

# BAB I PENDAHULUAN

## 1.1 Latar Belakang

Setiap tahun akademis banyak mahasiswa yang membuat proposal tugas akhir untuk diajukan sebagai Tugas Akhir mahasiswa sebagai salah satu persyaratan untuk memperoleh gelar Sarjana (S1). Seringkali ditemukan tingkat kemiripan topik atau isi dokumen, antara satu dokumen proposal dengan dokumen proposal mahasiswa lainnya. Dalam dokumen entri sering terjadi duplikasi, maka dari itu dapat menggunakan metode jaccard measure agar tahu tingkat kemiripan dokumen satu dengan dokumen lainnya. Perlu adanya sistem Auto Matching untuk mengetahui tingkat kemiripan dalam proposal salah satu metode kemiripan dokumen yang banyak digunakan adalah Jaccard Measure.

Algoritma *jaccard measure* merupakan metode yang memanfaatkan nilai kemiripan dokumen. Penggunaan metode *jaccard measure* dikarenakan nilai akurasi kemiripan dari metode *jaccard measure* yang mendekati nilai keakuratan sangat tinggi. *Jaccard measure* juga merupakan salah satu teknik yang dapat dipergunakan untuk melakukan analisis dalam menghitung tingkat kemiripan suatu dokumen dengan tujuan menghasilkan perolehan yang optimal. *Jaccard measure* mengklasifikasikan yaitu dengan cara memberi pembobotan pada suatu laporan praktek kerja lapangan menghitung *Term Frequency/Inverse Document Frequency (TF/IDF)*. (Widodo et.al,2012).

Auto Matching merupakan pencocokan suatu dokumen dengan dokumen lain seperti proposal tugas akhir yang dengan berbahasa Indonesia. Dan dapat mengetahui tingkat kemiripan dokumen satu dengan dokument lainnya. Agar dapat mengetahui tingkat kemiripan dokumen tersebut. Aplikasi pendeteksi kemiripan pada dokumen dapat mempermudah proses dalam membandingkan dokumen. Selain itu dapat memberikan kata, kalimat atau paragraf yang memiliki kemiripan pada dokumen proposal tugas akhir .

Algoritma *jaccard measure* merupakan metode yang memanfaatkan nilai kemiripan dokumen. Penggunaan metode *jaccard measure* dikarenakan nilai akurasi

kemiripan dari metode *jaccard measure* yang mendekati nilai keakuratan sangat tinggi. *Jaccard measure* juga merupakan salah satu teknik yang dapat dipergunakan untuk melakukan analisis dalam menghitung tingkat kemiripan suatu dokumen dengan tujuan menghasilkan perolehan yang optimal. *Jaccard measure* yaitu dengan cara memberi pembobotan pada suatu document proposal tugas akhir menghitung *Term Frequency/Inverse Document Frequency (TF/IDF)*. Metode Jaccard Measure merupakan salah satu metode dengan konsep similarity yang digunakan untuk menghitung kesamaan dari beberapa dokumen. Dengan menggunakan metode ini, dapat memaksimalkan penilaian dengan mengukur tingkat kesamaan antar dokumen – dokumen

Oleh karena itu, berdasarkan uraian diatas maka untuk membangun model atau sistem auto matching tersebut maka penelitian ini menggunakan judul “SISTEM AUTO MATCHING DOKUMEN PROPOSAL TUGAS AKHIR BERBAHASA INDONESIA MENGGUNAKAN METODE *JACCARD MEASURE*” penulis menyelesaikan permasalahan kemiripan suatu dokumen dari sebuah proposal tugas akhir menggunakan metode *jaccard measure* yang dapat menghitung tingkat kemiripan satu dokumen dengan dokumen lainnya yang dalam hal ini dari proposal tugas akhir prodi Teknik Informatika Universitas Muhammadiyah Jember.

## **1.2 Rumusan Masalah**

Berdasarkan latar belakang yang diuraikan sebelumnya, terdapat beberapa permasalahan yang akan diangkat dalam proposal tugas akhir ini, antara lain :

1. Bagaimana mengetahui jumlah tingkat kemiripan antara dokumen satu dengan dokumen lainnya.
2. Bagaimana melakukan pra-proses menggunakan *text mining* dan perhitungan TF/IDF.
3. Bagaimana penerapan metode *jaccard measure* untuk mendapatkan derajat kemiripan antara dokumen.

### **1.3 Batasan Masalah**

Agar tidak menyimpang jauh dari permasalahan, maka Tugas Akhir ini mempunyai batasan masalah sebagai berikut :

1. Dataset berupa 34 dokumen dari bab 1, 2, dan 3 Proposal Tugas Akhir prodi Teknik Informatika pada tahun 2008-2013 berbahasa Indonesia.
2. Text mining digunakan pada dokumen berbahasa Indonesia.
3. Metode yang digunakan adalah *jaccard measure*.
4. Kemiripan dokumen berdasarkan lampiran antara bab dan keselarasan dalam data semua bab.

### **1.4 Tujuan**

Tujuan dari Tugas Akhir ini antara lain :

1. Untuk mengetahui jumlah tingkat kemiripan antara dokumen satu dengan dokumen lainnya.
2. Mengolah data atau dokumen dengan *text mining* dan menghitung bobot dengan TF/IDF.
3. Untuk menerapkan metode *jaccard measure* agar mendapatkan derajat kemiripan antara dokumen.

### **1.5 Manfaat**

Manfaat dari penelitian tugas akhir yang dilakukan oleh penulis adalah sebagai berikut :

1. Mengetahui jumlah tingkat kemiripan satu proposal dengan proposal lainnya.
2. Tingkat kemiripan yang tinggi mengidentifikasi plagiasi.
3. Menghindari kesamaan yang terjadi antara satu dokumen proposal dengan dokumen lainnya.