

# Peningkatan Akurasi pada Prediksi Kepribadian MBTI Pengguna Twitter Menggunakan Augmentasi Data

Rizki Nurhaliza Harahap<sup>1</sup>, Kemas Muslim L.<sup>2</sup>

<sup>1,2</sup>Fakultas Informatika, Universitas Telkom, Bandung

<sup>1</sup>lizaharahap@students.telkomuniversity.ac.id, <sup>2</sup>kemasmuslim@telkomuniversity.ac.id

---

## Abstrak

Kepribadian suatu individu perlu diketahui untuk membantu seseorang dalam mempertimbangkan beberapa hal, salah satunya perekrutan karier. Pada umumnya, kepribadian dapat diketahui melalui metode wawancara, observasi, maupun survei kuesioner. Akan tetapi, metode konvensional tersebut dinilai kurang praktis dari segi waktu dan materi karena dibutuhkan waktu yang lama dan biaya yang cukup besar untuk mengolah data. Selain itu, penggunaan metode konvensional juga dapat menimbulkan bias karena melibatkan orang ketiga dalam pengolahan data. Penelitian ini mencoba memberikan solusi dengan membangun model yang dapat melakukan prediksi terhadap kepribadian seseorang berdasarkan analisis data dan informasi dari media sosial Twitter. Data dan informasi tersebut akan diproses sehingga didapatkan prediksi kepribadian orang tersebut. Teori klasifikasi kepribadian yang digunakan adalah teori *Myers-Briggs Type Indicator* (MBTI). Penelitian ini juga mencoba menerapkan teknik augmentasi data untuk meningkatkan performa dari *text mining task* yang memiliki dataset sedikit. Hasil terbaik didapatkan dengan metode *Random Forest* menggunakan pembobotan *Term Frequency-Inverse Document Frequency* (TF-IDF) dan fitur yang tersedia pada Twitter. Penggunaan teknik augmentasi dapat meningkatkan akurasi hingga 20% dari akurasi awal sehingga hasil penelitian menunjukkan bahwa penggunaan teknik augmentasi data dapat meningkatkan performa pada model prediksi kepribadian MBTI.

**Kata kunci :** *MBTI, random forest, twitter, augmentasi data*

---