Personalized Recommendations for Music Genre Exploration

Yu Liang, Martijn Willemsen

HTI Group, Eindhoven University of Technology
Jheronimus Academy of Data Science

Jheronimus Academy of Data Science

EINDHOVEN UNIVERSITY OF TECHNOLOGY

Liang, Yu, and Martijn C. Willemsen. "Personalized Recommendations for Music Genre Exploration." *Proceedings of the 27th ACM Conference on User Modeling, Adaptation and Personalization*. ACM, 2019.

Introduction

- Traditional recommender system
  - Predicting users’ current preference

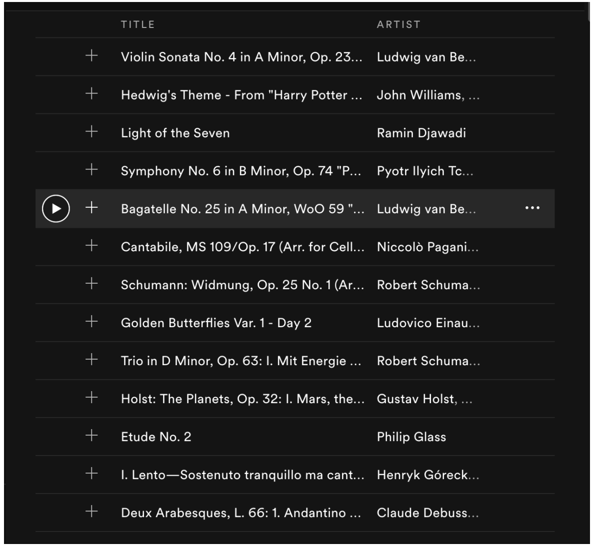


Fig. Martijn’s recommendations on Spotify

Fig. Yu’s recommendations on Spotify

[2]

Introduction

- Traditional recommender system
  - Predicting users' current preference

I want to explore some pop music!

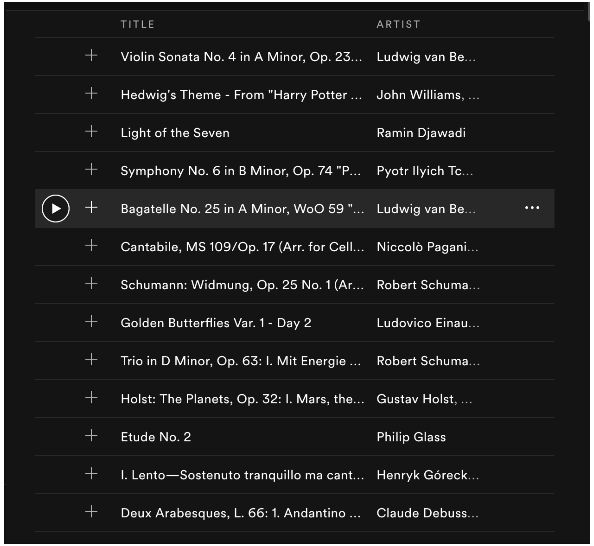


Fig. Martijn's recommendations on Spotify

Fig. Yu's recommendations on Spotify

[3]

Introduction

- Traditional recommender system
    - Predicting users' current preference

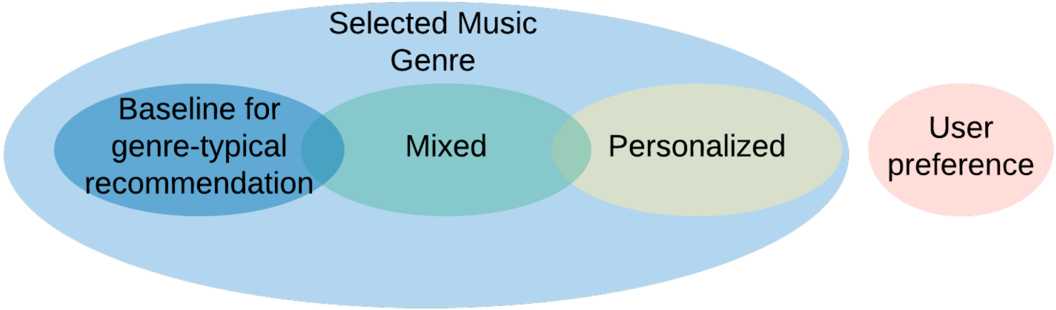
- How recommender system could help users with direct explorations from new music genres/tastes?

I want to start my exploration directly from the new genre "pop"!

[4]

Recommendation methods

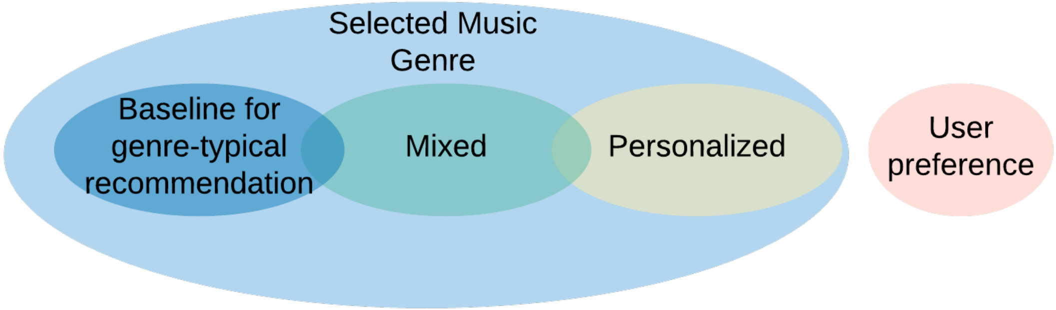
1. Recommend genre-typical tracks (representative)
   - The non-personalized method
2. Take into account users' current preferences (accurate and personalized)
   - The personalized method
3. Balance accuracy and representativeness
   - The mixed method



[5]

Research question

- Can we give more helpful recommendations than the genre-typical tracks from the non-personalized baseline?
  - Personalized method (accurate and personalized recommendations)
  - Mixed method (trade-off between accuracy and representativeness)



[6]

Content-based recommendation on audio features

- The recommendation is done in a content-based way by matching in terms of high-level audio features.
- Users' current preferences and genre space are represented by semantic audio features (acousticness, energy, valence, speechiness, liveness and danceability) retrieved from Spotify.

