

# Music Recommendation System

**TEAM – JARVIS**

Aakash Bhatia – 201711031

Krutika Mahavadia – 201711068

Akash Thakrar – 201711038

Chintan Sangani - 201711055

# Model

## 1. New User

- Popularity
- Content Based Filtering - CBF

## 2. Existing User

- Collaborative Filtering - CF
- Content Based Filtering - CBF
- Cosine Similarity
- Personalized Filtering – PF
- Popularity

# Decision Parameter      Rewarding & Penalizing

- Artist
- Album
- Release Year
- Listen Count

- Initial Score – 1.0
- Reward – 5%
- Penalty – 2%

## Existing User

PF	PF	PF	Popularity	Popularity	CF	CF	CBF	CBF	Cosine
----	----	----	------------	------------	----	----	-----	-----	--------

# Popularity

- Create song wise groups,
- Aggregate Listen count per group
- Derive proportion of aggregation with respect to total aggregation.

- Aggregation<sub>Per\_group</sub> :  $\sum_{1 \text{ to } n} (\text{Listen Count})$

- Total Aggregation :  $\sum_{1 \text{ to } n} (\text{Aggregation}_{\text{Per_group}})$

- Proportion<sub>Per\_group</sub> :  $(\text{Aggregation}_{\text{Per_group}} / \text{Total Aggregation}) * 100$

# Collaborative Filtering

- User based collaborative filtering.
- Co-occurrence Matrix
- Jaccard similarity

$$J(A, B) = \frac{|A \cap B|}{|A \cup B|} = \frac{|A \cap B|}{|A| + |B| - |A \cap B|}.$$

$$0 \leq J(A, B) \leq 1.$$

# Content Based Filtering

- Create matrix<sub>User X Album</sub>
- Create matrix<sub>User X Artist</sub>
- Get User row from matrix while login
- Fetch (highest listened artist & album)
- Create list<sub>Artist</sub> -> songs  $\in$  Artist
- Create list<sub>Album</sub> -> songs  $\in$  Album
- Extract Highest listened two songs

# Cosine Similarity

- A -> Attributes of current listened songs. (Like Artist, Album)
- B -> Attributes  $\in$  (All songs – A)

$$\text{similarity} = \cos(\theta) = \frac{A \cdot B}{\|A\| \|B\|} = \frac{\sum_{i=1}^n A_i B_i}{\sqrt{\sum_{i=1}^n A_i^2} \sqrt{\sum_{i=1}^n B_i^2}}$$

# Personalized Filtering

- Personalized\_Score  $\rightarrow \alpha * \beta$ 
  - $\alpha$  = Listen count
  - $\beta$  = Score
- Sort user list according to Personalized\_Score in Descending order
- Fetch ( Top few songs )



# Dataset

	A	B	C	D	E	F	G	H
1		user_id	song_id	listen_count	title	release	artist_name	year
2	4305	8dcb524ff75e0ad0f0a80aaccadf7bbaa3b89a41	SOSXPFU12AB017FE27	1	The Calculation (Album Version)	Far	Regina Spektor	2009
3	7316	94d5bdc37683950e90c56c9b32721edb5d347600	SOGJJON12A67AD9554	7	Why Do I Keep Counting?	Sam's Town	The Killers	2006
4	2787	b61afb42335287239bd40e1dea50d849cbf8a9a9	SOVEKMG12AB018881E	2	Blessa	Causers Of This	Toro Y Moi	2009
5	6738	e3937c7c32f5b68422808a854a4a7a824ee448a5	SOSXPFU12AB017FE27	1	The Calculation (Album Version)	Far	Regina Spektor	2009
6	9123	6beb4699102775dab57aa406c5ea1217c4ff4869	SOPAWLL12A6D4F8B42	1	Something (Radio Edit)	Positiva Presents	Lasgo	2002

	A	B	C	D	E
1	song_id	title	release	artist_name	year
2	SOQMMHC12AB0180CB8	Silent Night	Monster Ballad	Faster Pussy cat	2003
3	SOVFVAK12A8C1350D9	Tanssi vaan	KarkuteillÄä	Karkkiautomaatt	1995
4	SOGTUKN12AB017F4F1	No One Could Ever	Butter	Hudson Mohawk	2006
5	SOBNYVR12A8C13558C	Si Vos QuerÄ@s	De Culo	Yerba Brava	2003
6	SOHSBXH12A8C13B0DF	Tangle Of Aspens	Rene Ablaze Pr	Der Mystic	0

Thank you