

## Detecting Misinformation and its Sources on Social Media

Responsible:

Dimitris Kotzinos [Dimitrios.Kotzinos@cyu.fr](mailto:Dimitrios.Kotzinos@cyu.fr)

Recently a lot of discussion has emerged on the issue of “*fake news*” and “*misinformation*” in social media (mainly) and in other news and information sources. Either by accident, lack of proper knowledge or deliberately a lot of news’ items, posts or expressed facts that appear on social media and news sources on the web are simply wrong. The term “*misinformation*” is used to describe all these cases, while we refer to “*fake news*” when we know that the news item is deliberately false.

Our goal in this project is not to understand or judge anybody’s motives for sharing erroneous information but to detect that this information is wrong by any means possible, using other more credible sources of information available online. This means that we need to label first some of the existing sources as trustworthy.

Then we would to explore various sides of information available within and around a news item:

- We need to check the trustworthiness of the initial source of the item (for example in the case of Twitter we do not check how trustworthy is the one that retweets but the original poster of the tweet). We might also need to check this in comparison with other sites and not individually, e.g. site X is more trustworthy than site Y which is more trustworthy than site Z, etc.
  - We might also need to explore issues like the life span of the site, the history of the site (e.g. do we detect a lot of wrong information coming from this site) and so on.
  - We might also want to check if the news on a site are the same as the news on other sites, so as to group them together and deal with them as one.
- We need to check the type and context of information, for example usually fake news are discussed around names of politicians or controversial popular topics. We might want to detect if those are in our items.
- The context and the information in the item can also guide our decisions:
  - numerical values are easier to check for validity and the same is true for historically registered events
  - emerging news items could be more difficult to detect; a combination of criteria could help decide in this case.
- Finally, the time frame of the news and of the detection is important: we do not need to check every news item as it appears but we need to respond before it becomes important. So thresholds for starting the process should be specified and processes that would allow us to decide on whether a news item is significant enough to deal with should be specified.

So, the idea is to provide tools that will monitor an ever expanding set of news sources (including social media) in order to identify “*misinformation*” and alert the users, providing at the same time some information on why this is flagged as *false* and where the information for this decision came from.

The project is envisioned for 3 to 4 students. The students will acquire knowledge in the use of Natural Language Processing techniques, Text analysis techniques, Scoring functions, Machine Learning and Cloud Computing and Cloud Data Management.