

LONG TAIL MUSIC RECOMMENDATION WITH DEEP ARCHITECTURES

ORIOL NIETO

ONIETO@PANDORA.COM

INTERNATIONAL WORKSHOP ON DEEP LEARNING AND MUSIC

INTERNATIONAL JOINT CONFERENCE ON NEURAL NETWORKS

ANCHORAGE, AK, USA

MAY 18, 2017

pandora®

A photograph of Chris Cornell performing live on stage. He has long, curly, light brown hair and a beard. He is wearing a dark, textured jacket over a black shirt. He is holding a dark-colored electric guitar and singing into a microphone. The background is dark and hazy.

RIP CHRIS CORNELL

Outline

- Motivation: The Long Tail
- Background: Collaborative Filtering
- Recommending Music in the Long Tail:
 - Estimate Collaborative Factors from Audio
 - The Music Genome Project™
 - Multimodal Deep Architectures

Outline

- Motivation: The Long Tail
- Background: Collaborative Filtering
- Recommending Music in the Long Tail:
 - Estimate Collaborative Factors from Audio
 - The Music Genome Project™
 - Multimodal Deep Architectures

Now Playing My Stations

Create Station

pandora

Porcupine Tree Radio

Vast Radio

Katatonia Radio

Chaos Chaos Radio

Steven Wilson Radio

See All Stations

The Grudge

Tool

Lyrics

Wear the grudge like a crown of negativity.
Calculate what we will or will not tolerate.

3:20 | 8:36

Thumbs Down Thumbs Up

Search 11:44 AM 100% 🔋

My Stations Browse

Browse Genres View all >

Motown

6.6M listeners

Classic Metal

1.3M listeners

Disco

3.2M listeners

Smooth Jazz

10M listeners

Selenium Forest

Plini

THE ORIGINAL MOTION PICTURE SOUNDTRACK
STAYING ALIVE

NIKS COAST HIGHWAY

mariah carey & tammi terrell Complete Truth

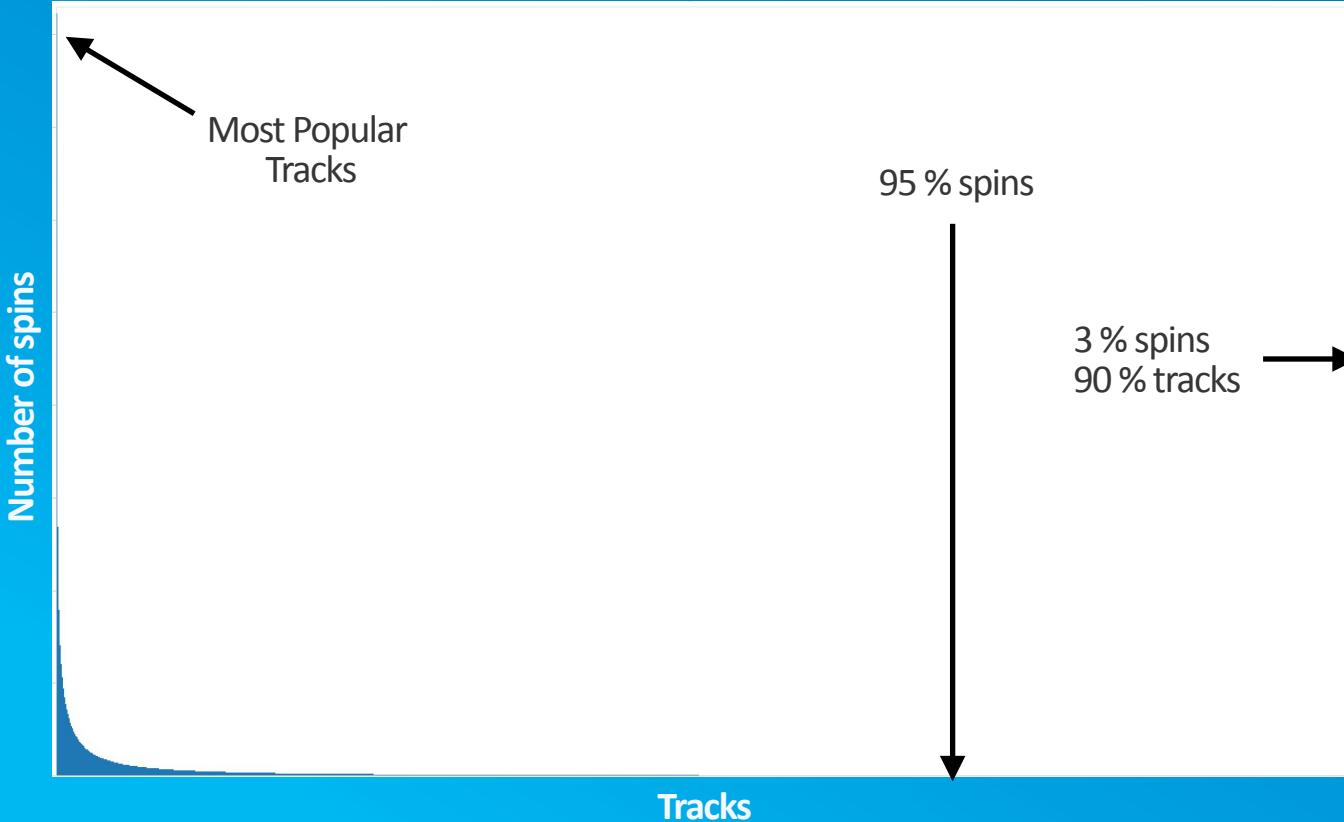
RIDE THE LIGHTNING

NIKS COAST HIGHWAY

pandora®

The Long Tail

(IT'S REALLY LONG)

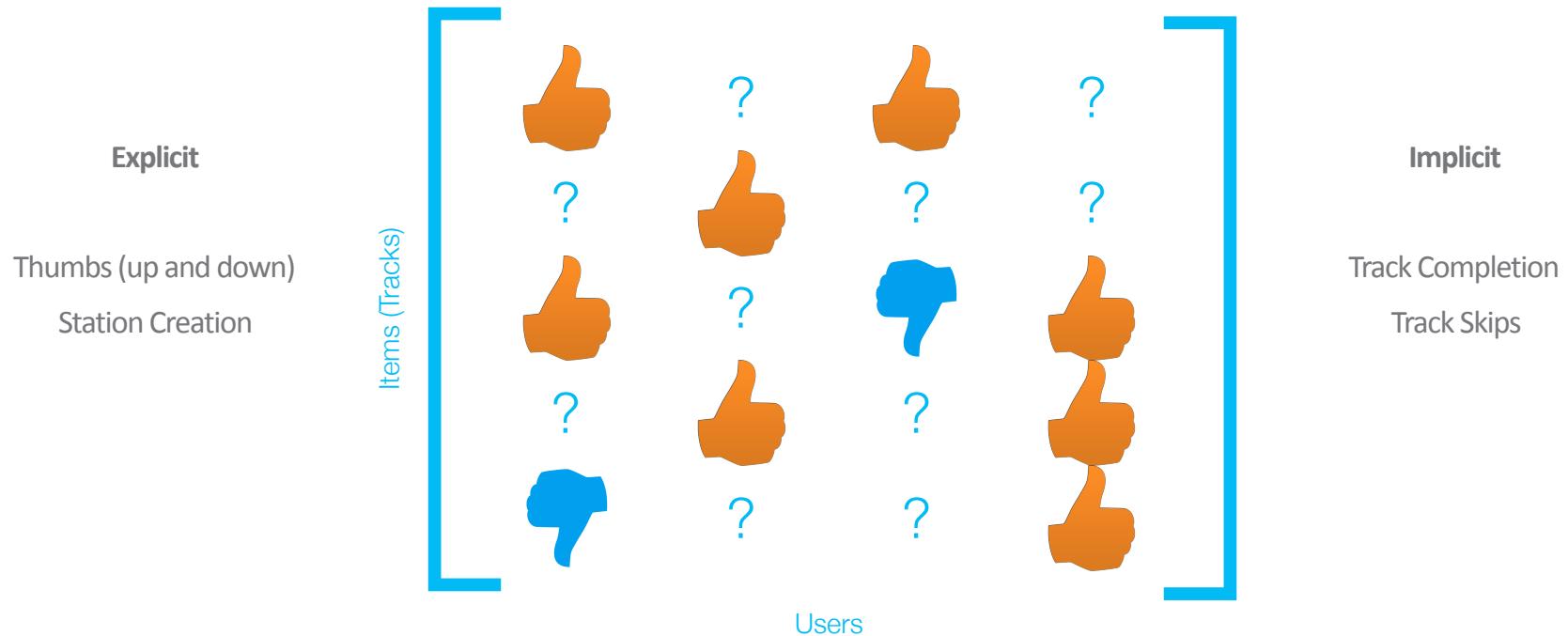


Outline

- Motivation: The Long Tail
- **Background: Collaborative Filtering**
- Recommending Music in the Long Tail:
 - Estimate Collaborative Factors from Audio
 - The Music Genome Project™
 - Multimodal Deep Architectures

