

BAB I

PENDAHULUAN

1.1 Latar Belakang.

Universitas Muhammadiyah Jember (UM Jember) merupakan salah satu perguruan tinggi swasta yang berada di Kota Jember. Universitas Muhammadiyah Jember memiliki sepuluh fakultas, salah satunya adalah Fakultas Teknik yang di dalamnya terdapat lima program studi yaitu Teknik Informatika, Teknik Sipil, Teknik Mesin, Teknik Elektro dan Sistem Informasi. Fakultas Teknik menerima mahasiswa baru setiap tahunnya. Banyaknya lulusan sarjana Fakultas Teknik yang akan dihasilkan tentunya berbanding lurus dengan banyaknya penelitian tugas akhir mahasiswa yang ada di Fakultas Teknik. Tugas akhir merupakan salah satu syarat yang harus dipenuhi oleh mahasiswa untuk menyelesaikan pendidikan di suatu universitas. Ilmu yang diperoleh selama perkuliahan akan diterapkan ke dalam suatu penelitian yang nantinya akan menghasilkan suatu keluaran berupa dokumen tugas akhir dan implementasi ilmu.

Seiring berkembangnya dunia informasi dan teknologi sering muncul permasalahan umum yaitu berkembang pesatnya jumlah data dan informasi yang ada. Permasalahan ini juga terjadi pada klasifikasi dokumen pada data abstrak tugas akhir mahasiswa di Fakultas Teknik Universitas Muhammadiyah Jember. Tujuan dari penelitian ini adalah mengetahui tingkat akurasi seleksi atribut dalam klasifikasi abstrak tugas akhir dengan menggunakan algoritma Multinomial Naive Bayes dan algoritma K-Nearest Neighbour pada lima kategori yaitu bidang teknik sipil, teknik elektro, teknik informatika, teknik mesin, dan manajemen informatika. Penelitian ini membandingkan kedua algoritma tersebut, dikarenakan belum ada penelitian yang membandingkan algoritma MNB dan KNN dalam kasus klasifikasi abstrak tugas akhir.

Berdasarkan uraian di atas, maka dibutuhkan sebuah metode klasifikasi dokumen / Document Classification. Permasalahan klasifikasi dokumen bisa diselesaikan dengan beragam metode, contohnya Multinomial Naive Bayes dan K-Nearest Neighbour. Multinomial Naive Bayes adalah model pembelajaran yang berorientasi dari Teorema Bayes. Multinomial Naive Bayes sangat berguna untuk

tugas-tugas pembelajaran yang melibatkan dimensi tinggi data, seperti klasifikasi teks & web mining. Sedangkan K-Nearest Neighbour adalah salah satu metode untuk melakukan klasifikasi berdasarkan data training atau data pembelajaran yang dilihat dari jarak yang paling dekat dengan objek berdasarkan banyaknya tetangga terdekat atau disebut juga dengan nilai K. Metode ini bertujuan mengklasifikasikan objek baru berdasarkan atribut dan training sample. Diberikan suatu titik query, selanjutnya akan ditemukan sejumlah K objek atau titik training yang paling dekat dengan titik query. Nilai prediksi dari query akan ditentukan berdasarkan klasifikasi tetangga (Tri, 2010).

Berdasarkan penelitian sebelumnya, yaitu Online News Classification Using Multinomial Naive Bayes oleh Rahman dkk (2017), membahas tentang bagaimana proses klasifikasi pada berita online. Didapatkan hasil perhitungan dari Algoritma Multinomial Naive Bayes dengan TF-IDF menghasilkan akurasi rata-rata tertinggi hingga 86,62% sedangkan Multinomial Naive Bayes yang tidak menggunakan TF-IDF Atau DF-Thresholding mencapai 86,28%, Multinomial Naive Bayes dengan DF-Thresholding-TFIDF hingga 86,15% dan Multinomial Naive Bayes dengan DF-Thresholding hingga 85,98%. Pada penelitian lainnya yaitu klasifikasi dokumen menggunakan Algoritma K-Nearest Neighbour oleh Nugraha (2018) membahas tentang klasifikasi dokumen untuk mendapatkan proses pencarian sebuah data atau dokumen yang valid dan sesuai dengan yang dibutuhkan. Berdasarkan hasil pengujian, algoritma KNN memiliki rata-rata akurasi sebesar 93,94438% pada klasifikasi dokumen. Perbandingan hasil klasifikasi dari algoritma Naive Bayes dan K-Nearest Neighbour pada penelitian Yusra (2016) yang membahas klasifikasi tugas akhir di jurusan Teknik informatika UIN Suska Riau didapatkan hasil bahwa algoritma Naive Bayes dapat menghasilkan akurasi yang lebih baik yaitu sebesar 87% sedangkan algoritma K-Nearest Neighbour menghasilkan nilai akurasi 84% dengan nilai K = 3, 85% dengan nilai K = 5, 86% dengan nilai K = 7 dan 84% dengan nilai k=9.

