

DEEP LEARNING FOR LARGE SCALE MUSIC RECOMMENDATION

ORIOL NIETO
ONIETO@PANDORA.COM

BIOSTAT SEMINAR
STANFORD, CA
OCTOBER 20, 2016

Outline

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

Outline

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

PANDORA | ONE™

Porcupine Tree Radio

Uri Nieto

+ Create Station



3:16

-2:45

Harvest

by Opeth

on <Q ch=0,1,3,11,14,19,23,24,37,...



Change Skin

Now Playing Music Feed My Profile

AC/DC Radio

Led Zeppelin Radio

Fryderyk Chopin Radio

Classical Radio

Suspense Radio

Dire Straits Radio

Sour Jazz Radio

The Waiting Is The H...

Spock's Beard Radio

Dream Theater Radio

Tool Radio

Opeth Radio

Riverside Radio

Steven Wilson Radio

Madee Radio

Mourn Radio

Monuments Radio

Metallica Radio

Chroma Key Radio

Meshuggah Radio

Porcupine Tree Radio

add variety

options ▾

Date

A - Z



on <Q ch=0,1,3,11,14,19,23,24,37,56 q=72 p=8.3% s=R144429 f=0> Blackwater Park

Harvest
by Opeth

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



80 M
MAUs



>3 M
Songs



24 H
Per Month



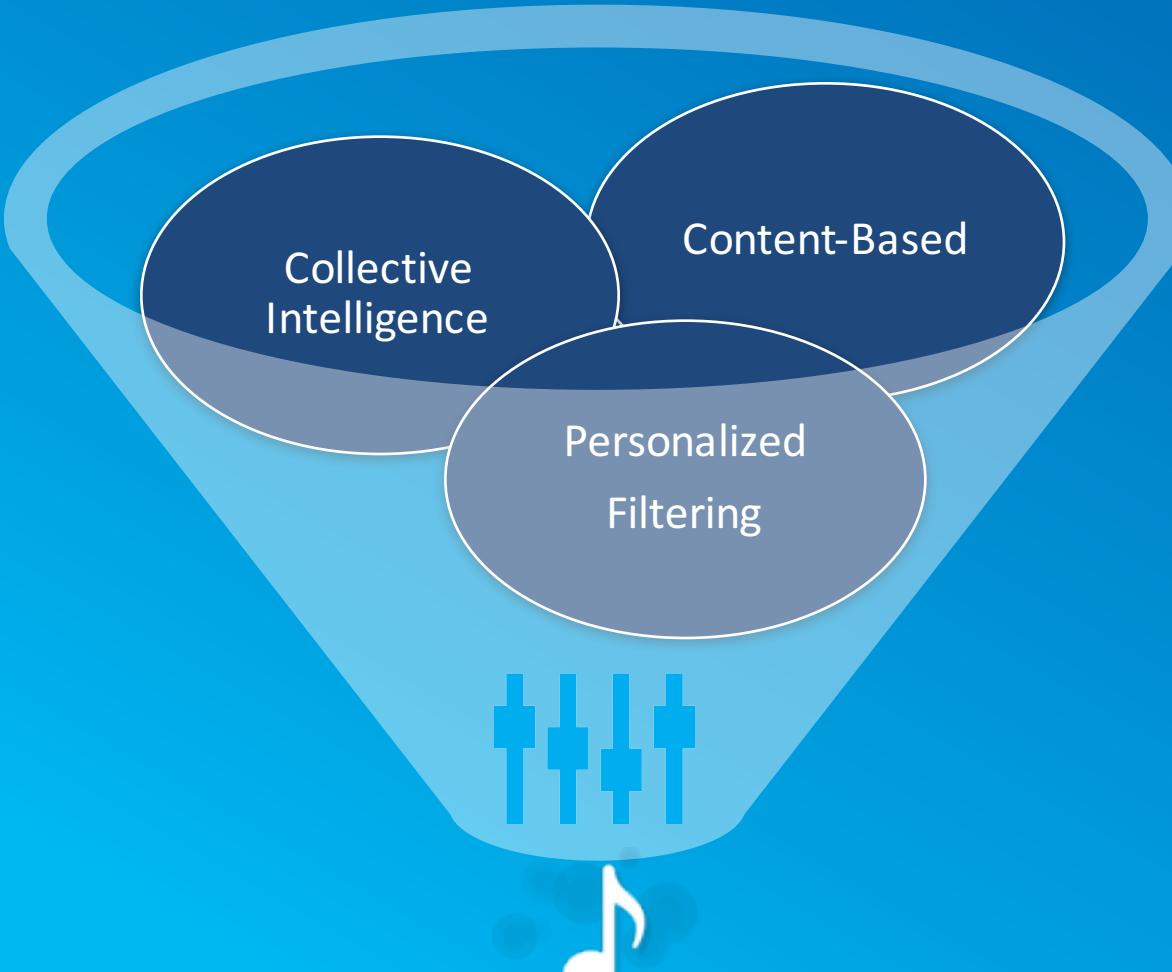
180 K
Artists spinning
every month



75 B
Thumbs



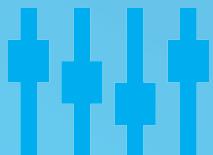
11 B
Stations



Collective
Intelligence

Content-Based

Personalized
Filtering

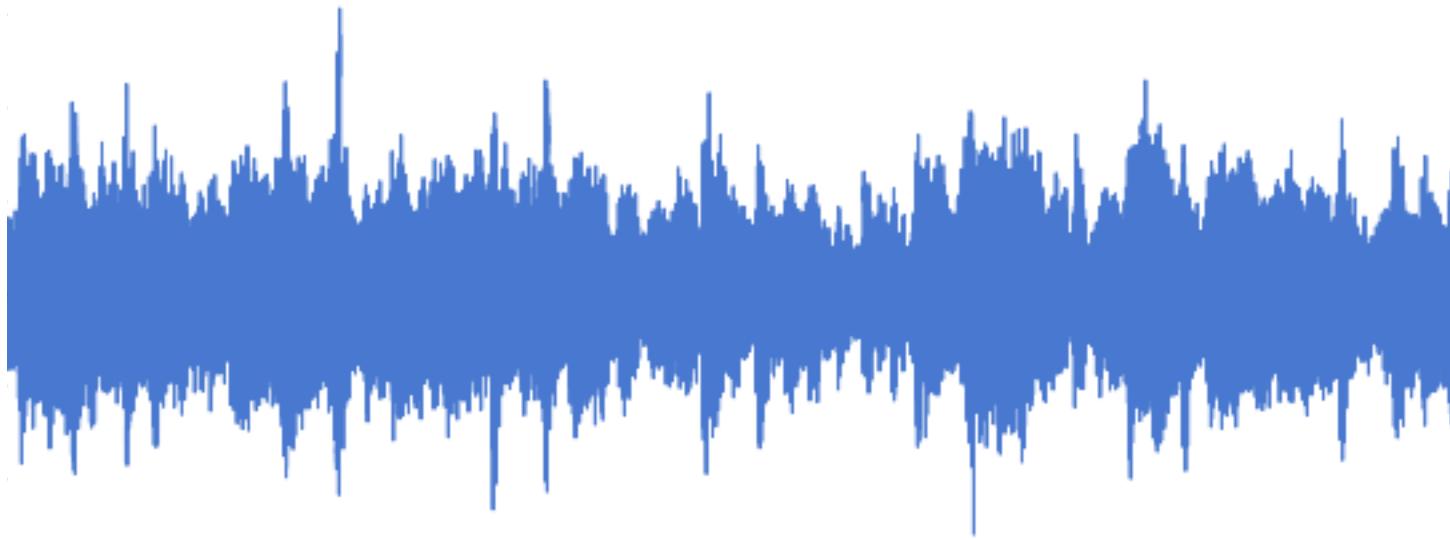


Outline

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

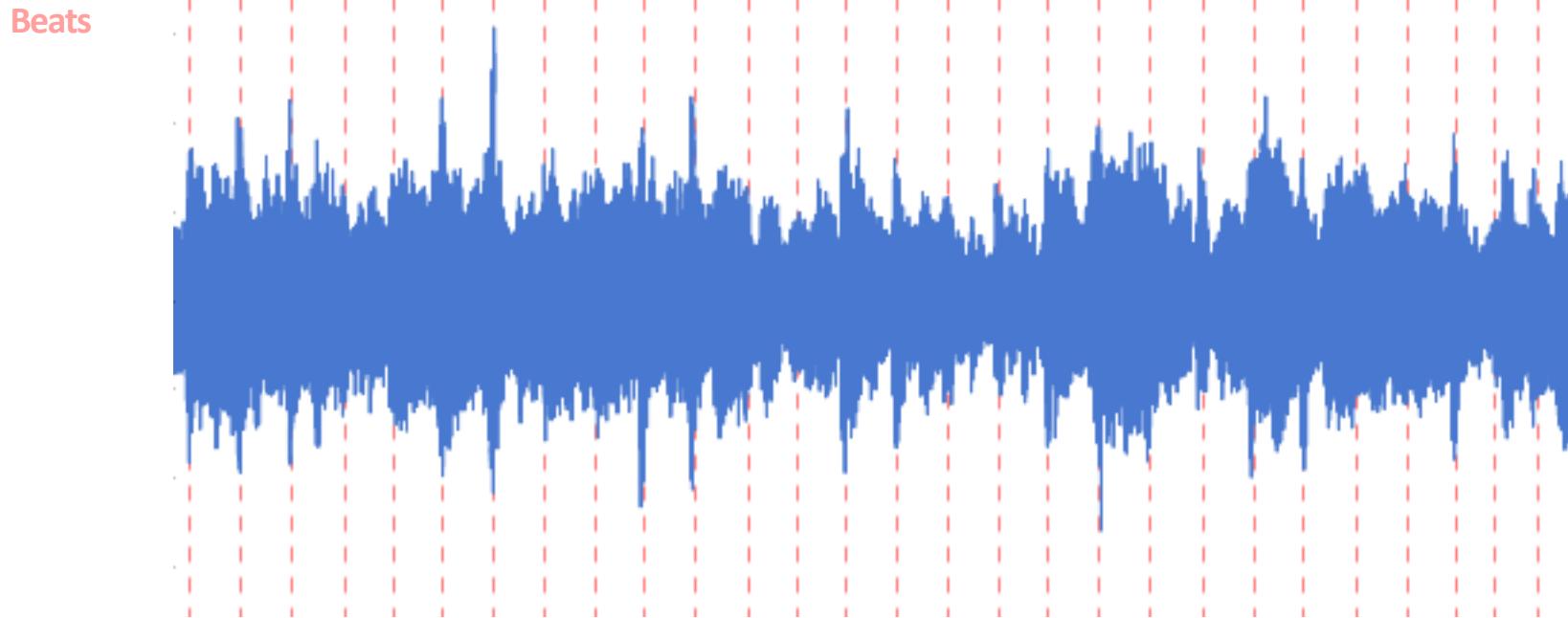
Music Informatics

(AKA MUSIC INFORMATION RETRIEVAL)



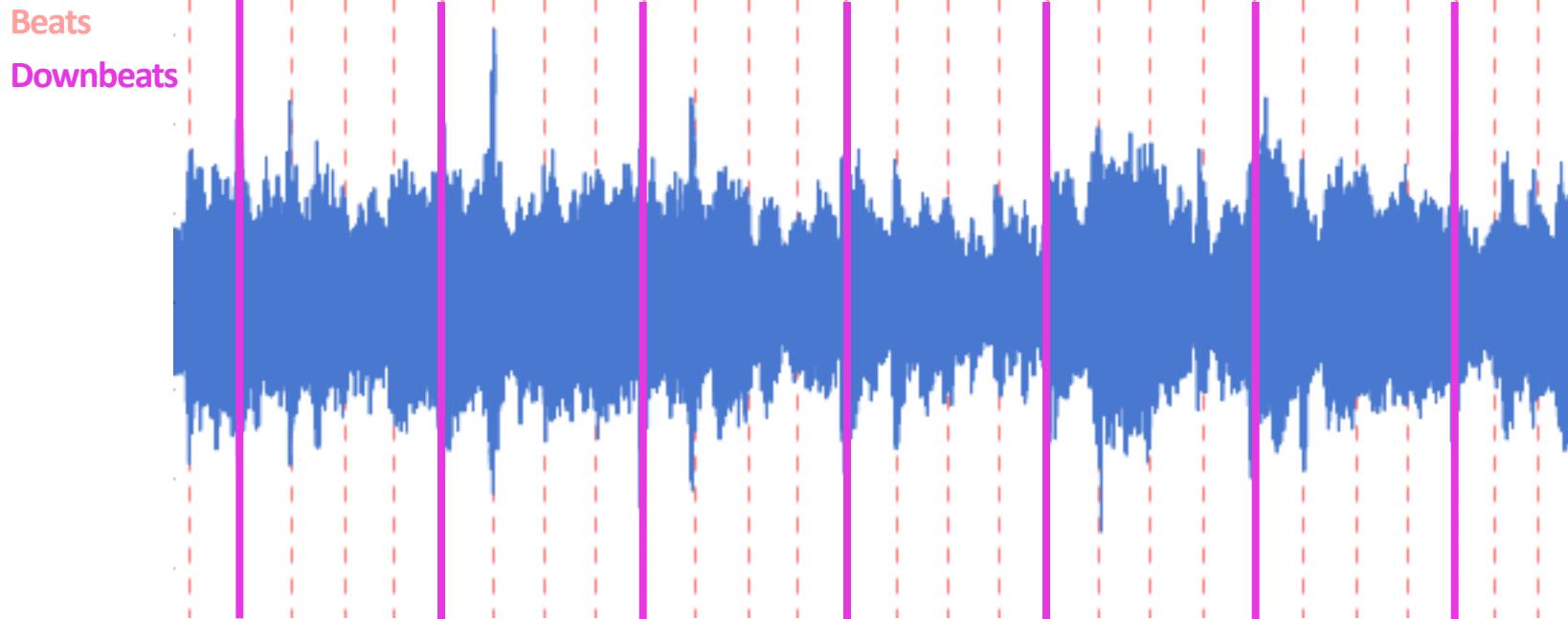
Music Informatics

(AKA MUSIC INFORMATION RETRIEVAL)



