Music Recommender using RNeo4j

Jeremy O'Brien and Kavya Beheraj

Song recommender system

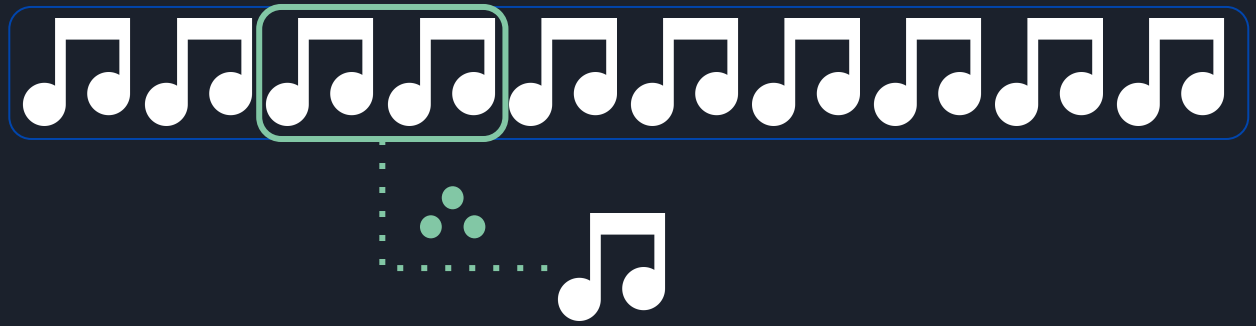
Using nearest-neighbor collaborative filtering

Research question: How reliable is the proportion of times that a user listened to a song in predicting the strength of their preference for another song?

Song recommender system

Using nearest-neighbor collaborative filtering

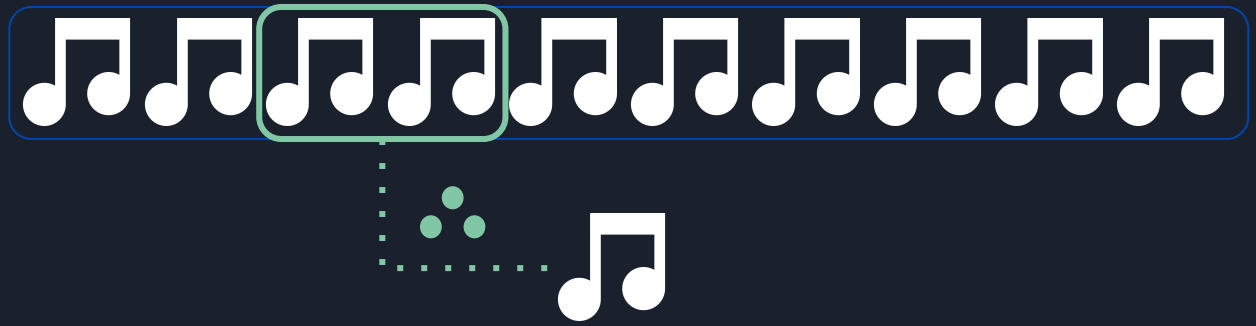
Research question: How reliable is the proportion of times that a user listened to a song in predicting the strength of their preference for another song?



