BAB 1
PENDAHULUAN

1.1 Latar Belakang Masalah

Dewasa ini perkembangan pertelevisian Indonesia sangatlah pesat, sehingga banyak bermunculan acara-acara baru yang bervariasi. Mulai dari acara *news*, *talk show*, *variety show* hingga acara lawak atau komedi. Salah satu acara komedi adalah *standup comedy*, yaitu salah satu genre profesi melawak yang pelawaknya membawakan lawakannya di atas panggung seorang diri, dengan cara bermonolog mengenai suatu topik. Orang yang melakukan kegiatan ini disebut pelawak tunggal atau komik[1].

Salah satu televisi yang mempunyai acara *standup comedy* yaitu kompastv. Tak sekedar memberi hiburan komedi dengan komik yang sudah ada, kompastv pun beberapa kali menggelar acara pencarian bakat *standup comedy*, yang audisinya tersebar di seluruh Indonesia. Oleh karena antusias masyarakat yang cukup besar terhadap *standup comedy* ini, maka kompastv pun memiliki media sosial sebagai sarana promosi acara *standup comedy*, salah satunya yaitu twitter. Twitter digunakan kompastv untuk menampilkan jadwal acara standup comedy yang akan ditayangkan sehingga masyarakat dapat menonton acara tersebut. Adanya akun resmi *standup comedy* kompastv di twitter, memungkinkan masyarakat yang menjadi followers akun tersebut melakukan *feedback* dengan menanggapi setiap postingan *tweet* dari akun tersebut. Karena ketersediaan sentimen untuk acara *standup comedy* kompastv sangat melimpah, sehingga *feedback* yang disampaikan oleh *followers* dapat digunakan untuk mengukur analisis sentimen *followers* terhadap acara *standup comedy* di kompastv.

Permasalahan analisis sentimen adalah bagaimana melakukan klasifikasi sentimen pada data twitter yang hasilnya dapat dimanfaatkan untuk evaluasi mengenai kualitas suatu acara televisi. Banyaknya *tweet* yang masuk ke akun twitter *standup comedy* kompastv dapat digunakan untuk mengetahui sentimen *followers* terhadap acara *standup comedy* kompastv, akan tetapi dalam penyampaian sentimen tersebut terdapat sentimen yang bukan merupakan opini,

maka dari itu sebelum sentimen diklasifikasikan ke dalam kategori positif dan negatif, sebuah data twitter harus dipisahkan ke dalam kategori relevan dan tidak relevan dengan menggunakan metode *Support Vector Machine*. Setelah data twitter masuk ke dalam kategori relevan, selanjutnya data twitter tersebut diklasifikasikan ke dalam kategori positif dan negatif dengan menggunakan metode *Naïve Bayes Classifier*. Penggunaan *Support Vector Machine* dan *Naïve Bayes Classifier* tersebut dikarenakan performasinya lebih baik untuk klasifikasi teks berbahasa Indonesia, sehingga analisis sentimen dapat diterapkan pada *tweet* bahasa Indonesia yang memiliki nilai subyektif[2].

Dari pemaparan di atas, studi kasus pada skripsi ini adalah analisis sentimen terhadap akun *standup comedy* kompastv. Data masukan analisis sentimen berasal dari postingan *tweet* akun *standup comedy* kompastv.

1.2 Rumusan Masalah

Berdasarkan uraian yang terdapat pada latar belakang masalah, maka rumusan masalah yang ada yaitu bagaimana melakukan analisis sentimen pengguna twitter terhadap acara *standup comedy* kompastv dengan menggunakan *Support Vector Machine* dan *Naïve Bayes Classifier*.

1.3 Maksud dan Tujuan

Berdasarkan permasalahan yang diteliti, maka maksud dari penulisan tugas akhir ini adalah melakukan analisis sentimen pengguna twitter terhadap acara *standup comedy* kompastv dengan menggunakan metode *Support Vector Machine* dan *Naïve Bayes Classifier.*

Adapun tujuan yang akan dicapai dalam tugas akhir ini adalah:

1. Untuk mengetahui sentimen pengguna twitter terhadap acara *standup comedy* kompastv dengan menggunakan metode *Support Vector Machine* dan *Naïve Bayes Classifier.*
2. Untuk mengetahui besarnya akurasi dari penggunaan metode *Support Vector Machine* dan *Naïve Bayes Classifier* dalam melakukan analisis sentimen pengguna twitter terhadap acara *standup comedy* kompastv.

1.4 Batasan Masalah

Batasan masalah dalam pembangunan aplikasi ini adalah sebagai berikut :

1. Twitter yang dianalisis yaitu twitter *standup comedy* kompastv "@StandUpKompasTV".
2. Tweet yang digunakan adalah teks *tweet* berbahasa Indonesia.
3. Praproses yang dilakukan adalah *Case Folding, Cleansing, Stopword Removal, Convert Emoticon, Convert* Negasi, dan *Tokenizer*.
4. Klasifikasi opini dalam bentuk relevan, tidak relevan, positif dan negatif.
5. Klasifikasi relevan dan tidak relevan menggunakan *Support Vector Machine.*
6. Klasifikasi positif dan negatif menggunakan Naïve Bayes *Classifier*.
7. Pengujian akurasi menggunakan *K-Cross Validation*.
8. Bahasa pemrograman dalam sistem ini menggunakan bahasa pemrograman PHP.
9. Sistem yang dibangun adalah website dengan menggunakan pendekatan *OOAD (Object Oriented Analysis and Design).*
10. Diagram UML yang digunakan yaitu *usecase diagram, activity diagram, sequence diagram, dan class diagram*.

