Knowledge Graphs with Large Language Models

Assignment 3: Applying Knowledge Graphs

Introduction

In this assignment you are asked to evaluate and improve the llm-based QA system that we saw in class and which answers questions by transforming them into Cypher queries and executing them against a knowledge graph.

Task 1 (50%)

Create a small evaluation set of 20-30 natural language questions that should be answerable by the Neo4j movie graph, feed them to the Text2Cypher pipeline, and evaluate their generated Cypher queries. Try to make your dataset as diverse as possible, covering simple and more complex questions. Also, identify and describe the most usual mistakes the LLM does in translating natural language questions to Cypher.

Task 2 (50%)

Using the analysis in task 1, try to improve the system by improving the Cypher generation prompt (zero-shot) and using question-query examples (few-shot).

Task 3 (50%) - Optional for extra credit

Try to further improve the system by implementing a pattern called dynamic few-shot prompting.

In this pattern, the system does not use static question-query examples but retrieves semantically similar questions and their corresponding Cypher queries from a vector index and uses them as context in the Cypher generation prompt. This retrieval strategy helps the chatbot generate more accurate queries and keeps the prompt small by only including examples that are relevant to the current input query.

Deliverables

- An implementation of the task 1, task 2 and task 3 systems, either in the form of a Colab notebook or any other way you prefer, with clear instructions on how to run and evaluate them.
- A report describing the development and evaluation process and results for each task.