

Lailatus Syarifah, 2019. *Text Mining untuk Pengklasifikasian Komentar Masyarakat dalam Media Center Surabaya dengan Metode Naïve Bayes Classifier*. Skripsi dibawah bimbingan Dr. Nur Chamidah, M.Si. dan Drs. H. Sediono, M.Si., Program Studi S-1 Statistika, Departemen Matematika, Fakultas Sains dan Teknologi, Universitas Airlangga, Surabaya.

ABSTRAK

Text mining merupakan proses menggali informasi menggunakan *tools* analisis dengan tujuan mendapatkan informasi yang berguna dari sekumpulan dokumen. Sumber data yang digunakan merupakan kumpulan teks yang memiliki format tidak terstruktur. Salah satu komponen dalam *text mining* adalah pengklasifikasian dokumen. Studi kasus dalam penelitian ini yaitu pengklasifikasian komentar masyarakat Surabaya dalam *website Media Center Surabaya* berdasarkan bidang pelayanan publik yaitu bidang infrastruktur dan administrasi. Salah satu metode *text mining* yang digunakan untuk mengklasifikasikan dokumen adalah *Naïve Bayes Classifier* (NBC). Metode NBC merupakan algoritma yang sering digunakan dalam pengkategorian teks, dimana konsep dasarnya adalah menggabungkan probabilitas kata-kata dan kategori sebuah dokumen. Sebagai salah satu metode komputasi yang efisien dan mempunyai *performance predictive* yang baik, NBC merupakan salah satu metode klasifikasi teks yang populer. Berdasarkan hasil perhitungan, didapatkan ketepatan klasifikasi pada data *testing* yaitu sebesar 98% dan pada data *testing* yaitu sebesar 97%. Hal ini menunjukkan bahwa metode *Naïve Bayes Classifier* sangat baik digunakan untuk pengklasifikasian komentar masyarakat dalam *Media Center Surabaya* berdasarkan bidang pelayanan publik.

Kata Kunci : Text Mining, Media Center Surabaya, Naïve Bayes Classifier (NBC).

Lailatus Syarifah, 2019. **Text Mining for Classification of Community Comments in Media Center Surabaya with the Naïve Bayes Classifier Method.** This *skripsi* is under supervised by Dr. Nur Chamidah, M.Si. and Drs. H. Sediono, M.Si., S-1 Statistics Courses, Matematics Departement, Faculty of Sains and Technology, Airlangga University, Surabaya.

ABSTRACT

Text mining is a process of digging information using analytical tools with the aim of getting useful information from a set of documents. The data source used is a collection of texts that have an unstructured format. One component in text mining is the classification of documents. The case study in this research is the classification of Surabaya community comments on the Surabaya Media Center website based on the field of public services, namely infrastructure and administration. One of the text mining methods used to classify documents is Naïve Bayes Classifier (NBC). The NBC method is an algorithm that is often used in categorizing text, where the basic concept is to combine the probability of words and the category of a document. As one of the efficient methods of computation and good predictive performance, NBC is one of the popular text classification methods. Based on the results of the calculation, obtained the classification accuracy on the testing data equal to 98% and on the testing data equal to 97%. This shows that the Naïve Bayes Classifier method is very well used for classifying public comments in Media Center Surabaya based on the field of public services.

Keywords : *Text Mining, Media Center Surabaya, Naïve Bayes Classifier (NBC).*