BAB I
PENDAHULUAN

I.1 Latar Belakang Masalah

Telkom Speedy adalah layanan broadband akses internet dari PT. Telekomunikasi Indonesia Tbk. yang berkualitas tinggi bagi perumahan serta SME (*Small Medium Enterprise*). Telkom Speedy menggunakan teknologi ADSL (*Asymmetric Digital Line Subsriber*), MSAN (*Multi Service Access Node*), dan GPON (*Gigabit Passive Optical Network*), yang menghantarkan sinyal digital berkecepatan tinggi melalui jaringan telepon secara optimal bagi keperluan konsumsi konten internet, dengan kecepatan data dari 384 kpbs hingga 100 Mbps. Pada tahun 2013, jumlah pengguna Telkom Speedy mencapai 2,4 juta orang [1]. Banyaknya jumlah pengguna ini membuat Telkom Speedy harus menyediakan layanan berupa *feedback* untuk penggunanya agar memudahkan dalam meninjau kembali produknya. Salah satu media yang digunakan sebagai layanan *feedback* adalah Twitter.

Pada jaman sekarang, *mikroblogging* menjadi sarana komunikasi *online* yang sangat popular di tengah masyarakat. Terlihat sangat jelas dari salah satu *mikroblogging* yang sangat popular yaitu Twitter, dengan jumlah pengguna aktif mencapai lebih dari 231,7 juta dengan rata-rata jumlah *tweets* perharinya mencapai 58 juta *tweets* [2]. Indonesia sendiri memiliki pengguna aktif Twitter sebanyak 19,5 juta [3]. Jumlah *tweets* berkembang dengan cepat karena kesederhanaan dan kemudahan dalam penggunaannya merupakan beberapa alasan mengapa Twitter lebih digemari masyarakat Indonesia dalam berkomunikasi. Tentu saja, informasi yang terkandung dalam *tweets* ini sangat berharga sebagai alat penentu kebijakan dan ini bisa dilakukan dengan menggunakan analisis sentimen.

Analisis sentimen atau disebut juga *opinion mining*, adalah bidang studi dalam menganalisis pendapat orang, evaluasi, penilaian, sikap dan emosi terhadap suatu entitas seperti produk, jasa, organisasi, individu, isu-isu, peristiwa dan topik. Fokus utama dari analisis sentimen adalah untuk menyatakan mana

yang termasuk opini positif dan mana yang termasuk opini negatif [4]. Salah satu contoh dari pengaplikasian analisis sentimen adalah ketika suatu perusahaan mengeluarkan suatu produk dan perusahaan tersebut menyediakan layanan untuk menerima opini dari konsumen mengenai produk tersebut. Analisis sentimen digunakan untuk mengelompokkan opini-opini positif dan negatif dari konsumen yang menggunakan produk tersebut sehingga mempercepat dan mempermudah tugas perusahaan untuk meninjau kembali kekurangan produknya.

Berdasarkan gambaran penjelasan di atas, disini penulis tertarik untuk mengaplikasikan sistem analisis sentimen terhadap *official* akun Twitter Telkom Speedy dengan menggunakan *Naïve Bayes Classifier* (NBC).

I.2 Perumusan Masalah

Berdasarkan uraian latar belakang diatas dapat dirumuskan masalahnya adalah bagaimana membangun suatu sistem yang dapat melakukan analisis sentimen pada *official* akun Twitter Telkom Speedy dengan menggunakan *Naïve Bayes Classifier* (NBC) ?

I.3 Maksud dan Tujuan

Berdasarkan permasalahan yang diteliti, maka maksud dari penelitian ini adalah untuk menerapkan metode *Naïve Bayes Classifier (NBC)* dalam membangun sebuah aplikasi yang mampu mengklasifikasikan opini sebagai opini positif ataupun opini negatif sebagai salah fungsi utama analisis sentimen.

