

BAB I

PENDAHULUAN

1.1 Latar Belakang

Skripsi adalah sebuah karya ilmiah yang dalam penulisannya harus mengacu pada ketentuan-ketentuan penulisan ilmiah seperti penggunaan bahasa yang baku dan efisien, kutipan yang harus ditulis rujukannya, serta adanya kesimpulan yang dibuat berdasarkan penalaran yang mengikuti hukum logika khususnya dalam bidang psikologi (Penyusun, 2014). Skripsi bertujuan agar mahasiswa mampu menyusun dan menulis suatu karya ilmiah, sesuai dengan bidang ilmunya.

Universitas Bina Darma Palembang merupakan perguruan tinggi swasta (PTS) yang berlokasi di Jalan Jenderal A. Yani Palembang, besarnya kampus tersebut membuat banyak jurusan yang membuat banyak skripsi yang dikeluarkan setiap tahunnya. Terutama di jurusan Teknik Informatika setidaknya dari tahun 2013 sampai dengan tahun 2016 hampir 400 judul skripsi yang sudah dibuat dan disimpan melalui *database* Universitas tersebut dan skripsi – skripsi yang telah dibuat oleh mahasiswa tersebut dapat tersimpan ke dalam media penyimpanan digital baik dalam bentuk proposal ataupun jurnal. Pada data skripsi dapat dikelompokkan berdasarkan kemiripan tema, judul, objek maupun metode penelitian. Oleh sebab itu banyak

mahasiswa yang mengambil penelitian yang mirip dari tema, judul dan isi dari penelitian sebelumnya.

Untuk mengatasi hal tersebut maka diperlukannya klasifikasi, Klasifikasi adalah suatu proses pengkategorian yang dilakukan terhadap sekumpulan dokumen. Klasifikasi sangat penting untuk kemudahan pengguna dalam melakukan pencarian dokumen (Indriani, 2014). Tujuan klasifikasi sendiri ialah untuk memisahkan sekumpulan dokumen ke dalam beberapa kelompok atau kelas dengan menilai kemiripan antar dokumen dari segi konten. Seperti kita ketahui skripsi yang dihasilkan dari satu jurusan untuk satu waktu kelulusan cukup banyak, guna klasifikasi sendiri untuk mempermudah pengkelompokan persamaan antar judul satu dengan judul lainnya sehingga dapat diketahui kemiripan yang dimiliki antar judul satu dengan judul lainnya. Maka dari itu dari hasil pengelompokan dapat menunjukkan kapan mahasiswa menentukan penelitian dalam mengambil materi yang sama.

Pengelompokan yang biasanya berbentuk teks dapat dilakukan dengan *text mining*. *Text mining* yang tujuannya adalah untuk menemukan kata-kata yang dapat mewakili isi dari dokumen sehingga dapat dilakukan analisa keterhubungan antar dokumen (Han, Kamber, & Pei, 2006). Terdapat beberapa metode *text mining* salah satunya adalah *clustering*. *Clustering* adalah suatu metode analisa data untuk memecahkan masalah pengelompokan data. Kategori algoritma *clustering* yang banyak dikenal adalah *Hierarchical Agglomerative*. Pada algoritma *clustering*, data

akan dikelompokkan menjadi cluster-cluster berdasarkan kemiripan satu data dengan yang lain. Prinsip dari *clustering* adalah memaksimalkan kesamaan antar anggota satu cluster dan meminimumkan kesamaan antar anggota cluster yang berbeda. *Hierarchical Clustering* adalah salah satu algoritma *clustering* yang dapat digunakan untuk meng-cluster dokumen (*document clustering*). Hasil keseluruhan dari algoritma *hierarchical clustering* secara grafik dapat digambarkan sebagai tree, yang disebut dengan dendogram. *Tree* ini secara grafik menggambarkan proses penggabungan dari cluster-cluster yang ada, sehingga menghasilkan cluster dengan *level* yang lebih tinggi.

Pada penelitian ini pengelompokkan dokumen skripsi dilakukan pada Universitas Bina Darma Palembang karena jumlah data penelitiannya banyak. Semakin bertambahnya penelitian skripsi dengan mata kuliah terbatas menyebabkan semakin banyak pula mahasiswa yang mengambil penelitian yang mirip tema, objek, atau metode penelitian dengan penelitian sebelumnya.

Berdasarkan latar belakang tersebut, hal ini menjadi topik menarik untuk diangkat menjadi tema sebuah penelitian. Maka penulis tertarik untuk melakukan penelitian dengan judul **“PENERAPAN METODE AGGLOMERATIVE HIERARCHICAL CLUSTERING UNTUK KLASIFIKASI DOKUMEN SKRIPSI”**.

