

DEEP LEARNING FOR LARGE-SCALE MUSIC RECOMMENDATION

ORIOL NIETO

ONIETO@PANDORA.COM

MUSIC 364 - DATA-DRIVEN RESEARCH IN MUSIC COGNITION

CCRMA, STANFORD UNIVERSITY

STANFORD, CA

FEBRUARY 24, 2017

pandora®

Outline

- Pandora Overview
- Large Scale Music Recommendation with Deep Learning:
 - Machine Listening
 - Collaborative Filtering

Outline

- **Pandora Overview**
- Large Scale Music Recommendation with Deep Learning:
 - Machine Listening
 - Collaborative Filtering

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, Refresh, Pause, Next, 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

pandora®

TODAY



78 M
Monthly Listeners



10 B
Stations



22 H
Average Listening per Month



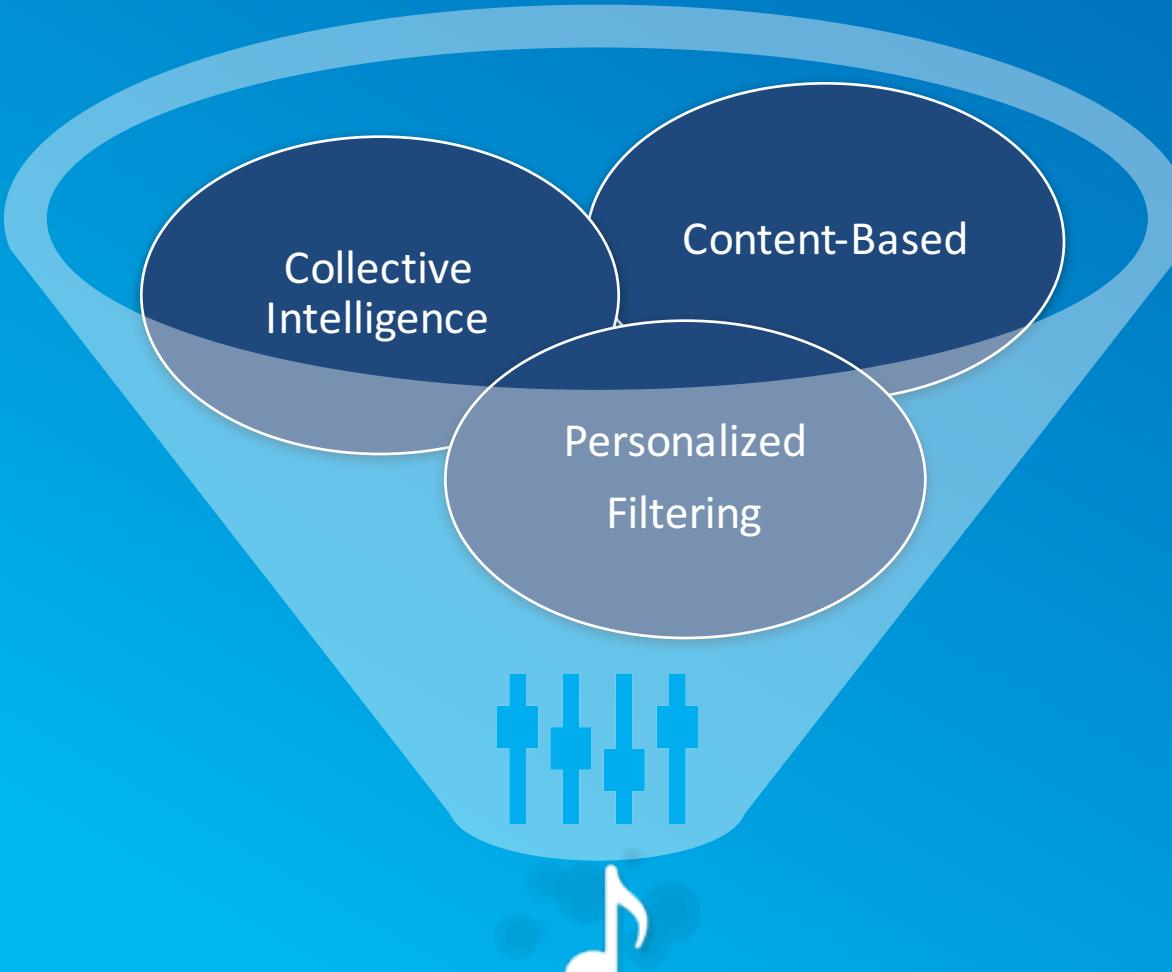
200 K +
Artists Spinning per Month



75 B +
Thumbs



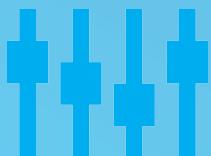
55 %
of all Music Streaming in USA



Collective
Intelligence

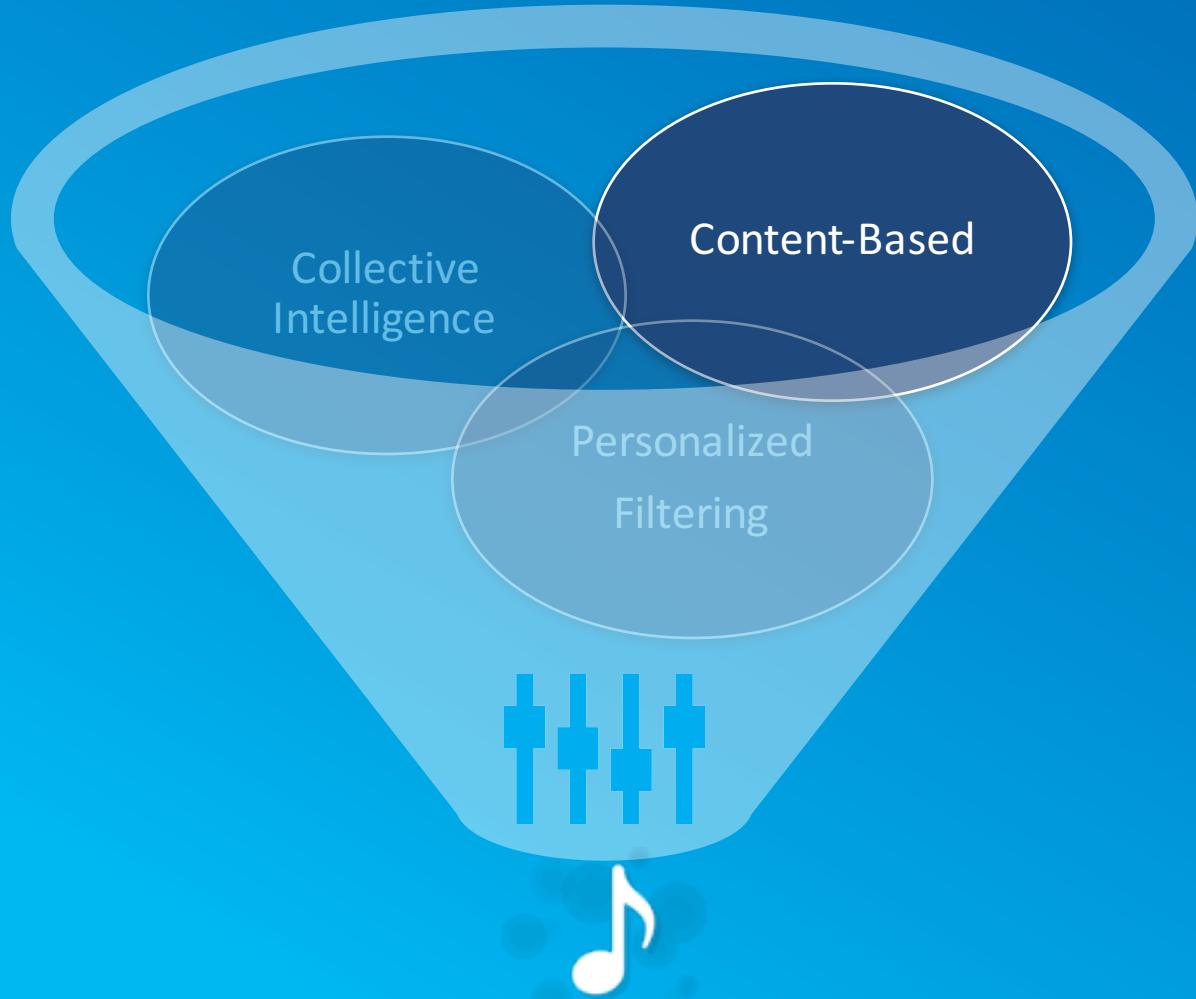
Content-Based

Personalized
Filtering



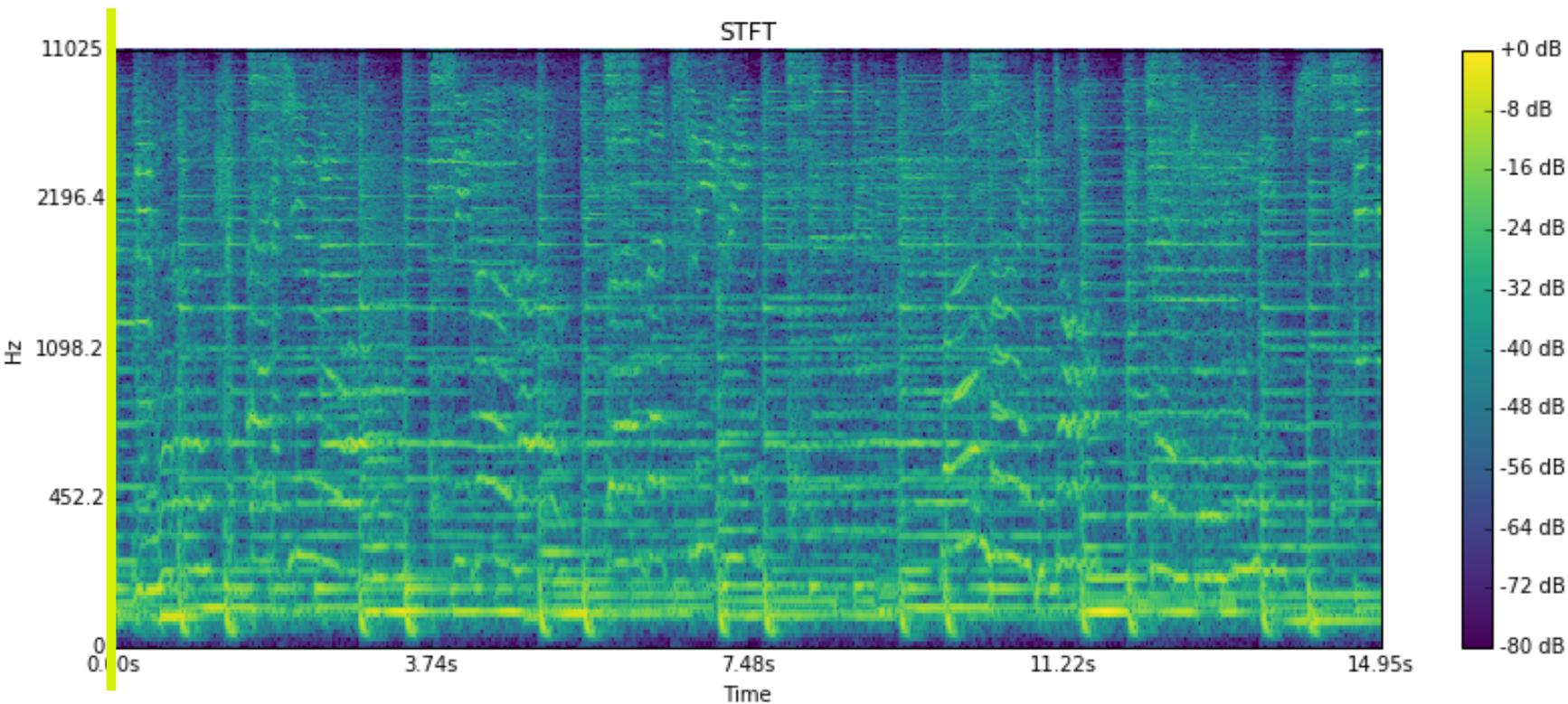
Outline

- Pandora Overview
- Large Scale Music Recommendation with Deep Learning:
 - Machine Listening
 - Collaborative Filtering



