

ABSTRAK

Komisi Pemberantasan Korupsi (KPK) merupakan lembaga resmi yang dibentuk dengan tujuan untuk menangani secara khusus kasus korupsi di Indonesia. Seiring berjalannya waktu setelah dibentuknya lembaga tersebut, kasus korupsi semakin meningkat yang mengakibatkan nilai fungsi dari lembaga tersebut semakin buruk dari penilaian masyarakat. Maraknya kasus korupsi membuat publik ikut serta dalam menyampaikan opininya terhadap kinerja KPK yang semakin dipertanyakan melalui berbagai media, terutama media sosial *twitter*. Berbagai macam opini yang disampaikan oleh masyarakat masih banyak kata yang ambigu atau belum dapat dipastikan apakah itu termasuk opini positif atau negatif. Oleh karena itu, pada penelitian ini dilakukan analisis sentimen terhadap opini masyarakat terhadap kinerja KPK melalui komentar di *twitter* dengan menggunakan metode *Naïve Bayes Classifier*. Penggunaan metode *Naïve Bayes Classifier* dianggap lebih unggul dibandingkan dengan metode lainnya karena lebih efektif dan sederhana, serta memiliki performa yang baik. Selain itu, metode TF-IDF juga digunakan dalam penelitian ini untuk pembobotan kata serta penggunaan SMOTE untuk mengatasi ketidakseimbangan kelas pada dataset yang didapatkan yaitu 503 data positif dan 336 data negatif. Hasil yang didapatkan pada penelitian ini dapat dilihat pada perhitungan *confusion matrix* pada penerapan *naive bayes* yang telah mengalami perlakuan SMOTE atau data *balance* yaitu 503 data positif dan 503 data negatif dengan pembagian data 80:20 untuk data latih dan data uji yang mana nilai akurasi tertinggi sebesar 82%, nilai *precision* sebesar 83%, nilai *recall* sebesar 82%, dan *f1-score* sebesar 82% dengan waktu 0,037 detik. Nilai spesifisitas yang didapatkan dari 20% data uji yaitu 91% atau nilai *True Negatif* sebanyak 93 data dan sensitivitas sebesar 73% atau nilai *True Positif* sebanyak 73 data yang artinya sentimen negatif lebih dominan dari pada sentimen positif tentang KPK di media sosial *twitter*.

Kata kunci: Analisis Sentimen, Twitter, KPK, *Naïve Bayes Classifier*,
Pembobotan TF-IDF, SMOTE.

ABSTRACT

The Corruption Eradication Commission (KPK) is an official institution established with the aim of dealing specifically with corruption cases in Indonesia. Over time after the establishment of the institution, corruption cases have increased which has resulted in the value of the function of the institution getting worse from the public's assessment. The rise of corruption cases has made the public participate in expressing their opinions on the KPK's performance, which is increasingly being questioned through various media, especially social media twitter. The various opinions conveyed by the public still have many ambiguous words or cannot be ascertained whether it is a positive or negative opinion. Therefore, in this study, a sentiment analysis of public opinion on the performance of the KPK through comments on twitter was carried out using the Naïve Bayes Classifier method. The use of Naïve Bayes Classifier method is considered superior to other methods because it is more effective and simple, and has good performance. In addition, the TF-IDF method is also used in this research for word weighting and the use of SMOTE to overcome class imbalance in the dataset obtained, namely 503 positive data and 336 negative data. The results obtained in this study can be seen in the calculation of the confusion matrix on the application of naive bayes which has undergone SMOTE treatment or data balance, namely 503 positive data and 503 negative data with a data division of 80:20 for training data and test data where the highest accuracy value is 82%, precision value is 83%, recall value is 82%, and f1-score is 82% with a time of 0.037 seconds. The specificity value obtained from 20% of the test data is 91% or a True Negative value of 93 data and a sensitivity of 73% or a True Positive value of 73 data, which means that negative sentiment is more dominant than positive sentiment about the KPK on twitter social media.

Keywords: Sentiment Analysis, Twitter, KPK, Naïve Bayes Classifier, TF-IDF Weighting, SMOTE.