1.2 Rumusan Masalah

Berdasarkan latar belakang diatas maka permasalahan yang akan di bahas dalam penelitian ini adalah bagaimana kinerja *Agglomerative Hierarchical Clustering* dalam pengklasifikasi dokumen skripsi mahasiswa yang ada di Universitas Bina Darma Palembang ?

1.3 Batasan Masalah

Beberapa batasan masalah yang akan digunakan acuan dalam penelitian ini ada sebebagai berikut :

1. Penelitian memanfaatkan judul dan abstrak sebagai *keyword* yang akan diberikan bobot yang lebih besar dibandingkan kata lain dalam skripsi.
2. Dokumen skripsi yang digunakan adalah skripsi mahasiswa Teknik Informatika dari tahun 2016 sampai dengan tahun 2018 yang terdapat pada Universitas Bina Darma Palembang.
3. *Tools* yang digunakan pada penelitian ini yaitu *Rapidminer* dengan *Agglomerative Hierarchical Clustering*.
4. *Text Mining* yang digunakan dalam penelitian ini adalah *Tokenize*, *Filter Stopwords* dan *Sterm*.

1.4 Tujuan dan Manfaat Penelitian

1.4.1 Tujuan Penelitian

Penelitian ini bertujuan untuk melakukan pengklasifikasian dokumen skripsi dengan menerapkan *Agglomerative Hierarchical Clustering*.

1.4.2 Manfaat Penelitian

Adapun manfaat yang diperoleh yaitu :

1. Memahami cara pengumpulan data dan pengklasifikasian data skripsi untuk kegiatan penelitian.
2. Memahami cara menggunakan *Agglomerative Hierarchical Clustering* dengan bantuan aplikasi Rapidminer.
3. Dapat digunakan sebagai referensi untuk menambah wawasan bagi mahasiswa khususnya bagi mahasiswa Universitas Bina Darma ataupun mahasiswa lainnya yang membutuhkan informasi terkait implementasi *Agglomerative Hierarchical Clustering* untuk klasifikasi dokumen skripsi.

1.5 Waktu dan Tempat Penelitian

Waktu penelitian dilaksanakan selama 5 bulan terhitung sejak bulan April 2019 hingga bulan Agustus 2019. Lokasi penelitian dilakukan pada Universitas Bina Darma Palembang yang beralamat di Jl. Jenderal Ahmad Yani No.3, 9/10 Ulu, Seberang Ulu I, Kota Palembang, Sumatera Selatan 30111.

