

BAB 1

PENDAHULUAN

1.1. Latar Belakang

Perkembangan teknologi telah mendorong banyak pebisnis hotel untuk menawarkan produk dan jasa mereka melalui internet, diantaranya menawarkan fasilitas reservasi secara online. Terdapat beberapa hal yang menjadi pertimbangan pelanggan dalam memilih hotel diantaranya adalah ulasan. Ulasan tersebut juga digunakan oleh hotel untuk evaluasi beberapa aspek yang dirasa kurang baik. Tetapi, karena banyaknya ulasan hotel menyebabkan banyak polaritas yang berbeda sehingga sulit menentukan polaritas sebuah ulasan. Maka dibutuhkan sistem yang dapat melakukan analisis sentimen pada sebuah ulasan dengan cara menentukan sentimen berdasarkan aspeknya.

Analisis sentimen adalah suatu bidang studi yang menganalisis dan mengidentifikasi sebuah opini, sentimen pujian, emosi seseorang terhadap entitas seperti produk, layanan, organisasi, kejadian, masalah termasuk ke dalam kategori positif, negatif, atau netral. Analisis sentimen dibagi menjadi tiga tingkat, yaitu level dokumen, level kalimat, dan level aspek dan entitas [1], [2]. Analisis sentimen level aspek menunjukkan performa yang lebih baik dibandingkan level dokumen dan level kalimat [3]. Hal ini dikarenakan, analisis sentimen menganalisis setiap ulasan yang diberikan konsumen dengan lengkap tidak hanya opini mengenai produk secara keseluruhan, namun mengulas beberapa aspek yang memiliki sentimen berbeda [4].

Ada banyak persiapan yang harus dilakukan untuk melakukan analisis sentimen diantaranya adalah memilih *classifier* yang akan digunakan. Pada penelitian ini, *classifier* yang dipilih adalah Naïve Bayes. Sebuah metode sederhana yang dikembangkan berdasarkan aturan Bayes dengan melihat kondisi-kondisi yang ada dan peluang-peluang setiap kondisinya. Salah satu hal yang menjadi masalah dalam analisis sentimen adalah kurang diperhatikannya gabungan dua kata atau lebih untuk hasil klasifikasi lebih maksimal [5].

Pada penelitian yang dilakukan oleh Shima Fanissa, M. Ali Fauzi, Sigit Adinugroho [5] yang berjudul “Analisis Sentimen Pariwisata di Kota Malang Menggunakan Metode Naive Bayes dan Seleksi Fitur Query Expansion Ranking” menghasilkan akurasi tertinggi yaitu 86,6 % dengan seleksi fitur 75%, sedangkan dengan seleksi fitur 25% dan 50% menghasilkan akurasi 83,3%. Penelitian ini menyarankan untuk memperhatikan gabungan dua kata atau lebih (bigram, trigram, n-gram) supaya hasil klasifikasi lebih sempurna. Penelitian ini juga menyebutkan bahwa proses seleksi fitur berpengaruh pada nilai akurasi yang didapatkan.

Berikutnya penelitian terkait dengan penggunaan fitur N-Gram yang dilakukan oleh Fitriyani [6] dengan judul “Penerapan Word N-Gram Untuk Sentiment Analysis Review Menggunakan Metode Support Vector Machine (Studi Kasus: Aplikasi Sambara)”. Penelitian ini melakukan perbandingan dalam penggunaan fitur N-Gram. Nilai akurasi Bi-Gram dengan jumlah data 1332 menghasilkan akurasi lebih tinggi dibandingkan dengan nilai akurasi Uni-Gram. Akurasi Bi-Gram mencapai 88,89% sedangkan Uni-Gram 88,21%. Nilai akurasi yang dihasilkan oleh Tri-Gram dengan jumlah data 1200 lebih tinggi dibandingkan dengan Uni-Gram yaitu 88,50% sedangkan Uni-Gram 87,58%. Hal tersebut membuktikan bahwa fitur N-Gram dapat digunakan untuk gabungan kata dan dapat meningkatkan akurasi.

Berdasarkan latar belakang yang telah diuraikan dan beberapa penelitian yang telah dilakukan, maka pada penelitian ini akan dilakukan analisis sentimen berdasarkan aspek pada ulasan hotel di dengan menggunakan metode Naive Bayes dan seleksi fitur Query Expansion Ranking dengan ekstraksi fitur N-Gram.

