



أكاديمية سدايا
SDAIA Academy



Elyas Maghrabi

Abstract

Recommendation Systems are very standard and popular all over the web. Companies like Netflix, Spotify, Apple, Pandora, and LinkedIn leverage the recommender systems to help users find new items quickly, creating a delightful user experience while driving incremental revenue.

ABOUT THE DATA

The Million Song Dataset, a collaborative project between The Echo Nest and LabROSA, is a freely available collection of audio features and meta data for a million modern popular music tracks.

Data Description



The Echo Nest, 2011

Size: 280GB (Subset)

1,000,000 unique tracks ID

Song Metadata, Artist similarity, Artist tags SQLite, Text Files

[Main Dataset Link](#)



musiXmatch

Size: 70MB

779K matches between of musiXmatch ID & MSD ID 210,519 BOW for training & 27,143 BOW for testing SQLite & Text Files

[Complementary Dataset Link](#)



Test Profile

Size: 488MB

48M user-song-play count triplets 1M unique users

380K unique songs

Tab-delimited

[Complementary Dataset Link](#)

Data Preprocessing

Data Consolidation:

(Lyrics, Listening History, MSD) - **Mapping**
MSD Ids to MusiXmatch Ids.

Extract Metadata:

Artists, Artist Similarity Songs, and Lyrics
from SQLite Database

Table Joins:

Combining Artists locations & Artists terms
(tag) .

Save the information separately into 6 files:

Listening_history.csv

Seed.Csv

Lyrics.Csv

Tracks_metadata.csv

Artist_similarity.csv

Artist.Csv

Clean Text Data:

