Mining public opinion using language models

Facilitator

Dr. Christina Haag, Digital & Mobile Health Group, Institute for Implementation Science in Health Care, University of Zurich

Description

Large-scale text data, such as social media or press articles, have become a highly relevant source of information for the analysis of public discourse and opinion. Online text data have become so influential that 'infodemiological' research is now investigating how information spreads in electronic media (especially the Internet) and among populations, and what factors determine the nature of this spread. A common interest in public health research is to examine how particular health policy decisions or recommendations are received by the general public – whether they are supported or whether they may cause fear or anger instead. This type of analysis is called 'opinion mining'. An example of this was the preventive measures taken during the pandemic or the vaccination campaigns. The way these issues were discussed on social media and covered in different types of media both reflected public opinion and, conversely, influenced public discourse.

But how do language models work? And how can they be used responsibly in research? This course will (1) cover the basics of how language models work from an applied research perspective, (2) explain how they can be used to answer common research questions aimed at mining public opinion, and (3) finally discuss challenges of using language models (e.g., underlying biases, stereotypes) and principles of 'responsible AI'.

Objectives

By the end of the course, participants will have a basic understanding of language models, how to apply them in the context of public health research, and what to consider in the responsible use of language models.

Specifically, participants will:

1. Discuss the potential and risks of using language models in public health research.

2. Evaluate and test how different language models

perform when mining public opinion using publicly available AI resources.

Ann Walser
Graduate Campus Manager
awalser@ssphplus.ch
https://ssphplus.ch/en/graduate-campus

3. Develop competencies for the responsible use and

critical evaluation of language AI

Date	**13 June 2024, 14:00 – 16:00**
Level	Basic
Eligibility	The course is open to PhD and MD students of the SSPH+ Inter-university Graduate Campus. It is aimed at PhD students who want to learn about language models and explore how they can be used for opinion mining in their own research. This is a basic level course, and no programming skills are required for course attendance. If you choose to complete the assignment, basic Python skills will be an advantage. You also can complete the assignment without any prior knowledge of Python as you will be provided with well commented sample code. However, the assignment is likely to take longer to complete if you have never used Python before.
Course Structure	The course will be delivered online and will consist of short lectures, group discussions and hands-on exploration of language models suitable for opinion mining. The short lectures will provide an introduction to language models in general and opinion mining in the context of public health research in particular. Participants will work in teams (breakout groups) to discuss questions about the potential, challenges, and responsible use of language models. As part of the hands-on approach of the course, participants will also explore publicly available pre-trained language models using the rich resources of the established AI platform Hugging Face (https://huggingface.co/). The website interface allows for convenient online testing of the language models during the course (e.g. https://huggingface.co/j-hartmann/sentiment-roberta-large-english-3-classes). The results and conclusions of the groups will be presented and discussed in plenary at the end of the 2-hour online course.
Work load	Preliminary Work: 2 hours; Contact time: 2 h
Credits	(1 ECTS for active participation in 4 2h-online courses and 1 passed assessment in 1 course)
Assessment	(voluntary)

Ann Walser
Graduate Campus Manager
awalser@ssphplus.ch
https://ssphplus.ch/en/graduate-campus

If participants choose to do an assessment, they will conduct an opinion mining analysis on a public dataset of their choice (public datasets containing, e.g., tweets on vaccination or Long Covid) and using a language model that they have chosen during the course. They will also produce a short report interpreting the results, including a critical evaluation of potential limitations of the chosen language model.

Location	Online course on zoom
Course Fees	SSPH+ PhD Students: 0.- CHF
Registration	www.conftool.com/ssph-phd-courses2024
Deadline for registration	13 May 2024

This course is an SSPH+ IGC initiative to provide post-doctoral scientists at SSPH+ network institutions to teach.

Ann Walser
Graduate Campus Manager
awalser@ssphplus.ch
https://ssphplus.ch/en/graduate-campus