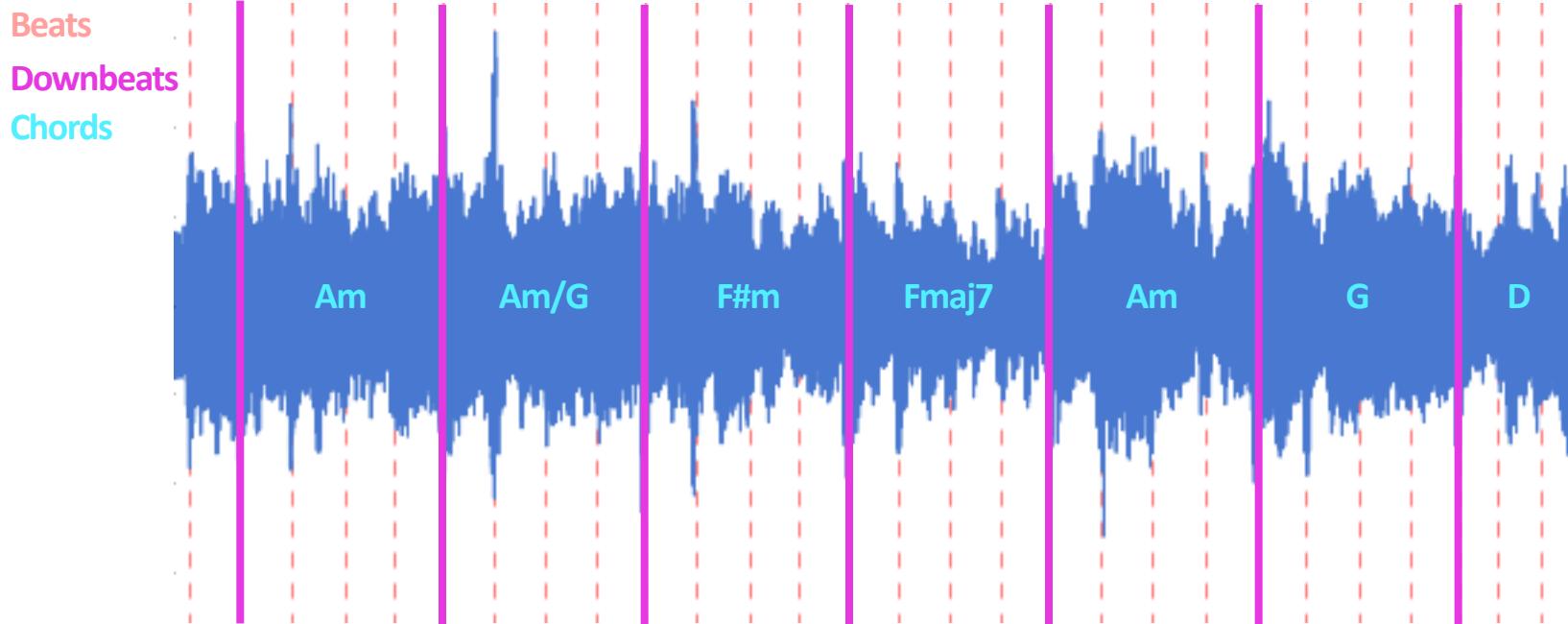
Music Informatics

(AKA MUSIC INFORMATION RETRIEVAL)



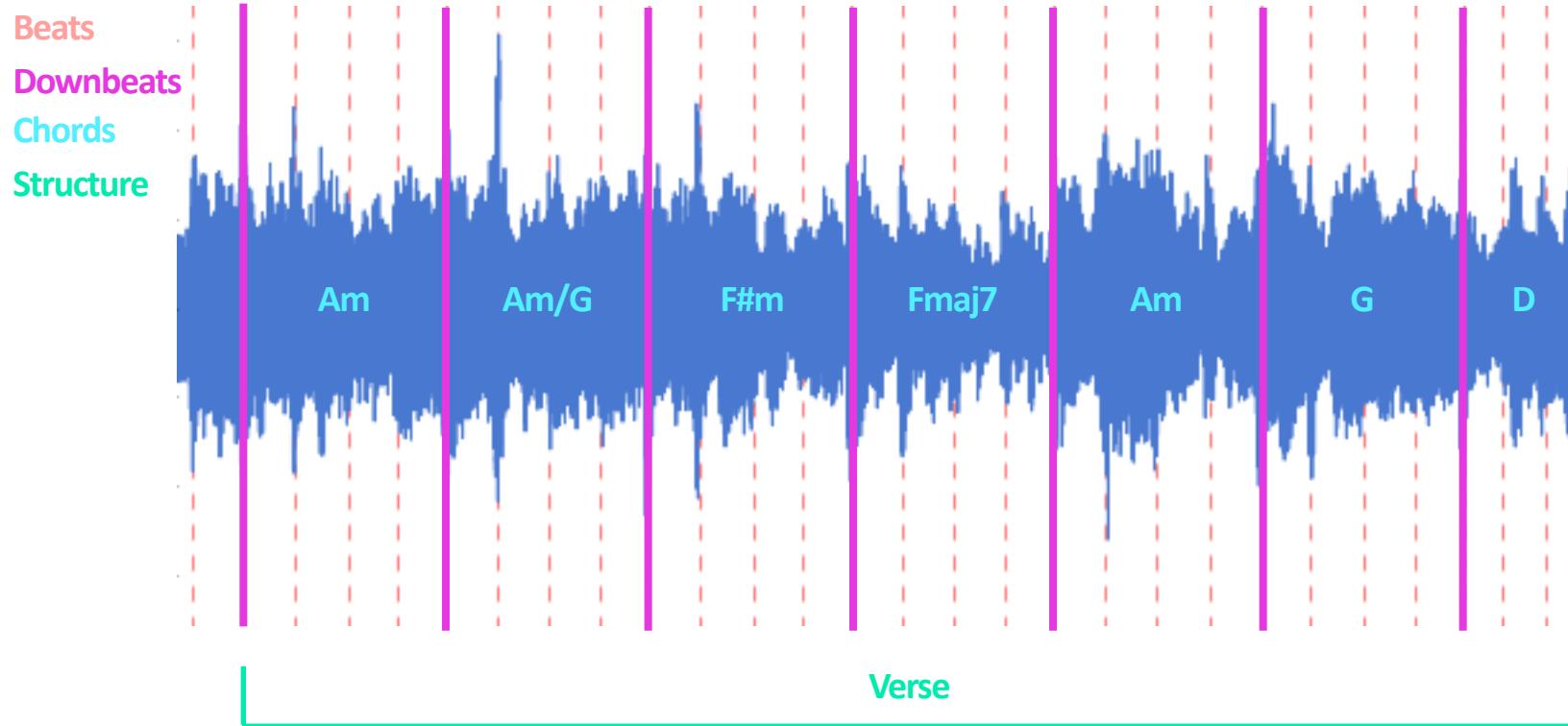
Music Informatics

(AKA MUSIC INFORMATION RETRIEVAL)



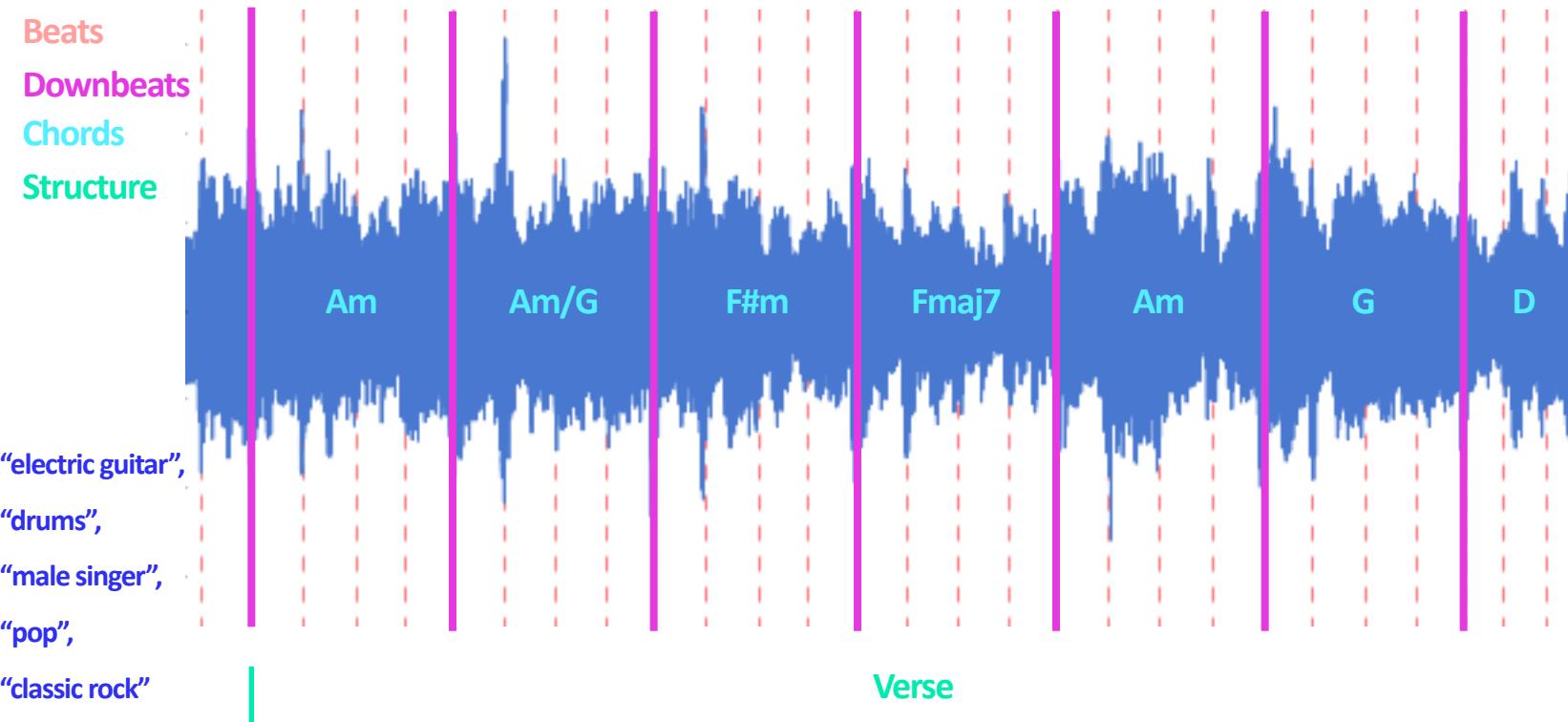
Music Informatics

(AKA MUSIC INFORMATION RETRIEVAL)



Music Informatics

(AKA MUSIC INFORMATION RETRIEVAL)



Music Informatics

THE BASICS



Music Informatics

THE BASICS

Discrete Fourier Transform

$$X = Wx$$



$$x \in \mathbb{R}^n$$

$$W \in \mathbb{C}^{n \times n}$$

$$X \in \mathbb{C}^n$$

Any sound can be built out of sinusoids

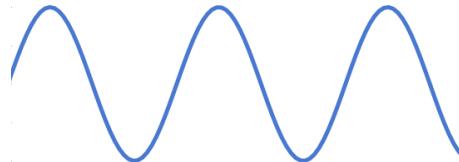


Discrete Fourier Transform

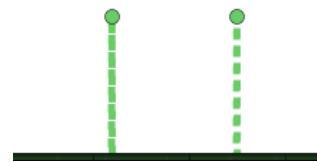
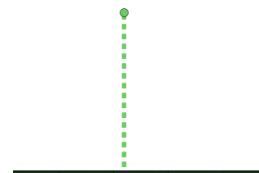
EXAMPLES

Any sound can be built out of sinusoids

x (time domain)

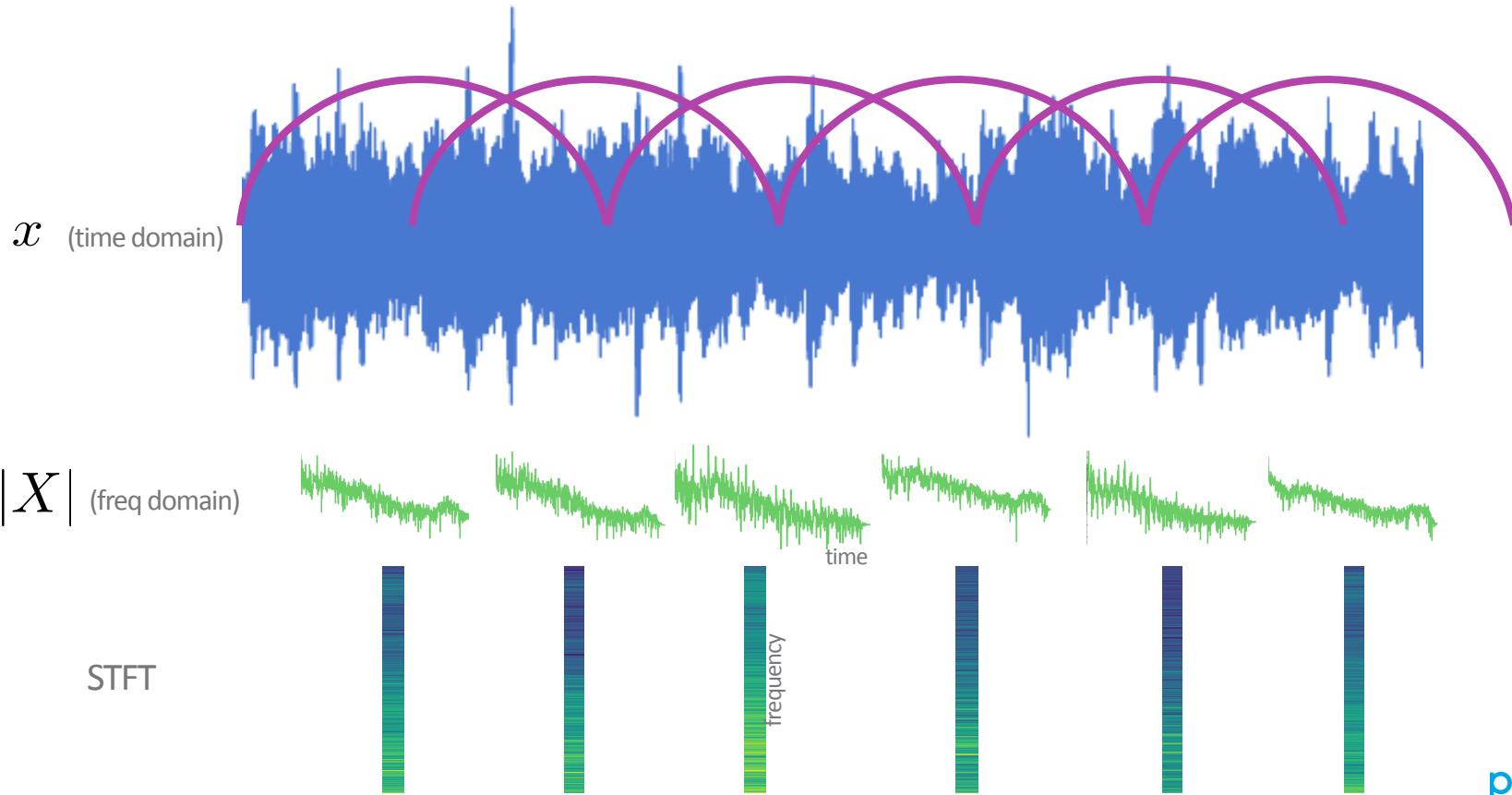


$|X|$ (frequency domain)



