

1. Pendahuluan

1.1 Latar Belakang Masalah

Mengevaluasi ringkasan merupakan hal yang paling penting dalam perkembangan *Automatic Text Summarization* (ATS) [15]. Dewasa ini, telah banyak penelitian yang dilakukan untuk mengotomatisasi proses mengevaluasi ringkasan tersebut. Salah satu teknik mengevaluasi ringkasan yang telah menunjukkan keterhubungan yang baik dengan manusia adalah ROUGE.

ROUGE (*Recall-Oriented Understudy for Gisting Evaluation*) merupakan suatu teknik yang digunakan untuk menentukan kualitas dari mesin peringkasan otomatis, dengan cara membandingkan ringkasan yang dihasilkan oleh mesin peringkasan teks otomatis (*candidate summary*) dengan ringkasan (ideal) lain yang dibuat oleh manusia (*references summary*) [5].

ROUGE mengadopsi beberapa macam algoritma yang dapat digunakan untuk mencari kesamaan kata yang terdapat diantara ringkasan referensi dan ringkasan kandidat. Dua diantaranya adalah algoritma *Longest Common Subsequences* (LCS) dan algoritma *Skip Bigram Co Occurences Statistics* (SB). Berdasarkan [14], kedua algoritma ini memiliki kesamaan dalam proses pencocokannya, yaitu: tidak membutuhkan pencocokan yang berurutan, hanya menghitung kata-kata yang sama namun tetap mencerminkan urutan kata yang terdapat dalam kalimat tanpa harus memberikan standar panjang yang diperlukan.

Longest Common Subsequences (LCS) merupakan algoritma yang digunakan dalam teknik ROUGE, yang lebih dikenal sebagai ROUGE-L. Dimana metode ini memandang kalimat-kalimat penyusun ringkasan sebagai rangkaian kata-kata. Dan dalam pencocokannya kata-kata yang terdapat pada *candidate summary* dapat langsung dicocokkan dengan *references summary*. Namun, sesuai dengan namanya, algoritma LCS hanya menghitung kemiripan kata-kata yang terdapat pada rangkaian terpanjang saja, dan mengabaikan kata-kata yang mirip namun tidak berada pada rangkaian tersebut

Skip-Bigram Co-Occurrence Statistics (SB) merupakan algoritma yang digunakan dalam teknik ROUGE, yang lebih dikenal sebagai ROUGE-S. Dalam pencocokannya metode ini menggunakan seluruh pasangan kata-kata yang terdapat pada kalimat-kalimat yang ada pada *candidate summary* dengan *references summary*. Apabila dibandingkan dengan algoritma LCS, algoritma SB dapat menghitung semua urutan kata yang memiliki kesamaan, walaupun tidak sesensitif algoritma LCS.

Dengan adanya kelebihan dan kelemahan yang terdapat pada algoritma LCS dan SB dalam mengevaluasi mesin peringkasan teks otomatis, maka berdasarkan [14] untuk memberikan nilai tambah terhadap kata-kata yang mirip namun tidak berada pada rangkaian LCS, dapat digunakan perhitungan dengan algoritma SB. Dengan harapan, melalui algoritma yang baru ini (algoritma LCS+SB), mampu menghasilkan suatu cara evaluasi yang memiliki tingkat keterhubungan yang baik dengan *human judgement*, karena informasi yang terambil menjadi semakin banyak.

1.2 Perumusan Masalah

Berdasarkan latar belakang tersebut, maka permasalahan yang dibahas adalah sebagai berikut:

1. Bagaimana cara mengimplementasikan algoritma LCS+SB dalam teknik ROUGE untuk mengevaluasi mesin peringkasan teks otomatis.
2. Apakah algoritma LCS+SB dalam teknik ROUGE mampu menghasilkan metode yang lebih baik dalam mengevaluasi mesin peringkasan teks otomatis, dengan melihat korelasinya terhadap penilaian manusia.
3. Faktor-faktor apa saja yang dapat mempengaruhi korelasi antara algoritma LCS+SB terhadap penilaian manusia seperti persentase ekstraksi dan jumlah referensi.