1.5 Metodologi Penelitian

Dalam penelitian tugas akhir ini, dilakukan beberapa metode untuk memperoleh data atau informasi dalam menyelesaikan permasalahan. Metodologi penelitian yang digunakan dalam penulisan tugas akhir ini adalah sebagai berikut:

1. Studi Literatur

   Dalam pengumpulan data untuk menyusun skripsi ini menggunakan studi literatur, yaitu sebuah penelitian ilmiah mengenai kesusastraan. Penulis melakukan studi pustaka dengan cara mempelajari literatur berupa *textbook*, jurnal elektronik, artikel ilmiah dan dokumen web yang relevan dengan topik penelitian, yaitu mengenai analisis sentimen dan ***Support Vector Machine***.

2. Pengumpulan Data Twitter

   Pengumpulan data menggunakan sumber data primer yang diambil secara langsung dari akun *standup comedy* kompastv dengan memanfaatkan *streamAPI (Application Programming Interface)* yang disediakan oleh twitter.

3. Metode Pembangunan Perangkat Lunak

   Dalam tahap ini, teknik analisis data menggunakan metode pengembangan perangkat lunak secara waterfall yang meliputi:

   a. Requirements analysis and definition

   Tahap ini merupakan tahap mengumpulkan kebutuhan secara lengkap kemudian dianalisis dan diidentifikasi kebutuhan yang harus dipenuhi oleh program yang akan dibangun.

   b. System and software design

   Tahap ini merupakan tahap desain yang dikerjakan setelah kebutuhan selesai dikumpulkan secara lengkap.

   c. Implementation and unit testing

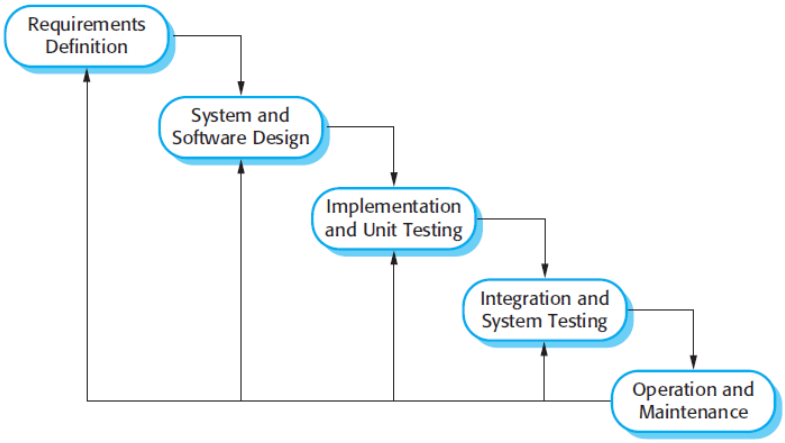
   Tahap ini merupakan tahap menerjemahkan desain program ke dalam kode-kode dengan menggunakan bahasa pemrograman yang sudah ditentukan. Program yang dibangun langsung diuji baik secara unit.

   d. Integration and system testing

   Tahap ini merupakan penyatuan unit-unit program kemudian diuji secara keseluruhan (*system testing*).

   e. Operation and maintenance

   Tahap ini merupakan tahap mengoperasikan program dilingkungannya dan melakukan pemeliharaan, seperti penyesuaian atau perubahan karena adaptasi dengan situasi sebenarnya. Untuk lebih jelasnya tahapan-tahapan tersebut dapat dilihat pada Gambar 1.1



Gambar 1. 1 Model Proses *Waterfall*[3]

4. Evaluasi Hasil

   Evaluasi performansi dilakukan dengan melihat nilai dari perhitungan *accuracy* yang didapat dengan menggunakan metode *k-fold cross validation*.

1.6 Sistematika Penulisan

Sistematika penulisan tugas akhir ini disusun untuk memberikan gambaran umum tentang penelitian yang dijalankan. Sistematika penulisan tugas akhir ini adalah:

BAB I PENDAHULUAN

Bab ini menguraikan tentang latar belakang masalah, rumusan masalah, maksud dan tujuan, batasan masalah, metodologi penelitian dan sistematika penulisan.

BAB II LANDASAN TEORI

Pada bab ini membahas berbagai konsep dasar dan teori-teori yang berkaitan dengan topik penelitian yang dilakukan. Landasan teori yang digunakan antara lain tentang twitter, analisis sentimen, teknik analisis sentimen menggunakan *Machine Learning*, pengklasifikasian menggunakan *Support Vector Machine* dan *Naïve Bayes Classifier*.

BAB III ANALISIS MASALAH DAN PERANCANGAN

Bab ini membahas mengenai penentuan atribut analisis sentimen, cara pengambilan data mentah dari twitter, serta implementasi *Support Vector Machine* dan *Naïve Bayes Classifier* dalam pengklasifikasian data.

BAB IV IMPLEMENTASI DAN PENGUJIAN

Bab ini membahas mengenai implementasi metode *Support Vector Machine* dan *Naïve Bayes Classifier* pada sistem yang dibuat, serta pengujian sistem beserta hasil pengujian dari analisis sistem yang ada pada Bab III kemudian melakukan analisis terhadap informasi dari hasil uji coba.

BAB V KESIMPULAN DAN SARAN

Bab ini berisi tentang kesimpulan-kesimpulan dari skripsi, dan terdapat pula saran yang sudah diperoleh dari hasil penulisan tugas akhir agar dapat digunakan di masa mendatang untuk pengembangan lebih lanjut analisis sentimen dengan menggunakan metode *Support Vector Machine* dan *Naïve Bayes Classifier*.