

BAB I

PENDAHULUAN

1.1. Latar Belakang

Transportasi adalah kegiatan pemindahan barang (muatan) dan penumpang dari suatu tempat ke tempat yang lain [1]. Terdapat berbagai jenis transportasi, baik roda dua maupun roda empat. Transportasi umum masih banyak digunakan oleh masyarakat sebagai sarana untuk menuju suatu tempat, beberapa jenis transportasi berbasis aplikasi yaitu ojek dan taksi *online*, dengan adanya transportasi *online* pengguna jasa tidak perlu repot menunggu lama untuk memanggil ojek atau taksi, cukup lewat aplikasi di *smartphone* pengguna jasa dapat memesan salah satu dari transportasi *online* ini, dengan menggunakan saran pengemudi/*driver* yang paling dekat dengan posisi pengguna yang telah ada di aplikasi. Ini menjadi keunggulan dari transportasi *online* bila dibandingkan dengan transportasi konvensional.

Dengan berkembangnya jasa transportasi *online*, banyak masyarakat menjadi sering membicarakannya berupa pendapat dan lainnya melalui berbagai media masa, salah satunya adalah media sosial Twitter. Perusahaan transportasi *online* memiliki akun twitter resmi untuk memberikan informasi *up-to-date* tentang layanan maupun *tweet* opini yang diberikan masyarakat terhadap jasa transportasi *online*, seperti masalah aplikasi, pelayanan dan pengemudi/*driver*.

Dwi Ardiansyah sebagai *Head Of Business Development Twitter South East Asia and Australia* mengatakan Indonesia merupakan salah satu negara dengan jumlah pengguna Twitter terbesar di dunia. Dari data yang dirilis Twitter Indonesia pada akhir 2016 lalu, disebutkan bahwa 77 persen pengguna Twitter di Indonesia merupakan pengguna aktif. Selain itu, pengguna Twitter di Indonesia juga termasuk

yang paling cerewet. Hal tersebut dapat dilihat dari jumlah *tweet* yang dihasilkan sepanjang 2016 yang mencapai 4,1 miliar *tweet* [2]. *Tweet* adalah teks status pengguna yang digunakan untuk memberikan informasi di Twitter, *tweet* bisa digunakan pengguna untuk memberitahu tentang apa yang sedang dilakukan atau dirasakan, berbagi informasi, dan pelaporan berita. Apabila diteliti lebih lanjut terhadap *tweet* tersebut maka akan didapatkan sebuah sentimen yang apabila dikumpulkan akan dapat dibuat kesimpulan jasa transportasi *online* manakah yang memiliki sentimen baik di Twitter. Analisis sentimen adalah studi komputasi dari opini-opini, sentimen, serta emosi yang diekspresikan dalam teks [3].

Ada berbagai macam teknik klasifikasi dokumen, di antaranya adalah *Naïve Bayes Classifier* dan *K-Nearest Neighbor* (KNN) dengan tingkat keakuratan yang baik. Penelitian menggunakan klasifikasi dokumen juga sudah dilakukan oleh beberapa peneliti. Pada kasus pengklasifikasian kualitas batik tulis rata-rata keberhasilan sistem klasifikasi dengan menggunakan metode KNN adalah 99.11% sedangkan dengan metode NCC keberhasilannya hanya mencapai 86.44% [4]. Penelitian yang dilakukan terhadap analisis sentimen review produk *Appstore* menggunakan metode *K-Nearest Neighbor* menghasilkan akurasi mencapai 74.50% dan nilai AUC sebesar 0.825 [5]. Hasil yang didapatkan dari analisis sentimen pada jasa ojek *online* dengan *Naïve Bayes* memperoleh akurasi ketepatan 80% [6]. Hasil pengamatan polaritas sentiment masyarakat terhadap pasangan calon presiden dan wakil presiden menggunakan metode NBC menunjukkan pasangan Joko Widodo – Jusuf Kala mendapatkan total 37.6% sentimen positif, 34.4% sentimen negatif, dan 27.9% sentimen netral [7]. Hasil penggunaan algoritma klasifikasi data mining pada prediksi penyakit kulit menunjukkan *Algoritma Naïve*

Bayes memberikan nilai persis sebesar 98.3% sedangkan SVM memiliki nilai persis 98.2% selisih 0.1%, jika dibandingkan berdasarkan Nilai, maka *Naïve Bayes* dianggap sebagai algoritma klasifikasi yang lebih baik[8]. Atas dasar tersebut, maka dalam penelitian ini peneliti mengambil judul **“Perbandingan Akurasi Algoritma *Naïve Bayes Classifier* dengan *K-Nearest Neighbor* (KNN) untuk Analisis Sentimen Masyarakat Terhadap Transportasi *Online* pada Twitter”**.

