



Hak cipta dan penggunaan kembali:

Lisensi ini mengizinkan setiap orang untuk menggubah, memperbaiki, dan membuat ciptaan turunan bukan untuk kepentingan komersial, selama anda mencantumkan nama penulis dan melisensikan ciptaan turunan dengan syarat yang serupa dengan ciptaan asli.

Copyright and reuse:

This license lets you remix, tweak, and build upon work non-commercially, as long as you credit the origin creator and license it on your new creations under the identical terms.

BAB III

METODOLOGI DAN PERANCANGAN

3.1 Metode Penelitian

Dalam melakukan penelitian ini metode yang dilakukan dapat dijabarkan sebagai berikut.

1. Studi Literatur

Pengumpulan referensi terkait dari berbagai media dan jurnal mengenai segala teori-teori yang dibutuhkan dalam membangun sistem ini. Teori yang akan dikumpulkan berupa teori bagaimana mengolah *natural language* agar dimengerti sistem, studi mengenai linguistik, dan penggunaan *Classification Tree* sebagai alat untuk mengklasifikasikan data.

2. Analisis Kebutuhan Sistem

Pengkajian atas kebutuhan yang digunakan sistem. Kebutuhan yang dikaji meliputi data yang akan digunakan sebagai *knowledge system* atas pengolah komentar dan mengkategorikannya sebagai positif atau negatif.

3. Perancangan Sistem

Perancangan yang dimaksud meliputi rancangan mengenai bagaimana alur dari sistem bekerja dalam mengolah data serta bagaimana tampilan yang baik untuk sistem. Dalam hal ini dibutuhkan perancangan dengan diagram agar dapat mengatur alur proses dengan baik.

4. Pembangunan Sistem dan Implementasi

Dengan teori yang sudah dipelajari dan terkumpulnya data maka akan dimulai pembangunan sistem yang sesuai dengan perancangan diatas untuk mencapai hasil yang sesuai dengan tujuan penelitian.

5. Pengujian

Hasil dari sistem yang telah dibuat akan dikaji kembali dengan teknik-teknik pengujian apakah hasil sudah sesuai dengan tujuan penelitian.

3.2 Analisis

Penelitian ini menggunakan data opini yang diambil dari media sosial Twitter. Untuk mengurangi beban olah dan meningkatkan keakuratan sistem maka setiap opini yang didapat perlu disaring dari penulisan yang tidak baku menjadi baku. Selain itu untuk mengolah kata per kata untuk mengklasifikasikan kalimat tersebut perlu digunakan juga teknik *stemming* untuk menghilangkan imbuhan pada kata-kata yang berimbuhan sehingga akan mudah disaring oleh sistem. Contohnya kata 'kebaikan' menggunakan awalan 'ke-' dan akhiran '-an' sehingga dapat diambil kata dasarnya yaitu 'baik'. Pengurangan imbuhan ini diperlukan untuk mengurangi jumlah *training data* dan waktu pengolahan kalimat. Algoritma Nazief Adriani digunakan sebagai teknik stemming yang digunakan dalam penelitian ini.

Aplikasi yang dibuat dalam penelitian ini menggunakan bahasa pemrograman web PHP, HTML, Javascript dan CSS. Selain itu dikarenakan aplikasi ini menggunakan media sosial Twitter sebagai sumber data sehingga diperlukan REST API oleh Twitter untuk mendapatkan data opini terkait.

Seluruh penyimpanan data seperti *training data* untuk kebutuhan pengklasifikasian akan disimpan ke dalam *database* MySQL.

