Rage Against the Machine Learning (RATML)

Master project for Digital Media & Computer Science
Winter Term 2020/21

Abstract

The project applies advanced visualization and machine learning techniques to systematically analyze the recommendations of YouTube's machine learning algorithm. The goal is to reveal possible algorithmic biases and to envision novel algorithmic transparency tools that enable users to influence their recommendations.

Description

Research showed that a large amount of global news consumption is due to YouTube's recommendations. With YouTube's growing importance as a news platform, its recommendation algorithm came under increased scrutiny. Media reports described machine learning algorithms as radicalization tools that lure users into a rabbit hole of conspiracy or hyperpartisan videos. One example of this are the 2018 Chemnitz protests in Germany, where a stabbing spawned large far-right demonstrations and rioting. According to the New York Times, this was fueled by YouTube recommendations. Users who wanted to inform themselves about the stabbing were directed towards extremist videos by the algorithm [1]. The Times cites an analysis of Chemnitz-related videos which suggests that YouTube's recommendation consistently directed users towards "predominantly conspiracy theorist or far-right" videos about the incidents.

The project will enable students to work on a challenging real-world problem with a large-scale dataset. A first step of the project will be collecting YouTube search results, video recommendations and advertisements for different topics, both for users with user accounts and for those without user accounts. We will then systematically analyze the recommendations by the machine learning systems. Informed by this analysis, we will design, implement, and evaluate algorithmic transparency tools that help users understand and influence the recommendations. We will also explore the commercial or political influence of machine learning algorithms that make recommendations. This connects to a large body of research that showed that the algorithmic biases in machine learning systems can lead to gender and ethnic discrimination. Such algorithms can also enact popularity biases and emotionality biases in the recommendations.

In the project, participants will learn how to formulate research questions, identify and analyze user requirements, and design and develop visualization and machine learning models. We follow a research-oriented learning paradigm, i.e. the project is self-organized. Participants will not only learn about research in data science but also gain experience in managing a long-term project. Programming experience is required to succeed in this project.

Preparatory courses:
Data Science (Summer Term 2020, Hendrik Heuer)

Contact:
Hendrik Heuer, hheuer@uni-bremen.de

References:
[1] Max Fisher and Katrin Bennhold. 2018. As Germans Seek News, YouTube Delivers Far-Right Tirades. The New York Times (Sept. 2018). https://www.nytimes.com/2018/09/07/world/europe/youtube-far-right-extremism.html