

## ABSTRAK

### IDENTIFIKASI DAN PREDIKSI *CYBERBULLYING* PADA MEDIA SOSIAL DENGAN *RANDOM FOREST*

**Muhammad Ghifar Rasyidi**

Program Studi Statistika, Fakultas MIPA

Universitas Islam Indonesia

Kebebasan berpendapat dan berekspresi mendukung terciptanya masyarakat dan negara Indonesia yang maju dan berkembang. Namun, kebebasan berpendapat selalu disertai dengan berbagai dampak. Salah satu dampak yang biasa dialami adalah perundungan atau *cyberbullying*. *Cyberbullying* merupakan perbuatan tercela yang dilakukan seseorang dengan memposting tulisan kejam atau berkomentar semena-mena tanpa memikirkan akibat pada orang lain. Tujuan dari penelitian ini adalah untuk mengidentifikasi komentar *cyberbullying* pada sosial media agar mengerti batasan seberapa jauh seseorang harus berkomentar yang melukai perasaan orang lain. Salah satu cara untuk mengidentifikasi komentar *cyberbullying* adalah dengan menggunakan metode *Random Forest*. Dimana input yang digunakan adalah komentar-komentar pada media sosial *Facebook* dan *Twitter*, sehingga nantinya akan didapatkan suatu model prediksi yang bermanfaat dalam memprediksi komentar *cyberbullying*. Didapatkan model dengan nilai akurasi sebesar 95.98% dan error sebesar 4.02%. Tahapan tahapan seperti TF-IDF dan model klasifikasi *Random Forest* dimasukkan dalam sebuah *pipeline*. Setelah itu dikembangkan menjadi sebuah *prototype web* aplikasi menggunakan *flask* dan *Heroku* yang dapat diakses secara umum untuk memprediksi komentar *cyberbullying*.

**Kata Kunci:** *Random Forest, Cyberbullying, Pipeline, Flask, Heroku*