

**PERBANDINGAN EKSTRAKSI FITUR *TF-IDF* DAN *WORD2VEC* PADA  
ANALISIS SENTIMEN MENGGUNAKAN MODEL *FINE-TUNING*  
INDOBERT UNTUK ULASAN GAME LOKAL DI STEAM**

Oleh

**I Gusti Ngurah Agung Pernata, NIM 2015091033**

**Jurusan Teknik Informatika**

**Program Studi Sistem Informasi**

**ABSTRAK**

Perkembangan industri game di Indonesia telah menghasilkan berbagai video game baru, mendorong pengembang untuk memperhatikan pendapat dan preferensi pengguna. Salah satu platform yang digunakan untuk mengevaluasi ulasan pengguna adalah Steam. Ulasan tersebut kemudian dapat dianalisis dengan menggunakan model seperti IndoBERT, yang saat ini populer dalam klasifikasi teks seperti analisis sentimen. Dalam proses memasukkan teks ke dalam model, ekstraksi fitur diperlukan, dengan TF-IDF dan Word2Vec menjadi dua metode ekstraksi fitur umum yang digunakan untuk analisis sentimen. TF-IDF fokus pada frekuensi kata, sementara Word2Vec menangkap hubungan makna kata. Tujuan dari penelitian ini yakni untuk membandingkan kinerja TF-IDF dan Word2Vec dalam analisis sentimen menggunakan model Fine-Tuning IndoBERT pada ulasan game lokal di platform Steam. Hasil penelitian menunjukkan bahwa TF-IDF memiliki kinerja yang lebih baik, dengan akurasi rata-rata sebesar 97%, presisi 86%, recall 70%, dan F1-Score 76%. Sebaliknya, Word2Vec menunjukkan akurasi rata-rata sebesar 96%, presisi 83%, recall 63%, dan F1-Score 68%. Selain itu, dari dataset yang digunakan, terlihat bahwa mayoritas ulasan pengguna cenderung bersifat positif, dengan 3.849 ulasan positif dari total 4.027 ulasan dalam dataset.

Kata Kunci: Steam, Analisis Sentimen, Ekstraksi Fitur, TF-IDF, Word2Vec, *fine-tuning* IndoBERT.

**COMPARISON OF TF-IDF AND WORD2VEC FEATURE EXTRACTION  
ON SENTIMENT ANALYSIS USING THE INDOBERT FINE-TUNING  
MODEL FOR LOCAL GAME REVIEWS ON STEAM**

**By**

***I Gusti Ngurah Agung Pernata, NIM 2015091033***

***Informatic Engineering***

***Information Systems Study Program***

**ABSTRACT**

*The growth of the gaming sector in Indonesia has produced a variety of new video games, encouraging developers to pay attention to user opinions and preferences. One platform used to evaluate user reviews is Steam. These reviews can then be analyzed using models such as IndoBERT, which is currently popular in text classification technique like sentiment analysis. In the process of inputting text into a model, feature extraction is necessary, with TF-IDF and Word2Vec being two common feature extraction methods used for sentiment analysis. TF-IDF focuses on word frequency, while Word2Vec captures word-meaning relationships. This research aims to compare the performance of TF-IDF and Word2Vec in sentiment analysis using the IndoBERT Fine-Tuning model on local game reviews on the Steam platform. The research results show that TF-IDF has better performance, with an average accuracy of 97%, precision of 86%, recall of 70%, and F1-Score of 76%. In contrast, Word2Vec shows an average accuracy of 96%, precision of 83%, recall of 63%, and F1-Score of 68%. In addition, from the dataset used, it can be seen that the majority of user reviews tend to be positive, with 3,849 positive reviews out of a total of 4,027 reviews in the dataset.*

*Keywords: Steam, Sentiment Analysis, Feature Extraction, TF-IDF, Word2Vec, fine-tuning IndoBERT.*