

## KLASIFIKASI DOKUMEN BERKATEGORI MENGGUNAKAN ALGORITMA NAIVE BAYES BERBASIS BERNOULLI

<sup>1</sup>Melisa Ayu Susanti (12 1065 1104), <sup>2</sup>Bagus Setya Ryantiarna, S.Si., M.Si,

<sup>3</sup>Daryanto M.kom

Jurusan Teknik Informatika Fakultas Teknik Universitas Muhammadiyah Jember

Email : melisaayus@gmail.com

### ABSTRAK

Dalam mengelola informasi dari sekumpulan dokumen dengan jumlah yang besar merupakan sebuah kesulitan untuk mengidentifikasi kata yang ada pada dokumen tersebut menurut masing-masing kategori dari dokumen tersebut diperlukan suatu metode. *Naive Bayes Classifier* merupakan salah satu metode machine learning yang menggunakan perhitungan probabilitas. Klasifikasi teks menggunakan Naive bayes ini ada salah satu model yang dapat membantu kita mengelompokkan dokumen yaitu Bernoulli NB. Penelitian ini berusaha untuk mengklasifikasi kategori dokumen dengan menggunakan algoritma *Naive Bayes* Berbasis *Bernoulli*. Klasifikasi ini ditekankan pada kategori dokumen diantaranya Ekonomi, Kesehatan, Hiburan dan Teknologi, untuk mengetahui nilai akurasi yang akan diukur menggunakan pembobotan proses Algoritma *Naive Bayes Classifier*. Metode Bernoulli NB merupakan metode yang digunakan untuk klasifikasi sebuah teks dari kategori dokumen. Hasil pengujian klasifikasi dokumen kategori dengan menggunakan metode *Naive Bayes* Berbasis *Bernoulli* dapat mengklasifikasikan dokumen kategori dengan tingkat *Presicion* sebesar 70%, *Accuracy* 65% dan *Recall* 70% dari nilai rata-rata keseluruhan dokumen percobaan dengan tingkat nilai berbeda. Hal ini menunjukkan bahwa metode *Naive Bayes* berbasis *Bernoulli* tingkat klasifikasi dalam mengelompokkan suatu dokumen belum optimal.

**Kata Kunci** : *Naive Bayes*, *Bernoulli*, Klasifikasi Dokumen Kategori

# DOCUMENT CLASSIFICATION USING CATEGORY BASED BERNOULLI NAIVE BAYES ALGORITHM

*Melisa Ayu Susanti (12 1065 1104)<sup>1</sup>, Bagus Setya Ryantiarna, S.Si.<sup>2</sup>, M.Si,  
Daryanto M.kom<sup>3</sup>*

*Departement of Informatics, Faculty of Engineering, Muhammadiyah University*

*Email: melisaayus@gmail.com*

## ABSTRACT

In managing information from a collection of documents with a large amount of a difficulty to identify words that exist in the document according to each category of the document need a method. Naive Bayes classifier is a machine learning method that uses probability calculations. Text classification using Naive Bayes is no one model that can help us classify documents that Bernoulli NB. This study sought to classify categories of documents using a Naive Bayes algorithm Based Bernoulli. This classification is emphasized in the document categories including Economy, Health, Entertainment and Technology, to determine the accuracy value will be measured using a weighting process Naive Bayes classifier algorithm. NB Bernoulli method is the method used for the classification of a text from the document category. The test results document classification category using the method Based Bernoulli Naive Bayes can classify documents by category Precision rate of 70%, 65% and Recall Accuracy 70% of the average value of the entire document experiment with different value levels. This shows that a method based on Bernoulli Naive Bayes classification level for classifying a document has not been optimal.

**Keyword:** Naive Bayes, Bernoulli, Document Classification Category