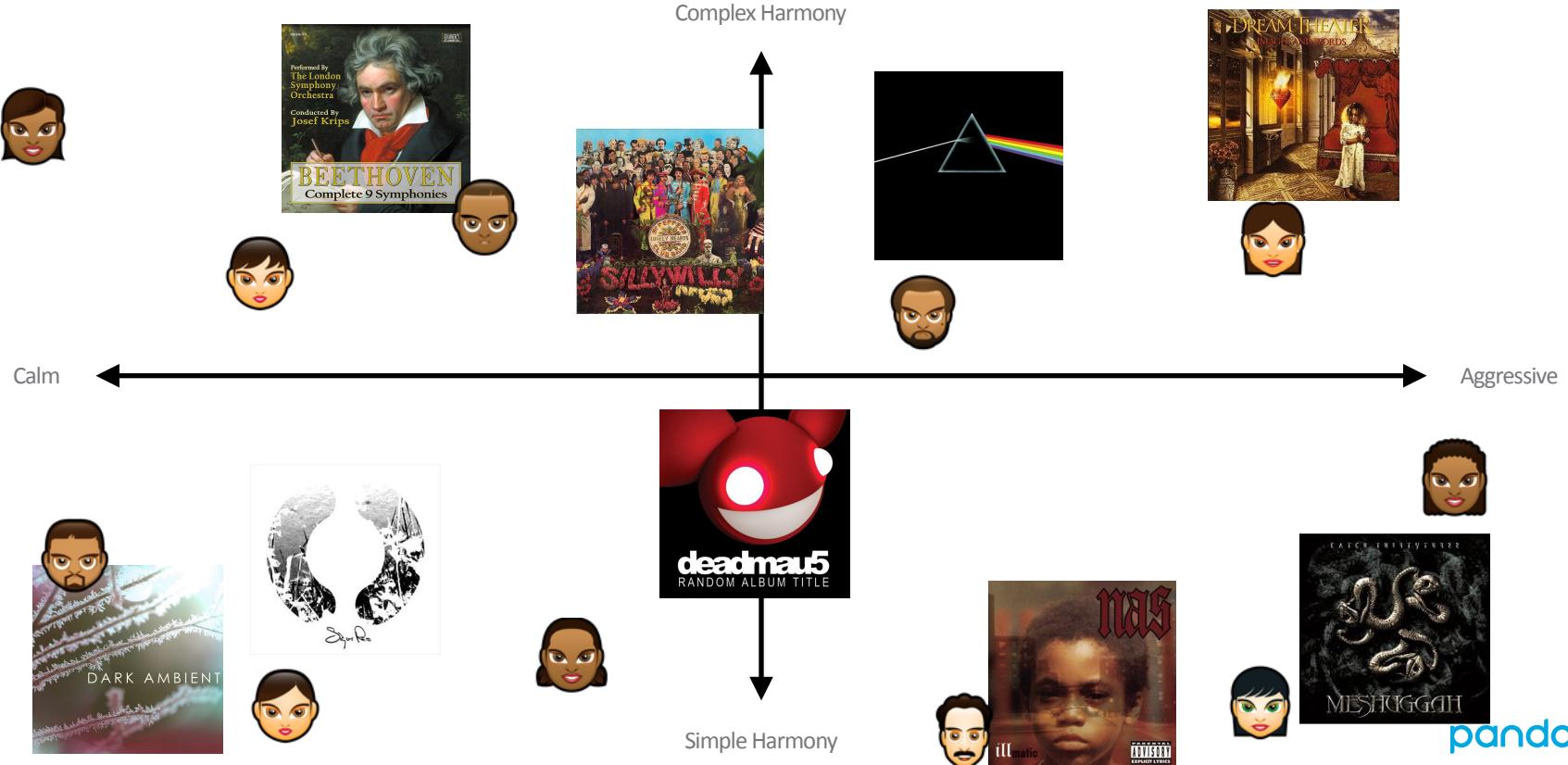
Collaborative Filtering

PROBLEM OVERVIEW



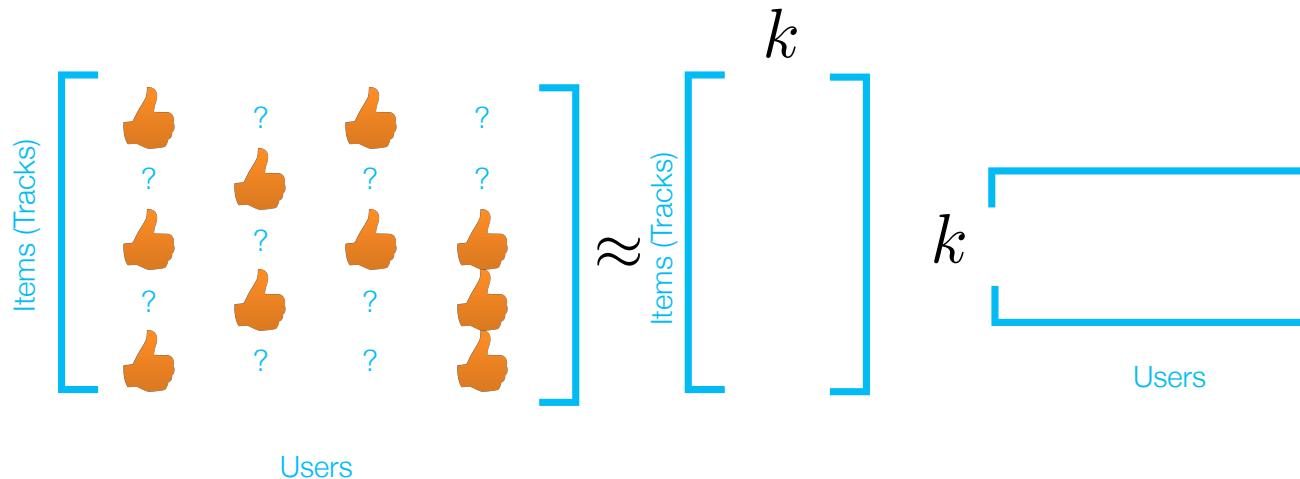
Collaborative Filtering

LATENT FACTORS



Collaborative Filtering

MATRIX FACTORIZATION



Recommend items that are near in the factorized space

Collaborative Filtering

EXAMPLE

	Artist	Title
Query Track	The Beatles	While My Guitar Gently Weeps
Ranked 1	The Beatles	A Day In The Life
Ranked 2	The Beatles	A Day In The Life (Love Version)
Ranked 3	The Beatles	Across The Universe

Collaborative Filtering

EXAMPLE

	Artist	Title
Query Track	The Beatles	While My Guitar Gently Weeps
Ranked 35	George Harrison	While My Guitar Gently Weeps (Live)
Ranked 82	George Harrison	My Sweet Lord (Live)
Ranked 91	Paul McCartney & Eric Clapton	Something (Live)
Ranked 158	Led Zeppelin	Tangerine

