

KLASIFIKASI KATA MENGGUNAKAN METODE NAÏVE BAYES UNTUK DOKUMEN BERBAHASA INDONESIA BERTEMA PENDIDIKAN

Tita Suharyati

Abstrak

Semakin berkembangnya teknologi menuntut agar pekerjaan dapat diselesaikan dengan cepat dan meminimalkan campur tangan dari manusia. Semakin banyak pekerjaan yang dikerjakan oleh suatu teknologi maka semakin mempermudah manusia dalam menyelesaikan pekerjaannya. Kategorisasi teks yang sering dijumpai masih banyak yang manual sehingga sistem hanya dapat menentukan kategori teks tersebut berdasarkan yang telah pengguna tetapkan sebelumnya. Oleh karena itu, dibutuhkan teknik khusus untuk mengidentifikasi kata yang relevan terhadap topik dalam sebuah dokumen berbahasa Indonesia. Dalam hal ini membahas tentang klasifikasi kata menggunakan metode Naïve Bayes. Tujuan dari penelitian ini untuk mengetahui metode Naïve Bayes dalam melakukan klasifikasi dokumen teks berbahasa Indonesia bertema pendidikan & Melakukan pengklasifikasian kata berdasarkan dokumen berbahasa Indonesia bertema pendidikan. Hal ini terlihat dari hasil perhitungan, yaitu kata yang paling banyak muncul adalah kata “program” dengan hasil probabilitas 13.4% dan kata yang paling sedikit muncul adalah kata “keilmuan” dengan hasil probabilitas 0.3% dengan jumlah dokumen yang diuji sebanyak 35 dokumen.

Kata Kunci : Klasifikasi, Naïve Bayes, Probabilitas

WORD CLASSIFICATION USING NAIVE BAYES METHOD FOR INDONESIAN LANGUAGE DOCUMENT FOR EDUCATION THEMED

Tita Suharyati

Abstract

Growing technology demanded that the job can be completed quickly and minimize the interference of man. The more work done by a technology the more ease in completing his work man. Text categorization often encountered many manual so that the system can only determine the category of the text based on the user who has set previously. Therefore, it takes a special technique to identify words that are relevant to the topic in a document language, Indonesia. In this case discusses the classification of the word Naïve Bayes method. The purpose of this research is to know the methods of Naïve Bayes document classification in doing text Indonesia-themed education based on words which do & document Indonesia education themed. This is apparent from the results of the calculation, i.e. the most words appear is the word "program" with the results of the probability of 13.4% and said that at least appears is the word "science" with the results of the probability of 0.3% with the number of the document that was tested as many as 35 of the document.

Keyword : Classification, Naïve Bayes, Probability