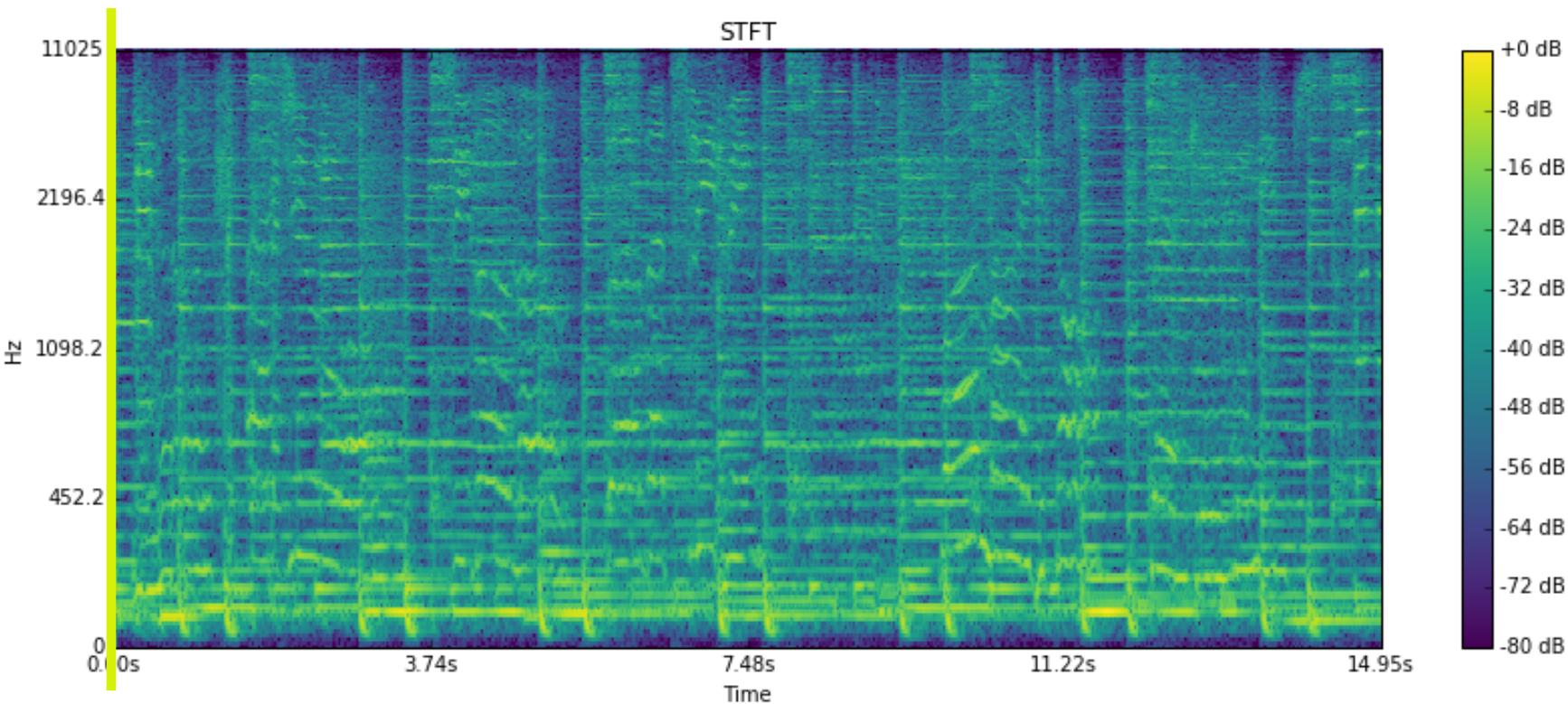
$$X = Wx$$

Short Time Fourier Transform



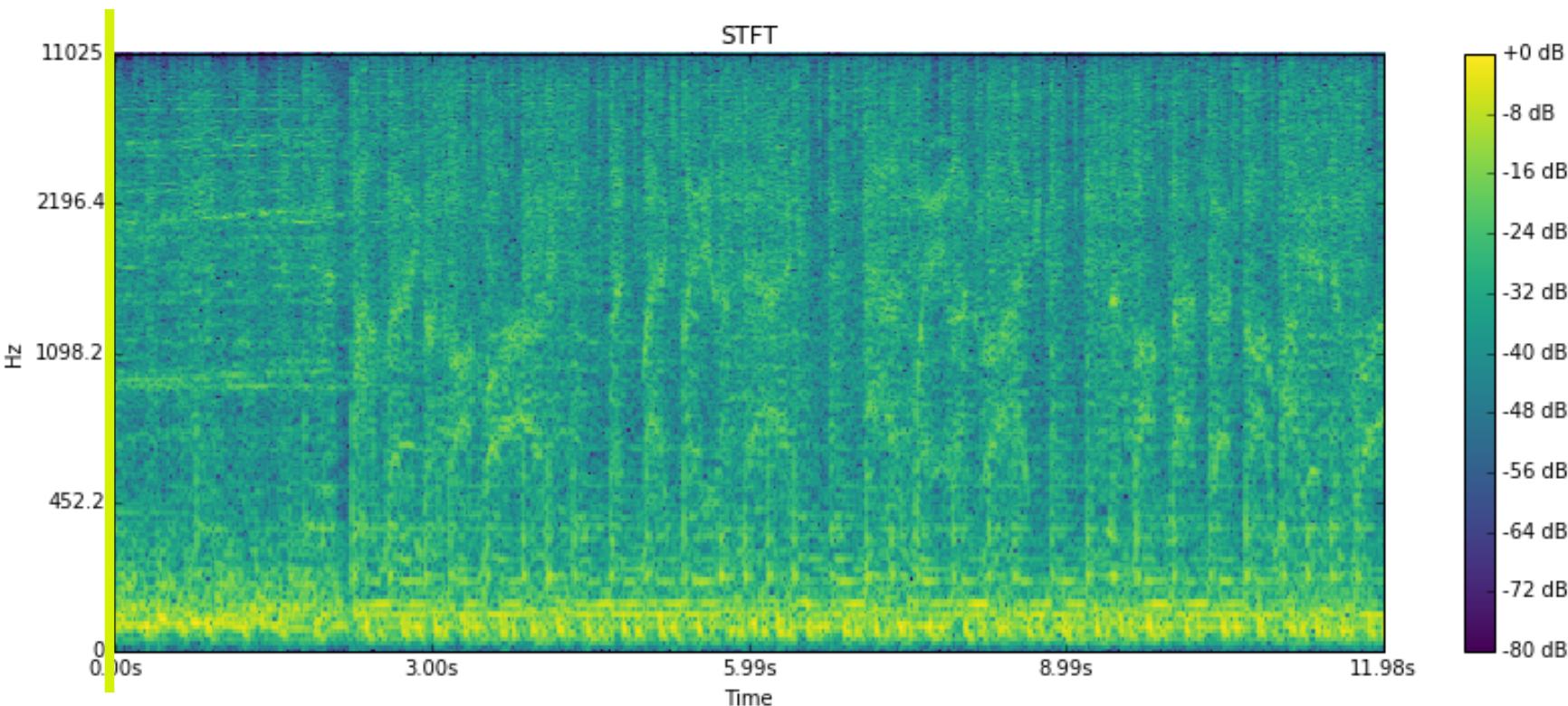
Short Time Fourier Transform

MY GUITAR GENTLY WEEPS - THE BEATLES



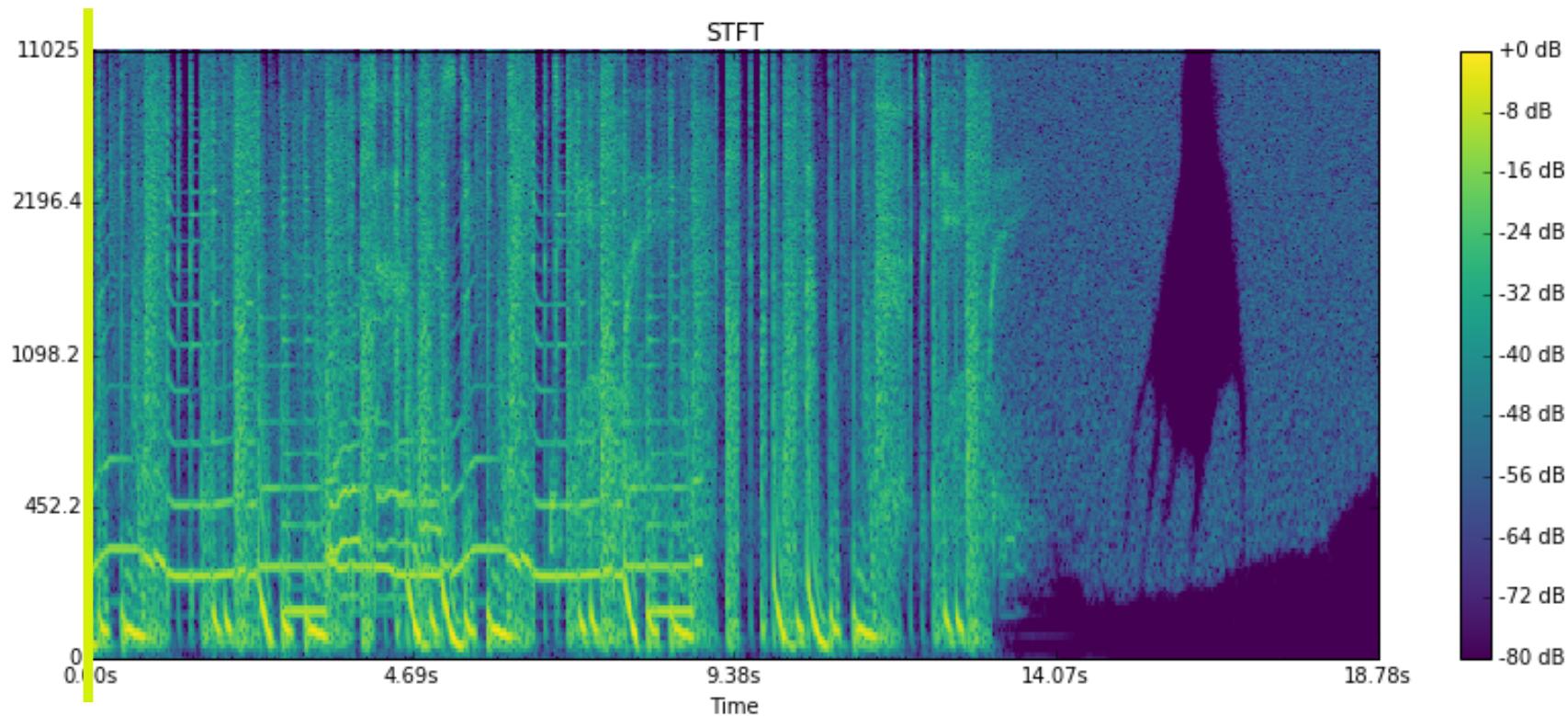
Short Time Fourier Transform

CLOCKWORK - MESHUGGAH

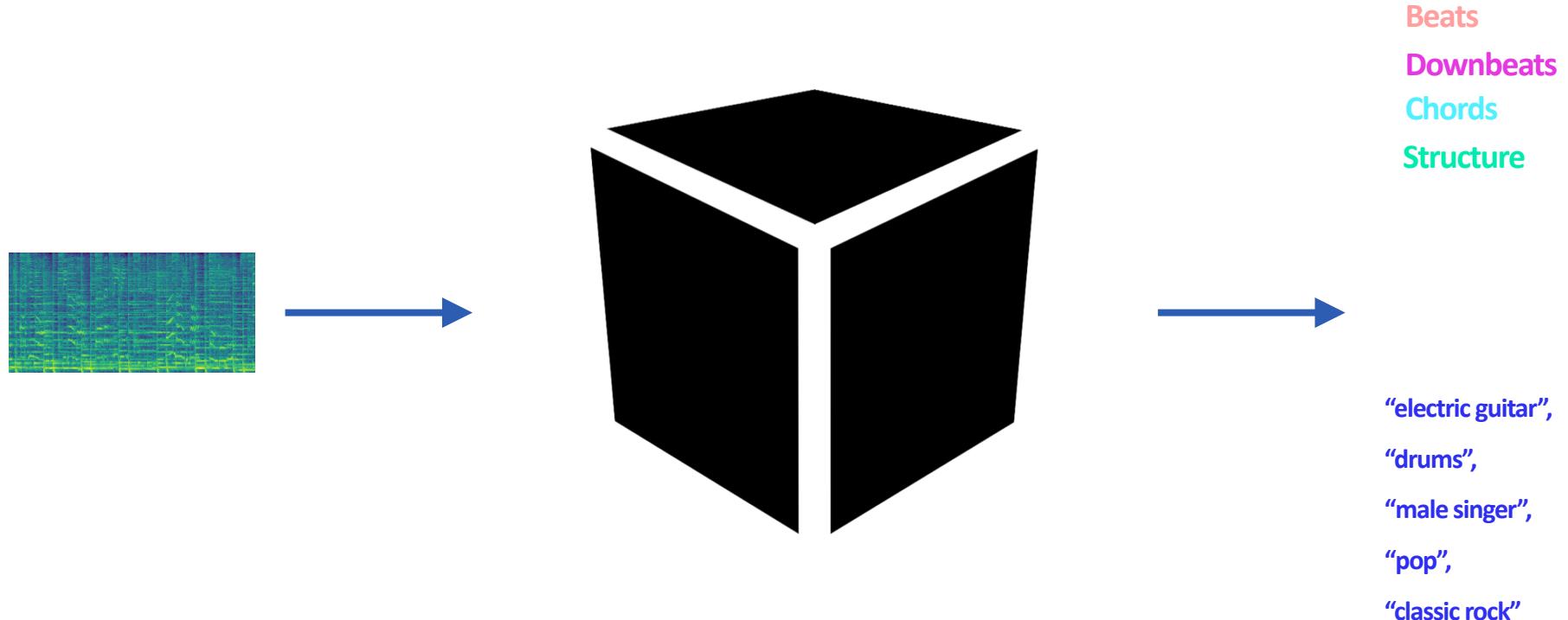


Short Time Fourier Transform

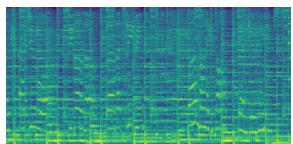
THE WARNING - NINE INCH NAILS



Extracting Information



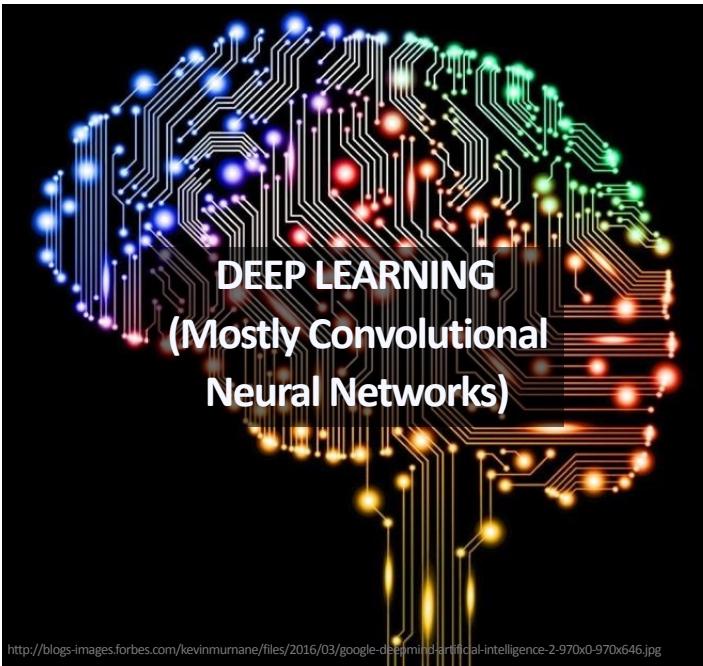
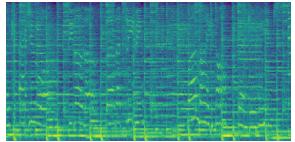
Hand Crafted Features



