

## INTISARI

### TEXT CLASSIFICATION WITH NAÏVE BAYES

**Diajukan Oleh:**

**EPRI WAHYUDI  
11312068**

Pengaruh gaya hidup masyarakat era *modern* dalam mengakses informasi lebih banyak dilakukan membaca secara *online* daripada membaca di media cetak. Akibatnya, setiap hari jumlah dokumen teks khususnya dokumen berita *online* di internet semakin meningkat dan sulit dalam memperoleh informasi berita yang diinginkan dengan cepat. Sehingga diperlukan sebuah teknik pengelompokan dokumen yang dapat mengelompokkan dokumen berita ke dalam kategori yang telah ditentukan. Teknik ini adalah klasifikasi dokumen. Salah satu metode klasifikasi yang baik dan populer adalah *Naïve Bayes* atau lebih dikenal *Naïve Bayes Classifier* (NBC). Metode probabilistik *Naïve Bayes Classifier* memiliki kelebihan keserhanaan dalam komputasinya. Penelitian ini mengkaji kinerja NBC untuk kategorisasi teks dokumen berita berbahasa Indonesia. Penelitian ini menggunakan data sebanyak 270 set data. Hasil penelitian yang dihitung menggunakan tabel Matriks Confusion (*Confusion Matrix*) menunjukkan tingkat akurasi maksimal dicapai 97%. Menggunakan pemrograman Java, program klasifikasi dengan *Naïve Bayes Classifier* berhasil membuktikan bahwa algoritma ini mampu mengelompokkan dokumen ke dalam kategori yang telah ditentukan.

**Kata kunci:** *Naïve Bayes Classifier*, klasifikasi, pemrograman java, *Confusion Matrix*.

## **ABSTRACT**

### **TEXT CLASSIFICATION WITH NAÏVE BAYES**

**Submitted By:**

**EPRI WAHYUDI  
11312068**

The influence of life style in modern era society to access information by reading online than reading in printed media. As a consequence, every day the total text document especially document news online in internet increased and difficulties to get news information quickly. Finally be needed a technique to classification document that can classified news document to category decided. This technique is document classification. One of the good method classification and popular is Naïve Bayes or in the other word Naïve Bayes Classifier (NBC). Probabilistic method Naïve Bayes Classifier has the advantages, simplicity in computation. The research inspected performance of NBC for categories news document text in Indonesia language. The research used the data as many as 270 set data. The result of research that counting using table Confusion Matrix showing level maximal 97%. By using Java programing, classification program with Naïve Bayes Classifier successful to prove this algorithm master in classified document in category that decided.

**The key words:** Naïve Bayes Classifier, Classification, Java Programming, Confusion Matrix