I am a fan of classical music and I prefer songs with low valence!

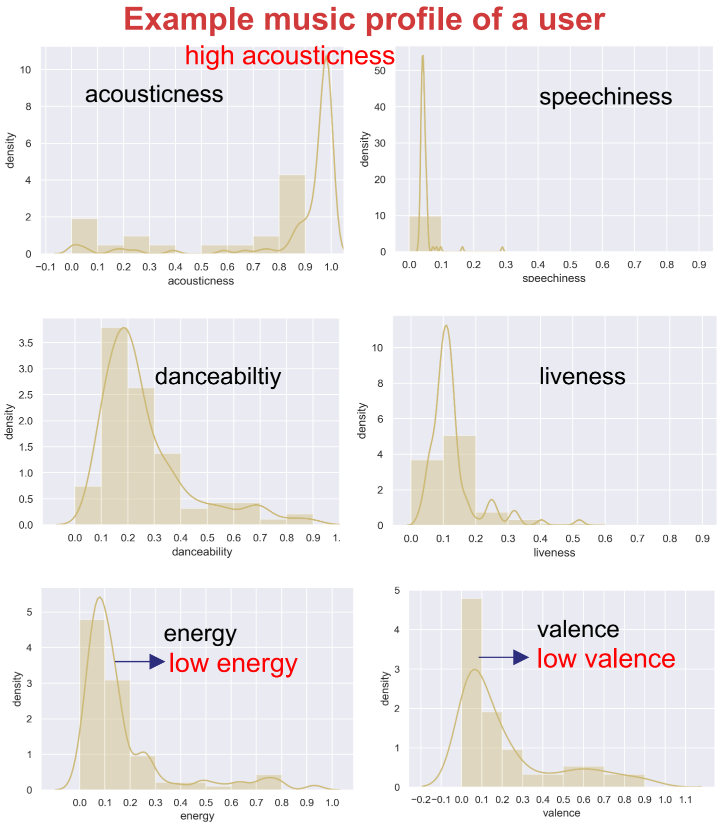
[7]

https://developer.spotify.com/documentation/web-api/

The personalized method

User Music Preference

- Model the user's music preferences with their top listened tracks from Spotify by Gaussian Mixture Model (GMM) in each feature dimension



[8]

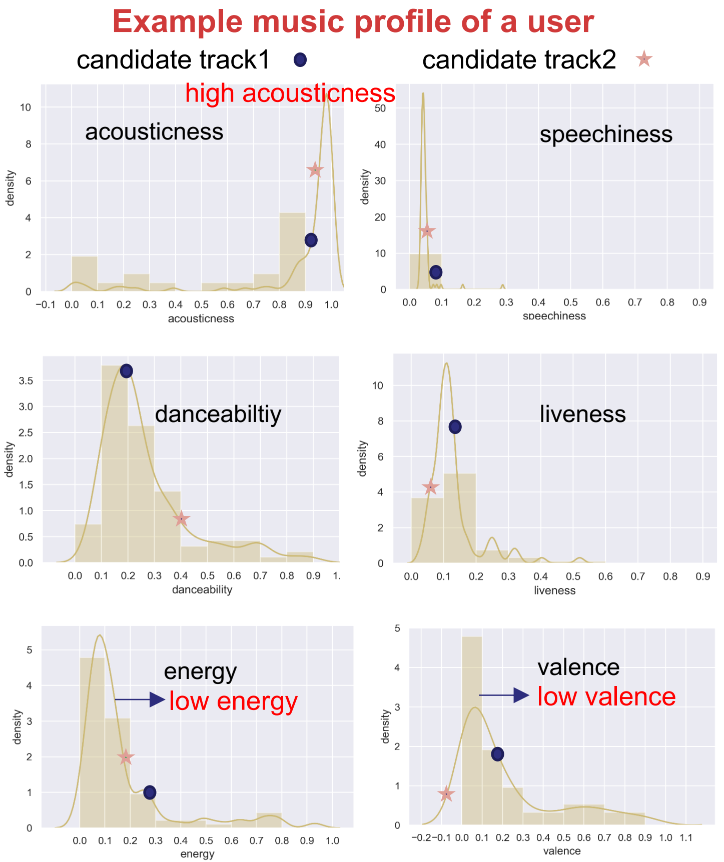
The personalized method

Music Preference Modeling

- Model the user's music preferences with their top listened tracks from Spotify by Gaussian Mixture Model (GMM) in each feature dimension

During recommendation

- In each feature dimension:
  - Map the candidate tracks from the recommendation dataset against the user model



[9]

The personalized method

ranking	track
1	track2
2	track1

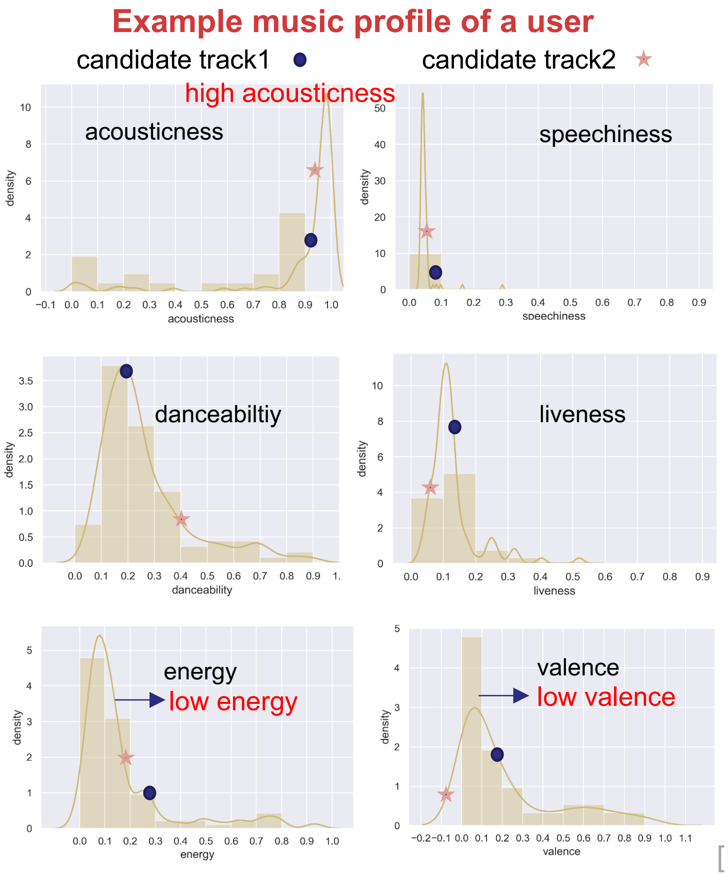
Music Preference Modeling

- Model the user's music preferences with their top listened tracks from Spotify by Gaussian Mixture Model (GMM) in each feature dimension