Lowercase

Tokenize

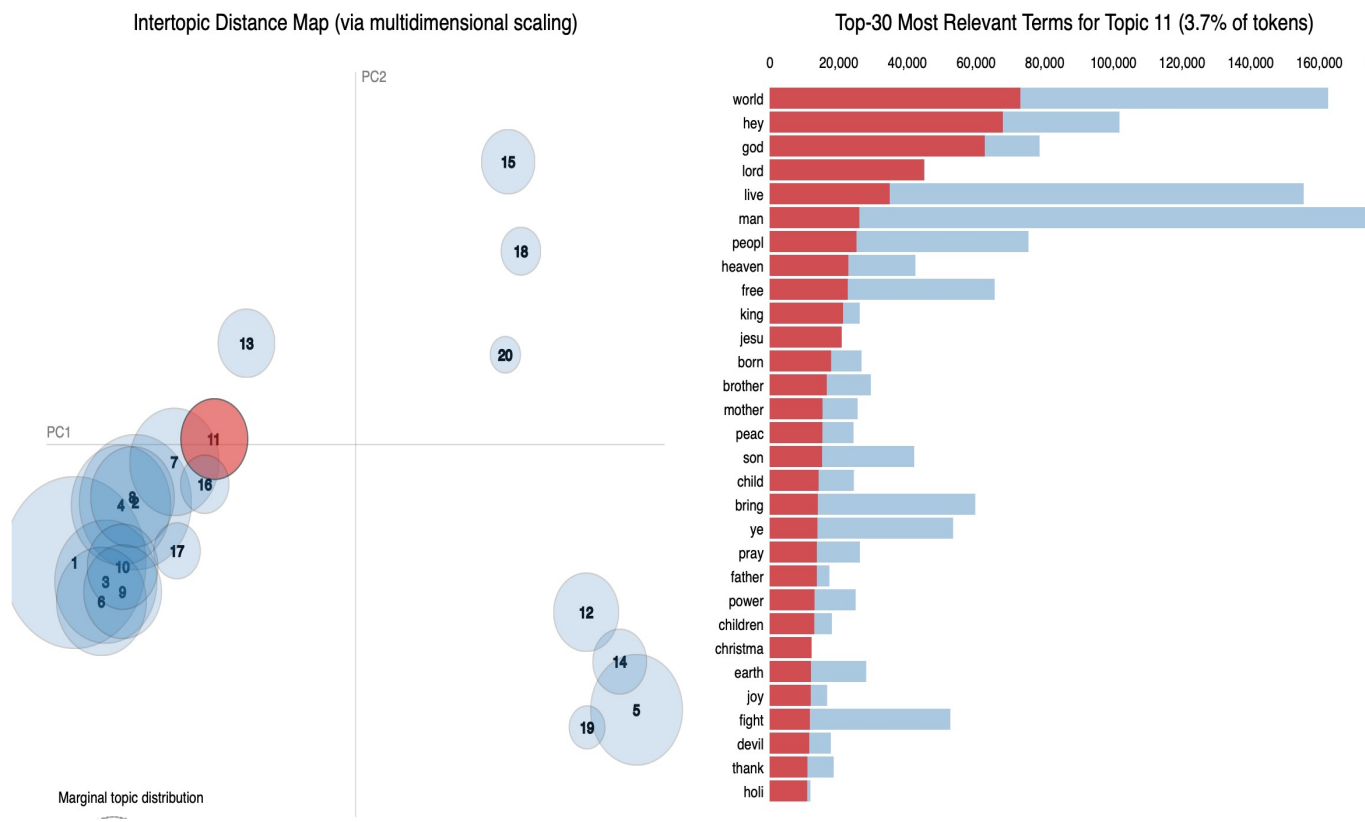
Strip out undesirable text

Remove Stop Words

Lemmatizer

Lyrics Analysis

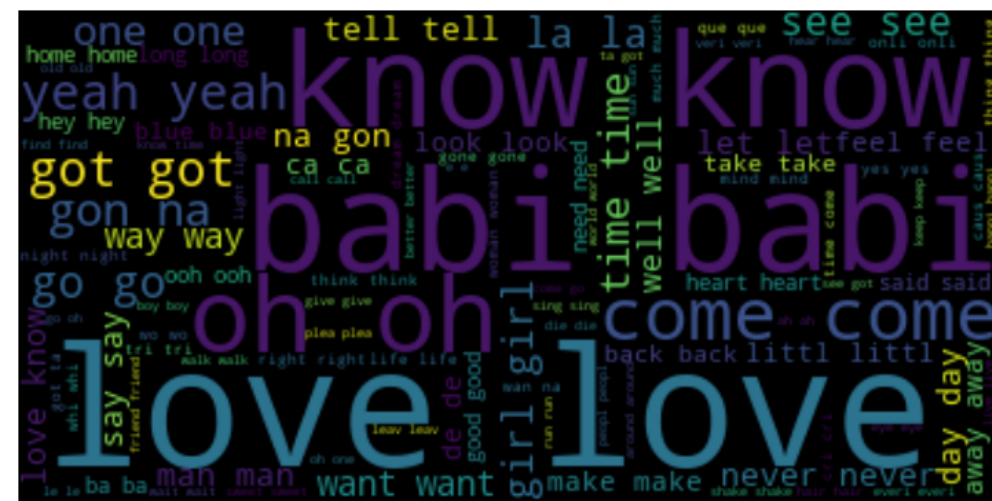
Topic Modeling (Cluster n=20)



Word Cloud Visualization

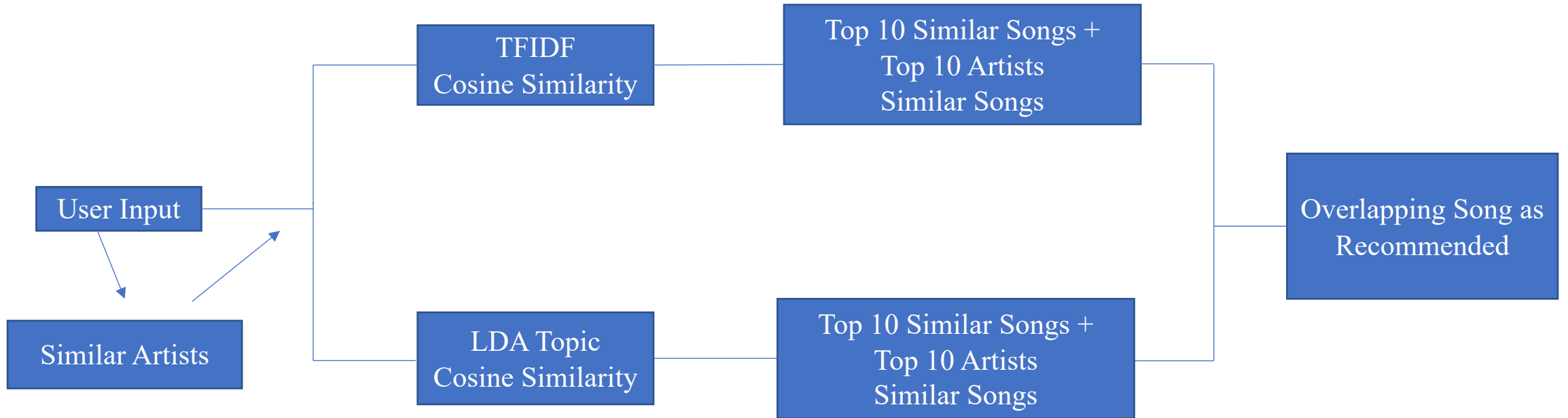


2000s



60s

Algorithm Development | Content-Based Filtering



Algorithm Development | Content-Based Filtering

Features Creation

Method 1: TFIDF for Lyrics

```
tfidf = tfidf_vectorizer.fit_transform(lyric['lyrics'])
```

```
In [92]: tfidf
```

```
Out[92]: <142263x4788 sparse matrix of type '<class 'numpy.float64'>'
         with 7669181 stored elements in Compressed Sparse Row format>
```