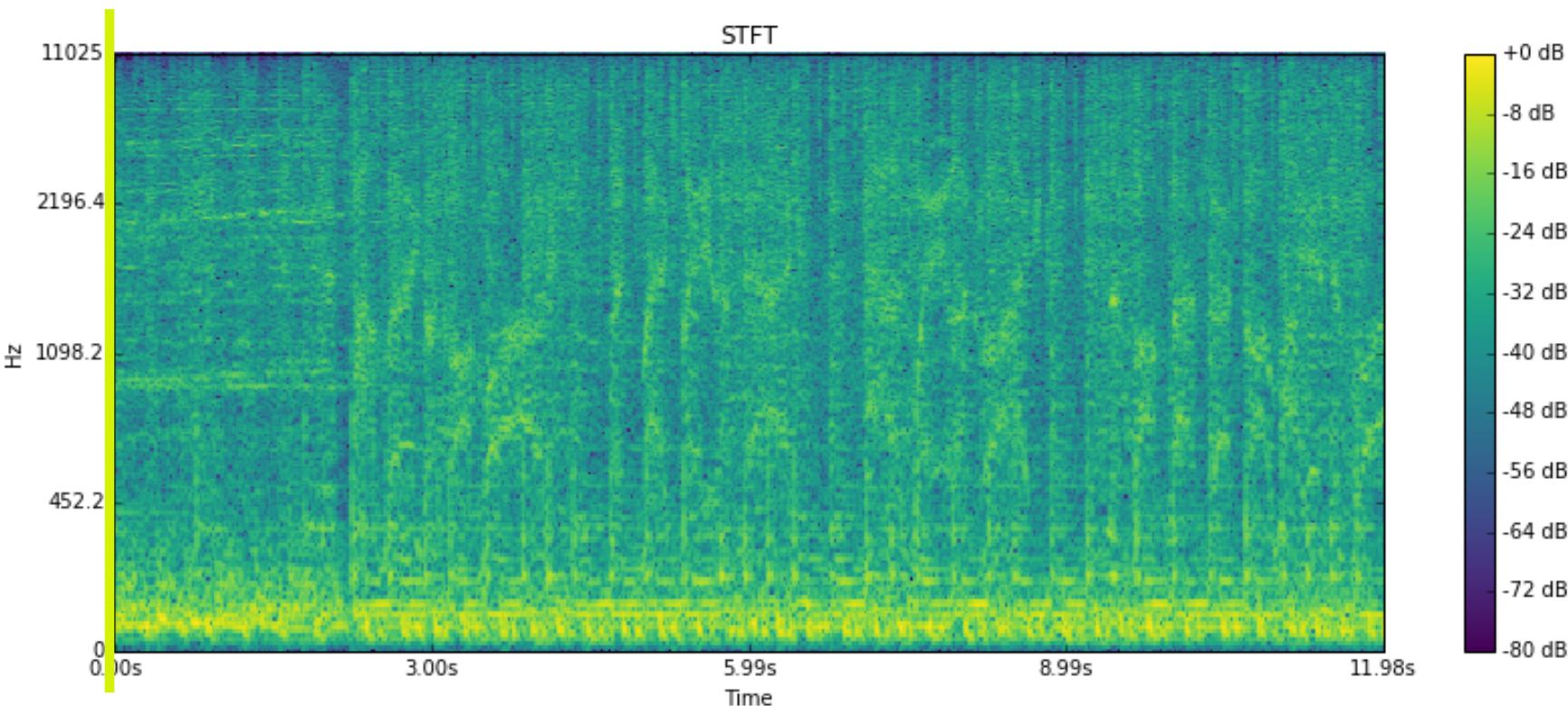
Short Time Fourier Transform

MY GUITAR GENTLY WEEPS - THE BEATLES



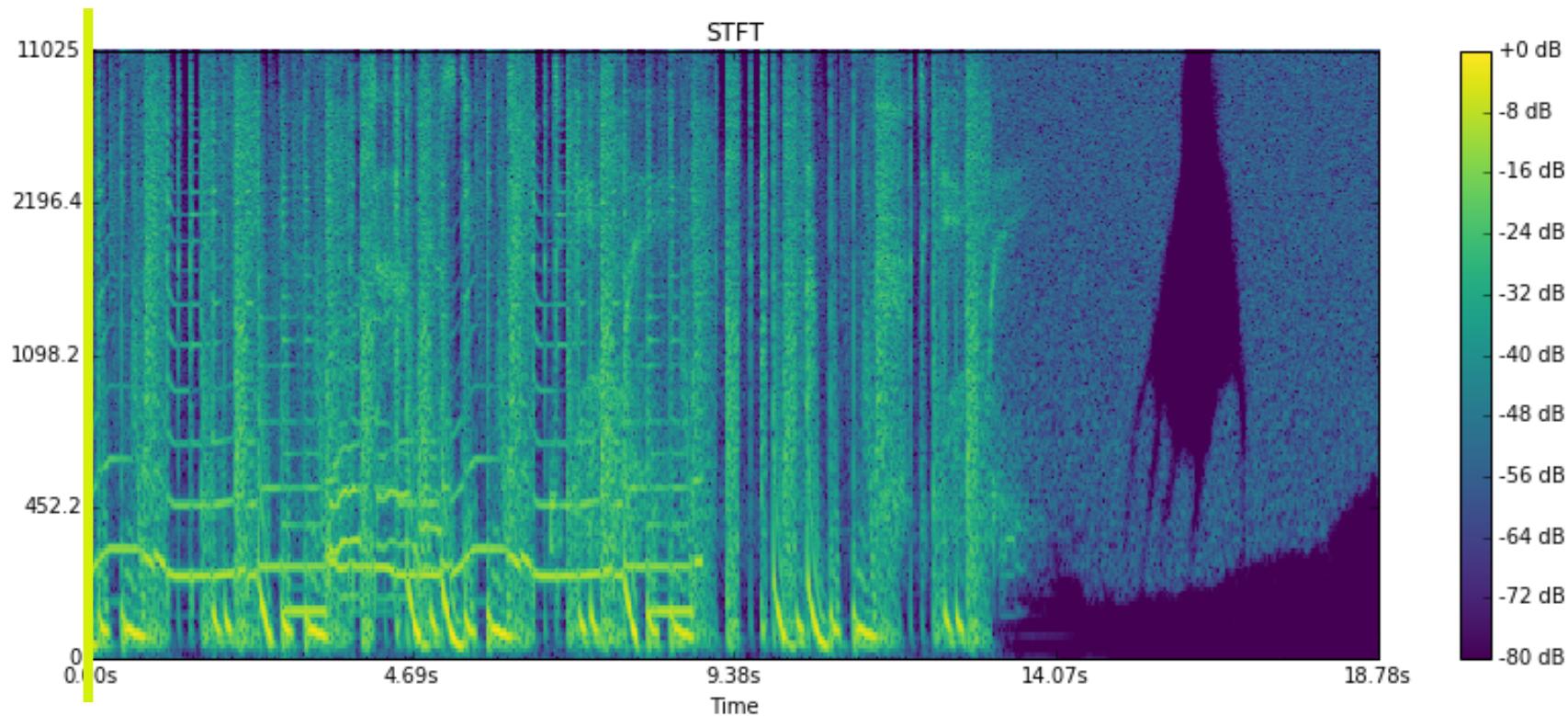
Short Time Fourier Transform

CLOCKWORK - MESHUGGAH

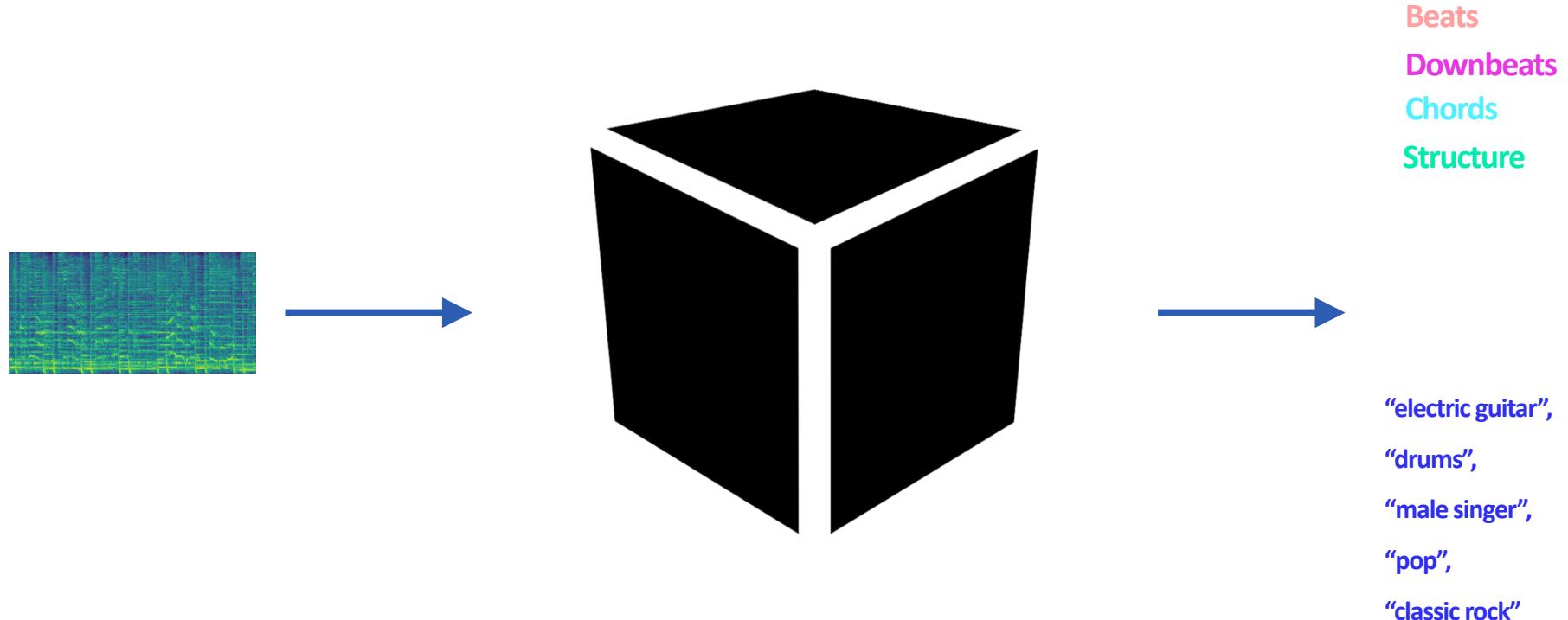


Short Time Fourier Transform

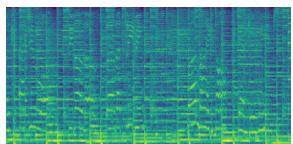
THE WARNING - NINE INCH NAILS



Extracting Information



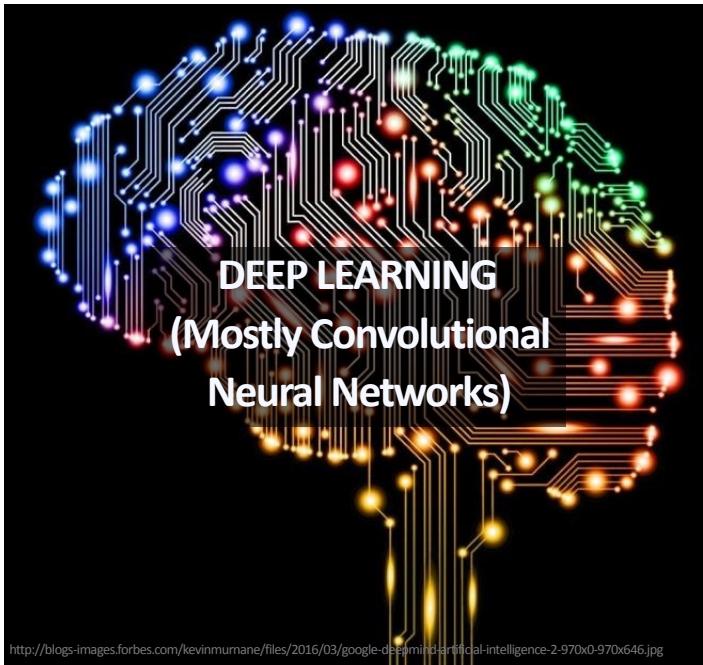
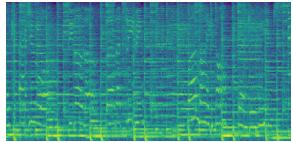
Hand Crafted Features



Beats
Downbeats
Chords
Structure

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

Learned Features



Beats
Downbeats
Chords
Structure

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

The Music *Genome* Project



Attribute Examples

Breathy Voice

Nasal Voice

Odd Meter

Has Banjo

