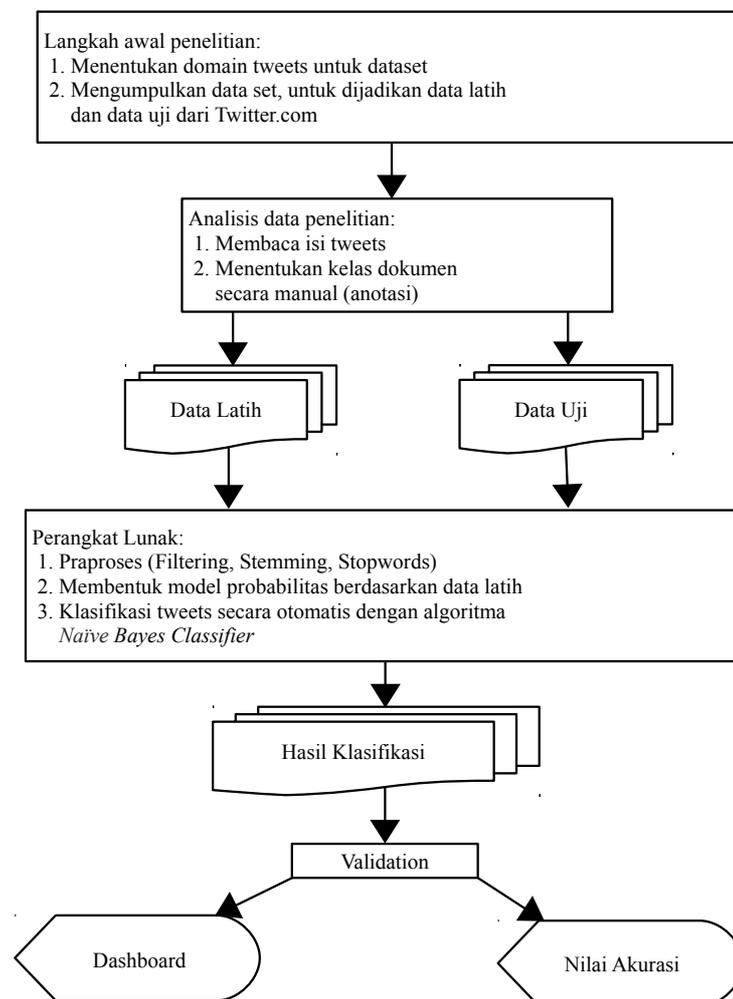


BAB III

METODOLOGI PENELITIAN

3.1 Desain Penelitian

Desain penelitian merupakan tahapan yang akan dilakukan penulis dalam melakukan penelitian. Tahapan penelitian yang digunakan dijabarkan pada Gambar 3.1.



Gambar 3.1 Desain Penelitian

Penjelasan dari Gambar 3.1 adalah sebagai berikut:

1. Menentukan domain

Domain *tweet* yang digunakan dalam penelitian ini adalah bpjs. *tweet* yang mengandung kata kunci “bpjs” akan otomatis dikumpulkan.

2. Pengumpulan Data

Data dalam penelitian ini bersumber dari *microblog* Twitter.com. Pengambilan data memanfaatkan Twitter API.

3. Anotasi

Anotasi atau menentukan kelas pada kumpulan data set. Data set yang telah dianotasi ini dijadikan data latih. Untuk mendapatkan data dengan *golden standard* maka anotasi dilakukan secara manual.

4. Praproses

Proses yang akan dilakukan pada tahap ini yaitu:

- a) *Cleaning*, yaitu proses membersihkan data *tweet* dari kata-kata yang tidak mendeskripsikan sesuatu untuk mengurangi noise. Beberapa contoh kata yang dihilangkan adalah hashtags (#), username (@username), dan url. Selain penghapusan kata, dilakukan penghapusan tanda baca seperti koma (,) dan titik (.).
- b) *Case folding*, yaitu penyeragaman bentuk huruf, sehingga data hanya menjadi huruf latin dari a sampai z.
- c) *Tokenizing*, yaitu proses memecah sebuah kalimat menjadi kata.
- d) *Convert Emoticon*, yaitu proses mengubah emoticon menjadi kata yang mewakili ekspresi. Contohnya “:)", “:D” menjadi “__senyuman__”, “:(“ menjadi “__cemberut__”

5. Pembentukan model probabilitas

Pada tahap ini seluruh *tweet* data latih yang telah terkumpul akan diubah menjadi kumpulan *term* yang memiliki nilai probabilitas terhadap kelas yang ditentukan.

6. Klasifikasi

Pada tahap ini akan dilakukan pengelompokkan *tweet* ke dalam kelas positif, negatif dan netral. Algoritma yang digunakan yaitu *Naïve Bayes Classifier*. Setelah diklasifikasi maka akan diuji kualitas dari model yang

dihasilkan dengan menghitung nilai akurasi.

