

# Computational analysis of social media

Week 3 (Regular course)

Duration: 12 hours/3 days

## Course Description

This workshop focuses on the application of the supervised and unsupervised texts analysis technique to real-time Twitter messages that can help social researchers to detect social trends or citizens' attitudes. Participants will learn the very basis of programming in Python and will be able to retrieve tweets with Streaming and Rest APIs of Twitter. In addition, they will learn the mathematical representation of text (BoW, Embeddings), 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 focus on supervised machine learning applied to unstructured data with different algorithms implemented in Scikit-Learn (naive bayes, logistic regression, support vector machines, etc.) and TensorFlow / Keras (RNN).

## Schedule

Day 1

- Introduction to Computational Methods in unstructured data
- Basic programming on Python 3.4
- Mathematical representation of text

Day 2

- Data retrieval of tweets with the Streaming and Rest APIs of Twitter
- Natural language processing with NLTK
- Lexicon-based sentiment analysis
- Topic modelling

Day 3

- Supervised machine learning
- Supervised sentiment analysis and visualization of tweets in real time
- Supervised text classification with deep learning
- Prerequisites
- Some bases in statistics is desirable.

## Software

We will use the language Python.

## References

- Van Atteveldt, W., Triling, D. & Arcila, C. (2022). Computational Analysis of Communication. Wiley.
- Bird, S., Klein, E., & Loper, E. (2009). Natural language processing with Python: analyzing text with the natural language toolkit. O'Reilly Media, Inc.
- Géron, A. (2019). Hands-on machine learning with Scikit-Learn, Keras, and TensorFlow: Concepts, tools, and techniques to build intelligent systems. O'Reilly Media, Inc..

## Short 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.