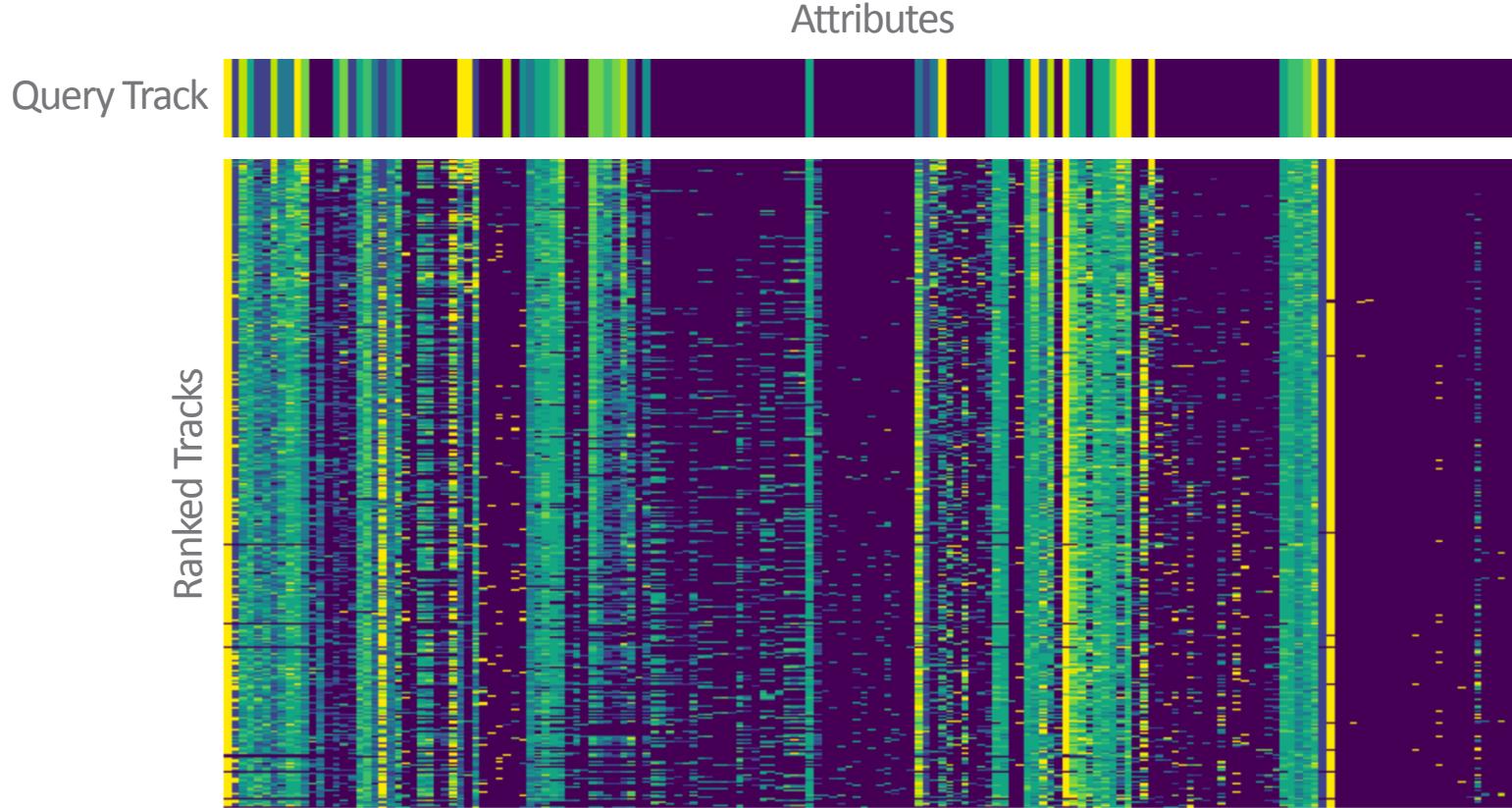
Joyful Lyrics

...

>1.5 Million tracks manually analyzed

~400 attributes per track

Recommending Music using the MGP



Recommending Music using the MGP

EXAMPLE

	Artist	Title
Query Track	Meshuggah	Bleed
Ranked 1	Meshuggah	Lethargica
Ranked 2	Meshuggah	Pravus
Ranked 5	Mithras	Into Black Holes Of Oblivion

Recommending Music using the MGP

EXAMPLE

	Artist	Title
Query Track	The Beatles	While My Guitar Gently Weeps
Ranked 1	Paul McCartney	Freedom
Ranked 2	Badly Drawn Boy	What Is It Now
Ranked 3	Jefferson Starship	Fading Lady Light

Recommending Music using the MGP

EXAMPLE

	Artist	Title
Query Track	The Beatles	While My Guitar Gently Weeps
Ranked 1	Paul McCartney	Freedom
Ranked 2	Badly Drawn Boy	What Is It Now
Ranked 3	Jefferson Starship	Fading Lady Light

