BAB I
PENDAHULUAN

I.1 Latar Belakang Masalah

Text Summarization adalah suatu proses untuk menghasilkan ringkasan dari suatu artikel, tetapi tetap memiliki gambaran atau dengan kata lain tidak menghilangkan inti informasi dari artikel. Ada beberapa metode untuk melakukan peringkasan teks, dari jurnal teknik informatika dengan judul "Analisis Perbandingan Peringkasan Teks Menggunakan Metode *Lexical chain* dan Metode *TF-IDF* dengan Ekstraksi Frase Utama" yang menyatakan dalam hasil analisis perbandingan dua metode tersebut bahwa Metode *Lexical chain* yang memiliki performansi lebih baik untuk peringkasan teks *single document* karena adanya pembentukan rantai leksikal pada setiap kalimat sehingga bobot kalimat untuk di ekstraksi ditentukan oleh rantai terkuat tidak tergantung pada judul, namun penggunaan Metode *Lexical chain* ini masih memiliki kelemahan yaitu tidak bisa mengidentifikasi kata yang berambigu dalam pembentukan rantai leksikal. Dimana rantai leksikal tersebut memiliki peran penting dalam proses selanjutnya yaitu menentukan bobot kalimat sebelum menghasilkan ringkasan. Maka dari itu jika terjadi ambiguitas dalam pembentukan rantai leksikal akan sangat berpengaruh kepada pembobotan kalimat dan hasil ringkasan yang dihasilkan. Dalam penelitian tersebut disarankan agar Metode *Lexical chain* dilengkapi dengan *Word sense disambiguation* [1].

Berdasarkan uraian tersebut, maka dalam penelitian tugas akhir ini akan memperbaiki kelemahan dalam penelitian sebelumnya yaitu menghilangkan ambiguitas dengan mengidentifikasi makna kata sebelum pembentukan rantai leksikal dan untuk mengetahui pengaruhnya terhadap hasil ringkasan. Dimana metode yang akan digunakan adalah Metode *Lexical chain* dengan *Word sense disambiguation* untuk peringkasan artikel . Metode *Lexical chain* adalah salah satu metode untuk meringkas teks yang membentuk rantai leksikal berdasarkan hubungan semantik antar kata dalam teks [2]. Sedangkan *Word sense*

disambiguation adalah suatu proses mengidentifikasi makna kata yang digunakan dalam kalimat tertentu ketika satu kata memiliki sejumlah makna yang berbeda [3]. Dimana dalam *Word sense disambiguation* algoritma yang akan digunakan adalah algoritma *Lesk*. Algoritma *Lesk* adalah algoritma yang bekerja untuk mengetahui makna dari kata yang memiliki makna ganda atau lebih dari satu dalam kalimat dengan memperhatikan kata yang ada disekitarnya [4].

Dalam penelitian ini artikel yang akan digunakan dipilih berdasarkan hasil dari kueisioner yang telah disebar untuk mengetahui artikel yang paling banyak disukai, hasil kueisioner dari 90 responden yang mengisi kuesioner didapat 56 repsonden menyukai artikel kesehatan, 17 responden menyukai artikel olahraga, 11 responden menyukai artikel pendidikan, dan 6 orang menyukai artikel politik. Oleh karena itu dalam penelitian ini Implementasi Metode *Lexical chain* dengan *Word sense disambiguation* untuk peringkasan artikel kesehatan. Artikel Kesehatan adalah suatu karangan faktual dengan pembahasan mengenai kesehatan secara lengkap dengan panjang tertentu, dibuat untuk dipublikasikan yang bertujuan untuk menyampaikan gagasan atau fakta menjadi sebuah infromasi.

I.2 Perumusan Masalah

Berdasarkan latar belakang yang telah dijelaskan, maka dirumuskan masalah pada penelitian ini adalah sebagai berikut:

1. Bagaimana membangun sistem peringkasan artikel kesehatan dengan mengimplementasikan Metode *Lexical chain* dengan *Word sense disambiguation.*
2. Bagaimana memperbaiki kelemahan penelitian sebelumnya menggunakan Metode *Lexical chain* dengan *Word sense disambiguation* untuk peringkasan artikel kesehatan.