Song recommender system

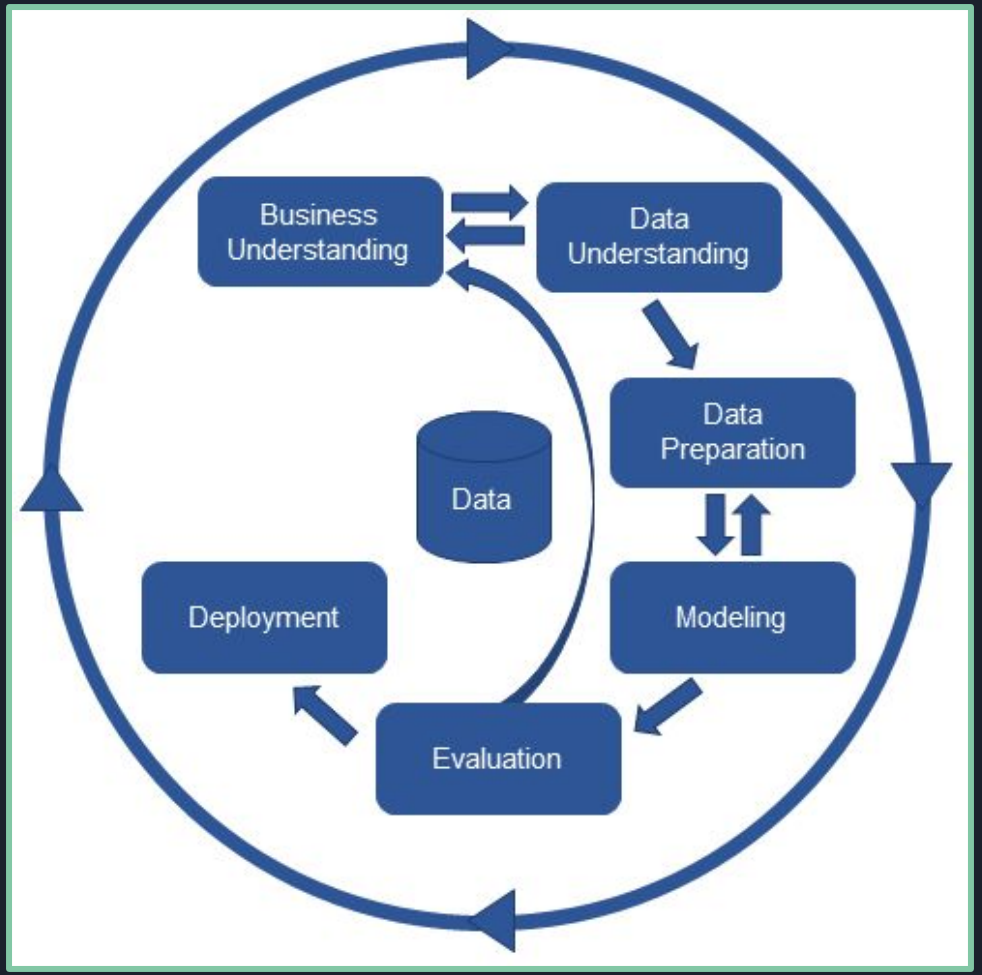
Using nearest-neighbor collaborative filtering

Research question: How reliable is the proportion of times that a user listened to a song in predicting the strength of their preference for another song?

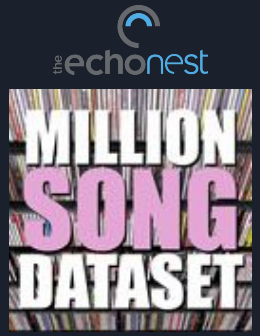


Generalized: How reliable are implicit ratings in predicting user preference?