ranking	track
1	track1
2	track2

During recommendation

- In each feature dimension:
    - Map the candidate tracks from the recommendation dataset against the user model
    - Get a ranked list based on the matching scores



[10]

The personalized method

ranking	track
1	track2
2	track1

Music Preference Modeling

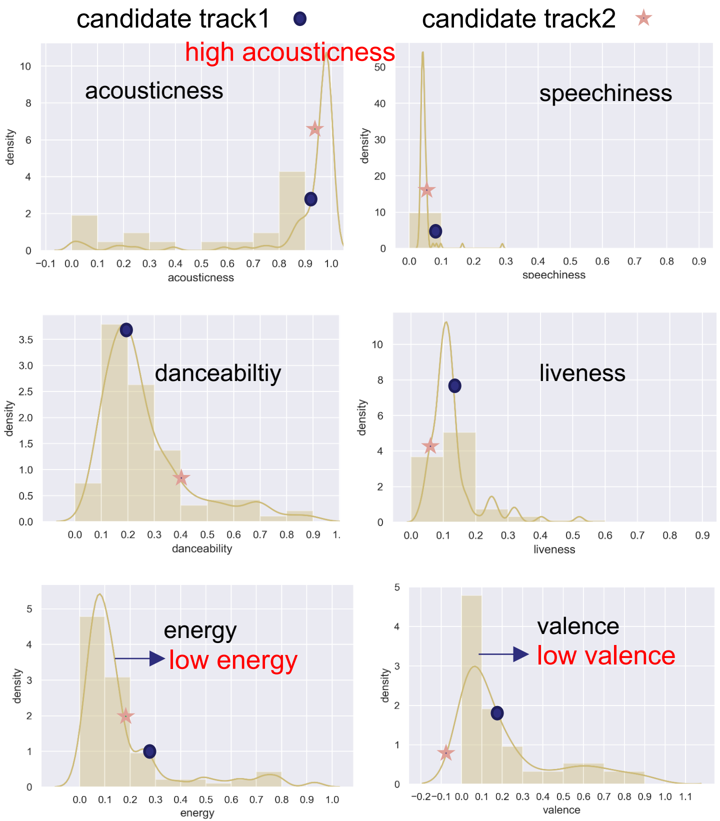
- Model the user's music preferences with their top listened tracks from Spotify by Gaussian Mixture Model (GMM) in each feature dimension

ranking	track
1	track1
2	track2

During recommendation

- In each feature dimension:
    - Map the candidate tracks from the recommendation dataset against the user model
    - Get a ranked list based on the matching scores
- Aggregate rankings from all feature dimensions and recommend the top 10 with the lowest ranking

Example music profile of a user



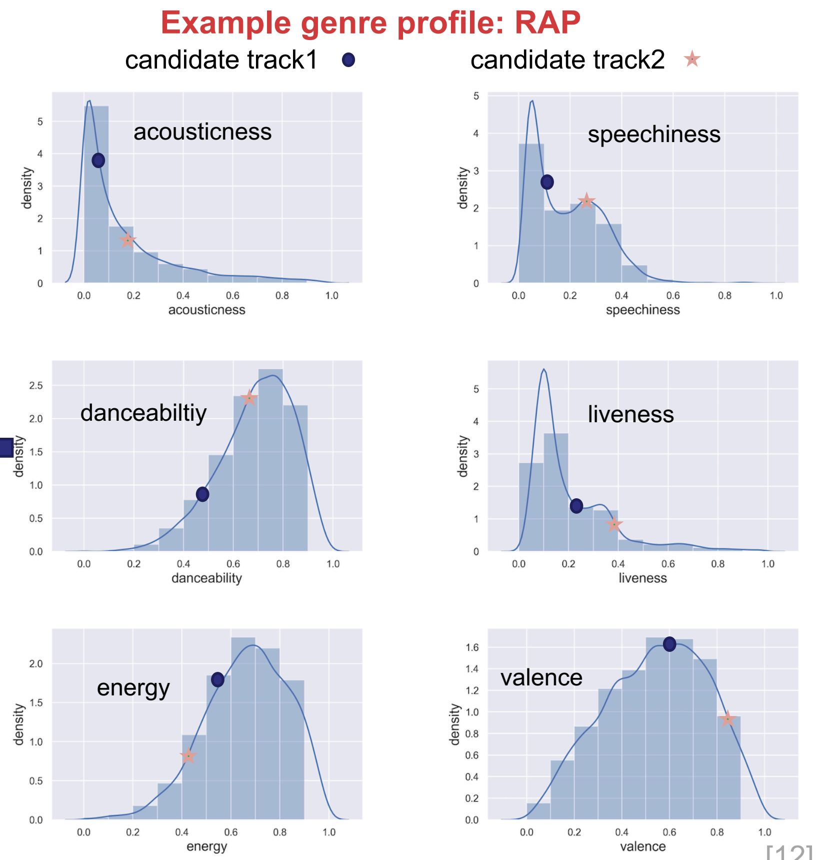
[11]

