

## ABSTRAK

### ANALISIS SENTIMEN MENGENAI KOMISI PEMBERANTASAN KORUPSI (KPK) PADA MEDIA SOSIAL TWITTER DENGAN MENERAPKAN ALGORITMA *NAÏVE BAYES CLASSIFIER*

Oleh

**HENDY SYUHADA**

Komisi Pemberantasan Korupsi (KPK) merupakan lembaga resmi di Indonesia yang ditugaskan secara khusus untuk menangani kasus korupsi. Seiring dengan maraknya kasus korupsi di Indonesia, publik turut menyampaikan opini terhadap kinerja KPK yang disampaikan salah satunya melalui Twitter. Namun opini tersebut masih samar bernada positif ataupun negatif. Oleh karena itu, pada penelitian ini dilakukan analisis sentimen terhadap opini tersebut dengan menggunakan algoritma *naïve bayes classifier* berbasis *machine learning*.

Data bersumber dari Twitter yang diambil dengan teknik *crawling* melalui API (*Application Programming Interface*). Data tersebut diolah melalui beberapa tahapan yaitu *preprocessing* yang meliputi menghilangkan tanda baca, menghapus kata yang berulang dan kata yang sering muncul tetapi tidak terlalu memiliki makna dalam kalimat. Tahap selanjutnya adalah *labeling* data yang dilakukan secara manual dengan memberikan label atau *class* pada data tersebut. Berikutnya adalah proses *modeling* yaitu proses untuk membangun sebuah model yang tepat untuk memprediksi probabilitas data yang akan masuk serta mengelompokkannya sesuai dengan perhitungan probabilitas sebelumnya. Data yang digunakan dalam proses *modeling* yaitu sebanyak 2055 data *tweet* yang dibagi menjadi *training set* dan *testing set* dengan perbandingan 80:20. Selanjutnya dilakukan *deployment system* dengan model yang dipilih untuk menganalisis sentimen terhadap KPK di Twitter.

Hasil dari penelitian ini menunjukkan dengan menggunakan *Naïve Bayes Classifier* model *multinomial* didapatkan nilai *precision* 0.69, *recall* 0.89, *F-1 Score* 0.74, dan akurasi sebesar 64%. Pada penelitian ini juga dikembangkan sebuah *website* untuk mengambil data baru yang kemudian secara otomatis mengklasifikasikannya ke dalam label positif, negatif, atau netral. *Website* ini juga menampilkan hasil dalam bentuk tabel dan grafik.

Kata kunci: Analisis Sentimen, KPK, *Multinomial model*, *Naïve Bayes Classifier*, Twitter

## **ABSTRACT**

### **SENTIMENT ANALYSIS OF THE KOMISI PEMBERANTASAN KORUPSI (KPK) ON SOCIAL MEDIA TWITTER BY APPLYING THE ALGORITHM NAÏVE BAYES CLASSIFIER**

**By**

**HENDY SYUHADA**

Komisi Pemberantasan Korupsi (KPK) is an official institution in Indonesia specifically assigned to handle corruption cases. Along with the rise of corruption cases in Indonesia, the public also expressed opinions on the performance of the KPK, which was conveyed through Twitter. However, this opinion is still vaguely positive or negative. Therefore, in this study, sentiment analysis was carried out on this opinion using the machine learning-based naïve bayes classifier algorithm.

Data comes from Twitter which is taken by crawling technique through API (Application Programming Interface). The data is processed through several stages, namely preprocessing which includes removing punctuation marks, removing repetitive words and words that often appear but do not really have meaning in sentences. The next stage is data labeling which is done manually by assigning a label or class to the data. Next is the modeling process, which is the process of building an appropriate model to predict the probability of incoming data and classifying them according to the previous probability calculations. The data used in the modeling process is 2055 tweet data which is divided into training sets and testing sets with a ratio of 80:20. Next, a system deployment with the chosen model was carried out to analyze sentiment towards the KPK on Twitter.

The results of this study indicate that using the multinomial Naïve Bayes Classifier model, the precision value is 0.69, the recall is 0.89, the F-1 Score is 0.74, and the accuracy is 64%. In this study, a website was also developed to retrieve new data which then automatically classified it into positive, or neutral labels. This website also displays the results in the form of tables and graphs.

**Keywords:** Sentiment Analysis, KPK, Multinomial model, Naïve Bayes Classifier, Twitter