I.3 Maksud dan Tujuan

Maksud dari penelitian ini yaitu membangun sistem yang mengimplementasikan Metode *Lexical chain* dengan *Word sense disambiguation* untuk peringkasan artikel kesehatan. Tujuan yang ingin dicapai dari penelitian ini adalah memperbaiki kelemahan pada penelitian sebelumnya yaitu menghilangkan

ambiguitas dengan mengidentifikasi makna kata sebelum pembentukan rantai leksikal dan untuk mengetahui pengaruh Metode *Lexical chain* dengan *Word sense disambiguation* terhadap hasil ringkasan pada peringkasan artikel kesehatan.

I.4 Batasan Masalah

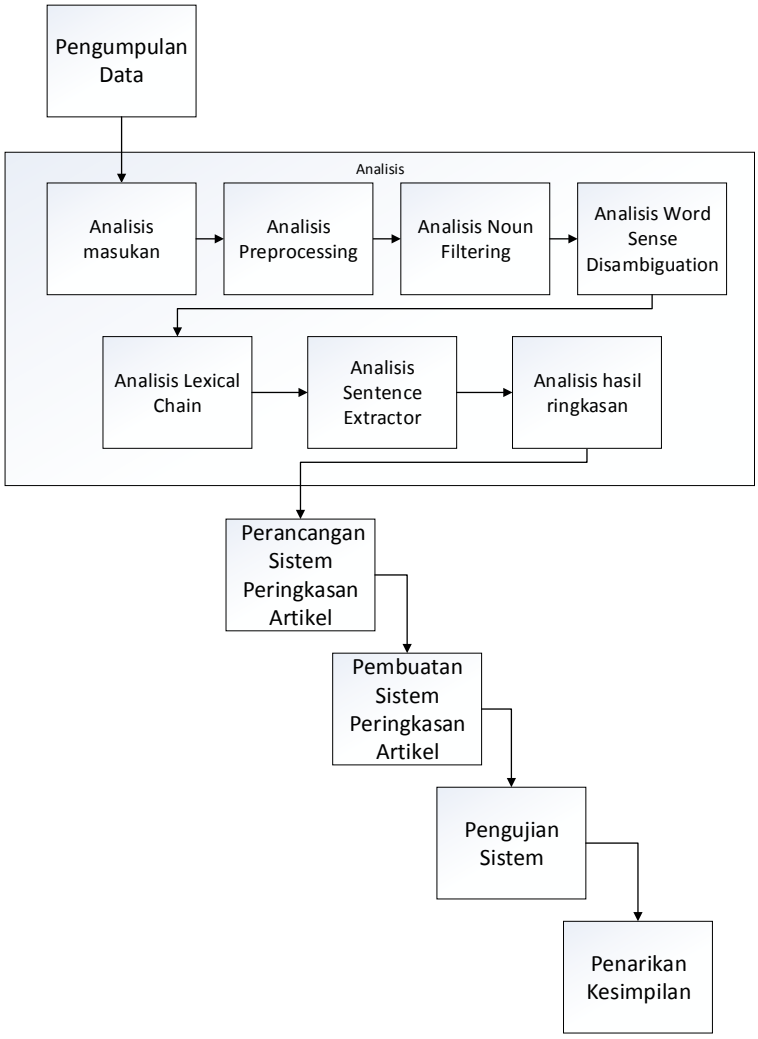
Berikut batasan masalah yang terdapat pada penelitian ini:

a. Data yang dimasukkan hanya berupa teks .txt atau .doc.
b. Data yang dimasukkan hanya artikel kesehatan dengan bahasa indonesia.
c. Artikel yang digunakan dari koleksi artikeltentangkesehatan.com.
d. Kamus makna dari Kamus Besar Bahasa Indonesia (KBBI)
e. Peringkasana hanya pada *single document*
f. *Range* hasil ringkasan 10% - 50% dari jumlah kalimat pada artikel

I.5 Metodologi Penelitian

Metodologi penelitian merupakan metode yang digunakan untuk pengumpulan data yang berhubungan dengan penelitian serta langkah yang dilakukan pada penelitian. Metode penelitian yang digunakan adalah metode deskriptif. Metode deskriptif merupakan metode penelitian yang menggambarkan suatu informasi, peristiwa, kejadian yang sedang terjadi saat sekarang secara sistematis, faktual dan akurat untuk mendapatkan gambaran yang jelas tentang hal-hal yang diperlukan [5]. Tahap-tahap penelitian dapat dilihat pada gambar I.1.

.



