

# BAB I

## PENDAHULUAN

### 1.1 Latar Belakang Masalah

Pada penelitian sebelumnya yang dilakukan oleh Jerry Satiamy Saputra, Muhammad Fachrurozi, Yunita pada tahun 2017 dengan judul Peringkasan Teks Berita Berbahasa Indonesia Menggunakan Metode *Latent Semantic Analysis* (LSA) dan Teknik *Steinberger & Jezek* menjelaskan bahwa dokumen berita merupakan dokumen yang memuat berbagai macam informasi. Dengan membaca keseluruhan dokumen berita tersebut akan memakan banyak waktu. Peringkasan dokumen secara otomatis merupakan solusi untuk membantu mendapatkan intisari dari dokumen secara cepat. Penelitian ini melakukan penerapan metode *latent semantic analysis* dan teknik *steinberger & jezek* yang digunakan untuk peringkasan teks otomatis. Jumlah data uji yang digunakan sebanyak 10 teks berita. Hasil penelitian yang telah dilakukan menghasilkan rata-rata recall 0.7027, precision 0.6973, dan f-measure 0.6974 (Saputra et al., 2017).

Pada penelitian lainnya yang dilakukan oleh Nurina Savanti Widya Gotami, Indriati, Ratih Kartika Dewi pada tahun 2018 dengan judul Peringkasan Teks Otomatis Secara Ekstraktif Pada Artikel Berita Kesehatan Berbahasa Indonesia Dengan Menggunakan Metode *latent semantic analysis*, pada penelitian ini menjelaskan bahwa data teks dokumen yang dapat berupa teks artikel berita dapat membuat sulit pembaca dalam memahami seluruh informasi sehingga mempengaruhi perolehan informasi yang akurat dan dibutuhkannya waktu yang lebih lama untuk mengekstraksi suatu informasi

pada sebuah dokumen. Sehingga perlu adanya sebuah sistem peringkasan teks otomatis secara ekstraktif menggunakan metode *latent semantic analysis* (LSA). Metode LSA yang digunakan adalah pendekatan aljabar *linear singular value decomposition* (SVD) dengan membentuk matriks representasi dari asosiasi term yang merupakan kata-kata pada dokumen yang berhubungan erat dari proses perhitungan TF-IDF. Sedangkan *Cross method* LSA digunakan untuk menyusun urutan ringkasan dalam tahap ekstraksi ringkasan. Penelitian ini memperoleh nilai rata-rata akurasi *precision*, *recall* dan *f-measure* secara berurutan pada *compression rate* 50% dengan nilai 0.668, 0.743, 0.700 dan 0.690 sedangkan pada *compression rate* 40% sebesar 0.696, 0.605, 0.642 dan 0.663 (Savanti et al., 2018).

Informasi yang tersedia saat ini berkembang pesat seiring dengan pertumbuhan informasi digital. Kebutuhan untuk memperoleh informasi berita secara praktis menjadi masalah yang harus diselesaikan seiring dengan berkembangnya *volume* berita yang dapat diakses secara *online*. Semakin banyak informasi yang tersedia dalam suatu dokumen berita maka semakin panjang pula sebuah dokumen tersebut. Untuk mendapatkan inti informasi dari suatu dokumen dapat dilakukan dengan membaca isi dokumen secara keseluruhan. Hal ini membutuhkan waktu yang cukup lama jika dibandingkan dengan membaca isi ringkasan dari suatu teks. Oleh sebab itu, dibutuhkan suatu sistem yang mampu meringkas dokumen untuk mendapatkan inti dari dokumen asli, sehingga memudahkan pengguna dalam menemukan inti dari informasi yang dicari.

Ringkasan merupakan bentuk dokumen yang lebih singkat dari sebuah teks dokumen yang dibuat dari satu dokumen atau lebih, dan terdapat informasi-informasi

penting yang terdapat pada dokumen asli. Peringkasan teks otomatis atau yang dikenal sebagai *Automatic Text Summarization* (ATS) merupakan bagian dari NLP (*natural language processing*) yang bisa digunakan untuk membuat sebuah sistem aplikasi yang memberikan informasi penting dari sumber dokumen asli, serta secara otomatis mampu membuat ringkasan teks dokumen otomatis (Sari & Fatonah, 2021). Peringkasan teks otomatis dilakukan dengan menghitung nilai tiap dokumen dengan melakukan perangkian pada setiap dokumen. Dokumen yang memuat informasi penting didalamnya akan ditempatkan pada rangking teratas yang akan dijadikan kalimat ringkasan. *Automatic Text Summarization* dapat digunakan sebagai solusi untuk peringkasan teks dokumen dan tetap mempertahankan kualitas dari hasil ringkasan tersebut (Setyadi et al., 2018).