The non-personalized method

Genre-typical profile

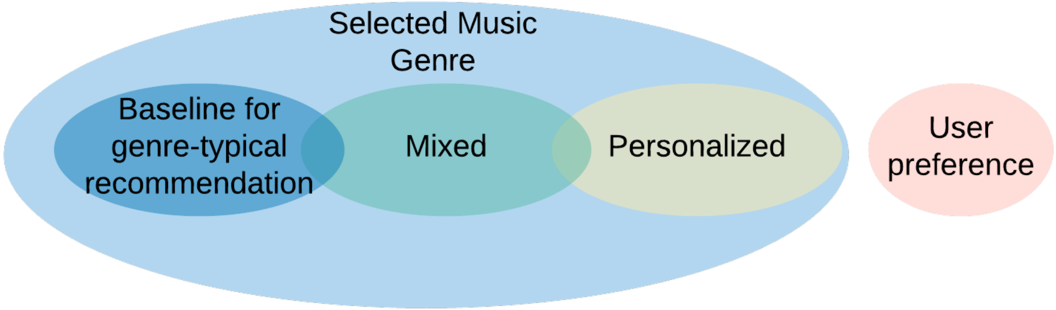
- Model genre-typical profile with the tracks from the genre highlighted artists

ranking	track
1	track2
2	track1



The mixed method

- Aggregate rankings from both personalized method and non-personalized method (weight=0.5)
  - $score_{mix} = weight * (n - r_{personal} + 1) + (1 - weight) * (n - r_{baseline} + 1)$



[13]

Online study

Consent forms

↓

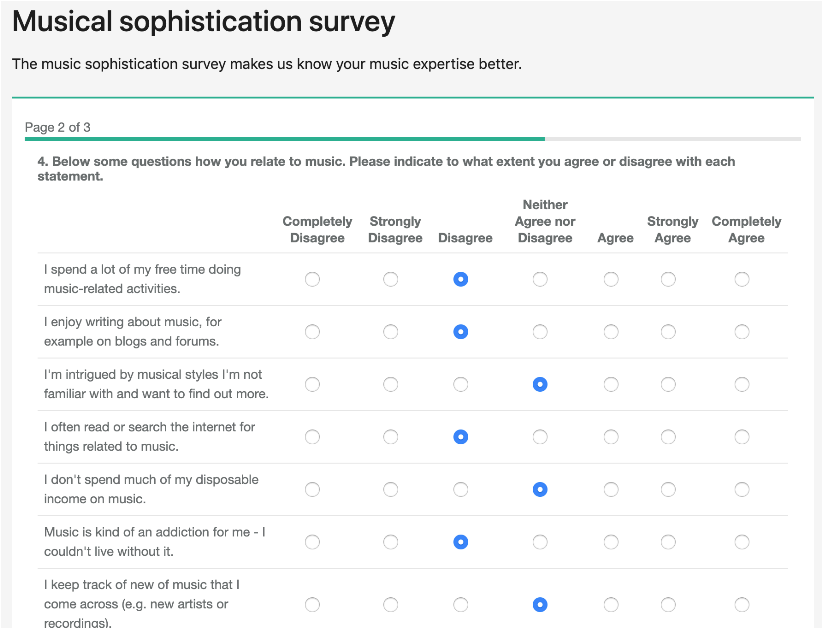
Login with Spotify Account

↓

Fill in the survey for Musical Sophistication

↓

Select a music genre to explore → Recommendations from different methods



Musical sophistication survey

The music sophistication survey makes us know your music expertise better.

Page 2 of 3

4. Below some questions how you relate to music. Please indicate to what extent you agree or disagree with each statement.

	Completely Disagree	Strongly Disagree	Disagree	Neither Agree nor Disagree	Agree	Strongly Agree	Completely Agree
I spend a lot of my free time doing music-related activities.	○	○	●	○	○	○	○
I enjoy writing about music, for example on blogs and forums.	○	○	●	○	○	○	○
I'm intrigued by musical styles I'm not familiar with and want to find out more.	○	○	○	●	○	○	○
I often read or search the internet for things related to music.	○	○	●	○	○	○	○
I don't spend much of my disposable income on music.	○	○	○	●	○	○	○
Music is kind of an addiction for me - I couldn't live without it.	○	○	●	○	○	○	○
I keep track of new of music that I come across (e.g. new artists or recordings).	○	○	○	●	○	○	○

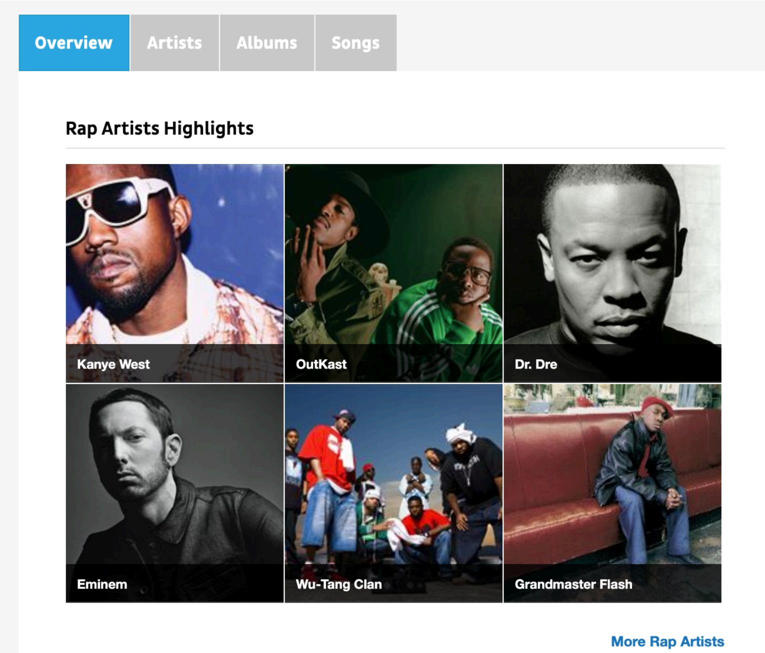
Daniel Müllensiefen, Bruno Gingras, Jason Musil, and Lauren Stewart. 2014. The musicality of non-musicians: an index for assessing musical sophistication in the general population. PloS one 9, 2 (2014), e89642.

