika tidak lanjutkan ke langkah 4b.
    b. For i = 1 to 3, tentukan tipe awalan lalu hapus awalan. Jika kata dasar belum juga ditemukan maka lanjutkan ke langkah 5, jika ditemukan maka algoritma berhenti pada tahap ini. Catatan: jika awalan pertama dan kedua sama, maka algoritma berhenti.
5. Lakukan perekaman data.
6. Jika semua proses sudah dilakukan tetapi masih tidak berhasil maka kata awal akan diasumsikan sebagai kata dasar. Proses selesai.

3.3 Sentence Scoring

Dalam peringkasan teks secara otomatis, pembobotan kalimat salah satu merupakan bagian penting. Pada 2013, Fereira dkk melakukan evaluasi terhadap 15 fitur pembobotan kalimat untuk menentukan fitur mana yang lebih optimal. Hasil penelitian menunjukan bahwa tiap fitur tersebut memiliki tingkat pengaruh yang berbeda-beda terhadap hasil ringkasan sistem. Selain itu, jumlah fitur yang banyak juga menyebabkan waktu komputasi lebih lama, sehingga perlu dipilih fitur-fitur yang dapat mempersingkat waktu sistem untuk menghitung bobot setiap kalimat (Ferreira, et al., 2013). Jika waktu penghitungan bobot tiap kalimat bisa ditekan menjadi lebih cepat, maka dapa berimplikasi pada total waktu peringkasan di setiap dokumen. Oleh karena itu, dalam penelitian ini akan digunakan 8 fitur pembobotan kalimat yang relevan dengan karakteristik dari sebuah teks berita. Fitur pembobotan yang akan digunakan dalam penelitian ini adalah TF/IDF, kalimat dengan huruf kapital, kalimat dengan kata benda, frasa penting, data angka, panjang kalimat, posisi kalimat, dan kemiripan kalimat dengan judul. Penjelasan dan rumus tiap-tiap fitur adalah sebagai berikut:

1. **TF/IDF (F1)**

   Pembobotan diperoleh berdasarkan jumlah kemunculan term atau kata dalam kalimat (TF) dan jumlah kemunculan term atau kata pada seluruh kalimat dalam dokumen (IDF).

$$TF(s,t) = \frac{term\ frequency(s,t)}{\max(term\ frequency(s,ti))} \quad .....(1)$$

$$Score\ f1(s,t) = TF(s,t) \times log\left(\frac{N}{sft}\right) \quad \text{.....(2)}$$

Dimana :

N = jumlah kalimat yang berisi term(t)

sft = jumlah kemunculan kata (term) terhadap kalimat

2. Kalimat Mengandung Huruf Kapital (F2)

Metode ini memberikan bobot yang lebih tinggi untuk kata-kata yang mengandung satu atau lebih huruf kapital, misalnya adalah nama orang, kota, negara, dan singkatan. Berikut rumus yang digunakan:

$$CW(s) = \frac{jumlah\ kata\ huruf\ besar\ dalam\ s}{jumlah\ kata\ dalam\ s} \quad \text{.....(3)}$$

$$Score\ f2(s) = \frac{CW(s)}{\max(CW(s))} \quad \text{.....(4)}$$

Dimana :

CW = Rasio total kata dengan huruf besar muncul dalam dokumen dengan jumlah semua kata

3. Kalimat Mengandung Kata Benda (F3)

Kalimat yang mengandung jumlah kata benda yang lebih banyak mendapat bobot lebih tinggi dan cenderung untuk dimasukkan dalam ringkasan dokumen. Kata benda seperti nama orang atau nama tempat.

$$Score\ f3(s) = \frac{jumlah\ kata\ benda\ dalam\ s}{jumlah\ kata\ dalam\ s} \quad \text{.....(5)}$$

4. Kalimat Mengandung Frasa Penting (f4)

Secara umum, kalimat dimulai dengan frasa "dengan demikian", "penyelidikan", dan menekankan seperti "yang terbaik", "paling penting", "menurut penelitian", "khususnya", serta frasa lainnya dapat menjadi indikator yang baik dari dokumen teks.

$$Score\ f4(s) = \frac{jumlah\ frasa\ dalam\ s}{jumlah\ frasa\ dalam\ dokumen} \quad .....(6)$$

