DIE PERFEKTE MUSIKEMPFEHLUNG ---**PERFEKT FUR WEN?**

PETER KNEES

FUTURE MUSIC CAMP 2019







ABOUT ME

- Music Information Retrieval researcher
 - Music search engines and interfaces
 - Music recommender systems
 - Recently: smarter tools for music creation
- PhD and PostDoc at JKU Linz (2005-2016)
- Since 2017: Assistant Professor at TU Wien

SESSION CONTENT

- Music Recommender Systems
 - Sources of data
 - Collaborative filtering
 - Content analysis
- Recommendation use cases
- Biases of platforms

RECOMMENDER CLASSIFICATION SCHEME

(based on users/community)

Collaborative Filtering (CF)

Context-aware Recommenders

(based on the usage context)

(a mixture of different approaches)

(based on item's content)

(product finder)

Content-based Recommenders

Knowledge-based Recommenders

Hybrid Recommenders



DATA FUELS RECOMMENDERS

- **Interaction Data**
- Listening logs, listening histories
- Feedback ("thumbs"), purchases
- **User-generated**
- Tags, reviews, stories
- **Curated collections**
- Playlists, radio channels
- CD album compilations









lost.fm









amazon



Epinions.com

DATA FUELS RECOMMENDERS

Content (audio, symbolic, lyrics)

- Machine listening/content analysis
- Human labelling

Meta-data

- Editorial
- Curatorial
- Multi-modal (album covers etc.)



COLLABORATIVE FILTERING

- Exploits interaction data
- "People who listened to track A, also listened to track B"
- Main underlying assumption: users that had similar taste in the past, will have similar taste in the future
- Typical methods
 - Comparing rows/columns in matrix
 - Matrix factorization



Items (~10's M)



THE INTERACTION MATRIX

Can contain number of plays, listening time, rating, etc.

Listening	Track 1	Track 2	Track 3	Track 4
User 1	3		2	3
User 2	4	3	4	
User 3	3	2	1	4
User 4		5	4	1
User 5	5		3	

Similar users found, e.g. by comparing user profiles

"user profile"

FACTORS HIDDEN IN THE DATA

Assumption of matrix factorization-based recommender systems:
Observed data are interactions of 2 factors: users and items

USERS

Calculate latent factors for users and items from the data



MATRIX FACTORIZATION

- Decompose rating matrix into user and item matrices of lower dimension k Learning factors from given ratings using stochastic gradient descent



Prediction of rating: inner product of vectors of user u and item i

Factors not necessarily interpretable (just capture variance in data)

$$-x_u^T y_i)^2 + \lambda(\|x_u\|^2 + \|y_i\|^2)$$



EXAMPLE OF COLLABORATIVE FILTERING OUTPUT

People who liked **Disturbed – The Sound of Silence**, also liked...

- 1. Bad Wolves Zombie
- 2. Five Finger Death Punch Bad Company
- 3. Disturbed The Light

4. Metallica – Nothing Else Matters

Effect/issue: popularity biases







FACTORS HIDDEN IN THE DATA

Assumption of matrix factorization-based recommender systems:
Observed data are interactions of 2 factors: users and items

USERS

Calculate latent factors for users and items from the data
But it's a bit more complex...



FACTORS HIDDEN IN THE DATA USERS

INTRINSIC "What?"

Listener Background

GOAL "Why?"

Listener Intent

CONTEXT "Where & How?"

Listener Context



AUDIO CONTENT ANALYSIS: SELECTED FEATURES



Disturbed The Sound of Silence

- Timbre
- Tonal features



Effect/issue: no popularity biases, but no personalization

▶ Beat/downbeat → Tempo: 85 bpm



e.g. for genre classification, "more-like-this" recommendations

e.g. for melody extraction, cover version identification

Different versions of this song: Simon & Garfunkel - The Sound of Silence Anni-Frid Lyngstad (ABBA) - En ton av tystnad etc.