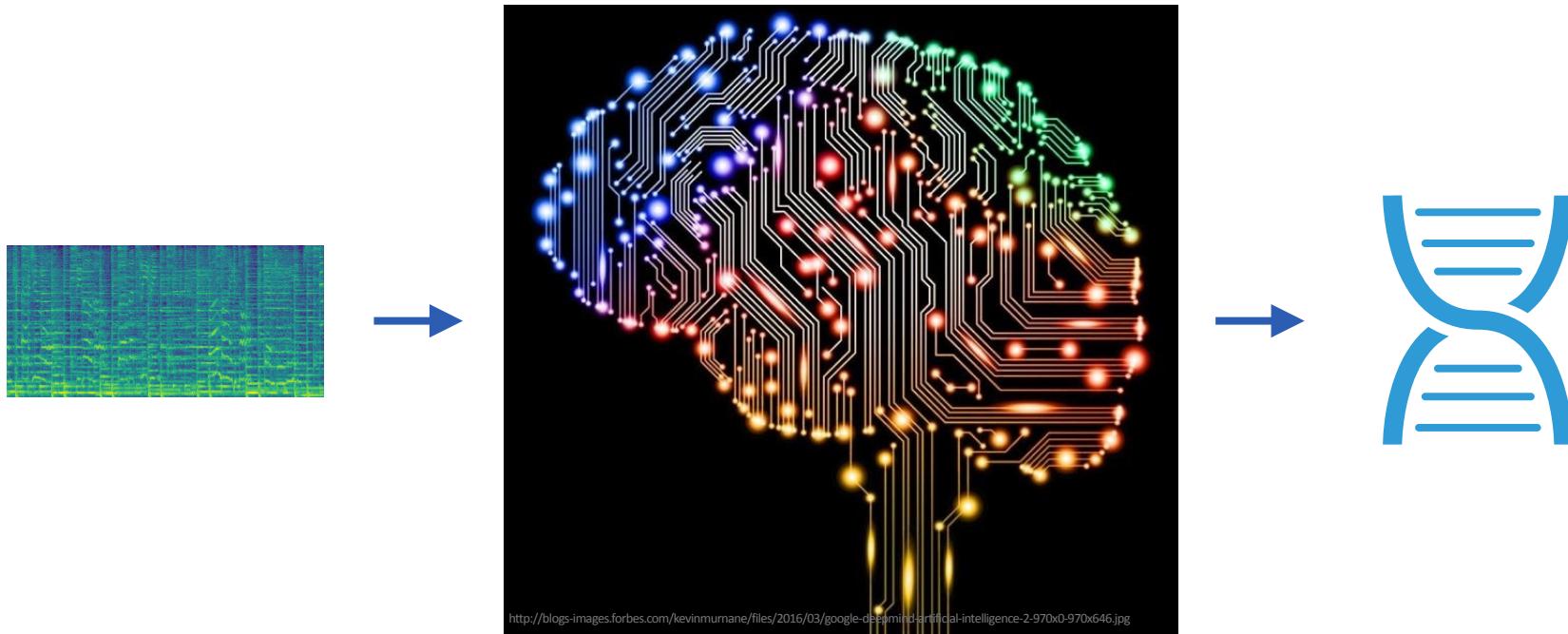
Recommending Music using the MGP

EXAMPLE

	Artist	Title
Query Track	The Beatles	While My Guitar Gently Weeps
Ranked 1	Paul McCartney	Freedom
Ranked 2	Badly Drawn Boy	What Is It Now
Ranked 3	Jefferson Starship	Fading Lady Light

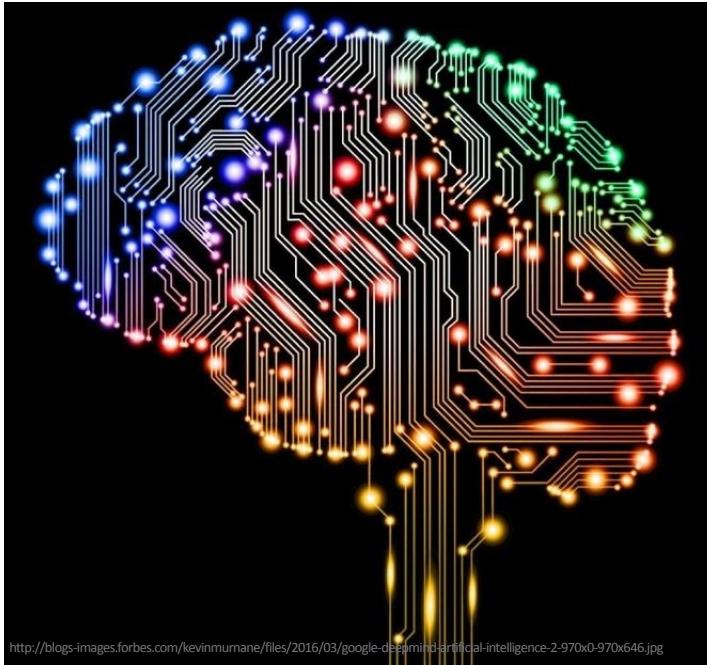
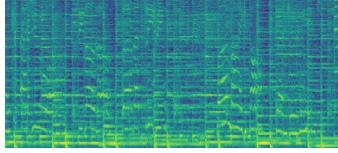
Estimating MGP Attributes

