

## ABSTRAK

Persepsi publik terhadap kebijakan strategis nasional seperti Rencana Pembangunan Jangka Panjang Nasional (RPJPN) 2025–2045 dapat diamati melalui media sosial, yang mencerminkan opini masyarakat secara *real time*. Penelitian ini bertujuan untuk mengevaluasi distribusi sentimen publik terhadap RPJPN dan mengkaji efektivitas kombinasi model IndoBERT dan algoritma *Support Vector Machine* (SVM) dalam klasifikasi sentimen. Pendekatan penelitian menggunakan kerangka metodologi SEMMA (*Sample, Explore, Modify, Model, Assess*). Data diperoleh melalui *crawling* dari media sosial X, disaring dan diberi label manual, kemudian dianalisis secara eksploratif untuk memahami distribusi awal sentimen. Data dimodifikasi dengan *preprocessing*, tokenisasi, dan *balancing* menggunakan *RandomOverSampler*. Model IndoBERT di *fine-tuning* untuk menghasilkan representasi *vektor* teks, kemudian digunakan sebagai *input* pada klasifikasi SVM yang dioptimalkan dengan *GridSearch*. Evaluasi model dilakukan menggunakan *confusion matrix* dan metrik klasifikasi seperti *accuracy*, *precision*, *recall*, dan *F1-score*. Hasil menunjukkan bahwa sentimen publik awal didominasi oleh opini positif 80%, namun setelah *balancing*, performa model meningkat secara signifikan. Model IndoBERT+SVM menghasilkan akurasi 98,75% dan meningkat menjadi 99,13% setelah *tuning* dioptimalkan dengan *GridSearch*. Hasil ini menunjukkan bahwa integrasi IndoBERT dan SVM dalam kerangka SEMMA mampu mengklasifikasikan sentimen positif, negatif, dan netral secara akurat, serta memberikan gambaran yang representatif tentang persepsi publik terhadap kebijakan RPJPN. Penelitian ini juga memperlihatkan bahwa pendekatan berbasis NLP dengan dukungan teori pemrosesan bahasa alami dan klasifikasi *machine learning* dapat memperkuat pengambilan kebijakan berbasis data.

**Kata Kunci :** IndoBERT, SVM, SEMMA, RPJPN, *GridSearch*

## ABSTRACT

*Public perception of national strategic policies such as the National Long-Term Development Plan (RPJPN) 2025–2045 can be observed through social media, which reflects public opinion in real time. This study aims to evaluate the distribution of public sentiment toward the RPJPN and assess the effectiveness of combining the IndoBERT model and Support Vector Machine (SVM) algorithm in sentiment classification. The research approach uses the SEMMA (Sample, Explore, Modify, Model, Assess) methodological framework. Data was obtained through crawling from social media platform X, filtered and manually labeled, then analyzed exploratively to understand the initial sentiment distribution. Data was modified through preprocessing, tokenization, and balancing using RandomOverSampler. The IndoBERT model was fine-tuned to generate text vector representations, then used as input for SVM classification optimized with GridSearch. Model evaluation was conducted using a confusion matrix and classification metrics such as accuracy, precision, recall, and F1-score. The results showed that initial public sentiment was dominated by positive opinions (80%), but after balancing, model performance improved significantly. The IndoBERT+SVM model achieved an accuracy of 98.75% and increased to 99.13% after optimization via GridSearch. These results indicate that the integration of IndoBERT and SVM within the SEMMA framework can accurately classify positive, negative, and neutral sentiments, providing a representative overview of public perception toward the RPJPN policy. This study also demonstrates that an NLP-based approach supported by natural language processing theory and machine learning classification can strengthen data-driven policy-making.*

**Keywords:** Sentiment Analysis, IndoBERT, SVM, SEMMA, RPJPN