3.3 Perancangan Sistem

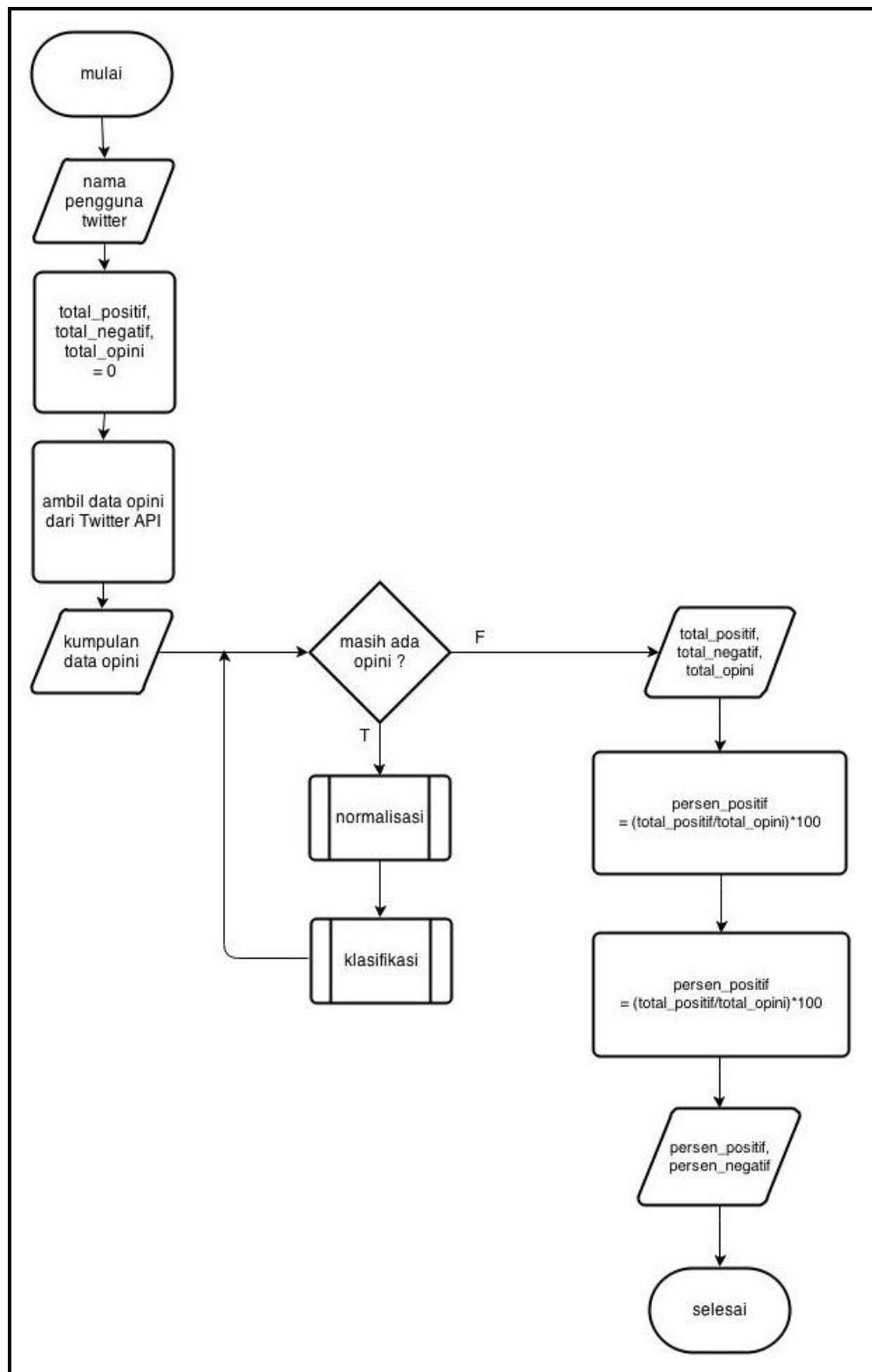
Sistem yang dikembangkan dapat dijelaskan melalui diagram sistem, fungsionalitas, serta masukan dan keluaran yang dibutuhkan yaitu sebagai berikut.