USING DEEP LEARNING



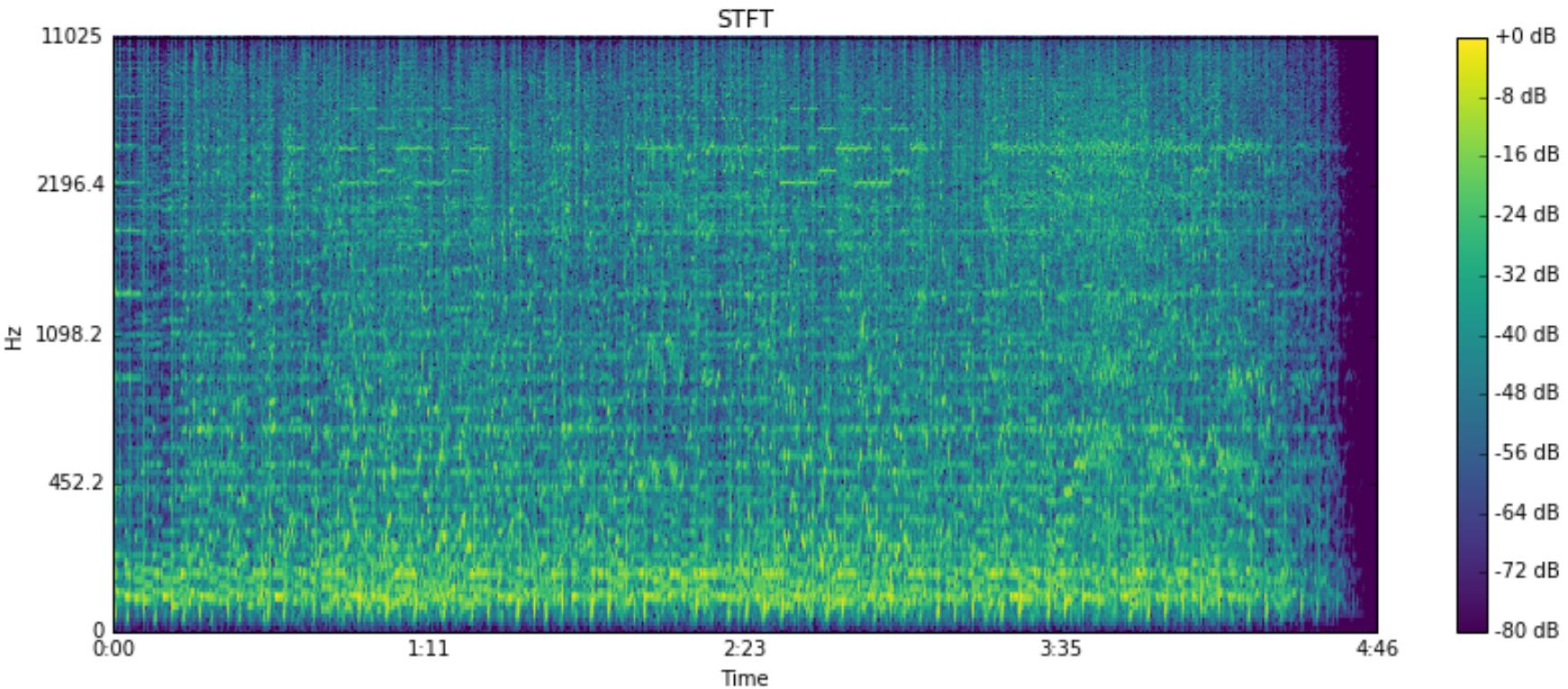
Estimating MGP Piano Attribute

USING DEEP LEARNING



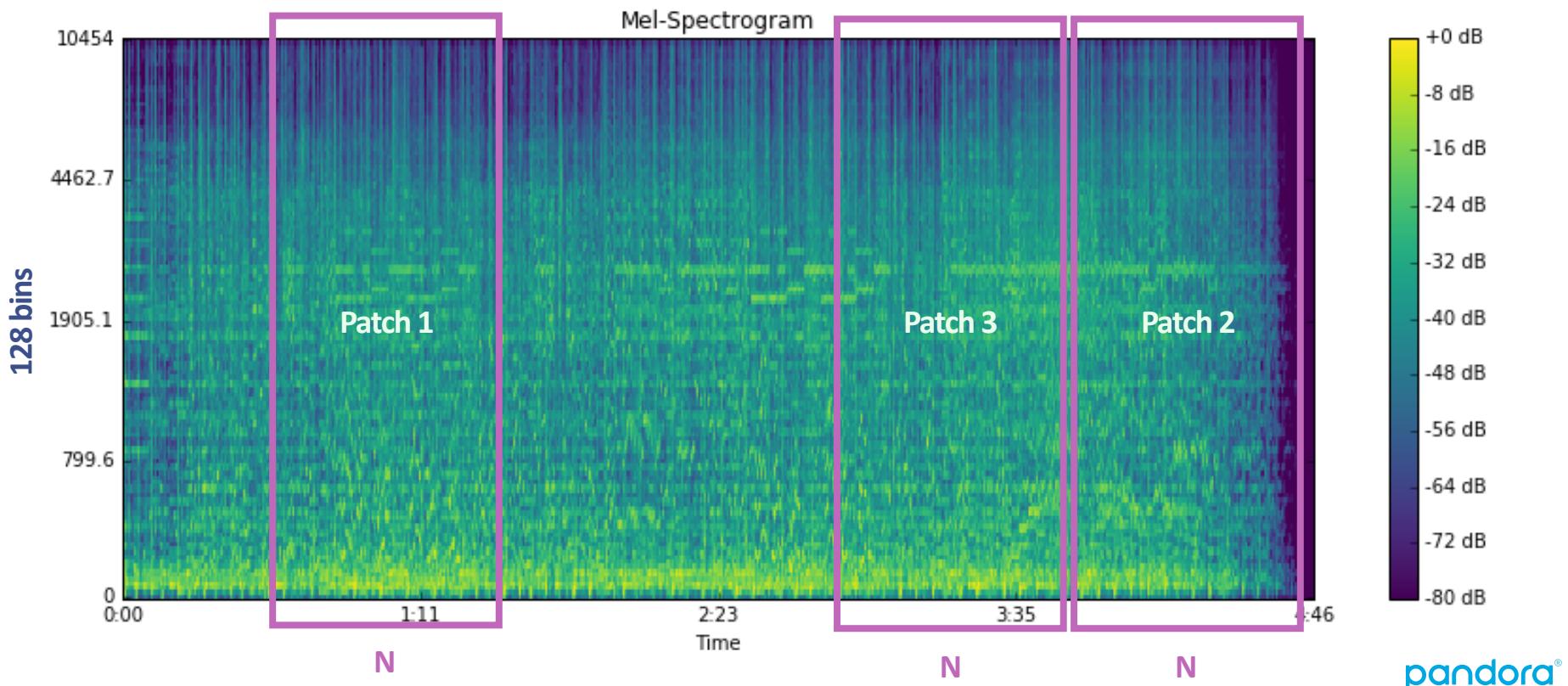
Piano Detector

USING DEEP LEARNING



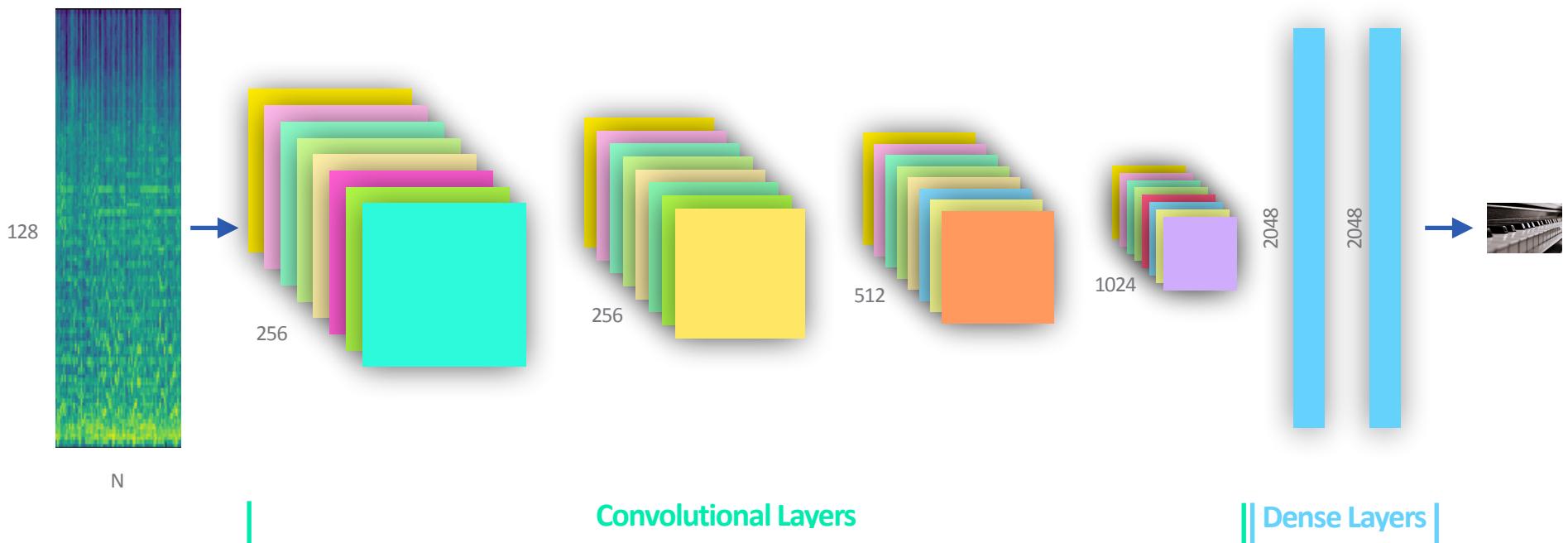
Piano Detector

USING DEEP LEARNING



Piano Detector

DEEP ARCHITECTURE



Piano Detector

TRAINING

Training

480k Patches

Validation

60k Patches

Test

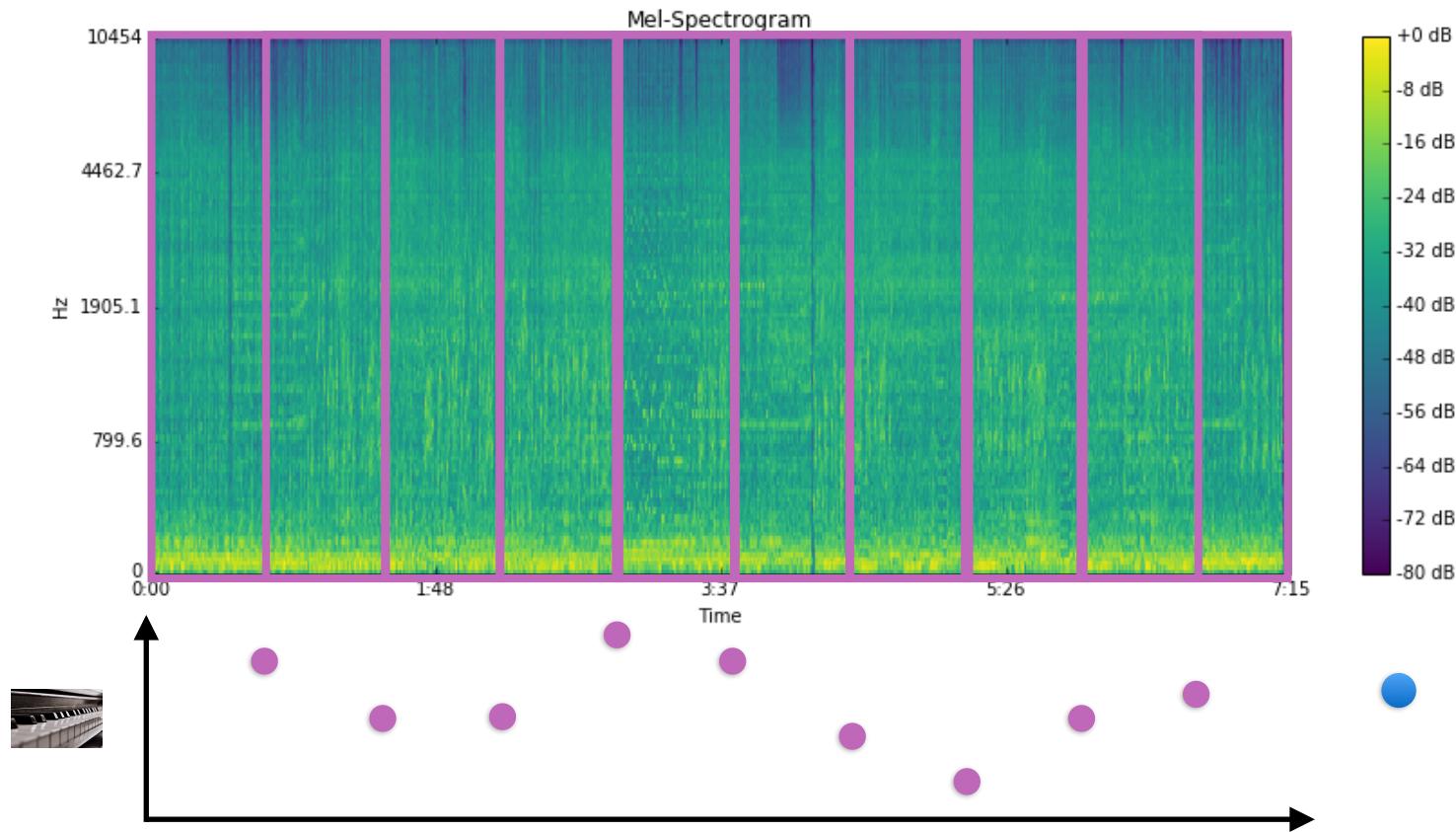
60k Patches

Architecture	seconds / epoch
CPU	~5700
GPU (CUDA)	100
GPU (CUDA + CuDNN)	70

(Different model)

Piano Detector

ESTIMATING FULL TRACKS



pandora®

Piano Detector

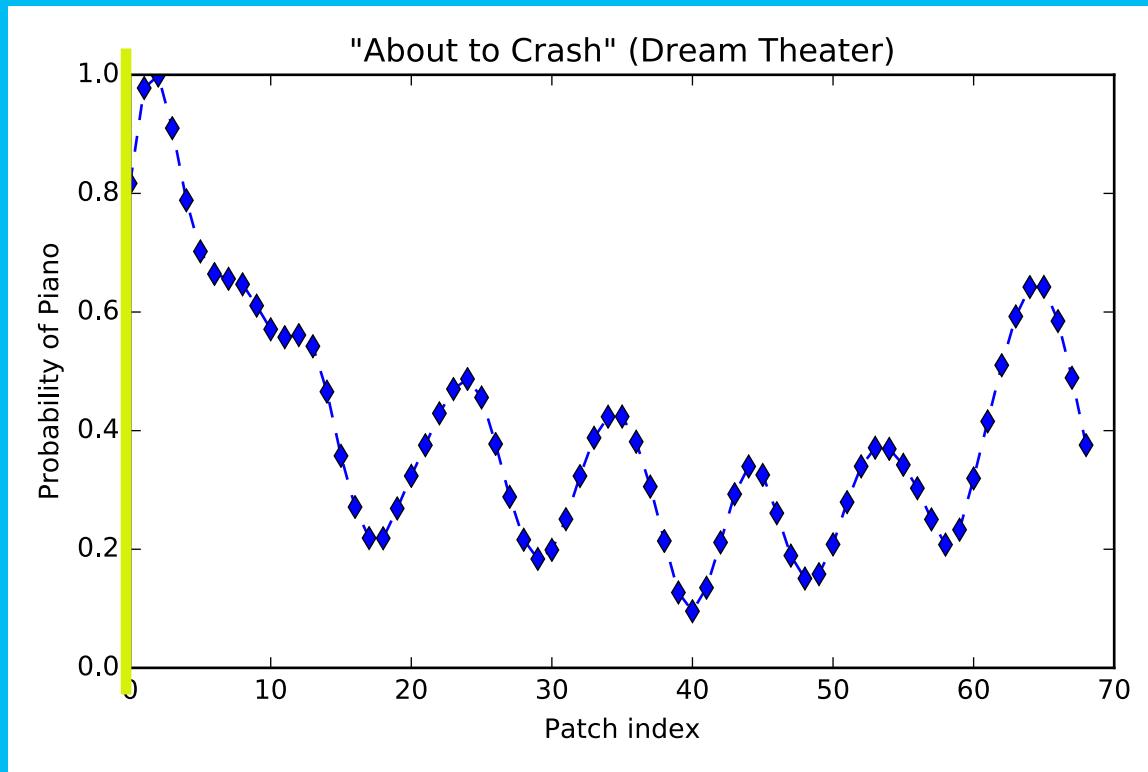
RESULTS

Method	AUC ROC	
Logistic Regression	~85%	
Gradient Boosted Trees	~87.5%	
Deep Learning	91.1%	