Genre dataset

- Retrieved genre highlighted artists from Allmusic.com
- Extended the dataset with Spotify API

Table 1: genre dataset

genre	# tracks	# artists
avant-garde	3307	349
blues	2489	252
classical	4116	444
country	2728	281
electronic	3818	388
folk	3101	317
jazz	3346	347
new-age	3897	395
rap	3351	345
r&b	3299	334



Highlighted artist from genre "rap" retrieved from allmusic.com (https://www.allmusic.com/genres)

Online experiment

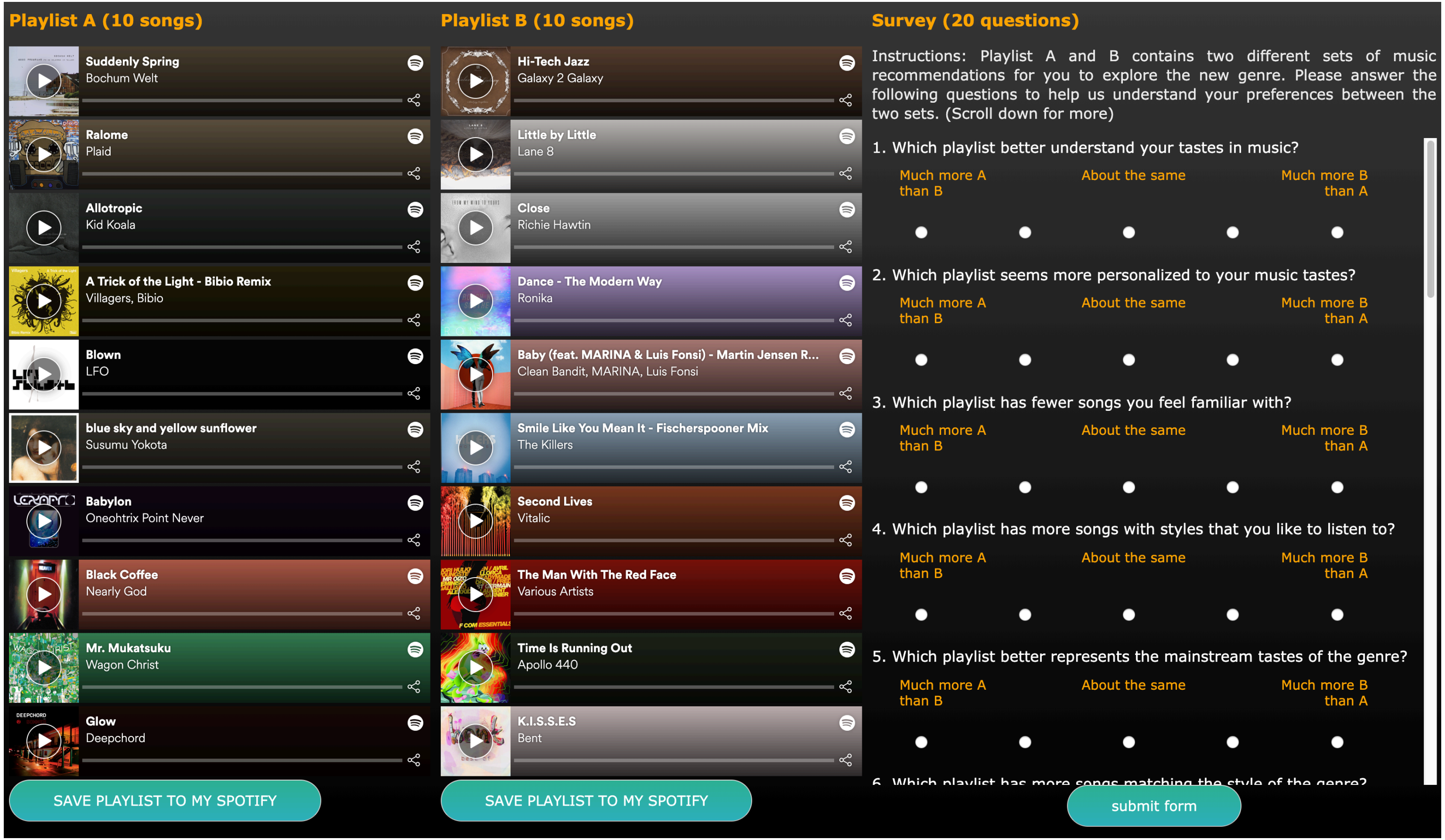
RQ: Can we give more helpful recommendations than the genre-typical tracks from the non-personalized baseline?

Comparative design

- Compare baseline with the personalized method
- Compare baseline with the mixed method.

- 156 validate response (78 females and 78 males)

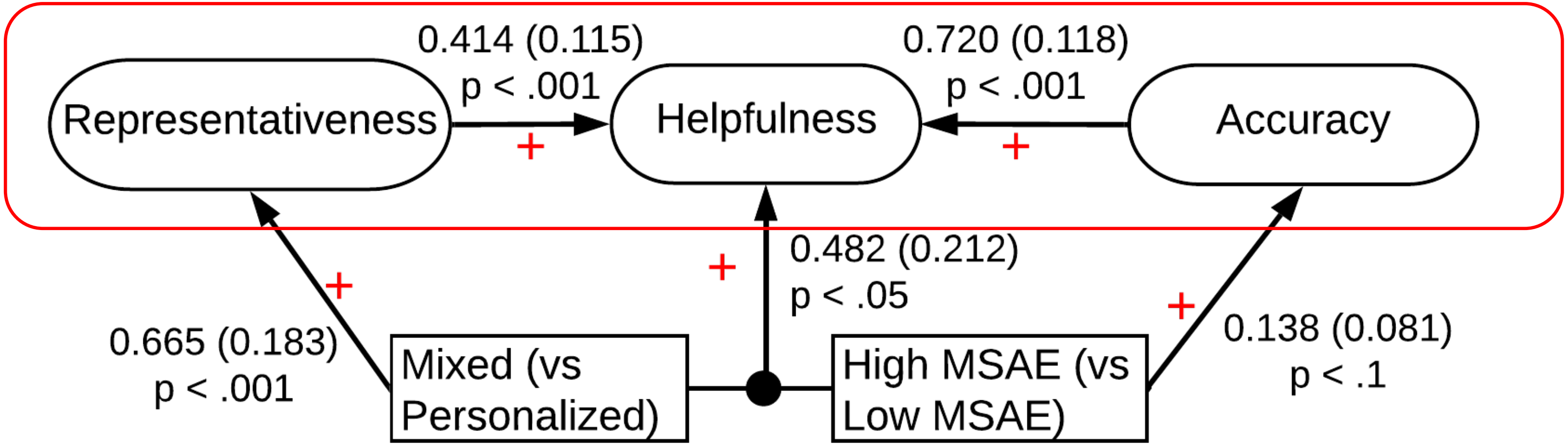
Michael D Ekstrand, F Maxwell Harper, Martijn C Willemsen, and Joseph A Konstan. 2014. User perception of differences in recommender algorithms. In Proceedings of the 8th ACM Conference on Recommender systems. ACM, 161–168.


