

# BAB I

## PENDAHULUAN

### 1.1 Latar Belakang

Khazanah keilmuan suatu Perguruan Tinggi sangat dipengaruhi oleh banyaknya jurnal dan hasil penelitian Ilmiah yang diterbitkan oleh para Dosen, Mahasiswa, maupun Lembaga Penelitian Kampus yang aktif mencari dan menuangkan narasinya dalam sebuah jurnal, makalah, tugas akhir, dan skripsi. Melalui perpustakaan kampus, sudah banyak jurnal, skripsi maupun karya diterbitkan. Hal ini berperan besar dalam meningkatkan kualitas dan akreditasi suatu Perguruan Tinggi. Semakin banyak jurnal, skripsi dan penelitian ilmiah yang terbit, juga harus diimbangi dengan kaya dan dinamis nya tema, studi kasus, maupun *platform* yang dipakai, sehingga sangat berpengaruh pada kualitas jurnal yang diterbitkan. Ini penting untuk menghindari adanya kemiripan bahkan plagiarisme di internal Perguruan Tinggi tersebut.

ISB Atma Luhur Pangkalpinang adalah Perguruan Tinggi Informatika dan Komputer di Provinsi Kepulauan Bangka Belitung. Kampus ini berdiri sejak tahun 2000. Sampai saat ini sudah ribuan jurnal, karya ilmiah dan penelitian yang diterbitkan oleh mahasiswa, dosen, dan lembaga Penelitian di internal kampus sendiri. Sampai Maret 2020, *repository* Atma Luhur sudah menerbitkan 2,227 judul skripsi dan 785 laporan kerja praktek dari Program Studi Teknik Informatika, Sistem Informasi, dan Manajemen Informatika. Hanya saja, keberagaman judul dan isi dari abstrak skripsi masih kurang heterogen, dan banyak kemiripan yang berindikasi ke arah plagiarisme.

Selama ini pihak kampus kesulitan untuk melakukan verifikasi terhadap mahasiswa yg mengajukan judul skripsi, saling tumpang tindih antara judul yang memiliki kesamaan konsep, metode, bahkan studi kasus dengan kemiripan yang cukup sulit untuk diperhitungkan titik acuannya. Jika ini tidak ditangani dalam waktu

dekat, maka pada selanjutnya akan banyak terjadi plagiarisme skripsi pada tingkat internal kampus Atma Luhur.

Algoritma *Cosine Similarity* merupakan metode untuk menghitung kesamaan antara dua buah objek yang dinyatakan dalam dua buah dengan menggunakan (kata kunci) dari sebuah dokumen sebagai ukuran<sup>[2]</sup>. Metode ini sudah umum digunakan untuk mendeteksi kemiripan dari sekumpulan kalimat dan kata. Sedangkan *Stemmer Library* Sastrawi merupakan sekumpulan kamus struktur Bahasa Indonesia yang dikembangkan oleh *Developer* Robbani dan Hanif Amal, dirangkai dan di dokumentasi dengan baik dalam bentuk File *Hypertext Preprocessor* (PHP), dan bersifat *free licence*.

Adapun dalam pembuatan laporan ini, penulis mengambil beberapa dari penelitian, Yaitu Penelitian oleh Sugiyamta pada Oktober tahun 2015 mengenai Sistem Deteksi Kemiripan Dokumen dengan Algoritma Cosine Similarity dan Single Pass Clustering<sup>[1]</sup>. Penelitian oleh Rizki Tri Wahyuni, Dhidik Prastiyanto, dan Eko Suprpto pada Bulan Januari-Juni tahun 2017 Penerapan Algoritma Cosine Similarity dan Pembobotan TF-IDF pada Sistem Klasifikasi Dokumen Skripsi<sup>[2]</sup>. Penelitian Pratama, Rito Putriwana Faisal, Muhammad Hanani, Ajib pada tahun 2019 mengenai Deteksi Plagiarisme pada Dokumen Jurnal Menggunakan Metode Cosine Similarity<sup>[3]</sup>. Ariantini, Dewa Ayu Rai, Lumenta, Arie S. M. dan Jacobus, Agustinus, dan Ajib pada tahun 2016 tentang pengukuran kemiripan dokumen teks bahasa Indonesia menggunakan metode cosine similarity<sup>[4]</sup>. Sejati, Fauzi Bayu Hendradi, Purwono Pujiarto, dan Bambang tahun 2019 mengenai Deteksi Plagiarisme Karya Ilmiah dengan Pemanfaatan Daftar Pustaka dalam Pencarian Kemiripan Tema Menggunakan Metode Cosine Similarity (Studi Kasus: Di Universitas Muhammadiyah Magelang)<sup>[5]</sup>.

