

TY B.Tech. (CSE) – II [2021-22]
5CS372 : Advanced Database System Lab.

Assignment No. 11

Neo4j Graph Database

Install and configure the Neo4j Graph database as discussed in class.

Consider the “**Research Papers Database**” scenario as follows :

The research papers have authors (often more than one). Most papers have a classification (what the paper is about). The classifications form a hierarchy in several levels (for example, the classification “Databases” has the sub-classifications “Relational” and “Object-Oriented”). A paper usually has a list of references, which are other papers. These are called citations.

1. Design/model the graph database using Neo4j for above scenario.
2. Download the raw data from **Cora Research Paper Classification Project** : <http://people.cs.umass.edu/~mccallum/data.html> The database contains approximately 25,000 authors, 37,000 papers and 220,000 relationships.
3. Load this data using Neo4j Data Browser
4. Design the python based desktop application for any kind of search on above database. The application should be able to answer queries like
 - a) Does paper A cite paper B? If not directly, does paper A cite a paper which in its turn cites paper B? And so on, in several levels.
 - b) Show the full classification of a paper (for example, Databases / Relational)

Note : Follow the submission guidelines.

Deadline : 01/05/2022

Dr. B. F. Momin
Course Coordinator