Beats
Downbeats
Chords
Structure

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

Learned Features

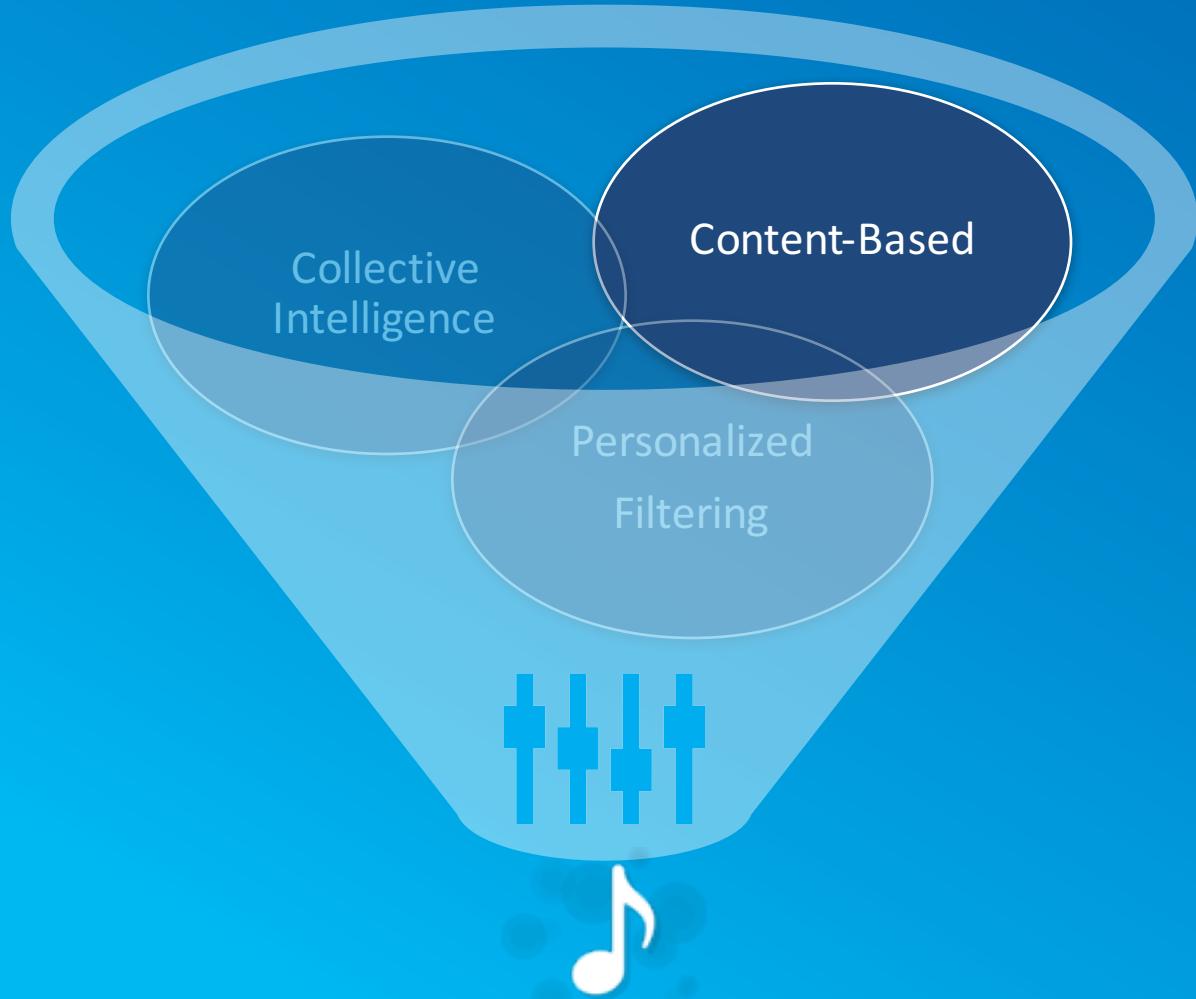


Beats
Downbeats
Chords
Structure

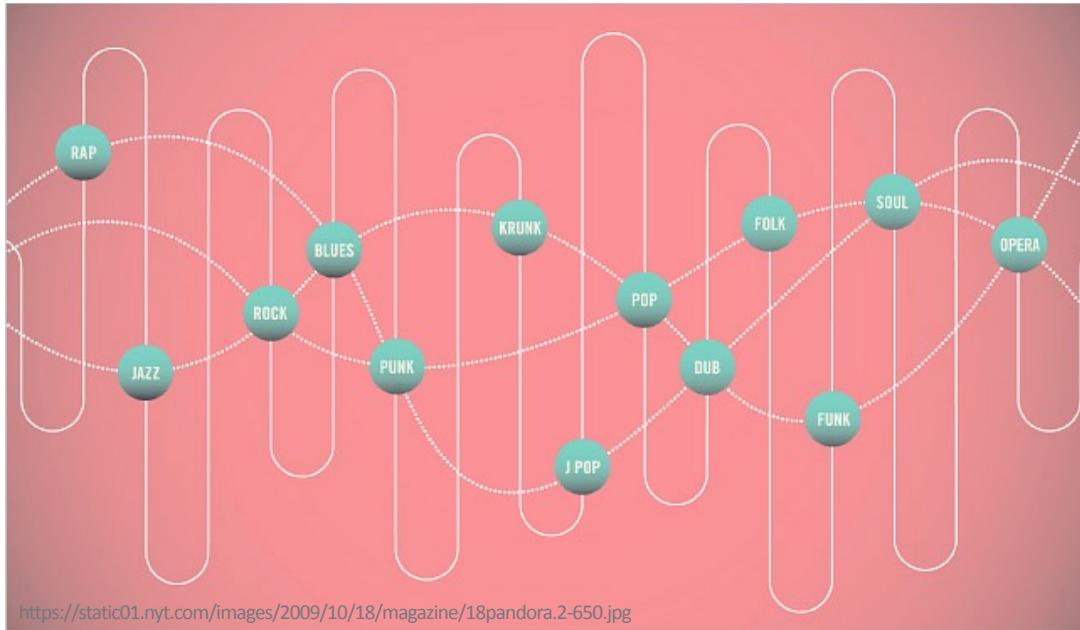
“electric guitar”,
“drums”,
“male singer”,
“pop”,
“classic rock”

Outline

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



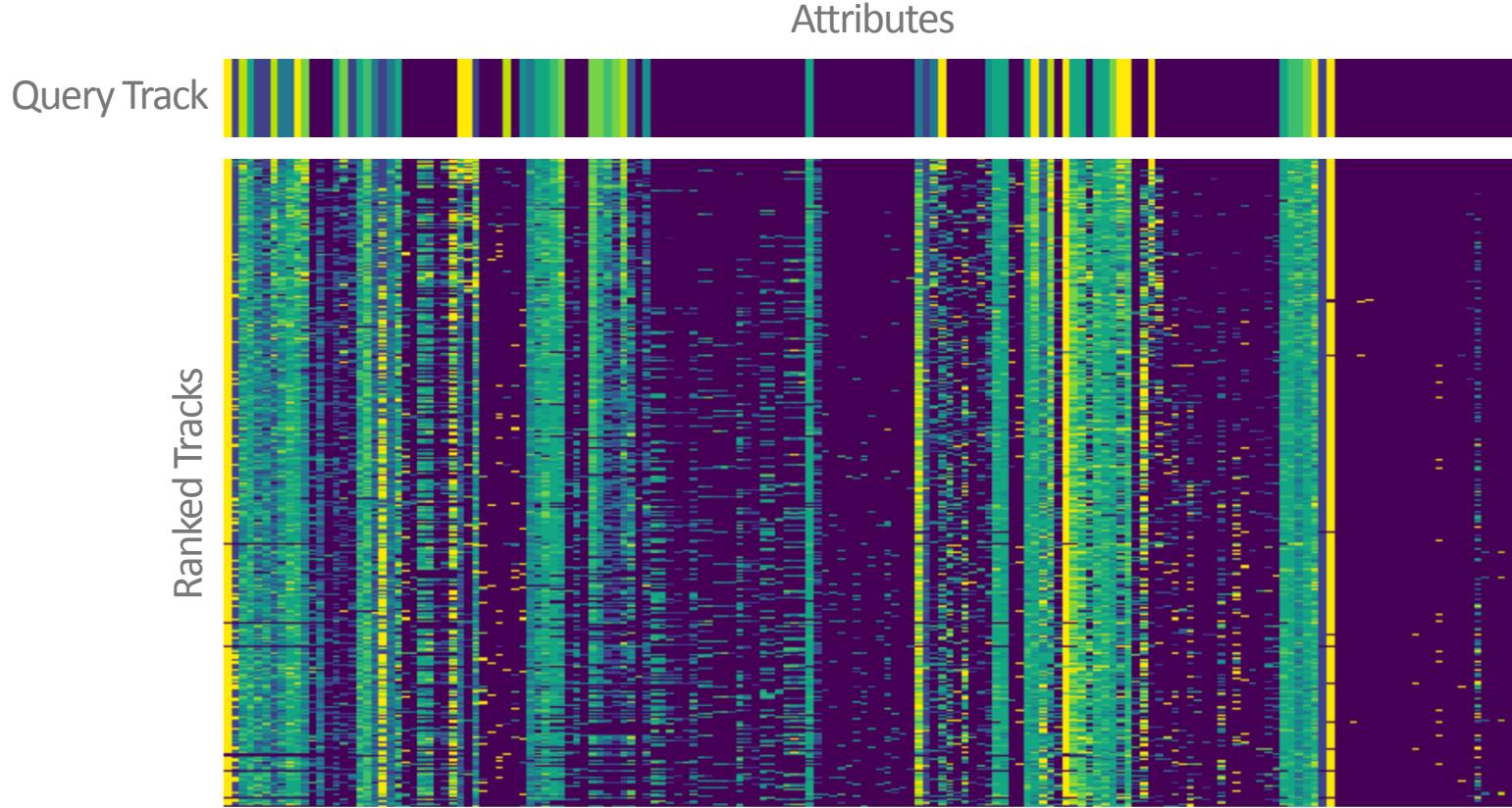
The Music *Genome* Project



>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

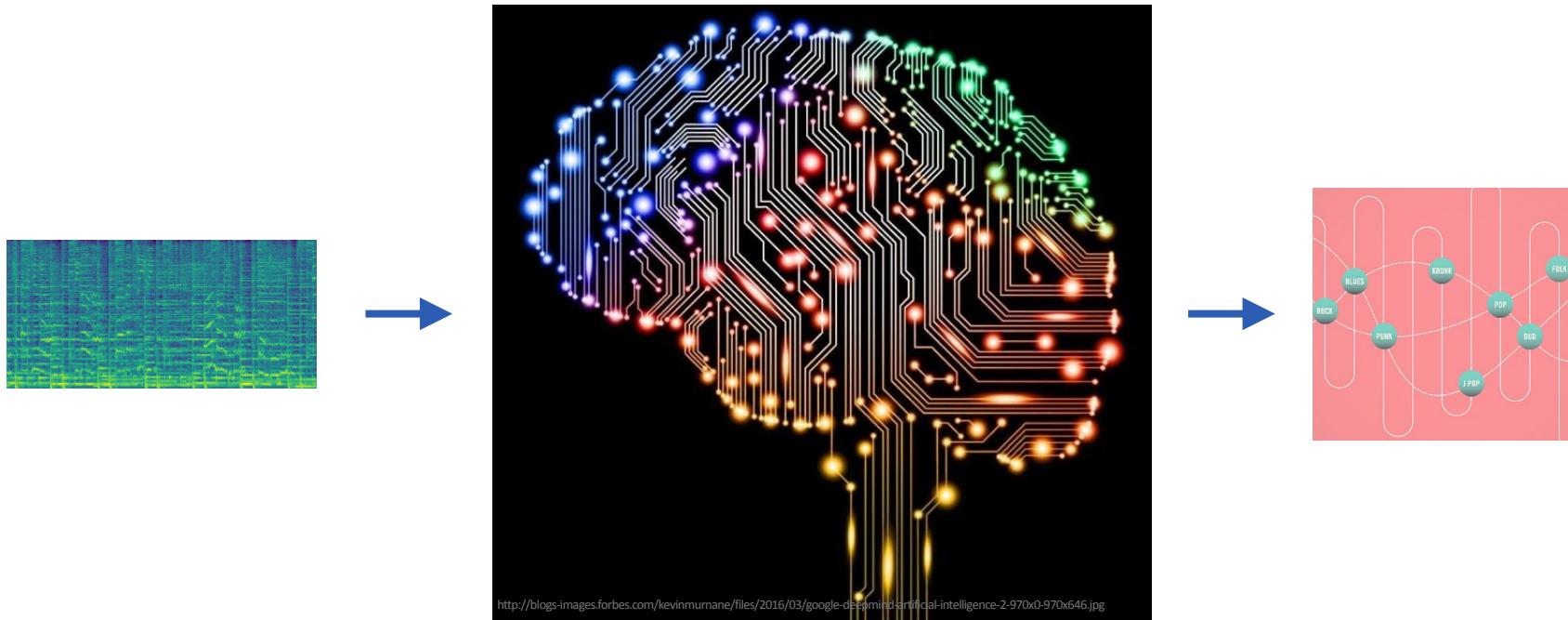
	Artist	Title
Query Track	The Beatles	While My Guitar Gently Weeps
Ranked 1	Paul McCartney	Freedom
Ranked 2	Badly Drawn Boy	What Is It Now
Ranked 3	Jefferson Starship	Fading Lady Light

Recommending Music using the MGP

EXAMPLE

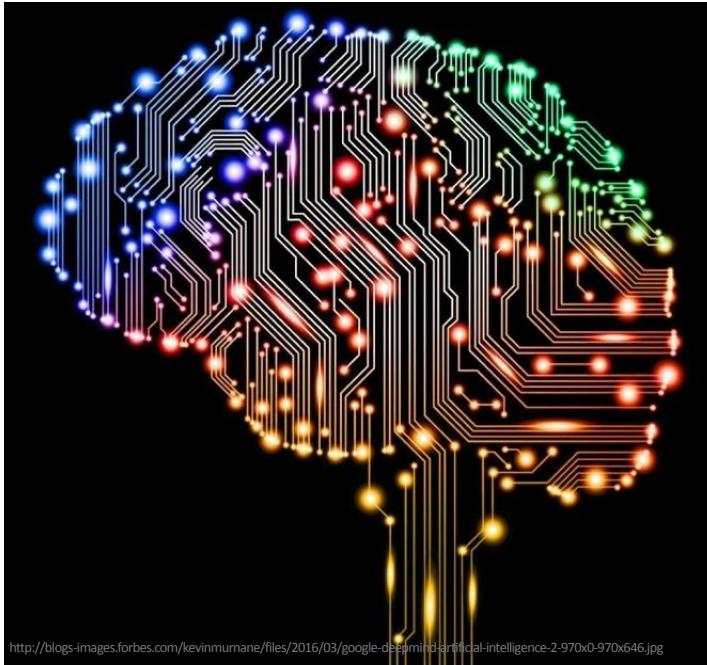
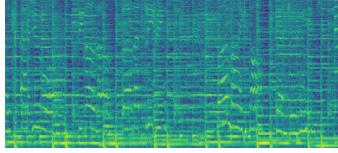
	Artist	Title
Query Track	The Beatles	While My Guitar Gently Weeps
Ranked 1	Paul McCartney	Freedom
Ranked 2	Badly Drawn Boy	What Is It Now
Ranked 3	Jefferson Starship	Fading Lady Light