Semantic categories via machine learning not_danceable, gender_male, mood_not_happy



A MIXTURE OF MANY THINGS

- Incorporation of different sources and complementary information
- Machine Learning to fit which recommender/information in which context
- E.g. to control for diversity, exploitation vs exploration, novelty, etc.
- Different types of recommenders and models for different features







Oramas et al., RecSys DLRS 2017

FOCUS ON: LEAN-IN EXPERIENCE

Lean in: Building Playlists

🔰 Тоо	much vocoder			F	
т	ITLE	ARTIST	ALBUM		Ŀ
+ 2	4K Magic	Bruno Mars	24K Magic	2017-03-15	3:46
+ F	ix	Blackstreet	Another Level	2017-03-15	4:05
+ G	Good Lovin'	Blackstreet	Another Level	2017-03-15	4:32
Recommend Based on the so	ded Songs ▲ ngs in this playlist				REFRESH
	D Back & Forth	Aaliyah		Age Ain't Nothing But A Nu	••• 3:51
AD	D Get It On Tonite	Montell Jorda	an	Get It OnTonite	4:36
AD	D Wifey - Club Mix/Dirty Ver EXPLIC	CIT Next		Work It Out!	4:02
AD	D Doin' It EXPLIC	CIT LL Cool J		Mr. Smith (Deluxe Edition)	4:54
AD	D Freek'n You	Jodeci		The Show, The After Party,	6:19



SOURCE: RECSYS'17 TUTORIAL ON MUSIC RECOMMENDATION

FOCUS ON: RE-DISCOVERY



Your Daily Mixes



Chris Cornell, Soundgarden, Red

Hot Chili Peppers and more

MADE FOR FABIEN

Daily Mix 1



Focus on stuff you know you like Personalized, leaning towards exploit

SOURCE: RECSYS'17 TUTORIAL ON MUSIC RECOMMENDATION

FOCUS ON: HYPER-PERSONALIZED DISCOVERY

<	Thumbprint Radio Station	Q		
Music inspi your station	red by your 1,285 thumbs from a	across all		
THUMBED U	THUMBED UP SONGS			
	Baiao Embolado Forro In The Dark 2:36			
	21st Century Red Hot Chili Peppers 4:22	RADIO		
FOOTFUHIES ONE DONE	Times Like These Foo Fighters 4:26			
<u> </u>	Tive Razao Seu Jorge	►		



About discovering Intended to feel lik curated. Just. For

Leaning towards

er.com/en/		☆	
	Your Vour Vour	EN DVET WE tape of fresh music. Enjoy ery Monday, so save your Gouyon by Spotify • 30 s FOLLOWING	new disco favourites! ongs, 2 hr
g new stuff. ike it's r. Me.	Q Filter TITLE + One Step Ahead + Not My Slave	ARTIST Split Enz Oingo Boingo	ALBUM Waiata Boi-Nge
explore	 + She Sheila + Drifting, Falling 	The Producers The Ocean Blue	You Ma The Oc



PETER KNEES (TU WIEN) – FMC 2019

ONE MORE THING...

INTRINSIC "What?"

Listener Background

GOAL "Why?"

Listener Intent

CONTEXT "Where & How?"

Listener Context



USERS



THE SERVICE INTRODUCES FURTHER BIASES



Service Context

- Which content is provided/recommended?
- e.g. Soundcloud recommends different content than Spotify
- Why is this service in place? What is the purpose/identified market niche? What are the identified use cases? (Discovery? Radio? Exclusives? Quality?) Do they push their own content (cf. Netflix)?

- How do catalog and service aims depend on context? Are there licensing issues/restrictions in particular countries? Is the service context-aware? (e.g. app vs desktop/browser)



IMPLICATIONS

- Different methods with different biases incorporated Algorithmic design choices to deal with biases
- Service design choices and restrictions introduce biases
- Feedback loops amplify popularity biases
- Platforms are in control and can shape recommendations





	ACCESS	MUSIC	PROFILE	STA
POPULAR	ACCESS	NIUSIC	PROFILE	SIA

https://artists.spotify.com/faq/promotion



ALL ABOUT THE MUSIC ...?

- Analysis of Spotify
 playlist dataset
- 1 million US playlists
- Webcrawler to identify record label of tracks
- Information for about
 50% of tracks



MAJOR LABEL (TRACKS)

ALL ABOUT THE MUSIC ...?

- Investigating playlist diversity wrt. labels
- Entropy-based
- Left: pure Right: diverse
- Any comments?

300000

250000

200000 -

150000 .

100000

50000 -

0







CONTACT

Dr. Peter Knees **Assistant Professor TU Wien, Institute of Information Systems Engineering**



peter.knees@tuwien.ac.at



https://www.ifs.tuwien.ac.at/~knees/



@peter knees



https://www.linkedin.com/in/peterknees/



FAKULTÄT FÜR INFORMATIK

Faculty of Informatics