

ANALISIS SENTIMEN PON PAPUA XX PADA PLATFORM X DENGAN METODE RANDOM FOREST

Zulfikar Yusuf Baiquni

(190411100120)

Pembimbing I : Dr. Bain Khusnul Khotimah, S.T., M.Kom

Pembimbing II : Husni, S.Kom., MT.

ABSTRAK

Penelitian ini bertujuan untuk melakukan analisis sentimen terhadap Pekan Olahraga Nasional (PON) Papua XX yang diselenggarakan pada tahun 2021, dengan memanfaatkan data dari platform media sosial X (Twitter). Dengan meningkatnya interaksi di media sosial, penting untuk memahami persepsi publik terhadap acara tersebut. Dataset yang digunakan terdiri dari 2000 tweet berbahasa Indonesia yang diklasifikasikan dengan *TextBlob* ke dalam tiga kelas sentimen: positif, negatif, dan netral. Untuk mengatasi masalah ketidakseimbangan kelas, metode SMOTE (Synthetic Minority Over-sampling Technique) diterapkan, yang menghasilkan sampel sintesis untuk kelas minoritas. Klasifikasi dilakukan menggunakan algoritma Random Forest, yang dikenal efektif dalam menangani data tekstual besar. Penelitian ini mengevaluasi kinerja model dengan membagi dataset menjadi beberapa skenario, yaitu 80:20, 70:30, 60:40, dan 50:50, serta membandingkan hasil klasifikasi antara model dengan dan tanpa penerapan SMOTE. Hasil evaluasi menunjukkan bahwa penggunaan SMOTE tidak selalu meningkatkan performa model, dengan akurasi tertinggi mencapai 71% pada skenario 60:40. Penelitian ini diharapkan dapat memberikan wawasan yang berguna bagi penyelenggara PON dalam meningkatkan kualitas acara di masa depan dan menjadi referensi bagi penelitian sejenis di bidang analisis sentimen.

Kata Kunci: Analisis sentiment, Pon Papua, *Random Forest*, SMOTE

ABSTRACT

This research aims to conduct sentiment analysis on the Papua XX National Sports Week (PON) held in 2021, utilizing data from the social media platform X (Twitter). With the increasing interactions on social media, it is essential to understand public perceptions of the event. The dataset consists of 2000 tweets in Indonesian, classified using TextBlob into three sentiment classes: positive, negative, and neutral. To address the issue of class imbalance, the SMOTE (Synthetic Minority Over-sampling Technique) method is applied, generating synthetic samples for the minority class. Classification is performed using the Random Forest algorithm, which is known to be effective in handling large textual data. This study evaluates the model's performance by splitting the dataset into several scenarios: 80:20, 70:30, 60:40, and 50:50, and compares the classification results between models with and without the application of SMOTE. Evaluation results indicate that the use of SMOTE does not always enhance model performance, with the highest accuracy reaching 71% in the 60:40 scenario. This research is expected to provide valuable insights for the PON organizers to improve the quality of future events and serve as a reference for similar research in the field of sentiment analysis.

Keywords: Sentiment analysis, Papua PON, Random Forest, SMOTE