Estimating MGP Attributes USING DEEP LEARNING



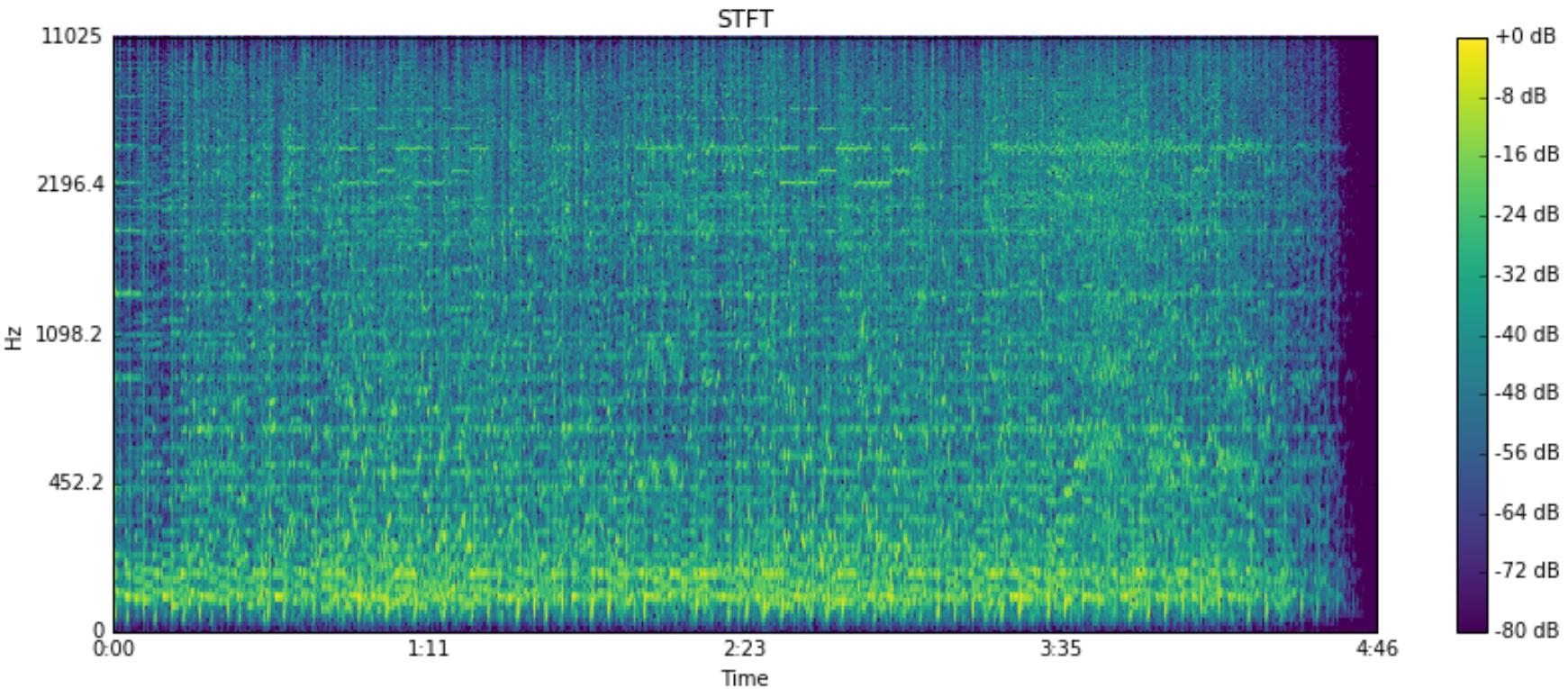
Estimating MGP Piano Attribute

USING DEEP LEARNING



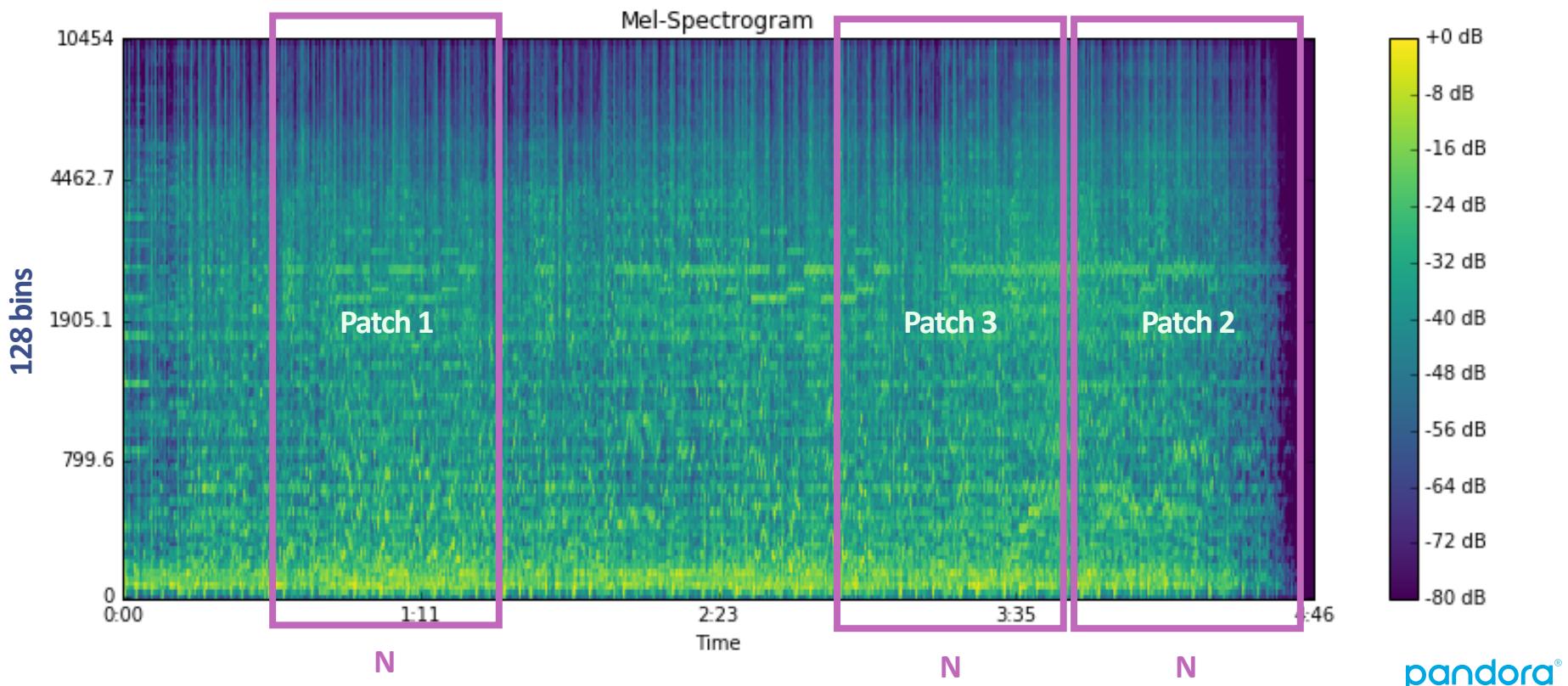
Piano Detector

USING DEEP LEARNING



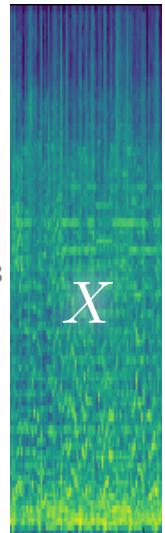
Piano Detector

USING DEEP LEARNING



Piano Detector

DEEP ARCHITECTURE

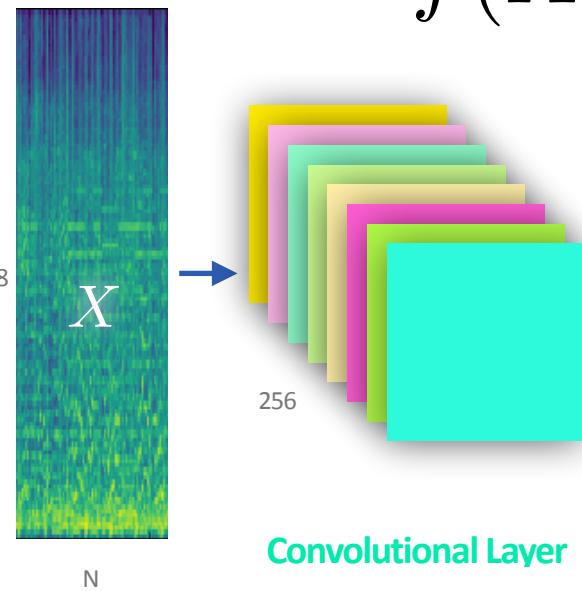


N

Piano Detector

DEEP ARCHITECTURE

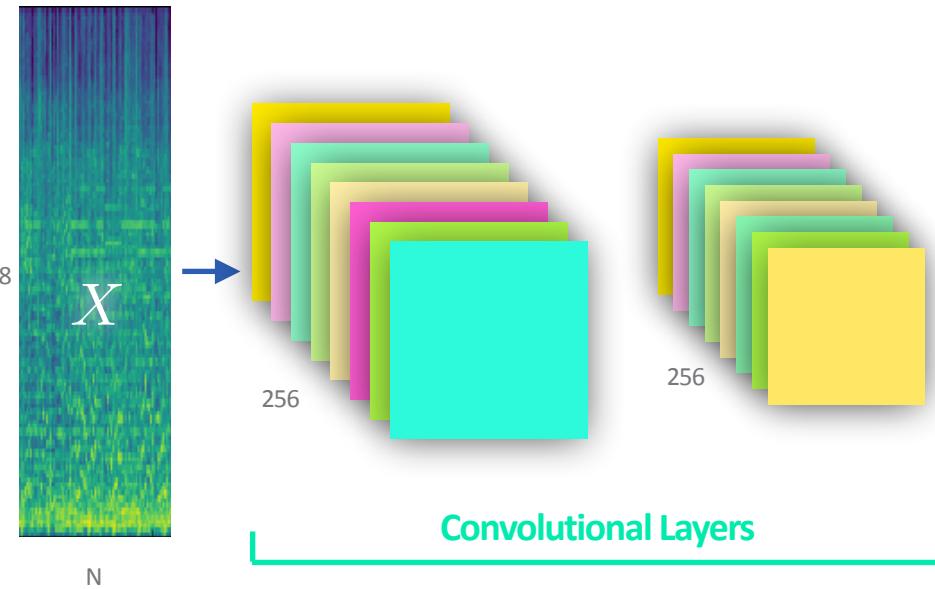
$$f(X, \theta) = \sigma(\mathcal{W} * X + b)$$



Piano Detector

DEEP ARCHITECTURE

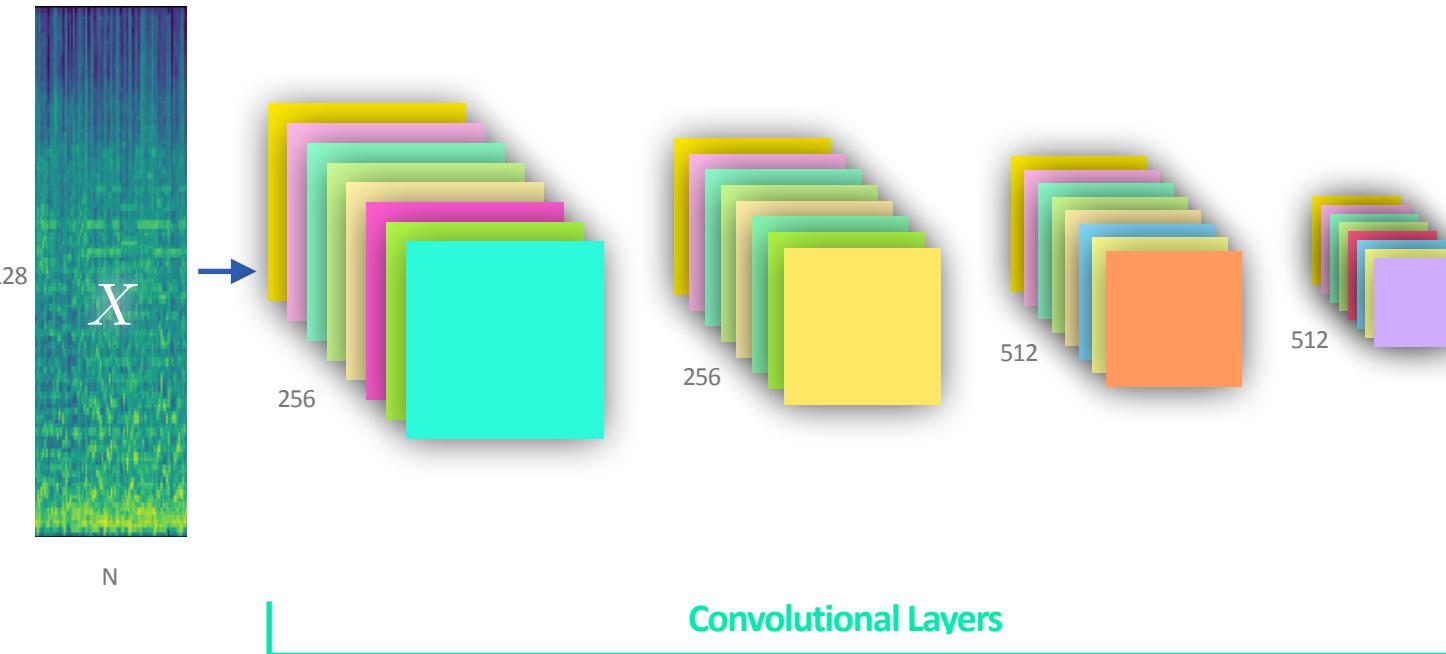
$$f(X, \theta) = \sigma_2(\mathcal{W}_2 * \sigma_1(\mathcal{W}_1 * X + b_1) + b_2)$$