Metode peringkasan teks otomatis terus mengalami perkembangan untuk menghasilkan ringkasan yang lebih baik. *Algoritma Cross Latent Semantic Analysis* (CLSA) merupakan salah satu algoritma yang merupakan pengembangan dari algoritma *Latent Semantic Analysis* (LSA) untuk peringkasan teks dokumen. CLSA merubah perhitungan pada LSA dengan melakukan proses silang untuk mempercepat proses pada LSA (Mandar & Gunawan, 2017). Pada metode LSA pemanfaatan *Singular Value Decomposition* (SVD) untuk membuat serta mengolah matriks hasil proses pembobotan kata pada dokumen untuk menemukan hubungan kesamaan antara kata dan kalimat. Pembuatan ringkasan digunakan dengan memperhitungkan panjang kalimat, tidak hanya berdasarkan pada kemiripan kata dan kalimat dalam suatu dokumen. Dalam penelitiannya (Mandar & Gunawan, 2017). mengungkapkan bahwa metode CLSA membuat hasil ringkasan lebih pendek dari LSA. CLSA mampu

meningkatkan nilai akurasi dari metode LSA walaupun dengan hasil akurasi antara LSA dan CLSA tidak terlalu jauh berbeda. Penelitian yang dilakukan Mandar dan Gunawan masih bisa ditingkatkan dengan menambahkan *similarity* antara judul dan dokumen berita karena judul berita merupakan bagian penting dari suatu berita.

Berdasarkan uraian di atas penulis akan melakukan penelitian dengan judul **“PENERAPAN *TEXT SUMMARIZATION* PADA PORTAL BERITA PADANGMEDIA.COM DENGAN METODE *CROSS LATENT SEMANTIC ANALYSIS*”**.

## **1.2 Perumusan Masalah**

Berdasarkan uraian dari latar belakang yang telah dipaparkan diatas, permasalahan yang dapat dirumuskan adalah sebagai berikut:

1. Bagaimana aplikasi ATS (*Automated Text Summarization*) pada berita Bahasa Indonesia di [padangmedia.com](http://padangmedia.com) untuk mempermudah dalam pencarian inti berita?
2. Bagaimana aplikasi ATS (*Automated Text Summarization*) dapat mengetahui performansi dari sistem peringkasan teks Bahasa Indonesia di [padangmedia.com](http://padangmedia.com) menggunakan metode CLSA?

## **1.3 Hipotesis**

Berdasarkan rumusan masalah diatas, penulis membuat suatu hipotesa, yaitu

1. Dengan penerapan CLSA pada peringkasan teks berita diharapkan mampu mempercepat pembaca dalam mencari inti dari berita pada [padangmedia.com](http://padangmedia.com).

2. Sistem yang dirancang dengan berbasis web diharapkan dapat mempermudah masyarakat untuk melakukan peringkasan berita pada [padangmedia.com](http://padangmedia.com).

#### **1.4 Batasan Masalah**

Pada penulisan penelitian ini, penulis hanya membatasi masalah pada proses peringkasan teks otomatis menggunakan metode *Cross Latent Semantic Analysis*. Inputan teks yang digunakan berupa text berita Bahasa Indonesia dari halaman [padangmedia.com](http://padangmedia.com). Sistem yang dibangun berbasis web dan dataset yang digunakan pada pengujian sistem menggunakan dataset yang diambil dari halaman berita [padangmedia.com](http://padangmedia.com).

#### **1.5 Tujuan Penelitian**

Melalui penelitian ini, maka tujuan yang ingin dicapai oleh penulis adalah :

1. Membuat sistem aplikasi ATS (*Automated Text Summarization*) berita Bahasa Indonesia untuk mempermudah dalam pencarian inti dari berita.
2. Dengan menerapkan sistem aplikasi ATS (*Automated Text Summarization*) di [padangmedia.com](http://padangmedia.com) dapat menambah minat pembaca berita di [padangmedia.com](http://padangmedia.com).
3. Mengetahui performansi dari sistem peringkasan teks Bahasa Indonesia menggunakan metode CLSA.