1.2. Perumusan Masalah

Berdasarkan latar belakang yang telah dijelaskan, rumusan permasalahan yang akan diselesaikan dalam penelitian ini yaitu :

1. Bagaimana klasifikasi sentimen negatif dan sentimen positif terhadap data Twitter dengan menggunakan metode *Naïve Bayes Classifier* dan *K-Nearest Neighbor* (KNN) ?
2. Bagaimana tingkat akurasi metode *Naïve Bayes Classifier* terhadap data Twitter pada pembentukan sentimen analisis dibandingkan dengan metode *K-Nearest Neighbor* (KNN) ?

1.3. Tujuan dan Manfaat

Penelitian ini dilakukan dengan tujuan sebagai berikut :

1. Melakukan klasifikasi terhadap data Twitter dengan menggunakan metode *Naïve Bayes Classifier* dan *K-Nearest Neighbor* (KNN).
2. Menguji akurasi metode yang digunakan yaitu *Naïve Bayes Classifier* dalam sentimen analisis dan membandingkannya dengan metode *K-Nearest Neighbor* (KNN).

Sedangkan manfaat dari kegiatan penelitian ini adalah sebagai berikut :

1. Mengetahui hasil klasifikasi analisis sentimen positif dan negatif metode *Naïve Bayes Classifier* dan *K-Nearest Neighbor (KNN)*.
2. Mengetahui tingkat akurasi antara metode *Naïve Bayes Classifier* dan *K-Nearest Neighbor (KNN)* pada persoalan analisis sentimen.

