Graph Functionalities for Data Platform

A SUMMER INTERNSHIP PROJECT REPORT

Submitted by

Tanish Khandelwal (201B283)

Under the Guidance: Chirayu Khandelwal (Lead Software Engineer)

Submitted in partial fulfilment for the award of the degree

of

BACHELOR OF TECHNOLOGY

IN

COMPUTER SCIENCE AND ENGINEERING

At

August & 2024

DECLARATION

I hereby declare that the project entitled "**Graph Functionalities for Data Platform**" was submitted for the B. Tech. (CSE) degree is my original work and the project has not formed the basis for the award of any other degree, diploma, fellowship or any other similar titles.

Signature of the Student

Place:

Date:

CERTIFICATE

Cloud Energy
Software Strategy & Services

Private & Confidential

Date: September 11, 2023

Name: Tanish Khandelwal

TO WHOMSOEVER IT MAY CONCERN

This is to certify that Tanish Khandelwal has successfully completed his internship project as Data Engineering Intern at CES.

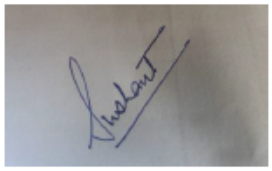
He has worked with us from 24th July 2023 to 11th September 2023. Tanish was associated with the project titled Graph Functionalities for Data Platform Kendra-Sustain, in which he contributed to Software Engineering and Product Development including supporting product manager, data science team and CEO.

During the internship program Tanish's work was satisfactory.

We wish him all the best in his future endeavors.

Yours faithfully,

For Cloud Energy Software Private Limited.



Manager Human Resource

Read and Accepted:

_____

Signature of Employee

Cloud Energy Software Private Limited

Mumbai – India
WeWork, 5th floor, Spectrum Tower, Mindspace, Chincholi Bunder Road, Malad (West), Mumbai 400064

Manchester – UK
Enterprise House, Manchester Science Park, Lloyd St N, Manchester M15 6SE | Tel: +44(0) 161 209 9090

CIN: U72300MH2014FTC254224

www.cloudenergysoftware.com

ACKNOWLEDGEMENT

I express my deep & sincere gratitude to Mr. Chirayu Khandelwal (Lead Software Developer at CES), Mr. Sushant Motghare (Manager Human Resource at CES) for their expert suggestion & encouragement which helped me to tide over the hard concepts encountered during the project.

My special thanks to Mr. Amit Kumar, for providing me an environment to learn deep learning concepts over SDKs so that I can perform medical Imaging and training exercises for completing my assignments. I like to express my sincere gratitude to them for providing me with the most valuable guidance. I would also like to extend our thanks to Jaypee University of Engineering and Technology, Guna for ensuring a good environment and providing the necessary experience and assistance for us to work on this project.

I am grateful for his guidance and support throughout the project. He was always available to answer my questions and provide me with helpful feedback. He also helped me to stay on track and to meet my deadlines.

Tanish Khandelwal

EXECUTIVE SUMMARY

In the era of big data, organizations face an ongoing challenge to extricate valuable insights from intricate, interconnected datasets. Graph databases have emerged as a potent tool for modelling and querying data relationships, offering distinct advantages over relational databases. This project investigates the integration of graph functionalities into a data platform, with a particular concentration on Neo4j, the industry-leading graph database management system. The primary objective of this project is to utilize Neo4j to improve the data platform's capacity to efficiently store, manage, and analyse interconnected data. Through the integration of Neo4j, the platform acquires the ability to represent data as nodes and relationships, enabling the modelling of complex relationships in a variety of domains, including social networks, recommendation systems, and fraud detection.

Data ingestion and synchronization with Neo4j, query optimization for graph-based queries, and the development of user-friendly interfaces for querying and visualizing graph data are among the key functionalities implemented. Additionally, this project investigates the integration of machine learning algorithms with Neo4j to unleash the potential of predictive analytics on graph data. The advantages of this integration are numerous. The data platform enables its users to uncover concealed patterns, execute complex traversals, and extract actionable insights from their data. In addition, this initiative contributes to the field of data management by demonstrating the practical benefits of integrating graph databases into existing data platforms.

List Of Figures