Where To Next? A Dynamic Model of User Preferences


Francesco Sanna Passino
Imperial College London


Lucas Maystre
Spotify


Dmitrii Moor
Spotify


Ashton Anderson
University of Toronto


Mounia Lalmas
Spotify

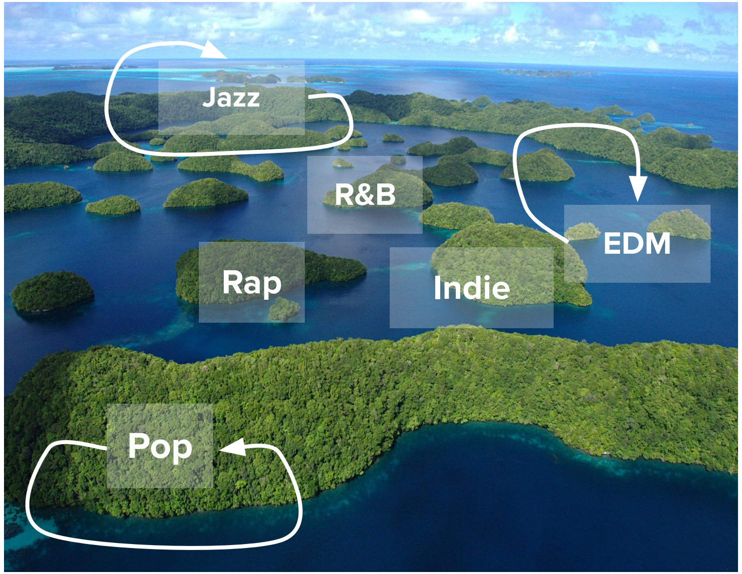
Idea

Most recommender systems attempt to capture ***simultaneous*** *preferences*:

"Users who like A tend to like B as well"

Recommender systems are effective at **exploiting** current preferences, but are less efficient at **exploring**.

Diversity of consumption is positively associated with ***user conversion and retention***.



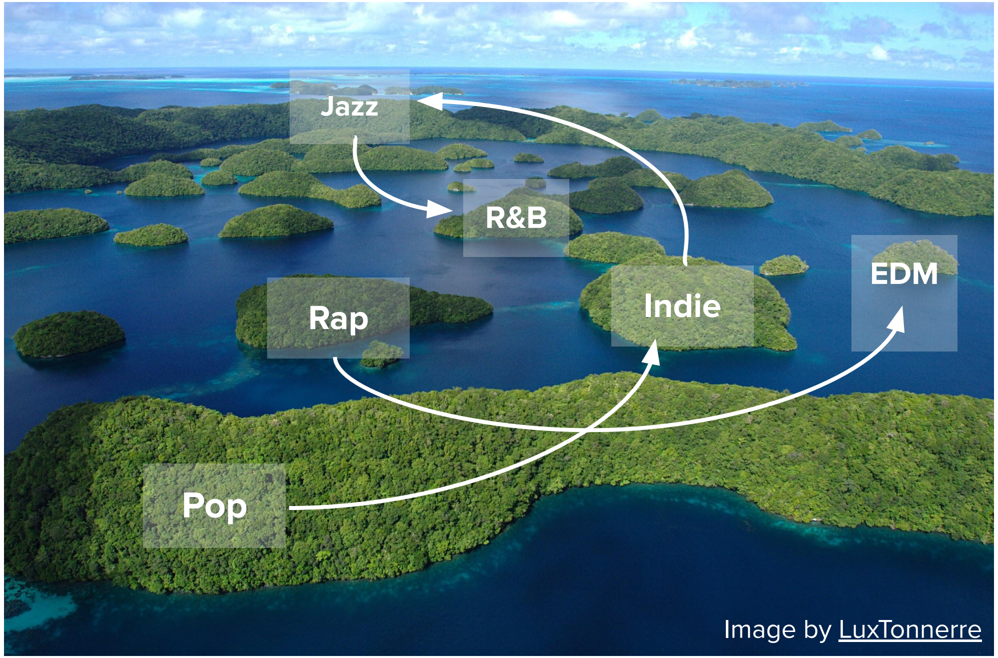
Ashton Anderson, Lucas Maystre, Ian Anderson, Rishabh Mehrotra, and Mounia Lalmas. 2020. Algorithmic Effects on the Diversity of Consumption on Spotify. The Web Conference 2020 (WWW '20).

