Team Name: Graph Heroes

Project Title: Movie Recommendations

Project summary (4-5+ sentences).

Fill in your problem and background/motivation (why do you want to solve it? Why is it interesting?). This should provide some detail (don't just say "I'll be working on object detection")

We want to learn about this novel way to approach recommendation systems. Currently the two main approaches to building recommendation systems are collaborative based and content based.

Collaborative based methods are beneficial in that they don't require much data on the things being recommended at the beginning and can improve significantly as the user interacts more with the system. The caveat is that these methods suffer from the cold start problem where initially there is no information on the user due to no logged interactions. Additionally, collaborative methods only leverage the interaction data, and can't use the rich user/item metadata.

Content based methods suffer less from the cold start problem than collaborative approaches. New things can be described by their characteristics and relevant suggestions can be generated for these new entities.

We hope that our graph neural network approach can have benefits from both approaches.

What you will do (Approach, 4-5+ sentences)

Be specific about what you will implement and what existing code you will use. Describe what you actually plan to implement or the experiments you might try, etc. Again, provide sufficient information describing exactly what you'll do. One of the key things to note is that just downloading code and running it on a dataset is not sufficient for a description or a project! Some thorough implementation, analysis, theory, etc. has to be done for the project.

We plan on leveraging many of the SOTA layers in our graph neural network as well as customize some specialized layer (pending research/understanding).

Specifically, beside using Pytorch, we will implement Graph neural network from scratch. Things that we want to explore in our project:

- Whether Graph neural networks can solve our given problem of link prediction.

- Whether Graph CNN is faster and more efficient than sklearn and filtering baselines
- Whether we can discover a graph structure in the data, with Graph neural network
- Whether we will achieve better explainability with this approach, compared to Deep CNN.
  - We may use Captum and integrated gradient for this.

Resources / Related Work & Papers (4-5+ sentences)

What is the state of art for this problem? Note that it is perfectly fine for this project to implement approaches that already exist. This part should show you've done some research about what approaches exist.

SOTA for this problem involves leveraging graph neural networks with GCN, SageConv, etc. layers in order to classify nodes and predict links. In our case, we will be focusing on link prediction in order to build a recommendation system. We will be leveraging PyG 2.0 for our research and implementations.

Here are the resources we have collected during the research of this proposal:

- *LightGCN: Simplifying and Powering Graph Convolution Network for Recommendation*. He 2020. https://arxiv.org/abs/2002.02126
- *Deep Learning on Graphs: A Survey.* Zhang 2018. https://arxiv.org/abs/1812.04202
- *Representation Learning on Graphs: Methods and Applications*. Hamilton 2017. https://arxiv.org/abs/1709.05584
- *The structure and dynamics of multilayer networks*. Boccaletti 2014. https://doi.org/10.1016/j.physrep.2014.07.001
- *Graph neural networks: A review of methods and applications*. Zhou 2020. http://arxiv.org/abs/1812.08434
- *A Gentle Introduction to Graph Neural Networks*. Lengeling 2021. 10.23915/distill.00033

We may add more resources as we explore more on the project.

Datasets

Deep learning is data-driven, so what datasets you use is crucial. One of the key things is to make sure you don't try to create and especially annotate your own data! Otherwise the project will be taken over by this.

We will be using the MovieLens dataset, provided at:

https://grouplens.org/datasets/movielens/1m/

List your Group members

- Quang Vu
- Ruiming Lu
- Michael Chen