

ABSTRAK

Topik pembicaraan yang sedang hangat dibicarakan oleh khalayak ramai memiliki informasi yang beragam. Media sosial *twitter* memiliki suatu sistem yang dapat menentukan topik yang sedang ramai dibahas oleh pengguna yang disebut *trending topic*. Namun *twitter* belum dapat mengklasifikasikan topik ke dalam beberapa kategori untuk mendapatkan informasi yang lebih baik, tujuan dari penelitian ini adalah untuk mengetahui pengaruh seleksi fitur *Chi-Square* terhadap sistem klasifikasi topik ke dalam beberapa kategori dari data *tweet* berbahasa Indonesia yang diambil dari *twitter*. *Chi-Square* adalah teknik seleksi fitur yang menghitung nilai ketergantungan fitur terhadap suatu kelas. Fitur yang telah terseleksi akan digunakan dalam proses klasifikasi dengan metode Naïve-bayes Multinomial (NBM) yang didesain khusus untuk klasifikasi teks. Pengujian dilakukan dalam beberapa skenario, dan skenario dengan hasil yang terbaik diperoleh dari kombinasi rasio dataset 80 : 20, kombinasi n-gram berupa unigram + bigram + trigram, dan tanpa seleksi fitur dengan nilai akurasi 74,05%. Sehingga dapat disimpulkan bahwa seleksi fitur *Chi-Square* tidak memberikan pengaruh yang signifikan terhadap proses klasifikasi *trending topic*.

Kata kunci : twitter, trending topic, seleksi fitur, chi-square, naïve-bayes multinomial, n-gram.