Idea

Users' consumption **changes over time**.

Can we learn **structure** underlying users' trajectories?

This could provide a map we can use to **introduce meaningful diversity**.

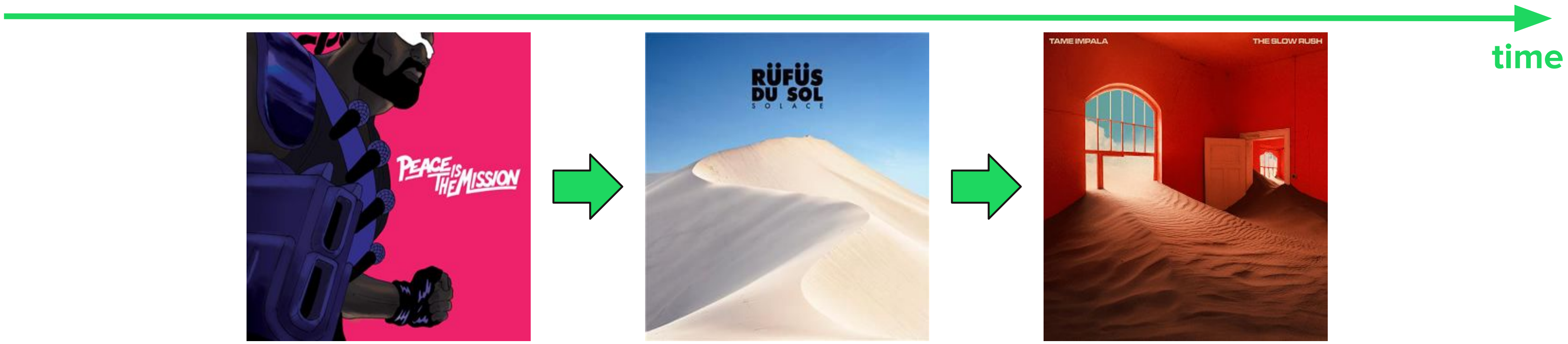


Idea

- Research question:

  “How well does preference for A at time t predict preference of B at time t+1?”

- Our proposal: **Preference Transition Model (PTM)**.

