ANALISIS SENTIMEN PROGRAM NATURALISASI TIM NASIONAL SEPAK BOLA INDONESIA PADA MEDIA SOSIAL X MENGGUNAKAN INDOBERT EMBEDDING DAN CONVOLUTIONAL NEURAL NETWORK

Oleh

I Putu Arya Vidyananta, NIM 2215091079

Jurusan Teknik Informatika

Program Studi Sistem Informasi

ABSTRAK

Program naturalisasi pemain Tim Nasional Sepak Bola Indonesia menimbulkan beragam tanggapan di kalangan masyarakat, khususnya di media sosial X. Perbedaan pandangan tersebut mendorong perlunya analisis sentimen untuk memahami kecenderungan opini publik terhadap program naturalisasi. Penelitian ini bertujuan untuk memodelkan analisis sentimen, mengevaluasi performa model klasifikasi, serta mengidentifikasi kecenderungan sentimen masyarakat terhadap program naturalisasi Tim Nasional Sepak Bola Indonesia. Data penelitian berupa tweet berbahasa Indonesia yang dikumpulkan dari platform X pada periode 1 Januari 2024 hingga 31 Desember 2025 dengan total 12.173 *tweet*. Berdasarkan hasil pelabelan, distribusi sentimen terdiri dari 48,7% sentimen negatif, 17,7% sentimen netral, dan 33,6% sentimen positif. Metode yang digunakan adalah Indo*BERT embedding* untuk merepresentasikan teks dengan memahami makna kata sesuai dengan posisinya dalam kalimat. Hasil tersebut kemudian diproses menggunakan *Convolutional Neural Network (CNN)* yang berfungsi sebagai model klasifikasi sentimen ke dalam tiga kelas, yaitu positif, negatif, dan netral. Evaluasi performa model dilakukan menggunakan skema *K-Fold Cross Validation* 5 dan 10 dengan variasi ukuran kernel 2, 3, dan 5 serta penerapan *Synthetic Minority Oversampling Technique* (SMOTE). Hasil penelitian menunjukkan bahwa model CNN dengan *kernel size* 2 dan 3 menghasilkan performa yang lebih baik dibandingkan *kernel size* 5. Performa terbaik tanpa penerapan SMOTE diperoleh pada konfigurasi *kernel size* 3 dengan 10-fold *cross validation*, dengan nilai akurasi sebesar 0,89304 dan *f1-score* sebesar 0,87496.

Kata kunci: Analisis Sentimen, Naturalisasi, IndoBERT Embedding, *Convolutional Neural Network*, Media Sosial X.

SENTIMENT ANALYSIS OF THE INDONESIAN NATIONAL FOOTBALL TEAM NATURALIZATION PROGRAM ON SOCIAL MEDIA X USING INDOBERT EMBEDDING AND CONVOLUTIONAL NEURAL NETWORK

By

I Putu Arya Vidyananta, NIM 2215091079

Informatics Engineering

Information Systems Study Program

ABSTRACT

The player naturalization program in the Indonesian National Football Team has generated various responses among the public, particularly on the social media platform X. These differing views encourage the need for sentiment analysis to understand the tendency of public opinion toward the naturalization program. This study aims to model sentiment analysis, evaluate the performance of the classification model, and identify the tendency of public sentiment toward the naturalization program of the Indonesian National Football Team. The research data consist of Indonesian-language tweets collected from the X platform during the period of January 1, 2024, to December 31, 2025, with a total of 12,173 tweets. Based on the labeling results, the sentiment distribution consists of 48.7% negative sentiment, 17.7% neutral sentiment, and 33.6% positive sentiment. The method used is IndoBERT embedding to represent text by understanding the meaning of words according to their positions in a sentence. The results are then processed using a Convolutional Neural Network (CNN), which functions as a sentiment classification model into three classes: positive, negative, and neutral. The model performance evaluation is conducted using 5-fold and 10-fold Cross Validation schemes with variations in kernel size and the application of the Synthetic Minority Oversampling Technique (SMOTE). The results of the study indicate that the CNN model with kernel sizes of 2 and 3 produces better performance compared to a kernel size of 5. The best performance without the application of SMOTE is obtained with a kernel size 3 configuration using 10-fold cross validation, achieving an accuracy value of 0.89304 and an f1-score of 0.87496.

Keywords: Sentiment Analysis, Naturalization, IndoBERT Embedding, Convolutional Neural Network, Social Media X.