

ABSTRAK

Agustina, Tria. 2025. Analisis Sentimen Komentar Youtube Terkait Kenaikan PPN 12% Dengan Multinomial Naïve Bayes. Tugas Akhir. Program Sarjana. Program Studi Teknik Informatika. Universitas Muhammadiyah Jember.

Pembimbing: Hardian Oktavianto, S.Si., M. Kom; Wiwik Suharso, M. Kom

Penerapan kebijakan PPN 12% oleh pemerintah Indonesia memunculkan berbagai tanggapan dari masyarakat, terutama di media sosial seperti *Youtube*. Penelitian ini bertujuan untuk melakukan klasifikasi terhadap komentar masyarakat di platform *Youtube* terkait kebijakan kenaikan PPN 12% dengan menggunakan *K-Fold Cross Validation* dan menggunakan algoritma *Multinomial Naïve Bayes*. Salah satu tantangan dalam klasifikasi teks adalah ketidakseimbangan kelas. Untuk mengatasi masalah tersebut digunakan teknik random oversampling guna menyeimbangkan distribusi kelas. Data yang digunakan merupakan komentar publik yang dikumpulkan dengan cara *crawling* dari video *youtube* yang membahas kenaikan ppn 12%, dengan total sebanyak 1690 data. Komentar kemudian diproses melalui *preprocessing* seperti *cleaning*, *case folding*, *tokenizing*, *normalisasi*, *stopword removal*, *stemming*. Fitur teks diekstraksi menggunakan *Term Frequency-Inverse Document Frequency (TF-IDF)* kemudian diklasifikasikan kedalam dua kategorisentimen yaitu positif dan negatif. Hasil klasifikasi menunjukkan bahwa model *Multinomial Naïve Bayes* memberikan performa yang cukup baik dengan akurasi sebesar 89.73%, presisi 88.98% dan recall 94.12%. Temuan ini menunjukkan bahwa algoritma *Multinomial Naïve Bayes* efektif dalam mengklasifikasikan sentimen publik terkait kebijakan kenaikan PPN 12% melalui analisis.

Kata kunci : Analisis Sentimen, Multinomial Naïve Bayes, Oversampling, TF-IDF, Youtube

ABSTRACT

Agustina, Tria. 2025. Sentiment Analysis Of Youtube Comments Related To The Increase Of Vat By 12% Using Multinomial Naïve Bayes. Undergraduate Thesis. Undergraduate program. Informatic Engineering Study Program. University of Muhammadiyah Jember.

Advisors: Hardian Oktavianto, S.Si., M. Kom; Wiwik Suharso, M. Kom

The implementation of the 12% VAT policy by the Indonesian government has sparked various responses from the public, especially on social media platforms like YouTube. This study aims to classify public comments on the YouTube platform regarding the 12% VAT increase policy using K-Fold Cross Validation and the Multinomial Naïve Bayes algorithm. One of the challenges in text classification is class imbalance. To address this issue, random oversampling techniques were used to balance the class distribution. The data used consists of public comments collected through crawling from YouTube videos discussing the increase of VAT by 12%, with a total of 1690 data points. Comments were then processed through preprocessing such as cleaning, case folding, tokenizing, normalization, stopword removal, and stemming. Text features were extracted using Term Frequency-Inverse Document Frequency (TF-IDF) and then classified into two sentiment categories: positive and negative. The classification results indicate that the Multinomial Naïve Bayes model performs quite well with an accuracy of 89.73%, precision of 88.98%, and recall of 94.12%. These findings demonstrate that the Multinomial Naïve Bayes algorithm is effective in classifying public sentiment regarding the policy of increasing VAT to 12% through analysis.

Keywords : Sentiment Analysis, Multinomial Naïve Bayes, Oversampling, TF-IDF, Youtube

