BAB 1
PENDAHULUAN

1.1 Latar Belakang Masalah

Skripsi adalah karangan ilmiah yang wajib ditulis oleh mahasiswa sebagai bagian dari persyaratan akhir pendidikan akademisnya [1]. Skripsi bertujuan agar mahasiswa mampu menyusun dan menulis suatu karya ilmiah, sesuai dengan bidang ilmunya. Pelaksanaan skripsi pada program studi Teknik Informatika UNIKOM setiap tahunnya berdampak pada laporan skripsi yang semakin bertambah dan mengakibatkan penumpukan dokumen yang banyak dalam penyimpanan. Kondisi tersebut bisa memberi kemudahan dalam mendapatkan informasi, namun disisi lain informasi akan semakin bervariasi sehingga kesalahan sering terjadi dalam hal menentukan kategori laporan sesuai topik keilmuan. Penentuan keilmuan dokumen skripsi bisa saja diklasifikasi secara manual, namun dengan jumlah dokumen yang banyak akan menimbulkan kesalahan (*human error)* dalam menentukan kelompok keilmuan pada laporan skripsi serta lebih lama dalam proses pengerjaannya. Faktor pemicu kesalahan dalam menentukan kelompok keilmuan karena masih berkaitannya satu kelompok keilmuan dengan kelompok keilmuan yang lain. Untuk melakukan identifikasi laporan skripsi dapat dilakukan dengan memanfaatkan metode klasifikasi.

Penelitian yang berkaitan dengan klasifikasi dokumen sudah pernah dilakukan sebelumnya. Penelitian yang dilakukan oleh Yosep Amin mengenai "Implementasi Algoritma *Fuzzy* K-*Nearest Neighbour* (*Fuzzy* K-NN) Untuk Klasifikasi Proposal Skripsi Berdasarkan Kelompok Keilmuan Di Teknik Informatika UNIKOM"[2], menunjukkan bahwa berdasarkan hasil penelitian, analisis, pelatihan dan pengujian algoritma *Fuzzy* K-*nearest neighbour* untuk klasifikasi proposal skripsi berdasarkan kelompok keilmuan di Teknik Informatika unikom dengan data latih yang digunakan sebanyak 150 buah dan data uji sebanyak 67 buah, didapatkan nilai keakuratan yang terbesar yakni 51% pada nilai K=7 dan K=8. Dari penelitian tersebut didapatkan akurasi yang masih rendah, sehingga masih dibutuhkan solusi lain agar mendapatkan akurasi yang lebih baik.

Pada penelitian ini dilakukan dengan menggunakan algoritma SVM dalam pengklasifikasiannya, dan penelitian yang terkait klasifikasi dokumen teks dengan algoritma SVM pernah dilakukan sebelumnya. Penelitian yang dilakukan oleh Arini Daribti Putri dengan judul "Klasifikasi Dokumen Teks Menggunakan Metode Support Vector Machine Dengan Pemilihan Fitur Chi-Square"[3] menunjukkan bahwa pemilihan algortima SVM memberikan hasil akurasi yang baik untuk klasifikasi dokumen teks dengan dua kelas. Koleksi dokumen yang digunakan sebanyak 457 dokumen dibagi menjadi 70% dokumen latih (320 dokumen) dan 30% dokumen uji (137 dokumen) yang merupakan dokumen berbahasa Indonesia dengan format XML. Dalam penelitian tersebut pada tahap preprocessing tidak digunakan *stemming* dan *casefolding*, Pada pengujian dengan menggunakan kernel linear dan polinomial dihasilkan nilai akurasi sama baik, yaitu 96.35% dan pada kernel RBF dihasilkan akurasi sebesar 95.62%. Dalam penelitian lain oleh E. Dianawati dan A. Ardyanti dan R. N. Dawayanti dalam penelitiannya mengenai "Klasifikasi Dokumen Teks Menggunakan Svm-Dta (*Support Vector Machines With Decission Tree Architecture*)"[4], menjelaskan untuk hasil pengklasifikasian berita dinilai tepat, dengan melihat nilai error rate yang cukup rendah yakni 4% pada kernel polynomial dan kernel RBF. Dalam penelitian tersebut menggunakan klasifikasi 3 kelas dan pada tahap *preprocessing* tidak menggunakan *casefolding*. Dari hasil studi literature yang sudah diuraikan diatas, penggunaan SVM dalam klasifikasi dokumen teks memiliki keunggulan dengan tingkat keakurasian yang tinggi. Dalam penelitian tersebut masih diperlukan pengembangan dari batasan yang telah ada sebelumnya dalam mengklasifikasikan dokumen teks menggunakan SVM.

