

## **ABSTRAK**

**Damai Arsila Salsabila.** ANALISIS DAN PERBANDINGAN STOPWORD TERHADAP AKURASI ANALISIS SENTIMEN TEKS DENGAN MENGGUNAKAN TF-IDF STUDI KASUS NLP (dibimbing oleh Fahrin Irhamna Rachman S.Kom., M.T. dan Titin Wahyuni S.Pd., M.T).

Dalam era digital yang berkembang pesat, jumlah data teks online meningkat signifikan, mencakup ulasan produk, komentar media sosial, dan artikel berita. Analisis sentimen penting untuk memahami opini masyarakat. Penelitian ini bertujuan untuk mengembangkan daftar stopwords yang lebih relevan menggunakan algoritma TF-IDF untuk meningkatkan representasi teks dalam analisis sentimen, serta mengevaluasi dan membandingkan pengaruh penggunaan stopwords yang dihasilkan oleh algoritma TF-IDF terhadap akurasi model analisis sentimen, dibandingkan dengan penggunaan stopwords Sastrawi. Hasil menunjukkan TF-IDF membantu mengidentifikasi kata-kata kurang penting, namun stopwords Sastrawi lebih baik mengenali konteks. Evaluasi dengan rasio pembagian data yang berbeda (90:10, 80:20, 70:30) menunjukkan akurasi tertinggi sebesar 0.789 pada rasio 80:20, meskipun ada ruang untuk peningkatan. Studi ini diharapkan dapat meningkatkan kinerja model analisis sentimen dengan daftar stopwords yang lebih sesuai.

**Kata Kunci :** Analisis Sentimen, TF-IDF, NLP, Stopwords

## ABSTRACT

**Damai Arsila Salsabilah.** *ANALYSIS AND COMPARISON OF STOPWORDS ON TEXT SENTIMENT ANALYSIS ACCURACY USING TF-IDF: A CASE STUDY IN NLP (supervised by Fahrin Irhamna Rahman S.Kom., M.T. and Titin Wahyuni S.Pd., M.T).*

*In the rapidly evolving digital era, the amount of online text data has significantly increased, encompassing product reviews, social media comments, and news articles. Sentiment analysis is crucial for understanding public opinion. This research aims to develop a more relevant stopword list using the TF-IDF algorithm to enhance text representation in sentiment analysis. Additionally, it evaluates and compares the impact of using stopwords generated by the TF-IDF algorithm on the accuracy of sentiment analysis models, compared to using Sastrawi stopwords. The results show that TF-IDF helps identify less important words, but Sastrawi stopwords are better at recognizing context. Evaluation with different data split ratios (90:10, 80:20, 70:30) showed the highest accuracy of 0.789 at the 80:20 ratio, although there is room for improvement. This study is expected to improve the performance of sentiment analysis models with a more suitable stopword list.*

**Keywords:** *Sentiment Analysis, TF-IDF, NLP, Stopwords.*