Playlist A (10 songs)
Suddenly Spring
Bochum Welt
Ralome
Plaid
Allotropic
Kid Koala
A Trick of the Light - Bibio Remix
Villagers, Bibio
Blown
LFO
blue sky and yellow sunflower
Susumu Yokota
Babylon
Oneohtrix Point Never
Black Coffee
Nearly God
Mr. Mukatsuku
Wagon Christ
Glow
Deepchord
SAVE PLAYLIST TO MY SPOTIFY
Playlist B (10 songs)
Hi-Tech Jazz
Galaxy 2 Galaxy
Little by Little
Lane 8
Close
Richie Hawtin
Dance - The Modern Way
Ronika
Baby (feat. MARINA & Luis Fonsi) - Martin Jensen R...
Clean Bandit, MARINA, Luis Fonsi
Smile Like You Mean It - Fischerspooner Mix
The Killers
Second Lives
Vitalic
The Man With The Red Face
Various Artists
Time Is Running Out
Apollo 440
K.I.S.S.E.S
Bent
SAVE PLAYLIST TO MY SPOTIFY
Survey (20 questions)
Instructions: Playlist A and B contains two different sets of music recommendations for you to explore the new genre. Please answer the following questions to help us understand your preferences between the two sets. (Scroll down for more)
1. Which playlist better understand your tastes in music?
Much more A than B
About the same
Much more B than A
2. Which playlist seems more personalized to your music tastes?
Much more A than B
About the same
Much more B than A
3. Which playlist has fewer songs you feel familiar with?
Much more A than B
About the same
Much more B than A
4. Which playlist has more songs with styles that you like to listen to?
Much more A than B
About the same
Much more B than A
5. Which playlist better represents the mainstream tastes of the genre?
Much more A than B
About the same
Much more B than A
6. Which playlist has more songs matching the style of the genre?
submit form

Considered aspects	Items	SEM Coef.
Accuracy	Which playlist has more songs that you find appealing?	0.949
Alpha: 0.96	Which playlist has more songs that you might listen to again?	0.942
AVE: 0.87	Which playlist has more obviously bad songs for you?	
	Which playlist has more songs that are well-chosen?	
Personalization (formerly)	Which playlist better understand your tastes in music?	0.933
	Which playlist seems more personalized to your music tastes?	0.876
	Which playlist has fewer songs you feel familiar with?	
	Which playlist has more songs with styles that you like to listen to?	0.947
Representativeness	Which playlist better represents the mainstream tastes of the genre?	
Alpha: 0.81	Which playlist has more songs matching the style of the genre?	0.818
AVE:0.65	Which playlist has fewer songs you would expect from the genre?	−0.772
	Which playlist seems less typical of the genre?	−0.779
Helpfulness	Which playlist better supports you to get to know the new genre?	0.716
Alpha: 0.77	Which playlist motivates you more to delve into the new genre?	
AVE: 0.61	Which playlist is more useful to explore a new genre?	0.626
	Which playlist has more songs that helps you understand the new genre?	0.402
Diversity	Which playlist has more songs that are similar to each other?	
Alpha: N.A.	Which playlist has a more varied selection of songs within the genre?	
AVE: N.A.	Which playlist would suit a broader set of tastes?	
	Which playlist has songs that match a wider variety of moods?	

[18]

Results - Structural Equational Model



Arrows represent the standardized coefficients with standard error between brackets and p-values.

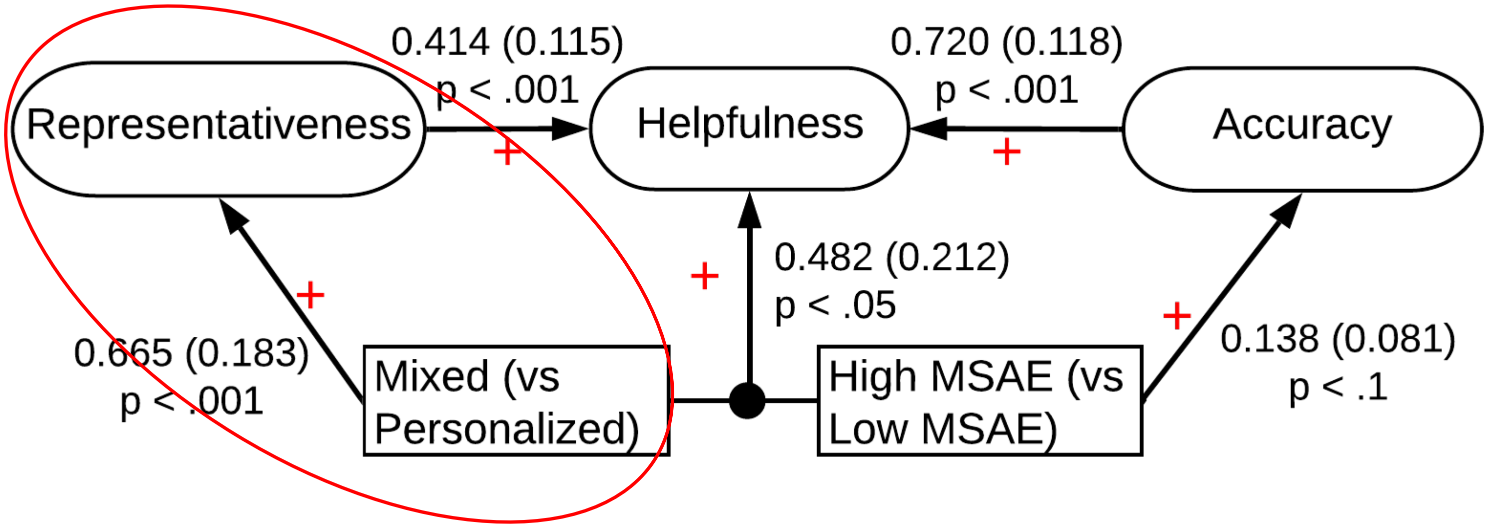
MSAE: Musical Sophistication Score for Active Engagement

[19]

Bart P Knijnenburg, Martijn C Willemsen, Zeno Gantner, Hakan Soncu, and Chris Newell. 2012. Explaining the user experience of recommender systems. User Modeling and User-Adapted Interaction 22, 4-5 (2012), 441–504.

