ABSTRAK

Sobaro, Nur Kholifatun Nisais. 2026. Analisis Perbandingan *Word2Vec* dan *N-Gram* pada Klasifikasi Sentimen Komentar Instagram Terkait Gaji Guru Honorer Menggunakan *Support Vector Machine*.

Pembimbing: (1) Moh. Dasuki, M.Kom.; (2) Ari Eko Wardoyo, S.T., M.Kom.

Penelitian ini bertujuan untuk membandingkan metode ekstraksi fitur *Word2Vec* dan *N-Gram* berbasis TF-IDF dalam klasifikasi sentimen komentar Instagram terkait gaji guru honorer menggunakan algoritma *Support Vector Machine* (SVM). Data penelitian berupa komentar Instagram yang telah melalui tahap *preprocessing* dan diekstraksi menggunakan *Word2Vec Skip-Gram*, *Continuous Bag of Words* (CBOW), TF-IDF Unigram, Bigram, dan Trigram. Untuk mengatasi ketidakseimbangan data digunakan metode *Synthetic Minority Over-sampling Technique* (SMOTE). Pengujian dilakukan menggunakan kernel *Linear*, *Polynomial*, *Radial Basis Function* (RBF), dan *Sigmoid* dengan evaluasi *5-Fold Cross Validation*. Hasil penelitian menunjukkan bahwa metode TF-IDF Unigram dengan kernel *Polynomial* memperoleh akurasi tertinggi sebesar 94,87%, diikuti oleh *Word2Vec* CBOW dengan kernel RBF sebesar 87,04%, TF-IDF Bigram dengan kernel RBF sebesar 86,35%, *Word2Vec Skip-Gram* dengan kernel *Polynomial* sebesar 85,78%, dan TF-IDF Trigram dengan kernel RBF sebesar 77,22%. Berdasarkan hasil pengujian, TF-IDF Unigram dengan kernel *Polynomial* menghasilkan performa klasifikasi terbaik dibandingkan metode ekstraksi fitur dan kernel lainnya. Selain menghasilkan akurasi tertinggi, model tersebut juga memperoleh nilai *precision, recall,* dan *F1-Score* masing-masing sebesar 95%.

Kata Kunci : *Analisis Sentimen, Word2Vec, N-Gram , TF-IDF, Support Vector Machine, Gaji Guru Honorer, Instagram*

ABSTRACT

Sobaro, Nur Kholifatun Nisais. 2026. *Comparative Analysis of Word2Vec and N-Gram for Sentiment Classification of Instagram Comments Related to Honorary Teacher Salaries Using Support Vector Machine.*

***Advisors*:** (1) Moh. Dasuki, M.Kom; (2) Ari Eko Wardoyo, S.T., M.Kom.

This study aims to compare the Word2Vec and TF-IDF-based N-Gram feature extraction methods for sentiment classification of Instagram comments related to honorary teacher salaries using the Support Vector Machine (SVM) algorithm. The research data consisted of Instagram comments that had undergone preprocessing and were represented using Word2Vec Skip-Gram, Continuous Bag of Words (CBOW), TF-IDF Unigram, Bigram, and Trigram feature extraction methods. To address the class imbalance problem, the Synthetic Minority Over-sampling Technique (SMOTE) was applied. The classification process was evaluated using Linear, Polynomial, Radial Basis Function (RBF), and Sigmoid kernels with 5-Fold Cross Validation. The results showed that the TF-IDF Unigram method with the Polynomial kernel achieved the highest accuracy of 94.87%, followed by Word2Vec CBOW with the RBF kernel at 87.04%, TF-IDF Bigram with the RBF kernel at 86.35%, Word2Vec Skip-Gram with the Polynomial kernel at 85.78%, and TF-IDF Trigram with the RBF kernel at 77.22%. Based on these results, TF-IDF Unigram with the Polynomial kernel demonstrated the best classification performance among all feature extraction methods and kernel combinations. In addition to achieving the highest accuracy, the model also obtained 95% precision, 95% recall, and a 95% F1-score..

Keywords : *Sentiment Analysis, Word2Vec, N-Gram , TF-IDF, Support Vector Machine, Honorary Teacher Salaries, Instagram*