## **1.6 Manfaat Penelitian**

Dengan menerapkan metode CLSA dalam peringkasan teks berita Bahasa Indonesia dapat menghasilkan ringkasan yang berisi intisari dari keseluruhan dokumen sehingga mempermudah pengguna sistem dapat membuat keputusan untuk melanjutkan membaca keseluruhan dokumen atau cukup dengan membaca ringkasan, membantu dan meningkatkan pengunjung di portal berita padangmedia.com.

## **1.7 Gambaran Umum Objek Penelitian**

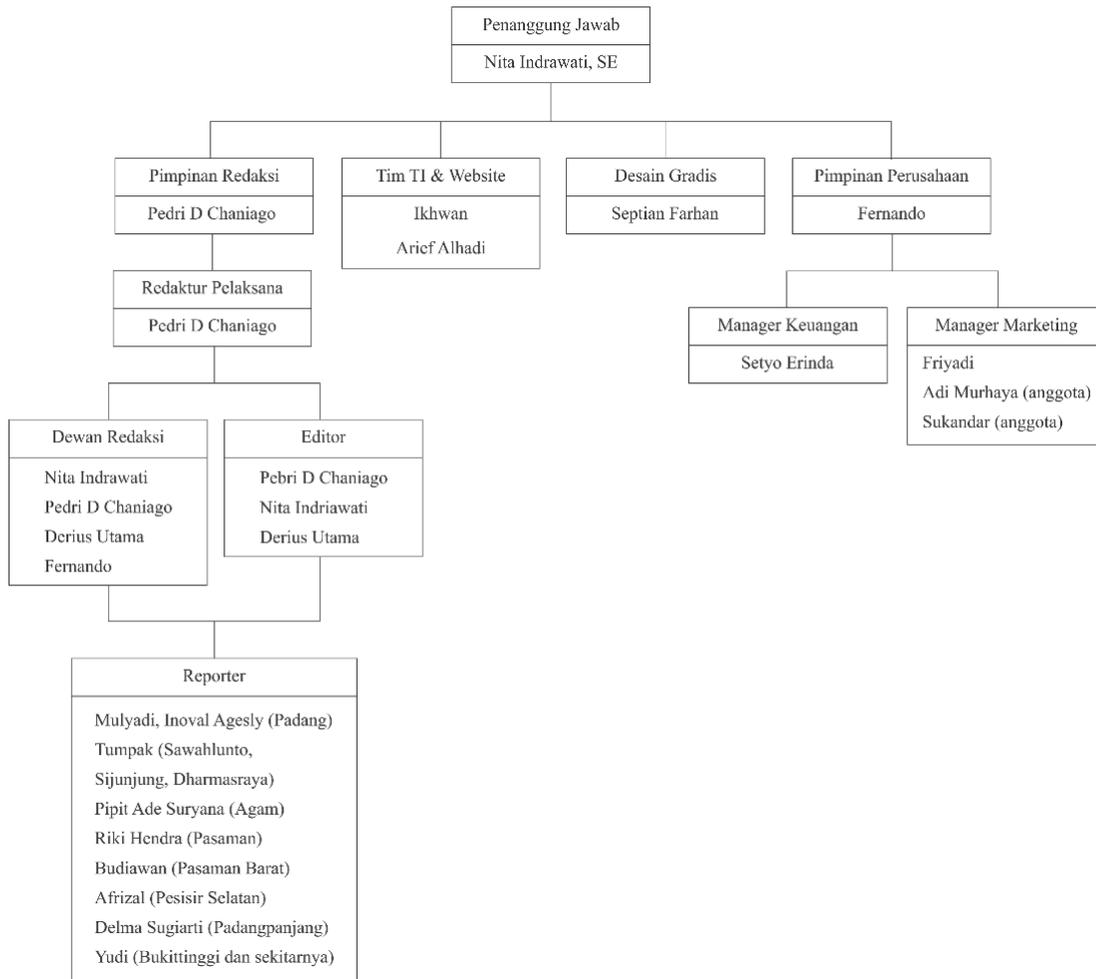
Objek penelitian yang digunakan didalam penelitian ini adalah perusahaan yang bergerak di bidang IT yaitu berita online. Situs halaman berita online ini adalah padangmedia.com yang diterbitkan oleh PT Padang Media Press.