5. Kalimat Mengandung Data Angka (F5)

Pada peringkasan teks mempertimbangkan data angka dalam dokumen, karena pada kalimat yang berisi data angka biasanya mengandung informasi yang penting.

$$Score\ f5(s) = \frac{data\ angka\ dalam\ s}{panjang\ s} \quad .....(7)$$

6. Panjang Kalimat (F6)

Kalimat yang panjang memiliki bobot yang lebih tinggi. Panjang kalimat dihitung berdasarkan jumlah kata dalam kalimat dikali dengan rata-rata panjang kalimat dalam dokumen.

$$Score\ f6(s) = panjang\ s * rerata\ panjang\ kalimat \quad .....(8)$$

7. Posisi Kalimat (F7)

Posisi kalimat adalah letak kalimat dalam sebuah paragraf. Asumsinya bahwa kalimat pertama pada tiap paragraf adalah kalimat yang paling penting.

$$Score\ f7(s) = \frac{posisi\ s\ dalam\ paragraf}{jumlah\ total\ kalimat\ dalam\ dokumen} \quad .....(9)$$

8. Kemiripan Kalimat Dengan Judul (F8)

Kalimat yang menyerupai judul dokumen adalah kata yang muncul dalam kalimat sama dengan kata yang ada dalam judul dokumen.

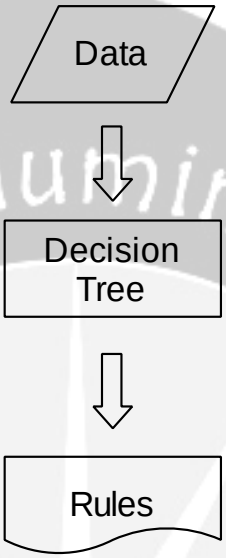
$$Score\ f8(s) = \frac{keyword\ dalam\ s \cap keyword\ dalam\ judul}{keyword\ dalam\ s \cup keyword\ dalam\ judul} \quad .....(10)$$

3.4 Decision Tree

Pohon Keputusan atau dikenal dengan *decision tree* adalah salah satu metode klasifikasi yang menggunakan representasi suatu struktur pohon, dimana setiap node pohon merepresentasikan atribut yang telah diuji, setiap cabang merupakan suatu pembagian hasil uji, dan node daun (*leaf*) merepresentasikan kelompok kelas tertentu (Wei & You, 2011). Struktur pohon ini menunjukkan faktor-faktor yang mempengaruhi hasil alternatif dari sebuah keputusan yang disertai dengan estimasi hasil akhir apabila kita mengambil keputusan tersebut.

Peranan pohon keputusan dalam hal ini berfungsi sebagai pendukung untuk membantu manusia dalam mengambil suatu keputusan. Manfaat dari pohon keputusan adalah untuk menjabarkan proses pengambilan keputusan yang

kompleks menjadi lebih sederhana sehingga pihak yang mengambil keputusan akan lebih menginterpretasikan solusi dari permasalahan.



