

Klasifikasi Hashtag Buzzer/Bot Menggunakan Algoritma Random Forest dengan Atribut Komunitas untuk Mengurangi Disinformasi Pada Twitter

Fadhil Nugroho Yudianto¹, Dr. Yuliant Sibaroni²

^{1,2,3}Fakultas Informatika, Universitas Telkom, Bandung

⁴Divisi Digital Service PT Telekomunikasi Indonesia

¹fadhiilnugroho@students.telkomuniversity.ac.id, ²yuliant@telkomuniversity.ac.id,

Abstrak

Saat ini persebaran informasi terjadi secara cepat pada sosial media. Namun, tidak semua informasi yang beredar pada sosial media belum tentu merupakan informasi yang benar. Twitter sebagai salah satu media sosial terbesar didunia juga mengalami masalah yang sama. Banyak sekali informasi dan juga hashtag yang tersebar pada twitter merupakan informasi dan hashtag yang disebarluaskan oleh buzzer. Dalam tugas akhir ini, penulis membangun sebuah sistem klasifikasi yang dapat membedakan antara sebuah hashtag natural dan hashtag yang dibuat oleh buzzer pada sosial media twitter. Penulis memakai algoritma *random forest* dengan memanfaatkan atribut komunitas untuk melakukan klasifikasi hashtag pada sistem yang dibuat. Penulis juga akan mencari tahu apa saja atribut yang paling berpengaruh dalam proses klasifikasi tersebut. Pada penelitian ini, didapatkan bahwa sistem mampu mendeteksi hashtag buzzer dengan tingkat akurasi sebesar 98% dan tingkat f-measure sebesar 98% berdasarkan dataset yang telah penulis dapatkan. Didapatkan juga bahwa atribut atribut yang berpengaruh dalam proses klasifikasi antara lain adalah atribut seperti `unique_user_pertweet` dan `q1_rasio`

Kata kunci : buzzer, *random forest*, hashtag, *trending topic*, twitter
