BAB 1
PENDAHULUAN

1.1 Latar Belakang Masalah

Layanan *social media* dan *microblogging* saat ini telah mengalami perkembangan yang cukup pesat. Salah satu layanan *microblogging* yang terkenal dan memiliki pengguna yang cukup besar adalah *Twitter*. Indonesia tercatat sebagai sumber pengguna *Twitter* terbanyak ketiga di dunia [1]. Pada *Twitter* terdapat istilah bernama *tweet* yang merupakan sebuah pesan atau status yang dibuat oleh penggunanya. Sebuah *tweet* dapat mengekspresikan sebuah perasaan atau keadaan dari pengguna *Twiter*. *Tweet* dapat mengandung sebuah opini dari penggunanya terhadap kejadian yang dialaminya. Opini tersebut dapat dimanfaatkan sebagai penilaian bagi sebuah perusahaan atau instansi.

Bagi suatu perusahaan atau instansi, penilaian konsumen terhadap produk atau jasa yang ditawarkannya sangatlah penting. Setiap penilaian ,saran, ataupun keluhan dari konsumen tentu harus ditanggapi serius agar tidak merusak reputasi perusahaan. *Twitter* banyak digunakan oleh instansi pemerintahan atau perusahaan sebagai media komunikasi dengan konsumen atau pengguna jasanya. Contohnya pada *provider* telekomunikasi di Indonesia, dimana terdapat beberapa *provider* telekomunikasi yang memiliki akun *twitter*, diantarnya seperti Telkomsel, XL, Indosat, Tri , dan lainnya. *Tweet-tweet* yang ditujukan kepada akun *official* dari *provider-provider* tersebut tentu akan dapat bermanfaat sebagai bahan penilaian terhadap layanan atau produk yang dikeluarkannya. Namun menentukan dan memilah apakah suatu *tweet* mengandung sebuah opini positif atau negatif tentu tidak mudah jika *tweet* yang diteliti sangat banyak jumlahnya. Masalah tersebut dapat diatasi dengan membangun sistem yang dapat melakukan analisis sentimen.

Analisis sentimen merupakan bagian dari *Natural Language Processing* (NLP) dan *Machine Learning* [2]. Cara kerjanya adalah dengan melakukan klasifikasi terhadap opini positif dan negatif. Analisis sentimen menganalisa pendapat orang, sentimen, evaluasi, penilaian, sikap, dan emosi terhadap entitas seperti produk, jasa, organisasi, individu, masalah, peristiwa atau suatu topik.

Tujuan lain dari analisis sentimen adalah untuk mengetahui apakah sebuah objek yang diteliti, dalam hal ini *tweet* mengandung opini atau tidak [3]. Klasifikasi sentimen akan sangat membantu dalam memberikan input dan *feedback* dari pelanggan dengan cepat [4].

Salah satu proses yang penting dalam analisis sentimen adalah klasifikasi. Metode klasifikasi yang digunakan sangat menentukan hasil dari analisis sentimen itu sendiri. Naive bayes classifier adalah metode klasifikasi probabilistik sederhana berdasarkan penerapan teorema *Bayes* (dari statistik *Bayesian*) dengan asumsi independen (*naif*) yang kuat. Kelebihan dari *naive bayes classifier* adalah hanya memerlukan sejumlah kecil data pelatihan untuk mengestimasi parameter (rata-rata dan varian dari variabel) yang diperlukan untuk klasifikasi [5]. Metode ini memiliki algoritma yang sederhana sehingga mudah diimplementasikan [3].

Berdasarkan masalah yang telah diuraikan, maka penelitian ini bermaksud untuk membangun Analisis Sentimen Pada Akun Twitter Provider Telekomunikasi.

1.2 Rumusan Masalah

Berdasarkan uraian pada latar belakang dapat dirumuskan masalahnya adalah bagaimana melakukan klasifikasi opini positif dan negatif terhadap *tweet* yang ditujukan kepada akun *twitter provider* telekomunikasi?

1.3 Maksud dan Tujuan

Maksud dari penelitian yang dilakukan adalah untuk mengimplementasikan analisis sentimen dengan menggunakan metode klasifikasi *naive bayes classifier*.

Tujuan yang diharapkan dengan dibangunnya sistem ini antara lain ,sistem dapat melakukan klasifikasi opini positif dan negatif terhadap *tweet* yang ditujukan kepada *akun twitter provider* telekomunikasi.

1.4 Batasan Masalah

Batasan masalah pada penelitian yang akan dilakukan antara lain :

1. Data latih dan data uji yang digunakan pada penelitian menggunakan Bahasa Indonesia.
2. Akun *twitter provider* telekomunikasi yang digunakan pada penelitian adalah akun *twitter* XL123.
3. Data latih dan data uji diperoleh dari *posting tweet* yang ditujukan kepada akun *twitter* XL123.
4. Klasifikasi *tweet* relevan dan tidak relevan menggunakan metode *Support Vector Machine*.
5. Klasifikasi sentimen menggunakan metode *Naive Bayes Classifier*.
6. Terdapat dua kelas yang diklasifikasikan yaitu positif dan negatif.
7. Sistem harus terkoneksi internet untuk mengumpulkan data latih dan data uji.
8. Sistem bersifat simulator

1.5 Metodologi Penelitian

Metodologi yang digunakan pada penelitian ini adalah metode deskriptif. Metode deskriptif adalah suatu metode dalam meneliti status,suatu objek, suatu set kondisi, suatu kelas peristiwa pada masa sekarang. Metode ini memiliki tujuan untuk membuat deskripsi atau gambaran secara sistematis, faktual dan akurat mengenai fakta-fakta, sifat-sifat serta hubungan antar fenomena yang diselidiki.