CRISP
Data
Mining
Approach

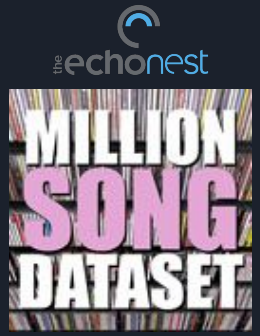


Our Toolbox



Open-source, large-scale song- and listener-level data

Our Toolbox

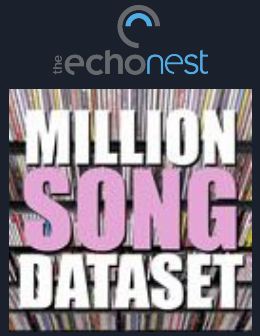


Open-source, large-scale song- and listener-level data

Data cleaning & preparation

Exploratory data analysis

Our Toolbox



Open-source, large-scale song- and listener-level data

Data cleaning & preparation

Exploratory data analysis

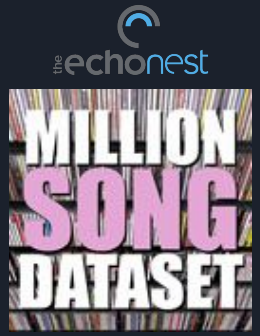
Nodal relationships

Similarity computation

Recommender logic

- recommenderLab -

Our Toolbox



Open-source, large-scale song- and listener-level data

Data cleaning & preparation

Exploratory data analysis

Nodal relationships

Similarity computation

Recommender logic

- recommenderLab -

User-friendly front-end GUI

- stretch goal -

Million Song Dataset

User

Song

“Rating” (0-100)

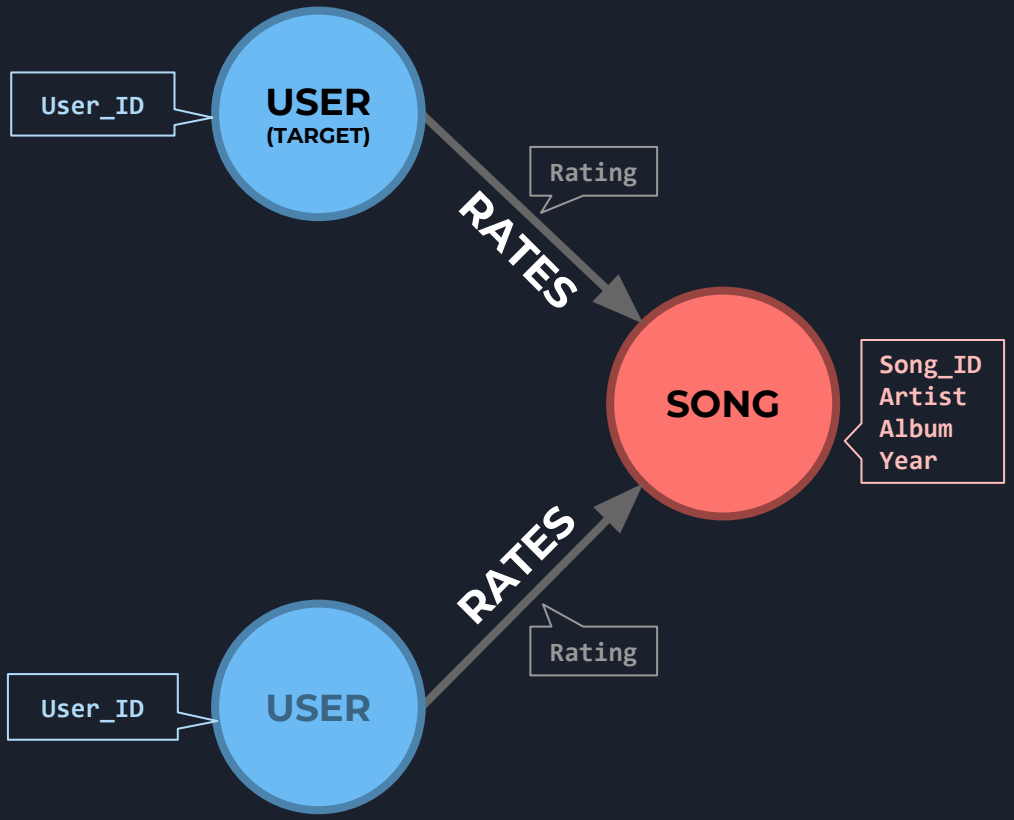
(Listen Count ÷ Sum of Song Listens) * 100