7. Dashboard

Kumpulan data yang telah diklasifikasi akan memiliki nilai informasi kemudian ditampilkan pada *dashboard* sehingga pengguna dapat melihat sentimen pada domain yang telah ditentukan.

3.2 Metode Penelitian

Untuk lebih jelasnya, metode penelitian yang dilakukan dijelaskan dalam sub-bab berikut:

3.2.1 Proses Pengumpulan Data

Data yang digunakan dalam penelitian ini merupakan data yang diambil dari Twitter melalui *Search* API dengan parameter query “bpjs”. Proses pengumpulan data dilakukan pada rentang bulan Februari 2015 dan Juni 2015. Data hasil *crawling* ini dimasukkan ke *database* untuk diproses lebih lanjut.

3.2.2 Proses Pengembangan Perangkat Lunak

Model proses pengembangan atau rekayasa perangkat lunak yang digunakan dalam penelitian ini adalah V Model. V Model merupakan perbaikan dari model *waterfall* (Balaji, 2012).

1. Analisis Kebutuhan, tahap pengumpulan data dan informasi terkait dengan metode yang akan digunakan dan berbagai kebutuhan penunjang lain untuk mendapatkan gambaran aplikasi. Pada tahap ini dilakukan pengumpulan data *tweet* dengan kata kunci bpjs dan menentukan algoritma apa yang digunakan untuk melakukan analisis sentimen.
2. Desain Sistem, tahap pembuatan desain rekayasa sistem yang akan dibuat sesuai dengan analisa kebutuhan. Pada tahap ini dibuat desain dari sistem yang akan dikembangkan untuk analisis sentimen (lihat gambar 3.1).
3. Desain Arsitektur/Program, tahap pembuatan desain rekayasa struktur program, dan algoritma. Pada tahap ini ditentukan modul apa saja yang akan dibuat untuk melakukan analisis sentimen, yaitu modul praproses, modul pembentukan model, modul klasifikasi, modul dashboard.
4. Desain Modul, tahap pembuatan modul-modul program hasil

implementasi dari tahap Desain Arsitektur/Program. Tahap ini membuat programmer dapat melakukan pemecahan unit-unit yang masing-masing unitnya memiliki fungsi yang sama. Pada tahap ini dibuat kerangka prosedur atau fungsi pada setiap modul praproses, modul pembentukan model, modul klasifikasi, dan modul dashboard.

5. Implementasi, tahap pembuatan aplikasi yang merupakan tindak lanjut dari perancangan program untuk mengolah data menjadi informasi berupa kode-kode program.
6. *Unit Testing*, tahap pengujian terhadap aplikasi yang telah dibuat. Jika terdapat kekurangan maka tahap ini berhubungan dengan tahap Desain Arsitektur/Program. Pada tahap ini diuji apakah ada *task* yang terlewat atau tidak ketika melakukan analisis sentimen.
7. *Integration Testing*, tahap melakukan integrasi dari modul-modul yang terpisah menjadi satu kesatuan yang utuh. Pada tahap ini setiap modul yang telah dibuat diuji fungsionalitasnya melalui antarmuka sistem.
8. *System Testing*, tahap pengujian sistem. Jika terdapat kekurangan maka tahap ini berhubungan dengan tahap Desain Sistem. Pada tahap ini dilakukan pengujian keseluruhan sistem dari mulai tahap praproses sampai ekstraksi informasi (dashboard).
9. *Acceptance Testing*, tahap pengujian yang bersifat menyeluruh, yakni berupa simulasi terhadap data sebenarnya. Jika terdapat kekurangan maka tahap ini berhubungan dengan tahap Analisis Kebutuhan. Pada tahap ini dilakukan pengecekan terhadap data hasil klasifikasi.

3.3 Alat dan Bahan Penelitian

Alat yang digunakan dalam penelitian ini adalah seperangkat komputer yang dilengkapi perangkat keras dan perangkat lunak pendukung. Sedangkan bahan yang digunakan adalah data yang diambil melalui *microblog* Twitter.

3.3.1 Alat Penelitian

Dalam penelitian ini digunakan perangkat keras komputer dengan spesifikasi sebagai berikut:

1. Processor Pentium(R) Dual-Core CPU T4500 @ 2.30GHz × 2
2. Memori 4 GB RAM
3. Hardisk berkapasitas 320 GB
4. Monitor 14” dengan resolusi 1366x768 pixel
5. Mouse dan Keyboard

Adapun perangkat lunak yang digunakan sebagai berikut:

1. Ubuntu GNOME 14.04 LTS
2. Sublime Text 2
3. Python 2.7.6
4. MySQL WorkBench

3.3.2 Bahan Penelitian

Adapun bahan penelitian yang digunakan berupa data yang diambil dari Twitter dengan kata kunci “bpjs”. Serta bahan lain berupa kumpulan kata *stopword* dan kumpulan data sinonim yang diambil *library tweet mining* berbahasa Indonesia yang dibuat oleh Yudi Wibisono di alamat: <http://code.google.com/p/id-tweetmining>.