

Factuality Challenges in the Era of Large Language Models

Prof. Preslav Nakov (Mohamed bin Zayed University of Artificial Intelligence, United Arab Emirates)

We discuss the risks, the challenges, and the opportunities that Large Language Models (LLMs) bring regarding factuality. We then present some recent work on using LLMs for fact-checking, on detecting machine-generated text, and on fighting the ongoing misinformation pollution with LLMs. Finally, we present a number of LLM fact-checking tools recently developed at MBZUAI: (i) LM-Polygraph, a tool to predict an LLM's uncertainty in its output using cheap and fast uncertainty quantification techniques, (ii) Factcheck-Bench, a fine-grained evaluation benchmark and framework for fact-checking the output of LLMs, (iii) Loki, an open-source tool for fact-checking the output of LLMs, developed based on Factcheck-Bench and optimized for speed and quality, (iv) OpenFactCheck, a framework for fact-checking LLM output, for building customized fact-checking systems, and for benchmarking LLMs for factuality, (v) LLM-DetectAIve, a tool for machine-generated text detection, and (vi) FRAPPE, a FRAming, Persuasion, and Propaganda Explorer.