

# BAB I

## PENDAHULUAN

### 1.1 Latar Belakang

Perkembangan teknologi yang semakin canggih telah membawa perubahan pada masyarakat Indonesia dengan adanya internet yang memungkinkan bagi seseorang untuk melakukan komunikasi dengan pihak lain tanpa di batasi waktu dan jarak. Hal ini juga yang mempengaruhi gaya hidup masyarakat Indonesia dengan adanya tren belanja online melalui internet. Dengan jumlah penduduk Indonesia yang mencapai 25 juta jiwa pada tahun 2008 dan diprediksikan akan meningkat sebesar 37% hingga tahun 2010 (Asmai Ishak, 2010). Banyaknya penjual yang menawarkan produk secara online memberi kesempatan konsumen untuk memilih sebelum memutuskan untuk membeli. Salah satunya adalah dengan membandingkan karakteristik satu situs penyedia produk dengan situs lain untuk produk yang sama. Berdasarkan kajian yang telah dilakukan (Chung and Shin, 2010).

Pada saat ini dimana toko online sedang populernya melakukan pendekatan kepada konsumen untuk menambah konsumen di toko online mereka. Hal ini dapat digunakan sebagai data untuk menganalisa sentimen terhadap toko online, dimana baik buruknya kinerja mereka langsung ditentukan oleh konsumen itu sendiri. Menurut (Liu. B, 2012) *sentiment analysis* (analisis sentimen) atau sering disebut juga dengan *opinion mining* (penambangan opini) adalah studi komputasi untuk mengenali dan mengekspresikan opini, sentimen, evaluasi, sikap, emosi, subjektifitas, penilaian atau pandangan yang terdapat dalam suatu teks yang dikemukakan oleh masyarakat.

Beberapa penelitian terkait dengan menggunakan metode naïve bayes classifier. Berdasarkan hasil yang diperoleh dapat disimpulkan bahwa kemunculan frekuensi fitur pada kategori yang diharapkan dan kategori yang tidak diharapkan memiliki peran penting dalam seleksi fitur *Chi Square*, oleh karena itu seleksi fitur *Chi Square* baik digunakan dalam penyeleksian fitur dibandingkan dengan

metode *frequency-based*. Serta pembangunan sistem analisis sentimen menggunakan metode NBC dengan Bahasa pemrograman Java memperoleh akurasi sebesar 83% dan rata-rata harmonik sebesar 90,713% (Juen Ling, I Putu Eka N. Kencana, Tjokorda Bagus Oka, 2014). Menurut (Faishal Nuruz Zuhri, Andry Alamsyah, S.Si., M.Sc, 2017) proses pengklasifikasian data uji menjadi 2 kelas, yaitu, sentimen positif dan negatif. Hasil klasifikasi data uji sangat dipengaruhi oleh model yang telah dibuat dari data latih. Dari proses evaluasi dan validasi data latih, didapatkan precision sebesar 98.42 %, recall sebesar 98.40 %, dan tingkat akurasi sebesar 98.40 %. Hal ini menunjukkan bahwa sistem dapat memisahkan data bersentimen positif dan negatif dengan sangat baik.