Berdasarkan uraian di atas, Algoritma Multinomial Naive Bayes dan K-Nearest Neighbour memiliki akurasi yang cukup baik dalam klasifikasi dokumen maka dari itu dalam penelitian ini akan dilakukan perbandingan algoritma

klasifikasi pada data abstrak tugas akhir yang ada di fakultas Teknik Universitas Muhammadiyah Jember. Perbandingan terhadap kedua algoritma ini dilakukan agar dapat diketahui algoritma yang memiliki akurasi yang lebih baik terhadap pengklasifikasian abstrak tugas akhir. Dari uraian tersebut maka penulis memutuskan penelitian ini berjudul “PERBANDINGAN AKURASI ALGORITMA MULTINOMIAL NAIVE BAYES DAN K-NEAREST NEIGHBOUR DALAM KLASIFIKASI ABSTRAK TUGAS AKHIR MAHASISWA FAKULTAS TEKNIK UNIVERSITAS MUHAMMADIYAH JEMBER”.

1.2 Rumusan Masalah

Berdasarkan latar belakang di atas, maka permasalahan yang akan dibahas dalam penelitian abstrak tugas akhir ini adalah berapa tingkat akurasi dari perbandingan algoritma *Multinomial Naive Bayes* dan *K-Nearest Neighbour* dalam klasifikasi abstrak tugas akhir Fakultas Teknik Universitas Muhammadiyah Jember?

1.3 Batasan Masalah

Agar permasalahan tidak menyimpang dari tujuan penelitian, maka berikut beberapa batasan yang perlu diperhatikan, yaitu :

1. Dataset yang digunakan pada penelitian ini adalah 100 data abstrak tugas akhir yang diambil dari *Repository* Universitas Muhammadiyah Jember.
2. Abstrak yang digunakan adalah abstrak tugas akhir mahasiswa Fakultas Teknik..
3. Jumlah data abstrak pada masing-masing kelas program studi yaitu 20 abstrak.
4. Metode pembobotan yang digunakan untuk dokumen abstrak pada penelitian ini adalah *TF/IDF*.
5. *Tools* yang digunakan pada penelitian ini adalah *Jupyter Notebook*.
6. Bahasa pemrograman yang digunakan pada penelitian ini adalah Bahasa pemrograman *Phyton*.

7. Output penelitian dikategorikan menjadi 5 kelas, yaitu : Prodi Teknik Elektro, Prodi Teknik Sipil, Prodi Teknik Informatika, Prodi Teknik Mesin, dan Prodi Manajemen Informatika

1.4 Tujuan Penelitian

Adapun tujuan dari penelitian ini adalah untuk membandingkan tingkat akurasi algoritma *Multinomial Naive Bayes* dan *K-Nearest Neighbour* dalam klasifikasi abstrak tugas akhir mahasiswa Fakultas Teknik Universitas Muhammadiyah Jember.

1.5 Manfaat Penelitian

Manfaat dari penelitian ini yang dilakukan sebagai berikut :

1. Bagi penulis
Hasil dari penelitian ini diharapkan dapat memberi kesempatan untuk menambah wawasan dan pengetahuan dalam bidang klasifikasi dokumen serta hasil penelitian dapat di jadikan sebagai syarat kelulusan penulis.
2. Bagi instansi
Diharapkan dapat memberikan kontribusi secara keilmuan berupa hasil perbandingan algoritma *Multinomial Naive Bayes* dan *K-Nearest Neighbour* dalam klasifikasi abstrak tugas akhir mahasiswa Fakultas Teknik Universitas Muhammadiyah Jember.
3. Bagi peneliti lain
Hasil penelitian ini dapat digunakan sebagai referensi untuk penelitian-penelitian berikutnya di bidang klasifikasi dokumen dengan metode *text mining*.