

BAB I

PENDAHULUAN

1.1 Latar Belakang

Dengan semakin berkembangnya teknologi informasi, sehingga membuat pembuatan karya tulis semakin mudah dan cepat. Hal tersebut dikarenakan informasi kini tersedia secara melimpah. Akan tetapi dikarenakan kemudahan dalam memperoleh informasi tersebut, pada pembuatan karya tulis sering ditemukan kesamaan dengan karya tulis orang lain sehingga kemudian menimbulkan isu plagiarisme.

Menurut *Tia Septiani Widi* (2012), plagiarisme adalah tindakan penyalahgunaan, pencurian/perampasan, penerbitan, pernyataan, atau menyatakan sebagai milik sendiri sebuah pikiran, ide, tulisan, atau ciptaan yang sebenarnya milik orang lain.

Jaro-Winkler distance merupakan varian dari *Jaro distance metric* yang merupakan sebuah algoritma untuk mengukur kesamaan antara dua string, biasanya algoritma ini digunakan di dalam pendeteksian duplikat dokumen. Penelitian ini akan membahas mengenai pendeteksian plagiarisme dari sebuah dokumen dengan melakukan perbandingan dengan dokumen lainnya yang memanfaatkan metode pencocokan string pada dokumen.

Berdasarkan masalah yang telah dikemukakan, maka peneliti mengambil penelitian Tugas Akhir dengan Judul **“Deteksi Plagiarisme Dokumen Bahasa Indonesia Dengan Algoritma Jaro-Winkler Distance”** dengan harapan mampu mengatasi kendala yang telah dikemukakan tersebut.

1..2 Rumusan Masalah

Berdasarkan uraian latar belakang masalah yang sudah dijelaskan di atas, permasalahan yang dapat dirumuskan adalah mengetahui presentase tingkat

kemiripan abstrak pada dokumen bahasa Indonesia dengan algoritma *Jaro Winkler Distance*.

1.3 Batasan Masalah

Agar pembahasan dalam penelitian ini dapat dilakukan secara terarah dan sesuai dengan yang diharapkan, maka perlu diterapkan batasan-batasan permasalahan yang akan dibahas, antara lain:

1. Pada tugas akhir ini hanya menyajikan informasi tentang prosentase tingkat kemiripan dokumen berbahasa Indonesia Program Studi Teknik Informatika dengan metode *Jaro-Winkler Distance*.
2. Penelitian ini tidak memperhitungkan perubahan makna, sinonim, antonim dan imbuhan
3. Data uji yang digunakan penelitian hanya abstrak dokumen Bahasa Indonesia dari jurnal penelitian tugas akhir.
4. Inputan berupa text dalam dokumen berformat docx.
5. Dataset yang digunakan sebanyak 100 dokumen abstrak dan 10 dokumen abstrak data testing.

1.4 Tujuan Penelitian

Adapun tujuan dari penelitian ini adalah untuk mengetahui tingkat kesamaan dokumen Bahasa Indonesia dengan menggunakan algoritma *Jaro-Winkler Distance*.

Tujuan yang ingin dicapai dalam pembuatan tugas akhir ini adalah:

1. Teciptanya aplikasi pendeteksi tingkat kemiripan dengan menggunakan algoritma *Jaro-Winkler Distance*.
2. Diketuainya persentase kesamaan antara abstrak yang diuji dengan algoritma *Jaro-Winkler Distance*.

1.5 Manfaat Penelitian

Manfaat yang diharapkan dari penelitian ini diharapkan dapat membantu dalam mengetahui persentase kesamaan abstrak pada dokumen Bahasa Indonesia dengan algoritma *Jaro-Winkler Distance*.