1.5.1 Metode Pengumpulan Data

Metode pengumpulan data yang digunakan pada penelitian ini adalah sebagai berikut :

1. Studi Literatur

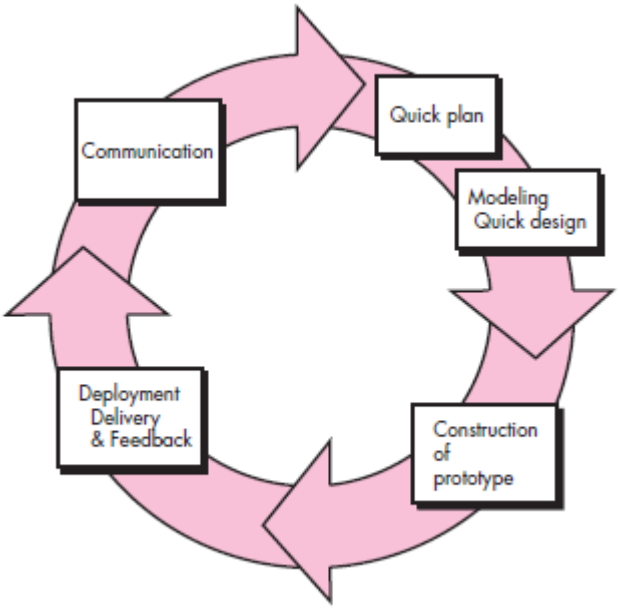
   Teknik pengumpulan data dengan cara mengumpulkan literatur, jurnal, *paper*, dan buku yang berkaitan dengan analisis sentimen.

2. Observasi

   Teknik pengumpulan data secara langsung dengan melakukan peninjauan terhadap *tweet-tweet* yang ditujukan kepada akun *twitter* XL

1.5.2 Metode Pembangunan Perangkat Lunak

Model proses pembangunan perangkat lunak yang digunakan dalam penelitian ini adalah model *prototype*. Tahapan-tahapan pada model *prototype* menurut Roger S. Pressman adalah sebagai berikut :



Gambar 1.1. Skema *Prototype [6]*

1. *Communication*

   Mendefinisikan objektif secara keseluruhan dan mengidentifikasi kebutuhan dari analisis sentimen.

2. *Quick plan*

   Merupakan tahapan menentukan rencana yang akan dilakukan setelah mengetahui kebutuhan apa saja yang diperlukan.

3. *Modeling Quick design*

   Merupakan tahap desain yang dikerjakan setelah kebutuhan selesai dikumpulkan secara lengkap.

4. *Construction of prototype*

   Merupakan tahap desain program yang diterjemahkan ke dalam kode-kode dengan menggunakan bahasa pemrograman yang sudah ditentukan dan *testing* yang disesuaikan dengan *modeling* yang telah selesai dibuat

5. *Deployment Delivery & Feedback*

   Prototype dievaluasi oleh pengguna dan digunakan untuk memperbaiki persyaratan perangkat lunak yang akan dikembangkan. Apabila *prototype* sudah sesuai dengan kebutuhan, maka perangkat lunak dapat diterima dan penelitianpun selesai. Jika terdapat suatu revisi yang harus dilakukan ataupun tidak sesuai dengan kebutuhan, maka kembali lagi ke proses *communication* dan melakukan revisi yang dibutuhkan.

1.6 Sistematika Penulisan

Sistematika penulisan laporan akhir penelitian ini disusun untuk mendeskripsikan secara umum tentang penelitian yang dilakukan. Sistematika penulisan tugas akhir ini adalah sebagai berikut :

BAB 1. PENDAHULUAN

Menjelaskan tentang latar belakang permasalahan, merumuskan inti permasalahan yang ada, menentukan tujuan dan kegunaan penelitian, yang kemudian diikuti dengan pembatasan masalah, asumsi, serta sistematika penulisan.

BAB 2. TINJAUAN PUSTAKA

Membahas konsep dasar dan teori-teori yang berkaitan dengan topik penelitian yang dilakukan dan hal-hal yang berguna dalam proses analisis permasalahan serta tinjauan terhadap penelitian.

Bab ini menjelaskan tinjauan mengenai analisis sentimen, *information retrieval*, *support vector machine, naive bayes classifier, confusion matrix* dan *unified modelling language.*

BAB 3. ANALISIS DAN PERANCANGAN SISTEM

Bab ini membahas analisis sistem, analisis masalah, analisis pengumpulan data, analisis *preprocessing*, analisis *support vector machine*, analisis *naive bayes classifier*, analisis kebutuhan non-fungsional, analisis kebutuhan fungsional, analisis perancangan antarmuka sistem dan jaringan semantik.

BAB I4. IMPLEMENTASI DAN PENGUJIAN SISTEM

Menjelaskan implementasi dan pengujian dari hasil analisis dan perancangan sistem yang dibangun, serta metode atau teknik analisis yang akan digunakan.

Bab ini menguraikan implementasi dari sistem, implementasi *preprocessing*, implementasi *support vector machine*, *naive bayes classifier* pada analisis sentimen, pengujian *black box*, pengujian *confusion matrix*.

BAB 5. KESIMPULAN DAN SARAN

Berisi kesimpulan dan saran yang diharapkan dapat menjadi masukan untuk pengembangan aplikasi di masa yang akan datang.