

BAB I

PENDAHULUAN

1.1 Latar Belakang

Saat ini bisnis perdagangan berkembang sangat pesat dan sudah memasuki era digital, dimana transaksi jual beli bisa dilakukan secara online dimanapun dan kapanpun. Seiring dengan berkembangnya transaksi jual beli online, banyak kostumer atau pembeli yang melakukan ulasan ataupun *review* kepada penjual di website penjualan mereka, akan tetapi karena banyaknya ulasan yang diberikan oleh kostumer membuat penjual terkadang tidak memberikan respon balasan kepada ulasan yang diberikan dan juga terkadang penjual tidak bisa memberikan kesimpulan yang tepat, dikarenakan untuk menyimpulkan ulasan yang begitu banyak dari pembeli itu sendiri membutuhkan waktu yang tidak sedikit. Penjual akan kesulitan untuk mengetahui apa saja hal yang sering dibahas dari produk mereka dan apakah reaksi positif atau negatif yang diberikan oleh kostumer.

Dalam mengatasi masalah tersebut maka dibutuhkan aplikasi yang bisa membantu untuk menganalisis sentiment yang ada pada setiap ulasan, Analisis sentiment sendiri merupakan studi komputasi dari opini-opini, sentiment, serta emosi yang diekspresikan dalam teks. Tugas dari analisis sentiment sendiri adalah mengelompokkan teks yang ada didalam dokumen, kalimat, atau pendapat yang memiliki aspek positif atau negative. Salah satu teknik yang paling banyak digunakan untuk analisis sentiment adalah *naïve bayes classifier* (NBC), dimana NBC adalah teknik pembelajaran yang berbasis probabilistic dengan metode yang sederhana tetapi memiliki akurasi serta performa yang tinggi dalam pengklasifikasian teks.

Naïve Bayes Classifier pada Analisis sentiment sendiri digunakan untuk penggolongan kata yang ada pada data ulasan yang telah dikumpulkan untuk di cari pada kelas manakah kata tersebut akan digolongkan, apakah kelas positif atau negative. Untuk meningkatkan tingkat akurasi dan efisiensi pada *naïve bayes*, maka

dibutuhkan metode untuk memfiltering kata fitur yang akan diklasifikasi, ada beberapa metode filtering fitur yang sering digunakan untuk klasifikasi yaitu *Information Gain*, *Chi-Square*, dan *Frequency Based*. Pada penelitian ini penulis akan menggunakan *Chi-Square* sebagai metode untuk filtering pada fitur.

Chi-Square merupakan metode filtering fitur yang sering digunakan untuk klasifikasi data atau dokumen dalam jumlah yang besar. *Chi-Square* sendiri akan mencari kata-kata yang paling dependent terhadap suatu kelas, semakin tinggi tingkat dependensi fitur terhadap suatu kelas maka semakin baik pula fitur tersebut digunakan untuk klasifikasi pada *naïve bayes*.

Berdasarkan hal diatas maka penulis memutuskan untuk membuat suatu aplikasi analisis sentiment terhadap toko online menggunakan seleksi fitur *Chi-Square* dan *Naïve bayes Classifier*, yang dimana diharapkan aplikasi ini akan dapat membantu para penjual untuk menganalisa ulasan-ulasan terhadap toko / produk mereka sehingga dapat membantu mengembangkan produk dan toko mereka menjadi lebih baik.

1.2 Rumusan Penelitian

Berdasarkan latar belakang yang telah dipaparkan diatas, maka diperoleh rumusan masalah adalah sebagai berikut :

1. Bagaimana menentukan fitur dari sebuah ulasan dengan *Chi-Square*?
2. Bagaimana menentukan penilaian positif dan negatif, pada setiap ulasan dengan *Naïve Bayes*?
3. Bagaimana menentukan penilaian akurasi, *presisi*, dan *recall* pada hasil kesimpulan menggunakan *confusion matrix*?
4. Bagaimana merancang bangun system analisis sentiment pada review toko online menggunakan metode *Naïve Bayes* dan seleksi Fitur *Chi-Square*?

1.3 Batasan Penelitian

Batasan yang digunakan dalam penelitian ini agar pembahasan yang dilakukan tidak terlalu luas yaitu :

1. Tidak membahas tentang keamanan sistem.
2. dokumen *ulasan* menggunakan bahasa indonesia.
3. penelitian ini dititik beratkan pada penerapan metode klasifikasi *Naïve bayes* dengan seleksi fitur *Chi-Square*.
4. penelitian ini lebih menitik beratkan pada *sentimen* negatif dan positif.

