

1. Pendahuluan

Kepribadian merupakan salah satu faktor yang memengaruhi perilaku manusia[1]. *American Psychological Association* (APA) mengemukakan bahwa kepribadian mengacu pada perbedaan individu dalam hal pola karakteristik berpikir, merasakan, dan berperilaku[2]. Tes kepribadian adalah sebuah tes yang dilakukan untuk menentukan tipe kepribadian seseorang[3]. Pengklasifikasian tipe kepribadian yang dilakukan dengan tes kepribadian saat ini sering kali masih menggunakan metode lama dan kurang memanfaatkan teknologi yang ada. Padahal pesatnya perkembangan teknologi memberikan manfaat dalam kehidupan manusia. Klasifikasi kepribadian telah banyak diaplikasikan pada berbagai jurnal ilmiah[4]. Penelitian ini menggunakan *twitter* sebagai media untuk menganalisis kepribadian seseorang[5]. *Twitter* telah menghasilkan 110 juta *tweet* setiap hari dan memiliki lebih dari 200 juta pengguna. Banyak pengguna *twitter* yang secara tidak sadar memberikan informasi tentang kepribadiannya melalui *tweets* atau posting yang mereka buat dengan bahasa alami[6].

Dalam proses menganalisis kepribadian seseorang melalui *twitter* dibutuhkan metodologi yang tepat untuk mendapatkan hasil yang akurat. *Tweet* atau *posting* pada *twitter* merupakan kumpulan kata yang tidak baku sehingga dibutuhkan perlakuan khusus untuk mendapatkan data yang bisa diproses. Oleh sebab itu, dalam pengolahan data diperlukan proses *pre-processing* terlebih dahulu yang kemudian bisa dilakukan pengklasifikasian[7]. Dalam melakukan klasifikasi pada penelitian kepribadian ini penulis menggunakan teori DISC. DISC dibuat untuk membantu individu mengeksplorasi kepribadian dan tipe tingkah laku mereka sehingga mereka dapat lebih baik dalam memahami diri mereka sendiri dan orang lain (Bonnstetter & Suiter, 2011 : 32)[8]. Metode DISC merupakan singkatan dari empat faktor kepribadian yaitu *Dominance*, *Influence*, *Steadness*, dan *Compliance*[9]. Pada penelitian tugas akhir ini, penulis melakukan analisis dalam penilaian kepribadian berdasarkan perilaku sosial dan pendekatan linguistik dari kata-kata *tweets* yang diunggah. Metode yang dipilih untuk melakukan analisis ini adalah *logistic regression* dengan pembobotan TF-IDF dan AHP.