

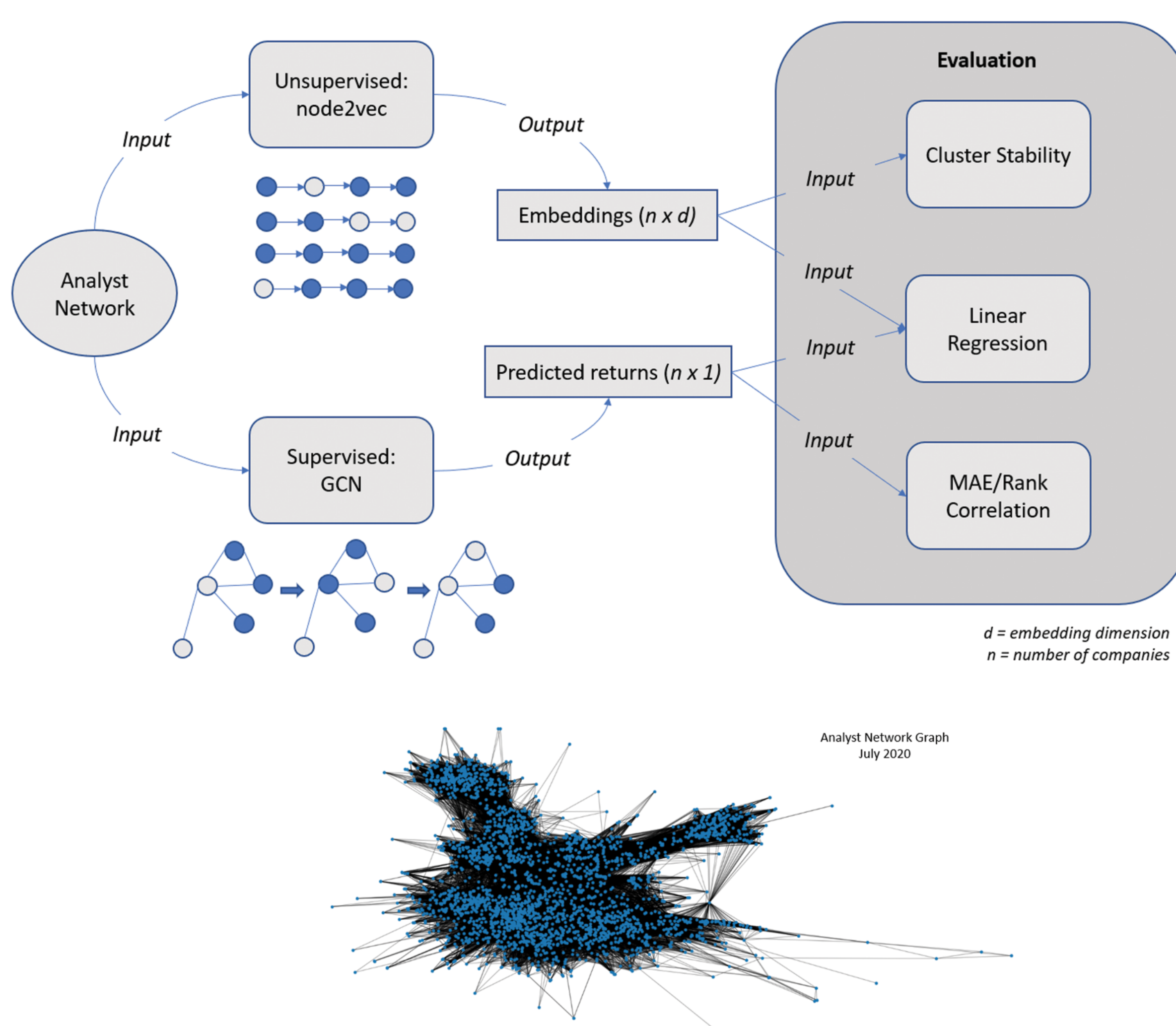
Exploring Graph Neural Network Methods on the Analyst Network Graph

Both Node2Vec and Graph Convolutional Networks can be used to exploit the topology of the Analyst Network to explain companies returns

Alisa Yang

Christina C. Christara
ACADEMIC SUPERVISOR

Yixing Zhang
INDUSTRY SUPERVISOR



PROJECT SUMMARY

Graph datasets can be a rich source of information in fields like Finance since they can capture relationships that might not be present in traditional tabular datasets. The Analyst Network (AN) is a graph dataset where nodes are companies, and edges represent the number of equity analysts working on both companies. We expect analysts to have domain expertise and work on companies that are related in some dimension. Being able to group companies is an important task in Finance since similar companies tend to have correlated returns. Graph Neural Networks (GNNs) are popular methods to analyze and make inferences on graphs since they can exploit the topology of graphs. We used both supervised (Graph Convolutional Network [1]) and unsupervised (node2vec [2]) techniques to extract information from the AN dataset. We find that the unsupervised technique encodes similar information to the Global Industry Classification Standard (GICS), is simpler to train and can create useful embeddings for a variety of downstream forecasting tasks. On the other hand, the supervised technique is target specific and can generate predictions directly but is harder to train.

REFERENCES

- [1] Thomas N Kipf and Max Welling. Semi-supervised classification with graph convolutional networks. In Proceedings of the 5th International Conference on Learning Representations, 2017
- [2] Aditya Grover and Jure Leskovec. node2vec: Scalable feature learning for networks. In Proceedings of the 22nd ACM SIGKDD international conference on Knowledge discovery and data mining, pages 855–864, 2016