1.2. Rumusan Masalah

Berdasarkan permasalahan yang ada pada latar belakang di atas dapat dirumuskan masalahnya yaitu, berapa nilai akurasi dari metode Naive Bayes dan Query Expansion Ranking dengan N-Gram dalam melakukan analisis sentimen berbasis aspek terhadap ulasan hotel?

1.3. Maksud dan Tujuan Penelitian

Maksud dari penelitian ini adalah untuk melakukan analisis sentimen pada ulasan hotel dengan menggunakan Naïve Bayes, seleksi fitur Query Expansion Ranking, dan N-Gram. Tujuan dari penelitian ini adalah untuk mengetahui nilai akurasi dari metode Naïve Bayes dan Query Expansion Ranking dengan N-gram dalam melakukan analisis sentimen berbasis aspek terhadap ulasan hotel.

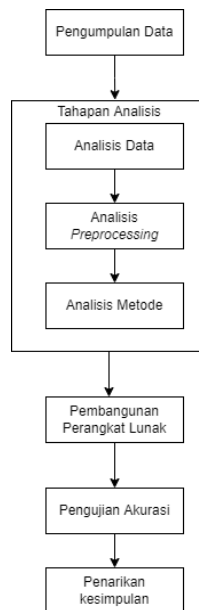
1.4. Batasan Masalah

Adapun Batasan masalah dalam penelitian ini adalah sebagai berikut :

1. Ulasan yang digunakan hanya ulasan yang berbahasa Indonesia pada beberapa hotel di Indonesia.
2. Aspek yang akan digunakan adalah layanan, fasilitas, kebersihan, makanan, dan yang tidak termasuk pada aspek yang ditetapkan.
3. Metode klasifikasi yang digunakan adalah Naïve Bayes.
4. Menggunakan ekstraksi fitur N-Gram dan Pembobotan TF.
5. Data yang telah diolah akan diklasifikasikan ke dalam kelas positif, negatif, dan netral.
6. Pelabelan akan dilakukan secara manual.
7. Stemming menggunakan *library* Sastrawi.

1.5. Metodologi Penelitian

Metodologi penelitian yang akan dilakukan pada penelitian dapat dilihat pada Gambar 1.1.



Gambar 1.1 Metode Penelitian

1.5.1. Metode Pengumpulan Data

Metode pengumpulan yang digunakan dalam penelitian ini adalah sebagai berikut:

1. Studi Literatur

Studi literatur merupakan teknik pengumpulan data dengan cara mengumpulkan jurnal maupun buku yang berkaitan dengan analisis sentimen berdasarkan aspek.

2. Pengumpulan Dataset

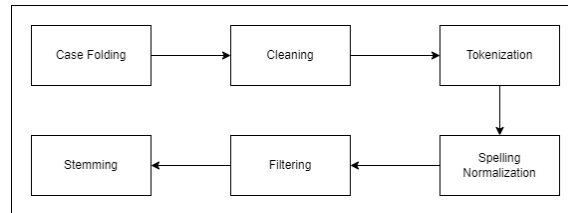
Ikter.cbeberapa hotel yang didapatkan dengan menggunakan *Web Scraping*.

1.5.2. Metode Tahapan Analisis

Tahapan analisis pada penelitian ini adalah sebagai berikut:

1. Analisis Data, merupakan proses untuk menganalisis pola kalimat dari data yang didapatkan.
2. Analisi *Preprocessing*, merupakan tahap untuk memproses data agar data menjadi lebih terstruktur dan lebih siap untuk diolah. Langkah yang dilakukan pada tahap preprocessing terdiri dari

case folding, cleaning, tokenization, spelling normalization, filtering, dan stemming.

