

## **BAB I**

### **PENDAHULUAN**

#### **1.1. Latar Belakang Penelitian**

Kemajuan di bidang ilmu pengetahuan dan teknologi mendorong pertumbuhan data di dunia. Berbagai fenomena alam semakin mudah diobservasi dan didokumentasikan. Keberadaan data sebagai nyawa dari tumbuh dan berkembangnya teknologi menjadi sangat mudah diperoleh. Apalagi perkembangan teknologi hardware dan software yang semakin pesat dan semakin murah, membuat proses dokumentasi menjadi semakin mudah. Faktor-faktor ini akhirnya menyebabkan terjadinya “ledakan data” di berbagai bidang.

Seperti halnya dalam bidang pendidikan yang secara khusus menyoroti masalah mahasiswa dan jurnal tugas akhirnya. Bagi perguruan tinggi yang sudah menerapkan basis IT (*Information of Technology*), adalah sebuah keharusan dan kebiasaan untuk seorang mahasiswanya untuk mendokumentasikan tugas akhirnya ke dalam bentuk *hardcopy* maupun *softcopy* agar bisa menjadi bukti fisik dalam bentuk dokumen akan keeksistensian jurusan tersebut di universitas maupun perguruan tinggi. Selain itu, pendokumentasian bisa juga dinikmati oleh khalayak ramai maupun mahasiswa lain yang sedang mencari referensi untuk tugas akhirnya.

Dalam suatu gelombang wisuda di suatu perguruan tinggi, akan ada banyak sekali mahasiswa yang melakukan uji sidang sampai dinyatakan lulus dan harus

menyerahkan tugas akhirnya berbentuk *softcopy* ke *digital library*. Untuk satu periode wisuda, sudah ada beberapa dokumen yang mengisi *digital library* dari berbagai jurusan. Bayangkan saja untuk berpuluh-puluh bahkan beratus-ratus periode wisuda. Akan sebanyak apa dokumen yang menumpuk dalam *digital library*.

Sejatinya, berkembangnya teknologi informasi akan menghasilkan berbagai jenis dokumen. Bila dokumen-dokumen jurnal tersebut tidak diatur berdasarkan kategorinya, dikhawatirkan dokumen-dokumen jurnal tersebut akan semakin menumpuk pada *digital library*. Dengan pengklasifikasian dokumen jurnal ke dalam kategori tertentu yang dituangkan dalam skripsi yang berjudul **“Penerapan Naïve Bayes Classifier Dalam Pengklasifikasian Jurnal Tugas Akhir”**, dokumen-dokumen jurnal di *digital library* akan tersusun rapi berdasarkan kategorinya dan pencarian informasi akan semakin mudah, efektif dan efisien.

## 1.2. Rumusan Masalah

Permasalahan yang timbul pada saat melakukan pengklasifikasian dokumen jurnal tugas akhir ini adalah:

1. Bagaimana hasil pengklasifikasian dokumen jurnal tugas akhir dengan masukan judul atau abstraknya?.
2. Seberapa akuratkah (%) proses klasifikasi dalam mengklasifikasikan jurnal tugas akhir (berdasarkan judul maupun abstraknya) yang belum memiliki kategori?.

3. Bagaimanakah pengaruh terhadap tingkat akurasi pengklasifikasian jurnal tugas akhir, jika terjadi pengurangan kuantitas data pelatihannya?.
4. Apakah pengklasifikasian jurnal tugas akhir lebih efektif berdasarkan judul atau abstraknya saja?.

### 1.3. Batasan Masalah

Dalam menyusun skripsi ini, penulis membatasi permasalahan dalam melakukan klasifikasi jurnal tugas akhir ini, hal ini dilakukan agar permasalahan dalam pengklasifikasian jurnal tugas ini tidak begitu meluas. Berikut adalah batasan masalahnya:

1. Pengklasifikasian jurnal berdasarkan pada frase atau kata-kata yang ada pada judul dan abstrak dari sebuah jurnal yang telah melalui proses *preprocessing*;
2. Kategori yang digunakan hanya tujuh dan kategori tidak dapat bertambah. ketujuh kategori tersebut antara lain: “Basis Data” (melingkupi: data *warehouse* dan *mining*), “Jaringan”, “Grafika” (melingkupi: pencitraan, penginderaan jarak jauh dan *wavelet*), “Sistem Informasi”, “Kecerdasan Buatan”, “Web/Multimedia” dan “Tak terkategori” (melingkupi jurnal-jurnal yang tidak dapat terkategori baik secara manual maupun secara otomatis);
3. Jurnal tugas akhir yang akan diklasifikasikan adalah jurnal tugas akhir yang telah melalui proses sidang dan dinyatakan lulus uji oleh kampus atau universitas tersebut;

4. Kategori setiap jurnal tugas akhir yang akan diklasifikasikan telah ditentukan sebelumnya secara manual, hal ini diperlukan untuk pembuatan data pelatihan dan untuk proses klasifikasi. Pengklasifikasian secara manual ini dapat dilakukan oleh pihak mahasiswa ataupun pihak jurusan. Nantinya, sistem akan mengecek hasil klasifikasi secara manual dengan hasil klasifikasi secara otomatis menggunakan sistem untuk mendapatkan angka akurasi pengklasifikasian;
5. Dalam melakukan proses pengujian (*testing*), data yang digunakan adalah data jurnal. Data pengujian adalah jurnal-jurnal yang belum diketahui kategorinya (belum dikategorikan oleh pihak mahasiswa ataupun pihak jurusannya). Jadi, jurnal-jurnal tugas akhir tersebut akan dikategorikan secara otomatis oleh sistem yang akan dikembangkan ini;
6. Jurnal yang akan digunakan untuk proses *testing* ini hanya dikategorikan pada satu kategori saja;
7. Sistem yang dikembangkan menangani masalah data *preprocessing*;
8. Jurnal-jurnal yang diklasifikasikan adalah jurnal-jurnal yang berhubungan dalam bidang informatika atau ilmu komputer.