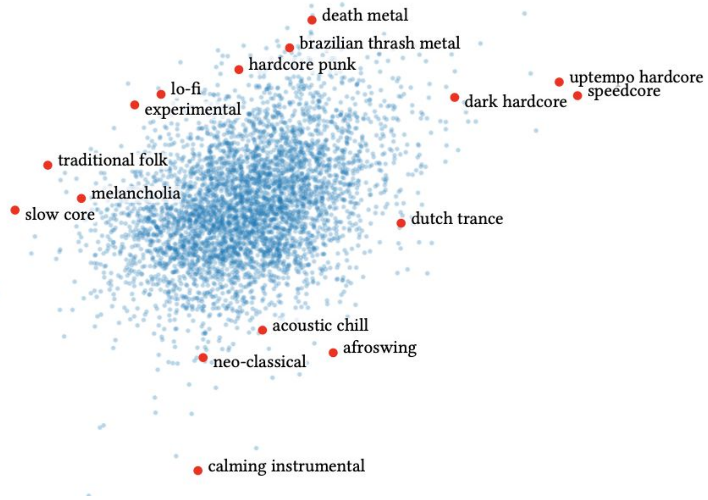


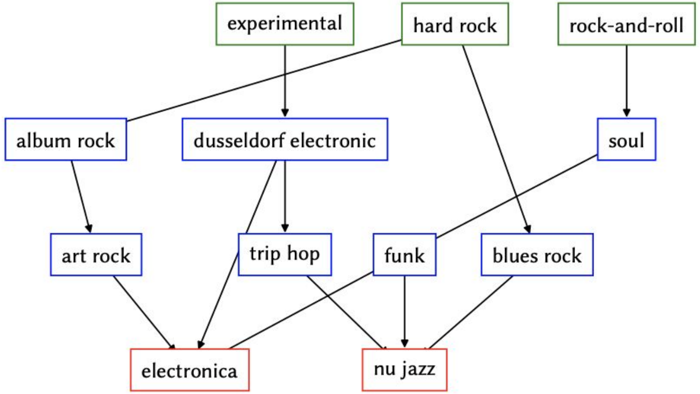
dance-pop indietronica neo-psychedelic

Album covers obtained from Spotify. Left: Major Lazer, “Peace Is The Mission”; Middle: Rüfüs Du Sol, “Solace”; Right: Tame Impala, “The Slow Rush”.

Background

- *Dynamic collaborative filtering*, for example dynamic matrix factorisation methods, or tensor-based approaches.
- *Sequential recommender systems*. In this work, we focus on obtaining interpretable insights into how user preferences change i the long term.
- *Diversity in recommender systems*.
- *Estimation of relationships between multivariate processes*.





Spotify

Dataset

Microgenres

> 4000 **fine-grained musical genres**. We associate each song to a single genre.

Longitudinal consumption traces

Stream counts for 100k UK Premium subscribers. **Spans 2016-2020,** aggregated per quarter.

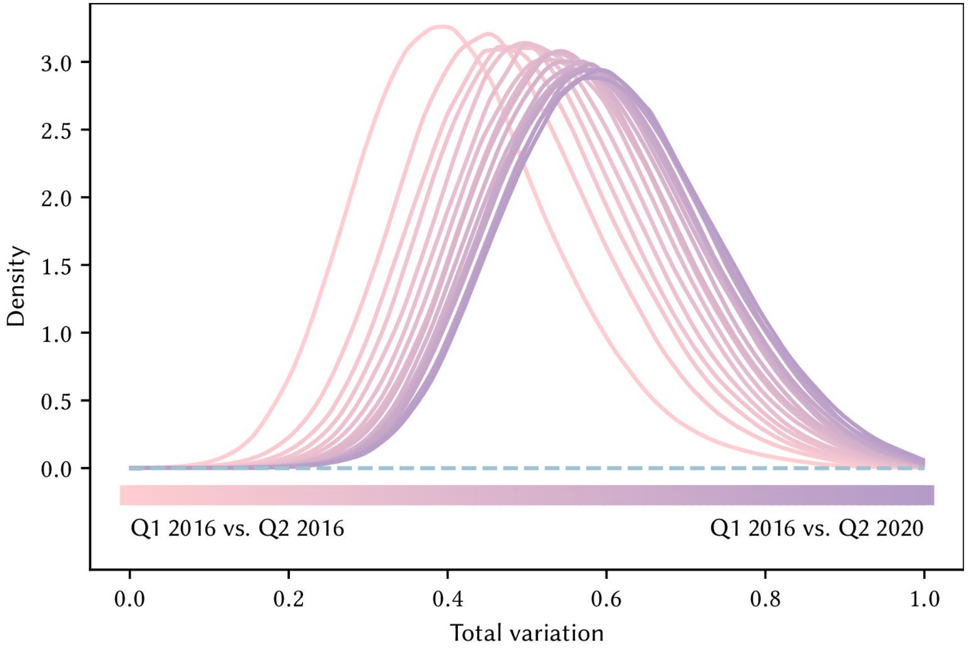
user 3	***t*** **= 1**	***t*** **= 2**	**...**	***t*** **= T**

user 2	***t*** **= 1**	***t*** **= 2**	**...**	***t*** **= T**

user 1	***t*** **= 1**	***t*** **= 2**	**...**	***t*** **= T**
pop	40	0	...	3
rock	3	4	...	23
...	...	...	...	...
edm	0	0	...	61

Spotify

Preferences Are Shifting



For each user, we measure the distance between streams in 2016 Q1 and streams in future periods.