3.3.1 Diagram Sistem

Secara umum alur kerja sistem diawali dengan *input* pengguna yang merupakan *username* suatu akun dari Twitter. Setelah pengguna memasukan *username* tersebut maka aplikasi akan menarik data dari Twitter API untuk mendapatkan *tweet* atau opini terkait yang kemudian akan diolah. Pengolahan data dilakukan secara berulang opini per opini.

Dalam tahap pengolahan data akan melalui beberapa fase. Fase pertama yaitu normalisasi atau penyaringan tanda baca dan angka serta menghilangkan imbuhan pada kata. Sedangkan fase kedua adalah klasifikasi teks yang akan menghasilkan analisa sentimen terhadap teks tersebut dan memberikan kesimpulan apakah teks tersebut dikategorikan sebagai teks bersifat positif atau negatif.

Setelah semua opini sudah diklasifikasikan maka jumlah opini positif dan opini negatif akan menjadi keluaran sistem. Jumlah tersebut dapat digunakan berikutnya untuk melihat perbandingan jumlah komentar positif dan negatif dari kumpulan opini tersebut. Untuk lebih jelasnya akan dijelaskan pada gambar 3.1.



Gambar 3.1 Diagram alir aplikasi

3.3.2 Fungsionalitas Sistem

Fungsionalitas yang dimiliki oleh sistem adalah sebagai berikut.

1. Melakukan pengambilan kumpulan data opini dari Twitter menggunakan Twitter API. Data opini dapat diperoleh dengan mengakses API tersebut menggunakan *username* Twitter sebagai parameter. Setelah API diakses maka akan mendapatkan kumpulan opini terkait *username* Twitter tersebut.
2. *Stemming* atau menghilangkan imbuhan pada setiap kata yang sudah dinormalisasi menjadi kata dasar yang sesuai dengan *training data* menggunakan algoritma Nazief dan Adriani.
3. Melakukan pemrosesan klasifikasi data terhadap setiap kalimat opini yang didapatkan menggunakan metode *text classification*. Metode *text classification* yang digunakan adalah metode *classification tree*.

3.3.3 Masukan dan Keluaran Sistem

Dalam aplikasi ini masukan yang diperlukan cukup *username* Twitter dari pengguna yang ingin diklasifikasikan opini terkait pengguna tersebut. Sedangkan keluaran dari aplikasi ini adalah sebagai berikut.

1. Potongan kata-kata dari kalimat opini yang telah dinormalisasi dan berbentuk kata dasar.
2. Kesimpulan hasil klasifikasi dari setiap kalimat opini berupa negatif atau positif.
3. Persentase komentar positif dan negatif dari opini yang didapatkan.

