בס"ד

Software Engineering Department

Braude College

Capstone Project Phase B – 61999

Citation networks evolution using Dynamic Network Embeddings

User Guide

25-1-R-16

Students:

 Avishay Bar 316165984

Shay Bistrizky 206622086

Supervisors:

Prof. Zeev (Vladimir) Volkovich

Dr. Renata Avros

GIT Repository Link

1. System Purpose

The system analyzes the dynamic development of academic papers using learned representations (embeddings) of citation networks over time.
The goal is to classify papers based on their behavior across years:

- **Rising Stars: Consistently increasing influence**

- **Emerging: Recently published papers that suddenly gained attention**

- **Falling Stars: Previously influential papers that are losing relevance**

- **Steady: Foundational works with stable long-term influence**

2. Prerequisites

- **No local installation required**

- **The only thing users need to do is run the Google Colab notebook provided.**

- **Access to .jsonl files with the appropriate format (usually organized by year, e.g., papers_2000.jsonl, papers_2001.jsonl, etc.)**

- **GPU (Google Colab provides an A100 GPU by default – recommended)**

3. Running the System

1. **Open the notebook (e.g., Dynamic_Citation_Network_Analysis.ipynb) in Google Colab**

2. **Upload the data directory or mount your Google Drive (if the data is stored there)**

3. **Ensure that the following fields exist in the input files:**

   o **id – unique identifier for each paper**

   o **year – year of publication**

   o **references – list of paper IDs this paper cites**

4. **Configure the parameters at the beginning of the notebook (e.g., year range, data path, etc.)**

5. **Run all cells (Runtime > Run all)**

4. System Output

After full execution, the system will generate:

- **Embedding files by year (M_<year>.pt)**

- **Shared transformation matrix (alpha.pt)**

- **Yearly graph snapshots (G_<year>.pkl)**

- **Paper classification output (CSV)**

5. Common Issues

Issue	Solution
File loading error	Make sure the data_dir is correctly defined and contains .jsonl files
Memory error	Make sure a GPU is active (Colab menu: Runtime > Change runtime type > GPU)
Execution stopped mid-run	Use the --resume parameter to continue from the last completed year

6. Creating Graphs and Analyzing Results (Optional)

- **You can run dedicated cells at the end of the notebook to generate visualizations that demonstrate how papers change over time.**

- **All papers will automatically be labeled based on their influence trajectory across years.**