

ANALISIS SENTIMEN OPINI MASYARAKAT TENTANG KEBOCORAN DATA PDN BERDASARKAN TWITTER MENGUNAKAN *NAIVE BAYES*

Zukrufi Mas Jaudah

(180411100153)

ABSTRAK

Kebocoran data, khususnya data pemerintah, dapat menimbulkan berbagai *sentiment negative* dan positif baik secara sosial, ekonomi, maupun politik. Sentimen masyarakat terhadap insiden ini dapat bervariasi maka diperlukan penelitian yang mampu menangkap dan menganalisis sentimen publik secara komprehensif. Penelitian ini menggunakan metode *Naïve Bayes* merupakan algoritma yang umum dipakai untuk klasifikasi teks, dan juga merupakan metode machine learning yang menggunakan perhitungan klasifikasi probabilitas. *Naïve Bayes* diusulkan Thomas Bayes, yang merupakan seorang ilmuwan inggris, yaitu melakukan prediksi peluang masa depan berdasarkan pengalaman masa lalu dengan menggunakan statistik untuk melakukan perhitungan probabilitas hipotesis. Data yang dipakai dalam penelitian ini di dapatkan dengan melakukan crawling ulasan data *Twitter* dengan topik kasus kebocoran data PDN. Data sentimen yang dilakukan pengambilan merupakan ulasan yang menggunakan bahasa indonesia. Data *tweet* yang relevan dengan *keyword* topik dikumpulkan melalui proses *crawling* dari platform menggunakan *Twitter API*. Jumlah data awal yang dihasilkan dari *crawling* sebanyak 1.262. Tujuan dari penggunaan metode *Naïve Bayes Classifier* dengan jumlah data yang digunakan 1.440 ulasan setelah melakukan *SMOTE*. Selanjutnya dilakukan skenario ujicoba menggunakan *k-fold cross validation* dengan k-5 menghasilkan nilai akurasi 99.77%. Ketika menggunakan *split validation* pembagian data latih dan data uji (80%:20%), (70%:30%), (60%:40%), dan (50%:50%). Akurasi mencapai 99.67%, 99.78%, 99.80%, dan 99.78%. Sehingga pada penelitian ini diketahui bahwa penggunaan teknik *split validation* menghasilkan akurasi yang lebih tinggi ataupun lebih baik yaitu dengan akurasi 99.80%.

Kata kunci : *Naïve bayes, Analisis Sentimen, K-Fold Validation, Split Validation*