Gambar I.1 Tahap-tahap penelitian

1. Pengumpulan data

   Metode pengumpulan data digunakan untuk mengumpulkan data yang terkait dengan penelitian. Metode pengumpulan data yang dilakukan pada penelitian ini adalah Kueisioner dan Studi literatur. Kueisioner dilakukan untuk mendapatkan infromasi artikel yang paling diminati oleh masyarakat. Sedangkan Studi literatur ini dilakukan dengan cara mempelajari jurnal-jurnal ilmiah, artikel-artikel dan buku-buku yang berkaitan dengan *Text Summarization,* Metode *Lexical chain, Word sense disambiguation, Metode Lexical chain* dengan *Word sense disambiguation,* dan Algoritma *Lesk.*

2. Analisis

   Dalam tahap ini mempelajari konsep dan menganalisis mulai dari analisis masukan, analisis *preprocessing*, analisis *noun filtering*, analisis *text Summarization,* analisis *word sense disambiguation*, analisis l*exical chain,* analisis *sentence extractor,* sampai analisis hasil ringkasan.

3. Melakukan perancangan sistem peringkasan artikel

   Pada tahap ini melakukan perancangan sistem peringkasan artikel kesehatan yang akan dibuat sesuai dengan analisis yang telah dilakukan pada tahap sebelumya yaitu merancang masukan berupa artikel kesehatan dengan format .txt atau .doc, merancang proses dari metode *lexical chain* dengan *word sense disambiguation* mulai dari *preprocessing*, proses *noun filtering*, *word sense disambiguation*, *lexical chain* sampai dengan *sentence extractor*, kemudian merancang keluaran dari sistem.

4. Pembuatan sistem peringkasan artikel

   Pada tahap ini yaitu menerapkan atau mengimplementasikan metode *lexical chain* dengan *word sense disambiguation* ke dalam sistem peringkasan artikel.

5. Melakukan pengujian terhadap sistem

   Dalam tahapan ini yaitu melakukan pengujian dengan memasukan data uji berupa artikel kesehatan .txt dan menganalisis apakah keluaran dari sistem sudah sesuai dengan tujuan.

6. Penarikan Kesimpulan

   Pada tahap ini dilakukan penarikan kesimpulan dari hasil pengujian untuk mengetahui hasil dari penelitian yang telah dilakukan.

I.6 Sistematika Penulisan

Berikut adalah sistematika penulisan yang digunakan dalam penelitian ini:

BAB I PENDAHULUAN

BAB I menjelaskan mengenai latar belakang masalah penelitian, perumusan masalah, maksud dan tujuan dilakukannya penelitian, batasan pada penelitian, metodelogi penelitian baik itu metode pengumpulan data serta metode penyelesaian masalah dan sistematika penulisan laporan penelitian.

BAB II LANDASAN TEORI

BAB II mendeskripsikan teorI.teori yang berhubungan dengan penelitian ini untuk diimplementasikan pada penelitian seperti konsep dan teorI.teori mengenai *Text Summarization,* Metode *Lexical chain, Word sense disambiguation, Metode Lexical chain* dengan *Word sense disambiguation,* dan Algoritma *Lesk* .

BAB III ANALISIS DAN PERANCANGAN

BAB III mendeskripsikan analisis terhadap objek penelitian yang selanjutnya dilakukan perancangan perangkat lunak yang menjadi solusi dari objek penelitian tersebut. Dimana analisis yang dilakukan berupa analisis sistem, analisis kebutuhan fungsional dan analisis kebutuhan non fungsonal. Sedangkan perancangan yang dilakukan berupa perancangan data dan perancangan antarmuka untuk sistem yang akan dibangun.

BAB IV IMPLEMENTASI DAN PENGUJIAN

BAB IV mendeskripsikan implementasi pembangunan perangkat lunak yang sesuai berdasarkan hasil dari perancangan yang telah dijelaskan, setelah implementasi selesai, selanjutnya dilakukan pengujian terhadap sistem yang telah dibangun.

BAB V KESIMPULAN DAN SARAN

BAB V menjelaskan kesimpulan yang didapat dari penelitian yang dilakukan apakah telah berhasil memenuhi tujuan yang ditetapkan serta saran agar penelitian ini dapat dikembangkan lebih lanjut kedepannya dan dilakukan perbaikan untuk menutupi kekurangan dari penelitian ini.