Dari rangkaian latar belakang diatas maka peneliti tertarik untuk mengusulkan judul: Penerapan Algoritma Cosine Similarity Dan Optimasi Stemmer Library Sastrawi Dalam Menghitung Persentase Kemiripan Judul Dan Abstrak Skripsi (Studi Kasus: *Repository* Atma Luhur Pangkalpinang).

## 1.2 Rumusan Masalah

Berdasarkan uraian pada latar belakang yang telah dijelaskan, maka permasalahan yang akan dikaji dalam penelitian ini dapat dirumuskan sebagai berikut:

1. Bagaimana cara menghindari plagiarisme judul skripsi di lingkungan ISB Atma Luhur?
2. Bagaimana cara menentukan titik acuan kemiripan judul skripsi bagi mahasiswa dosen ataupun kaprodi sebelum melakukan verifikasi proposal skripsi?

## 1.3 Tujuan dan Manfaat Penelitian

Adapun tujuan dan manfaat dari merancang dan membuat WEB sistem deteksi kemiripan judul dalam penelitian ini adalah:

### 1.3.1 Tujuan

Tujuan penulisan skripsi ini dapat dilampirkan dalam beberapa poin, diantaranya adalah :

1. Mendeteksi judul skripsi yang akan diajukan oleh Mahasiswa akhir dalam bentuk persentase kemiripan pada *repository* Atma Luhur Pangkalpinang.
2. Membantu pihak kampus, Dosen maupun Kaprodi dalam menentukan judul mahasiswa berdasar kesamaan judul hasil perhitungan *Cosine Similarity*.

### 1.3.2 Manfaat

Beberapa manfaat yang ingin dicapai dalam melakukan penelitian skripsi ini sebagai berikut:

1. Deteksi dini dan menghindari plagiarisme bagi mahasiswa semester akhir Atma Luhur yang ingin mengajukan judul skripsi.
2. Agar tidak terjadi tumpang tindih judul atau tema yg sama dengan tahun berikutnya.

3. Sebagai acuan sekunder bagi Dosen Pembimbing, maupun Kaprodi dalam menimbang, menyetujui, dan memberi masukan judul yang diajukan oleh mahasiswa bimbingannya.

#### **1.4 Batasan Masalah**

Dalam merancang dan membuat website sistem deteksi kemiripan judul dan abstrak dalam penelitian ini diberikan batasan masalah sebagai berikut:

1. Rancangan Sistem ini berbasis Web, menggunakan MySQL sebagai *database*, dan *composer* untuk memanggil *Stemmer Library* Sastrawi.
2. *Stemmer Library* Sastrawi hanya dipakai untuk *preprocessing* sampai *stemmer* kata imbuhan dan optimasi kata kunci, dan tidak dibahas secara rinci struktur datanya.
3. Sistem pendeteksi ini hanya berlaku pada *Repository* Atma Luhur Pangkalpinang

#### **1.5 Sistematika Penulisan**

Untuk mengetahui kerangka keseluruhan penulisan, penulis menjabarkan sistematika penulisan sebagai berikut :

### **BAB I PENDAHULUAN**

Berisi tentang latar belakang, perumusan masalah, tujuan dan manfaat penelitian, ruang lingkup penelitian, dan sistematika penulisan.

### **BAB II LANDASAN TEORI**

Bab ini berisikan teori yang berupa definisi model pengembangan perangkat lunak, definisi metode pengembangan perangkat lunak, definisi tools pengembangan perangkat lunak, teori pendukung, dan penelitian terdahulu.



### **BAB III METODOLOGI PENELITIAN**

Bab ini berisikan metodologi penelitian yang akan dipakai sebagai *tool* untuk menjabarkan hasil dan pembahasan pada BAB IV, meliputi model pengembangan sistem, metode pengembangan sistem, *tools* pengembangan sistem, *Text Preprocessing*, Algoritma *Cosine Similarity*.

### **BAB IV HASIL DAN PEMBAHASAN**

Bab ini berisi profil dan struktur Perpustakaan Atma Luhur, dilanjutkan pada Analisis masalah meliputi analisis kebutuhan dan sistem berjalan, Perancangan sistem meliputi identifikasi system usulan rancangan system dan rancangan layar, dan implementasi meliputi tampilan layar, dan pengujian.

### **BAB V PENUTUP**

Bab ini berisi kesimpulan dan saran yang berkaitan dengan analisa dan optimalisasi sistem berdasarkan yang telah diuraikan pada bab sebelumnya.

