

## BAB V

### KESIMPULAN DAN SARAN

#### 5.1. Kesimpulan

Pada penelitian ini berdasarkan hasil dari analisis, terhadap sistem pendeteksi informasi *hoax* pada aplikasi *Twitter* dengan dikembangkannya algoritma Apriori dan algoritma *Random Forest*, maka dapat disimpulkan:

- 1) Hasil dari proses mengintegrasikan algoritma Apriori pada proses *training data*, dapat menyelesaikan permasalahan banyaknya *dataset* pada teks informasi *hoax (tweet)* dalam menambang *itemset* yang mana hasil dari proses algoritma Apriori dijadikan parameter utama untuk proses klasifikasi, dan proses algoritma Apriori mampu menghasilkan data lebih terstruktur dan mudah dipahami untuk program bekerja. Sehingga pada proses *testing data* dilakukan, algoritma *Random Forest* mampu membangun dan mengklasifikasikan beberapa model *decision tree* dengan cepat dan lebih optimal.
- 2) Sistem ini membentuk dua *class* utama sebagai pendeskripsian hasil dari proses prediksi, dengan asumsi kondisi penilaian *mean absolute error* yaitu, pertama yaitu *class* benar proses prediksi maka nilai dari MAE yaitu 0 dan *class* kedua dengan kondisi salah memprediksi maka nilai dari MAE yaitu 1, dimana *mean absolute error* (MAE) menyatakan rata-rata *error* dalam satu data.
- 3) Nilai *mean absolute error* (MAE) dalam pengamatan pada proses klasifikasi untuk masing-masing pengamatan terhadap berita *hoax* dan berita *non-hoax* yaitu 0, dan untuk nilai *mean absolute error* (MAE) untuk program utama dalam memproses semua informasi pada aplikasi *Twitter* mencapai 0.01.
- 4) Proses klasifikasi yang tidak sesuai pada penelitian ini, dapat dipengaruhi oleh beberapa faktor seperti, nilai *minimum support* yang tidak sesuai, *dataset training* yang digunakan kurang tepat, pola dari proses Apriori

yang terlewatkan dan kemungkinan dari ada perbedaan sumber data latih dan data prediksi. Namun upaya untuk menemukan nilai *minimum support* dan model sistem yang tepat, harus dibarengi dengan penelitian atau pemeriksaan secara manual untuk dapat memberikan hasil yang paling akurat.

- 5) Tingkat akurasi dalam mendeteksi informasi berita sebagai mencapai 99.4% dan untuk tingkat akurasi dalam mendeteksi informasi berita pada aplikasi *Twitter*. Tingkat akurasi disini memiliki makna yaitu tingkat persentase berita yang benar diprediksi *hoax* dan *non-hoax* dari keseluruhan berita, dengan nilai *minimal support 2*.

## 5.2. Saran

Menurut penulis, penelitian ini masih mengalami beberapa kendala dan kekurangan. Saran untuk penelitian selanjutnya adalah sebagai berikut:

- 1) Hasil dari proses penelitian ini menggunakan dua sumber data yang berbeda, maka untuk menguji model dari sistem perlu dilakukannya percobaan dengan menggunakan sumber data yang sama antara proses *learning* dan proses *prediction*.
- 2) Sebagai salah satu upaya dalam mempercepat dan membuat program lebih presisi diperlukannya perhitungan atau menentukan nilai *confidence* dari proses algoritma Apriori, dengan tujuan agar didapatnya nilai *minimum support* yang tepat benar.
- 3) Proses *cleansing* data, yaitu dengan tujuan untuk menghapus pengulangan kata (*repetition removal*) dan penanganan terkait singkatan serta terkait terjemahan bahasa *non-formal* (*Abbreviations and non-formal language translation*), termasuk bahasa daerah. Tahapan ini dibutuhkan karena ada banyak kemungkinan suatu berita informasi dalam *Twitter* dalam Bahasa Indonesia yang kurang sesuai. Oleh karena itu, dibutuhkannya satu tahapan *cleansing* data pada *text preprocessing*.