Text Mining Algorithm for Large-Scale Social Media Data

Authors : Alexander A. Kharlamov, Maria Pilgun

Abstract : This paper presents the validation results of a text mining algorithm applied to urban system development in the transportation sector, leveraging user-generated content from social media platforms. The study employed sentiment analysis, aggression detection, semantic network and core formation, associative search, associative network development, and word association analysis. Data collection was conducted using the Brand Analytics social media and media monitoring system. Data analysis and interpretation utilized TextAnalyst 2.32, GPT-3.5, GPT-4, and GPT-4o, while Tableau was used for interactive visualization and analytics. Social tension levels were assessed through calculated indices of social stress and well-being. Based on the findings, recommendations were proposed to improve project effectiveness by integrating residents' perspectives.

Keywords : social media, text mining, neural network technologies., large-scale data
Conference Title : ICEMNLP 2025 : International Conference on Empirical Methods in Natural Language Processing
Conference Location : Copenhagen, Denmark
Conference Dates : June 12-13, 2025