Piano Detector

DEEP ARCHITECTURE

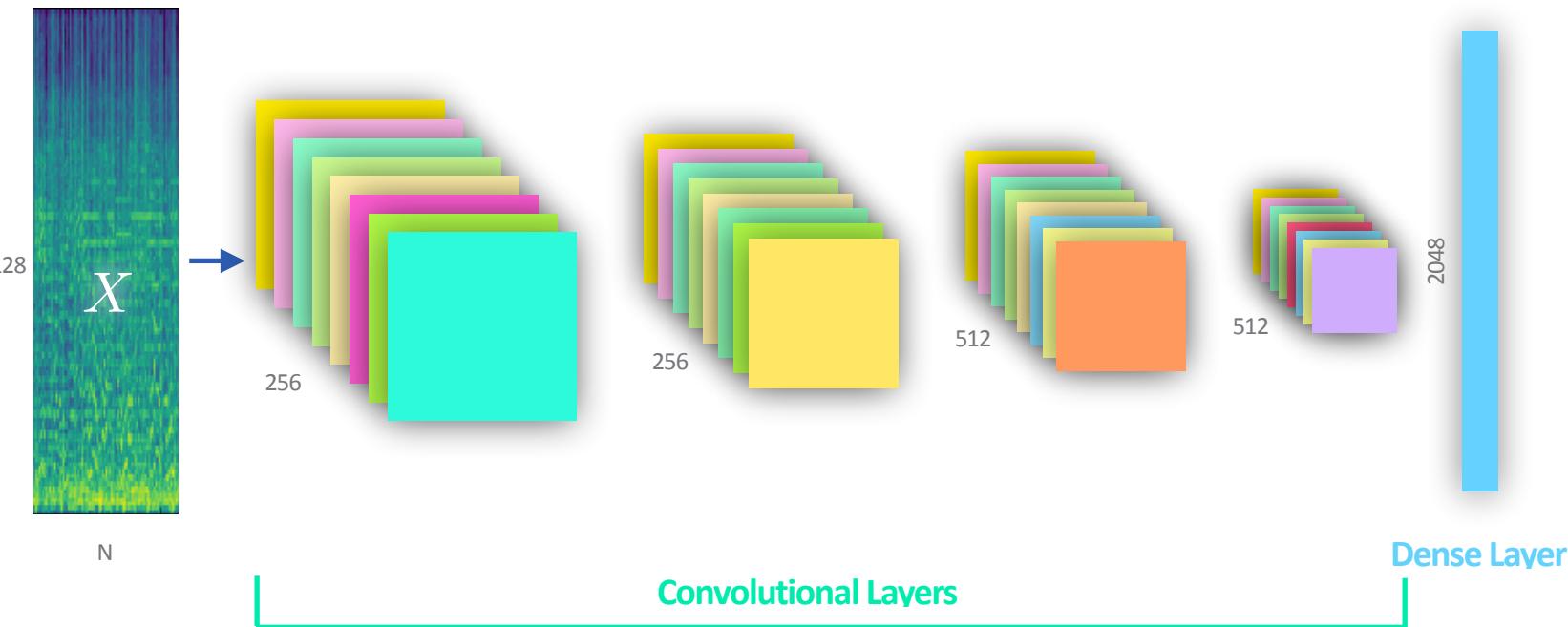
$$f(X, \theta) = \sigma_4(\mathcal{W}_4 * \dots * \sigma_1(\mathcal{W}_1 * X + b_1) + \dots + b_4)$$



Piano Detector

DEEP ARCHITECTURE

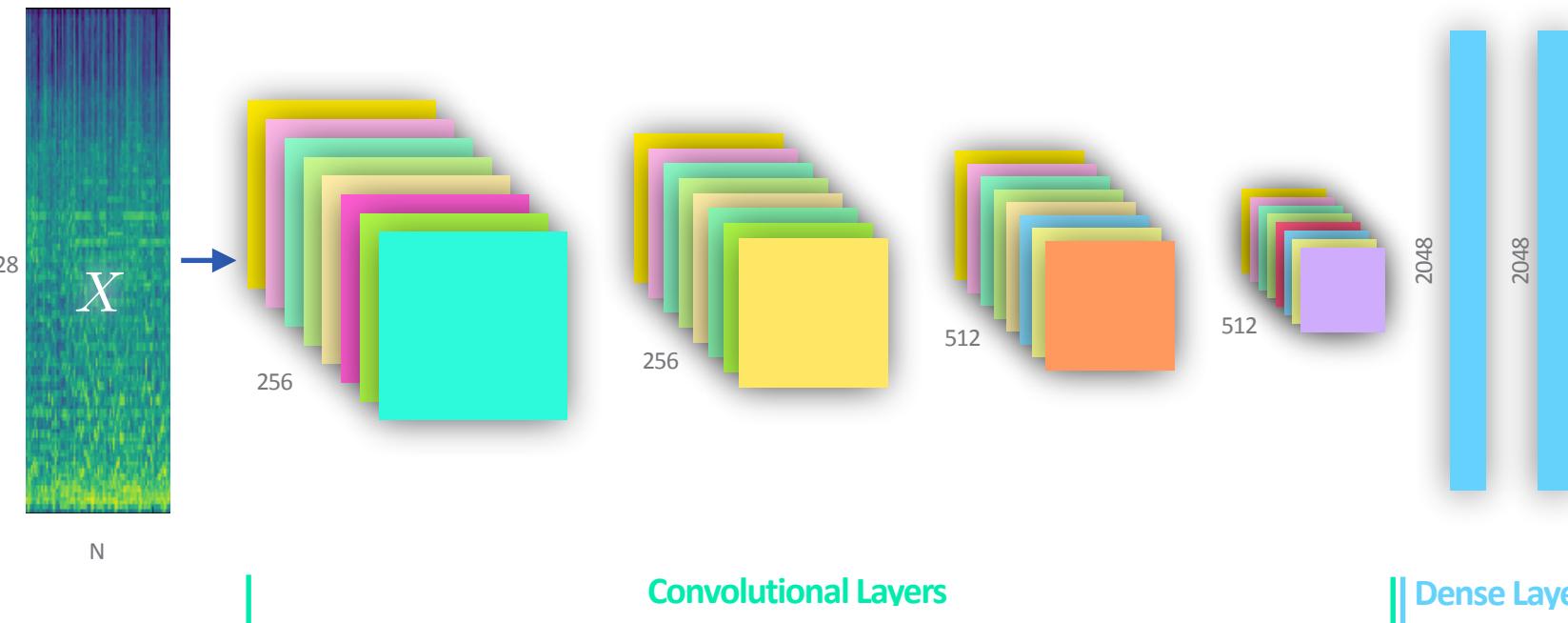
$$f(X, \theta) = \sigma_5(W_1\sigma_4(\mathcal{W}_4 * \dots * \sigma_1(\mathcal{W}_1 * X + b_1) + \dots + b_5)$$



Piano Detector

DEEP ARCHITECTURE

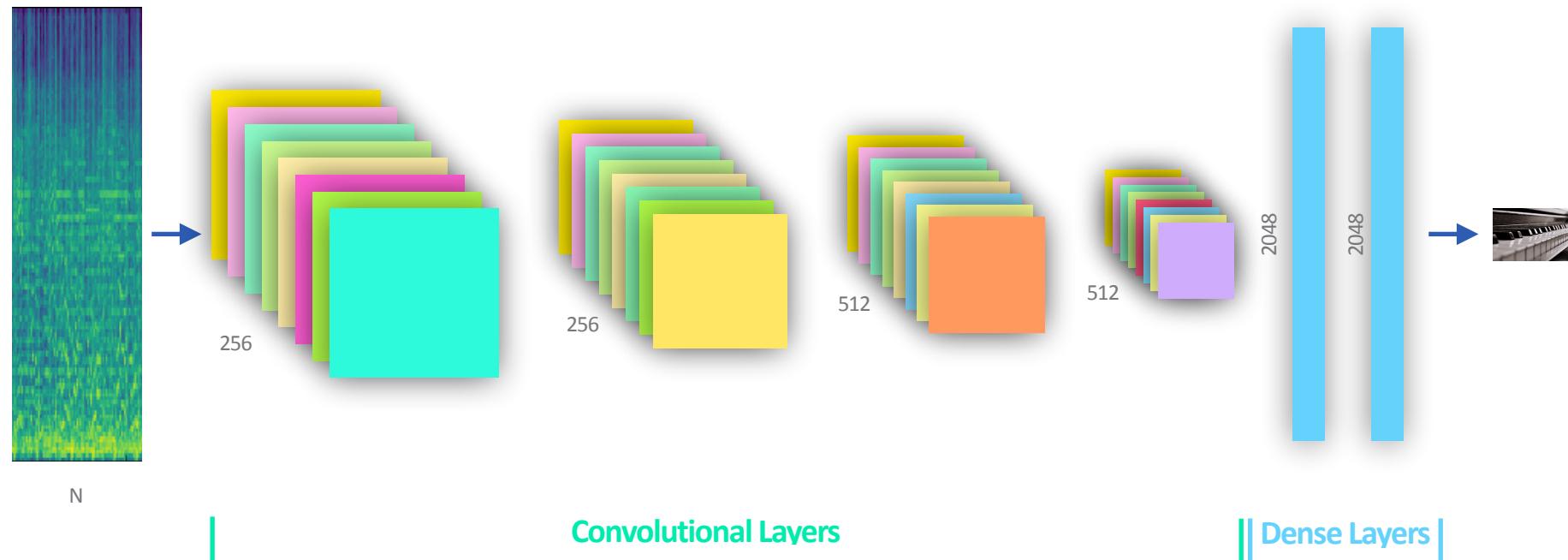
$$f(X, \theta) = \sigma_6(W_2\sigma_5(W_1\sigma_4(\mathcal{W}_4 * \dots * \sigma_1(\mathcal{W}_1 * X + b_1) + \dots b_6)$$



Piano Detector

DEEP ARCHITECTURE

$$f(X, \theta) = \sigma_7(W_3\sigma_6(W_2\sigma_5(W_1\sigma_4(\mathcal{W}_4 * \dots * \sigma_1(\mathcal{W}_1 * X + b_1) + \dots + b_7)$$



Piano Detector

TRAINING

Training

480k Patches

Validation

60k Patches

Test

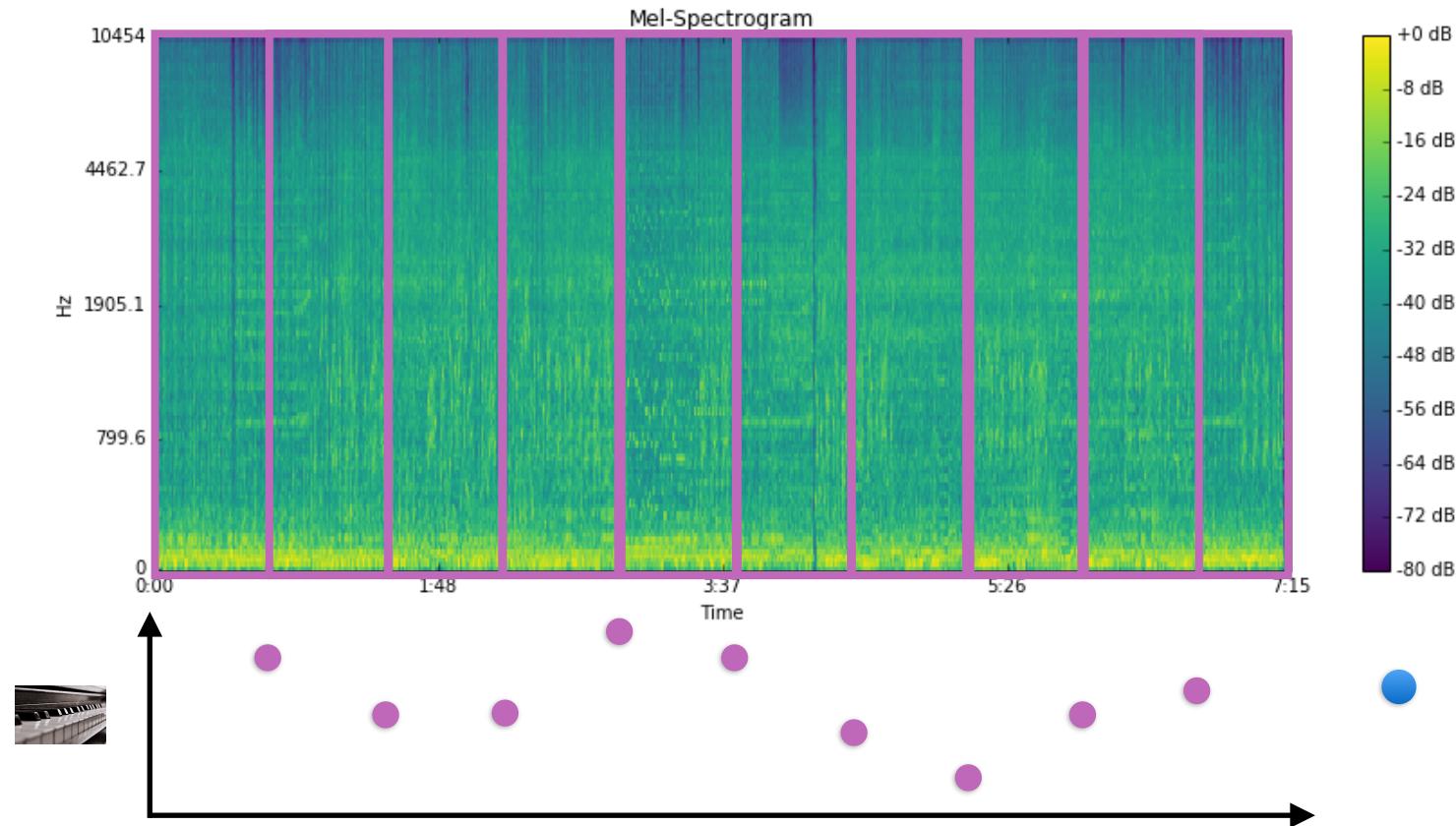
60k Patches

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

(Different model)

Piano Detector

ESTIMATING FULL TRACKS



Piano Detector

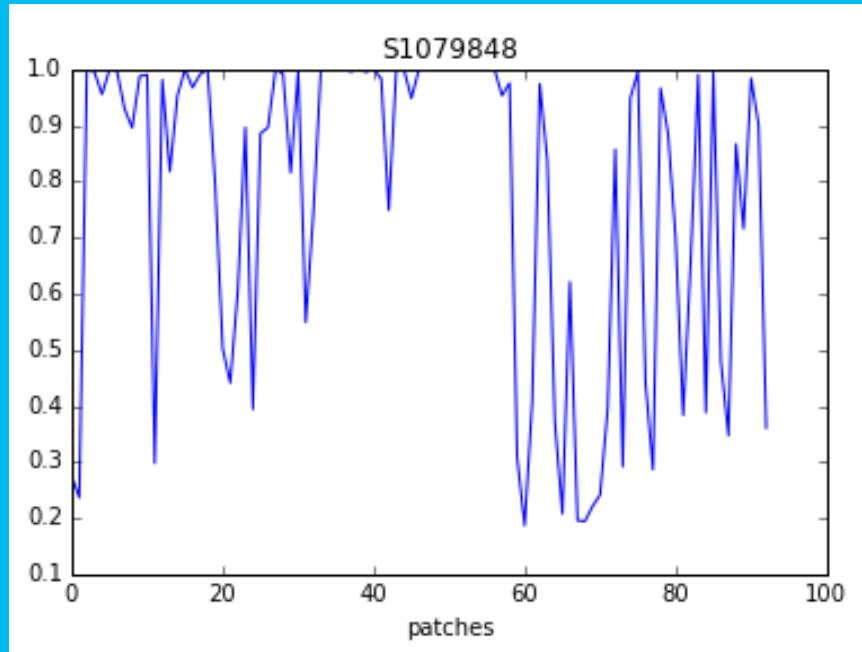
RESULTS

Method	AUC ROC	
Logistic Regression	~85%	
Gradient Boosted Trees	~87.5%	
Deep Learning	90.08%	

Piano Detector

OUTPUT EXAMPLES

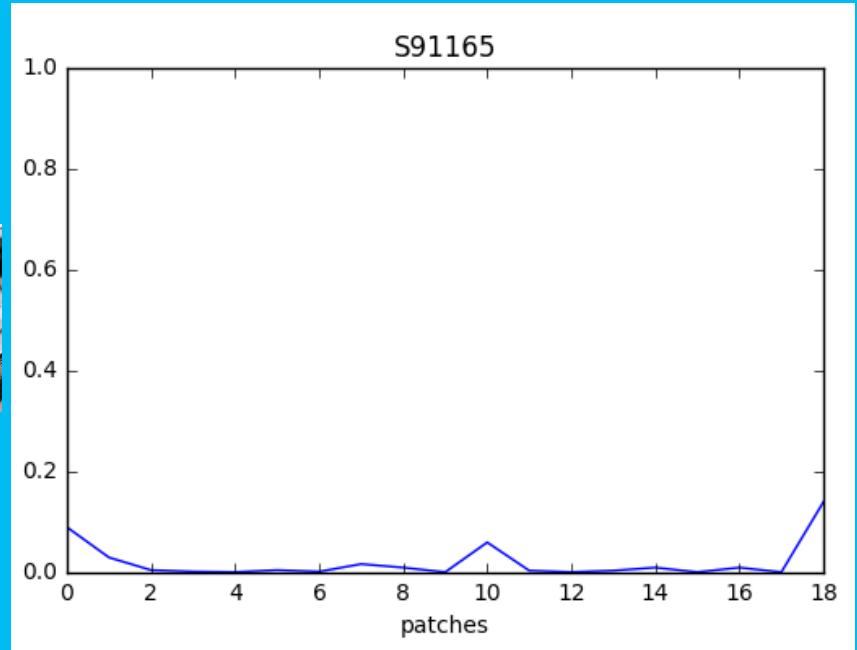
Then, by Dewey Redman Quartet.



