



Contribution ID : 146

Type : **Oral presentation**

Enhancing Sentiment Analysis Using Formal Linguistic Tools

Generative Artificial Intelligence (GAI) text production is crucial to research fields as Data Science (DS) and Network Textual Data Analysis (NTDA), the main purposes of GAI being to simulate human language production, exploiting both Machine Learning (ML) and Large Language Models (LLMs).

However, as pre-trained probabilistic models, LLMs are biased when built on non-perfectly balanced data as for retrieval sources, taxonomy, ontology interconnections and linguistic inference. This is most relevant to DS and NTDA, as it can contribute in social media to spreading fake news, conspiracy theories, counterproductive narratives, and online hate speech. Equally relevant is GAI being devoid of a reality formal model, causing GAI to have no ethics, as it cannot identify and correct its inaccuracies. This brings LLMs and GAI to suffer from effectiveness and reliability issues, showing tendency to prompt incorrect and discriminatory information, and hallucinations.

Newborn Neuro-Symbolic Artificial Intelligence (NSAI) tries to cope with these issues building elementary ontologies to integrate human symbolic reasoning principles with ML and Artificial Neural Networks (ANNs). Here we will demonstrate that better results come integrating also formalized morphosyntactic and semantic information, as those relating to Italian negation grammar. Therefore, to tackle on-line hate speech, we propose here a method of Sentiment Analysis (SA) that uses NooJ software to build formal ontologies and syntactic grammars within graphs representing finite state automata/transducers. While ontologies will conceptualize sets of word having contiguous contextualized meanings, syntactic grammars will parse texts using Italian formalized morphosyntax and semantics.

Keywords/Topics

Generative Artificial Intelligence, Data Science, Network Textual Data Analysis, Large Language Models, Rule-Based Natural Language Processing, Finite-State Automata and Transducers, NooJ, NooJ Grammars

Primary author(s) : MONTELEONE, Mario (Dipartimento di Scienze Politiche e della Comunicazione, Università degli Studi di Salerno; X23 Science in Society, Bergamo.)

Session Classification : Networks of Hate: Analysing User and Language Patterns in Racist Discourse

Track Classification : Networks of Hate: Analysing User and Language Patterns in Racist Discourse