Tujuan yang akan dicapai dalam penelitian ini yaitu:

1. Membangun suatu sistem analisis sentimen pada *official* akun Twitter Telkom Speedy dengan menggunakan *Naïve Bayes Classifier (NBC),* yang dapat membantu Telkom Speedy dalam menganalisis sentimen para penggunanya.
2. Menentukan kata yang menjadi pemicu sentimen berdasarkan *tweets* perhari.

I.4 Batasan Masalah

Adapun batasan masalah dari penelitian ini adalah:

1. Dataset yang digunakan berasal dari *tweets* yang menggunakan bahasa Indonesia yang ditujukan kepada akun @TelkomSpeedy.
2. Metode pengklasifikasiannya menggunakan *Naïve Bayes Classifier (NBC).*
3. Metode pembobotan yang digunakan adalah *Term Frequency – Inverse Document Frequency* (TF-IDF).
4. *Tweets* yang diklasifikasi, diharapkan berupa *tweets* yang mengandung opini.
5. *Tweets* diklasifikasikan berdasarkan 2 kelas, yaitu kelas sentimen positif dan kelas sentimen negatif.

I.5 Metodologi Penelitian

Metodologi penelitian merupakan suatu proses untuk memecahkan suatu masalah yang logis, dan memerlukan data-data untuk mendukung terlaksananya suatu penelitian. Metode yang digunakan dalam penelitian ini adalah metode deskriptif. Metode deskriptif merupakan sebuah metode yang menggambarkan fakta-fakta dan informasi secara sistematis, faktual dan akurat.

I.5.1 Metode Pengumpulan Data

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

a. Studi Literatur

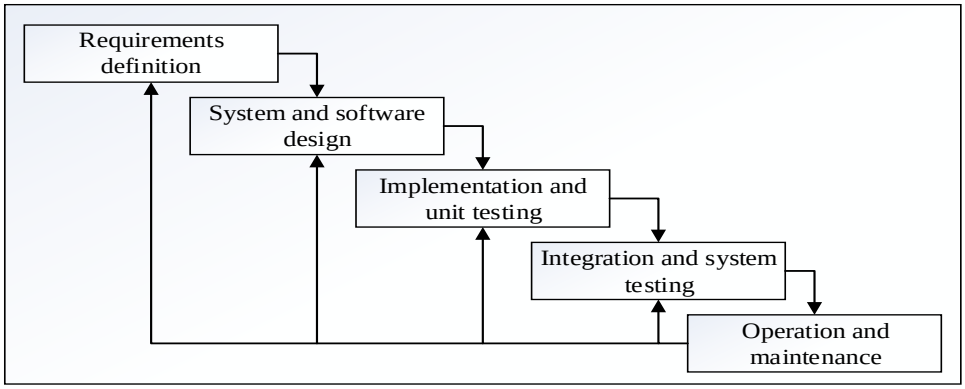
   Pengumpulan data dengan cara mengumpulkan referensi seperti jurnal, *paper* dan bacaan-bacaan yang ada kaitannya dengan judul penelitian. Mempelajari literatur dan teori pendukung penelitian mengenai klasifikasi, khususnya *naïve bayes classifier*.

b. Observasi

   Teknik pengumpulan data dengan mengadakan penelitian dan peninjauan langsung terhadap permasalahan yang diambil.

I.5.2 Metode Pembangunan Perangkat Lunak

Metode pembangunan dalam pembuatan perangkat lunak ini menggunakan paradigma *waterfall*. Fase-fase dalam model *waterfall* menurut referensi Ian Sommerville terdapat pada Gambar I.1 berikut :



Gambar I.1 Skema *Waterfall* [5]

Yang meliputi beberapa proses diantaranya :

a. *Requirements Analysis and Definition*

   Tahap ini merupakan kegiatan pengumpulan kebutuhan secara lengkap kemudian dianalisis dan didefinisikan kebutuhan yang harus dipenuhi oleh program yang akan dibangun. Fase ini harus dikerjakan secara lengkap untuk bisa menghasilkan desain sistem yang lengkap.