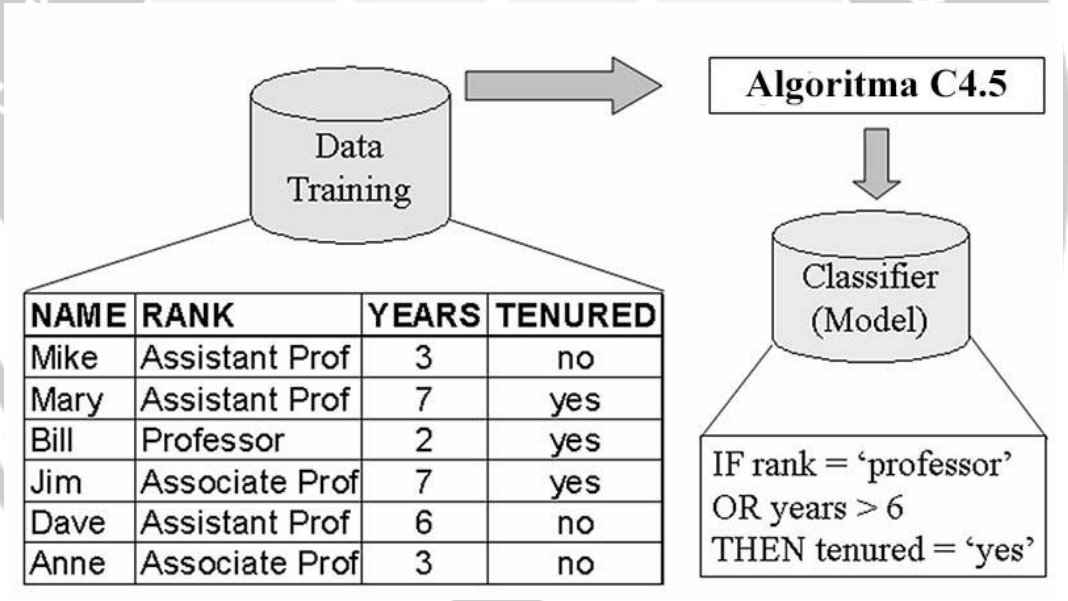
Gambar 2 Konsep Decision Tree

Gambar 2 diatas menunjukan proses dimulai dengan data acak yang diklasifikasi menjadi aturan-aturan keputusan (*rule*). Pada umumnya *decision tree* melakukan strategi pencarian secara *top-down* untuk solusinya. Pohon dibangun dengan cara membagi data secara rekursif sampai setiap bagian terdiri dari data yang berasal dari kelas yang sama (Hssina, Merbouha, Ezzikouri, & Erritali, 2014). Decision tree menggunakan sebuah struktur hierarki untuk pembelajaran supervised. Proses dari decision tree dimulai dengan root node hingga leaf node yang dilakukan secara rekursif. Di mana setiap percabangan menyatakan suatu kondisi yang harus dipenuhi dan pada setiap ujung pohon menyatakan kelas dari suatu data. Pada *decision tree,* struktur terdiri dari tiga bagian utama yaitu:

1. *Root node*, merupakan node yang terletak paling atas dari suatu pohon.

2. *Internal node*, merupakan node percabangan, dimana hanya terdapat satu masukan tetapi mempunyai minimal dua keluaran.
3. *Leaf node*, adalah node akhir, hanya memiliki satu masukan, dan tidak memiliki keluaran.

Algoritma C4.5 merupakan salah satu metode untuk membuat *decision tree* berdasarkan training data yang telah disediakan. Algoritma C4.5 juga dapat menangani data numerik (kontinyu) dan data diskret. Gambaran alur algoritma kerja C4.5 dapat dilihat sebagai berikut:



NAME	RANK	YEARS	TENURED
Mike	Assistant Prof	3	no
Mary	Assistant Prof	7	yes
Bill	Professor	2	yes
Jim	Associate Prof	7	yes
Dave	Assistant Prof	6	no
Anne	Associate Prof	3	no

Gambar 3 Model pohon keputusan dengan algoritma C4.5

Pada gambar 3 dapat dijelaskan proses pengolahan data training menggunakan algoritma C4.5 yang pada akhirnya menghasilkan aturan-aturan (*rules*). Perhitungan dalam algoritma C.45 menggunakan rasio perolehan (*gain ratio*). Sebelum menghitung gain ratio, perlu dilakukan perhitungan terlebih dahulu nilai informasi dalam satuan bits dari sebuah kumpulan objek.

Perhitungannya dilakukan dengan menggunakan konsep entropi. Secara matematis, untuk mendapatkan nilai *entropy* dapat dihitung menggunakan formula sebagai berikut :

$$Entropi(S) = \sum_{i=1}^{n} - pi * log_2\ pi \qquad .....(11)$$

Dimana S adalah himpunan kasus, n adalah jumlah nilai yang ada pada atribut target (jumlah kelas), sedangkan pi adalah jumlah sampel pada kelas i. Dari rumus diatas dapat kita perhatikan bahwa apabila hanya terdapat dua kelas dan dari kedua kelas tersebut mempunyai jumlah komposisi sampel yang sama, maka nilai entropi adalah nol. Ketika nilai entropi sudah didapatkan, maka langkah berikutnya adalah menghitung nilai *information gain*. Berdasarkan perhitungan matematis *information gain* dari suatu atribut A dapat diformulasikan sebagai berikut :

$$Gain(S, A) = Entropi(S) \sum_{i=1}^{n} \frac{|S_i|}{|S|} * Entropi(S_i) \qquad .....(12)$$

