

# MLNS: MACHINE LEARNING IN NETWORK SCIENCE

## CENTRALESUPÉLEC

### Kaggle Challenge

Instructor: Fragkiskos Malliaros  
TA: Vahan Martirosyan

Due: **March 24, 2024 at 23:00**

## Description

**Goal:** Predicting missing links in an actor co-occurrence network. Edges have been deleted at random from a graph. Your mission is to accurately reconstruct the initial network using both graph-theoretical and node feature information.

In this competition, we define an actor network where nodes represent actors and edges between two nodes stand for co-occurrence on the same Wikipedia page. This is a proxy for their joint participation in the same movie, for their personal relation, for their level of fame, etc. In addition to the graph structure, each node (i.e., actor) is also associated with textual information processed from its Wikipedia page, from which some keywords were extracted. Only the processed features are available here. Your goal is to utilize information from both the underlying actor network and the processed Wikipedia description in order to accurately predict missing edges.

## Evaluation and Rules of the Competition

The assignment is a Kaggle competition. Detailed information about the competition, the rules, the evaluation and the data can be found in the Kaggle website:

<https://www.kaggle.com/t/84f2376058844771ae4589add897d5ce>

Please read in detail the rules of the competition and the evaluation (grading) process.

## How to Submit

Please complete the second assignment in groups of **3-4** students (preferably, the same team as in the project of the course). No late assignments will be accepted.

1. **Kaggle submission:** submission of your solution in the kaggle platform (team submission – pick also a name for your team).
2. **Report:** *typeset* your report (PDF file only). **In your report, you should mention the name of your team in Kaggle.** The submission of the report (max **3 pages**) should be made on gradescope (Kaggle Challenge; Entry Code: 6GYREX). **Only one member of each team will do the submission, adding the rest of the students as group members.**
3. **Code:** prepare a .zip file (`code_name_of_your_team.zip`) containing the code that is needed to reproduce your submission. The file should be sent by email to: `netsci.class.centralesupelec@gmail.com`, subject: “MLNS - Kaggle Challenge Code - Name of Your Team”.