

## ABSTRAK

Lirik merupakan komponen penting untuk mengetahui pesan yang terkandung dari sebuah lagu. Lirik tersebut dapat digunakan sebagai objek penelitian dalam *text categorization*. Saat ini sistem pengkategorian lagu pada aplikasi pemutar musik seperti iTunes pada iPhone, laya pada Android sudah mampu mengkategorikan lagu berdasarkan *genre*, *artist*, dan *album*. Namun belum terdapat pengkategorian lagu berdasarkan topik lagu, sehingga untuk mengkategorikan lagu berdasar topik ini masih dilakukan secara manual. Penelitian ini bertujuan untuk membuat sistem yang mampu mengkategorikan lagu berdasarkan topik lagu dengan menggunakan lirik lagu sebagai objek penelitian. Penelitian ini diawali dengan mengumpulkan lirik lagu dari berbagai *website* penyedia lirik lagu yang dijadikan sebagai *dataset* penelitian, kemudian dilakukan *preprocessing* data yang terdiri dari *case folding*, *tokenization*, *stopword removal* dan *stemming*. Tahap selanjutnya yaitu melakukan *feature selection* dengan menggunakan metode *chi-square*. Selanjutnya hasil *feature selection* tersebut dilakukan *feature extraction* dengan menggunakan metode *bag of words*. Selanjutnya dilakukan pembangunan *classifier* yang nantinya digunakan pada saat proses klasifikasi. Tahap selanjutnya yaitu melakukan klasifikasi lirik lagu pada data *testing* kedalam sebuah kelas topik tertentu dengan menggunakan metode *multinomial naïve bayes*. Tahap terakhir pada penelitian ini yaitu mengevaluasi sistem dengan menghitung *accuracy*, *precision*, *recall*, dan *f1-measure*. Kelas topik lagu yang dihasilkan pada sistem ini yaitu cinta, persahabatan, nasionalisme, keluarga, religi dan konten negatif. Performansi sistem yang dihasilkan pada penelitian ini didapatkan nilai *f1-measure* sebesar 88.91% dan nilai akurasi 96.00%.

Kata kunci: klasifikasi, lirik lagu, *chi-square*, *multinomial naïve bayes*