Jheronimus Academy of Data Science

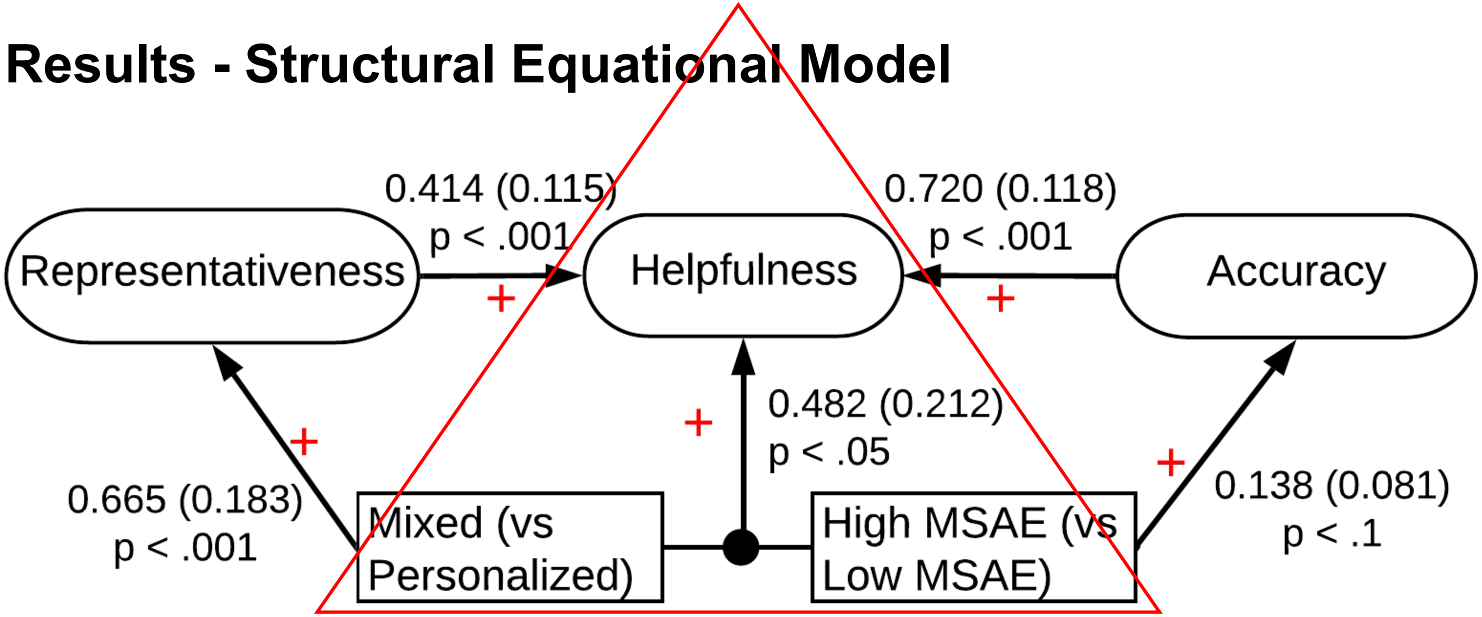
Results - Structural Equational Model



Arrows represent the standardized coefficients with standard error between brackets and p-values.

MSAE: Musical Sophistication Score for Active Engagement

Results - Structural Equational Model

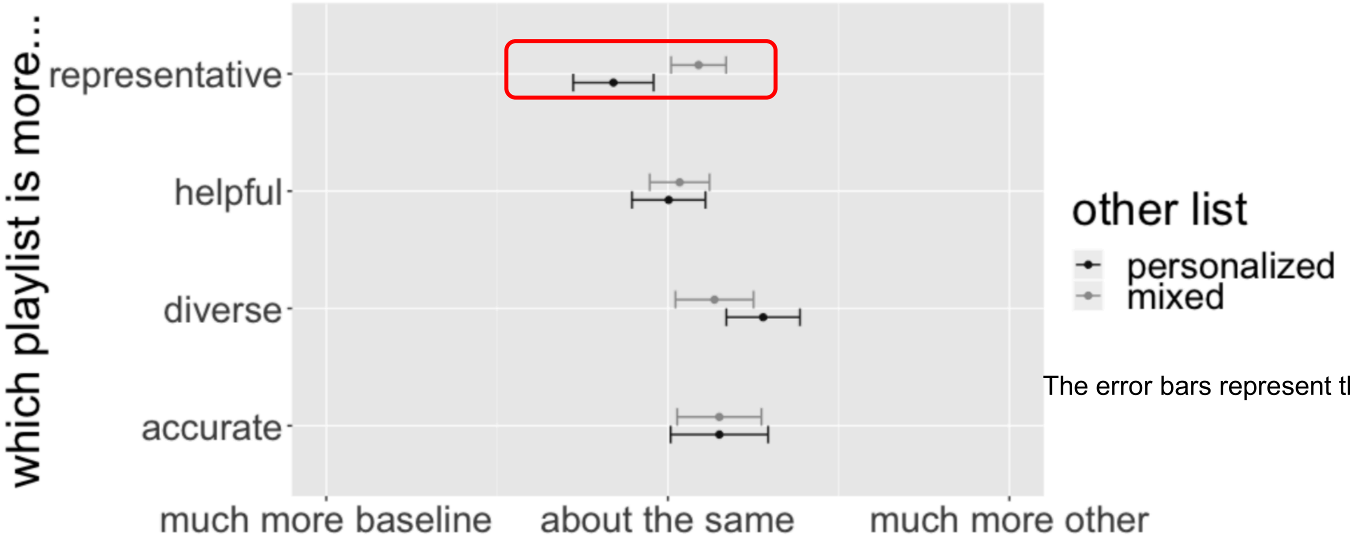


Arrows represent the standardized coefficients with standard error between brackets and p-values.