3.4 Desain Sistem

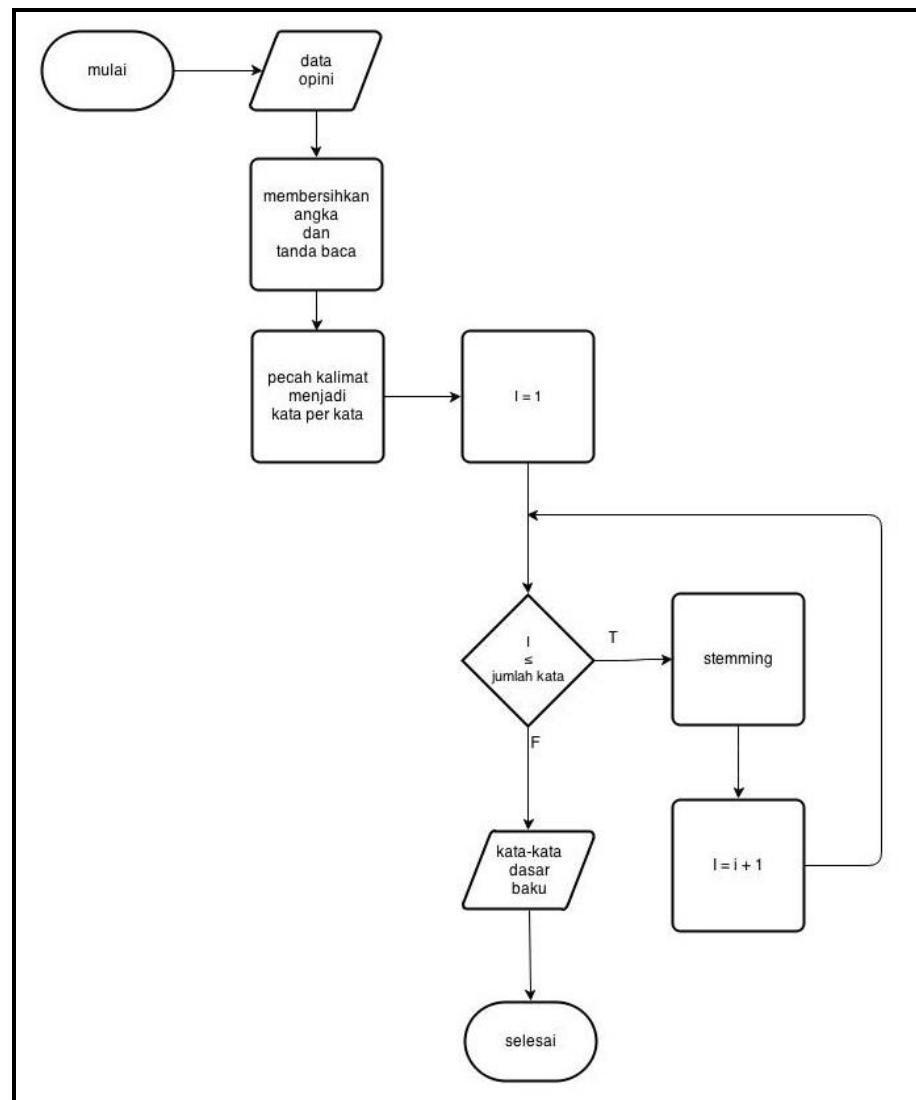
Sistem yang akan dirancang dan diimplementasikan dalam penelitian ini akan melalui fase normalisasi dan fase klasifikasi. Penjelasan dari masing-masing fase akan digambarkan dalam diagram alir beserta tampilan antarmuka aplikasi berikut.

3.4.1 Desain Modul dan Subroutine

Fase yang berupa modul yang telah disebutkan sebelumnya akan dijelaskan menggunakan diagram alir sebagai berikut.

1. Desain modul normalisasi

Pada fase normalisasi, kalimat opini yang sudah didapat dari Twitter API akan diolah agar dapat mudah diklasifikasikan di tahap berikutnya. Tahap pertama adalah menghilangkan angka dan tanda baca, kemudian kalimat opini tersebut dipecah menjadi kata per kata dan diolah kembali menjadi kata dasar atau *stemming* menggunakan algoritma Nazief dan Adriani. Dengan terbentuknya kata dasar yang baku maka akan mudah untuk melakukan perbandingan dengan *training data* yang sudah disusun sebelumnya dan memperpendek waktu perbandingan. Penggunaan algoritma Nazief dan Adriani pada penelitian ini diambil dari penelitian yang sudah dilakukan sebelumnya yang dicantumkan pada bab tinjauan pustaka sebelumnya. Tahap normalisasi ini dapat dijelaskan melalui diagram alir pada gambar 3.2.