Internal Dataset of 50k tracks

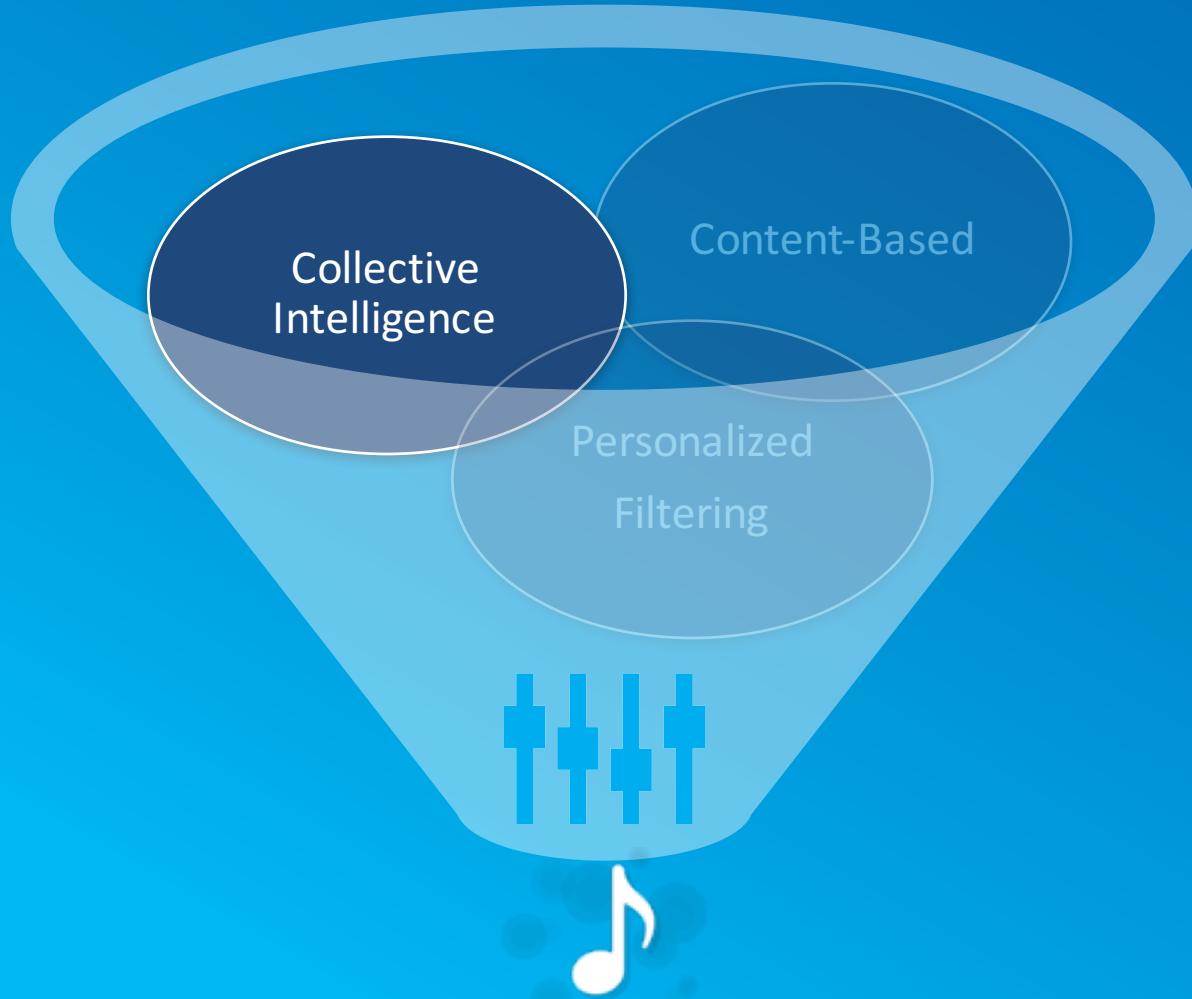
Piano Detector

OUTPUT EXAMPLE



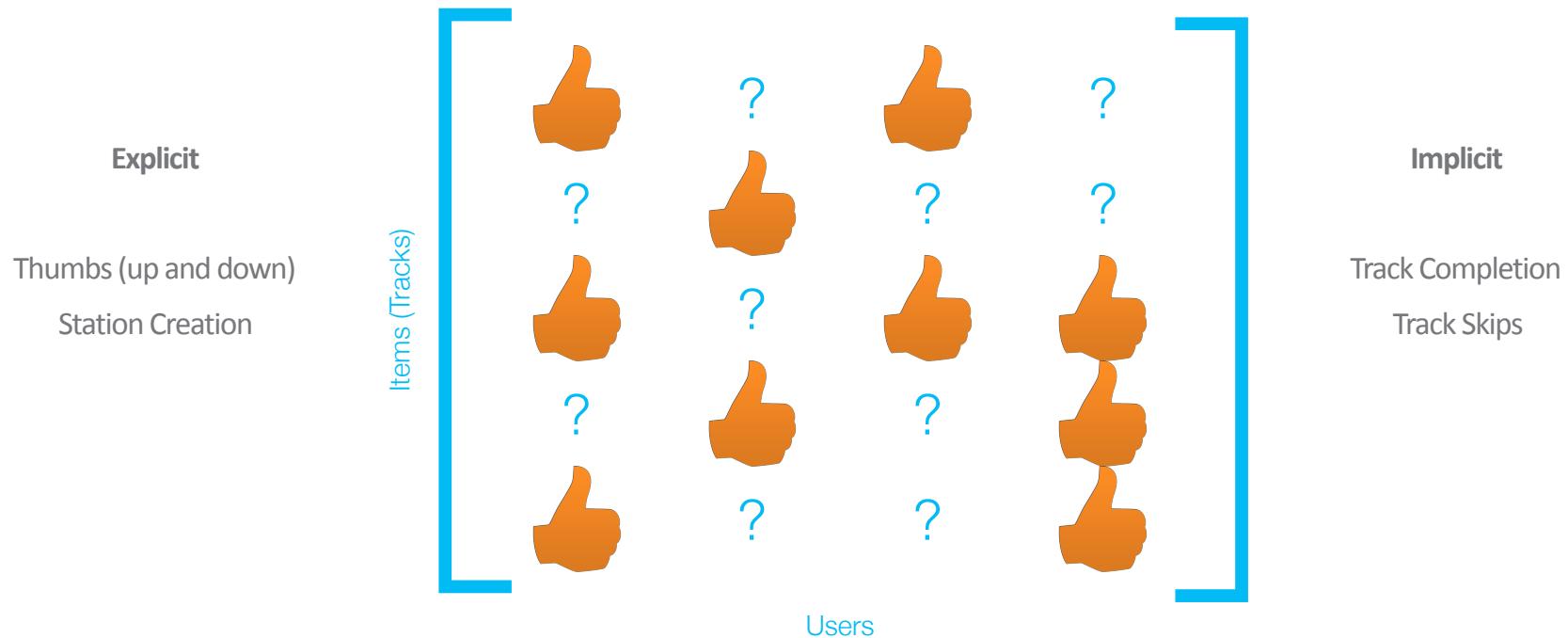
Outline

- Pandora Overview
- Large Scale Music Recommendation with Deep Learning:
 - Machine Listening
 - **Collaborative Filtering**



