

BAB I

PENDAHULUAN

1.1 Latar Belakang Masalah

Menurut kamus besar bahasa Indonesia, kata informasi berasal dari bahasa Perancis kuno yakni *informacion* yang diambil dari bahasa Latin *informationem* yang berarti garis besar, konsep, atau ide. Informasi juga merupakan kata benda dari *informare* yang berarti pengetahuan yang telah disebarkan. Pengetahuan tentang suatu peristiwa atau situasi tertentu yang telah dikumpulkan atau diterima melalui proses komunikasi, pengumpulan intelegen, ataupun didapatkan dari berita juga disebut informasi.

Saat ini informasi menjadi kebutuhan utama bagi sebagian besar orang. Dengan informasi seseorang menjadi “kaya”, dengan informasi seseorang menjadi tahu. Karenanya manusia menciptakan berbagai media dengan tujuan untuk mempercepat penyampaian informasi. Media-media tersebut diantaranya adalah media cetak contohnya adalah majalah atau koran dan media elektronik seperti televisi dan radio.

Manusia dapat berinteraksi dan bertukar informasi melalui bahasa. Bahasa merupakan teknologi, bahasa memungkinkan seseorang memahami informasi yang disampaikan oleh orang lain. Informasi yang disampaikan dari mulut ke mulut hanya bertahan sebentar saja, yaitu hanya pada saat pengirim menyampaikan informasi melalui ucapannya itu saja. Setelah ucapan itu selesai, maka informasi yang berada di tangan penerima tidak dapat disimpan lama. Selain

itu jangkauan suara juga terbatas. Pada jarak tertentu, meskipun masih terdengar, informasi yang disampaikan lewat bahasa suara akan terdegradasi bahkan hilang sama sekali.

Seiring berjalannya waktu teknologi SMS mulai digunakan oleh radio sebagai salah satu fasilitas yang memudahkan proses penerimaan dan penyampaian informasi disamping fasilitas lainnya seperti telepon atau e-mail. SMS merupakan *feature* dari *handphone* yang berfungsi untuk mengirim dan menerima sebuah informasi berbasis teks dengan panjang teks dan tujuan tertentu. SMS dipilih sebagai fasilitas pengirim dan penerima informasi berdasarkan asumsi bahwa hampir setiap orang di Indonesia memiliki *handphone* dan setiap *handphone* memiliki *feature* SMS. [Wahidin, 2010]

Pikiran Rakyat (PR) FM adalah sebuah radio swasta yang terdapat pada frekuensi 107.5 FM dan beralamat di Jl. Braga No.5 Bandung. PR FM menerima ratusan SMS berupa berita, keluhan dan komentar dari pendengar setiap harinya. Dari sekian banyak pesan, sering ditemukan kata-kata yang tidak layak. Firman Allah SWT dalam surat Al-Ahzab ayat 70:

سَدِيدًا قَوْلًا وَقُولُوا لِلَّهِ اتَّقُوا ءَامِنُوا الَّذِينَ يَتَأْتِيهَا

Artinya: *Hai orang-orang yang beriman, bertakwalah kamu kepada Allah dan Katakanlah Perkataan yang benar* (QS. Al-Ahzab: 70)

Maka diperlukansebuah aplikasi yang dapatmendeteksi kata tidak layak atau kata kotor pada berita yang dikirim pendengar melalui SMS. Kemudian berita tersebut diklasifikasi menjadi berita layak (bersih) dan tidak layak(kotor).

Beberapa penelitian sebelumnya, yakni Sri Nurhayati pada tahun 2010 mampu mengklasifikasi kesenian Jawa Barat dengan metode Naive Bayes *Classifier* dengan cukup baik (Sri, 2010). Selain itu, Yudi mengklasifikasi berita berbahasa Indonesia berdasarkan topik dengan metode Naive Bayes *Classifier* dan hasilnya mampu mengklasifikasi berita secara efektif (Yudi, 2005). Penelitian Yudi diteruskan oleh Aziz pada tahun 2009 dengan menambahkan proses *tokenizing* dan *filtering* sehingga menambah nilai keakuratan proses klasifikasi [Aziz, 2009].

Berdasarkan latar belakang, maka dilakukan penelitian dan penyusunan tugas akhir yang berjudul “Pemanfaatan *Text Mining* pada Berita dengan Metode Naive Bayes *Classifier* untuk Deteksi Kata Tidak Layak”.

1.2 Rumusan Masalah

Berdasarkan latar belakang, terdapat beberapa masalah yang dapat diidentifikasi, diantaranya adalah:

1. Bagaimana cara untuk menentukan layak atau tidak layak berita yang dikirimkan oleh pendengar untuk dibacakan?
2. Bagaimana cara untuk meningkatkan kecepatan dan keakuratan dalam pendeteksian kata tidak layak tersebut?

1.3 Tujuan

Berdasarkan identifikasi masalah di atas, dapat disimpulkan bahwa penelitian ini bertujuan untuk:

1. Membuat aplikasi *text mining* yang berfungsi untuk mendeteksi kata kotor pada SMS yang berasal dari pendengar lalu menentukan layak atau tidak layak makna dari kata kotor tersebut untuk dibacakan secara *onair*.
2. Menerapkan *Naïve Bayes Classifier* untuk mengoptimalkan proses pendeteksian kata tidak layak.