Users' consumption is drifting over the years.

$$\text{User } (1) \quad \overbrace{\begin{bmatrix}1 & 0 & 0\end{bmatrix}}^{\tilde{\boldsymbol{\pi}}_1^{(t-1)}} \cdot \overbrace{\begin{bmatrix}0.5 & 0.4 & 0.1\\ 0.2 & 0.4 & 0.4\\ 0.2 & 0.3 & 0.5\end{bmatrix}}^{\boldsymbol{A}} = \overbrace{\begin{bmatrix}0.5 & 0.4 & 0.1\end{bmatrix}}^{\hat{\boldsymbol{\pi}}_1^{(t)}}$$

soul, new age, blues

predicted distribution at next step

"If you listen to *soul*, you will likely listen to *new age* next"

$$\text{User } (2) \quad \overbrace{\begin{bmatrix}0 & 1 & 0\end{bmatrix}}^{\tilde{\boldsymbol{\pi}}_2^{(t-1)}} \cdot \overbrace{\begin{bmatrix}0.5 & 0.4 & 0.1\\ 0.2 & 0.4 & 0.4\\ 0.2 & 0.3 & 0.5\end{bmatrix}}^{\boldsymbol{A}} = \overbrace{\begin{bmatrix}0.2 & 0.4 & 0.4\end{bmatrix}}^{\hat{\boldsymbol{\pi}}_2^{(t)}}$$

"If you listen to *new age*, you will likely listen to *blues* next"

Relations between genres are not necessarily *symmetric*

The matrix **A** encodes the graph structure between the genres

Weighted adjacency matrix of the **Genre Interaction Graph (GIG)**

Preference Transition Model (PTM): key features

Notation

Genre counts at time *t*

$$\boldsymbol{n}_i^{(t)} = \left(n_{i1}^{(t)}, \ldots, n_{iN}^{(t)}\right)$$

User activity at time *t*

$$\xi_i^{(t)} = \sum_{j=1}^{N} n_{ij}^{(t)}$$

Genre distribution at *t*

$$\boldsymbol{\pi}_i^{(t)} = \boldsymbol{n}_i^{(t)} / \xi_i^{(t)}$$

Preference transition matrix

$$\boldsymbol{A} = \{a_{jk}\} \in [0, 1]^{N \times N}$$

$$\boxed{\begin{aligned} \xi_i^{(t)} &\sim \text{Poisson}\left(\xi_i^{(t-1)}\right), \\ \boldsymbol{n}_i^{(t)} \mid \xi_i^{(t)}, \tilde{\boldsymbol{\pi}}_i^{(t-1)} &\sim \text{Multinomial}\left(\xi_i^{(t)}, \tilde{\boldsymbol{\pi}}_i^{(t-1)} \boldsymbol{A}\right) \end{aligned}}$$

Key feature 1: Exponentially weighted moving average distribution

$$\tilde{\boldsymbol{\pi}}_i^{(t)} = (1 - \gamma) \tilde{\boldsymbol{\pi}}_i^{(t-1)} + \gamma \boldsymbol{\pi}_i^{(t)}$$

