

BAB V

PENUTUP

5.1 Kesimpulan

Penelitian ini menggunakan data akun *Twitter* berjumlah 120 akun. Akun *Twitter* tersebut telah divalidasi sebelumnya secara manual oleh seorang pakar atau dalam hal ini adalah seorang psikolog yang mampu menganalisis karakter seseorang dilihat dari *Tweet* dari masing-masing akun secara manual.

Berdasarkan pada bab-bab sebelumnya, kesimpulan yang didapat pada penelitian ini adalah sebagai berikut :

1. Berdasarkan perumusan masalah pada penelitian ini, maka disimpulkan bahwa Algoritma *Naive Bayes Classifier* dan *K – Nearest Neighbor* telah berhasil mengklasifikasi profil akun *Twitter* kedalam karakter DISC (*Dominance, Influence, Steadiness, Compliance*) meskipun dengan tingkat akurasi yang cukup rendah.
2. Hasil perhitungan evaluasi yang telah dilakukan dengan *confusion matrix* menyatakan bahwa tingkat akurasi yang didapat dari membandingkan hasil validasi manual dari pakar dan hasil klasifikasi dari sistem untuk algoritma *Naive Bayes Classifier* adalah 34.16 % sedangkan untuk algoritma *K – Nearest Neighbor* adalah 28.33 %. Dan dapat disimpulkan bahwa algoritma *Naive Bayes Classifier* terbukti lebih unggul dibandingkan dengan algoritma *K – Nearest Neighbor* dalam melakukan klasifikasi teks dengan pembobotan TF-IDF menggunakan pendekatan

DISC dengan perbedaan sebesar 5.83% dalam tingkat akurasi, Sedangkan dengan pengujian *3 - Fold Cross Validation* algoritma *Naive Bayes Classifier* mendapat akurasi sebesar 26.66% sedangkan *K - Nearest Neighbor* sebesar 29.16% sehingga dapat disimpulkan dengan pengujian *3 - Fold Cross Validation* algoritma *K - Nearest Neighbor* lebih unggul dibandingkan dengan algoritma *Naive Bayes Classifier* dalam melakukan klasifikasi teks dengan pembobotan TF-IDF menggunakan pendekatan DISC dengan perbedaan sebesar 2.5% dalam tingkat akurasi.

5.2 Saran

Pada penelitian ini terdapat beberapa kekurangan dan keterbatasan. Kekurangan dan keterbatasan ini dapat dijadikan acuan dan pertimbangan untuk penelitian selanjutnya. Adapun saran yang dihasilkan dari penelitian ini yaitu sebagai berikut :

1. Dengan digunakannya algoritma *Naive Bayes Classification* dan *K - Nearest Neighbor*, maka perlu dicoba algoritma lain untuk dijadikan bahan komparasi untuk mengetahui mana algoritma terbaik untuk melakukan klasifikasi karakter individu dengan metode DISC.
2. Dengan data yang digunakan untuk klasifikasi dengan pelabelan yang kurang seimbang pada tiap labelnya, diharapkan untuk kedepannya bisa menambah data-data agar pelabelan menjadi seimbang pada tiap labelnya.

3. Dengan digunakanya pembobotan TF-IDF pada penelitian ini, maka perlu dicoba menggunakan metode pembobotan lain untuk mengetahui manakah metode pembobotan yang paling optimal untuk digunakan pada masing-masing algoritma.

