

ABSTRACT

Based on a survey by the Indonesian Internet Service Providers Association (APJII), the number of internet users in Indonesia reached 215.63 million people in the 2022-2023 period, an increase of 2.67% compared to the previous period. This increase has caused a surge in the number of online news which requires better data management, especially in dealing with data set class imbalances in data classification. This research aims to overcome this problem by applying the SMOTE technique, which generates new samples for minority data set classes to improve data representation. In addition, the KNN algorithm is used to evaluate the effect of the combination of SMOTE and KNN on the performance of the classification model. Evaluation is carried out using accuracy, precision, recall and F1-Score metrics. The research results show that the application of SMOTE succeeded in improving the performance of the classification model. The best combination was obtained at the parameter value $k=1$, with an accuracy of 62.50%, precision of 58.39%, recall of 86.96%, and F1-Score of 69.87%. Compared with the model before SMOTE implementation, there was an increase in accuracy performance of 58.33%, precision of 49.56%, and F1-Score of 63.28%, while maintaining a recall of 87.50%. This research proves that SMOTE is effective in dealing with data set class imbalance, resulting in more accurate and balanced model predictions. The research results contribute to the management of online news data to support better classification quality.

Keywords : AI, KNN , Online News, SMOTE