Piano Detector

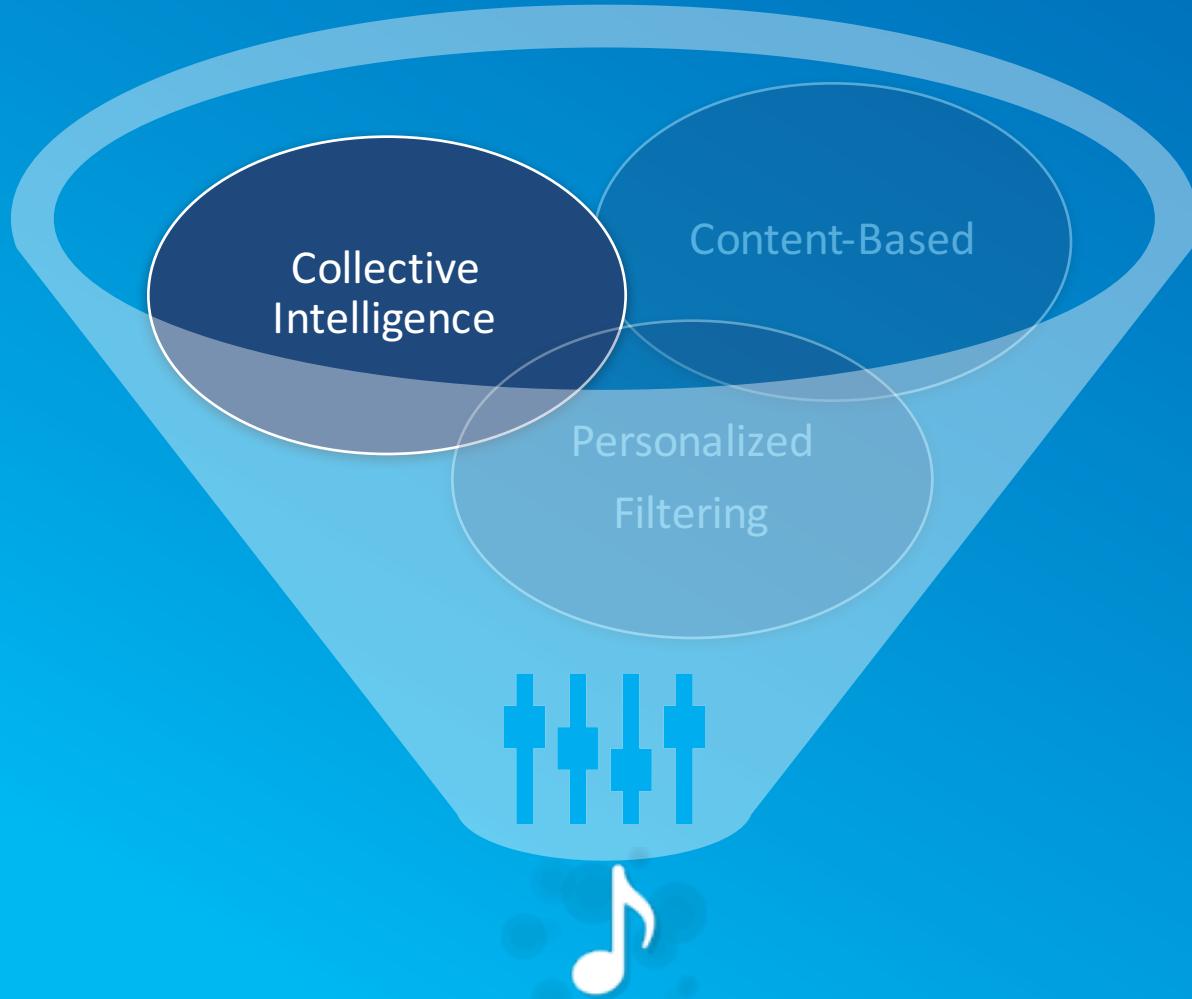
OUTPUT EXAMPLES

Contract On The World Love Jam
by Public Enemy



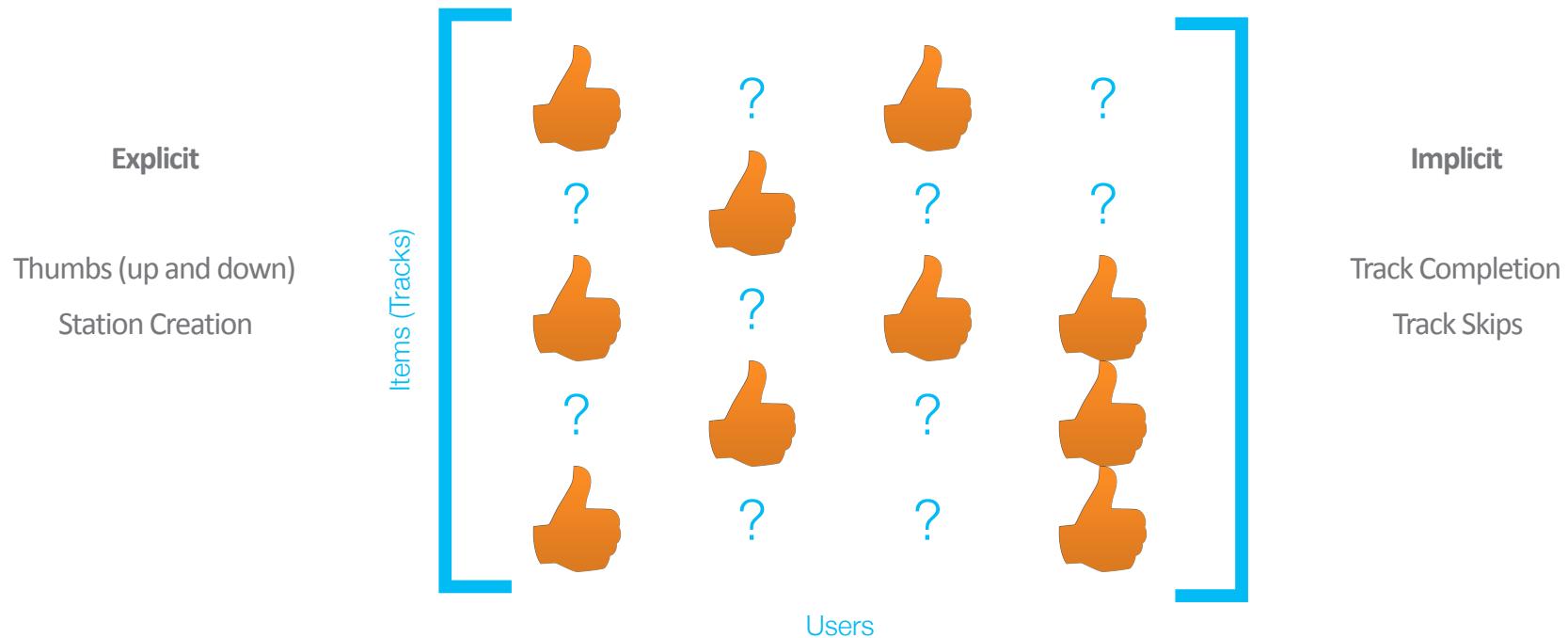
Outline

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



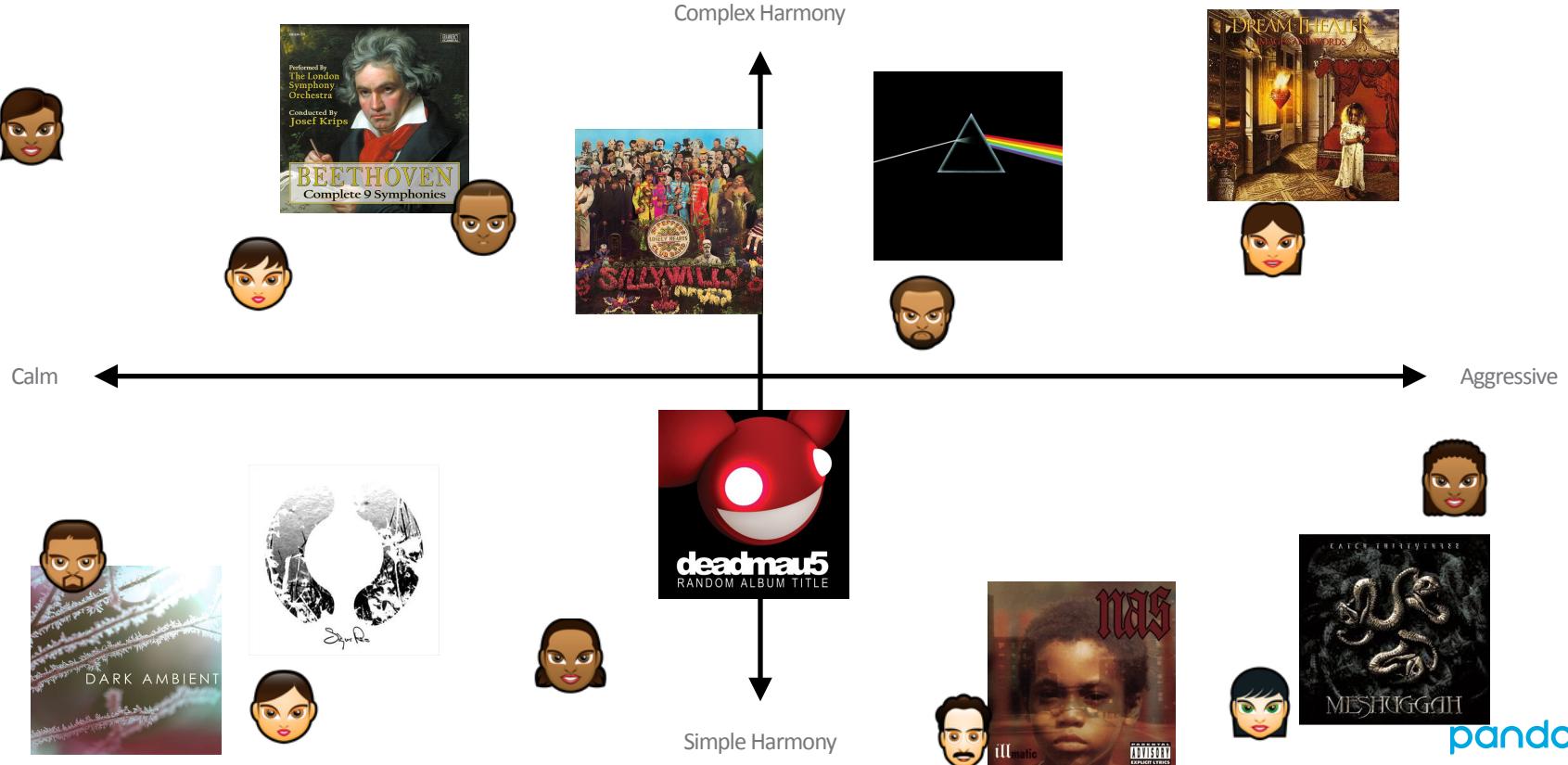
Collaborative Filtering

PROBLEM OVERVIEW



Collaborative Filtering

LATENT FACTORS



Collaborative Filtering

PROBLEM FORMULATION

Given Item i and User u :

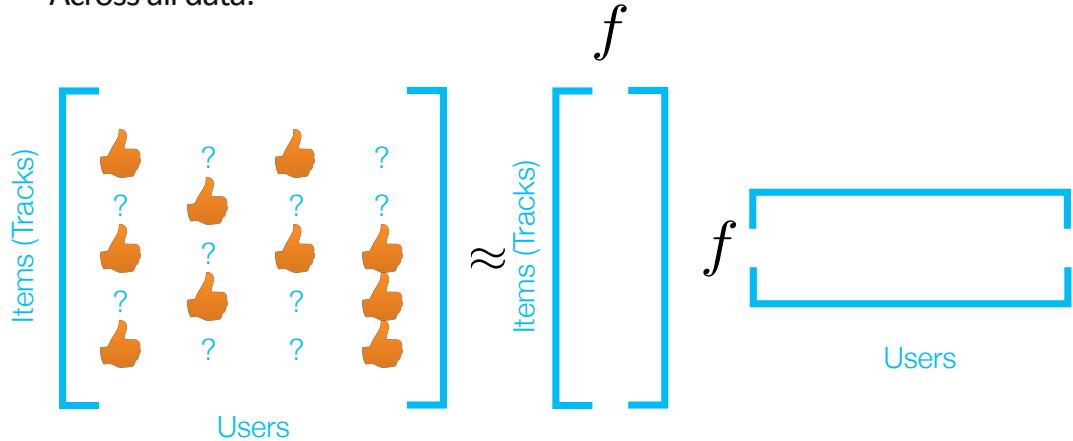
Rating: r_{ui}

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

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

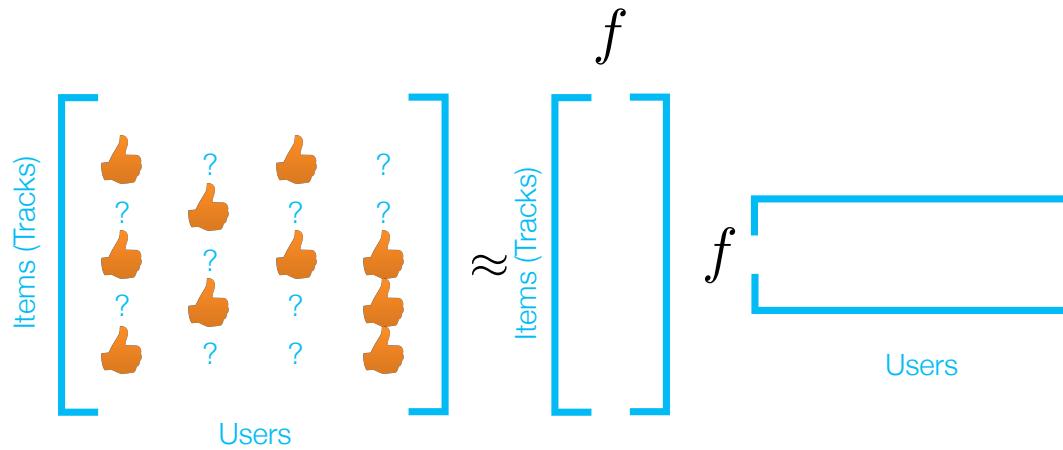
Rating Approximation: $\hat{r}_{ui} = q_i^T p_u$