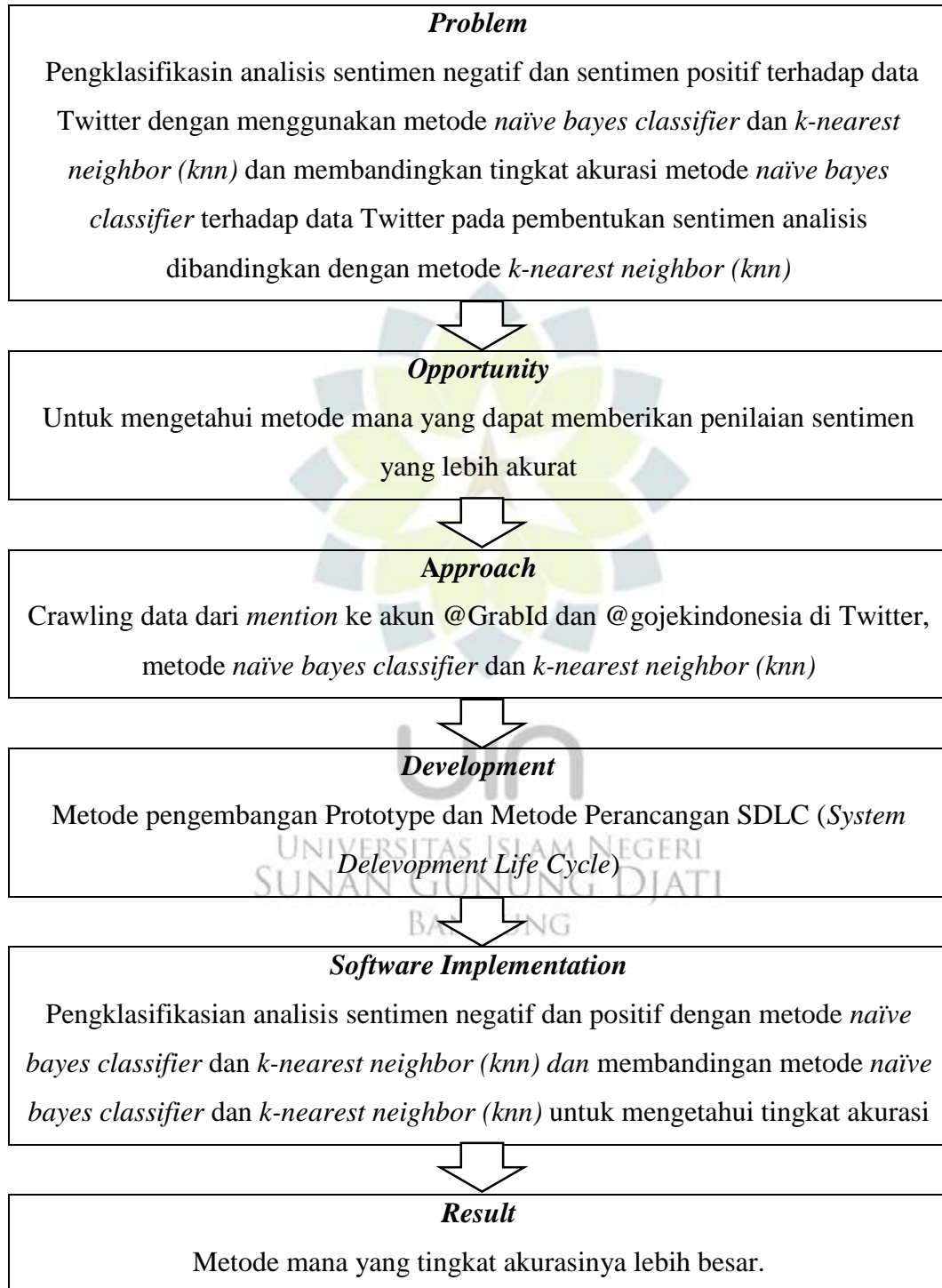
1.4. Batasan Masalah

Berdasarkan rumusan masalah yang telah ditetapkan, maka batasan masalah penelitian ini adalah sebagai berikut :

1. Dalam penelitian ini metode *text mining* yang digunakan adalah *Naïve Bayes Classifier* dan *K-Nearest Neighbor (KNN)*.
2. Data yang di analisa hanya bersumber dari *tweet* mention akun @GrabID dan @gojekindonesia.
3. Hanya menggunakan sentimen positif dan sentimen negatif.
4. Hanya menggunakan teks berbahasa Indonesia.
5. Pelabelan kata positif dan negatif secara manual oleh pakar atau admin sistem.

1.5. Kerangka Pemikiran

Kerangka Pemikiran untuk sistem yang akan dikembangkan dapat dilihat pada **Gambar 1.1** di bawah ini.



Gambar 1.1 Kerangka Pemikiran

1.6. Metodologi Penelitian

Metode pengembangan perangkat lunak yang digunakan untuk penelitian ini adalah model prototipe. Prototipe merupakan suatu metode dalam pengembangan sistem yang menggunakan pendekatan untuk membuat sesuatu program dengan cepat dan bertahap sehingga segera dapat dievaluasi oleh pengguna [9].

Dengan model prototipe ini perancang dan pengguna bertemu untuk mendefinisikan secara obyektif dengan perangkat lunak, mengidentifikasi kebutuhan yang diinginkan. Kemudian dilakukan perancangan kebutuhan sistem setelah mendengarkan keinginan pengguna. Setelah semuanya selesai pelanggan menguji coba kebutuhan pengembangan perangkat lunak yang telah di buat.

1.7. Sistematika Penulisan

Sistematika penulisan yang digunakan dalam penyusunan laporan ini dibedakan menjadi beberapa sub bab sebagai berikut :

BAB I PENDAHULUAN

Bab yang berisikan latar belakang permasalahan, perumusan masalah, tujuan dan manfaat penelitian, batasan masalah, metodologi pengerjaan tugas akhir dan sistematika penulisan.

BAB II LANDASAN TEORI

Pada bab ini berisikan konsep dasar yang berkaitan dengan penelitian yang diangkat. Teori atau konsep dasar dijadikan acuan untuk kegiatan penelitian, sehingga penelitian menjadi tepat dan terarah.

BAB III ANALISIS PERANCANGAN SISTEM

Bab yang berisikan pemodelan sistem usulan dengan menggunakan diagram perancangan sistem *Unified modelling Language (UML)*, *Conceptual Data Model (CDM)*, *Physical Data Model (PDM)* dan desain tampilan antar muka sistem.

BAB IV IMPLEMENTASI DAN PENGUJIAN SISTEM

Bab yang berisikan hasil pengkodean sistem yang ditampilkan dalam bentuk gambar sistem yang telah dijalankan dan diuji coba.

BAB V PENUTUP

Pada bab ini berisikan berisikan kesimpulan akhir penelitian dan saran yang dapat dijadikan untuk pengembangan sistem pada penelitian lain.

