

Klasifikasi Dokumen Teks Bahasa Jawa dengan Menggunakan Metode Naive Bayes

WILDAN AULIA ZAKKI

(Pembimbing : Wijanarto, M.Kom)

Teknik Informatika - S1, FIK, Universitas Dian Nuswantoro

www.dinus.ac.id

Email : 111201206643@mhs.dinus.ac.id

ABSTRAK

Seiring perkembangan jaman dan teknologi yang semakin pesat ternyata dapat mempengaruhi tingkat tutur pada generasi muda, khususnya bahasa Jawa. Lamanya memahami bahasa Jawa dikarenakan seringnya berkomunikasi menggunakan bahasa Indonesia ketimbang bahasa Jawa. Sulitnya memahami bahasa Jawa dilihat dari dua segi, yaitu komunikasi langsung (lisan) dan komunikasi tidak langsung (tulisan). Pada penelitian ini difokuskan untuk mengklasifikasikan komunikasi tidak langsung (tulisan), kedalam tiga kategori yang terdiri dari ngoko lugu, krama madya, dan krama inggil. Metode yang digunakan untuk klasifikasi adalah Naive Bayes. Uji coba dilakukan dengan menggunakan sampel dokumen teks yang diambil dari media massa elektronik berbasis web. Hasil eksperimen menunjukkan bahwa nilai accuracy dari metode Naive Bayes tidak konsisten, yang ditunjukkan pada hasil pengujian skenario 1 sampai 10. Ternyata nilai accuracy konsisten ditunjukkan pada hasil pengujian skenario 1 sampai 5 dengan nilai (0.5774- 12.5722), sedangkan nilai accuracy tidak konsisten ditunjukkan pada hasil pengujian skenario 6 sampai 10 dengan nilai (7.0351- 8.8445). Pada penelitian ini tidak menggunakan stemming pada tahap pre-processing.

Kata Kunci : Bahasa Jawa, klasifikasi teks, preprocessing, naive bayes

Classification of Javanese Text Document Using Naive Bayes Method

WILDAN AULIA ZAKKI

(Lecturer : Wijanarto, M.Kom)

*Bachelor of Informatics Engineering - S1, Faculty of Computer
Science, DINUS University*

www.dinus.ac.id

Email : 111201206643@mhs.dinus.ac.id

ABSTRACT

Along with the development of the era and technology increasingly rapidly it can affect the level of speech in the younger generation, especially the Javanese language. The duration of understanding Javanese is due to frequent communication using Indonesian language rather than Javanese. The difficulty of understanding Javanese language is seen from two aspects, namely direct communication (spoken) and indirect communication (writing). In this research is focused to classify indirect communication (writing), into three categories consisting of ngoko lugu, krama madya, and krama inggil. The method used for the classification is Naive Bayes. The test is done by using a sample of text documents extracted from web-based electronic mass media. The experimental results show that the accuracy value of the Naive Bayes method is inconsistent, which is shown in the test results of scenarios 1 through 10. The accuracy value shown is shown in the test results of scenarios 1 to 5 with values (0.5774- 12.5722), while the accuracy value is not consistent shown in Test scenario 6 to 10 with value (7.0351-88445). In this study did not use stemming at the pre-processing stage.

Keyword : Javanese language, text classification, preprocessing, naive bayes