user_id	song_id	listen_count	title	release	artist_name	year	number_songs	mean_listen_count	sum_listen_count	rating
969cc6fb74e076a68e36a04409cb9d3765757508	SOABRAB12A6D4F7AAF	2	A Pain That Im Used To	Playing The Angel	Depeche Mode	2005	19	3.31579	63	3.17
969cc6fb74e076a68e36a04409cb9d3765757508	SOAOQFD12A6D4FAAA9	1	Medicating	Trainwreck	Boys Night Out	2005	19	3.31579	63	1.59
969cc6fb74e076a68e36a04409cb9d3765757508	SOBFPJC12A58A7D1AB	9	Crystal Blue Persuasion	Anthology	Tommy James And The Shondells	1968	19	3.31579	63	14.29
969cc6fb74e076a68e36a04409cb9d3765757508	SOBZZDU12A6310D8A3	2	Don't Dream It's Over	Recurring Dream_ Best Of Crowded House (Domestic Only)	Crowded House	1986	19	3.31579	63	3.17

Neo4j Recommender

STEPS

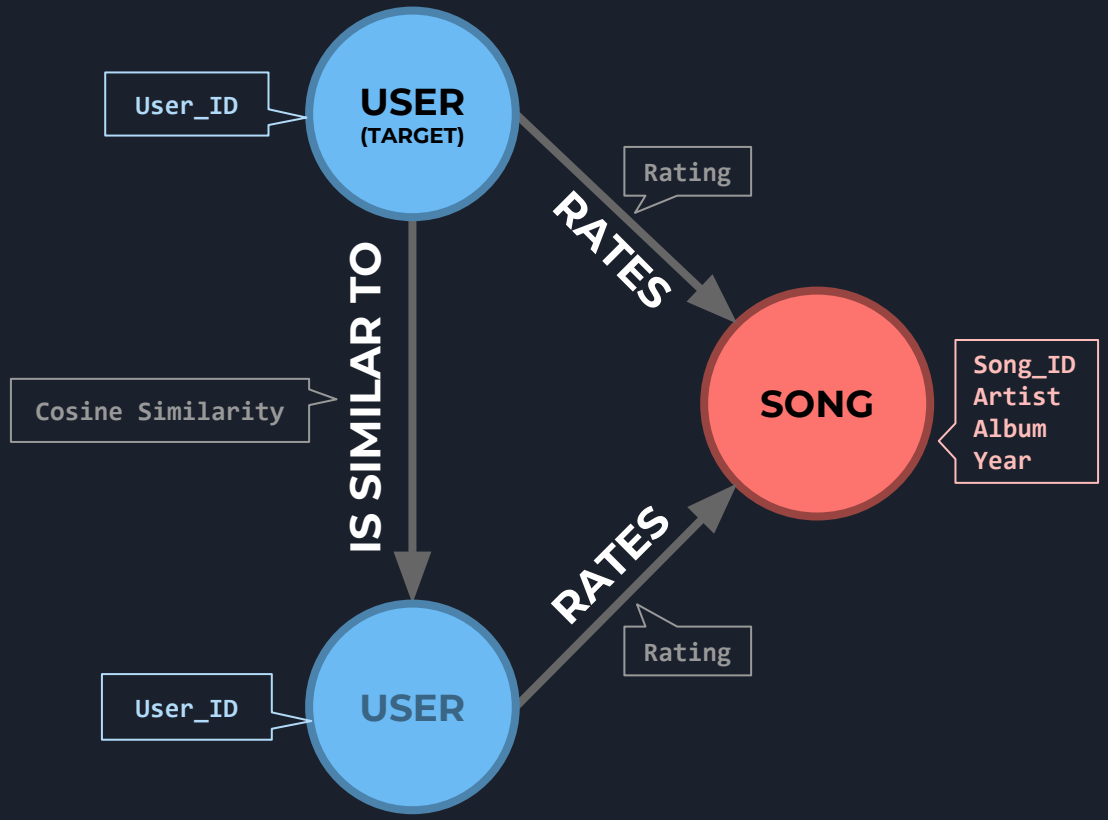
1. Create nodes, set relationships, and add properties