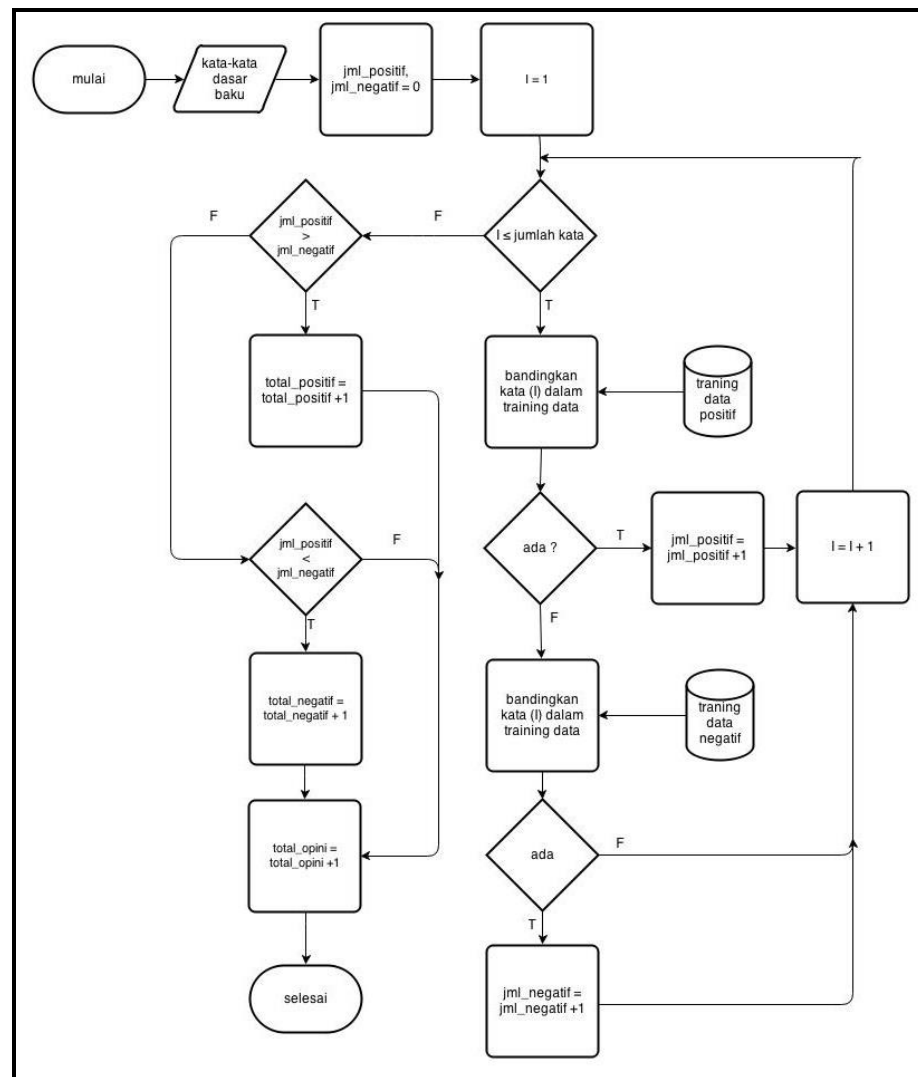


Gambar 3.2 Diagram alir normalisasi

2. Desain fase klasifikasi

Setelah mendapat kata-kata dasar yang baku dari kalimat opini sistem akan membandingkan setiap kata dengan *training data* yang sudah disusun sebelumnya. Sesuai dengan algoritma *classification tree*, pengklasifikasian ini akan bercabang menjadi kelas positif dan negatif. Masing-masing kelas telah memiliki *training data* yang sudah terkumpul dalam database sehingga pengklasifikasian akan menjadi mudah dengan membandingkan kata per kata.

Setelah perbandingan masing-masing kata telah selesai maka dapat disimpulkan kalimat berikut dikategorikan sebagai positif atau negatif dari jumlah kata positif dan kata negatif pada kalimat tersebut. Menurut algoritma *classification tree* yang telah dijelaskan pada tinjauan pustaka, kalimat tersebut dikatakan positif atau negatif menurut jumlah kata yang paling banyak antar kelasnya. Apabila jumlah kata positif lebih banyak dari jumlah kata negatif maka kalimat tersebut dapat dinilai positif. Sebaliknya bila kata negatif lebih banyak dari kata positif kalimat tersebut dinilai negatif. Apabila jumlah kata dalam kedua kelas sama atau tidak ada sama sekali maka kalimat tersebut dapat dikatakan netral atau tidak memiliki sentimen. Untuk lebih jelas fase ini dijelaskan pada gambar 3.3.