b. *System and Software Design*

   Merupakan tahap menganalisis hal-hal yang diperlukan dalam pelaksanaan proyek pembuatan perangkat lunak.

c. *Implementation and Unit Testing*

   Desain program diterjemahkan ke dalam kode-kode dengan menggunakan bahasa pemrograman yang sudah ditentukan. Program yang dibangun langsung diuji secara unit.

d. *Integration and System Testing*

   Penyatuan unit-unit program kemudian diuji secara keseluruhan (*system testing*).

e. *Operation and Maintenance*

Mengoperasikan program dilingkungannya dan melakukan pemeliharaan, seperti penyesuaian atau perubahan karena adaptasi dengan situasi sebenarnya.

I.5.3 Metode Pembangunan Sistem Analisis Sentimen

Sistem analisis sentimen yang akan dibangun akan melewati tahapan-tahapan pada Gambar I.2.



Gambar I.2 Alur Pembangunan Sistem Analisis Sentimen

Tahapannya terdiri dari:

a. Analisis Sumber Data

Data yang digunakan adalah kumpulan *tweets* bahasa Indonesia yang diambil dari *official* akun Twitter Telkom Speedy (@TelkomSpeedy). Data *tweets* ini diperoleh dengan membuat program *crawling* yang secara otomatis akan mengambil data *tweets* yang mengandung kata "TelkomSpeedy", "speedy reguler", "speedy instant", dan "speedy gold".

b. *Data Preprocessing*

Pada tahap ini, *tweets* yang terkumpul akan diproses sehingga data yang didapat menjadi lebih terstruktur dan mudah untuk diolah. Langkah-langkah *preprocessing* terdiri dari *tokenizing*, normalisasi fitur, *case folding, convert emoticon, convert negation, stemming* dan *stopword removal*.

c. Klasifikasi Sentimen

Langkah selanjutnya adalah proses pengklasifikasian yang akan diproses menggunakan metode *Naïve Bayes Classifier* untuk menentukan mana yang termasuk opini positif dan mana yang termasuk opini negatif.

d. Pembobotan Kata

Pada tahap ini akan dilakukan proses pengekstrakan *keyword* menggunakan nilai TF-IDF (*Term Frequency – Inverse Document Frequency*). *Term*

(kata) yang memiliki bobot TF-IDF terbesarlah yang akan merepresentasikan topik pemicu nilai sentimen di selang waktu tersebut (diambil perhari).

e. Visualisasi

Hasil dari tahap klasifikasi sentimen akan digambarkan dalam bentuk diagram *pie*. Data diambil dari jumlah opini positif dan opini negatif lalu ditampilkan dalam bentuk persentase pada diagram.

I.6 Sistematika Penulisan

Sistematika penulisan penelitian ini disusun untuk memberikan gambaran umum tentang kasus yang akan dipecahkan. Sistematika penulisan penelitian ini adalah sebagai berikut :

BAB I PENDAHULUAN

Bab ini menjelaskan mengenai latar belakang masalah, perumusan masalah, maksud dan tujuan, batasan masalah, metodologi penelitian, serta sistematika penulisan.

BAB II TINJAUAN PUSTAKA

Membahas berbagai 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 III ANALISIS DAN PERANCANGAN SISTEM

Menganalisis masalah dari model penelitian untuk memperlihatkan keterkaitan antar variabel yang diteliti serta model matematis untuk analisisnya.

BAB IV IMPLEMENTASI DAN PENGUJIAN SISTEM

Bab ini membahas tentang implementasi dari hasil tahapan analisis dan perancangan aplikasi yang dibangun. Serta berisi uji coba dan hasil pengujian sistem.

BAB V KESIMPULAN DAN SARAN

Berisi kesimpulan yang sudah diperoleh dari hasil pengujian sistem dan berisi saran untuk pengembangan selanjutnya.