Graph Analytics for Python Developers

Assignment 5

Version 1.0

Assignment 5

The goal of this assignment is to understand the graph embedding techniques we discussed during the fifth lecture. The task will be performing either **node classification** or **link prediction** on a dataset of your choosing.

The requirements are:

- Choose an **appropriate dataset consisting of at least 100 nodes and 1000 edges**. It should represent a real-world network so the machine learning tasks can have a higher chance of generating meaningful classifications or predictions. We suggest you use a SNAP large network.

- Perform **node classification** or **link prediction** on the chosen dataset utilizing node embeddings generated with the **node2vec** algorithm.

Resources

StellarGraph node2vec	StellarGraph Reference Guide Link
Stanford Large Network Dataset Collection	SNAP Datasets

How to submit the assignment?

Upload your code to a GitHub repository and send us the link to ivan.despot@memgraph.com.

Because all assignments will be submitted using GitHub, we recommend creating one repository where you will upload all of the assignments.