

## **BAB V**

### **PENUTUP**

#### **5.1 Kesimpulan**

Dari hasil uji coba klasifikasi tweet pada Twitter menggunakan metode Naïve Bayes dapat disimpulkan sebagai berikut:

1. Algoritma Naïve Bayes dengan pembobotan kata TF-IDF terbukti dapat melakukan klasifikasi data tweet dalam kategori prostitusi dan bukan tweet prostitusi. Uji coba yang dilakukan menggunakan 40000 dataset yang dibagi menjadi 32000 data training dan 8000 untuk data testing.
2. Berdasarkan hasil pengujian yang dilakukan oleh peneliti terhadap model klasifikasi tweet menggunakan metode Naïve Bayes didapatkan tingkat akurasi sebesar 98,41%, dengan nilai precision sebesar 98%, nilai recall sebesar 98%, dan nilai f1-measure sebesar 98%.
3. Validasi model klasifikasi menggunakan teknik K-Fold Cross Validation menunjukkan hasil nilai rata-rata akurasi sebesar 98.42%, nilai rata-rata precision sebesar 98.65% dan nilai rata-rata recall sebesar 98.19%.
4. Pembagian data training terhadap pada model klasifikasi Naïve Bayes menunjukkan bahwa pembagian data training 90% memiliki nilai akurasi tertinggi yaitu sebesar 98.55%.

## 5.2 Saran

Selama melakukan penelitian dan uji coba, peneliti menemukan beberapa kekurangan yang perlu diperbaiki agar mendapatkan hasil klasifikasi yang optimal. Saran yang diberikan sebagai berikut:

1. Pada tahap *preprocessing data* perlu dilengkapi dengan *term normalization* yaitu proses pemetaan beberapa variasi kata menjadi kata tunggal yang memiliki arti sama atau perbaikan kata yang memiliki penulisan salah (*typo*) agar dapat meningkatkan konsistensi.
2. Sistem ini belum dapat menangani negasi, sehingga perlu penambahan metode untuk mendeteksi dan menangani negasi untuk meningkatkan akurasi dalam melakukan klasifikasi data.

