Learning Shared Vector Representations of Lyrics and Chords in Music

Timothy Greer, Karan Singla,
Benjamin Ma, Shri Narayanan

University of Southern California

Paper ID 1934
ICASSP 2019
May 15th 2019, Brighton, UK

Motivation

- Music is something we all enjoy
- Lyrics and music affect us differently
- Can we capture this joint effect?
- What other musical phenomena can we find?

Short Demonstration

- “Sad” song with “sad” lyrics:

  “I hurt myself today, to see if I still feel”

- "Happy” song with “sad” lyrics:

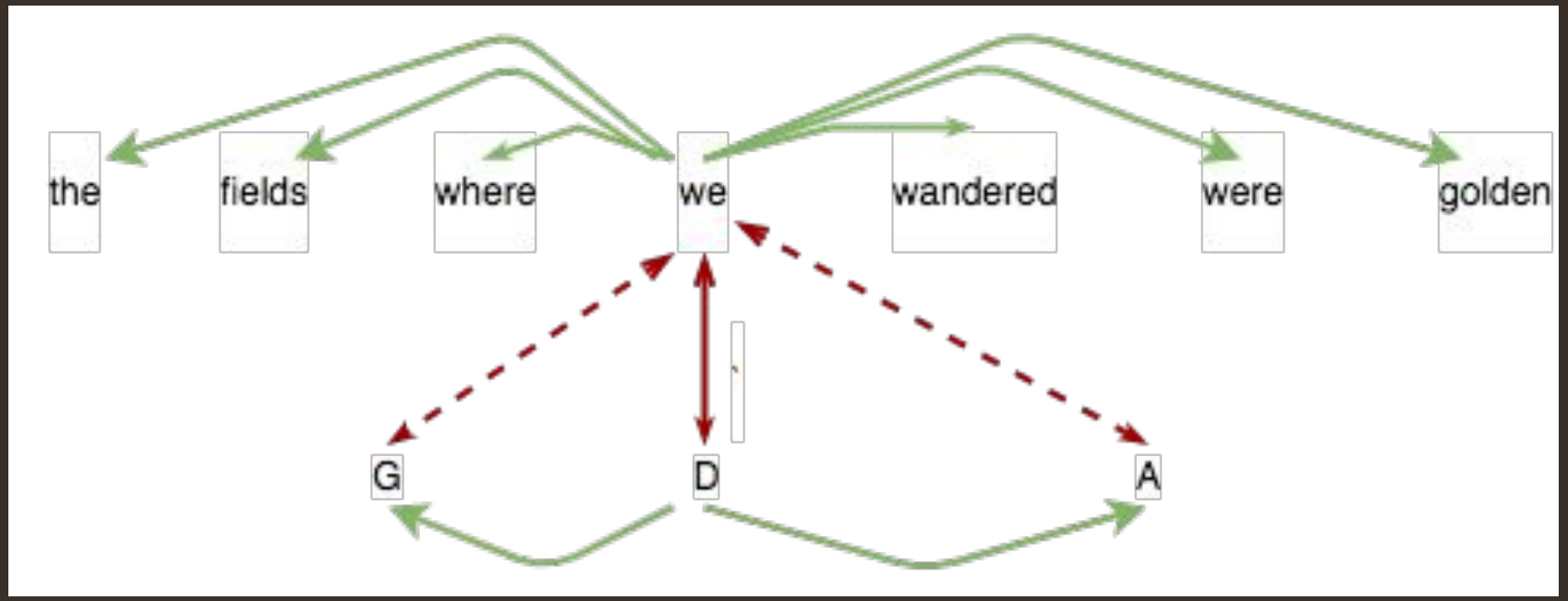
  “I’ve been lonely and sad about the things that you do”

Johnny Cash, “Hurt” *American IV: The Man Comes Around*, American, Lost Highway. *Spotify*, https://open.spotify.com/track/28cnXtME493VX9NOw9cIUh?si=lj4D0y8aQz2Rmw-eNW6OBQ.
ABBA, “Mamma Mia” *ABBA*, Polar. *Spotify,* https://open.spotify.com/track/22NN4BS1AlqVbyKIWExgON?si=9Fsgknt2RIOGtNpg_fSbLQ

Methods

- Word embeddings model (Word2Vec)
- Use two “languages” to find context
  - Chords and lyrics
- Use this model on pilot task
- Make music-theoretical observations