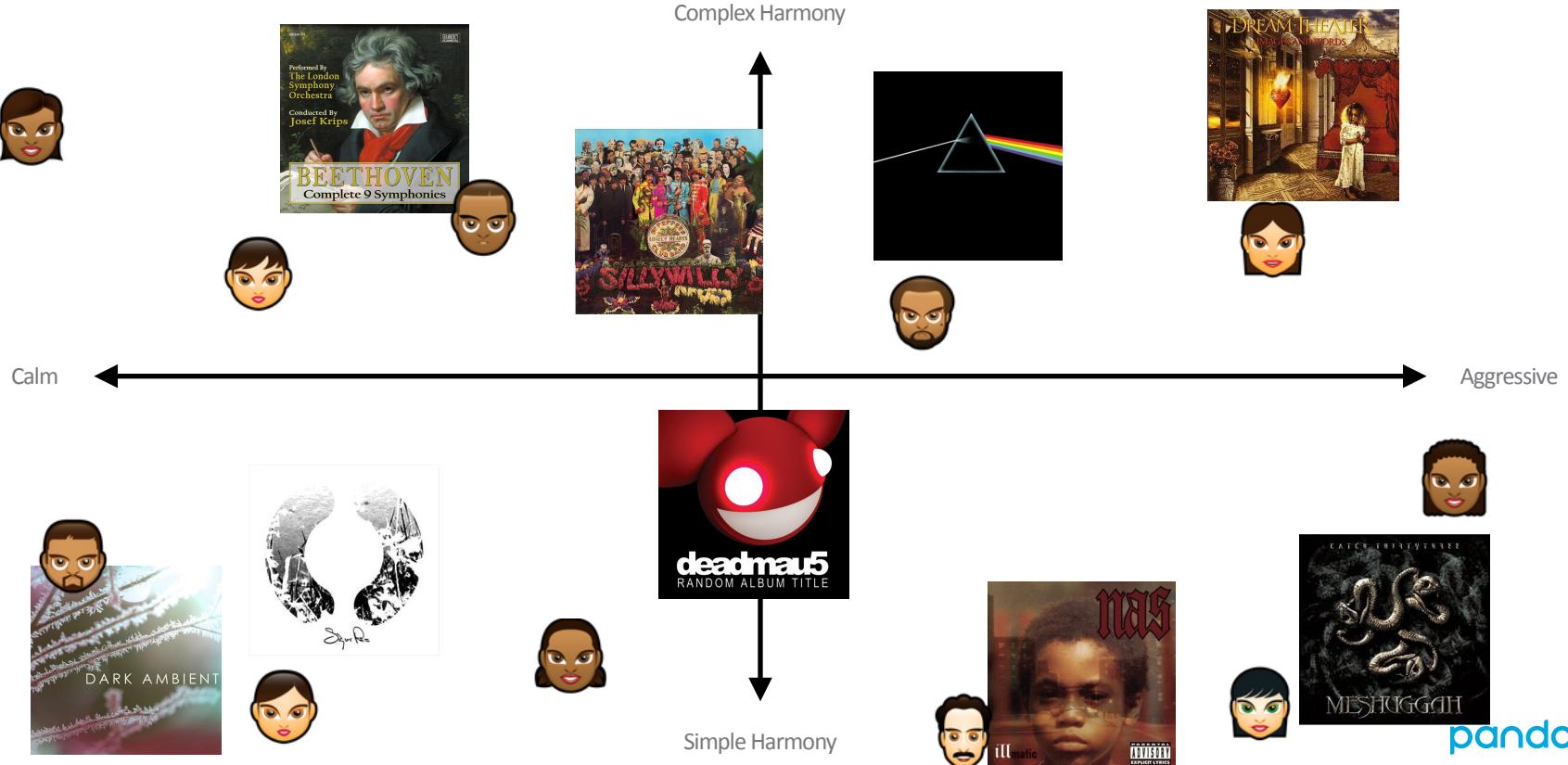
Collaborative Filtering

PROBLEM OVERVIEW



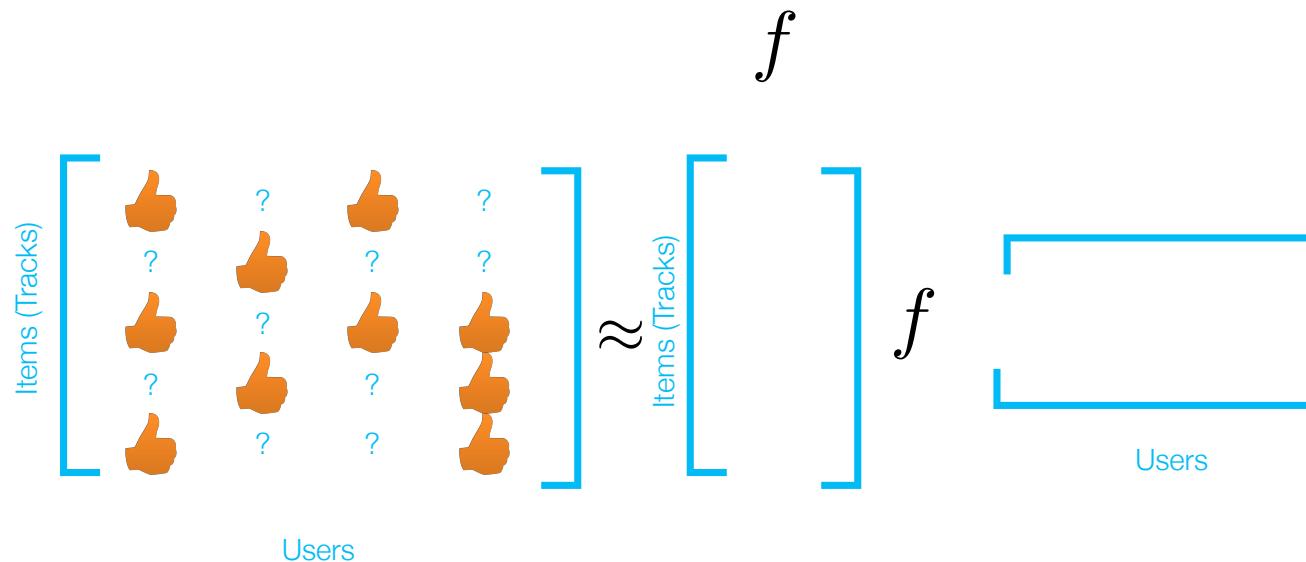
Collaborative Filtering

LATENT FACTORS



Collaborative Filtering

MATRIX FACTORIZATION



(Koren et al., 2009)

pandora®

Collaborative Filtering

PROBLEM FORMULATION

Given Item i and User u :

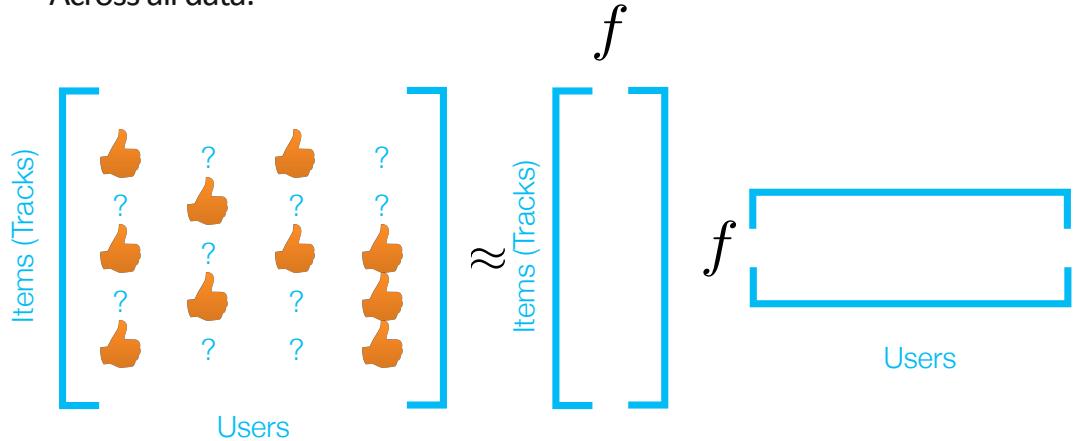
Rating: r_{ui}

Item Latent Factor: $q_i \in \mathbb{R}^f$

User Latent Factor: $p_u \in \mathbb{R}^f$

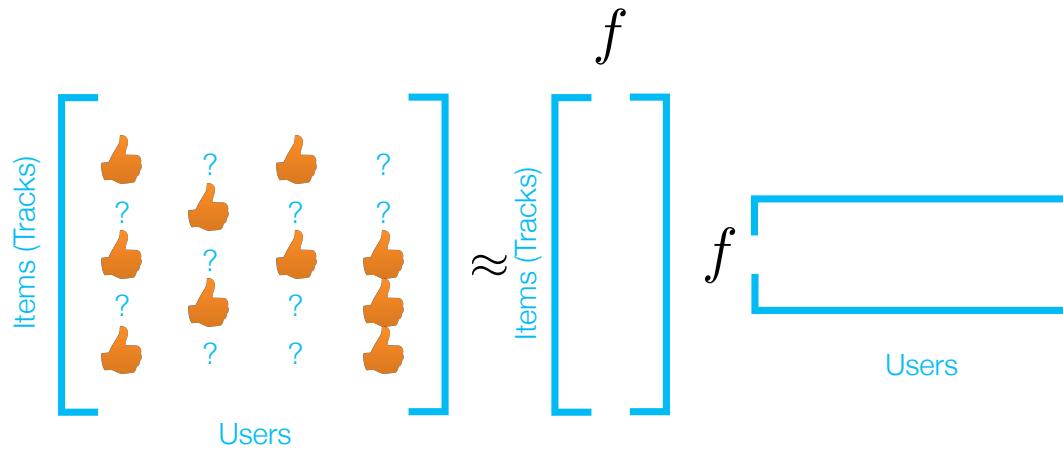
Rating Approximation: $\hat{r}_{ui} = q_i^T p_u$

Across all data:



Collaborative Filtering

MATRIX FACTORIZATION



$$\operatorname{argmin}_{q^*, p^*} \sum_{u, i \in \mathcal{S}} (r_{ui} - q_i^T p_u)^2 + \lambda(||q_i||^2 + ||p_u||^2)$$

(Koren et al., 2009)

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	While My Guitar Gently Weeps (Love Version)

Collaborative Filtering

THE GOOD AND THE BAD

Rich preference-driven similarity space

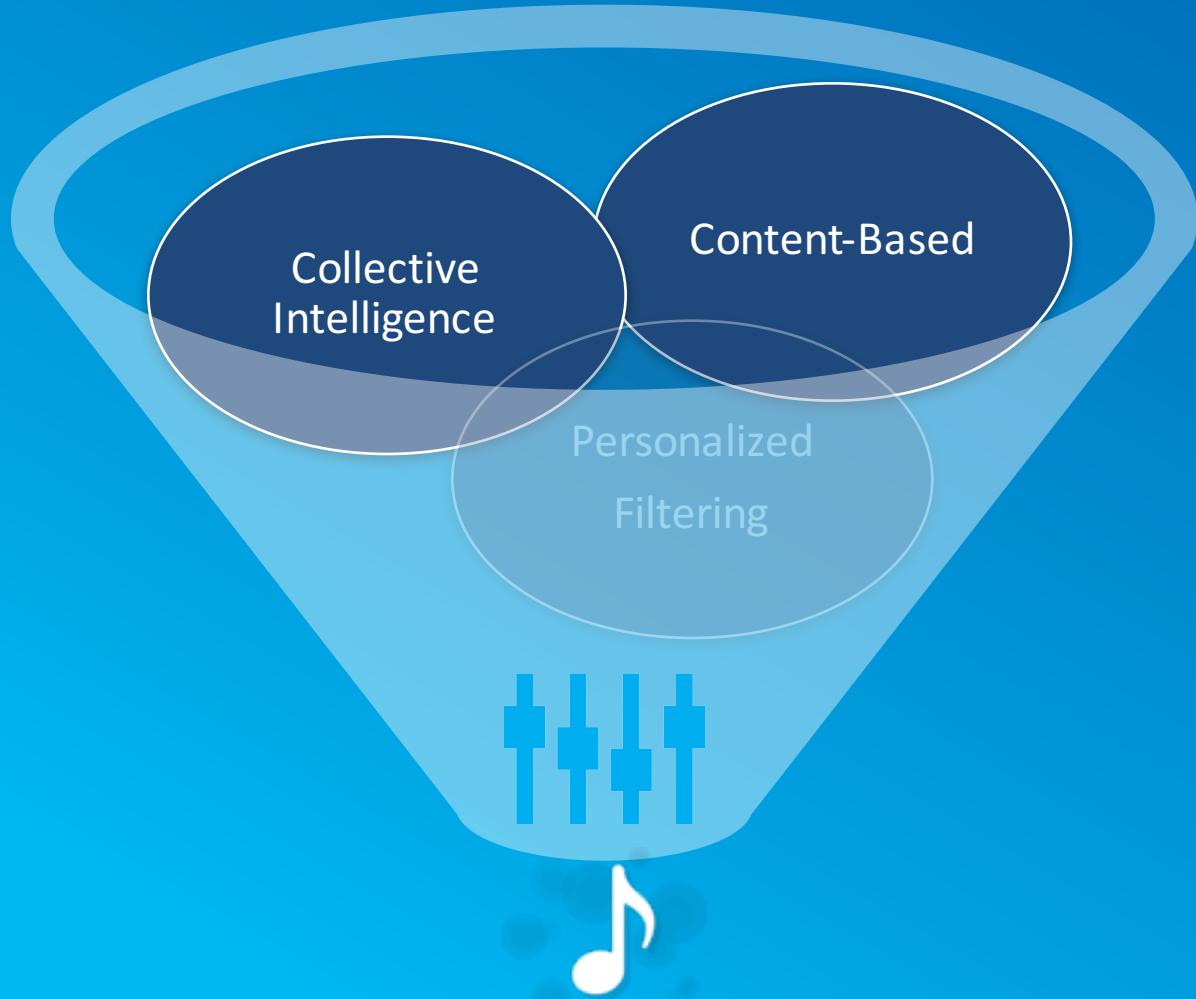
Powerful at matching the right song
with the right listener