Method 2: LDA for Lyrics

```
# Fit lda model
lda = models.LdaModel(lyric2['corpus'], id2word=dictionary, num_topics=10)
# Topic matrix (V matrix)
lda.print_topics(10)
```

```
/Users/elyasm/opt/anaconda3/lib/python3.8/site-packages/ipykernel/ipkernel.py:287: DeprecationWarning: `should_run_async` will not call `transform_cell` automatically in the future. Please pass the result to `transformed_cell` argument and any exception that happen during the transform in `preprocessing_exc_tuple` in IPython 7.17 and above.
    and should_run_async(code)
```

```
[(0,
  '0.055*"ich" + 0.042*"und" + 0.040*"die" + 0.025*"da" + 0.024*"du" + 0.023*"der" + 0.021*"nicht" + 0.020*"ist" + 0.019*"es" + 0.017*"ein"'),
 (1,
  '0.030*"time" + 0.019*"life" + 0.017*"day" + 0.012*"live" + 0.010*"never" + 0.009*"still" + 0.009*"noth" + 0.008*"take" + 0.008*"wait" + 0.008*"away"'),
 (2,
  '0.041*"know" + 0.022*"want" + 0.020*"ca" + 0.018*"never" + 0.018*"feel" + 0.017*"see" + 0.016*"way" + 0.015*"say" + 0.013*"think" + 0.012*"could"'),
 (3,
  '0.047*"le" + 0.029*"la" + 0.028*"je" + 0.026*"et" + 0.021*"en" + 0.020*"pa" + 0.016*"du" + 0.016*"de" + 0.015*"que" + 0.014*"cest"'),
 (4,
  '0.075*"la" + 0.069*"que" + 0.033*"el" + 0.026*"en" + 0.024*"mi" + 0.024*"lo" + 0.022*"te" + 0.021*"tu" + 0.017*"se" + 0.016*"un"'),
 (5,
  '0.053*"danc" + 0.041*"ah" + 0.041*"ooh" + 0.036*"alright" + 0.032*"shake" + 0.028*"move" + 0.018*"u" + 0.018*"im" + 0.017*"bodi" + 0.015*"beat"'),
 (6,
  '0.010*"eye" + 0.009*"light" + 0.008*"world" + 0.008*"god" + 0.007*"soul" + 0.007*"sky" + 0.007*"burn" + 0.006*"see" + 0.006*"night" + 0.006*"die"'),
 (7,
  '0.056*"e" + 0.036*"da" + 0.030*"di" + 0.028*"che" + 0.022*"non" + 0.019*"se" + 0.019*"la" + 0.017*"eu" + 0.015*"mi" + 0.015*"il"'),
 (8,
  '0.012*"man" + 0.008*"know" + 0.008*"cau" + 0.008*"fuck" + 0.007*"nigga" + 0.007*"shit" + 0.006*"rock" + 0.006*"big" + 0.005*"yo" + 0.005*"put"'),
 (9,
  '0.102*"love" + 0.048*"babi" + 0.025*"girl" + 0.023*"let" + 0.018*"know" + 0.016*"hey" + 0.014*"night" + 0.013*"give" + 0.013*"heart" + 0.012*"good"')]
```

Content-Based Filtering

Simulation of A User

```
title = input('Please enter the song name:')
```

Please enter the song name:Soul Deep

```
artist = input('Please enter the artist name:')
```

Please enter the artist name:The Box Tops

```
# Recommend based on tfidf
recommend_tfidf = recommend_title(title, artist, cosine_sim_tfidf)
```

```
# Recommend based on topic
recommend_lda = recommend_title(title, artist, cosine_sim_lda)
```

Playlist with TFIDF

```
recommend_tfidf
/Users/elyasm/opt/anaconda3/lib/python3.8/site-packages/ipykernel/_eventloop_impl.py:10: DeprecationWarning: This function is deprecated. Please pass the result in IPython 7.17 and above.
  and should_run_async(code)
```

	artist_name	title
0	The Box Tops	Soul Deep
1	K'Naan	Is It A Myth?
2	K'Naan	Be Free
3	Bananarama	Rapture
4	Sanctus Real	Dear Heart
5	The Bears	Nobody's Fool
6	702	No Way
7	Lissy Trullie	Forget About It
8	Leigh Nash	Angel Tonight
9	Avril Lavigne	Mobile
10	Eels	My Descent Into Madness
11	Joe Cocker	There Goes My Baby
12	Ray Charles	Let's Go
13	The Association	Like Always (Album Version)
14	The Kingsmen	Rock Of Ages
15	The Hollies	The Woman I Love (2003 Digital Remaster)
16	Joe Cocker	Just To Keep From Drowning
17	The Kingsmen	Amazing Grace
18	Everly Brothers	Rockin' Alone
19	Four Tops	L'Arcobaleno (Walk Away Renee)
20	The Hollies	Words Don't Come Easy

