Tu chiamale se vuoi emozioni: The impact of emotions on disinformation and sexism identification

Paolo Rosso[1]

[1]Universitat Politècnica de València, Spain

1. Abstract

Harmful information in social media, both disinformation and hate speech, is pervasive. Emotional content can be exploited by malicious agents to provoke inflammatory responses and exacerbate polarisation in online debate. Disinformation often is conveyed through text with a high emotional valence. Emotions have in fact a key role: fake news tend to trigger different emotions and of different intensity to the users compared to texts containing truthful information. Readers prefer high arousal news, which catches more attention than regular news and sticks easily on mind. Moreover, the emotional bias may distort the perception of argumentation, for instance bringing one person to refute to acknowledge facts that upset them. In this talk I will mainly focus on disinformation and how this problem is perceived and addressed at EU level, being disinformation campaigns (often from abroad) considered as a form of information warfare and an attempt to trigger negative emotions and create polarization in our democracies. I will also show what kind of emotional responses may trigger narratives where conspiracy theories are mentioned. The language of these narratives is inherently different, and it works on a subconscious, emotional layer, exploiting the readers' biases that may distort the perception of argumentation, for instance bringing one person to refute to acknowledge facts that upset them. In the last part of the talk, I will address the problem of sexism identification in memes, showing what kind of emotions are triggered and how emotional signals can be extracted with physiological data collected via eye-tracking, heart rate, and EEG signals.

2. Short Bio

Paolo Rosso studied Computer Science in Italy at the University of Pisa, did his PhD in Ireland at the Trinity College Dublin, and ended up in Spain at the Technical University of Valencia where he's Full Professor and a member of the Pattern Recognition and Human Language Technology (PRHLT) Research Center and the Valencian Graduate School and Research Network of Artificial Intelligence (ValgrAI). He was advisor of 30 PhD students and at the moment he is of seven. He published more than 400 peer-reviewed papers in conferences and journals. During the last years, he was the PI of several research projects, such as FairTransNLP_Stereotypes: Fairness and Transparency for equitable NLP applications in social media–Identifying stereotypes and prejudices and developing equitable systems, FAKEnHATE-PoC: FAKE news and HATE speech, and XAI-DisInfodemics: eXplainable AI for disinformation and conspiracy detection during infodemics, funded by MCIN/AEI and by European Union NextGenerationEU/PRTR. His research interests focus on AI for social media related tasks such as the detection of harmful information, both fake news and hate speech. He collaborated with the Spanish National Security Department and with the Science and Technology Office (Oficina-C) of the Spanish Congress of Deputies in topics related to disinformation campaigns and AI, and with the Spanish Observatory on racism and xenophobia (OBERAXE). He helped in the organisation of more than 50 shared tasks in evaluation forums such as SemEval, CLEF, FIRE, IberEval/Iberlef and of course Evalita

EVALITA 2026: 9th Evaluation Campaign of Natural Language Processing and Speech Tools for Italian, Feb 26 – 27, Bari, IT

prosso@dsic.upv.es (P. Rosso)

(MultiPRIDE, AMI, SardiStance, IronITA, SENTIPOLC). At the moment he is actively involved in the analysis of physiological signals (EEG variables, heart rate, blinks, fixations, saccades) collected during the viewing of sexist memes and videos. He currently serves as Associate Editor at IEEE Transactions on Affective Computing.

Personal Webiste. https://personales.upv.es/prosso/