

Emotional Analysis of Education and Gender Discussions in Dari and Pashto Twitter Conversations Using Topic Modeling

Abdul Razaq Vahidi
Karabuk University, Turkey

Abstract

Afghanistan, with its diverse linguistic and cultural landscape, has witnessed increasing polarization in online public discourse since 2021. This study focuses on analyzing Dari and Pashto Twitter conversations concerning "education" and "gender" following the Taliban's rise to power. This study aims to provide a comprehensive view of Dari and Pashto-speaking Twitter users' opinions on these topics. Using of approximately 50,000 Dari and Pashto tweets, this research employs keyword frequency analysis, clustering algorithms, and trend analysis over time. By leveraging Natural Language Processing (NLP) and unsupervised Machine Learning (ML), emotional analysis is conducted to classify tweets into various emotional categories related to "education" and "gender." The study examines September 2021 to May 2023, revealing thematic and discussion changes. Particularly after the Taliban's rise to power, social media activity intensified, with widespread campaigns on education and women's rights. This research highlights the central issues in Afghanistan's public discourse and the polarization trends in the online public sphere. Emotional analysis is based on Ekman's six categories (happiness, sadness, fear, surprise, disgust, anger) and Other in the months following the fall of the Republic and the Taliban's ascent. This analysis provides a deeper understanding of the socio-political dynamics and contributes to the broader discourse on political polarization in Afghanistan. Considering numerous campaigns on women's education, we analyze and compare findings in both Dari and Pashto.

Keywords: Social networking, public discourse, polarization, socio-political dynamics, women's rights