



Thesis Project Form

Title (tentative): Graph-based analysis of temporal evolution in Multiple Sclerosis scientific research

Thesis advisor(s): Barla Annalisa, Ilaria Stanzani

E-mail: Annalisa.Barla@unige.it

Address:

Phone: (+39) 010 353 6602

Description

Motivation and application domain

This thesis is cast in the artificial intelligence domain.

Analyzing the temporal evolution of MS research through graph-based methods can reveal trends, gaps, and advances, aiding in better understanding and future breakthroughs.

General objectives and main activities

General Objectives:

The thesis revolves around the understanding of how MS research evolves over time:

1. Learn and apply advanced graph network analytics methods to scientific research data.
2. Develop machine learning models to analyze the temporal evolution of Multiple Sclerosis (MS) research.
3. Make use of data from OpenAlex to explore trends, collaborations, and knowledge diffusion in MS studies.
4. Identify key research milestones and emerging topics in MS research through graph analysis.

Main Activities:

1. Study advanced graph network analytics techniques, including node centrality, clustering, and community detection.
2. Learn and apply machine learning methods suitable for temporal and graph-structured data.
3. Clean and preprocess the OpenAlex dataset for analysis, focusing on MS-related publications.
4. Construct and analyze co-authorship networks from the dataset.
5. Interpret findings and identify gaps or opportunities for future MS research directions.

Training Objectives (technical/analytical tools, experimental methodologies)

1. Machine Learning: Understand and apply machine learning techniques, including supervised and unsupervised learning, for graph and temporal data analysis.
2. Graph Analytics: Master advanced graph theory concepts, such as centrality, community detection, and network dynamics, and apply them to analyze scientific research networks.
3. Language Models: Learn how to use language models for processing and analyzing textual data related to scientific publications.
4. Data Cleaning: Acquire skills in cleaning, transforming, and structuring complex datasets, particularly from the OpenAlex repository, for analysis.
5. Evaluation Metrics: Develop an understanding of key evaluation metrics used to assess the performance of machine

learning models and graph analytics methods.

6. **OpenAlex APIs**: Gain proficiency in using OpenAlex APIs to extract and handle relevant data, focusing on publications and authors in the field of Multiple Sclerosis.

7. **Python**: Enhance Python programming skills, particularly in the context of data science, machine learning, and graph-based analysis, leveraging libraries like NetworkX, Scikit-learn, and PyTorch.

Place(s) where the thesis work will be carried out: Malga

Additional information

Maximum number of students: 1