The Model



- Created embeddings similar to Luong, 2015

Luong, H Pham, and C D Manning, "Bilingual word representations with monolingual quality in mind," in Proc. of the 1st Workshop on Vector Space Modeling for Natural Language Processing, 2015, pp. 151–159

Learning Embeddings

- Monolingual loss term:

$$MONO = \frac{1}{T}\sum_{t=1}^{T}\sum_{-l\le j\le l, j\neq 0}\log(p(w_{t+j}|w_t))$$

- Cross-lingual loss term:

$$CROSS_{WC} = \frac{1}{T}\sum_{t=1}^{T}\sum_{-l\le j\le l}\log(p(c_{k+j}|w_t))$$

- Training objective function:

$$\mathcal{L} = MONO_W + MONO_C + CROSS_{CW} + CROSS_{WC}$$

The Dataset

- Ukutabs.com
- Chords and lyrics are aligned

```
D          A          D                    G
I've been cheated by you, since I don't know when
D    A          D                    G
So I made up my mind, it must come to an end
```

The Pilot Task

- Two annotators labeled song segments
- Labeled negative, positive, or neutral
- 929 segments were agreed-upon
- Used embeddings to predict emotion

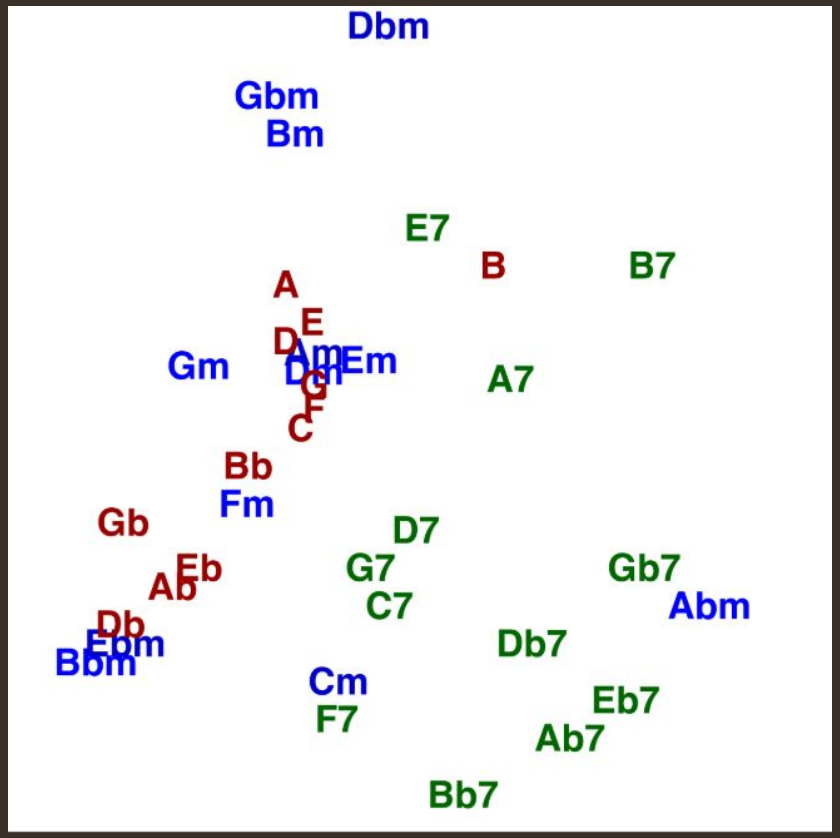
Results

Model	Accuracy	p-value
Baselines		
Majority Class	55.32 %	N/A
Chord n-grams	57.83 %	.138
Lyrics n-grams	57.31 %	.194
Our models		
Chords only	59.74 %	.027
Lyrics only	60.52 %	.012
Chords & Lyrics	62.28 %	.001