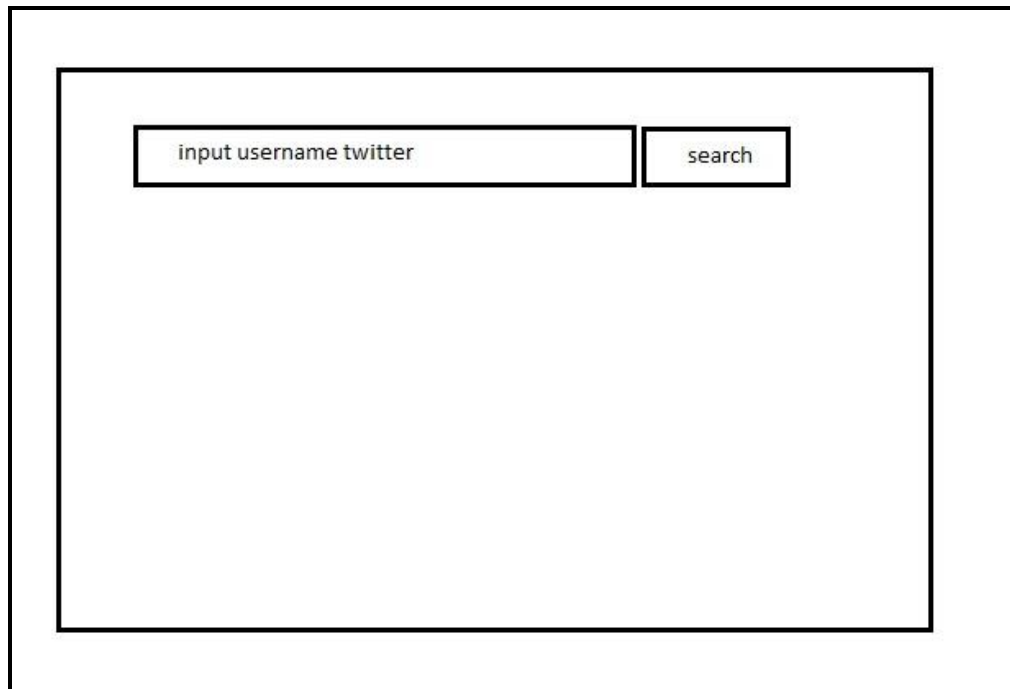


Gambar 3.3 Diagram alir klasifikasi

3.4.2 Tampilan Antar Muka

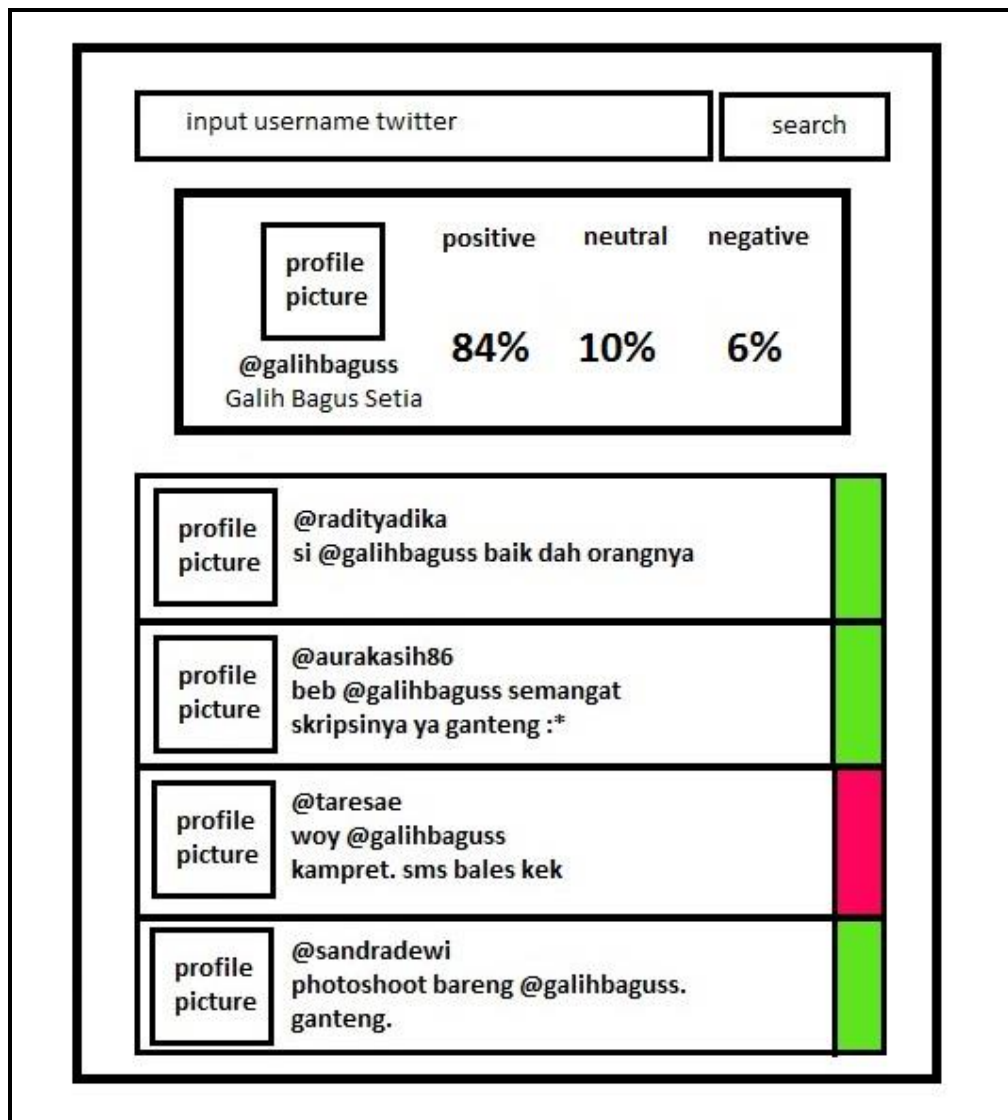
Aplikasi dalam penelitian ini menggunakan bahasa pemrograman berbasis *web* sehingga aplikasi dijalankan di dalam *web browser*. Aplikasi hanya berjalan di sebuah halaman *web* tanpa pindah ke halaman lain. Namun tampilan halaman sebelum dan sesudah proses akan berbeda.

Tampilan antar muka awal ketika halaman dibuka akan menampilkan sebuah *form* yang terdiri dari *text field* untuk memasukkan *username* Twitter dan tombol yang berfungsi untuk *submit*.

The image shows a simple web form layout. It consists of a large outer rectangular border. Inside this border, there is a smaller rectangular area. Within this inner area, there are two elements: a text input field on the left with the placeholder text "input username twitter" and a rectangular button on the right with the text "search".

Gambar 3.4 Tampilan antar muka awal

Ketika pengguna sudah memasukkan *username* Twitter dan menekan tombol *search* maka akan ditampilkan seluruh opini terkait dari akun Twitter tersebut beserta penilaian atas opininya tersebut. Pada masing-masing opini akan diberikan tanda dibagian kanan berupa warna dimana warna hijau merepresentasikan opini positif dan warna merah merepresentasikan opini negatif. Selain itu juga terdapat perhitungan prosentase atas keseluruhan opini yang didapat dari perbandingan jumlah opini negatif dan opini positif.



Gambar 3.5 Tampilan antar muka hasil analisa sentimen