MSAE: Musical Sophistication Score for Active Engagement

Results - Absolute difference

The recommendations from the baseline method are perceived more representative than the personalized method, but less representative than the mixed method



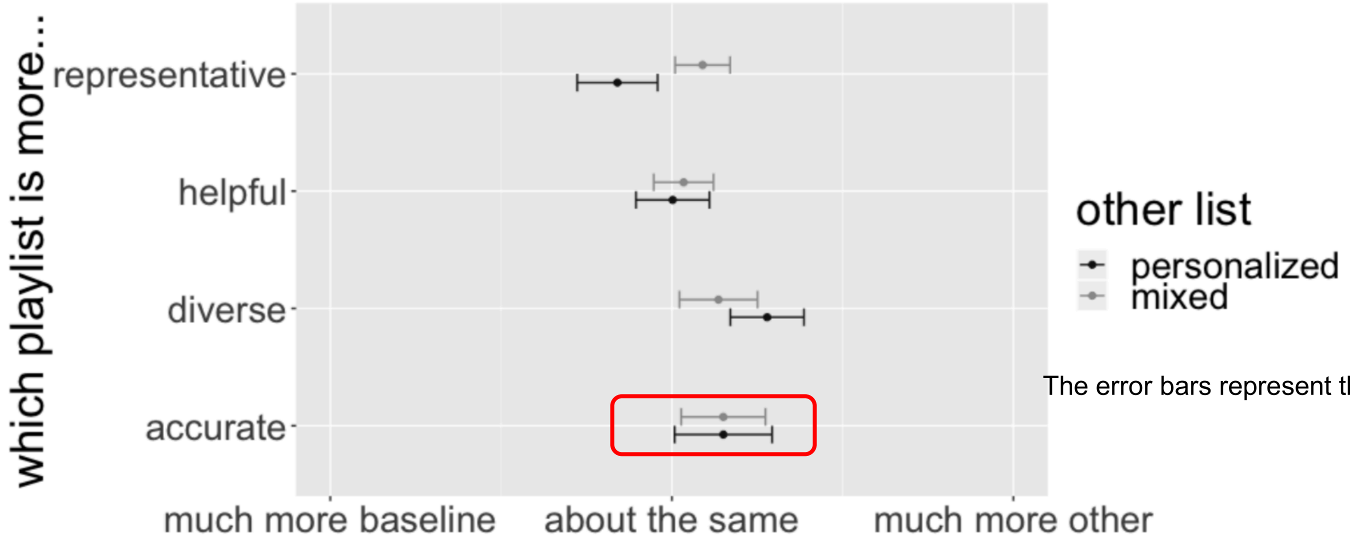
The error bars represent the 95% confidence interval

[22]

Jheronimus Academy of Data Science

Results - Absolute difference

The recommendations from both personalized and the mixed method are perceived more accurate than those from the baseline

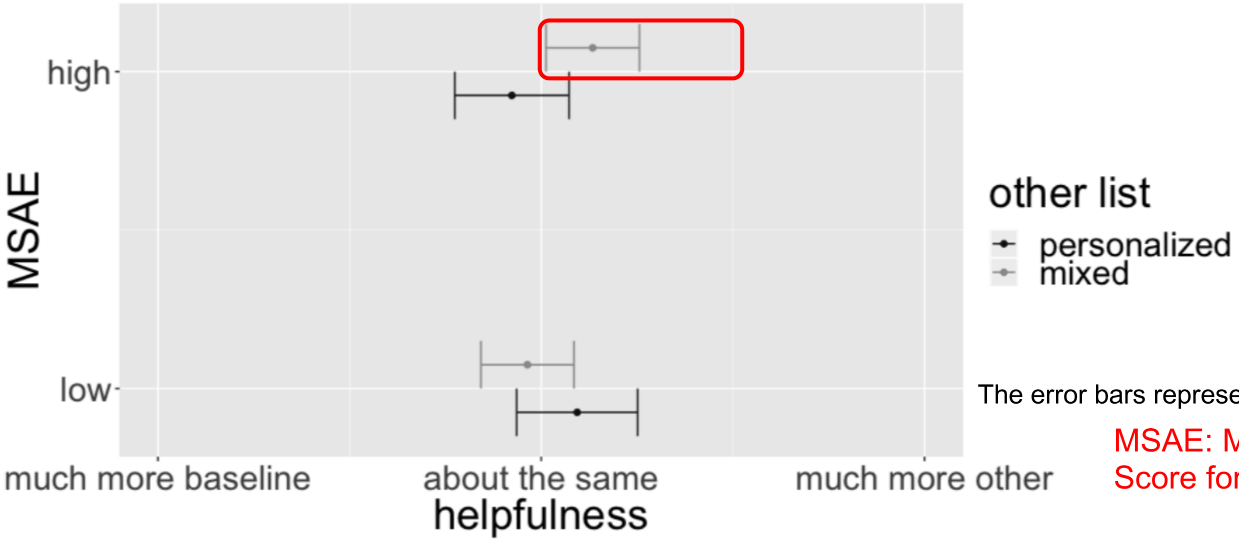


The error bars represent the 95% confidence interval

[23]

Which method is more helpful?

Users with high MSAE perceived the mixed method to be more helpful than the purely personalized method



The error bars represent the 95% confidence interval

MSAE: Musical Sophistication Score for Active Engagement

Conclusions and Future work

- In general, we found that both methods (*the personalized and the mixed*) are not perceived more helpful than the baseline.
- Perceived helpfulness is positively related to both perceived accuracy and representativeness
- Users with high MSAE perceived the mixed method to be more helpful
  - balance the perceived accuracy and representativeness
  - provide different methods for users with different musical expertise

- Follow up (more interaction and understandability)
  - Visualization (*improve perceived understandability*)
  - Addition of mood control (*improve perceived control*)

Thanks! Q & A?