

INTISARI

PERBANDINGAN METODE *RANDOM FOREST* DAN *SUPPORT VECTOR MACHINE* DALAM KLASIFIKASI JUDUL BERITA *CLICKBAIT*

(Studi Kasus : Berita Online Selebriti di Berbagai Media Massa)

Apriyani Bella Thania

Program Studi Statistika, Fakultas Matematika dan Ilmu Pengetahuan Alam
Universitas Islam Indonesia

Tiap tahunnya jumlah pengguna internet di Indonesia berkembang pesat. Akibatnya media *online* yang bermunculan yang populer dikalangan masyarakat. Karakteristik umum dari berita *online* adalah sifatnya yang *real time*. Dengan kata lain berita dipublikasikan pada saat kejadian berlangsung. Media selalu menyajikan berita paling *update* setiap harinya, apalagi informasi mengenai selebriti yang tidak ada habisnya. Berita kehidupan selebriti diminati oleh pembaca. Namun dengan adanya *clickbait*, pemberitaan seringkali mengecoh, menyesatkan, dan mengelirukan pembaca karena judul yang dibuat hiperbolis tidak selalu mencerminkan isi berita. Oleh karena itu, dilakukan penelitian untuk mengklasifikasikan judul berita ke dalam kategori bukan *clickbait*, *clickbait* tidak parah, dan *clickbait* parah dengan melakukan perbandingan metode *random forest* dan *support vector machine*. Performa *random forest* lebih bagus dalam melakukan klasifikasi judul berita *clickbait* ini dan mampu menghasilkan *error* yang lebih rendah. Dari model yang didapatkan, dibangun *website* sebagai *user interface model*. Data yang dikumpulkan berupa judul berita dari berbagai portal berita *online* sebanyak 2022. Kemudian diterapkan metode SMOTE untuk mengatasi *imbalanced data* sehingga data menjadi 4890. Hasil yang didapatkan dengan perbandingan data 80% data *training* dan 20% data *testing* didapatkan nilai akurasi sebesar 96%, *precision* sebesar 96%, dan *recall* sebesar 100%.

Kata Kunci : Selebriti, *Clickbait*, *Random Forest*, *Support Vector Machine*, SMOTE