## **1.8 Sekilas Tentang Objek Penelitian**

Padangmedia.com adalah sebuah portal berita online yang berdiri sejak 07 Juli 2007. Padangmedia.com terletak di kota Padang yang beralamat di Jalan Kimis No 10 F, Komplek PGRI Gunung Pangilun Padang. Sebagai media berita online pertama di Ranah Minang, padangmedia.com ingin menjadi bagian dalam percepatan kemajuan informasi dan teknologi di Sumatera Barat. Padangmedia.com telah terdaftar di Dewan Pers dan sekitar 75% redaktur dan wartawannya telah mengikuti dan lulus Ujian Sertifikasi Wartawan (UKW) dari Dewan Pers. Motto dari padangmedia.com adalah 'Terbetik Terberita'. Dilandaskan dengan semangat dalam memberi informasi secepat mungkin.

### 1.1.1 Struktur Organisasi

Struktur perusahaan PT Padang Media Press dapat dilihat pada gambar 1.1 sebagai berikut.



Sumber : PT Padang Media Press

Gambar 1.1 Struktur Perusahaan PT Padang Media Press

### 1.1.2 Tugas dan Tanggung Jawab

#### a. Penanggung Jawab

Penanggung Jawab adalah orang yang bertanggung jawab terhadap jalannya perusahaan dan redaksional. Ia akan secara berkala menerima

laporan perkembangan perusahaan dari pemimpin perusahaan dan penanggung jawab redaksi. Dengan data-data dari laporan-laporan itu, setiap penanggung jawab umum akan mengetahui perkembangan perusahaan.

b. Pemimpin Redaksi

Pemimpin Redaksi bertanggung Jawab atas jalannya semua instrumen dalam satu kali masa liputan hingga edisi terbit/disiarkan. Bertanggung jawab terhadap keberhasilan dan kegagalan sebuah pemberitaan, baik dari sisi hukum atau lainnya. Ia harus menguasai semua teknis dan non-teknis pemberitaan, gaya, jenis dan metode bagaimana menjadikan sebuah berita/program layak jual atau layak terbit.

c. Tim TI & Website

Dipimpin seorang kepala yang bertanggung jawab mengenai segala hal tentang komputerisasi. Seorang operator k omputer harus memahami sistem administrator local area network (LAN) untuk membagi data spatial (peta) dan non spatial (dokumen) dari semua komputer. Membuat situs website dapat terus berjalan membuat tampilan website semenarik mungkin.

d. Desain Grafis

Desain Grafis harus mampu membuat sebuah desain yang bisa memberikan / menyampaikan informasi dan pesan secara tepat sasaran, singkat padat dan jelas namun lebih menarik perhatian pembaca.

e. Pimpinan Perusahaan

Pimpinan Perusahaan bertanggung jawab mengendalikan dan mengkoordinasikan kebijakan di lingkungan perusahaan yang nantinya akan bertanggung jawab kepada Penanggung jawab Umum.

f. Manager Keuangan

Manager Keuangan bertanggung jawab mengenai semua hal yang berhubungan dengan keputusan investasi dan pembiayaan perusahaan yang berpengaruh pada laju pertumbuhan perusahaan.

g. Manager Marketing

Manager Marketing bertanggung jawab menyusun, mengatur, menganalisis, mengimplementasi dan mengevaluasi manajemen pemasaran, penjualan dan promosi.

h. Dewan Redaksi

Dewan Redaksi harus dapat berkomunikasi dengan baik dengan reporter di lapangan. Harus dapat menjawab pertanyaan reporter atau membantu reporter jika sewaktu-waktu mereka menemui kendala teknis di lapangan.

i. Redaktur / Editor

Redaktur berita bertugas memberikan TOR/outline kepada reporter sesuai hasil rapat redaksi. Setiap editor harus memberikan panduan teknis lapangan ke reporter sebelum bertugas meliput suatu isu.

j. Reporter

Reporter bertugas melakukan liputan sesuai hasil rapat redaksi (inline). Pelaksanaan liputan mengacu pada peran editor, yakni berupa penugasan (term of reference, TOR/outline), pengusulan tunggal, dan isu hangat.