Global exploration parameter

$$\gamma \in [0, 1]$$

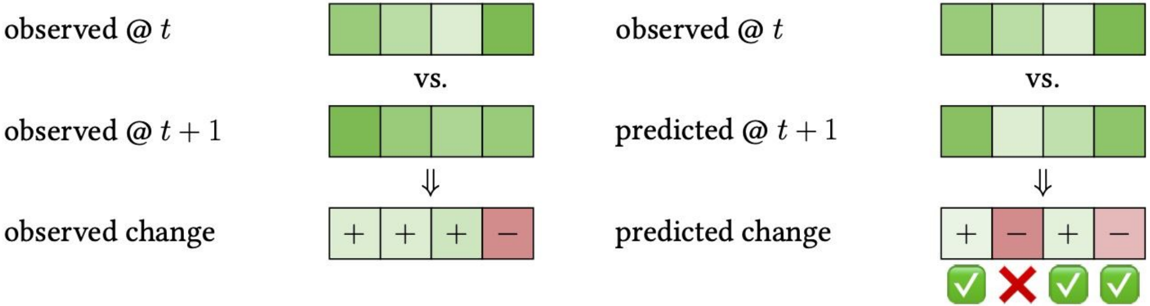
Key feature 2: Poisson-Multinomial two-stage distribution

User activity and genre distributions are modelled **separately**.

Model evaluation: three tasks

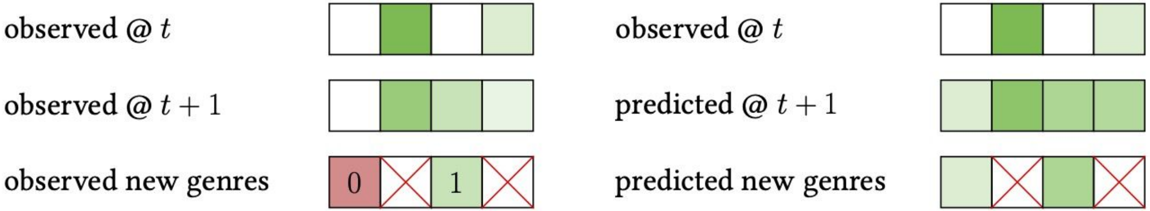
1. Minimise **total variation** between observed and predicted genre distributions.

2. Predict if the consumption of each genre **increases or decreases (+/-)**.



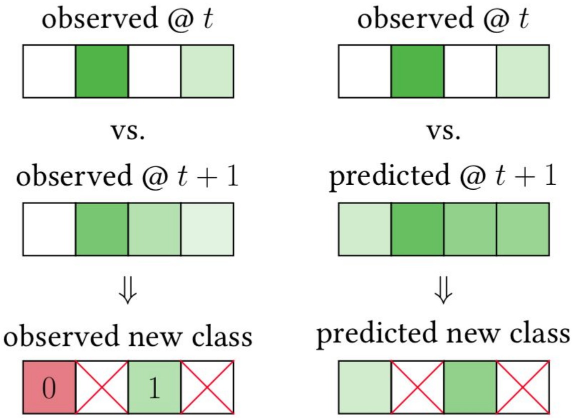
3. Predict which **new genres** will be streamed.

with respect to *t-1* or entire streaming history



Predictive Performance

Based on streams up to time *T*, can we **predict discoveries** at time *T*+1?



	New classes (all)	
Model	ROC-AUC	PR-AUC
PTM ($\gamma = 0.360$)	**0.889**	**0.039**
Poisson AR	0.854	0.033
DPF, $K = 5$	0.849	0.016
NMF, $K = 50$	0.853	0.025
Previous obs.	—	—