Collaborative Filtering

THE GOOD AND THE BAD

Rich preference-driven similarity space

Powerful at matching the right song
with the right listener



Latent space is generally not interpretable

Can only recommend items that
have already been rated

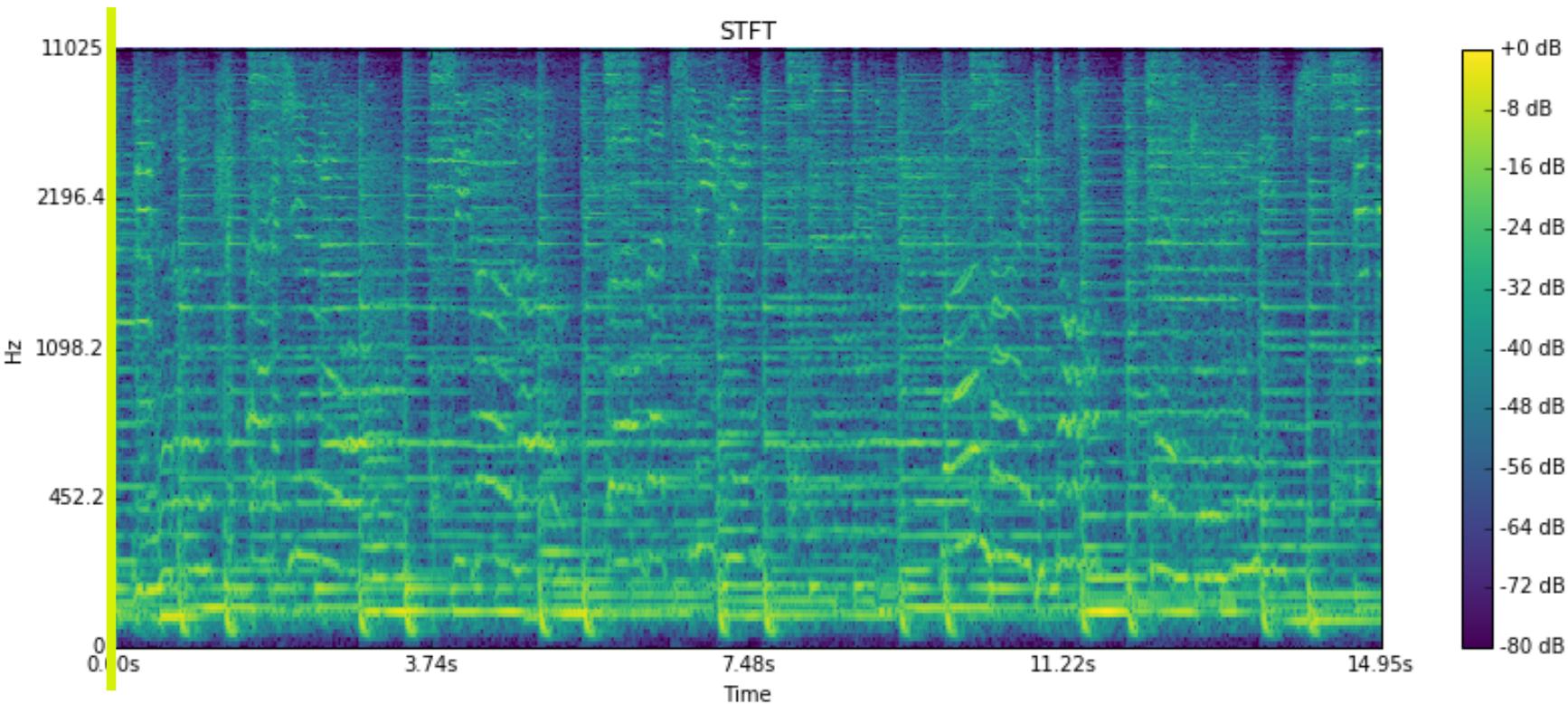


Outline

- Motivation: The Long Tail
- Background: Collaborative Filtering
- **Recommending Music in the Long Tail:**
 - **Estimate Collaborative Factors from Audio**
 - The Music Genome Project™
 - Multimodal Deep Architectures

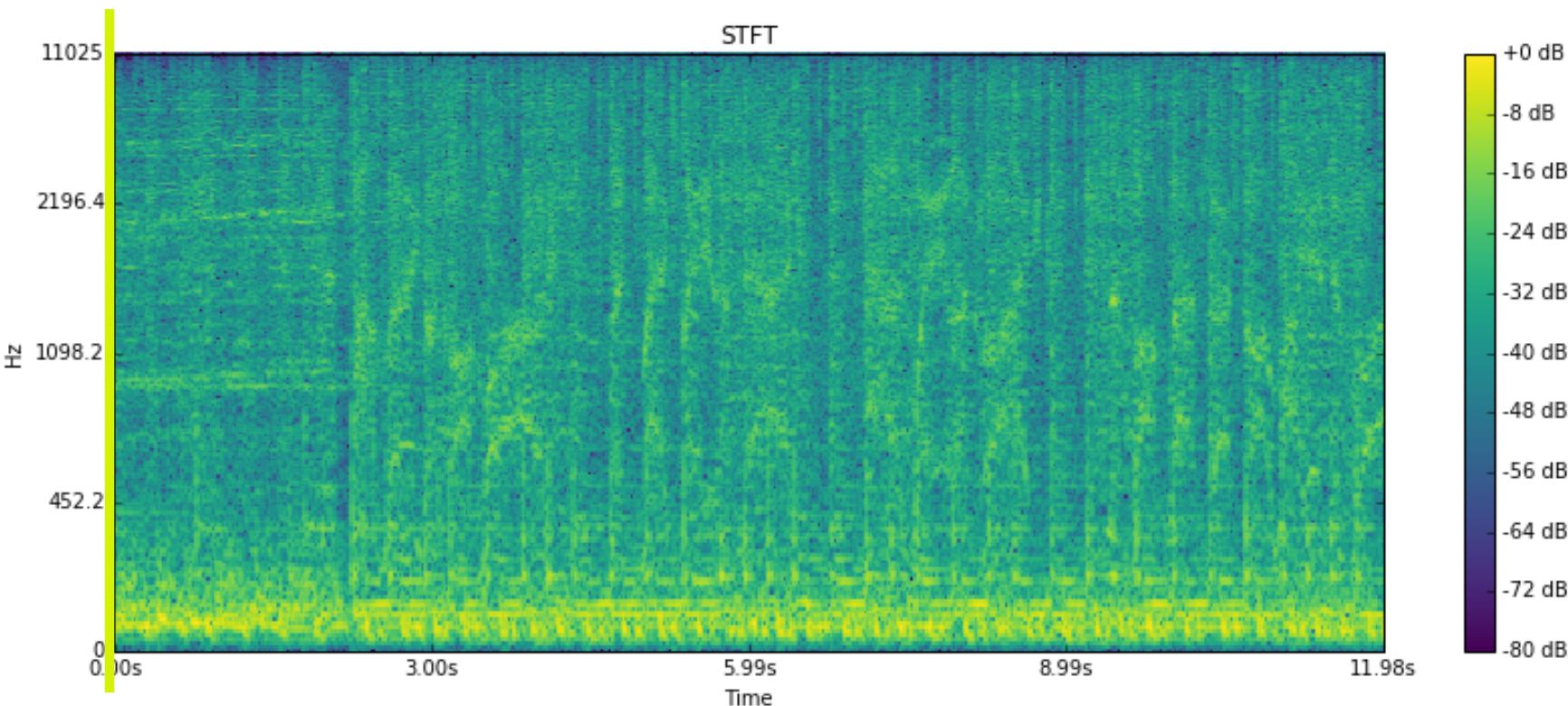
Short Time Fourier Transform

MY GUITAR GENTLY WEEPS - THE BEATLES

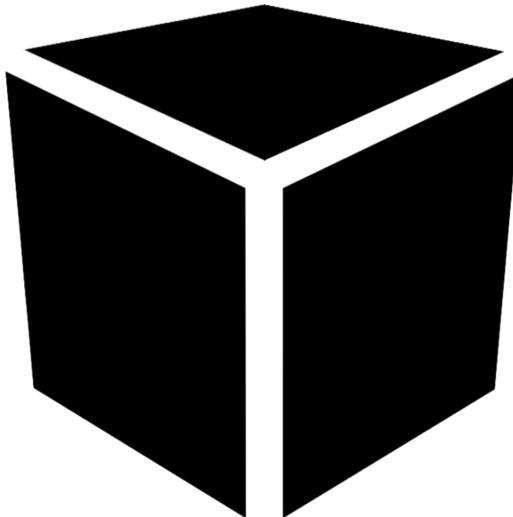
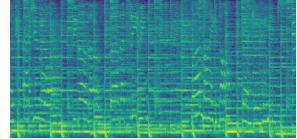


