

BAB V

KESIMPULAN DAN SARAN

1.1 Kesimpulan

Kesimpulan dari penelitian ini adalah sebagai berikut :

1. Pola kata kunci terbaik yang dapat digunakan untuk mengekstrak informasi *gender*, usia, dan status pekerjaan merupakan kombinasi 7 kategori kata kunci yang meliputi :

- a. Gender

- i. Fashion
- ii. Olah Raga
- iii. Kata Emosi
- iv. Kata Panggilan
- v. Emoticon

- b. Usia

- i. Kata Singkatan
- ii. Kata Non Formal
- iii. Kata Partikel
- iv. Kata Penghubung
- v. Emoticon

- c. Status Pekerjaan

- i. Kata Kantor
- ii. Kata Kerja
- iii. Kata Pin

- iv. Kata Info
 - v. Kata Ganti Subjek Pertama
 - vi. Kata Asing Daerah
2. Pola kata kunci tersebut digunakan sebagai bahan pembentukan model klasifikasi dengan menggunakan algoritma *decision tree* C4.5. Model klasifikasi yang terbentuk akan menjadi acuan untuk mengklasifikasikan *gender*, usia, dan status pekerjaan.

1.2 Saran

Untuk memprediksi karakteristik demografi seseorang berdasarkan tutur kata yang digunakan pada tulisannya memang sulit untuk mencapai ketepatan atau tingkat akurasi 100%. Tanpa sistem komputerisasi pun memprediksi demografi seseorang berdasarkan tutur kata yang digunakan pada tulisannya masih sering terjadi kesalahan prediksi. Meskipun begitu, penelitian untuk menghasilkan sistem prediksi karakteristik demografi tetap dapat dilakukan. Penambahan pengenalan pola dan daftar kata ke dalam model klasifikasi sangat mempengaruhi tingkat akurasi. Secara logika, manusia pun akan dapat mengerti arti dari kata-kata yang mereka dengar dan atau tulisan-tulisan yang mereka baca dengan melalui proses pembelajaran dari pola dan daftar kata yang digunakan pada kata-kata atau tulisan-tulisan tersebut. Oleh karena itu, penulis menyarankan agar penelitian ini dapat terus dilakukan dan dikembangkan baik dari aspek model klasifikasi, pola, dan juga daftar kata yang digunakan pada saat penelitian dilakukan.