

BAB V

KESIMPULAN DAN SARAN

Pada bab ini berisi kesimpulan dari keseluruhan penelitian yang telah dilakukan, serta saran dari penulis untuk kegiatan penelitian selanjutnya terkait dengan topik yang sedang dibahas.

5.1 Kesimpulan

Adapun kesimpulan pada penelitian ini adalah sebagai berikut:

Algoritma *Naïve Bayes* untuk pengklasifikasian teks pada kasus SMS spam *filtering* layak digunakan. Aplikasi tidak langsung melakukan klasifikasi SMS menggunakan *Naïve Bayes*. Terlebih dahulu dilakukan praproses pada data, agar algoritma *Naïve Bayes* dapat digunakan. Praproses juga meningkatkan kinerja klasifikasi karena praproses akan memperkecil dimensi data sehingga lebih berkualitas.

Evaluasi dan validasi dilakukan menggunakan *10-fold cross validation*. Klasifikasi dilakukan dua tahap. Klasifikasi pertama akan mengklasifikasi data dengan hasil kelas spam atau bukan. Klasifikasi kedua akan mengklasifikasi data dengan hasil kelas penipuan atau promosi.

Pada penelitian ini data *training* yang digunakan sebanyak 1140 SMS. Klasifikasi pertama memperoleh rata-rata akurasi sebesar 94,12%. Klasifikasi kedua memperoleh rata-rata akurasi 94,56%. Pengukuran kinerja sistem berdasarkan ketepatan pengklasifikasian data, terhadap masing-masing kelas menggunakan hasil prediksi dari model, diukur menggunakan perhitungan *precision*. Sedangkan tingkat keberhasilan hasil klasifikasi dapat diukur dengan menggunakan perhitungan *recall*.

Pembangunan perangkat lunak menggunakan metode *linear sequential model* atau *waterfall model*. Tahapan yang dilakukan diantaranya melakukan rekayasa informasi atau sistem, analisis kebutuhan, perancangan, *coding*, dan uji coba. Terdapat dua perangkat lunak yang dibangun secara *independence* yaitu

aplikasi *desktop* dan aplikasi *mobile*. Aplikasi *desktop* akan membangun model yang akan digunakan pada aplikasi *mobile*. Pada aplikasi *mobile* akan melakukan klasifikasi menggunakan model yang dihasilkan oleh aplikasi *desktop*.

5.2 Saran

Adapun saran yang diberikan pada penelitian ini untuk pengembangan selanjutnya adalah sebagai berikut:

1. Perbaiki praproses agar *noise* yang dihasilkan sekecil mungkin, salah satu caranya adalah memperbaiki daftar sinonim dan daftar *stopword*.
2. Menggunakan teknik *oversampling* untuk membuat data baru pada kasus tertentu. Contohnya, jika probabilitas suatu kata terhadap kelas sangat besar. Maka sertakan beberapa data yang mengandung kata tersebut yang memiliki karakteristik pada kelas lain. *Oversampling* dapat dilakukan dengan mereplika, merekayasa, atau menyesuaikan dengan data yang sudah ada.
3. Aplikasi dapat melakukan *report* ketika terjadi kesalahan klasifikasi dan memperbaiki model, sehingga akurasi semakin baik. Dengan demikian pengguna dapat memberikan preferensi terhadap SMS masing-masing. Pembahasan ini terkait materi *relevant feedback*.
4. Aplikasi memiliki kemampuan untuk dapat dijadikan *default messaging app*. Sehingga bisa menampilkan daftar SMS bersih yang bebas dari SMS yang diprediksi spam. Hal ini dapat dilakukan salah satunya dengan melakukan *update* status pada atribut *report_spam* dalam *database* SMS android yang hanya dapat dilakukan oleh *default messaging app*.