

ABSTRAK

Skripsi ini bertujuan untuk menerapkan metode *Term Frequency-Inverse Document Frequency* (TF-IDF) dan algoritma *K-Nearest Neighbor* (K-NN) dalam pengembangan sistem deteksi *clickbait* pada judul berita *online* berbahasa Indonesia. Pendekatan ini mencakup fase *input* dengan pengumpulan data judul berita, proses transformasi menggunakan TF-IDF untuk menganalisis frekuensi kata, dan penerapan K-NN untuk klasifikasi *clickbait*. Hasilnya adalah sebuah program web yang dapat mendeteksi *clickbait*, disertai dengan *file* teks yang berisi data hasil deteksi. Evaluasi sistem menunjukkan tingkat akurasi sebesar 1.0, *precision* sebesar 1.0, *recall* sebesar 1.0, dan *f1-score* sebesar 1.0, menunjukkan hasil yang cukup baik dalam mengenali *clickbait* pada judul berita *online* berbahasa Indonesia. Penelitian ini memberikan kontribusi pada pemrosesan bahasa alami dan pengembangan sistem deteksi *clickbait* pada lingkungan berbahasa Indonesia, serta dapat diaplikasikan pada jurnal dengan merinci variabel, metode, dan hasil evaluasi secara rinci.

Kata Kunci: TF-IDF, *k-Nearest Neighbor*, *Clickbait*, Judul Berita Online, Bahasa Indonesia.

ABSTRACT

This study aims to implement the Term Frequency-Inverse Document Frequency (TF-IDF) method and the K-Nearest Neighbor (K-NN) algorithm in developing a clickbait detection system for online news headlines in the Indonesian language. The approach encompasses an input phase involving the collection of news headline data, a transformation process using TF-IDF to analyze word frequencies, and the application of K-NN for clickbait classification. The result is a web program capable of detecting clickbait, accompanied by a text file containing the detection system's data. System evaluation indicates an accuracy rate of 1.0, precision of 1.0, recall of 1.0, and an F1-score of 1.0, demonstrating fairly good results in identifying clickbait in Indonesian-language online news headlines. This research contributes to natural language processing and the development of clickbait detection systems in the Indonesian language environment, and it can be applied to journals by detailing variables, methods, and evaluation results comprehensively.

Keywords: TF-IDF, k-Nearest Neighbor, Clickbait, Online News Headlines, Indonesian.