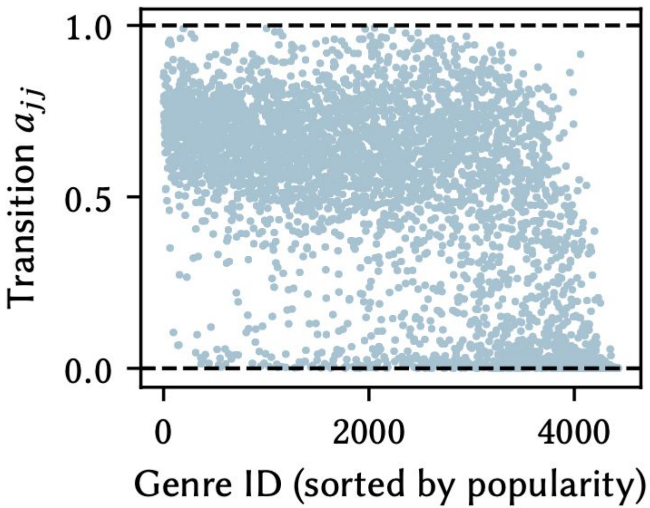
Predictive Performance

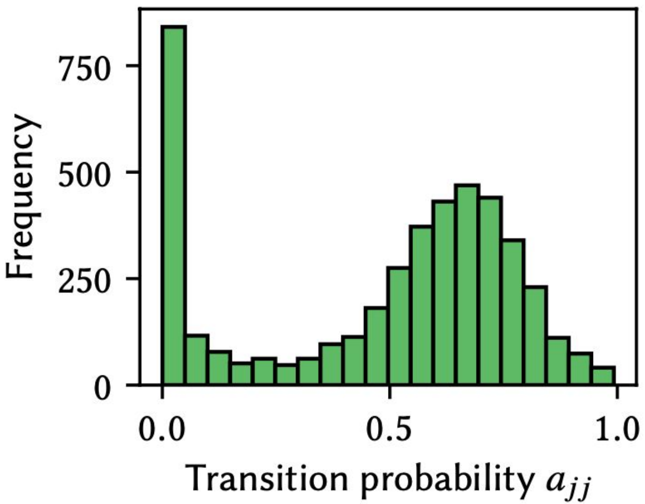
The PTM has the best performance for all tasks.

	Total variation	Plus-minus (+/-)	New classes (t)		New classes (all)	
Model	TV	ROC-AUC	ROC-AUC	PR-AUC	ROC-AUC	PR-AUC
PTM ($\gamma = 0.360$)	**0.378**	**0.698**	**0.944**	**0.267**	**0.889**	**0.039**
Poisson AR	0.416	0.663	0.915	0.190	0.854	0.033
DPF, $K = 5$	0.909	0.512	0.849	0.051	0.849	0.016
NMF, $K = 50$	0.509	0.646	0.914	0.189	0.853	0.025
Previous obs.	0.389	0.581	0.639	0.221	—	—

