Algorithm: **Graph Neural Network**, and another set of classification, improving accuracy (Boosting, random forest, k-nearest neighborhood, Logistic Regression, feature-based classification)
Topic: Classification
Data format: Json, csv, database files….

Topics:
Problem: Detecting a malicious account, URL on Twitter… social network
Step: (Pipeline)

- Research examples of malicious accounts on twitter or any other social network.
- Data (Twitter profile) scrape using Twitter API, and/or python libraries.
- Possibly set up a database (Probably non-relationship) → Exploratory data analysis
- Machine learning.
- Result - It can be a application (Dashboard), API

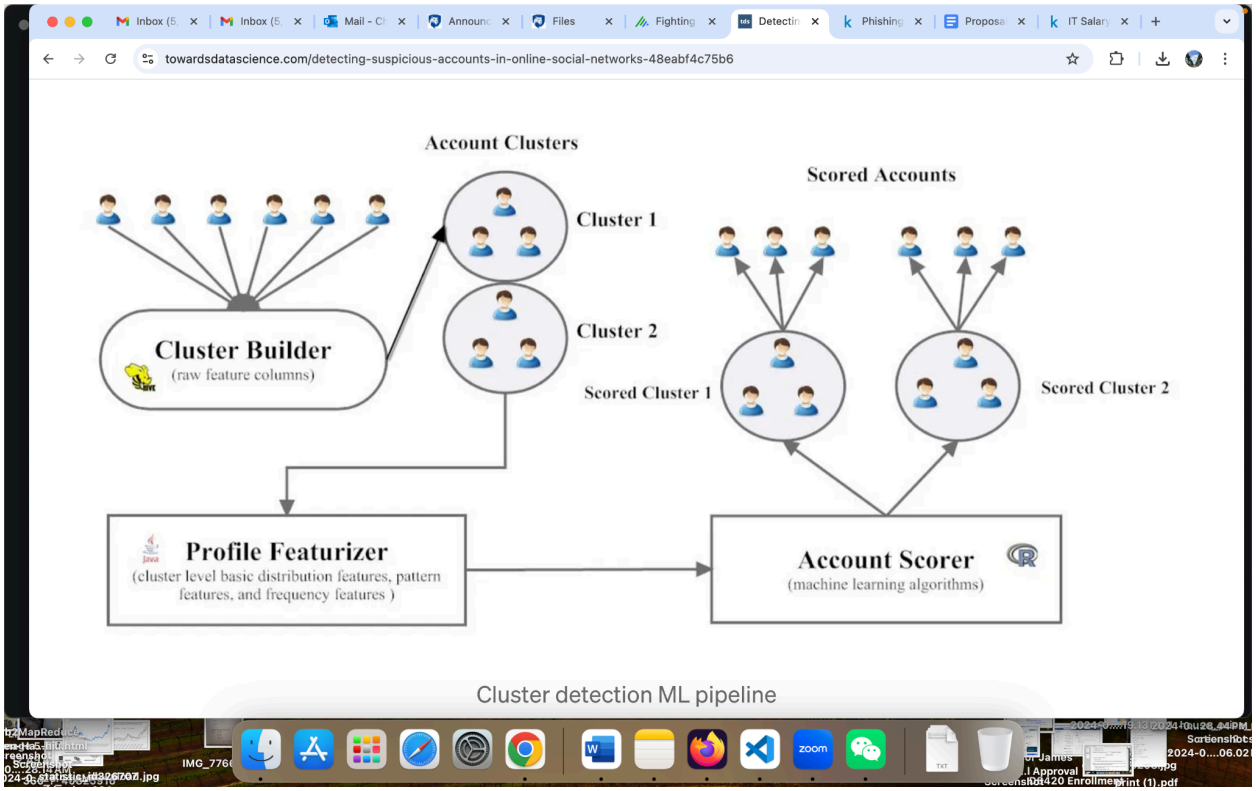
Problem type: novelty detection

Introduction:
Direct, undirected, DFS, BFS, multigraph, simple graph, signed graph, attributed graph, long-tail distribution, heterogeneous information network, adjacency matrix, positive (false positive)

What we need to research:

- Scholarly article regarding GNN
- Different types of graphs.
- Data scrape off Twitter.

Possible final result:



Possible topic: Detecting phishing websites.
Datatset: https://www.kaggle.com/datasets/shashwatwork/phishing-dataset-for-machine-learning

Scrape twitter data, different types of tweet and tag, response by other users. Classification - classified different groups based on their tweet. (This is more of a clustering problem).