

# 1. Pendahuluan

## 1.1. Latar Belakang

Perkembangan informasi saat ini menyebabkan pembaca sukar menemukan dan menyaring informasi yang sesuai dalam suatu teks. Suatu teks dengan ukuran yang panjang dapat menyulitkan pembaca dalam membaca dan menyerap semua informasi dari teks tersebut. Sedangkan informasi dari suatu teks merupakan komponen yang penting dari keseluruhan isi teks. Permasalahan itu timbul karena informasi yang didapat sulit untuk dipahami dan tidak mudahnya mendapatkan intisari dari suatu informasi. Peringkasan teks dapat digunakan untuk mempermudah menemukan informasi yang relevan tanpa menghilangkan informasi penting dari teks tersebut.

Peringkasan teks adalah proses mengurangi dokumen teks dengan komputer untuk menciptakan sebuah ringkasan yang mempertahankan poin penting dari dokumen asli[1]. Dalam peringkasan teks tersebut akan menimbulkan permasalahan seperti ambiguitas kata dan redundansi yang mempengaruhi kualitas dari hasil peringkasan teks tersebut. Ambiguitas kata merupakan permasalahan yang memiliki sejumlah makna kata yang berbeda. Sedangkan redundansi adalah permasalahan yang muncul pada sejumlah kalimat yang berulang atau ganda. Sehingga dibutuhkan metode untuk mengurangi tingkat redundansi dan ambiguitas dari peringkasan teks tersebut.

*Word sense disambiguation* adalah suatu proses mengidentifikasi makna kata yang digunakan dalam kalimat tertentu ketika kata memiliki sejumlah makna yang berbeda[2]. Masalah ambiguitas kata pada proses peringkasan teks inilah yang akan diatasi, sehingga dibutuhkan proses untuk mengurangi masalah ambiguitas dengan menerapkan *word sense disambiguation* pada peringkasan teks.

Metode yang digunakan untuk menghasilkan ringkasan teks dengan redundansi minimum yaitu, *maximal marginal relevance*. Metode ini menggunakan teknik ekstraksi yang digunakan untuk mengurangi redundansi kalimat dengan cara menghitung *similarity* antar kalimat dan kalimat dengan kalimat lain yang terpilih sebagai ringkasan[3].

Tugas akhir ini merupakan pengembangan dari tugas akhir yang telah ada dengan menggunakan metode *maximal marginal relevance*. Pada tugas akhir sebelumnya, metode MMR digunakan untuk dapat mengurangi redundansi dan akan dikembangkan dengan menerapkan *word sense disambiguation* dengan algoritma *lesk* pada peringkasan teks untuk mengurangi masalah ambiguitas. Untuk mengetahui kualitas dari hasil peringkasan teks tersebut akan dievaluasi menggunakan pengukuran *F-Measure* serta dilihat dari hasil *survey* pembaca terhadap hasil peringkasan teks dengan memberikan nilai. Dan melakukan analisis dengan membandingkan hasil peringkasan teks dengan menggunakan WSD dan hasil peringkasan teks tanpa menggunakan WSD .

## 1.2. Perumusan Masalah

Berdasarkan pada latar belakang diatas, rumusan masalah yang akan diuraikan dan diteliti adalah sebagai berikut:

1. Bagaimana implementasi dari *Word Sense Disambiguation* pada peringkasan teks dalam menangani masalah ambiguitas?
2. Bagaimana implementasi metode *Maximal Marginal Relevance* pada peringkasan teks dalam mengurangi redundansi?

3. Bagaimana mengukur performansi sistem peringkasan teks?
4. Bagaimana hasil peringkasan teks dengan menggunakan *Maximal Marginal Relevance* dan *Word Sense Disambiguation*?

Dalam penelitian Tugas Akhir ini, terdapat beberapa batasan masalah dalam pembahasan dan pengerjaannya, diantaranya :

1. Data yang digunakan adalah single dokumen berbahasa Inggris.
2. Menggunakan dataset *CNN Corpus* artikel berita dengan format *STORY*.
3. Metode peringkasan yang digunakan adalah peringkasan ekstraksi.
4. Menggunakan *library* pada proses *word sense disambiguation*.
5. Evaluasi dilakukan dengan menggunakan *ROUGE evaluation toolkit*.
6. Menggunakan algoritma *Cosine Lesk* untuk *Word Sense Disambiguation*.
7. Nilai *compression rate* yang digunakan 10%-50%.
8. Nilai parameter  $\lambda$  yang digunakan 0.1, 0.3, 0.5, 0.7, 0.9.