Structure of the Transition Matrix

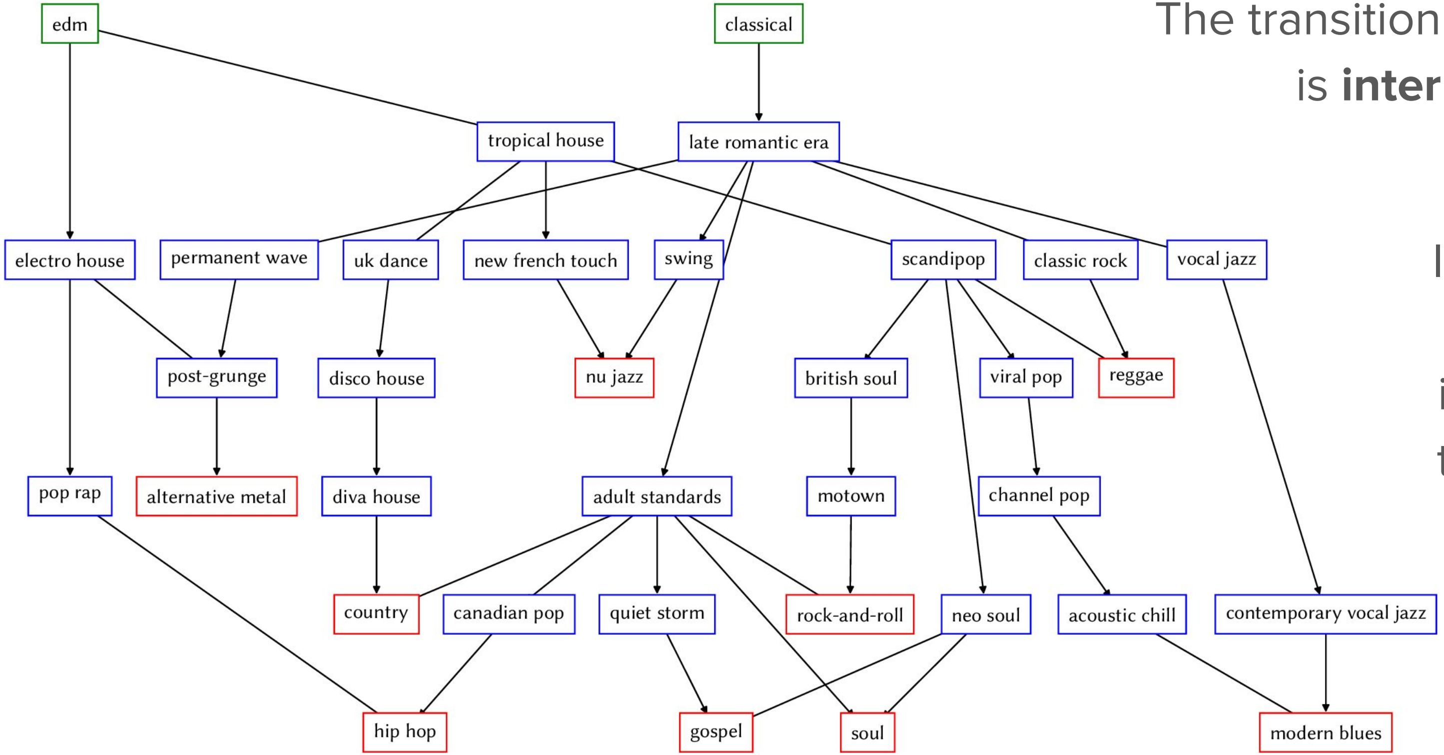
The transition matrix we learn is **interpretable**: **diagonal elements** tend to be large, since users that liked a given genre will probably like it also in the near future.





Spotify

Structure of the Transition Matrix



The transition matrix we learn is **interpretable**.

It could be used as a **map** to introduce users to a new genre.

Dafna Shahaf and Carlos Guestrin. 2010. Connecting the Dots between News Articles. ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD ’10).

Other datasets

The PTM has also been tested for:

- Prediction of movie tags (MovieLens),
- Prediction of music genres (Last.fm),
- Restaurant recommendations (Yelp).

Dataset	M	N	T	Interval	Start	End
Spotify	10000	4430	18	Quarter	Q1 2016	Q2 2020
Last.fm	450	1500	6	Quarter	Q3 2007	Q4 2008
MovieLens	1320	1128	5	Year	2015	2019
Yelp	1808	192	5	Year	2015	2019

		Total variation	Plus-minus (+/-)	New classes (t)		New classes (all)	
Dataset	Model	TV	ROC-AUC	ROC-AUC	PR-AUC	ROC-AUC	PR-AUC
MovieLens	PTM ($\gamma = 0.495$)	**0.601**	**0.685**	**0.800**	**0.269**	**0.790**	**0.134**
	Poisson AR	0.637	0.661	0.740	0.207	0.704	0.090
	DPF, $K = 5$	0.791	0.548	0.736	0.155	0.724	0.072
	NMF, $K = 50$	0.637	0.650	0.782	0.226	0.769	0.108
	Previous obs.	0.618	0.558	0.622	0.179	—	—

Spotify

Takeaways

There are **consistent patterns** explaining how user preferences shift over time.

The proposed **Preference Transition Model (PTM)** provides a **simple** and **interpretable** statistical framework for estimating users’ trajectories.

This could be used to **increase diversity** within recommender systems.

Thanks!
Questions?

Spotify®