Neo4j Recommender

STEPS

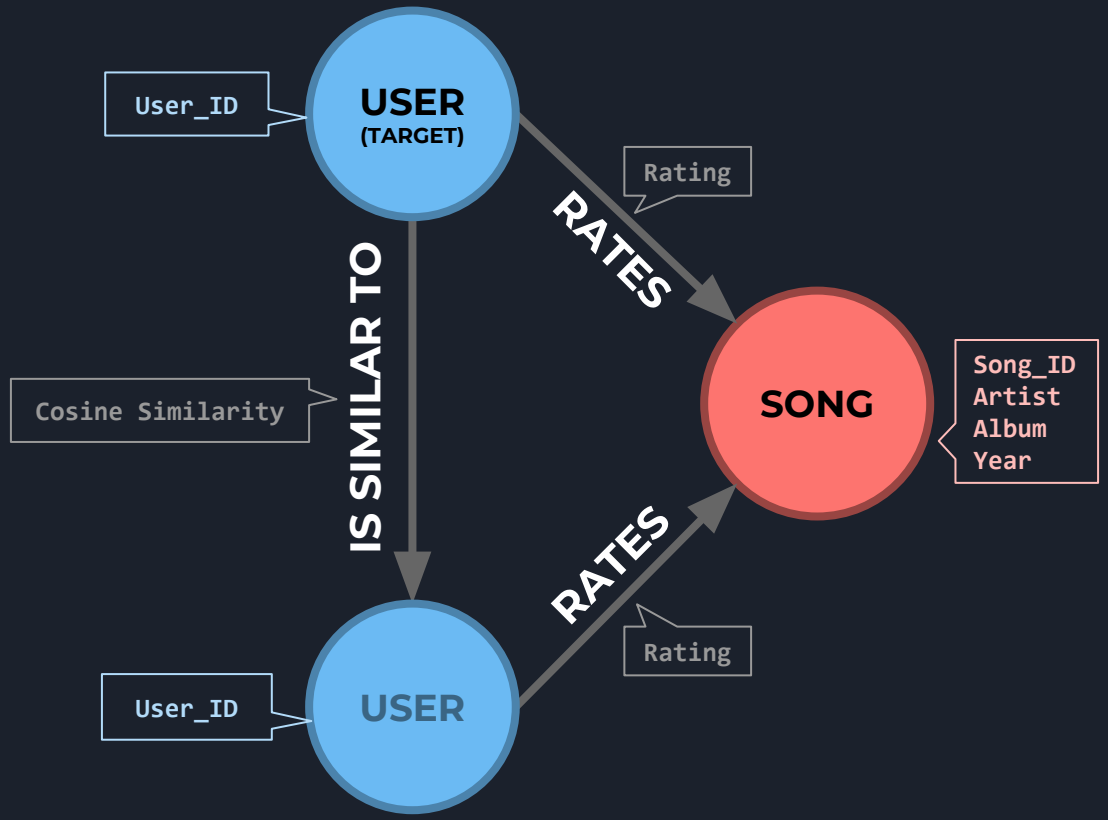
1. Create nodes, set relationships, and add properties
2. Determine nearest neighbors from user cosine similarity