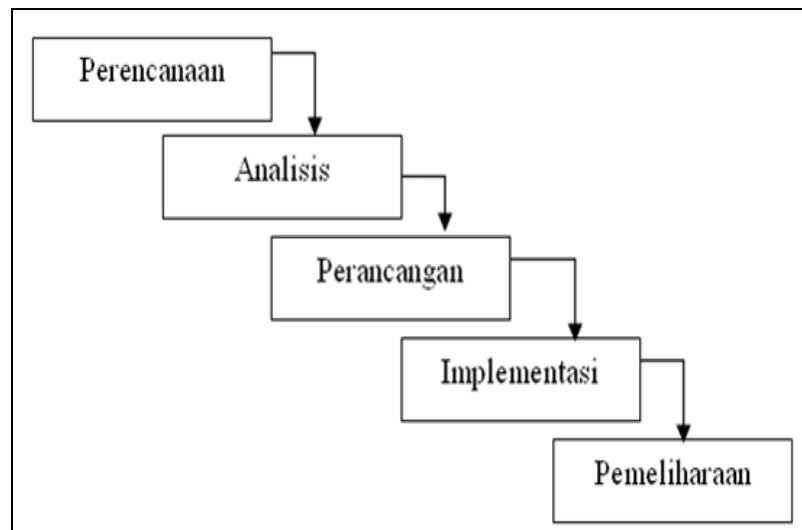


Gambar 1.2 Preprocessing

3. Analisis Metode, pada tahap ini akan dilakukan klasifikasi dengan menggunakan metode Naïve Bayes. Meliputi, penggunaan Fitur N-Gram, Pembobotan TF, dan Seleksi Fitur menggunakan Query Expansion Ranking.

1.5.3. Metode Pembangunan Perangkat Lunak

Metode pembangunan perangkat lunak yang digunakan dalam penelitian ini adalah model *waterfall*. Tahapan-tahapan pada metode *Waterfall* dapat dilihat pada Gambar 3.



Gambar 1.3 Metode Pembangunan Perangkat Lunak

1. Perencanaan

Pada tahap ini dibuat rencana tentang kebutuhan dari perangkat lunak yang akan dibuat agar perangkat lunak yang dibuat sesuai dengan kebutuhan pengguna.

2. Analisis

Pada tahap dilakukan analisa kebutuhan apa saja yang diperlukan dalam pembangunan aplikasi secara lengkap berdasarkan masalah yang telah disampaikan.

3. Desain

Pada tahap ini setelah selesai menganalisa kebutuhan secara lengkap, pengembang melakukan pembuatan desain sistem secara keseluruhan dan menentukan alur dari perangkat lunak yang akan dibangun. Desain yang dibangun berupa alur data dan juga tampilan antar muka.

4. Implementasi

Pada tahapan ini semua data informasi yang telah dikumpulkan dan didesain oleh pengembang akan diterjemahkan kedalam bahasa pemrograman dan diintegrasikan sehingga nantinya akan terbentuk aplikasi.

5. Pemeliharaan

Proses ini merupakan tahapan perbaikan apabila terdapat kesalahan yang ditemukan pada aplikasi setelah digunakan oleh user.

1.5.4. Penarikan Kesimpulan

Pada tahapan ini dilakukan penarikan kesimpulan berdasarkan hasil dari tahapan pengujian dalam penerapan metode Naïve Bayes dan Query Expansion Ranking dengan N-Gram pada analisis sentimen berdasarkan aspek.

1.6. Sistematika Penulisan

Sistematika penulisan yang digunakan pada penelitian ini adalah sebagai berikut:

BAB 1 PENDAHULUAN

Bab 1 menguraikan tentang latar belakang yang terdiri dari permasalahan yang berkaitan dengan penelitian, rumusan masalah, maksud dan tujuan penelitian, batasan masalah, metode penelitian, dan sistematika penulisan.

BAB 2 LANDASAN TEORI

Bab 2 membahas dan menjelaskan mengenai landasan teori yang mendukung penulisan skripsi.

BAB 3 ANALISIS DAN PERANCANGAN

Bab 3 berisikan tentang analisis terhadap sistem yang akan dibangun meliputi analisis masalah, analisis data masukan, dan analisis pada pemrosesan serta klasifikasi.

BAB 4 IMPLEMENTASI DAN PENGUJIAN

Bab ini berisikan hasil implementasi dari analisis dan pengujian terhadap sistem yang sudah dibangun.

BAB 5 KESIMPULAN DAN SARAN

Bab ini berisi kesimpulan mengenai hasil penelitian dan saran yang diharapkan akan berguna untuk penelitian lebih lanjut pada kasus yang sama.