#### 1.4. Maksud dan Tujuan Penelitian

Maksud dari penelitian ini adalah untuk memperoleh, mengolah dan menganalisis data menjadi informasi yang diperlukan dalam penulisan skripsi dan pengembangan sistem.

Adapun tujuan dari penelitian ini adalah:

1. Merancang model pengklasifikasian dokumen jurnal tugas akhir;
2. Mengimplementasikan hasil perancangan model pengklasifikasian dokumen jurnal tugas akhir.

### 1.5. Manfaat Penelitian

Ada beberapa manfaat dalam pengklasifikasian jurnal tugas akhir ini. Berikut adalah beberapa manfaatnya:

1. Bagi pihak pencari informasi jurnal tugas akhir

Kemudahan dalam mencari dan mendapatkan informasi mengenai sebuah jurnal dan kategorinya.

2. Bagi pihak *stakeholder*, dalam hal ini adalah pihak universitas atau perguruan tinggi

Efisiensi kapasitas penyimpanan data di *digital library* karena jurnal-jurnal tersebut sudah dikategorikan ke dalam kategori-kategori tertentu sehingga jurnal-jurnal tersebut tidak begitu menumpuk dalam *digital library*.

3. Bagi penulis

Hasil penelitian ini diharapkan dapat menambah wawasan penulis dan dapat menjawab keingintahuan penulis terhadap masalah yang penulis teliti sehingga

penulis dapat lebih memahami mengenai masalah pengklasifikasian jurnal tugas akhir.

#### 4. Bagi pihak lain

Hasil penelitian ini diharapkan dapat dijadikan sebagai bahan perbandingan bagi pihak lain yang ingin melakukan penelitian lebih lanjut dan diharapkan penelitian ini dapat menambah wawasan bagi para pembaca.

### 1.6. Metode Penelitian

Pengerjaan pengklasifikasian jurnal tugas akhir ini dilakukan dalam beberapa metode penelitian, yaitu:

#### 1. Metode Pengumpulan Data

Penulis menggunakan konsep studi pustaka dengan cara mencari berbagai referensi dan acuan dalam mempelajari teori, konsep, dan algoritma dari salah satu algoritma *classification* dalam *data mining*, yaitu *Naïve Bayes Classifier*. Penulis juga melakukan pembelajaran dan pencarian terhadap data-data jurnal (judul dan abstrak) dan kategorinya untuk dijadikan sebagai data masukan dalam mengembangkan sistem.



## 2. Metode Pengembangan Perangkat Lunak

### a) Metode Pendekatan Perangkat Lunak

Metode yang digunakan adalah metode pendekatan terstruktur, yakni analisis yang terfokus pada aliran data. Pendekatan terstruktur mengenalkan beberapa alat untuk mengembangkan sistem yang terstruktur. Alat-alat tersebut diantaranya, *data dictionary* (Kamus Data), *Entity Relationship Diagram* (ERD), *Data Flow Diagram* (DFD), *State Trasisition Diagram* (STD), *Process specification* (Pspec). Namun, dalam pengembangan sistem ini, penulis tidak menggunakan *State Trasisition Diagram* (STD) karena sistem yang akan dikembangkan bukan sistem yang *real time*.

### b) Model Proses

Model proses dalam pengembangan perangkat lunak ini menggunakan model *Sequence Linier*. Dibagi ke dalam beberapa tahapan, yaitu: Analisis (*Analysis*), Desain (*Design*), Implementasi (*Code*), Pemeliharaan dan Pengujian (*Testing*).

Penjelasan mengenai point-point dalam metode penelitian ini seperti studi pustaka, metode pendekatan perangkat lunak dan tahapan model proses, ada pada BAB III dan dokumen teknis yang disertakan bersama skripsi ini.

## 1.7. Sistematika Penulisan

Dalam penulisan skripsi ini, sistematika penulisan dibagi ke dalam beberapa tahapan, yaitu:

## BAB I PENDAHULUAN

Bab ini meliputi pembahasan masalah secara umum meliputi latar belakang masalah, identifikasi masalah yang di jelaskan ke dalam rumusan masalah dan batasan masalah, maksud dan tujuan penelitian, manfaat penelitian, metode penelitian sampai pada sistematika penulisan.

## BAB II TINJAUAN PUSTAKA

Bagian ini memuat landasan teori yang berfungsi sebagai sumber atau alat dalam memahami permasalahan yang berkaitan dengan teori jurnal, *data mining*, *classification*, *preprocessing* dan algoritma *classification*, yaitu: *Naïve Bayes Classifier* (NBC).

## BAB III METODOLOGI PENELITIAN

Bab ini merupakan penjabaran dari metode pengembangan sistem yang digunakan yaitu metode berorientasi proses. Dijelaskan pula simbol-simbol yang akan digunakan untuk memodelkan sistem.

## BAB IV HASIL PENELITIAN DAN PEMBAHASAN

Pada bagian ini akan dikupas secara mendalam hal-hal yang akan menjawab apa yang sudah dirumuskan dalam rumusan masalah.



## BAB V KESIMPULAN DAN SARAN

Kesimpulan merupakan jawaban atas rumusan masalah dalam penelitian dan juga intisari dari BAB IV. Saran atas kesimpulan serta rekomendasi pengembangan sistem penulis utarakan pada subbab saran.