Short Time Fourier Transform

CLOCKWORK - MESHUGGAH



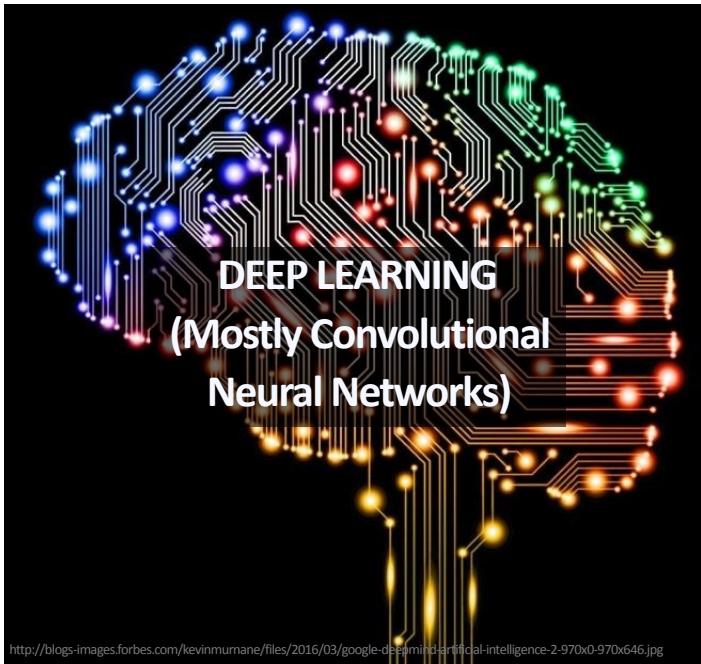
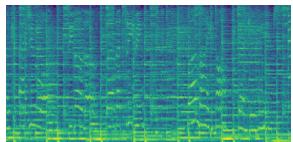
Extracting Information from Audio



Beats
Downbeats
Chords
Structure

“electric guitar”,
“drums”,
“male singer”,
“pop”,
“classic rock”

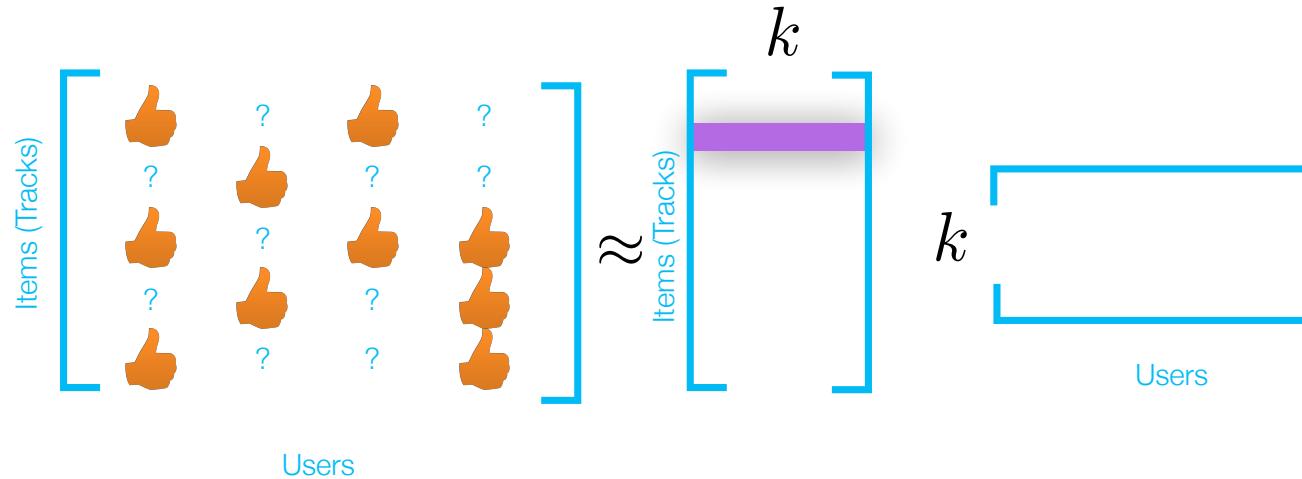
Extracting Information from Audio



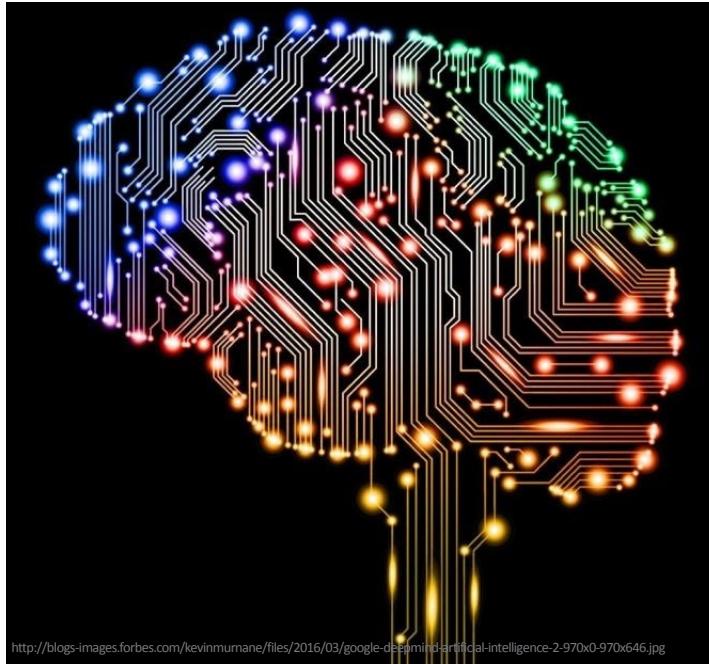
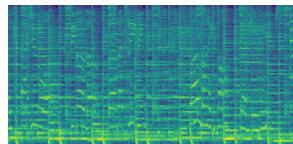
Beats
Downbeats
Chords
Structure

“electric guitar”,
“drums”,
“male singer”,
“pop”,
“classic rock”

