

1. Pendahuluan

1.1 Latar Belakang

Dalam pemahaman isi dari suatu teks, kata kunci dapat membantu para pembaca untuk memahami teks tanpa membaca keseluruhan isinya. Kata kunci tersebut dapat menggambarkan konten yang sedang dibicarakan [15]. Menemukan kata kunci secara manual sangat sering digunakan, tetapi hal itu akan sulit jika dokumen berjumlah banyak hal tersebut akan memakan waktu yang lama. Agar kata kunci pada dokumen yang berjumlah banyak dapat diperoleh dengan cepat dibutuhkan ekstraksi kata kunci secara otomatis.

Penggunaan ekstraksi kata kunci otomatis pada dokumen teks *social media* adalah contoh penggunaan yang sangat menguntungkan karena perkembangan dari *social media* yang begitu pesat. Saat ini sudah banyak *social media* yang bermunculan seperti Facebook, Twitter, Plurk, LinkedIn, dan lain sebagainya yang sudah memiliki jutaan pengguna di seluruh dunia. Ekstraksi kata kunci pada dokumen *social media* ini dapat digunakan untuk *advertising*, *search*, dan *content filtering* [9].

Tidak seperti dokumen tradisional, teks pada *social media* biasanya sangat pendek dan cenderung tidak formal [9]. Dengan teks yang seperti itu, ekstraksi kata kunci pada dokumen *social media* memiliki perbedaan dengan ekstraksi kata kunci pada umumnya.

Ada banyak metode yang dapat digunakan dalam ekstraksi kata kunci. Salah satu metode yang dapat digunakan adalah KEA (*Keyphrase Extraction Algorithm*). Pada paper yang diacu menyatakan bahwa KEA merupakan metode yang sederhana dan efektif dalam penggunaannya [15]. KEA hanya menggunakan dua *feature* yang sederhana dalam penerapannya pada dokumen teks baik yang panjang ataupun pendek. Oleh karena itu, ekstraksi kata kunci pada dokumen *social media* yang memiliki teks pendek ini menggunakan KEA.

1.2 Perumusan Masalah

Berdasarkan latar belakang yang penulis uraikan sebelumnya, permasalahan yang akan diteliti dalam tugas akhir ini adalah sebagai berikut.

1. Bagaimana melakukan ekstraksi kata kunci pada dokumen *social media* yang bentuk teks nya relatif pendek.
2. Bagaimana mengetahui keakuratan KEA untuk menghasilkan kata kunci pada dokumen *social media* berbahasa Indonesia.

Dan adapun batasan masalah untuk tugas akhir ini adalah:

1. *Social media* yang digunakan adalah Twitter
2. Data yang dianalisis berupa *plain text* dari Twitter tidak termasuk foto, video, *link web*, dan lokasi
3. Data akan dianalisis secara *offline learning*

1.3 Tujuan

Tujuan yang diharapkan oleh penulis dalam pengerjaan tugas akhir ini adalah:

1. menentukan kata kunci pada dokumen *social media* menggunakan KEA
2. mengevaluasi akurasi dari hasil ekstraksi kata kunci menggunakan KEA

1.4 Metodologi Penyelesaian Masalah

Metode yang akan digunakan untuk menyelesaikan tugas akhir ini adalah sebagai berikut.

1. Studi Litelatur

Mempelajari literatur-literatur meliputi *crawling tweet using Twitter API, text mining*, dan ekstraksi frasa kunci terutama dengan metode KEA.

2. Pengumpulan Data

Mengumpulkan twit untuk dianalisis.

3. Perancangan Sistem

Melakukan analisis kebutuhan yang diperlukan dalam sistem serta merancang sistem yang sesuai dengan identifikasi kebutuhan tersebut.

4. Implementasi

Melakukan implementasi berdasarkan rancangan yang telah dibuat sebelumnya

5. Analisis Hasil

Analisis akurasi terhadap hasil pengujian yang didapat berdasarkan parameter uji.

6. Pengambilan Kesimpulan dan Pembuatan Laporan

Membuat kesimpulan dari analisa yang telah dilakukan dan membuat laporan tugas akhir yang mendokumentasikan tahap-tahap penelitian.