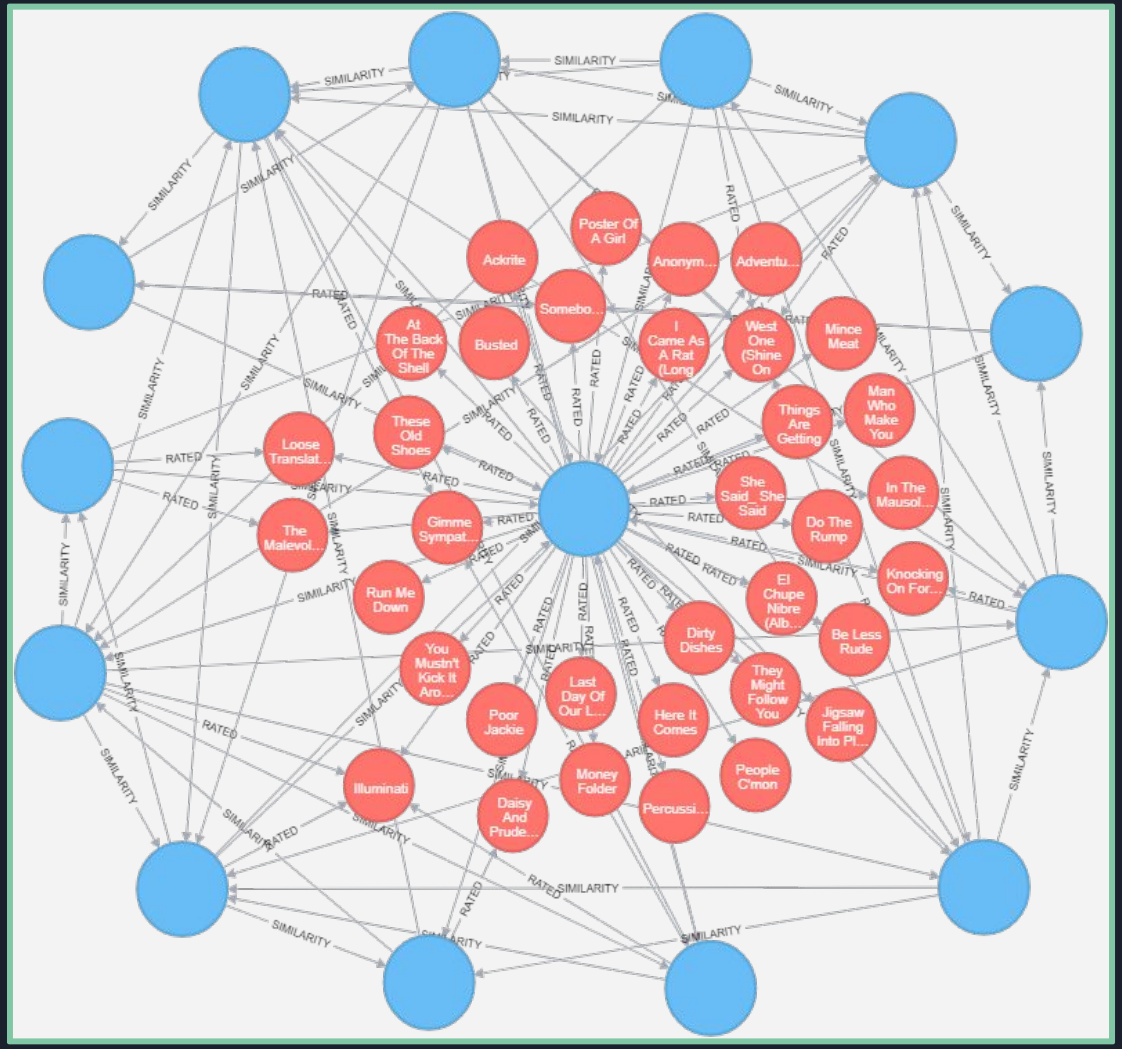


Neo4j Recommender

STEPS

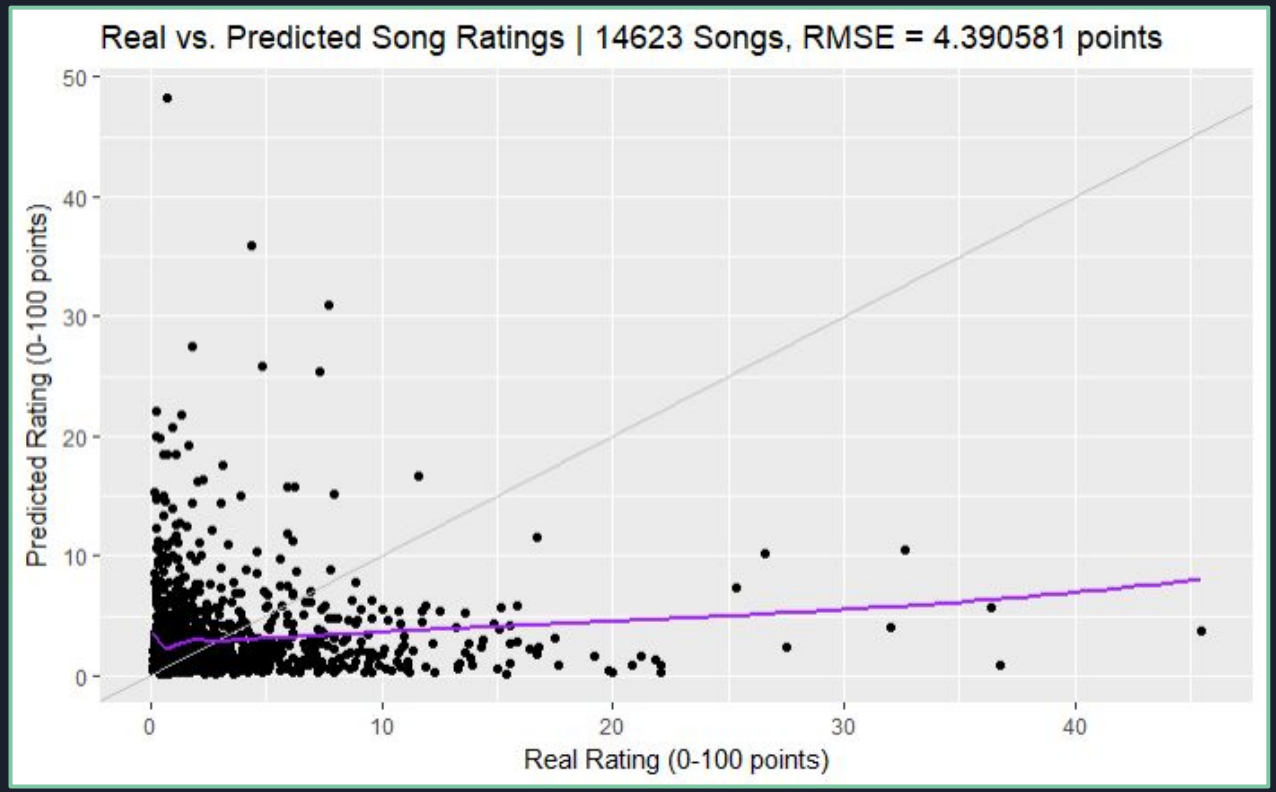
1. Create nodes, set relationships, and add properties
2. Determine nearest neighbors from user cosine similarity
3. Pull top-rated songs from nearest neighbors as recommendation




SIMILARITY
RATED
Poster Of A Girl
Ackrite
Anonym...
Adventu...
Somebo...
At The Back Of The Shell
Busted
I Came As A Rat (Long
West One (Shine On
Mince Meat
Man Who Make You
Things Are Getting
These Old Shoes
Loose Translat...
She Said_ She Said
In The Mausol...
Do The Rump
The Malevol...
Gimme Sympat...
Knocking On For...
Run Me Down
El Chupe Nibre (Alb...
Be Less Rude
Dirty Dishes
You Mustn't Kick It Aro...
Last Day Of Our L...
They Might Follow You
Poor Jackie
Here It Comes
Jigsaw Falling Into Pl...
Illuminati
Money Folder
People C'mon
Daisy And Prude...
Percussi...

