

Abstrak

Buku pelajaran dan buku cerita adalah salah satu sumber pengetahuan bagi anak-anak. Ketika buku itu dibaca, tentu anak akan berusaha memaknai setiap kata dan kalimat di dalamnya. Namun, muncul masalah ketika ditemukan kesalahan konten pada buku tersebut. Konten yang dimaksud yaitu kata-kata vulgar dan kalimat tak senonoh. Bagi anak-anak di tingkat SD konten tersebut menjadi hal yang bermakna gereflektif (tidak boleh atau tabu). Berdasarkan masalah tersebut, maka dilakukan penelitian tugas akhir terhadap cerita anak yang diambil dari buku fiksi dan buku pelajaran. Penelitian ini dilakukan dengan membangun sistem untuk mendeteksi konten gereflektif pada teks cerita yang dijadikan sebagai data set. Penelitian dilakukan dengan membangun model menggunakan algoritma klasifikasi teks Naïve Bayes Classifier (NBC). Sistem kemudian dievaluasi dengan dua skenario menggunakan metrik pengukuran akurasi, presisi, dan *recall* karena karakteristik data set yang *imbalanced* atau tidak rata dengan jumlah data pada kelas negatif lebih besar dibandingkan dengan data pada kelas positif. Berdasarkan hasil evaluasi, skenario pengujian menghasilkan rata-rata presisi yang tinggi yaitu 99,01%, sedangkan untuk nilai *recall* memiliki rata-rata di atas 50%. Dari kedua nilai tersebut, dapat disimpulkan bahwa model yang dibangun oleh sistem belum dapat mendeteksi kelas dengan baik, tetapi sangat terpercaya ketika sistem dapat membangun model dengan baik.

Kata kunci : cerita anak, konten gereflektif, Naïve Bayes Classifier