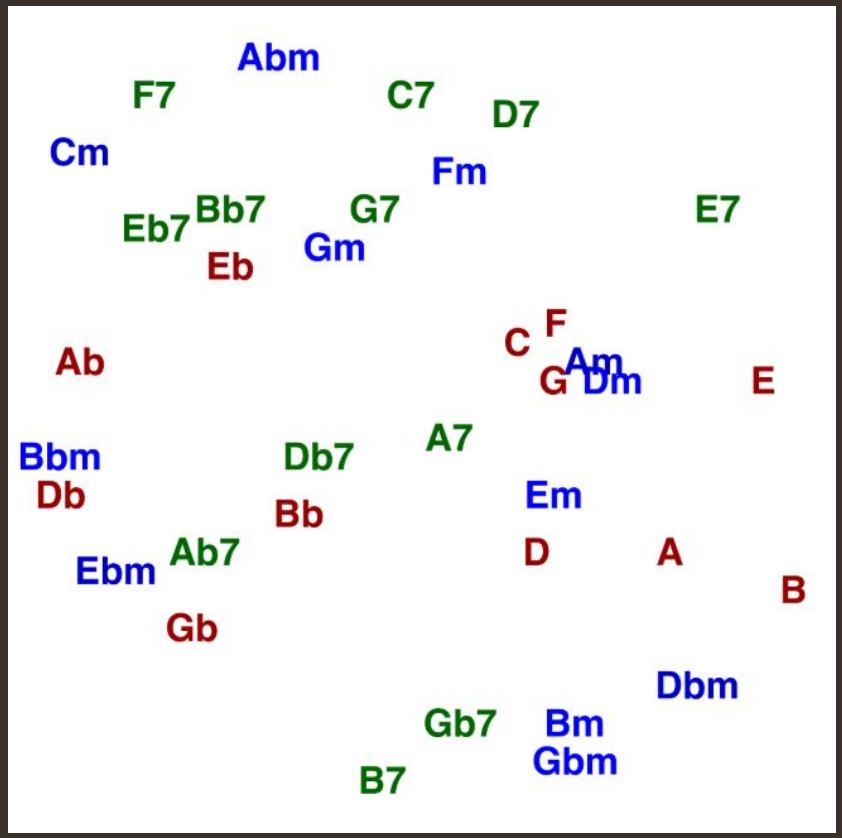
- Used Logistic Regression classifier
- Chords-and-lyrics model performs best

