

ABSTRAK

Isu SARA (suku, ras, agama dan antar golongan) merupakan sentimen diskriminatif menyangkut keturunan, agama, etnis dan kebangsaan atau kesukuan. Isu SARA menjadi salah satu faktor utama penyebab terjadinya konflik di media sosial dalam masyarakat multikultural. Di Indonesia, Twitter merupakan salah satu media sosial yang paling banyak untuk menyebarkan ujaran kebencian dalam bentuk isu SARA. Masyarakat bisa melaporkan konten SARA di Kepolisian terdekat. Sedangkan Polsek Pati mengalami beberapa kendala dalam menangani kasus kejahatan siber tersebut antara lain, kemampuan SDM yang masih terbatas, sarana komputerisasi yang belum bisa mengklasifikasi pesan SARA secara otomatis dan pemahaman masyarakat tentang pesan yang mengandung unsur SARA.

Penelitian ini menggunakan algoritma *random forest decision tree* yang dikombinasikan dengan pembobotan fitur TF-IDF. Metode tersebut menggunakan data latihan sebanyak 1707 dokumen dengan pembagian 1158 dokumen non-sara dan 549 dokumen pesan sara untuk dilakukan pembuatan model algoritma. Model algoritma *random forest* dibuat menggunakan data numerik hasil pembobotan fitur TF-IDF terhadap data latihan. Penelitian ini menggunakan model algoritma yang terbuat dari 3000 pohon keputusan. Model algoritma melakukan pengklasifikasian pesan dari hasil *scraping* url di Twitter apakah pesan tersebut masuk ke dalam kategori SARA atau tidak.

Penelitian ini menghasilkan sistem *web* yang dapat membantu Polsek Pati dalam menerima aduan konten SARA, namun untuk proses klasifikasi pesan yang mengandung SARA atau tidak masih belum bisa dikatakan berhasil. Hal ini dikarenakan algoritma *random forest decision tree* digunakan untuk komputasi data numerik dan tidak sesuai jika digunakan untuk data berupa teks atau kata-kata. Sehingga akurasi klasifikasi tergantung pada ketersediaan kata yang ada di data latihan. Pengujian pada penelitian ini menggunakan *confusion matrix* yang mendapatkan nilai akurasi sebesar 93%.

Kata Kunci : Polsek Pati, konten SARA, Twitter, TF-IDF, *random forest*