Analisis Kepribadian DISC Pengguna Twitter dengan Fitur Pembobotan TF-IDF dan TF-CHI-SQUARE menggunakan Algoritma C4.5

Mohammad Abbiyu[1], Erwin Budi Setiawan[2]

[1,2,3]Fakultas Informatika, Universitas Telkom, Bandung
[4]Divisi Digital Service PT Telekomunikasi Indonesia
[1]biyubiu@students.telkomuniversity.ac.id, [2]erwinbudisetiawan@telkomuniversity.ac.id,

---

Abstrak
Twitter merupakan salah satu perkembangan media sosial berbasis *microblogging* yang saat ini banyak digunakan di seluruh dunia. Dengan perkembangan teknologi saat ini, menentukan kepribadian seseorang melalui media sosial seperti *twitter* adalah hal yang diunggulkan. Permasalahannya adalah bagaimana mengklasifikasi teks yang ada di media sosial *twitter* ke dalam kelas-kelas yang akan dibuat untuk mendapatkan nilai kinerja yang baik. Penelitian ini bertujuan untuk membangun sistem untuk mengklasifikasi kepribadian pengguna *twitter* dengan menggunakan metode klasifikasi Algoritma C4.5 (*Decission Tree* C4.5) serta metode pembobotan TF-IDF dan TF *Chi-Square*. Masih sedikitnya penelitian mengenai penggunaan model kepribadian DISC sebagai alat bantu dalam menilai kepribadian, penggunaan dua metode pembobotan serta penambahan fitur berdasarkan perilaku sosial seperti *follower*, *following*, *retweet* dan lainnya dapat menjadi pembeda dari penelitian sebelumnya. Dari hasil percobaan didapatkan akurasi rata-rata terbaik sebesar 43.60% pada perbandingan data latih dan data uji sebesar 90:10 dan pendekatan linguistik dengan menggunakan TF-*CHI-Square* dengan pemilihan 50 kata terbaik.

Kata Kunci: *Twitter*, DISC, Algoritma C 4.5, TF-IDF, TF-*CHI-Square*.