Berdasarkan masalah yang sudah dipaparkan, maka penulis mencoba mengimplemantasikan Metode *Support Vector Machine* (SVM) untuk mendapatkan hasil akurasi yang lebih baik dalam kasus klasifikasi dokumen Laporan Skripsi (studi kasus :Teknik Informatika UNIKOM).

1.2 Identifikasi Masalah

Berdasarkan latar belakang masalah yang telah disampaikan, dapat dirumuskan masalah yang timbul, yaitu Masih rendahnya akurasi dari klasifikasi laporan Skripsi Teknik Informatika UNIKOM.

1.3 Maksud dan Tujuan

Maksud penelitian ini adalah sebagai salah syarat kelulusan Program S1 Prodi Teknik Informatika UNIKOM, di samping itu penelitian ini dilakukan untuk merancang Sistem Pengklasifikasian Laporan Skripsi Prodi Teknik Informatika UNIKOM.

Sedangkan tujuan dari penelitian tugas akhir ini adalah untuk Mengetahui tingkat akurasi dari klasifikasi laporan Skripsi Teknik Informatika UNIKOM dengan penggunaan Support Vector Machine (SVM).

1.4 Batasan Masalah

Batasan masalah yang dibuat meliputi :

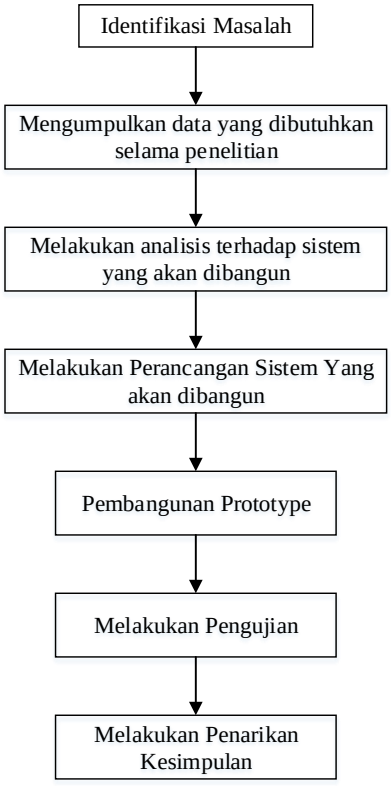
1. Kategori yang ditentukan berdasarkan kelompok keilmuan Teknik Informatika, yaitu :
   a. Kelompok A (Sistem Informasi)
   b. Kelompok B (RPL, Data dan BI)
   c. Kelompok C (Jaringan Komputer dan Multimedia)
   d. Kelompok D (Teknologi Internet dan Mobile)
   e. Kelompok E (Komputasi Perseptual Dan Kecerdasan Buatan)
2. Dokumen input yang digunakan adalah file-file abstrak Laporan Skripsi Program Studi Teknik Informatika Unikom yang diambil dari *elib.unikom.ac.id*.
3. Bentuk teks masukan dalam Bahasa Indonesia. Dengan format dokumen :*.xls dan tidak memperhatikan kesalahan penulisan (typo) pada dokumen input.
4. Laporan skripsi yang digunakan kelompok ilmuannya tidak memiliki banyak irisan dengan kelompok lain.
5. Data yang digunakan berjumlah 150 abstrak dimana 100 abstrak menjadi data latih dan 50 abstrak menjadi data uji.

6. Penelitian dibatasi pada tahap klasifikasi dokumen dan penilaian akurasi.
7. Menggunakana bahasa pemrograman *Java* dengan IDE Netbean berbasis *Dekstop* dan database yang digunakan yaitu MySQL.

