Data Mining

HW 4

Ashkan Ansarifard

1970082

Table of Contents

List of Figures

List of Tables

1 Introduction

1.1 Project Structure

In this section, we show how the project is structured. In the following sections, we reflect to this structure and explain what each module does.

For now, you can find the overall project structure as following:

```
project_root/
+-- main.py
|
+-- data/
|   +-- dataset_loader.py
|   +-- prepare_link_prediction_data.py
|
+-- embeddings/
|   +-- analyze_embeddings.py
|   +-- visualize_embeddings.py
|
+-- evaluation/
|   +-- evaluator.py
|   +-- cross_dataset_evaluator.py
|   +-- link_prediction_evaluator.py
|
+-- models/
|   +-- gnn_model.py
|   +-- generalized_gnn.py
|   +-- gnn_explainer.py
|   +-- node2vec_model.py
|   +-- link_prediction_gnn.py
|
+-- saved_models/
|   +-- generalized_<timestamp>/
|       +-- model.pth
|       +-- metadata.json
|
+-- training/
|   +-- trainer.py
|   +-- trainer_link_prediction.py
|
+-- utils/
|   +-- config.py
|   +-- dimensionality_handler.py
|   +-- device.py
|   +-- find_best_models.py
|   +-- load_generalized_gnn.py
|   +-- model_saver.py
|
+-- README.md
|
+-- requirements.txt
+-- setup.bat
+-- setup.sh
```

1.2 Environment Setup

To execute this project, follow these steps:

1. Ensure that Python 11.3 is installed on your system. You can download it from the official Python website.

2. Run the provided setup script to create and configure the virtual environment:

   - On Windows, run:

     ```
     setup.bat
     ```

   - On macOS/Linux, run:

     ```
     bash setup.sh
     ```

3. After the setup script completes, activate the virtual environment:

   - On Windows:

     ```
     <venv_name>\Scripts\activate
     ```

   - On macOS/Linux:

     ```
     source <venv_name>/bin/activate
     ```

4. Run the project using:

   ```
   python main.py
   ```

After completing these steps, your environment will be properly configured for executing the project.