Recommender Performance

Root mean square error from 3 random samples of 100 users



GUI Proof-of-Concept

Leveraging scenario design for better UX with RShiny

> **Goal:** reduce friction and improve user interaction with MSD recommender

GUI Proof-of-Concept

Leveraging scenario design for better UX with RShiny

Goal: reduce friction and improve user interaction with MSD recommender

MVP: potential further enhancements

GUI Proof-of-Concept

Leveraging scenario design for better UX with RShiny

Goal: reduce friction and improve user interaction with MSD recommender

Challenge: non-trivial to integrate Neo4j and RShiny

MVP: potential further enhancements

GUI Proof-of-Concept

Leveraging scenario design for better UX with RShiny

Goal: reduce friction and improve user interaction with MSD recommender

Challenge: non-trivial to integrate Neo4j and RShiny

MVP: potential further enhancements

Caveat: POC not web-enabled, just sandbox

Music Recommender

Kavya Beheraj & Jeremy O'Brien

Please give us your name, select three songs, and rate them on a scale of 1-10

Your name

Jeremy

First song

Thong Song

0 100

Second song

Comfort Eagle

0 75 100

Third song

Holiday

0 63 100

Here we go!

Here are the songs you rated:

Title	Rating
Thong Song	100
Comfort Eagle	75
Holiday	63

Based on your ratings, here are some songs you might also like:

Title	Artist
Deathzone	Apocalyptica
On My Way Back Home	Band Of Horses
Personal Jesus	Depeche Mode
ObservÃ¡ndonos (SatÃ©lites)	Soda Stereo
Savin' Me	Nickelback

Music Recommender

Kavya Beheraj & Jeremy O'Brien

Please give us your name, select three songs, and rate them on a scale of 1-10

Your name

Jeremy

First song

Thong Song

0 100

Second song

Comfort Eagle

0 75 100

Third song

Holiday

0 63 100

Here we go!

Here are the songs you rated:

Title	Rating
Thong Song	100
Comfort Eagle	75
Holiday	63

Based on your ratings, here are some songs you might also like:

Title	Artist
Deathzone	Apocalyptica
On My Way Back Home	Band Of Horses
Personal Jesus	Depeche Mode
ObservÃ¡ndonos (SatÃ©lites)	Soda Stereo
Savin' Me	Nickelback

Music Recommender

Kavya Beheraj & Jeremy O'Brien

Please give us your name, select three songs, and rate them on a scale of 1-10

Your name

Jeremy

First song

Thong Song

0 100

Second song

Comfort Eagle

0 75 100

Third song

Holiday

0 63 100

Here we go!

Here are the songs you rated:

Title	Rating
Thong Song	100
Comfort Eagle	75
Holiday	63

Based on your ratings, here are some songs you might also like:

Title	Artist
Deathzone	Apocalyptica
On My Way Back Home	Band Of Horses
Personal Jesus	Depeche Mode
ObservÃ¡ndonos (SatÃ©lites)	Soda Stereo
Savin' Me	Nickelback

Music Recommender

Kavya Beheraj & Jeremy O'Brien

Please give us your name, select three songs, and rate them on a scale of 1-10

Your name

Jeremy

First song

Thong Song

0 100

Second song

Comfort Eagle

0 75 100

Third song

Holiday

0 63 100

Here we go!

Here are the songs you rated:

Title	Rating
Thong Song	100
Comfort Eagle	75
Holiday	63

Based on your ratings, here are some songs you might also like:

Title	Artist
Deathzone	Apocalyptica
On My Way Back Home	Band Of Horses
Personal Jesus	Depeche Mode
ObservÃ¡ndonos (SatÃ©lites)	Soda Stereo
Savin' Me	Nickelback

Summary

Question

Can a recommender system built in R with a graph database use an implicit measure of preference to reliably predict song ratings?

Result

Our recommender system predicted song ratings for a sample set of users that were within ~5 points of their real rating.

Conclusion

An implicit rating can reliably signal strength of user preference and serve as a basis for useful recommendations.