Computational Analysis of Social Media

Week 3 (regular course)

Duration: 12 hours/ 3 days

Format: In person and online

Instructor: Carlos Arcila Calderón

Course description:

This workshop focuses on the application of supervised and unsupervised analysis techniques to social media messages that can help social researchers to detect social trends or citizens' attitudes. Participants will learn the mathematical representation of text (BoW, Embeddings, BERT), the basis of natural language processing (tokenizing, stemming, lemmatizing, stop words, frequency of words, etc.) and will learn lexicon-based sentiment analysis and topic modelling. The workshop will cover supervised machine learning applied to textual data with shallow and deep learning. It will also introduce the use of network and multimedia data to the analysis of social media. All the examples will be implemented in R.

Day 1	9 – 10.45	Introduction to Computational Methods in unstructured data Preprocessing text
	10.45 – 11.15	Break
	11.15 - 13	Text as data
Day 2	9 – 10.45	Social Media Sources for data analysis Dictionary Approaches to text analysis

	10.45 – 11.15	Break
	11.15 - 13	Unsupervised text analysis: Topic modelling
Day 3	9 – 10.45	Supervised text analysis I. Shallow learning Supervised text analysis II. Deep learning
	10.45 – 11.15	Break
	11.15 - 13	Network data and Multimedia data

Prerequisites: Some bases in statistics is desirable.

Software: R

Readings: It is recommended to read the following paper as introductory reading.

Van Atteveldt, W., Triling, D. & Arcila, C. (2022). Computational Analysis of Communication. Wiley.

Short instructor biography:

Dr. Carlos Arcila Calderón, Associate Professor at the Department of Sociology and Communication at the University of Salamanca, Spain. PhD in Communications and Master in Data Science.

Specialist in computational methods in social sciences.