1.4 Tujuan Penelitian

Adapun tujuan penelitian yang ingin dicapai oleh penulis adalah sebagai berikut:

1. Menentukan fitur dari sebuah *ulasan* menggunakan *Chi-Square*.
2. Menentukan penilaian positif dan negatif menggunakan *Naïve bayes Classifier*.
3. Menentukan penilaian akurasi, *presisi*, dan *recall* menggunakan confusion Matrix.
4. Merancang bangun *system* analisis *sentiment* pada *review* toko online menggunakan metode *Naïve Bayes* dan seleksi Fitur *Chi-Square*.
5. Sebagai syarat untuk menyelesaikan Strata-1 jurusan Informatika Universitas Amikom Yogyakarta.

1.5 Manfaat Penelitian

Adapun manfaat yang diharapkan dari penelitian adalah sebagai berikut:

- a. Manfaat yang didapatkan bagi penulis adalah :
 1. Menambah pengetahuan dalam bidang sentiment analisis menggunakan *Chi-Square* dalam pencarian Fitur dan *Naïve Bayes* dalam penilaian sentiment positif dan negative.
 2. Sebagai bahan penyampaian ilmu yang telah didapatkan penulis kepada masyarakat.
- b. Manfaat yang didapatkan bagi para pemilik toko online.
 1. Sistem dapat membantu dalam memberikan kesimpulan dari apa saja hal yang sering dibahas dari ulasan setiap kostumer dan penilaian positif atau negatif yang diberikan oleh para kostumer.
 2. Dapat membantu mengembangkan toko online menjadi lebih baik.

1.6 Metodologi Penelitian

Penulis menjabarkan cara-cara memperoleh data-data yang akan digunakan Untuk kebutuhan penelitian.

1.6.1 Metode Pengumpulan data

Penulis akan mengambil ulasan-ulasan yang ada di toko online menggunakan aplikasi parsehub yang kemudian keluaran dari aplikasi tersebut akan di inputkan ke dalam sistem, berdasarkan toko atau barang yang memiliki lebih dari 100 ulasan dari kostumer.

1.6.2 Analisis dan Perancangan

Pada tahap ini penulis melakukan analisis terhadap metode perhitungan *Chi-Square* bagaimana cara metode tersebut melakukan perhitungan untuk mendapatkan fitur yang akan diklasifikasi oleh *naïve bayes*, dan juga cara perhitungan metode *naïve bayes* untuk dapat melakukan klasifikasi pada fitur yang

telah didapatkan. Kemudian penulis akan melakukan perancangan kerangka system atau pemodelan system menggunakan diagram-diagram UML seperti use case diagram, Sequence diagram, dan class diagram.

1.6.3 Metode Implementasi

Pada tahap ini penulis akan membuat system menggunakan PHP dan MySQL, *Chi-Square* untuk menentukan fitur yang berguna untuk meningkatkan efisiensi pada metode *Naïve bayes* ketika melakukan proses klasifikasi. *Naïve Bayes* untuk menentukan sentiment positif atau negatif pada dokumen yang telah dikumpulkan.

1.6.4 Pengujian

Pada tahap ini penulis akan melakukan pengujian evaluasi terhadap system yang telah dibuat dengan menggunakan *confussion matrix*.

1.7 Sistematika Penulisan

BAB I PENDAHULUAN

Bab ini menjelaskan mengenai deskripsi topik kajian dan latar belakang, rumusan masalah, tinjauan pustaka, serta tujuan dari analisis sentiment pada *review* toko online menggunakan metode *Naïve Bayes* dan seleksi Fitur *Chi-Square*.

BAB II LANDASAN TEORI

Bab ini memuat tentang teori-teori yang dipakai serta yang berhubungan dengan perancangan dan pembuatan sistem.

BAB III ANALISIS DAN PERANCANGAN

Bab ini menjelaskan hasil analisis, perancangan sistem, hingga pada perancangan antarmuka aplikasi.

BAB IV IMPLEMENTASI DAN PENGUJIAN

Bab ini memuat uraian dan langkah-langkah implementasi sistem yang telah dibangun beserta dengan pengujian evaluasi sistem terhadap analisis sentiment pada *review* toko online menggunakan metode *Naïve Bayes* dan seleksi Fitur *Chi-Square*.

BAB V KESIMPULAN DAN SARAN

Bab ini berisi kesimpulan yang merupakan rangkuman hasil analisis kerja secara keseluruhan serta saran-saran penulis untuk pengembangan.

