

## BAB I

### PENDAHULUAN

#### 1.1 Latar Belakang

Dalam dunia pendidikan naskah publikasi dari suatu penelitian merupakan hal yang lumrah dikerjakan bagi kebanyakan civitas akademika. Pasalnya naskah publikasi ini menjadi syarat untuk memperoleh gelar terutama pada strata pendidikan, magister dan doktoral. Kemampuan dalam menyesuaikan jurnal yang sesuai dengan naskah yang telah dikerjakan agar dapat terpublikasi di jurnal tersebut. [1] Pembahasan mengenai penulisan naskah ini menjadi persoalan serius terkait maraknya kasus plagiarisme yang terjadi dalam penulisan karya ilmiah ini. Hal ini disebabkan karena kemajuan teknologi informasi membuat seorang akademisi dengan mudahnya mendapatkan segala informasi yang dibutuhkan untuk menyelesaikan naskahnya. Kemudahan dalam mendapatkan informasi inilah yang membuat seorang akademisi menghasilkan karya dengan cara menyalin atau menjiplak hasil karya orang lain sehingga terjadi banyak kesamaan data diantara akademisi.

Kasus pagiarisme sangat bisa ditanggulangi dengan cara pencegahan dan penanganan. Sebenarnya sudah jelas ditulis didalam Undang-Undang No. 20 tahun 2003 yang mengatur tentang sanksi bagi orang-orang yang melakukan plagiat, khususnya yang terjadi dalam lingkungan akademik. Penggunaan alat deteksi *similarity* secara umum dapat menjadi langkah yang cukup efektif dalam

menangani plagiarisme yang berfungsi untuk mendeteksi plagiarisme secara otomatis. [2]

Untuk mengetahui seberapa besar persamaan suatu dokumen teks dengan dokumen teks yang lainnya dapat menggunakan pendekatan *string metric* yaitu melakukan perbandingan *string* dengan memasukkannya kedalam fungsi matematis tertentu. Metode ini sering digunakan dalam mendeteksi sebuah dokumen yang diperkirakan sama dengan dokumen lain. Semakin tinggi tingkat *similarity* yang ditemukan didalam dalam dua *string*, maka semakin mirip juga *string* tersebut. [3]

Berdasarkan permasalahan tersebut dibutuhkan sistem yang dapat digunakan untuk pendeteksi *similarity* dokumen naskah publikasi. Dengan menggunakan pendekatan *string* metrik, dimana dengan melakukan proses perbandingan *string* dan memasukkan dalam fungsi matematis. Terdapat beberapa algoritma yang berdasarkan *string matrix* yaitu *TF/IDF*, *Needleman-Wunsch Distance*, *Levenshtein Distance*, *Jaro-Winkler Distance*, dan lain sebagainya. Dari beberapa algoritma yang tersebut, penulis memilih algoritma *Jaro-Winkler Distance* karena algoritma ini dapat memeriksa kemiripan hanya dengan melewati sedikit proses sehingga meminimalkan waktu dan proses yang relatif lebih cepat, algoritma *Jaro-winkler* sangat akurat dan cocok untuk digunakan dalam perbandingan *string* pendek, seperti abstraksi pada naskah publikasi.

## 1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah dijelaskan diatas dapat diambil suatu rumusan masalah, yaitu sebagai berikut:

1. Metode seperti apa yang cocok untuk menguji kemiripan suatu dokumen naskah publikasi?
2. Bagaimana cara mengimplementasikan algoritma *Jaro Winkler Distance* untuk mendeteksi *similarity* pada abstraksi dokumen naskah publikasi?
3. Bagaimana cara pengujian *similarity* dokumen teks dengan menggunakan *stremmer* sastra?
4. Bagaimana cara kerja aplikasi pendeteksi *similarity* dokumen naskah publikasi dengan cara membandingkan dua abstraksi naskah publikasi?

### 1.3 Batasan masalah

Beberapa batasan masalah yang akan digunakan dalam penelitian ini adalah sebagai berikut:

1. Aplikasi hanya dapat mendeteksi teks yang bersifat *plain text* seperti huruf dan angka, tidak termasuk gambar, diagram, dan simbol.
2. Aplikasi hanya untuk mendeteksi *similarity text* yang berbahasa Indonesia.
3. Data yang diuji merupakan teks abstrak naskah publikasi program studi informatika Universitas Amikom Yogyakarta.
4. Aplikasi tidak memperhatikan kesalahan penulisan atau ejaan pada dokumen tersebut, serta tidak memperhatikan sinonim dari kata tersebut.

5. Implementasi sistem yang dibuat menggunakan bahasa pemrograman berbasis HTML, Javascript dan PHP yang akan berjalan melalui *browser*.
6. Pembangunan aplikasi menggunakan *software* Notepad++ dan Sublime text, dengan menggunakan *Bootstrap* untuk mempercantik tampilan aplikasi.

#### **1.4 Maksud dan Tujuan Penelitian**

##### **1.4.1 Maksud Penelitian**

Maksud penelitian dengan judul “Implementasi Algoritma Jaro Winkler Distance untuk Sistem Pendeteksi Similarity Pada Dokumen Naskah Publikasi” adalah untuk memenuhi persyaratan dalam mencapai gelar sarjana pada program studi S1 Informatika di Universitas Amikom Yogyakarta.

##### **1.4.2 Tujuan Penelitian**

Tujuan dari penelitian ini adalah sebagai berikut:

1. Merancang aplikasi yang dapat mendeteksi *similarity* teks dengan menggunakan algoritma *Jaro Winkler Distance* berbasis website.
2. Mendeteksi *similarity* dokumen teks pada abstraksi naskah publikasi yang diharapkan dapat memberikan solusi yang lebih baik dari permasalahan yang sebelumnya.