1.4 Batasan Masalah

Permasalahan yang ditemukan selama penelitian ini dibatasi oleh hal-hal yang tercantum berikut ini:

1. Pendeteksian kata kasar atau kata tidak layak dalam bahasa Indonesia yang tidak terdapat penyingkatan kata ataupun penambahan angka.
2. Metode klasifikasi yang digunakan adalah *Naïve Bayes Classifier (NBC)*.
3. Sistem yang akan dibangun berbasis web dan berjalan pada sistem lokal.
4. Pendeteksian kata tidak layak tanpa menggunakan pemecahan imbuhan dan kata dasar (*stemming*) bahasa Indonesia.

1.5 Metodologi Penelitian

Penelitian ini menggunakan metodologi penelitian sebagai berikut:

1. *Text mining* dan metode klasifikasi

Berita yang diterima akan diproses aplikasi *text mining* yang memiliki beberapa tahapan yakni:

a. *Text preprocessing*

Text preprocessing adalah tahap dimana seluruh karakter huruf diubah menjadi huruf besar (*uppercase*), kemudian dipecah menjadi kata, kemudian kata-kata tersebut dimasukkan ke dalam sebuah *array*.

b. *Pattern discovery*

Pattern discovery adalah proses pembentukan pola baru yang dimulai dari pencocokan kata-kata pada *array* dengan *database* kata kotor. Kemudian menentukan jenis kata yang terletak sebelum dan yang terletak setelah kata kotor yang terdeteksi menjadi jenis “kata kerja” atau “lainnya”. Sehingga terbentuk pola baru yang berisi kata kotor, kata yang terletak sebelum kata kotor, dan kata yang terletak setelah kata kotor.

c. *Naive Bayes Classifier*

Naive Bayes Classifier adalah metode klasifikasi yang akan digunakan untuk menghitung nilai *class* dari pola baru yang telah terbentuk pada proses *pattern discovery* dengan pola lama yang terdapat pada *database* aturan. Apabila nilai akumulasi *class* “bersih” lebih besar dari nilai akumulasi *class* “kotor”, maka *class* berita adalah “bersih”. Namun jika Apabila nilai akumulasi *class* “bersih” lebih kecil dari nilai akumulasi *class* “kotor”, maka *class* berita adalah “kotor”.

2. Metode pengembangan perangkat lunak

Menurut Pressman, tahap-tahap dalam penggunaan metode *prototype* adalah:

a. Langkah pertama

Mendefinisikan kebutuhan dan garis besar aplikasi yang akan dirancang.

b. Langkah kedua

Membangun aplikasi yang bersifat sementara namun berdasarkan kebutuhan dan keinginan *user*.

c. Langkah ketiga

Mengizinkan *user* menggunakan aplikasi tersebut. *User* akan memerikan masukan berupa saran atau kritik. Hal ini dimaksudkan untuk mengetahui apakah *prototype* sudah sesuai keinginan *user*.

d. Langkah keempat

Mengimplementasikan masukan dari *user*.Lalu kembali ke langkah ketiga hingga aplikasi sesuai dengan kebutuhan dan keinginan *user*. [Pressman, 2002]

3. Metode pengumpulan data

a. Wawancara

Wawancara dilakukan kepada Pemimpin Redaksi, Redaktur Pelaksana, dan beberapa staf PRFM.

b. Observasi

Observasi dilakukan dengan cara pemantauan secara langsung proses penerimaan SMS dan kemudian penyuntingan berita.

c. Studiliteratur

Studi literatur dilakukan dengan membandingkan jurnal dan tugas akhir yang menggunakan metode Naive Bayes *Classifier* untuk mengetahui gambaran dan mengetahui seberapa berbeda penelitian yang akan dilakukan.

1.6 Sistematika Penulisan

BAB I PENDAHULUAN

Bab I berisi tentang latar belakang, identifikasi masalah yang dihadapi, tujuan penelitian, batasan masalah yang dihadapi, metodologi penelitian, teknik pengumpulan data serta sistematika penulisan.

BAB II TINJAUAN PUSTAKA

Menjelaskan tentang teori-teori yang digunakan dalam tugas akhir ini serta untuk menjelaskan dan menyelesaikan permasalahan yang akan dikaji.

BAB III TINJAUAN PERUSAHAAN

Bab III menjelaskan mengenai perusahaan/organisasi/institusi yang menjadi *study* kasus dalam penelitian sebagai bahan dari tugas akhir ini.

BAB IV ANALISIS DAN PERANCANGAN

Bab IV membahas tentang analisis dan rancangan aplikasi.

BAB V IMPLEMENTASI DAN PENGUJIAN

Membahas proses implementasi dan pengujian perangkat lunak secara detail.

BAB VI PENUTUP

Membahas tentang kesimpulan dan saran yang diperoleh dari penulisan tugas akhir ini.