Back to Collaborative Filtering



Approximate Item Factors using Audio

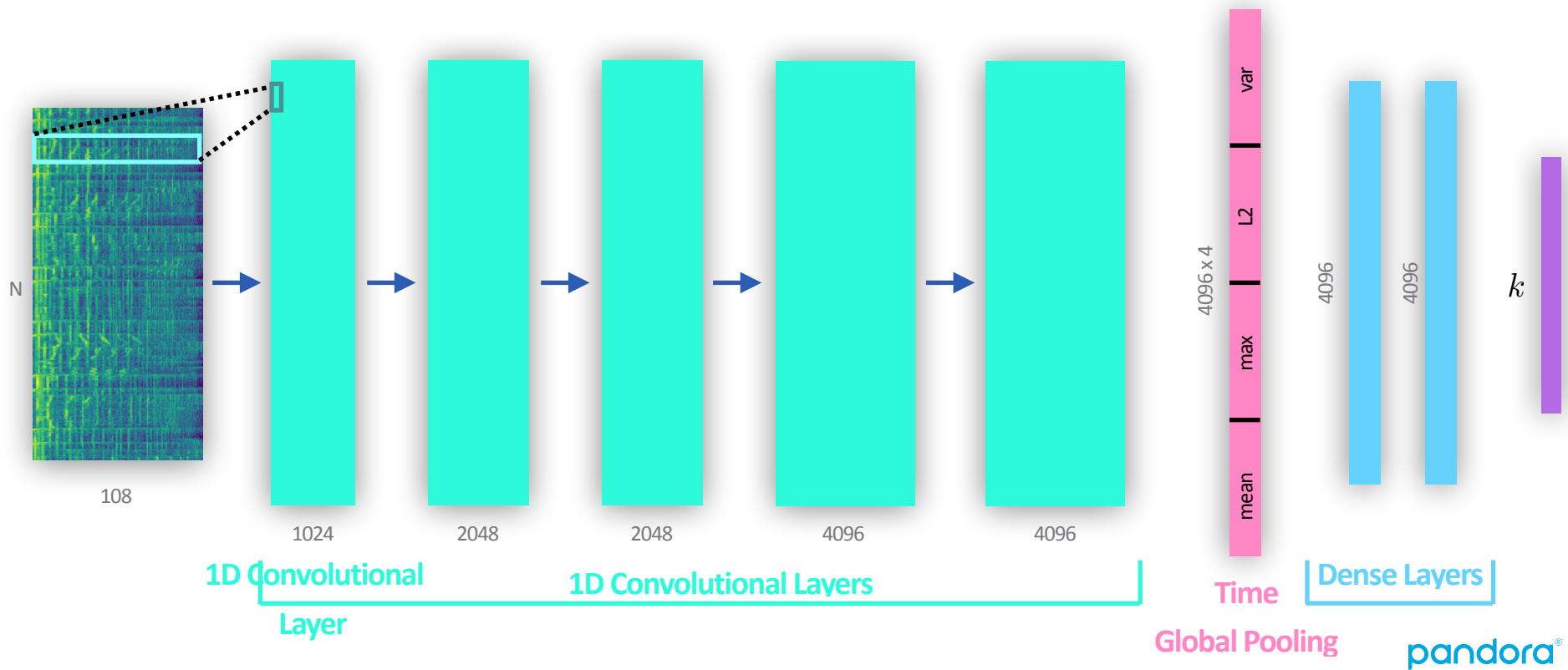


k

Oord, A. Van Den, Dieleman, S., & Schrauwen, B. (2013). Deep Content-based Music Recommendation. Advances in Neural Information Processing Systems, 2643–2651.

pandora®

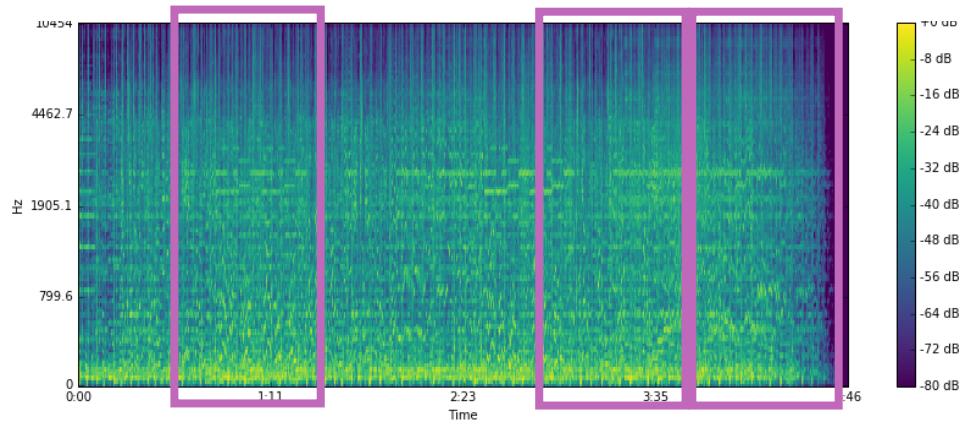
Approximating Factors using Audio WITH DEEP LEARNING



Approximating Item Factors using Audio

TRAINING DATA

- (Small) Data set $\{\mathbb{X}, \mathbb{Y}\}$:
 - 83k tracks
 - 3 patches of 35 seconds per track (251k patches = M)
 - (Patches only for training!)
 - Splits:
 - Train: 80%
 - Validation: 10%
 - Test: 10%

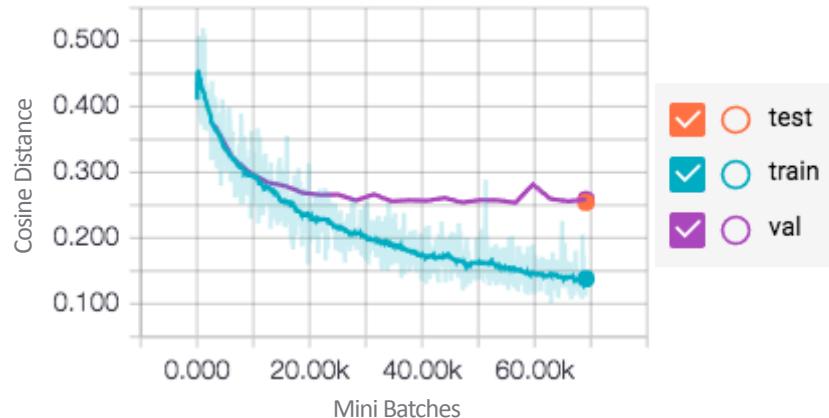


Approximating Item Factors

TRAINING

- Loss function:
 - Cosine Distance
- Optimization:
 - Adam (default params)
 - 50% Dropout on Dense Layers
 - Early Stopping
 - Mini-batches of 64 examples

$$\mathcal{L}(\theta) = 1 - \frac{1}{M} \sum_{X \in \mathbb{X}, \mathbf{y} \in \mathbb{Y}} \frac{f(X; \theta)^T \mathbf{y}}{\|f(X; \theta)\|_2 \|\mathbf{y}\|_2}$$



Approximating Item Factors using Audio

RESULTS

Input	Cos Distance	# Epochs	Time / Epoch
Audio (35s Patches)	0.25	22	~2h

Approximating Item Factors using Audio

RESULTS

Input	Cos Distance	# Epochs	Time / Epoch
Audio (35s Patches)	0.25	22	~2h
Audio (Full Tracks)	0.21	-	-

Outline

- Motivation: The Long Tail
- Background: Collaborative Filtering
- **Recommending Music in the Long Tail:**
 - Estimate Collaborative Factors from Audio
 - **The Music Genome Project™**
 - Multimodal Deep Architectures

The Music Genome Project™



Attribute Examples

Breathy Voice

Nasal Voice

Odd Meter

Has Banjo

Joyful Lyrics

...

>1.5 Million tracks manually analyzed

~400 attributes per track

Recommending Music using the MGP™

EXAMPLE

	Artist	Title
Query Track	The Beatles	While My Guitar Gently Weeps
Ranked 1	IV Thieves	The Sound And The Fury
Ranked 2	Journey	Too Late
Ranked 3	Albert Lee	Look Out Cleveland

Recommending Music using the MGP™

EXAMPLE

	Artist	Title
Query Track	The Beatles	While My Guitar Gently Weeps
Ranked 1	IV Thieves	The Sound And The Fury
Ranked 2	Journey	Too Late
Ranked 3	Albert Lee	Look Out Cleveland

Recommending Music using the MGP™

EXAMPLE

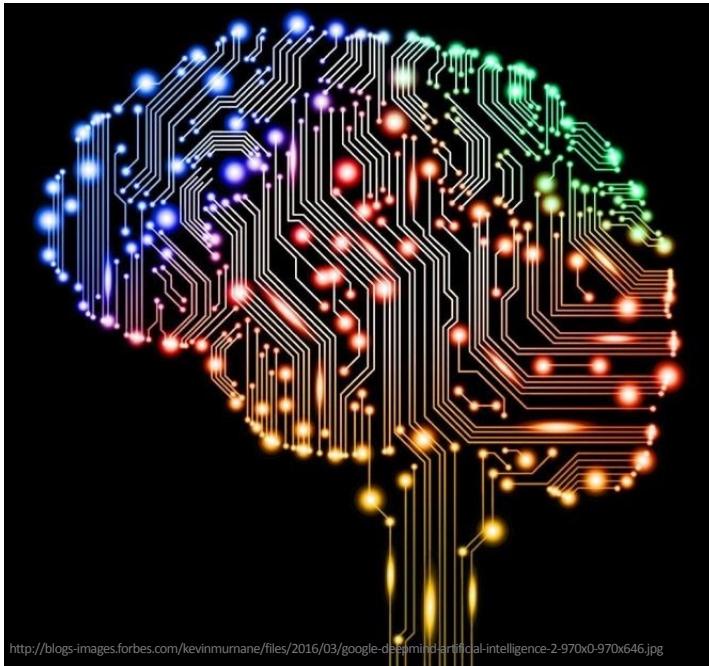
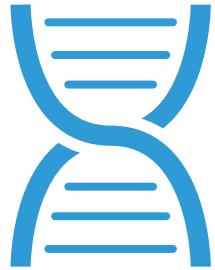
	Artist	Title
Query Track	The Beatles	While My Guitar Gently Weeps
Ranked 1	IV Thieves	The Sound And The Fury
Ranked 2	Journey	Too Late
Ranked 3	Albert Lee	Look Out Cleveland