Menurut (Ghulam Asrofi Buntoro, 2016), Dari penelitian yang telah dilakukan, maka dapat ditarik kesimpulan bahwa Analisis sentimen tweet Bahasa Indonesia dengan tagar Hatespeech dapat membantu menentukan sentimen pada twit opini Bahasa Indonesia yang ada di Twitter. Setelah dilakukan analisis sentimen terlihat banyak tweet opini yang sebenarnya tidak masuk kategori Hatespeech tapi diberi tagar Hatespeech. Nilai akurasi tertinggi didapat dengan metode klasifikasi *Support Vector Machine (SVM)* dengan tokenisasi *unigram*, *stopword list* Bahasa Indonesia dan *emoticons*, dengan nilai rata-rata akurasi mencapai 66,6%, nilai presisi 67,1%, nilai *recall* 66,7% nilai *TP rate* 66,7% dan nilai *TN rate* 75,8%. Menurut (Syahmia Gusriani, Kartina Diah Kusuma Wardhani<sup>2</sup> dan Muhammad Ihsan Zul, 2016), Kesimpulan yang diperoleh dari penelitian ini ialah *Naïve bayes* dapat dijadikan metode klasifikasi untuk analisis sentimen dengan keakuratan 93.7%. Teknik klasifikasi dokumen dengan NBC dan pembobotan fitur metode *tf-idf* menghasilkan nilai yang pasti dan akurasi yang baik karena bobot memperkecil kemungkinan kesalahan pada pengklasifikasian, fitur yang mempunyai frekuensi tertentu dapat mempengaruhi keakuratan dalam klasifikasi bergantung pada frekuensi fitur dan dokumen yang mengandung fitur tersebut. Hasil dari klasifikasi dokumen menggunakan NBC. Pada penelitian ini dengan data *training* sebanyak 260 dokumen politik dan 222 dokumen ekonomi menggunakan 40 data *testing* menunjukkan nilai akurasi yang baik pada

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.

2. Dilarang mengemukakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

keseluruhan klasifikasi, dengan akurasi keseluruhan klasifikasi sebesar 85% (Akhmad Pandhu Wijaya & Heru Agus Santoso, 2016),

Pada penelitian analisis sentimen ini akan dibahas bagaimana mengimplementasikan analisis sentimen untuk menentukan pandangan konsumen atau publik terhadap toko online menggunakan metode klasifikasi NBC.

## 1.2 Rumusan Masalah

Berdasarkan penjelasan yang telah diuraikan pada latar belakang, maka dapat dirumuskan suatu permasalahan yaitu ;

1. Bagaimana menerapkan metode NBC pada klasifikasi sentimen terhadap review bukalapak.
2. Bagaimana mengukur tingkat akurasi pada penerapan metode NBC pada klasifikasi sentimen terhadap review bukalapak.

## 1.3 Batasan Masalah

Berdasarkan rumusan masalah diatas kemampuan peneliti dan kepentingan peneliti, maka permasalahan dibatasi pada hal-hal berikut :

1. Review toko online yang digunakan adalah bukalapak.
2. Klasifikasi sentimen yang digunakan dibagi menjadi 2 kelas yaitu : sentimen positif dan sentimen negative.
3. Data yang diambil hanya komentar yang menggunakan bahasa Indonesia.

## 1.4 Tujuan

Tujuan dari penelitian ini adalah :

1. Menerapkan metode NBC pada klasifikasi sentimen terhadap review bukalapak.
2. Menguji tingkat akurasi terhadap penerapan metode NBC pada klasifikasi sentimen pada review bukalapak.

## 1.5 Sistematika Penulisan

Penulisan dalam penelitian ini terdiri dari 6 Bab yaitu sebagai berikut:

## BAB I PENDAHULUAN

Berisi tentang deskripsi secara umum dari penelitian ini, yang terdiri dari latar belakang, rumusan masalah, batasan masalah, tujuan penelitian, dan sistematika penulisan.

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

## **BAB II LANDASAN TEORI**

Pada bab ini dibahas beberapa teori yang akan mendukung pembahasan pada bab selanjutnya.

## **BAB III METODOLOGI PENELITIAN**

Memuat mengenai tahap-tahap yang dilakukan pada penelitian berupa teknik pengumpulan data, analisa sistem, perancangan, implementasi, pengujian, kesimpulan dan saran.

## **BAB IV ANALISA DAN PEMBAHASAN**

Memuat pembahasan mengenai analisa dan pembahasan metode yang digunakan pada penelitian.

## **BAB V IMPLEMENTASI DAN PENGUJIAN**

Memuat implementasi dari hasil analisa dan perancangan serta pengujian metode yang telah diterapkan dan membandingkannya.

## **BAB VI PENUTUP**

Memuat kesimpulan dari hasil penelitian yang di implementasikan dan di uji serta saran untuk penelitian lebih lanjut.