Dimana S adalah himpunan kasus, A adalah atribut, n adalah jumlah partisi atribut A, sedangkan i menyatakan suatu nilai yang mungkin untuk atribut A dan Si adalah jumlah kasus partisi ke i. Selanjutnya untuk menghitung rasio perolehan perlu diketahui suatu term baru yang disebut pemisahan informasi (*SplitInformation*). Pemisahan informasi dihitung dengan cara :

$$SplitInformation(S, A) = -\sum_{t=1}^{c} \frac{S_i}{S} log_2 \frac{S_i}{S} \qquad .....(13)$$

Dimana S 1 sampai S c adalah c subset yang dihasilkan dari pemecahan S dengan menggunakan atribut A yang mempunyai banyak C nilai. Selanjutnya rasio perolehan (*gain ratio*) dihitung dengan cara :

$$GainRatio(S,A) = \frac{Gain(S,A)}{SplitInformation(S,A)} \quad .....(14)$$

Dari formula perhitungan *gain ratio* diatas, dapat dijelaskan bahwa rasio perolehan didapat dari pembagian nilai gain dengan hasil pemisahan informasi (*SplitInformation*).

3.5 Bahasa Indonesia

Bahasa Indonesia merupakan salah satu bahasa yang mempunyai sejarah panjang dalam pembentukannya, baik lisan maupun pada tulisan. Bahasa Indonesia berasal dari bahasa melayu yang sudah berada di Nusantara sejak zaman kerajaan-kerajaan yang ada di Nusantara. Bahasa Indonesia mengalami perkembangan yang cukup pesat baik dalam bentuk lisan maupun tulisan sejak era penjajahan sampai era globalisasi pada saat ini. Berkembang dari ejaan Van Ophuijsen, Soewandi, Melindo sampai Ejaan Yang Disempurnakan yang kita pakai pada saat ini. Kegiatan-kegiatan yang berkaitan dengan bahasa Indonesia terus ditingkatkan seperti penelitian bahasa, seminar bahasa sampai dengan kongres bahasa Indonesia yang dilaksanakan setiap satu tahun sekali, hal ini membuktikan batapa pentingnya kedudukan dan fungsi bahasa di mata pemerintahan dan masyarakat Republik Indonesia (Nugroho, 2015).

Penggunaan bahasa Indonesia pada bentuk tulisan yang formal seperti dalam artikel ilmiah diharuskan mengikuti syarat-syarat tertentu. Pertama, secara morfologis bahasa pada artikel ilmiah harus lengkap. Dalam hal ini perwujudan setiap kata yang akan digunakan harus mengandung afiksasi yang lengkap contohnya: ditambahkan, mempertimbangkan, mempunyai dan lain sebagainya. Kedua, jika dilihat secara sintaksis bahasa dalam sebuah artikel ilmiah harus lengkap yaitu memuat unsur-unsur subjek, predikat, dan objek yang dinyatakan secara eksplisit. Banyak ditemukan dalam karya ilmiah bentuk penghilangan unsur subjek dalam kalimat yang kompleks padahal jika dilihat secara sintaksis subjek tersebut tidak memiliki rujukan yang sama dengan subjek pada kalimat induknya atau subjek kedua ini telah jauh terpisah dari subjek petamanya. Ketiga, bahasa dalam karya ilmiah harus tepat makna dan memiliki arti tunggal. Penulis karya ilmiah harus menimbang-nimbang secara seksama penggunaan setiap kata, ungkapan dan bentuk sintaksis sehingga apa yang dimengerti oleh pembaca sama dengan yang dimaksud oleh penulis.

Secara umum, setiap paragraf dalam penulisan artikel yang baik biasanya hanya memberikan satu gagasan utama. Terkait dengan metode Deduktif-Induktif dalam tata bahasa Indonesia, dimana kalimat pertama dan terakhir dalam sebuah paragraf biasanya dianggap sebagai kalimat utama yang mengekspresikan gagasan utama paragraf (Akhadiah, Arsyad, & Ridwan, 1985). Ada empat lokasi untuk memasukkan kalimat utama:

1. Kalimat utama pada awal paragraf (Deductive Paragraph)
2. Kalimat utama pada akhir paragraf (Inductive Paragraph)
3. Kalimat utama pada awal dan akhir paragraf (Deduktif-Induktif)
4. Tidak ada kalimat utama

Dengan demikian, kita harus mempertimbangkan untuk memberi nilai lebih tinggi pada kalimat pertama dan terakhir dari setiap paragraf dibandingkan dengan kalimat lainnya.

