INTERPRETABLE SENTIMENT ANALYSIS VIA EEG–TEXT COGNITIVE GRAPH ALIGNMENT

Author's M. Nishar Fathima[1], G. Thilipkumar[2], K. Sujith[3], N. Vanjulavalli[4]

Abstract

Understanding human sentiment requires integrating both linguistic information and cognitive signals from the brain. This project presents an interpretable multimodal framework titled "Interpretable Text Sentiment Determination by Aligning Cognition Between EEG-Derived Brain Graph and Text-Derived Knowledge Graph." The proposed system combines natural language processing with brain signal analysis to improve sentiment classification accuracy while providing meaningful explanations of model decisions. Text data are first processed using standard preprocessing techniques such as tokenization and normalization. From the processed text, a Text-Derived Knowledge Graph (KG) is constructed to represent semantic relationships between words based on co-occurrence patterns. In parallel, synthetic EEG signals representing multiple brain channels and frequency bands are generated and transformed into an EEG-Derived Brain Graph (BG) using functional connectivity measures.Graph neural network (GNN) encoders are then applied separately to the knowledge graph and the brain graph to obtain structural embeddings that capture semantic and cognitive patterns. To model the relationship between linguistic understanding and neural responses, a contrastive alignment mechanism is introduced. This mechanism learns a shared latent representation that aligns the embeddings of the knowledge graph with those of the brain graph, enabling the model to capture correlations between textual sentiment and EEG-based cognitive responses. The aligned representations are fused and passed through a neural classification layer to determine the sentiment category (positive, neutral, or negative.