BAB I

PENDAHULUAN

1.1 Latar Belakang

Twitter merupakan salah satu jejaring sosial yang cukup populer di kalangan masyarakat sekarang ini. Banyak dari mereka yang cukup aktif di media sosial bentukan Jack Dorsey, Evan William serta Biz Stone ini [1], untuk sekedar menuliskan kicauan pribadinya maupun mendapatkan beragam informasi dari pengguna twitter lainnya. Tak hanya sebatas itu, *Micro-Blogging* yang berdiri sejak Maret 2006 ini, kerap kali digunakan berbagai organisasi, perusahaan maupun instansi pendidikan untuk menyajikan informasi yang mereka miliki. Salah satunya Universitas Dian Nuswantoro Semarang (UDINUS), Perguruan Tinggi Swasta ternama di Jawa Tengah ini, juga memiliki akun twitter untuk memberikan informasi akademik maupun non akademik mereka kepada para mahasiswanya.

Banyaknya informasi tekstual yang disampaikan melalui akun twitter yang dimiliki Universitas Dian Nuswantor Semarang (UDINUS), menuai berbagai tanggapan dari para *followers-nya* yang kebanyakan memang Mahasiswa UDINUS. Pemberitaan mengenai UDINUS dari berbagai media juga kerap dituliskan dalam akun *twitter* milik mereka. Ada yang menuliskan tanggapan ataupun berita secara positif maupun secara negatif. Dengan *Sentiment Analysis* kita dapat mengetahui berbagai tanggapan orang lain dalam menyikapi suatu permasalahan [3].

Sentiment Analysis menitik beratkan pada penerapan pengolahan bahasa alami, komputasi secara linguistik serta analisis teks yang digunakan untuk mengidentifikasi dan mengekstrak informasi subyektif yang terkandung dalam sumber informasi [3]. Pada umumnya, *Sentiment Analysis* dibutuhkan guna mengetahui sikap seorang pembicara maupun penulis mengenai berbagai topik atau polaritas kontekstual dokumen secara menyeluruh [3]. Semua berita maupun opini mengenai UDINUS nantinya akan diklasifikasikan, agar terlihat komentar tersebut akan masuk ke dalam kelompok *sentiment* mana. Oleh karena itu dibutuhkan sebuah metode klasifikasi guna menganalisis tanggapan tersebut. Dari berbagai metode pengklasifikasian, penulis memilih metode *K-Nearest Neighbor (KNN)* dikarenakan ketangguhan algoritma ini

terhadap data *training* yang *noisy* serta efektif terhadap data latih yang besar [6], sehingga nantinya diharapkan metode ini akan sesuai dengan permasalahan ini. *K-Nearest Neighbor (KNN)* merupakan suatu metode yang menerapkan *Supervised Algorithm*, dimana algoritma ini mengklasifikasikan hasil dari sampel uji yang baru berasarkan atribut dan sampel latih [3]. Dalam pencocokannya pengklasifikasian ini tidak menggunakan metode apapun, hanya berdasarkan pada memori [3]. Diberikan titik uji yang nantinya akan ditemukan sejumlah K objek (titik latih) terdekat dengan titik uji. Pemilihan terbanyak di antara klasifikasi dari K objek digunakan pada klasifikasi. Sebagai nilai prediksi dari sample uji yang baru, Algoritma *KNN* ini menggunakan klasifikasi ketetanggaan. Jauh maupun dekatnya ketetanggaan biasanya didapat dari perhitungan berdasarkan jarak Eucledian [3].

Berdasarkan latar belakang diatas maka dibuatlah tugas akhir *sentiment analysis* mengenai Universitas Dian Nuswantoro Semarang dengan metode *K-Nearest Neighbor (KNN)* ini yang diharapkan nantinya dapat diketahui permasalahan apa saja yang ada pada Universitas Dian Nuswantoro itu sendiri berdasarkan opini yang disampaikan mahasiswa atau bahkan masyarakat umum, sehingga dapat dijadikan saran serta masukan untuk peningkatan mutu dan kualitas pelayanan Universitas Dian Nuswantoro Semarang.

1.2 Perumusan Masalah

Mengacu pada latar belakang yang sudah diuraikan diatas maka rumusan masalah dapat disusun sebagai berikut:

1. Bagaimana mengimplementasikan Algoritma *K-Nearest Neighbor* untuk menganalisa sentiment masyarakat terhadap Universitas Dian Nuswantoro Semarang,
2. Bagaimana tingkat opini masyasrakat terhadap Universitas Dian Nuswantoro Semarang berdasarkan intensitas *term frequency*,
3. Bagaimana tingkat akurasi yang dihasilkan dari percobaan pada tugas akhir *sentiment analysis* ini.

1.3 Tujuan Penelitian

Berdasarkan rumusan masalah yang telah disusun maka dapat dituliskan tujuan dari penelitian ini sebagai berikut:

1. Melakukan implementasi dan pengembangan aplikasi *sentiment analysis* terhadap Universitas Dian Nuswantoro Semarang dengan metode *K-Nearest Neighbor (KNN),*
2. Mengetahui tingkat opini masyarakat terhadap Universitas Dian Nuswantoro Semarang berdasarkan intensitas *term frequency*,
3. Mengetahui akurasi yang dihasilkan Algoritma *K-Nearest Neighbor* dari percobaan *sentiment analysis* pada tugas akhir ini.

1.4 Batasan Masalah

Adapun batasan masalah dalam penelitian ini sebagai berikut:

1. Data yang digunakan berupa *tweets* Berbahasa Indonesia mengenai Universitas Dian Nuswantoro Semarang,
2. Dokumen diklasifikasikan kedalam dua kelas *sentiment* yakni positif serta negatif,
3. Data yang digunakan dalam penelitian sejumlah 1230 *tweets* yang diambil secara periodik menggunakan Program *R GUI* dengan kata kunci terkait,
4. Aplikasi yang dibangun dalam penelitian ini menggunakan Bahasa Pemrograman R, PHP, MySQL, dan Microsoft Excel.
5. Pelabelan data dilakukan dengan proses semi-manual

1.5 Manfaat Penelitian

Diharapkan dengan adanya penelitian ini dapat memberikan manfaat diantaranya:

1. Bagi Penulis
   a. Menerapkan Ilmu yang telah diperoleh selama menempuh pendidikan Teknik Informatika di Fakultas Ilmu Komputer Universitas Dian Nuswantoro Semarang serta menyelesaikan tugas akhir sebagai syarat kelulusan,
   b. Mendapatkan pemahaman tentang perancangan serta pengembangan penelitian *Twitter data sentiment analysis* mengenai Universitas Dian Nuswantoro Semarang dengan metode *K-Nearest Neighbor (KNN).*

2. Bagi Universitas
   a. Sebagai saran serta masukan untuk meningkatkan mutu dan kualitas pelayanan Universitas Dian Nuswantoro Semarang.

b. Untuk menambah bahan bacaan pustaka kampus mengenai *sentiment analysis* dengan menggunakan *twitter data.*