

ABSTRAK

Penelitian ini bertujuan untuk melakukan analisis sentimen publik terhadap implementasi Kurikulum Merdeka melalui data komentar YouTube menggunakan algoritma klasifikasi K-Nearest Neighbors (K-NN) dan Naïve Bayes. Kurikulum Merdeka merupakan program pendidikan yang dicanangkan oleh pemerintah Indonesia pada tahun 2020 sebagai upaya untuk memperbaiki kualitas pendidikan di Indonesia. Metode penelitian yang digunakan adalah kuantitatif dengan menganalisis data komentar YouTube terkait Kurikulum Merdeka. Tahapan yang dilakukan meliputi pengumpulan data, preprocessing data (case folding, cleaning, tokenisasi, normalisasi, stopword removal, dan stemming), dan pemodelan klasifikasi sentimen menggunakan algoritma K-NN dan Naïve Bayes. Evaluasi model dilakukan dengan mengukur akurasi, presisi, recall, dan nilai AUC (Area Under Curve). Hasil penelitian menunjukkan bahwa algoritma KNN memiliki performa yang lebih baik dibandingkan algoritma *Naïve Bayes* dalam mengklasifikasikan sentimen publik terhadap implementasi Kurikulum Merdeka. Algoritma *Naïve Bayes* mendapatkan nilai akurasi, presisi, *recall*, dan *f1-score* adalah sebesar 53.82%, 72%, 53.82% dan 59%. Sedangkan pada metode K-Nearest Neighbor didapatkan nilai k terbaik adalah 11 dengan menghasilkan nilai akurasi, presisi, *recall*, dan *f1-score* adalah sebesar 79.70%, 77.20%, 79.70% dan 77.65%, dengan nilai AUC sebesar 0,75 pada algoritma *gaussian naïve bayes* sehingga klasifikasi nya masuk kedalam kategori cukup sedangkan untuk algoritma KNN mendapatkan nilai AUC sebesar 0,91 maka hasil klasifikasi algoritma KNN masuk kedalam kategori sangat baik. pembandingan, menghasilkan 19.979 komentar dengan sentimen positif dan 5.108 komentar dengan sentimen negatif sedangkan untuk sentiment netral berjumlah 1.398 komentar

Kata Kunci: Analisis Sentimen, Kurikulum Merdeka, K-Nearest Neighbors, Naïve Bayes, Youtube

ABSTRACT

This study aims to analyze public sentiment towards the implementation of the Merdeka Curriculum through YouTube comment data using the K-Nearest Neighbors (K-NN) and Naïve Bayes classification algorithms. The Merdeka Curriculum is an educational program launched by the Indonesian government in 2020 to improve the quality of education in Indonesia. The research method used is quantitative, analyzing YouTube comment data related to the Merdeka Curriculum. The steps taken include data collection, data preprocessing (case folding, cleaning, tokenization, normalization, stopword removal, and stemming), and sentiment classification modeling using the K-NN and Naïve Bayes algorithms. Model evaluation is done by measuring accuracy, precision, recall, and AUC (Area Under Curve) values. The results show that the K-NN algorithm has better performance compared to the Naïve Bayes algorithm in classifying public sentiment towards the implementation of the Merdeka Curriculum. The Naïve Bayes algorithm achieved accuracy, precision, recall, and f1-score of 53.82%, 72%, 53.82% and 59% respectively. Meanwhile, the K-Nearest Neighbor method with the best k value of 11 achieved accuracy, precision, recall, and f1-score of 79.70%, 77.20%, 79.70% and 77.65%, with an AUC value of 0.91, indicating very good classification performance. In comparison, the analysis of 26,539 comments resulted in 19,979 positive, 5,108 negative, and 1,398 neutral sentiments, showing a dominant positive sentiment. The lower performance of the Naïve Bayes algorithm is due to the violation of the independence assumption of features and the class imbalance in the dataset, causing Naïve Bayes to be biased towards the majority class.

Keywords: *Sentiment Analysis, Merdeka Curriculum, K-Nearest Neighbors, Naïve Bayes, YouTube*

Keywords : *KNN; K-Nearest Neighbor; Data Mining, OpenSID*