Recommending Music using the MGP™

EXAMPLE

	Artist	Title
Query Track	The Beatles	While My Guitar Gently Weeps
Ranked 1	IV Thieves	The Sound And The Fury
Ranked 2	Journey	Too Late
Ranked 3	Albert Lee	Look Out Cleveland

Approximate Factors using the MGP™

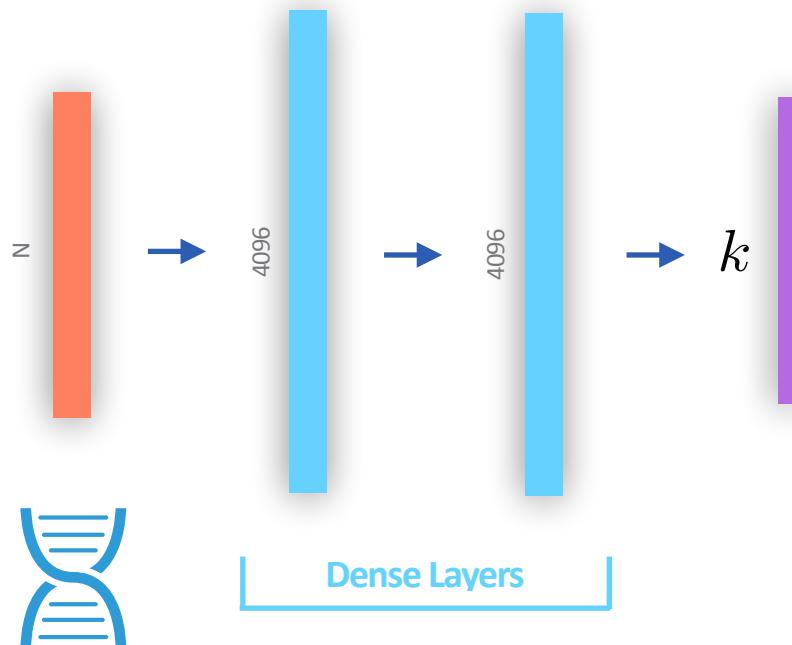


k



Approximate Factors using the MGP

DEEP ARCHITECTURE



Approximating Item Factors using the MGP™

TRAINING DATA

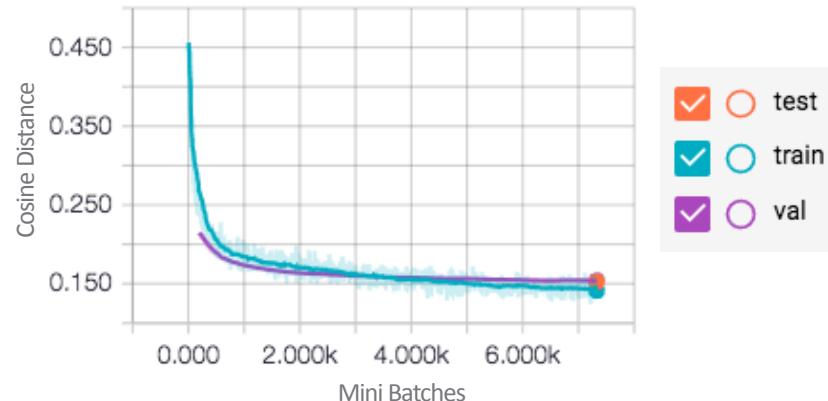
- (Small) Data set $\{\mathbb{X}, \mathbb{Y}\}$:
 - 83k tracks (M)
 - Splits:
 - Train: 80%
 - Validation: 10%
 - Test: 10%

Approximating Item Factors using the MGP™

TRAINING

- Loss function:
 - Cosine Distance
- Optimization:
 - Adam (default params)
 - 50% Dropout on Dense Layers
 - Early Stopping
 - Mini-batches of 256 examples

$$\mathcal{L}(\theta) = 1 - \frac{1}{M} \sum_{\mathbf{x} \in \mathbb{X}, \mathbf{y} \in \mathbb{Y}} \frac{f(\mathbf{x}; \theta)^T \mathbf{y}}{\|f(\mathbf{x}; \theta)\|_2 \|\mathbf{y}\|_2}$$



Approximating Item Factors

RESULTS

Input	Cos Distance	# Epochs	Time / Epoch
Audio (35s Patches)	0.25	22	~2h
Audio (Full Tracks)	0.21	-	-
MGP	0.15	37	7s

