1. Pendahuluan

1.1 Latar Belakang

Semakin berkembangnya teknologi saat ini, kebutuhan manusia akan teknologi dan informasi semakin meningkat. Informasi yang diperoleh pun jumlahnya banyak. Hal ini juga mengakibatkan penumpukan informasi (baik *digital* maupun cetak), sehingga orang akan merasa bingung untuk memilih dan menemukan informasi yang sesuai. Tetapi permasalahan muncul ketika bagaimana mencari informasi yang berguna dan terkait diantara sekian banyak sumber informasi yang didapatkan. Salah satu cara untuk menyelesaikan permasalahan ini adalah dengan bantuan kata kunci yang diberikan oleh setiap dokumen.

Ekstraksi kata kunci adalah metode untuk mengidentifikasi kata atau sekumpulan kata/frase dari sebuah dokumen yang dapat menjelaskan makna dari sebuah dokumen. Kata kunci adalah kata atau kumpulan kata pada dokumen yang mampu memberikan deskripsi dari isi sebuah dokumen. Maka, dengan adanya kata kunci pembaca/orang yang ingin mencari informasi yang saling terkait dapat menemukan tanpa merasa kebingungan. Cara untuk menentukan kata kunci adalah dengan cara ekstraksi kata kunci. Dengan adanya kata kunci akan memudahkan proses *information retrieval* lainnya, seperti klasterisasi dokumen, klasifikasi dokumen, ringkasan dokumen, dan masih banyak lainnya.

Ada banyak metode yang dapat digunakan untuk mengekstraksi kata kunci, diantaranya yaitu *Naïve Bayes, Text Rank, Support Vector Machin (SVM), Lexical Chain, Hidden Markov Model (HMM),* dll. Dalam Tugas Akhir ini, metode yang digunakan adalah *Support Vector Machine* (SVM). SVM adalah salah satu metode *supervised learning* yang telah mencapai kinerja *state-of-the-art* pada berbagai *classification task*, termasuk *named entity recognition* [10]. Berdasarkan penelitian-penelitian yang sudah ada, SVM merupakan metode terbaik dengan menghasilkan akurasi diatas 70% dalam kasus *opinion mining*. Berdasarkan latar belakang SVM akan diimplimentasikan pada studi kasus lain yaitu ekstraksi kata kunci dan melihat bagaimana SVM diimplementasikan untuk kasus ini.

1.2 Perumusan Masalah

Masalah yang ditangani dalam Tugas Akhir ini adalah:

1. Bagaimana menentukan kata kunci/*keyword* suatu dokumen berbahasa Indonesia menggunakan metode *Support Vector Machine* (SVM).
2. Parameter apa yang dapat mempengaruhi akurasi dalam pencarian kata kunci/*keyword* pada dokumen berbahasa Indonesia.

Adapun batasan masalah dalam Tugas Akhir ini adalah:

1. Dataset berupa abstrak Tugas Akhir yang berasal dari digilib.ittelkom.ac.id.
2. Dataset yang digunakan menggunakan bahasa Indonesia.
3. Menggunakan proses *POS Tagging* yang sudah ada.
4. Frase *list* didapatkan secara manual.

5. LibSVM digunakan untuk membangun model *Support Vector Machine* dengan menggunakan *polynomial kernel*.

1.3 Tujuan

Tujuan dari Tugas Akhir ini adalah:

1. Menentukan kata kunci/*keyword* suatu dokumen berbahasa Indonesia menggunakan metode *Support Vector Machine* (SVM).
2. Mengetahui dan menganalisis parameter yang dapat mempengaruhi akurasi dalam pencarian kata kunci/*keyword* pada dokumen berbahasa Indonesia.

1.4 Metodologi Penyelesaian Masalah

Metodologi yang digunakan dalam memecahkan permasalahan-permasalahan dalam Tugas Akhir ini terdiri dari 6 tahap, yaitu:

1. Tahap Studi Literatur
   Pada tahap ini, dilakukan studi literatur dan mencari literatur yang berhubungan dengan topik Tugas Akhir baik *offline* maupun *online*. Memahami konsep dan teori dari *Support Vector Machine*.
2. Tahap Pengumpulan Data
   Pada tahap ini, dilakukan pengumpulan data berupa abstrak Tugas Akhir berbahasa Indonesia yang berasal dari digilib.ittelkom.ac.id yang dilakukan secara manual.
3. Tahap Perancangan Sistem
   Pada tahap ini, dilakukan perancangan sistem *keyword extraction* pada abstrak Tugas Akhir berbahasa Indonesia dengan metode *Support Vector Machine*.
4. Tahap Implementasi Metode
   Pada tahap ini, dilakukan implementasi berdasarkan rancangan yang telah dibuat sebelumnya.
5. Tahap Pengujian
   Pada tahap ini, dilakukan pengujian pada data yang sudah dikumpulkan terhadap sistem yang telah dibuat dalam *keyword extraction* dengan metode *Support Vector Machine*.
6. Tahap Analisis Hasil Pengujian
   Pada tahap ini, dilakukan analisis terhadap akurasi yang dihasilkan dari pengujian dan implementasi sistem.
7. Tahap Pembuatan Laporan
   Pada tahap ini, dilakukan penyusunan laporan dari hasil penelitian Tugas Akhir ini yang berfungsi sebagai dokumentasi.