#### **1.5 Manfaat Penelitian**

Hasil dari penelitian ini diharapkan dapat bermanfaat untuk:

1. Bagi Peneliti

Sebagai sarana untuk melatih dan mengembangkan kemampuan dalam bidang penelitian, serta menambah pengetahuan dan wawasan dalam penerapan teori-teori yang sudah diperoleh di bangku kuliah.

## 2. Bagi Universitas Amikom Yogyakarta

Menambah referensi ilmiah dalam bentuk skripsi bagi mahasiswa yang sedang melakukan penelitian dan yang sedang menyusun skripsi atau tugas akhir di Perpustakaan Amikom Yogyakarta.

## 3. Bagi Masyarakat

Aplikasi ini diharapkan dapat membantu mengatasi berbagai permasalahan dan sebagai bahan pertimbangan dalam melakukan pengujian similarity teks dan plagiarisme pada kehidupan bermasyarakat serta menambah wawasan masyarakat tentang kajian ilmiah berdasarkan teori yang telah diuji kebenarannya.

### 1.6 Metode Penelitian

Pada penelitian ini, peneliti menggunakan beberapa metode penelitian sebagai berikut:

#### 1.6.1 Metodologi Pengumpulan Data

Pengumpulan data yang dilakukan untuk memperoleh data-data yang dilakukan dalam penelitian ini yaitu:

##### 1. Studi Literatur

Pengambilan sumber data teori dasar dari penelitian-penelitian sebelumnya, mempelajari teori - teori similarity, plagiarisme, dan algoritma Jaro Winkler. Sumber bacaan berupa jurnal-jurnal, e-book,

artikel yang diperoleh dari internet yang berguna untuk menunjang penelitian, dan skripsi yang diperoleh di perpustakaan Universitas Amikom Yogyakarta.

## 2. Metode Penelitian

Mengumpulkan sampel naskah publikasi untuk dilakukan percobaan secara kualitatif.

### 1.6.2 Analisis Sistem

Analisis sistem disini adalah penguraian dari suatu sistem informasi yang utuh kedalam bagian-bagian komponennya dengan maksud untuk mengidentifikasi dan mengevaluasi permasalahan, kesempatan, hambatan yang terjadi dan kebutuhan-kebutuhan yang diharapkan sehingga dapat diusulkan perbaikannya. [19]

### 1.6.3 Metode Pengembangan Aplikasi

Pada penelitian ini, peneliti menggunakan metode waterfall. Berikut adalah penjelasan dari tahapan-tahapan metode ini:

#### 1. Requirement Analysis

Pada tahap *Requirement Analysis* ini peneliti mengeksplorasi dan melakukan pengumpulan data-data dari pustaka, pengetahuan saat kuliah, serta mempelajari referensi dari beberapa buku, jurnal, serta artikel dari internet yang berhubungan dengan *similarity* dan algoritma *jaro winkler distance*.

#### 2. System Design

Pada Tahap *System Design* ini peneliti merancang kebutuhan system perangkat yang akan digunakan dan merancang alur blok diagram yang digunakan dalam penelitian ini.

### 3. *Implementation*

Pada tahap ini peneliti merancang aplikasi sesuai dengan skematik diagram yang telah dibuat serta memprogram sistem menggunakan bahasa *php*, *javascript* dan *html* melalui comand line di *notepad++* dan *sublime text*.

### 4. *Integration & Testing*

Pada tahap ini peneliti melakukan pengintegrasian kedalam sistem dan melakukan pengujian terhadap sistem yang telah dibuat, peneliti juga melakukan pengujian aplikasi yang telah dibuat, apakah sudah berjalan sesuai dengan yang diinginkan.

### 5. *Operation and Maintenance*

Pada tahap ini peneliti menjalan perangkat sistem serta melakukan pemeliharaan. Pemeliharaan termasuk dalam memperbaiki kesalahan yang tidak dilakukan pada langkah sebelumnya.

## 1.7 Sistematika Penulisan

Guna memahami lebih jelas laporan Skripsi ini, maka laporan ini dikelompokan kedalam beberapa sub bab dengan sistematika penulisan sebagai berikut:

## **BAB I** Pendahuluan

Bab ini berisi uraian tentang latar belakang masalah, rumusan masalah, batasan masalah, maksud dan tujuan, metode penelitian & sistematika penulisan.

## **BAB II Landasan Teori**

Berisikan berbagai dasar teori yang mendukung dan berisi tentang pengujian-pengujian yang telah dilakukan oleh orang lain yang nantinya akan mendukung dan mendasari penulisan skripsi ini.

## **BAB III Analisis Dan Perancangan**

Berisikan tentang penjabaran tentang tahapan-tahapan berupa penelitian, pengumpulan data, analisa kebutuhan sistem, perancangan, perangkat lunak, implementasi, pengujian sistem dan analisa akhir.

## **BAB IV Implementasi dan Pembahasan**

Bab ini akan membahas mengenai implementasi sistem pendeteksi tingkat plagiarisme abstraksi naskah publikasi menggunakan algoritma *Jaro Winkler Distance* serta disajikan evaluasi dan kesimpulan dari pengujian yang telah dilakukan terhadap sistem.

## **BAB V Kesimpulan dan Saran**

Bab ini berisi kesimpulan dari keseluruhan penelitian dan saran-saran yang membangun untuk pengembangan serta perbaikan dari pembuatan sistem yang sudah dibuat.