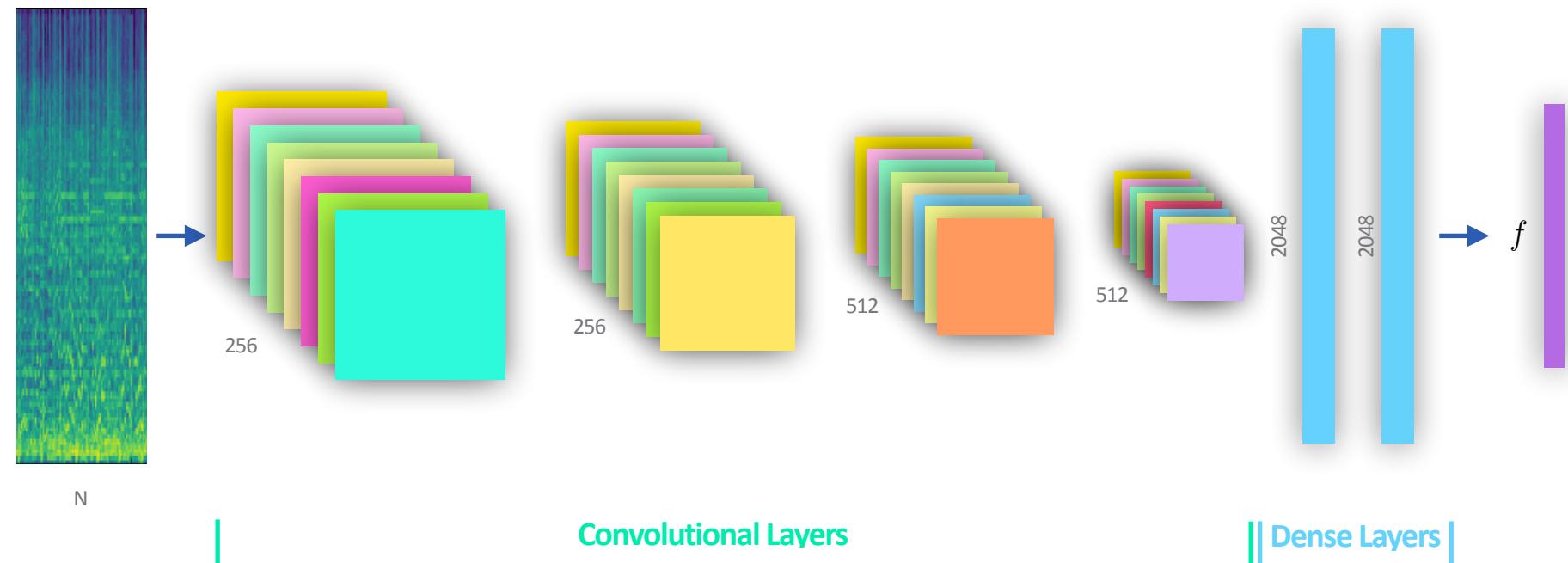
Latent space is generally not interpretable

Can only recommend items that
have already been rated





Approximating Item Factors WITH DEEP LEARNING

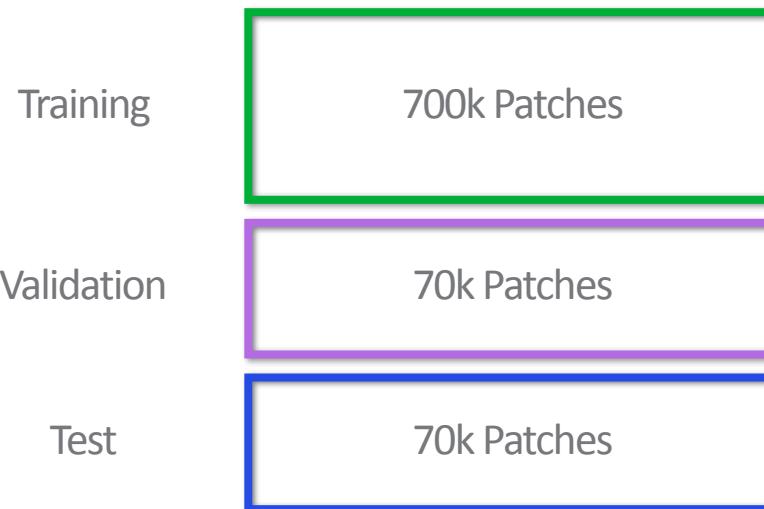


(van den Oord et al., 2013)

pandora®

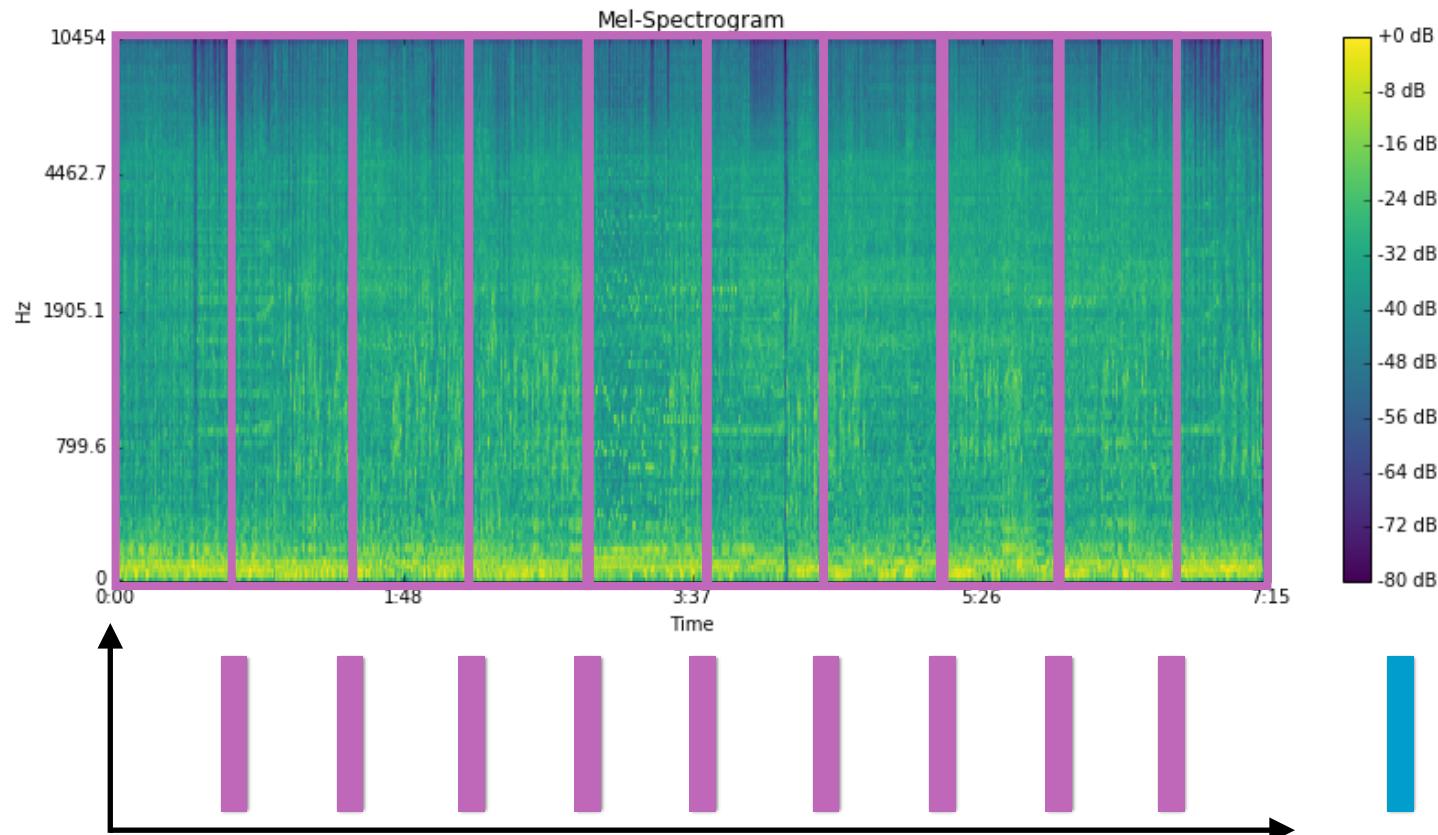
Approximating Item Factors

TRAINING



Approximating Item Factors

ESTIMATING FULL TRACKS



Recommending from Learned Factors

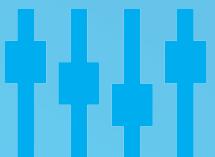
EXAMPLE

	Artist	Title
Query Track	La Bossa d'Urina	Mama i Papa
Ranked 1	Johnny Cash	I'm Going To Memphis
Ranked 2	Merle Haggard	The Woman Made A Fool Out Of Me
Ranked 3	Waylon Jennings	Lonesome On'ry And Mean

Recommending from Learned Factors

EXAMPLE

	Artist	Title
Query Track	Sargon	Vida
Ranked 1	Escape The Fate	Remember Every Scar
Ranked 2	Five Finger Death Punch	Coming Down
Ranked 3	Bullet For My Valentine	Breaking Out, Breaking Down



Content-Based

Collective
Intelligence

Personalized
Filtering



**ENSEMBLE OF RECOMMENDERS MAY PRODUCE
OPTIMAL RECOMMENDATIONS**



MAN vs MACHINE?



MAN + MACHINE



MAN + MACHINE
“Mix of Art and Science”

THANKS!

ONIETO@PANDORA.COM

pandora®

References

- Van Den Oord, A., Dieleman, S., Zen, H., Simonyan, K., Vinyals, O., Graves, A., Kavukcuoglu, K. (2016). Wavenet: A Generative Model for Raw Audio, 846–849.
- Simonyan, K., & Zisserman, A. (2015). Very Deep Convolutional Networks for Large-Scale Image Recognition. International Conference on Learning Representations, 1–14. <http://doi.org/10.1101/10.1016/j.infsof.2008.09.005>
- Mcfee, B., Humphrey, E. J., & Bello, J. P. (2015). A Software Framework For Musical Data Augmentation. In Proc. of the 16th International Society for Music Information Retrieval Conference (pp. 248–254). Málaga, Spain.
- Koren, Y., Bell, R., & Volinsky, C. (2009). Matrix Factorization Techniques for Recommender Systems. Computer, 42(8), 42–49. <http://doi.org/10.1109/MC.2009.263>
- Oord, A. Van Den, Dieleman, S., & Schrauwen, B. (2013). Deep Content-based Music Recommendation. Advances in Neural Information Processing Systems, 2643–2651. Retrieved from <http://papers.nips.cc/paper/5004-deep-content-based-music-recommendation>