3.6 Berita

Berita (news) adalah laporan mengenai suatu peristiwa atau kejadian yang terbaru (aktual); laporan mengenai fakta-fakta yang aktual, menarik perhatian, dinilai penting, atau luar biasa (Djuroto, 2004).

3.6.1 Bagian-bagian Berita

Seperti tubuh manusia, sebuah dokumen berita juga memiliki bagian-bagian diantaranya adalah berikut ini:

1. *Headline*

   Merupakan bagian berita yang mewakili isi dari berita yang ingin disampaikan. Judul berita biasanya memiliki daya tarik yang kuat sehingga pembaca tertarik untuk membacanya.

2. *Dateline*

Adalah bagian yang terdiri dari nama media yang memberitakan, tempat dan tanggal kejadian. Tujuan bagian ini adalah untuk menunjukkan waktu, tempat kejadian dan inisial media.

3. *Lead* atau *Intro*

   Adalah bagian yang biasanya ditulis pada paragraf pertama pada sebuah berita. *Intro* merupakan bagian yang paling penting dari berita yang menentukan apakah isi berita menarik untuk dibaca atau tidak.

4. *Body*

   Merupakan bagian dari berita yang isinya menceritakan peristiwa yang dilaporkan dengan bahasa yang singkat, padat, dan jelas baik yang sudah dikemukakan dalam teras maupun yang belum diungkapkan.

3.6.2 Nilai Berita

Beberapa hal yang perlu diperhatikan pada saat menulis berita adalah terkait nilai berita itu sendiri (Djuroto, 2004). Beberapa nilai berita dikelompokkan agar menjadi acuan dalam sebuah penulisan berita. Nilai-nilai berita tesebut adalah sebagai berikut:

1. *Magnitude* (pengaruh)

   Artinya seberapa luas pengaruh suatu berita yang disampaikan terhadap masyarakat.

2. *Significant* (arti)

   Artinya seberapa penting arti dari suatu peristiwa bagi khalayak.

3. *Actuality* (aktualitas)

   Artinya seberapa besar tingkat aktualitas dari suatu peristiwa yang terjadi.

4. *Proximity* (kedekatan)

   Artinya sebuah bertia tentang peristiwa pada suatu daerah tertentu lebih pas jika diberitakan di daerah bersangkutan.

5. *Prominence* (keakraban)

   Artinya akrabnya suatu peristiwa yang diberitakan terhadap masyarakat.

6. *Surprise* (kejutan)

   Artinya tingkat pengraruh terhadap pembaca berita.

7. *Clarity* (kejelasan)

   Kejadian atau peristiwa yang diberitakan.

8. *Impact* (dampak)

   Artinya apakah berita tersebut berdampak terhadap khalayak.

9. *Conflict* (konflik)

   Artinya apakah apa diberitakan dapat menimbulkan konflik.

10. *Human Interest* (menarik)

    Artinya kemampuan suatu peristiwa menyentuh perasaan khalayak.

3.6.3 Unsur-unsur Berita

Dalam penulisan sebuah berita, harus dipahami unsur dari suatu berita agar memberi kemudahan bagi kita dalam mendeskripsikan berita tersebut dan berita yang dibuat supaya mudah dipahami oleh pembaca (Olii, 2007). Unsur-unsur tersebut biasa dikenal dengan istilah 5W1H adalah sebagai berikut:

1. What (apa) artinya Peristiwa apa yang tengah terjadi.
2. Who (siapa) artinya siapa saja yang terlibat dalam peristiwa itu.
3. Where (dimana) artinya dimana lokasi terjadinya peristiwa itu.
4. When (kapan) artinya kapan perisitiwa itu berlangsung.
5. Why (mengapa) artinya mengapa kejadian itu bisa terjadi.
6. How (bagaimana) artinya bagaimana kejadian itu bisa berlangsung.