

ABSTRAK

Twitter telah menjadi salah satu media sosial utama dalam diskursus politik di Indonesia., menghasilkan data teks tidak terstruktur dalam jumlah besar yang sulit diklasifikasikan secara manual. Penelitian ini bertujuan membangun representasi data Twitter dalam bentuk tiga jenis graf, merancang model *Multi-View Graph Neural Network* (GNN) untuk klasifikasi topik politik, serta menganalisis pola penyebaran dan keterkaitan antar pengguna berdasarkan hasil klasifikasi. Data tweet direpresentasikan sebagai *Semantic Content Graph* menggunakan embedding IndoBERT dan kemiripan kosinus, *User Interaction Graph* berdasarkan relasi mention, reply, dan retweet, serta *Temporal Propagation Graph* berdasarkan urutan waktu propagasi. Ketiga graf berbagi 15.131 node tweet dengan jumlah edge masing-masing 155.428, 184.751, dan 103.055. Model *Multi-View Graph Attention Network Fusion* dikembangkan untuk mengintegrasikan ketiga graf menggunakan mekanisme attention yang dipelajari dari data secara adaptif, dan dilatih dalam kerangka semi-supervised dengan 1.677 tweet berlabel dari total 15.131 node. Hasil pengujian menunjukkan bahwa model mencapai *Accuracy* 84,23%, *Macro Precision* 82,52%, *Macro Recall* 83,76%, dan *Macro F1-Score* 83,04%. Analisis attention mengungkapkan bahwa *Semantic Content Graph* memberikan kontribusi terbesar sebesar 54,53%, diikuti *Temporal Propagation Graph* (26,99%) dan *User Interaction Graph* (18,48%). Dari 15.131 tweet yang dianalisis, 7.996 tweet (52,84%) diklasifikasikan sebagai konten politik, sementara sisanya merupakan konten non-politik yang mencakup beragam topik termasuk aktivitas komersial dalam jumlah signifikan, menunjukkan heterogenitas percakapan dalam ekosistem Twitter Indonesia.

Kata Kunci: *Graph Neural Network*, klasifikasi topik politik, *multi-view learning*, IndoBERT,

Twitter

ABSTRACT

Twitter has become a primary platform for political discourse in Indonesia, generating large volumes of unstructured text data that are difficult to classify manually. This study aimed to construct Twitter data representations as three graph types, develop a Multi-View Graph Neural Network (GNN) model for political topic classification, and analyze propagation patterns and user relationships based on classification results. Tweet data were represented as a Semantic Content Graph using IndoBERT embeddings and cosine similarity, a User Interaction Graph based on mention, reply, and retweet relationships, and a Temporal Propagation Graph based on chronological propagation sequences. All three graphs shared 15,131 tweet nodes with 155,428, 184,751, and 103,055 edges respectively. A Multi-View Graph Attention Network Fusion model was developed to integrate the three graphs using a fully learned attention mechanism, trained under a semi-supervised setting with 1,677 labeled tweets out of 15,131 total nodes. The model achieved an Accuracy of 84.23%, Macro Precision of 82.52%, Macro Recall of 83.76%, and Macro F1-Score of 83.04%. Attention analysis revealed that the Semantic Content Graph contributed most at 54.53%, followed by the Temporal Propagation Graph (26.99%) and User Interaction Graph (18.48%). Of 15,131 analyzed tweets, 7,996 (52.84%) were classified as political content, while the remainder covered diverse non-political topics including a significant volume of commercial activities, reflecting the heterogeneity of discourse in the Indonesian Twitter ecosystem.

Keywords: *Graph Neural Network, political topic classification, multi-view learning, IndoBERT, Twitter*