### 1.3. Tujuan

Adapun tujuan yang ingin dicapai dalam pengerjaan Tugas Akhir ini adalah sebagai berikut:

1. Mengimplementasikan *Word Sense Disambiguation* pada peringkasan teks dalam menangani masalah ambiguitas.
2. Mengimplementasikan Metode *Maximal Marginal Relevance* pada peringkasan teks dalam mengurangi redundansi.
3. Mengukur performansi sistem yang dibangun dengan menggunakan *ROUGE evaluation toolkit* dengan mengukur *precision*, *recall*, dan *F-Measure*.
4. Melakukan analisis terhadap hasil sistem peringkasan teks.

### 1.4. Metodologi Penyelesaian Masalah

Rencana kegiatan yang akan dilaksanakan dalam Tugas Akhir ini adalah sebagai berikut:

1. Kajian Pustaka  
Pada tahap ini, akan dilakukan pencarian dan pengumpulan berbagai macam referensi tentang Analisis *Summarization*, *Word Sense Disambiguation*, dan *Maximal Marginal Relevance* yang terdapat pada jurnal, paper, atau mempelajari materi tersebut.
2. Pengumpulan dan Analisis Data  
Pada tahap ini bertujuan untuk lebih mendalami materi yang telah didapatkan pada tahap sebelumnya yaitu tahap Kajian pustaka. Pada tahap ini juga akan dilakukan analisis permasalahan yang dijadikan topik penelitian dan analisa data yang akan digunakan pada penelitian ini. Hasil dari tahap ini yaitu telah ditentukan metode yang akan digunakan dalam penyelesaian masalah yang diangkat serta mendapatkan data yang akan digunakan pada penelitian ini, data yang digunakan adalah dokumen tunggal artikel berita bahasa inggris *CNN Corpus* dengan format *STORY*.
3. Analisis dan Perancangan Sistem  
Pada tahap ini, akan dilakukan perancangan sistem dengan membuat flowchart dari setiap tahapan yang dilakukan, dengan membuat perancangan sistem ini akan mudah memahami alur yang terjadi didalam sistem yang akan dibangun.
4. Implementasi Sistem  
Pada tahap ini, akan dilakukan implementasi dari setiap alur yang terdapat pada perancangan sistem yang telah dibuat pada tahap sebelumnya, implementasi pada Tugas akhir ini dibangun menggunakan bahasa pemrograman python dengan tools *JetBrains Pycharm Community Edition 2016.3.1*.

5. Analisis dan Pengujian  
Pada tahap ini, akan dilakukan pengujian terhadap sistem yang telah dibangun untuk mengetahui hasil implementasi yang telah dilakukan serta menganalisis dari hasil pengujian yang dilakukan.
6. Pembuatan Laporan Tugas Akhir  
Pada tahap ini, akan dilakukan penyusunan laporan dan pengumpulan dokumentasi berdasarkan hasil dari penelitian.

## **1.5. Sistematika Penulisan**

Penulisan Tugas Akhir ini akan dibagi menjadi 5 bab yaitu:

- 3.1. Bab 1 : Pendahuluan  
Pada Bab ini akan membahas latar belakang, Tujuan, Batasan masalah, hipotesa, metodologi penyelesaian masalah, serta sistematika penulisan dari Tugas Akhir yang dibangun.
- 3.2. Bab 2 : Studi Literatur  
Pada Bab ini akan membahas teori-teori yang terkait dengan pengerjaan tugas akhir. Teori-teori yang akan dibahas yaitu Peringkasan Teks, *Maximal Marginal Relevance*, *Preprocessing*, *Cosine Similarity*, TF-IDF, Evaluasi, WordNet, *Part of Speech Tagging*, dan *Word Sense Disambiguation*.
- 3.3. Bab 3 : Perancangan Sistem  
Pada Bab ini membahas Perancangan Sistem untuk membangun sistem yang meringkas teks dengan mengurangi redundansi dan meminimalisir ambiguitas pada kata. Pada bab ini juga dibahas alur dari setiap proses yang akan dilalui beserta contoh dari proses yang dilalui serta Skenario Pengujian yang digunakan pada bab 4.
4. Bab 4 : Pengujian dan Analisis  
Pada Bab ini membahas hasil pengujian berdasarkan skenario pengujian yang dituliskan pada bab Perancangan Sistem. Selain itu, pada bab ini juga dijelaskan analisis terhadap hasil pengujian tersebut. Hasil dari kegiatan analisis ini menjadi dasar pengambilan kesimpulan.
5. Bab 5 : Kesimpulan dan Saran  
Pada Bab ini berisi tentang kesimpulan beserta saran dari keseluruhan hasil pengerjaan Tugas Akhir .