1.5 Metodologi Penelitian

Dalam penelitian ini, metode yang digunakan yaitu deskriptif. Metode deskriptif adalah metode yang digunakan untuk menggambarkan atau menganalisis suatu hasil penelitian tetapi tidak digunakan untuk membuat kesimpulan yang lebih luas [5]. Metode deskriptif dapat diartikan sebagai prosedur pemecahan masalah yang diselidiki dengan menggambarkan keadaan subjek atau objek dalam penelitian berdasarkan fakta-fakta yang tampak atau apa adanya.

Langkah-langkah yang dilakukan selama melakukan penelitian dapat dilihat pada Gambar 1.1



Gambar 1. 1 Langkah-langkah Penelitian

1.5.1 Metode Pengumpulan Data

Studi Literatur

Studi Literatur yang dilakukan dalam penelitian ini yaitu mengumpulkan dan mempelajari sumber-sumber yang diperlukan berupa e-book, jurnal, tutorial dan beberapa informasi yang relevan dengan penelitian.

1.5.2 Metode Pembangunan Perangkat Lunak

Metode yang digunakan pada penelitian ini adalah model *Prototype* [6]. Berikut tahapan-tahapan yang dilakukan dalam penelitian ini :

1. Analisis

   Analisis masalah dilakukan untuk memahami masalah yang timbul dan mencari solusi untuk memecahkan masalah dalam menghasilkan sebuah klasifikasi dari beberapa kumpulan dokumen.

2. Kebutuhan data

   Pada tahap ini peneliti akan mengumpulkan data laporan skripsi berupa file abstrak untuk data masukan sistem.

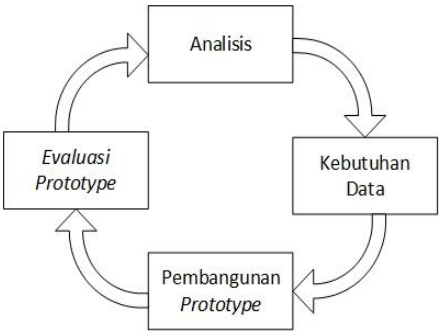
3. Pembangunan *Prototype*

   Merupakan tahap pengimplementasian dari proses analisis dan kebutuhan sistem yang sudah didapatkan. Pada tahap ini pengembang mencoba mengimplementasikan metode SVM ke dalam laporan skripsi teknik Informatika ke dalam logika-logika program.

4. Evaluasi *Prototype*

   Program akan diuji, dimana uji coba dilakukan untuk mengetahui kekurangan pada program. Jika masih ada kekurangan, maka *prototype* direvisi dengan tahapan-tahapan yang sebelumnya telah dilakukan.

   Tahapan *prototype* yang dilakukan pada penelitian ini akan dijelaskan pada gambar 1.2.



Gambar 1. 2 Model *Prototype*

1.6 Sistematika Penulisan

Penyusun membagi sistematika penulisan ini ke dalam :

BAB I PENDAHULUAN

Pada bab ini diuraikan mengenai latar belakang masalah, identifikasi masalah, maksud dan tujuan, batasan masalah mengenai klasifikasi dokumen dan metodologi penelitian *prototype* serta sistematika penulisan.

BAB II LANDASAN TEORI

Pada bab ini membahas mengenai landasan teori yang digunakan untuk menganalisis masalah dan teori yang digunakan dalam penelitian. yaitu teori mengenai Klasifikasi berdasarkan *Computing Curricula 2005*, *Preprocessing*, Metode SVM, dan teori mengenai software pembangun sistem.

BAB III ANALISIS DAN PERANCANGAN SISTEM

Pada bab ini membahas mengenai analisis metode SVM pada pengklasifikasian dokumen skripsi Teknik Informatika UNIKOM berdasarkan *Computing Curricula 2005*.

BAB IV IMPLEMENTASI SISTEM

Pada bab ini membahas mengenai pengujian metode SVM pada pengklasifikasian dokumen skripsi Teknik Informatika UNIKOM dari hasil analisis dan perancangan sistem.

BAB V KESIMPULAN DAN SARAN

Pada bab ini membahas tentang kesimpulan dari hasil penelitian serta saran untuk perbaikan sistem pengklasifikasian dokumen skripsi Teknik Informatika UNIKOM.