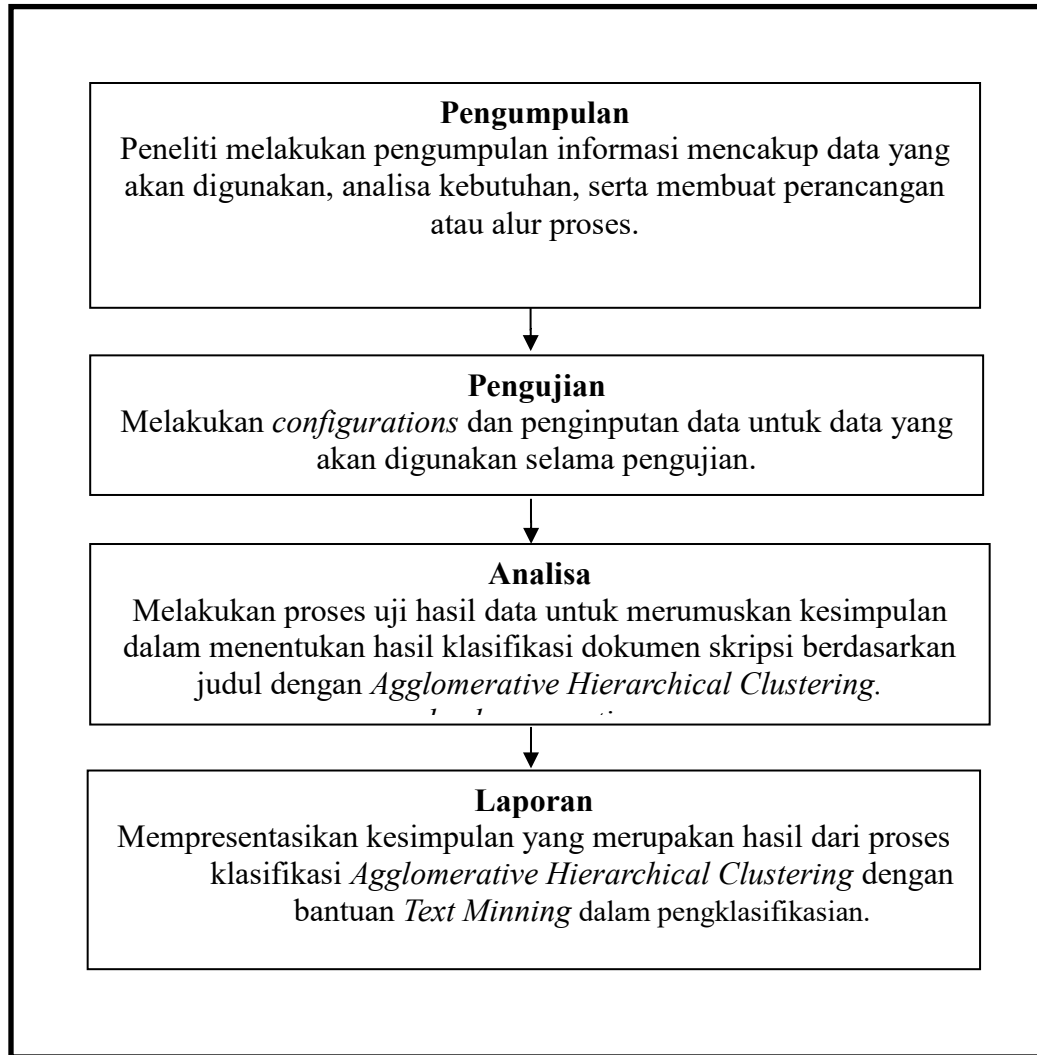
1.6 Metodologi

1.6.1 Metode Penelitian

Menurut Sulianta (2014 : 86), sebuah metode untuk mengarahkan penelitian maka haruslah sistematis dan dikerjakan dalam suatu proses yang logis dan akurat maka pada penelitian ini menggunakan 4 tahapan forensik dalam metode analisis forensik yaitu :

1. Pengumpulan (*Collection*)
2. Pengujian (*Examination*)
3. Analisa (*Filtering*)
4. Laporan (*Presentation*)

Berikut diperlihatkan pada Gambar 1.1 model kerangka penelitian yang disesuaikan dengan metode diatas :



Gambar 1.1 Kerangka Metode Penelitian

1.6.2 Metode Pengumpulan Data

Metode pengumpulan data yang dilakukan dalam penelitian ini dengan cara sebagai berikut:

1. Studi Pustaka

Data diperoleh dengan cara melakukan pencarian dan pengumpulan melalui buku-buku, data-data serta jurnal sebagai bahan referensi

2. Wawancara

Merupakan teknik pengumpulan data dengan mengadakan tanya jawab langsung dengan narasumber secara langsung, dan menjadi sumber adalah pihak dari Universitas Bina Darma.

1.7 Sistematika Penulisan

Berikut ini adalah isi dari sistematika penulisan laporan skripsi dapat dijabarkan sebagai berikut :

BAB I PENDAHULUAN

Pada bab ini berisi tentang latar belakang masalah yang diangkat, rumusan masalah, ruang lingkup, tujuan dan manfaat, metodologi penelitian, serta sistematika penulisan laporan pada skripsi ini.

BAB II TINJAUAN PUSTAKA

Pada bab ini menguraikan mengenai teori pendukung dalam penelitian dan penelitian-penelitian sebelumnya yang berkaitan dengan penelitian yang akan dilakukan. Teori pendukung yang digunakan ialah Klasifikasi, definisi Dokumen Skripsi, Mahasiswa, *Text Mining*, *Clustering*, *Agglomerative Hierarchical Clustering*, Rapidminer, serta teori mengenai objek yang diteliti.

BAB III ANALISA DAN PERANCANGAN

Pada bab ini membahas tentang analisis dan perancangan, baik dalam analisa kebutuhan maupun juga pembuatan kerangka kerja penelitian, adapun tahapan penelitian tersebut, yaitu identifikasi masalah, studi literatur, pengumpulan data, pemilihan data, melakukan penginputan data di aplikasi Rapidminer, proses selanjutnya yaitu pengujian serta penganalisisan ke tahap selanjutnya.

BAB IV HASIL DAN PEMBAHASAN

Pada bab ini membahas tentang hasil pengujian *Agglomerative Hierarchical Clustering* dengan algoritma bantuan *text mining* dan Rapidminer untuk pengklasifikasian judul skripsi jurusan Teknik Informatika di Universitas Bina Darma Palembang. Selain itu juga membahas tentang hasil dari pengklasifikasian tersebut.

BAB V KESIMPULAN DAN SARAN

Pada bab ini merupakan bagian akhir dari penulisan laporan penelitian, yang berisi kesimpulan dari hasil penelitian dan saran yang dapat digunakan untuk pengembangan penelitian selanjutnya.