UNIVERSITY OF GENOA
DEPARTMENT OF INFORMATICS, BIOENGINEERING, ROBOTICS AND SYSTEMS ENGINEERING
MASTER'S PROGRAM IN BIOENGINEERING

Thesis Project Form

Title (tentative): Depicting the temporal evolution of Multiple Sclerosis scientific research with NLP and Machine Learning

Thesis advisor(s): Barla Annalisa

E-mail: Annalisa.Barla@unige.it

Address:

Phone: (+39) 010 353 6602

Description

Motivation and application domain

Exploring the application of NLP and machine learning techniques to analyze and understand trends, patterns, and key insights in Multiple Sclerosis scientific research. By leveraging advanced data analytics, this project aims to enhance our comprehension of the evolving landscape in MS studies, identifying gaps, trends, and future directions for more effective research and treatment strategies.

General objectives and main activities

Objectives:

1. Analyze the temporal evolution of Multiple Sclerosis (MS) research using NLP and machine learning.
2. Identify key trends, milestones, and emerging topics in MS scientific research.
3. Develop models to track changes in research focus over time.
4. Visualize and interpret the insights for improved understanding of MS research dynamics.

Main Activities:

1. Collect and preprocess MS-related research data from OpenAlex.
2. Apply NLP techniques to extract relevant information from scientific papers.
3. Use machine learning to model and analyze trends in the dataset.
4. Create visualizations and reports based on the findings.

Training Objectives (technical/analytical tools, experimental methodologies)

NLP Techniques: Learn to apply NLP methods for extracting and analyzing information from scientific texts.
Machine Learning: Develop skills in machine learning to model and analyze temporal data trends.
Data Processing: Gain expertise in data preprocessing and cleaning for effective analysis.
Visualization: Learn to create meaningful visualizations to interpret and communicate findings.
API Utilization: Master the use of OpenAlex APIs to gather and handle research data efficiently.
Programming: Master Python language to implement Machine Learning pipelines

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

Additional information

Maximum number of students: 1