Research Plan for CSE3000 Research Project

Improving Music Recommendation Systems For Youngsters

Borislav Semerdzhiev

April 28, 2024

Background of the research

Music plays a crucial role in people's lives and can significantly impact a child's childhood development. Most of the music nowadays is streamed through online platforms which use recommendation systems to aid in finding new tracks to listen to, however, it is important to note that those systems are designed to cater mostly to the adults and so is most of the literature written about music recommendation. Researches have been previously concluded which try to explore what are the factors that affect children's taste in music[4] [3] and how they differ from the adults'. The conclusion so far is that there is no single best strategy that would work on all children and that tailoring a collaborative filtering system to users <= 18 years is indeed beneficial.

Research Question

The main question driving this project is "Which user/item characteristics and/or preferences inferred from youngsters' listening behavior on music platforms can be used to model a user profile which could be then used to make new recommendations". In order to accomplish that, we should be able to extract features from the songs which could potentially influence whether or not a user likes a given song. I think it is reasonable to believe that this work could be completed within the duration of the research project, as there are numerous researches I could draw inspiration from regarding recommendation system.

The following sub-questions can be derived:

- Determine which features of a song influence a listener's opinion about a song.
- Determine how to represent these features as a continuous/discrete variable.
- Approximate the user given rating to a song by knowing how often this song has been listened by them.
- Build a profile for a user using those approximated ratings and song features.
- Predict a rating to a song that the user hasn't heard using the constructed profile.

Method

In order to determine the features that impact the likelihood of a user liking a song, I intend to read over previous researches that have been concluded on this topic, and then determine which of those features I want to include in my research [1].

To represent a feature of a song as a continous/discrete variable I intend to use popular online music streaming API such as Spotify's, which provides information about some features of interests(energy, dancability, etc...) and potentially scale or normalize them. Furthermore I might use additional datasets/API that provide information about more features if I determine that the one Spotify provides are not sufficient.

Since the current data set I have does not provide us with a rating the user has given to a song but only contains the amount of times he has listened to the track I need to find a method to approximate a rating he would give the song. I need to take into account that a lot of the listening events could be generated by accident(the user had accidentally clicked the song) or a song might always start after another one and the user skips it instantly. Approximating ratings given implicit feedback in recommendation systems is a researched topic which I plan on delving deeper [2].

To compute ratings to new songs I intend to use either collaborative filtering or content-based filtering depending on the listening habits of the youngsters about which I need to read more before making the decision.

Finally, to test how accurate the model is I intend on splitting the data at the beginning into train data and test data and in the end comparing the actual ratings to the predicted ratings the model has generated.

Planning of the research project

Date	Deadline	duration
week 1	define research question/setup and start reading	
23 April	Create a plan for the first week	1 day
24 April	Meeting peer group & supervisor	1.25h
28 April	Research plan (create slides for the presentation)	2 days
week 2	Define experiment setup	10h
week 2	Read literature	30h
week 2	Determine features that influence user opinion about a song	(Reading literature time)
week 2	Responsible research lecture	2h
1 May	Meeting peer group & supervisor	1.25h
week 3	Write intro, finish experiment setup	9h
week 3	Read literature	30h
week 3	Collect song features and represent them as variables	10h
week 3	Responsible research lecture	2h
8 May	Meeting peer group & supervisor	1.25h
8 May	ACS assignment 1	2h?
week 4	Experiment iterations & writing (intro, experiment, ...)	40h
week 4	Approximate user ratings by implicit feedback	15h
week 4	Find date for final presentation	0.5h
15 May	Meeting peer group & supervisor	1.25h
15 May	ACS assignment 2a & 2b	2h?
week 5	Experiment iterations & writing (intro, experiment, ...)	40h
week 5	Midterm presentation	2 days preparation
week 5	Generate user ratings for new songs	15h
22 May	Meeting peer group & supervisor	1.25h
24 May	ACS assignment 3	3h?
week 6	Experiment iterations & writing (intro, experiment, ...)	40h
week 6	Test model on test data	5h
29 May	Meeting peer group & supervisor	1.25h
week 7	Stop experiment & writing (results, discussion, ...)	
3 June	Paper v1 draft	throughout the research
5 June	Paper v1 peer review	3 hours
5 June	Meeting peer group & supervisor	1.25h
week 8	start poster, writing (discussion, conclusion, ...)	40h
11 June	Paper v2 draft	throughout the research
12 June	Meeting peer group & supervisor	1.25h
week 9	Finish poster, writing (discussion, conclusion, ...)	40h
19 June	Meeting peer group & supervisor	1.25h
23 June	Submit final paper	throughout research
24 June	Submit poster	past 2 weeks (few hours total)
26 June	Meeting peer group & supervisor	1.25h
week 10	Poster presentation	1h presentation, 2 day preparation
week 10	Final presentation	1h presentation, week preparation

References

[1] Yashar Deldjoo et al. "Recommender Systems Leveraging Multimedia Content". In: *ACM Comput. Surv.* 53.5 (Sept. 2020). ISSN: 0360-0300. DOI: 10.1145/3407190. URL: https://doi.org/10.1145/3407190.

[2] Gawesh Jawaheer, Martin Szomszor, and Patty Kostkova. "Comparison of implicit and explicit feedback from an online music recommendation service". In: *Proceedings of the 1st International Workshop on Information Heterogeneity and Fusion in Recommender Systems*. HetRec '10. Barcelona, Spain: Association for Computing Machinery, 2010, 47â51. ISBN:

9781450304078. DOI: 10.1145/1869446.1869453. URL: https://doi.org/10.1145/1869446.1869453.

[3] Markus Schedl and Christine Bauer. *Online Music Listening Culture of Kids and Adolescents: Listening Analysis and Music Recommendation Tailored to the Young*. 2019. arXiv: 1912.11564 [cs.IR].

[4] Lawrence Spear et al. "Baby Shark to Barracuda: Analyzing Childrenâs Music Listening Behavior". In: *Proceedings of the 15th ACM Conference on Recommender Systems*. RecSys '21. Amsterdam, Netherlands: Association for Computing Machinery, 2021, 639â644. ISBN: 9781450384582. DOI: 10.1145/3460231.3478856. URL: https://doi.org/10.1145/3460231.3478856.