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

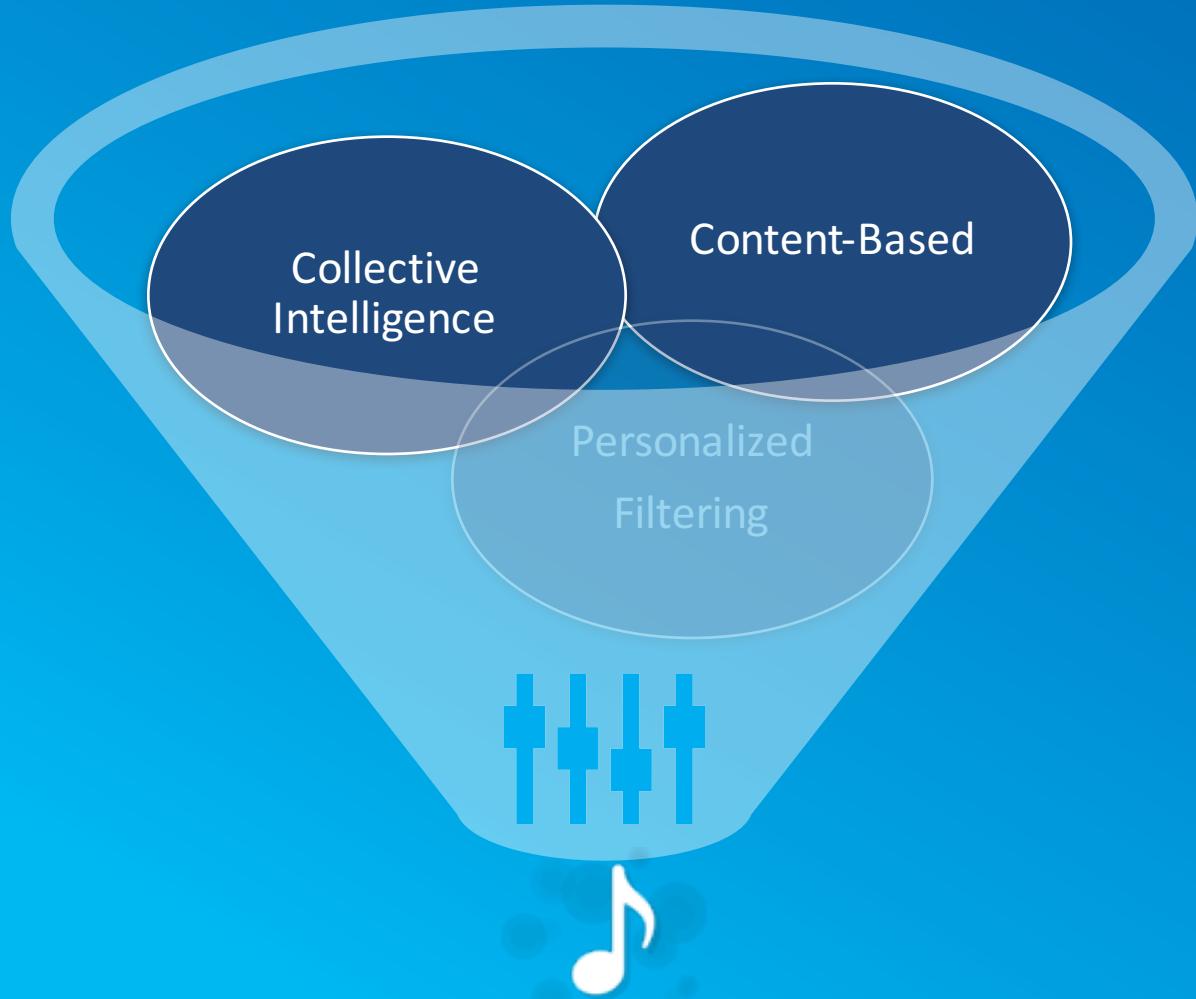
Powerful at matching the right song
with the right listener



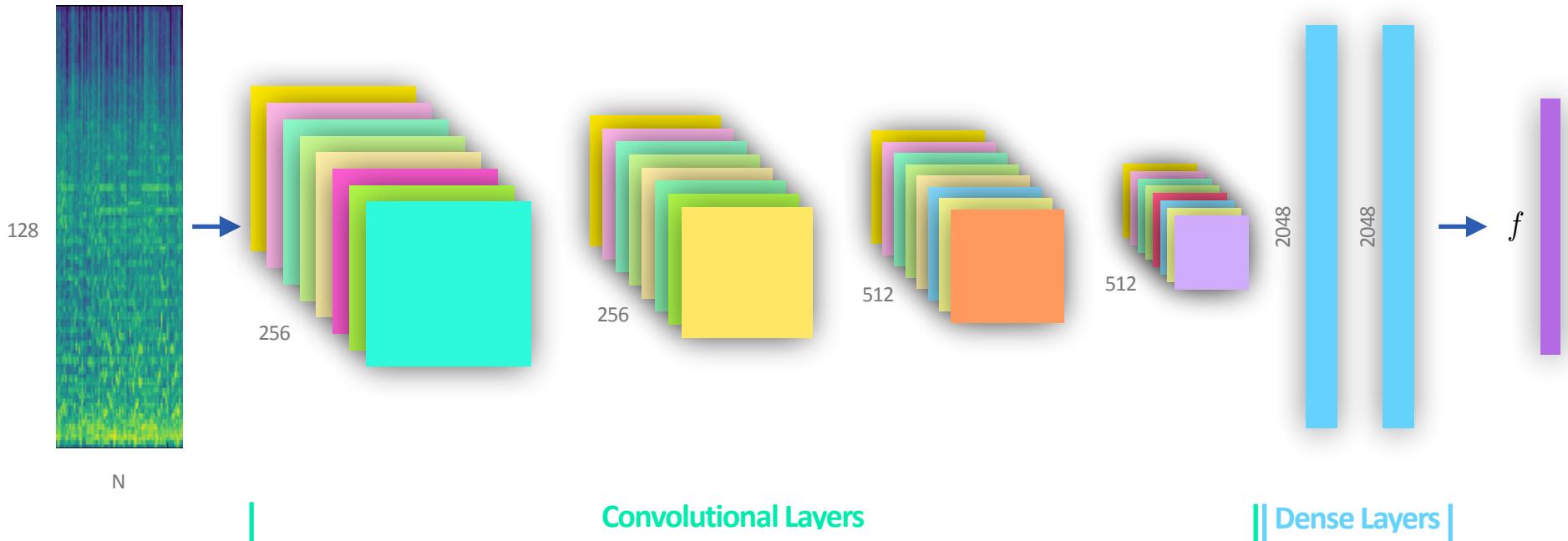
Latent space is generally not interpretable

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

Training

700k Patches

Validation

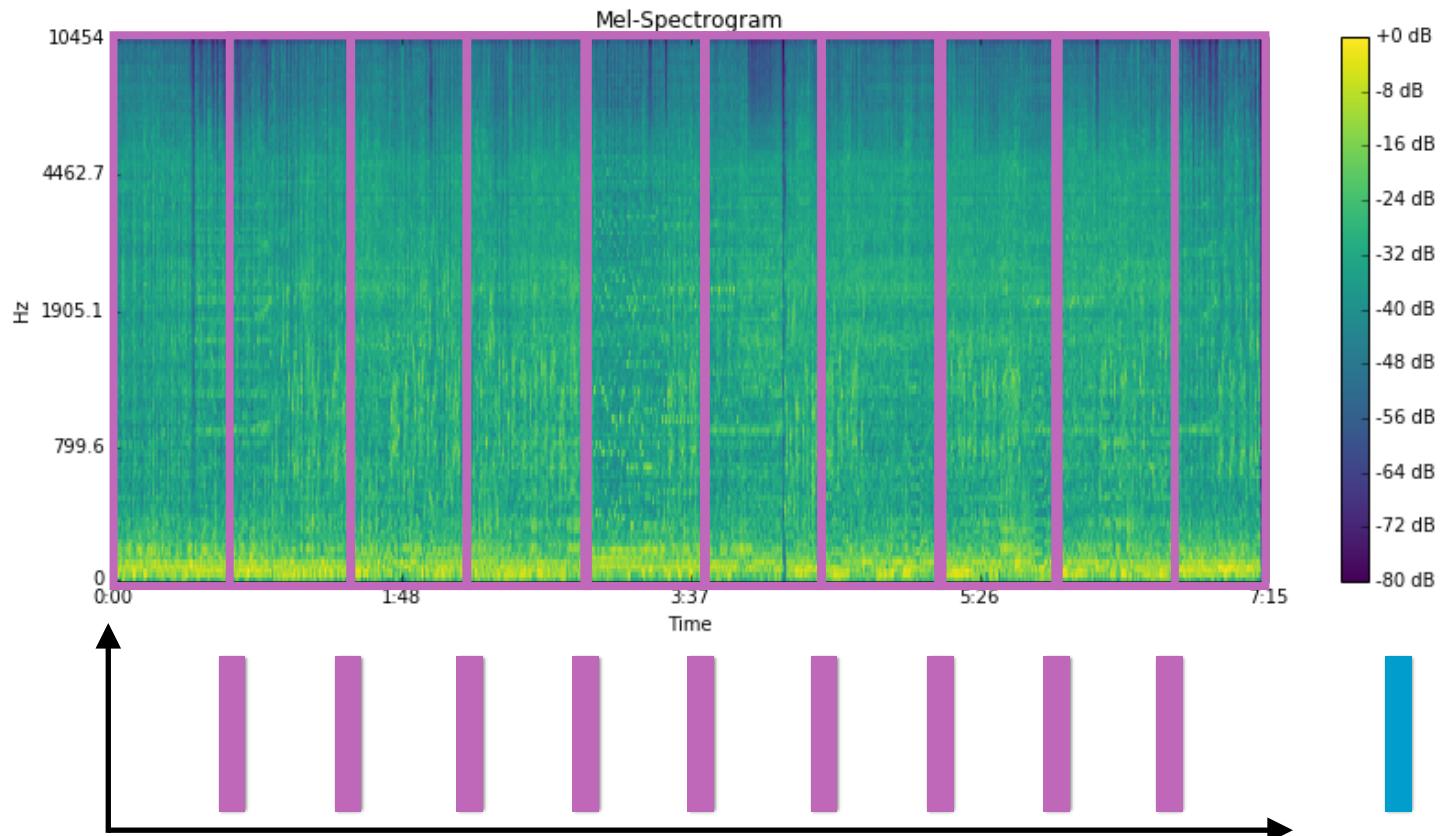
70k Patches

Test

70k Patches

Approximating Item Factors

ESTIMATING FULL TRACKS



Recommending from Learned Factors

EXAMPLE

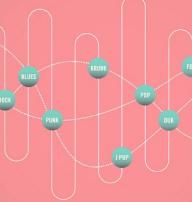
	Artist	Title
Query Track	La Bossa d'Urina	Mama i Papa
Ranked 1	Johnny Cash	I'm Going To Memphis
Ranked 2	Merle Haggard	The Woman Made A Fool Out Of Me
Ranked 3	Waylon Jennings	Lonesome On'ry And Mean



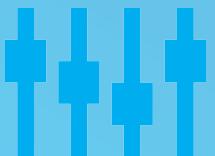
Collective
Intelligence



Content-Based



Personalized
Filtering





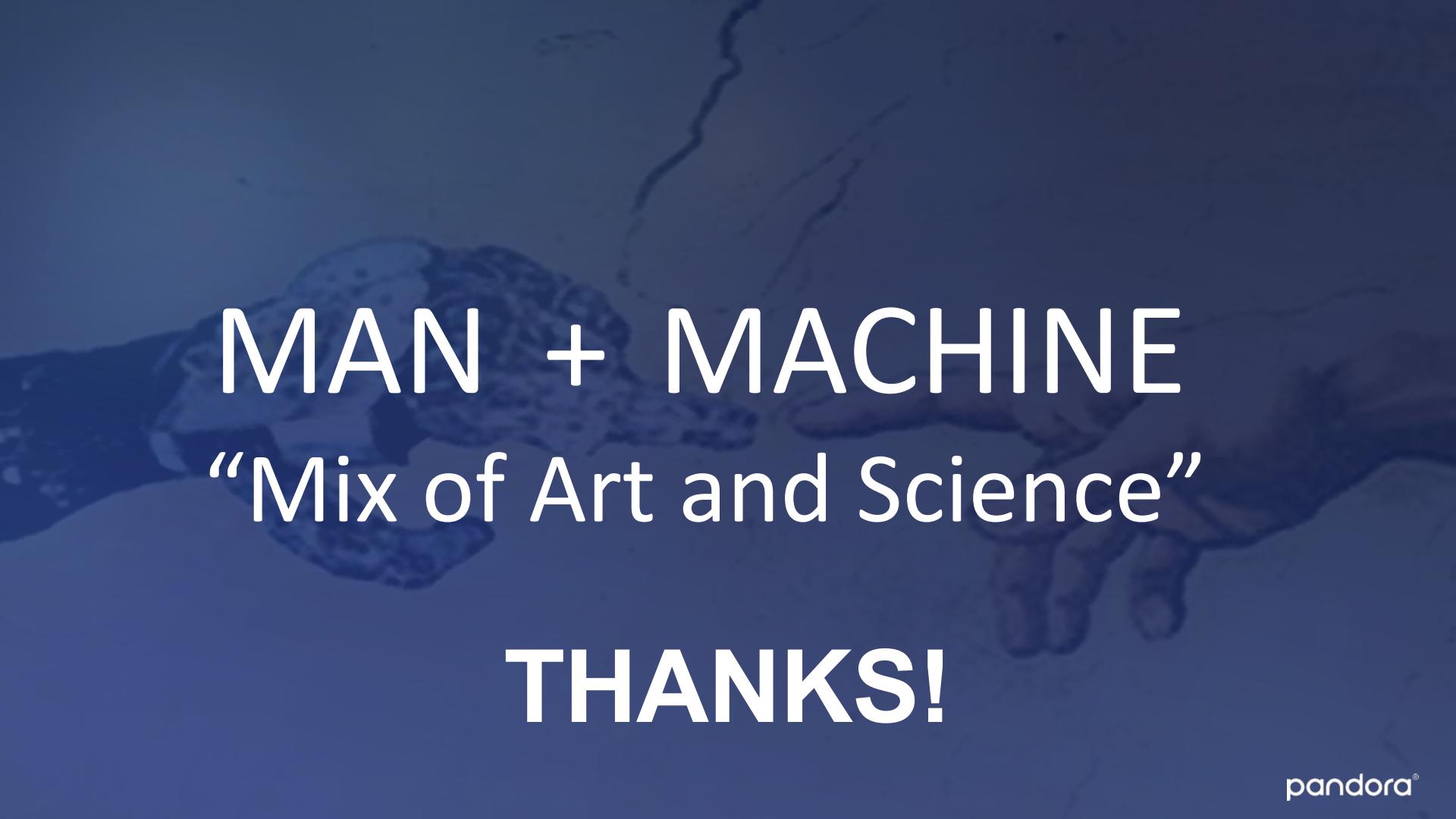
ENSEMBLE OF RECOMMENDERS MAY PRODUCE OPTIMAL RECOMMENDATIONS



MAN vs MACHINE?



MAN + MACHINE



MAN + MACHINE
“Mix of Art and Science”

THANKS!

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/10.1101/10.1101/2015.05.09.005>
- 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>