MLNS
PROJECT PROPOSAL

Antonin Roche, Maximilien Huet & Louis Funck-Brentano

1 MOTIVATION AND PROBLEM DEFINITION

Music is present in our every day life and recently platforms have provided users (us) with an overwhelming amount of artists and songs to choose from [1]. It then becomes important to have a strong recommendation system for users to remain engaged with a specific platform. Graph methods have a lot of potential for those applications as the interactions between songs, users and artists are naturally represented as a graph. Most methods used in the industry, like collaborative filtering and content-based filtering [2], do not leverage the full information contained in the graph. Moreover, the classical methods, due to their inherent design, have limitations that prevents solving some problems posed in the context of music recommendation such as cold start, scalability, recommendation novelty and others [3].

The problem is as follows: Given an open music dataset (as presented in [1]), construct a knowledge graph and make recommendation (probability of connection between an user and an artist or a song), similar to the formulation in [4]. Our aim in this project is to re-implement different methods to solve this problem (from simple to more complex methods) and then compare them based on properties we would want of the recommendation system. For the sake of simplicity, we will focus on 1 property of interest specific to music recommendation (like novelty, recommending sequences, cold-start) and focus the comparaison on that aspect, where surveys like in [3] can provide reference papers.

This task of recommendation is a common one in graph related problems (such as social networks for friend recommendation) but in the context of music recommendation there are a few subtleties that make this problem very challenging (like the cold-start problem and need for novelty). It could then naturally be extended to any other recommendation system where some of the properties highlighted above are required.

2 METHODOLOGY

To accomplish the task we have described above, we will need to go through a few steps that we will highlight below.
First, we have the task of collecting data (even if we have public datasets, we will still need to build a representative graph from it). Then we will implement known algorithms using different techniques and compare them. Our baseline will be the Collaborative filtering approach as it is one of the most used method [3]. We will implement 2 types of graph-based methods, namely methods based on knowledge graphs and classical methods (like graph embedding in [1] or [2]) and methods based on Deep Learning methods (most used architecture seems to be GAT-like in [4] and [5].

3 EVALUATION

For our task we have a choice over 2 main open datasets described in [1]: the MusicBrainz dataset which has information about music items (like artists, songs and music labels) and the Last.fm dataset which also contains information about users patterns.
To quantify our work, many metrics ca be used, with some of the most popular ones being metrics based on the top-K recommendation ([5] like recall and Normalized Discounted Cumulative Gain (ndcg). There are also metrics specific to certain tasks like novelty as described in [6] and [7] based on self-information of an object.

References

[1] N. Bertram, J. Dunkel, and R. Hermoso, "I am all ears: Using open data and knowledge graph embeddings for music recommendations," *Expert Systems with Applications*, vol. 229, p. 120 347, 2023, ISSN: 0957-4174. DOI: `https://doi.org/10.1016/j.eswa.2023.120347`. [Online]. Available: `https://www.sciencedirect.com/science/article/pii/S0957417423008497`.

[2] X. Liu, Z. Yang, and J. Cheng, "Music recommendation algorithms based on knowledge graph and multi-task feature learning," *Scientific Reports*, vol. 14, no. 1, p. 2055, 2024. [Online]. Available: `https://doi.org/10.1038/s41598-024-52463-z`.

[3] Y. Deldjoo, M. Schedl, and P. Knees, "Content-driven music recommendation: Evolution, state of the art, and challenges," *Computer Science Review*, vol. 51, p. 100 618, 2024, ISSN: 1574-0137. DOI: `https://doi.org/10.1016/j.cosrev.2024.100618`. [Online]. Available: `https://www.sciencedirect.com/science/article/pii/S1574013724000029`.

[4] Z. Tao, Y. Wei, X. Wang, X. He, X. Huang, and T.-S. Chua, "Mgat: Multimodal graph attention network for recommendation," *Information Processing & Management*, vol. 57, no. 5, p. 102 277, 2020, ISSN: 0306-4573. DOI: `https://doi.org/10.1016/j.ipm.2020.102277`. [Online]. Available: `https://www.sciencedirect.com/science/article/pii/S0306457320300182`.

[5] X. Wang, X. He, Y. Cao, M. Liu, and T.-S. Chua, "Kgat: Knowledge graph attention network for recommendation," in *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, ser. KDD '19, ACM, Jul. 2019, 950–958. DOI: `10.1145/3292500.3330989`. [Online]. Available: `http://dx.doi.org/10.1145/3292500.3330989`.

[6] M. Zhao, L. Wu, Y. Liang, *et al.*, "Investigating accuracy-novelty performance for graph-based collaborative filtering," in *Proceedings of the 45th International ACM SIGIR Conference on Research and Development in Information Retrieval*, ser. SIGIR '22, ACM, Jul. 2022, 50–59. DOI: `10.1145/3477495.3532005`. [Online]. Available: `http://dx.doi.org/10.1145/3477495.3532005`.

[7] T. Zhou, Z. Kuscsik, J.-G. Liu, M. Medo, J. R. Wakeling, and Y.-C. Zhang, "Solving the apparent diversity-accuracy dilemma of recommender systems," *Proceedings of the National Academy of Sciences*, vol. 107, no. 10, 4511–4515, Feb. 2010, ISSN: 1091-6490. DOI: `10.1073/pnas.1000488107`. [Online]. Available: `http://dx.doi.org/10.1073/pnas.1000488107`.