

# 1. Pendahuluan

## 1.1 Latar Belakang

Pemanfaatan teknologi *website* merupakan salah satu teknologi yang sedang marak digunakan, baik untuk melakukan promosi, review produk tertentu maupun untuk melakukan pemasaran suatu produk dalam jual beli secara online (*e-commerce*). Berdasarkan survey yang dilakukan oleh comScore (2007) dan Horrigan (2008) ditemukan fakta bahwa 81% pengguna teknologi *website* di US menggunakan internet untuk melakukan pencarian terhadap produk yang ingin dibeli dan lebih dari 30% pengguna memberikan review terhadap sebuah produk yang dibeli [1].

Saat ini, jumlah pelanggan toko online (*onlineshop*) semakin meningkat seiring peningkatan penggunaan layanan *e-commerce* tersebut. Dalam rangka meningkatkan kepuasan pelanggan, pelanggan diijinkan untuk memberikan opini berupa saran, pengalaman atau sekedar tanggapan tentang produk didalam komentar yang disediakan pada ulasan tersebut. Opini biasanya merupakan ekspresi subjektif yang mendeskripsikan sentimen ataupun perasaan seseorang terhadap suatu ulasan. Opini tersebut dapat dimanfaatkan sebagai salah satu parameter analisis untuk melihat tingkat minat masyarakat terhadap suatu produk sehingga dapat menjadi parameter untuk menentukan jumlah produk yang akan diproduksi. Opini berupa pengalaman juga dapat membantu pelanggan lainnya untuk memutuskan pembelian suatu produk. Namun dengan peningkatan jumlah review terhadap suatu produk dapat membuat pelanggan bingung harus membaca serta menganalisis review satu per satu. Permasalahannya, review yang diberikan bersifat ambigu dimana suatu review memiliki sentiment positif dan negatif secara bersamaan.

Oleh sebab itu, diperlukan suatu aplikasi yang dapat memberikan hasil ikhtisar analisis sentimen dari opini yang diberikan reviewer terhadap produk ataupun topik. Untuk itu perlu dilakukan analisis, identifikasi, ekstraksi fitur serta pengklasifikasian ke dalam kelas ekspresi sentiment baik positif maupun negatif terhadap opini dari *reviewer*. Sebelumnya, terdapat beberapa penelitian [13], metode yang digunakan untuk mengklasifikasikan opini ke dalam kelas ekspresi sentimen antara lain : metode *unsupervised learning* dengan menghitung *Pointwise Mutual Information* (PMI) [19], *dictionary-based* atau *lexicon based* [3], dan *supervised learning* dengan metode *machine learning* seperti *Naive-Bayes Classifier*, *Maximum Entropy*, dan *Support Vector Machine* (SVM) [13]. Pada penelitian [13], metode SVM memperoleh performa yang terbaik, jika dibandingkan dengan metode machine learning lainnya.

Untuk itu pada Tugas akhir kali ini akan digunakan metode Nu-SVM yang merupakan salah metode turunan dari metode SVM. Nu-SVM digunakan untuk meminimalkan kesalahan karena Nu-SVM memiliki tambahan batas atas fraksi kesalahan marjin dan batas bawah pada fraksi vektor dukungan (yaitu kekurangan dari solusi). Hal itu dapat dilihat dari percobaan yang dilakukan pada SVM dan Nu-SVM untuk melihat performansi diantara keduanya. Dalam percobaan tersebut kita akan menggunakan parameter terbaik bagi masing-masing metode dan diperoleh hasil pengujian bahwa pada SVM kemampuan untuk

mengklasifikasikan kelas positif yaitu sebesar 86,78% untuk data yang 100% berlabel dan untuk klasifikasi kelas negatif sebesar 68,46 dengan waktu 2,62 detik. Sedangkan pada metode nu-SVM kemampuan untuk mengklasifikasikan kelas positif sebesar 85,447% dan untuk kelas negatif sebesar 76,64% dengan waktu 0,404 detik. Dengan hasil tersebut kita lihat bahwa besar performansi rata-rata pada Nu-SVM memperoleh nilai yang lebih besar daripada SVM yaitu sebesar 81.0435 dan juga waktu eksekusi yang lebih cepat dibanding SVM. Dengan demikian, perumusan Nu-SVM akan berpotensi lebih berguna daripada *Support Vector Machine* (SVM).

## 1.2 Perumusan masalah

Berdasarkan pada latar belakang di atas, masalah yang akan diuraikan dan diteliti pada Tugas Akhir ini adalah:

1. Bagaimana implementasi metode Nu-SVM untuk mengekstrak fitur yang diulas dari opini *reviewer*?
2. Bagaimana performansi Nu-SVM terhadap penentuan nilai sentiment berdasarkan parameter degree, gamma, *coefisien*, nu, kernel, rasio *labeled* data dan *unlabeled* data berdasarkan evaluasi *F1-measure*?

## 1.3 Batasan masalah

Adapun batasan masalah pada tugas akhir ini antara lain :

1. Data yang digunakan adalah data teks berbahasa Inggris.
2. Data yang digunakan berasal dari komentar *reviewer* pada web review <http://www.cnet.com> yang berfokus pada produk telepon seluler.
3. Penentuan orientasi opini menggunakan *engine Sentiword*.

## 1.4 Tujuan

Tujuan dari pembuatan Tugas Akhir ini adalah sebagai berikut :

1. Melakukan implementasi dan mengekstrak fitur produk dengan menggunakan metode Nu-SVM?
2. Menganalisis performansi Nu-SVM berdasarkan parameter degree, gamma, *coefisien*, nu, kernel, rasio *labeled* data dan *unlabeled* data berdasarkan evaluasi *F1-measure*.

## 1.5 Metodologi penyelesaian masalah

Metodologi penyelesaian yang digunakan dalam mengerjakan sistem ini terdiri dari beberapa tahapan, diantaranya :

1. Tahap Studi Literatur dan Identifikasi Masalah

Pada tahap ini, akan dilakukan pengumpulan literatur, informasi yang berhubungan dengan pembuatan aplikasi seperti *crawling*, *cleaning data*, analisis sentimen, *opinion retrieval*, *Support Vector Machine* dan *Nu Support Vector Machine*, kemudian membaca literatur yang ada dan mencari literatur tambahan yang dibutuhkan dalam pendalaman materi terhadap konsep dan teori analisis sentimen.

2. Tahap Pengumpulan Data  
Pada tahap ini, akan dilakukan pengumpulan data yang dibutuhkan untuk keperluan pembuatan sistem. Data-data yang akan dikumpulkan antara lain :
  - a. Data opini dari *reviewer*.
  - b. Daftar kelas data (*part-of-speech/POS*)
  - c. Daftar Stopword
  - d. Data probabilitas bobot sentimen kata yang diambil dari *Sentiword*
3. Tahap Perancangan Sistem  
Pada tahap ini, akan dilakukan perancangan sistem analisis sentimen. Rancangan tersebut antara lain rancangan proses dan rancangan sistem.
4. Tahap Implementasi  
Pada tahap ini, akan dilakukan implementasi berdasarkan rancangan yang telah dibuat sebelumnya menjadi suatu aplikasi klasifier opini berdasarkan fitur produk.
5. Tahap Pengujian (testing) dan Analisis Hasil Pengujian  
Pada tahap ini, akan dilakukan pengujian terhadap sistem yang dibangun dan analisis terhadap output yang dihasilkan dari tahap pengujian sebelumnya.
6. Tahap Pembuatan Dokumentasi  
Pada tahap ini, dibuat dokumentasi dari sistem yang dibangun.