Beyond the MGP™



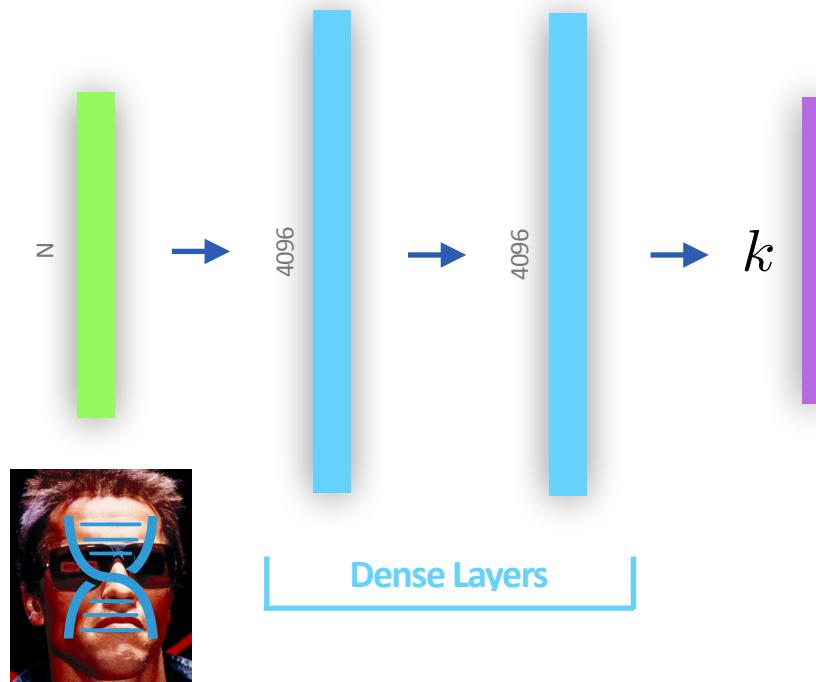
MACHINE LISTENING GENES

APPROXIMATE MGP WITH MACHINE LISTENING

pandora®

Approximate Factors using MLG

DEEP ARCHITECTURE



Approximating Item Factors

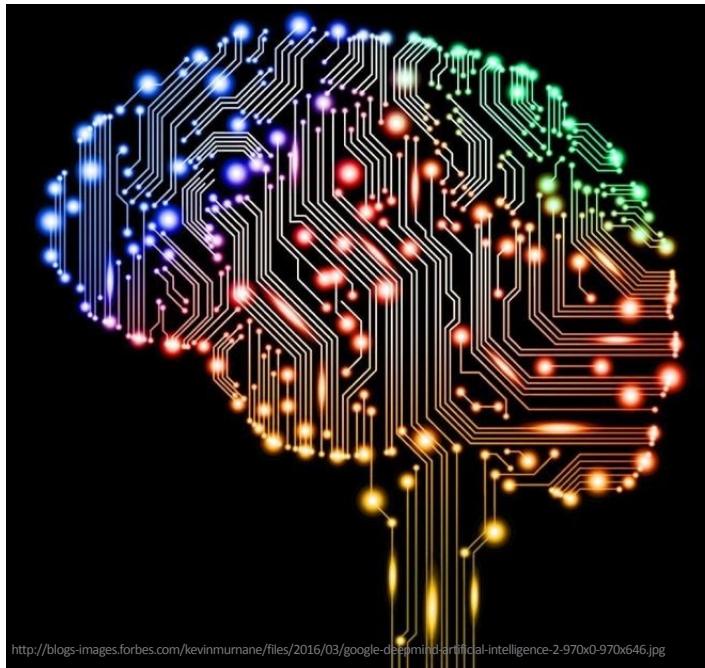
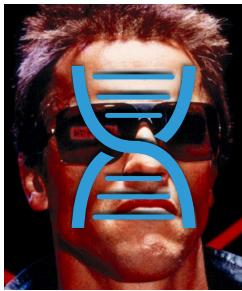
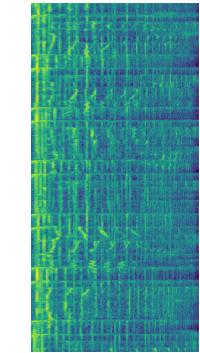
RESULTS

Input	Cos Distance	# Epochs	Time / Epoch
Audio (35s Patches)	0.25	22	~2h
Audio (Full Tracks)	0.21	-	-
MGP	0.15	37	7s
MLG	0.22	37	7s

Outline

- Motivation: The Long Tail
- Background: Collaborative Filtering
- **Recommending Music in the Long Tail:**
 - Estimate Collaborative Factors from Audio
 - The Music Genome Project™
 - **Multimodal Deep Architectures**

