Degree Project in Computer Science and Engineering

Second cycle, 30 credits

Lightweight Memory Networks for Link Prediction

Generalization of the LiMNet Architecture

TITOUAN MAZIER

Lightweight Memory Networks for Link Prediction

Generalization of the LiMNet Architecture

TITOUAN MAZIER

Abstract

Lorem ipsum dolor.

Keywords

Sammanfattning

Lorem ipsum dolor.

Nyckelord

Graph Representation Learning, Temporal Interaction Networks, Link Prediction, Data Mining, Machine Learning, Recommendations

Acknowledgments

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do.

Lorem.

Lorem ipsum dolor.

Stockholm, April 2025
Titouan Mazier

Contents

Acronyms and Abbreviations

Chapter 1
Introduction

Lorem ipsum dolor sit amet.

1.1. Problem

Lorem ipsum.

1.2. Purpose

epoch Lorem ipsum dolor sit amet.

1.3. Goals

Lorem.

1.4. Structure of the Thesis

Lorem ipsum dolor sit amet.

Chapter 2
Background

2.1. Link Prediction

- There are too many information, thus we need smarter way to select it.

- A popular solution nowadays is through recommendations: each user is being recommended personalized suggestions that are tailored to fit their needs, based on their past interactions.

- The problem of recommendation is often formulated as for each user, create a ranking of the items. Past interactions are represented as edges in a graph were users and items are nodes.

- Collaborative filtering is the process of exploiting not only the user's past interactions but also the interactions of similar users. Example, movies.

- A classical way to tackle the problem is to use graph-based metrics to measure how much related each item is to the user's preferences. Examples of this are distance, Adamic-Adar distance, etc.

- These solutions are quite limited because they rely solely on the structure of the interactions, disregarding any additional knowledge on the users, items or interactions. This lead to the use of machine learning approaches.

- One can use the graph-based metrics as features for a traditional ML pipeline.

- ML opened a new opportunity: computing scores for all pairs is expensive, ML can be used to produced embeddings of users and items separately and use their distances as scores instead. Allow for pre-computing embeddings of items and much faster recommendations. (and better scalability).

2.2. Graph Representation Learning

- The task of learning embeddings from graphs is called GRL and have applications outside link-prediction: node-level, edge-level, graph-level (molecule graphs, ask gpt).

- Two approaches have been used so far: random walk based and GNNs. Random walk are used to create node sequences that are then fed to sequence based models of the same kind as Language models. GNNs are NN using the structure of the graphs to pass information through nodes.

(- Practical examples of GNNs in for link prediction: GraphSAGE)

- Recently, researchers have been trying to exploit even more information with the addition of dynamic graphs. Until there, the temporality of the interactions did not have impact on the recommendations. e.g. listening session on spotify.
- There are 2 way of representing dynamic graphs: DTDG and CTDG. We are interested in the latter since it better represent the reality.
- CTDG can be seen as a stream of interactions instead of a stable graph structure. Thus, graphs concepts such as neighborhood become blurred.
- Approaches to cope with this added dificulty include using a window of past interactions to feed into a ML system (e.g. cross-attention with DeePRed).
- A common solution since deepcoevolve is to use cross-RNN. The system keeps a memory of each node and will have these memories "interact" to update each other whenever an interaction is observed.
- This approach has the benefit of conserving the causality of interactions into the embeddings.

2.3. LiMNet

- LiMNet is such a solution that aims at being as lightweight and simple as possible, using only one RNN cell to compute the embeddings. This has the double benefit of making it very cheap to run but also very flexible with node insertion and deletion being trivial operations.
- Description of LiMNet Architecture
- LiMNet has proven effective on the task of botnet and fraud detection but was not initially designed to tackle link prediction. Which is what this work aims to do.

2.3.1. Test figure

Chapter 3
Method or Methods

Lorem ipsum dolor sit amet.

3.1. Research Process

Lorem ipsum dolor.

3.2. Data Collection

Lorem.

3.2.1. Sampling

Lorem ipsum.

3.2.2. Sample Size

Lorem ipsum.

3.2.3. Target Population

Lorem ipsum dolor.

3.3. Evaluation Framework

Lorem ipsum.

References

8 | References

Appendix A
Usage Instructions

Hmm...

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do.

Appendix B
Something Else

Whoa!

www.kth.se