Playlist with LDA

```
recommend_lda
/Users/elyasm/opt/anaconda3/lib/python3.8/site-packages/ipykernel/_eventloop_impl.py:10: DeprecationWarning: This function is deprecated. Please pass the result in IPython 7.17 and above.
  and should_run_async(code)
```

	artist_name	title
0	The Box Tops	Soul Deep
1	3 Doors Down	It's Not Me
2	The Subways	With You
3	X Marks The Pedwalk	Never look Back
4	B.C. Camplight	Couldn't You Tell
5	Incognito	Stone Cold Heart
6	New Order	Everything's Gone Green
7	Eagle-Eye Cherry	If You Don't Know By Now
8	Unwound	Lady Elect
9	Destroy The Runner	Isabella's
10	Kyuss	Fatso Forgotso
11	Joe Cocker	There Goes My Baby
12	Ray Charles	Let's Go
13	The Association	Like Always (Album Version)
14	The Kingsmen	Rock Of Ages
15	The Hollies	The Woman I Love (2003 Digital Remaster)
16	Joe Cocker	Just To Keep From Drowning
17	The Kingsmen	Amazing Grace
18	Everly Brothers	Rockin' Alone
19	Four Tops	L'Arcobaleno (Walk Away Renee)
20	The Hollies	Words Don't Come Easy

Recommendation | Results

	artist_name	title
0	The Box Tops	Soul Deep
1	Joe Cocker	There Goes My Baby
2	Ray Charles	Let's Go
3	The Association	Like Always (Album Version)
4	The Kingsmen	Rock Of Ages
5	The Hollies	The Woman I Love (2003 Digital Remaster)
6	Joe Cocker	Just To Keep From Drowning
7	The Kingsmen	Amazing Grace
8	Everly Brothers	Rockin' Alone
9	Four Tops	L'Arcobaleno (Walk Away Renee)
10	The Hollies	Words Don't Come Easy

Results | Content-Based Filtering



Lyrics

Darlin' I don't know much
I know I love you so much
A lot depends on your touch
My love is a river running soul deep
A way down inside me it's a soul deep
Too big to hide, can't be denied
Love is a river running soul deep

I worked myself to euphoria
Just to show I adore ya
There's nothing I wouldn't do for ya
Cause my love is a river running soul deep
A way down inside me it's a soul deep
Too big to hide, can't be denied
Love is a river running soul deep

All I ever, ever hoped to be
Depends on your love for me
If you believe me, if you should leave me
I'd be nothing but a jilted male
I know darned well, I could tell, but

I don't know much
I know I love you so much
A lot depends on your touch
My love is a river running soul deep
A way down inside me it's a soul deep
Too big to hide, can't be denied
Love is a river running soul deep
My love is a river running soul deep
A way down inside me it's a soul deep
My love is a river running soul deep
A way down inside me it's a soul deep
My love is a river running soul deep
A way down inside me it's a soul deep

Source: LyricFind

Songwriters: Per Gessle

Soul Deep lyrics © Kobalt Music Publishing Ltd.

Lyrics

You keep running away
Though I beg you not to leave
But still you won't stay
Darlin' you keep running away
Tear my heart apart every step of the way

You're here today and gone tomorrow
Leavin' this heart of mine in sorrow
Now you come around every now and then
Long enough to hurt me, and then you're gone again

Darlin' you, you keep running away
Oh, I begged you not to leave, you never stay
Now you, you keep running away
Leavin' me here to face another lonely day

To you all of this is just a game
But each time you came here, I feel the pain
But I've got so much love for you
I keep wanting you, no matter what you do

All I want to do is take care of you
Everything I have in my life, I'll share with you
This soul of mine has been possessed by you
Darling my heart has been obsessed with you
Just look at me, I'm not the man, I used to be

I used to be proud, I used to be strong
But all of that's changed girl, since you come along
Your lovin' sweetness is my weakness
Though I need you, dear, I just can't keep you near

Running away, running away, running away
Running away, running away, running away

Each time you go, the hurt comes callin'
My days become nights, darlin'
My nights become so much longer
You're in my life, you're in my heart

But I can't get you, get you into my arms
Darlin' you, you keep runnin' away
Darlin' you, you just keep runnin' away

Source: LyricFind

Songwriters: Jr. / Brian Holland / Edward Holland / Edward / Jr. Holland / Lamont Dozier / Lamont Herbert Dozier

You Keep Running Away lyrics © Sony/ATV Music Publishing LLC



Thank You