Semantic Data Management Project

OBJECTIVES

This is a research-oriented project where students must put into practice the lessons learnt in this course. The project has a strong component based on critical reasoning and then a practical aspect to make you realise how to use graphs in real Data Science projects.

In particular, the goal is to either use a property graph or a knowledge graph to model your data, and exploit the resulting graph with data analysis (via graph embeddings) in real-world use cases (e.g., recommendations).

PROJECT STATEMENT

The project topic is Real use cases of property/knowledge graphs and their exploitation via data analysis.

The project has two components: a research aspect and a practical aspect to understand the potential of graphs in real use cases. You are expected to do a search on the topic, summarize what you found, and apply the approach to a real use case (as realistic as possible). More precisely, you are expected to:

- Learn and understand what graph embeddings are. Graph embeddings provide a vector representation of graphs (node-based or edge-based) that can be later used to perform data analysis on them. A starting point for this task could be: https://arxiv.org/abs/1709.07604. You need to summarize the main findings about graph embeddings in at most 5 pages.
- Put into practice what you learnt. Given a real graph (e.g., one of those generated in the labs, or any other you may have access to), briefly describe it to understand its meaning, propose and implement an embedding strategy, and run a chosen machine learning or data mining algorithm with a specific purpose. You need to summarize the practical part in at most 5 pages.

DELIVERABLES

By the deadline stated in this event, one person of the group must upload a zip file including the document in pdf format, as well as the code implementing your use case (graph embedding + execution of an analytics algorithm) in Python or Java.

The document must be 10 pages long. This is a hard constraint, as part of the project's objective is to synthesize and effectively summarize your findings.

PROJECT TEAMS

Use the team creator event to register your project. The team must be of 2 people and can be a repetition of a lab team. So be sure to choose a teammate you feel comfortable working with.

EVALUATION CRITERIA

The project will be evaluated according to the following criteria:

Conciseness

The document fits in 10 pages and explains all the details requested.

Understandability

You provide enough details as to assess your solution.

Soundness

There are no contradictions about the choices made and the inherent advantages of the topic.

Proof of Concept

A proof of concept has been developed and applied to a graph in order to obtain some meaningful result.