GATTO: Can Topological Information Improve Node Classification via GAT?

Midterm Report for Learning from Network's project

Francesco Biscaccia Carrara *(2120934)*, Riccardo Modolo *(2123750)*,
Alessandro Viespoli *(2120824)*

Master Degree in Computer Engineering
University of Padova

I. INTRODUCTION

The scope of this document is to:

- Clarify all points not properly explained in the proposal paper
- Summarize all work done
- Define the missing work
- Estimate time and effort to finish the project.

TODO: inserire riassuntino di quello che scriviamo dopo

II. CLARIFICATION POINTS

One of the most important clarification point is about data: we didn't provide an extensive explanation, in the proposal paper, about how data are built and how we want to use them. The graphs we choose in SNAP[1] are undirected. Each node contain community as parameter.

III. WORK DONE

We've already build the **precomputation module** (explained in the project proposal) and the code to automate the testing phase inside cluster. We've decided the hyperparameters about node2vec.

IV. WORK IN PROGRESS

We nedd to build the **GAT module** (Viespoli) and do the tests on **CAPRI**[2]. Also writing the paper (for obvious reason) is a missing step.

V. ESTIMATION

Since we need our collegue (Viespoli) to build the GAT, we estimate to finish the project on 7th January. With this estimation we have 8 days for eventual issue and refinements.

REFERENCES

[1] Jure Leskovec and Andrej Krevl. *SNAP Datasets: Stanford Large Network Dataset Collection*. `http://snap.stanford.edu/data`. June 2014.

[2] DEI University of Padova. *CAPRI: Calcolo ad Alte Prestazioni per la Ricerca e l'Innovazione*. 2017. URL: `https://capri.dei.unipd.it`.