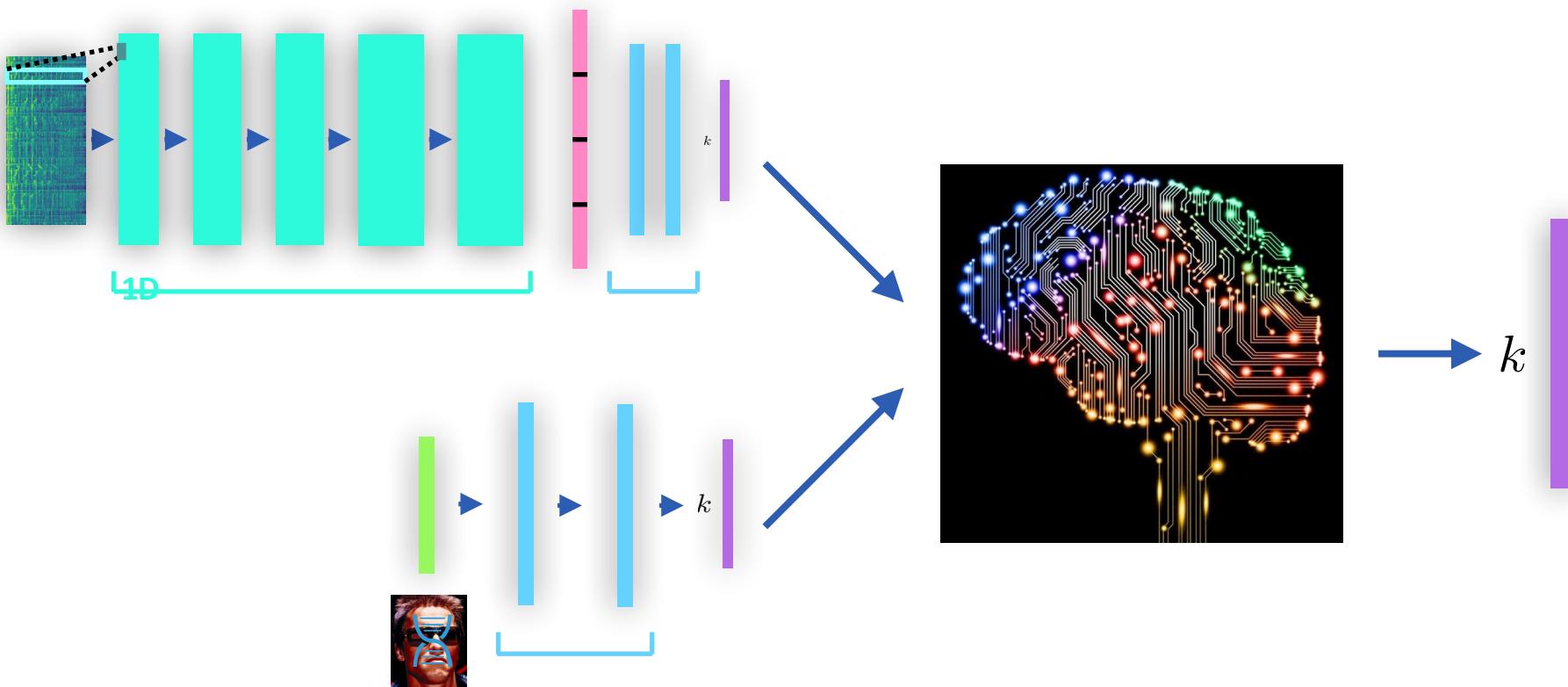
Combine Methods to Approximate Factors



k

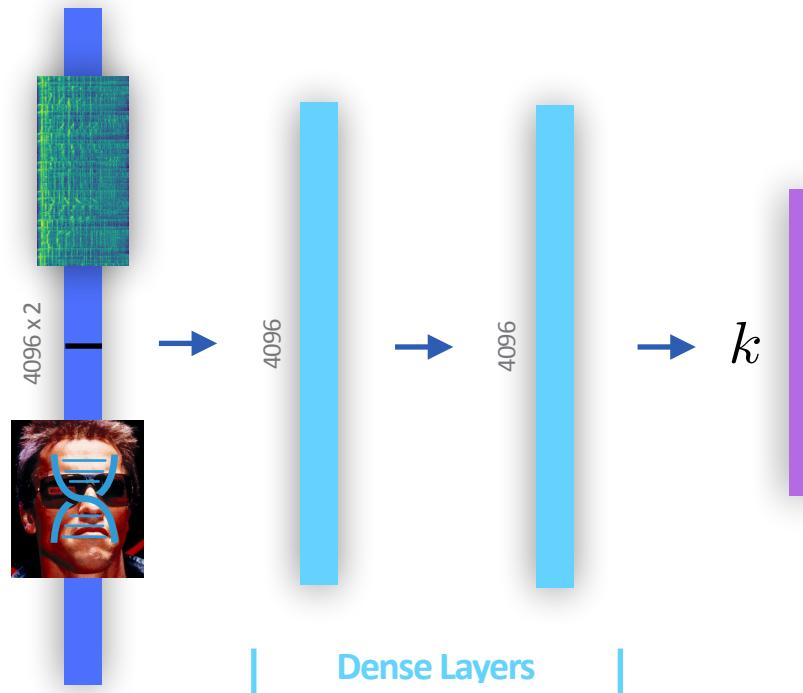


Combine Methods to Approximate Factors



Combine Methods to Approximate Factors

LATE-FUSION DEEP ARCHITECTURE

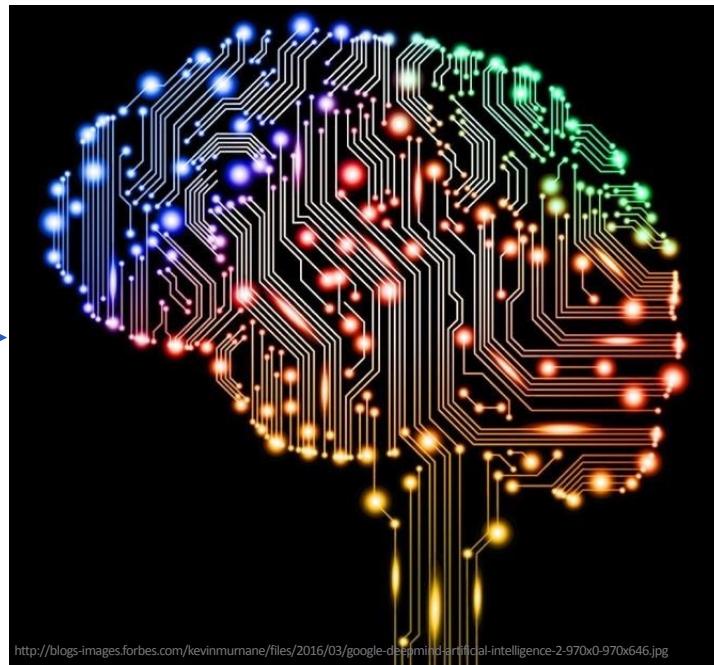
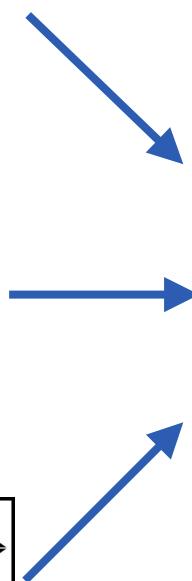
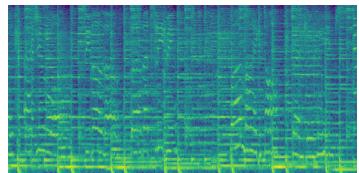


Approximating Item Factors

RESULTS

Input	Cos Distance	# Epochs	Time / Epoch
Audio (35s Patches)	0.25	22	~2h
Audio (Full Tracks)	0.21	-	-
MGP	0.15	37	7s
MLG	0.22	37	7s
Audio + MLG	0.19	37	7s

Further Multimodality to Approximate Factors



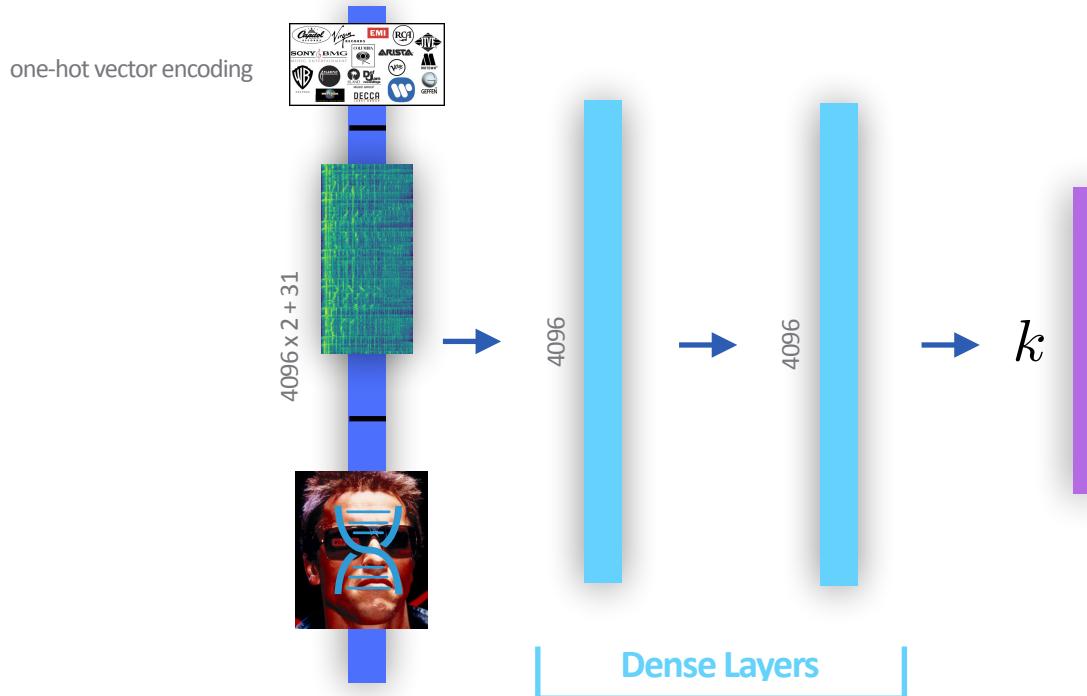
k



pandora®

Further Multimodality to Approximate Factors

LATE-FUSION DEEP ARCHITECTURE



Approximating Item Factors

RESULTS

Input	Cos Distance	# Epochs	Time / Epoch
Audio (35s Patches)	0.25	22	~2h
Audio (Full Tracks)	0.21	-	-
MGP	0.15	37	7s
MLG	0.22	37	7s
Audio + MLG	0.19	37	7s
Audio + MLG + genres	0.16	37	7s

Recommendation Examples

	Artist	Title
Query Track	The Beatles	While My Guitar Gently Weeps
Ranked 1	Bob Dylan	Knockin' On Heavens Door
Ranked 2	Neil Young	Heart Of Gold
Ranked 3	The Rolling Stones	Angie

Recommendation Examples

	Artist	Title
Query Track	The Beatles	While My Guitar Gently Weeps
Ranked 1	Bob Dylan	Knockin' On Heavens Door
Ranked 2	Neil Young	Heart Of Gold
Ranked 3	The Rolling Stones	Angie

Recommendation Examples

	Artist	Title
Query Track	The Beatles	While My Guitar Gently Weeps
Ranked 1	Bob Dylan	Knockin' On Heavens Door
Ranked 2	Neil Young	Heart Of Gold
Ranked 3	The Rolling Stones	Angie

Recommendation Examples

	Artist	Title
Query Track	The Beatles	While My Guitar Gently Weeps
Ranked 1	Bob Dylan	Knockin' On Heavens Door
Ranked 2	Neil Young	Heart Of Gold
Ranked 3	The Rolling Stones	Angie

Recommendation Examples

	Artist	Title
Query Track	Sargon	Continuarà
Ranked 1	Mudvayne	Happy?
Ranked 2	Mudvayne	Forget To Remember
Ranked 3	Stone Sour	Hell & Consequences

Recommendation Examples

	Artist	Title
Query Track	Sargon	Continuarà
Ranked 1	Mudvayne	Happy?
Ranked 2	Mudvayne	Forget To Remember
Ranked 3	Stone Sour	Hell & Consequences

Recommendation Examples

	Artist	Title
Query Track	Sargon	Continuarà
Ranked 1	Mudvayne	Happy?
Ranked 2	Mudvayne	Forget To Remember
Ranked 3	Stone Sour	Hell & Consequences

Recommendation Examples

	Artist	Title
Query Track	Sargon	Continuarà
Ranked 1	Mudvayne	Happy?
Ranked 2	Mudvayne	Forget To Remember
Ranked 3	Stone Sour	Hell & Consequences

Recommendation Examples

	Artist	Title
Query Track	La Bossa d'Urina	El Tiempo
Ranked 1	Il Divo	Hallelujah
Ranked 2	Sarah Brightman & The London Symphony Orchestra	Time To Say Goodbye
Ranked 3	Andrea Bocelli	Amapola

Recommendation Examples

	Artist	Title
Query Track	La Bossa d'Urina	El Tiempo
Ranked 1	Il Divo	Hallelujah
Ranked 2	Sarah Brightman & The London Symphony Orchestra	Time To Say Goodbye
Ranked 3	Andrea Bocelli	Amapola

Recommendation Examples

	Artist	Title
Query Track	La Bossa d'Urina	El Tiempo
Ranked 1	Il Divo	Hallelujah
Ranked 2	Sarah Brightman & The London Symphony Orchestra	Time To Say Goodbye
Ranked 3	Andrea Bocelli	Amapola

Recommendation Examples

	Artist	Title
Query Track	La Bossa d'Urina	El Tiempo
Ranked 1	Il Divo	Hallelujah
Ranked 2	Sarah Brightman & The London Symphony Orchestra	Time To Say Goodbye
Ranked 3	Andrea Bocelli	Amapola



**ENSEMBLE OF RECOMMENDERS MAY PRODUCE
OPTIMAL RECOMMENDATIONS**



MAN vs MACHINE?



MAN + MACHINE



MAN + MACHINE
“Mix of Art and Science”

THANKS!

ONIETO@PANDORA.COM

pandora®