ABSTRACT

WULAN FADHILAH, NATURAL LANGUAGE PROCESSING TO ANALYZE PUBLIC SENTIMENT TOWARDS THE 2024 GENERAL ELECTION ON TWITTER USING THE SUPPORT VECTOR MACHINE ALGORITHM

The digital era and advances in information technology have given people easy access to voice opinions and interact via social media such as Twitter. Social media is crucial in capturing public sentiment regarding the 2024 General Election. By conducting sentiment analysis, it is possible to avoid the spread of fake news (hoaxes), attempts to influence negative opinions, and hate speech that can damage a healthy political atmosphere so that the election process can take place with integrity. and transparency. This study uses sentiment analysis using the Support Vector Machine algorithm to categorize negative, positive and neutral sentiment with a dataset containing 4,438 tweets. The dataset was collected from September 2023 to January 2024. Data preprocessing and processing was then carried out by labeling it using Inset Lexicon and the word weighting process using TF-IDF. This dataset is divided into 80% training data and 20% test data. The processed training data will later be used to train the Support Vector Machine model. The results of the classification with this model obtained 694 positive sentiment data, 188 negative data, and 14 neutral data. Model evaluation using the test set resulted in an accuracy of 79.5759%, recall of 82.1492%, precision of 58.3696%, and F1-Score of 60.5354%. This study is also implemented into a website that can display public sentiment and its potential for broader topic analysis. Overall, this research reveals the public's views on the 2024 General Election and can be used as material for future evaluation to understand the public's perspective via Twitter.

Keywords: Sentiment analysis, General Election 2024, Natural Language Processing, Support Vector Machine, Twitter

ABSTRAK

WULAN FADHILAH, NATURAL LANGUAGE PROCESSING UNTUK MENGANALISIS SENTIMEN MASYARAKAT TERHADAP PEMILIHAN UMUM TAHUN 2024 DI TWITTER MENGGUNAKAN ALGORITMA SUPPORT VECTOR MACHINE

Era digital dan kemajuan teknologi informasi telah memberikan masyarakat akses mudah untuk menyuarakan pendapat dan berinteraksi melalui media sosial seperti Twitter. Media sosial ini menjadi krusial dalam menangkap sentimen publik terhadap Pemilihan Umum 2024. Dengan dilakukan analisis sentimen dapat untuk menghindari penyebaran berita palsu (*hoax*), upaya memengaruhi opini negatif, serta ujaran kebencian yang dapat merusak atmosfer politik yang sehat sehingga proses pemilu dapat berlangsung dengan integritas dan transparansi. Studi ini menggunakan analisis sentimen menggunakan Algoritma *Support Vector Machine* untuk mengategorikan sentimen negatif, positif, dan netral dengan dataset berisi 4.438 *tweets*. Dataset dikumpulkan dari bulan September 2023 hingga Januari 2024. Kemudian dilakukan *Preprocessing* data dan *processing* dengan memberi label menggunakan *Inset Lexicon* dan proses pembobotan kata menggunakan TF-IDF. Dataset ini dibagi menjadi 80% data latih dan 20% data uji. Data pelatihan yang telah diproses nantinya digunakan untuk melatih model *Support Vector Machine*. Hasil klasifikasi dengan model tersebut diperoleh sentimen positid sebanyak 694 data, negatif 188 data, dan netral 14 data. Evaluasi model dengan menggunakan set pengujian menghasilkan akurasi sebesar 79,5759 %, *recall* sebesar 82,1492%, presisi sebesar 58,3696%, dan F1-Skor sebesar 60,5354%. Studi ini juga mengimplementasikan ke dalam sebuah website yang dapat menampilkan sentimen publik dan potensinya untuk analisis topik yang lebih luas. Secara keseluruhan, penelitian ini mengungkapkan pandangan masyarakat terhadap Pemilihan Umum 2024 dan bisa menjadi bahan evaluasi ke depannya untuk memahami perspektif masyarakat melalui Twitter.

Kata kunci: Analisis sentimen, Pemilihan Umum 2024, *Natural Language Processing*, *Support Vector Machine*, Twitter