

Gunther Martens (Ghent University)

Proof or spoof. Experimenting with diarizing vs. generating online oral debates

Abstract

DH2024

Recent research (Pianzola et al. 2020, Rebora 2021, Salgaro 2022) has explored the possibility of utilizing cultural analytics to ingest and chart online data related to literary taste. Now that we have arrived at a more thorough knowledge of how the opinions of platform-based critics compare to those of professional literary experts in the context of literary prizes (De Greve 2024), further pretraining and/or finetuning is needed to tackle full-scale literary criticism. Unlike our previous data sources, however, this type of criticism is largely proprietary. This paper aims to explore to what extent the precision of the language model we trained for parsing "casual" and oral literary criticism can be improved by using either synthetically generated data or by stop-gapping via transcribed oral data. The annotated corpus for this experiment will consist of automatically diarized transcriptions of online literary criticism. Its analysis will allow for more targeted detection of sentiments expressed about specific aspects (e.g. contender, nominated book, jury) on different platforms and by different groups. While the demography of those engaging in the online evaluation cannot be fully grasped due to anonymity, we will present an attempt to sharpen the profile of the audience by contrasting it with a parallel corpus of online comments on a popular casting show (DSDS), revealing more striking differences in habitus, expressivity and choice of criteria. The paper discusses the results of the annotation process, the annotation system, and the technical challenges faced by the project. With the arrival of LLM (Large Language Models), the slow process of annotation has received a boost from the ability to generate fully annotated "mock data". It will be explored to what extent both real and artificial training data can be helpful in the process of shedding light on the evolving role of traditional gatekeepers and the emergence of new literary gatekeepers, providing insights into the evaluative criteria used by different groups to evaluate literature. While review-aggregating platforms for films and music (like "Rotten Tomatoes") are commonly consulted for aggregated review scores (Seaver 2022), the field of literature still prefers to rely on individual gatekeepers rather than the collective wisdom of the cloud or recommendation systems, despite the untapped potential of this data."