Hipotesa awal dari penelitian ini adalah algoritma LCS+SB dalam teknik ROUGE dapat menghasilkan metode evaluasi yang lebih baik dalam mengevaluasi mesin peringkasan teks otomatis.

Batasan masalah untuk penelitian ini adalah:

1. Dokumen yang dievaluasi adalah artikel berita berbahasa Indonesia dan Inggris.
2. Mesin peringkasan teks otomatis yang digunakan sebanyak dua macam, yaitu: ATS Lexrank dan ATS Mead.
3. *Candidate* dan *reference summary* memiliki jumlah kalimat yang sama.
4. Untuk mengatasi ketidak-konsistenan manusia, maka maksimal *references summary* dan *human judgement* yang digunakan adalah sebanyak lima orang (dengan latar belakang pendidikan yang sama, yaitu: mahasiswa maupun alumni: IT TELKOM)
5. Output yang dihasilkan berupa nilai dari metode ROUGE yang digunakan.
6. Parameter yang dianalisis adalah faktor struktur informasi, cakupan informasi, keterhubungan informasi, dan redundansi informasi.

1.3 Tujuan

Tujuan dengan dilakukannya penelitian ini adalah:

1. Merancang dan membangun aplikasi yang menerapkan algoritma LCS+SB dalam teknik ROUGE untuk mengevaluasi mesin peringkasan teks otomatis.
2. Menganalisis algoritma LCS+SB dalam teknik ROUGE untuk mengevaluasi mesin peringkasan teks otomatis, dengan melihat korelasinya terhadap penilaian manusia.
3. Menganalisis faktor-faktor yang dapat mempengaruhi korelasi antara algoritma LCS+SB terhadap penilaian manusia seperti persentase ringkasan dan jumlah referensi.

1.4 Metodologi Penyelesaian Masalah

Metodologi yang dilakukan untuk menyelesaikan permasalahan adalah sebagai berikut:

1. Studi literatur

Melakukan pencarian serta mempelajari informasi dan pembelajaran tentang evaluasi mesin peringkas teks otomatis, khususnya mengenai konsep dan cara kerja ROUGE dalam mengevaluasi mesin peringkas teks otomatis.

2. Pengumpulan data-data

Melakukan pencarian data yang digunakan untuk penelitian Tugas Akhir ini. Data yang dicari adalah artikel berita, baik artikel berita berbahasa Indonesia maupun berbahasa Inggris, yang nantinya diringkas dengan menggunakan mesin peringkas teks otomatis.

3. Analisis modifikasi ROUGE

Melakukan analisis kemungkinan modifikasi atau pengembangan terhadap teknik ROUGE, berdasarkan kelebihan dan kekurangan yang terdapat pada metode ROUGE-L dan ROUGE-S.

4. Analisis dan perancangan aplikasi

Melakukan analisis dan perancangan aplikasi teknik ROUGE berdasarkan algoritma LCS, SB, dan LCS+SB sehingga dapat digunakan untuk mengevaluasi mesin peringkas teks otomatis dengan pendekatan *summary-level*.

5. Implementasi aplikasi

Melakukan implementasi aplikasi sesuai dengan hasil analisis dan perancangan dari teknik ROUGE berdasarkan algoritma LCS, SB, dan LCS+SB untuk mengevaluasi mesin peringkas teks otomatis dengan pendekatan *summary-level*.

6. Pengujian aplikasi

Melakukan pengujian aplikasi dan menganalisis hasil keluaran aplikasi, sejauh mana hasil yang diberikan tersebut mampu menggambarkan kualitas dari mesin peringkas teks otomatis dengan melihat korelasinya terhadap penilaian manusia.

7. Pembuatan laporan Tugas Akhir

Melakukan penyusunan laporan hasil penelitian yang telah dilakukan serta memberikan kesimpulan dari hasil penelitian tersebut.