Observations from Embeddings

Embedding with Chords Only

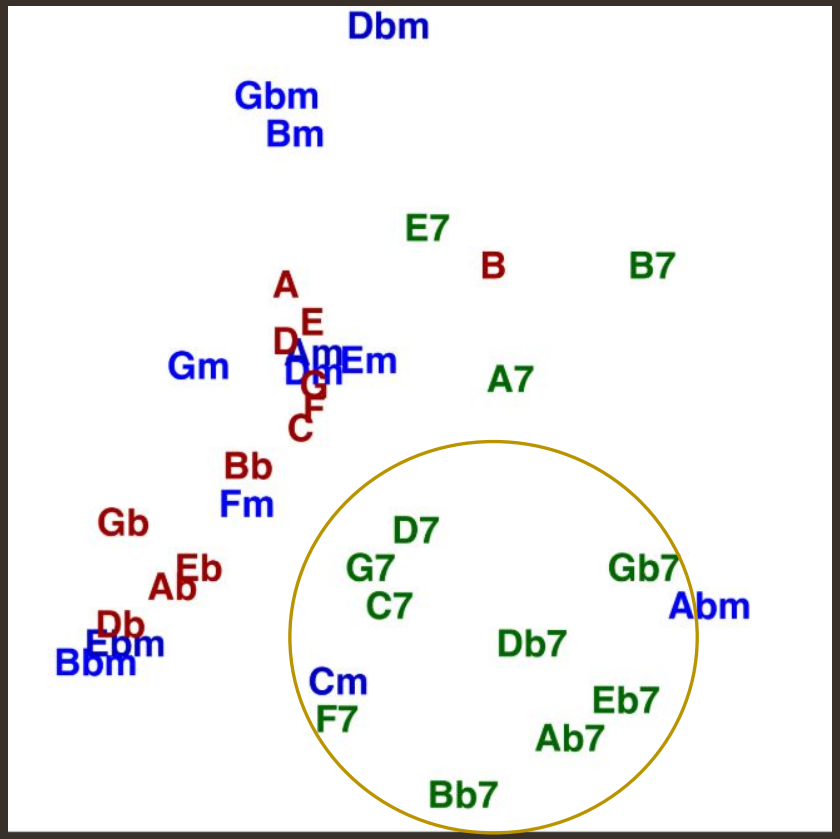


Embedding with Chords and Lyrics

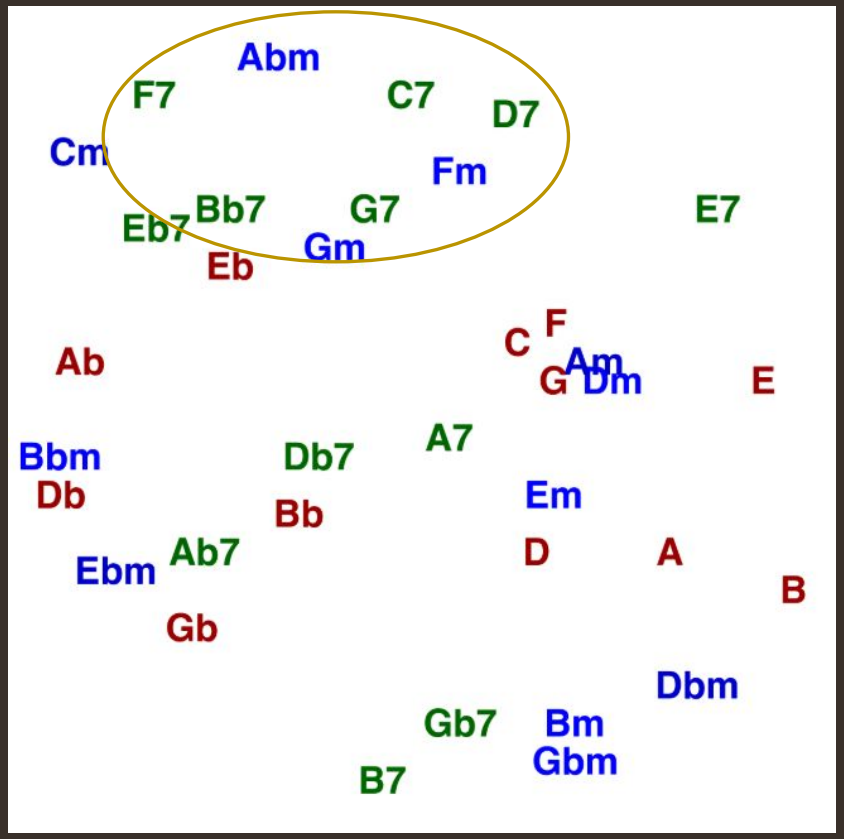


Observations from Embeddings

Embedding with Chords Only



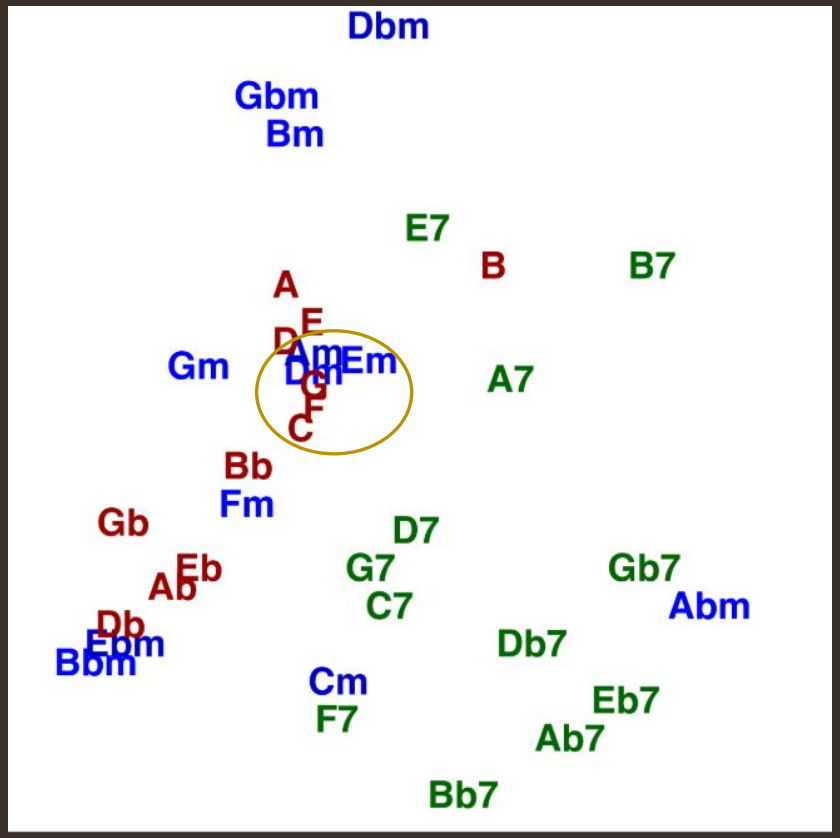
Embedding with Chords and Lyrics



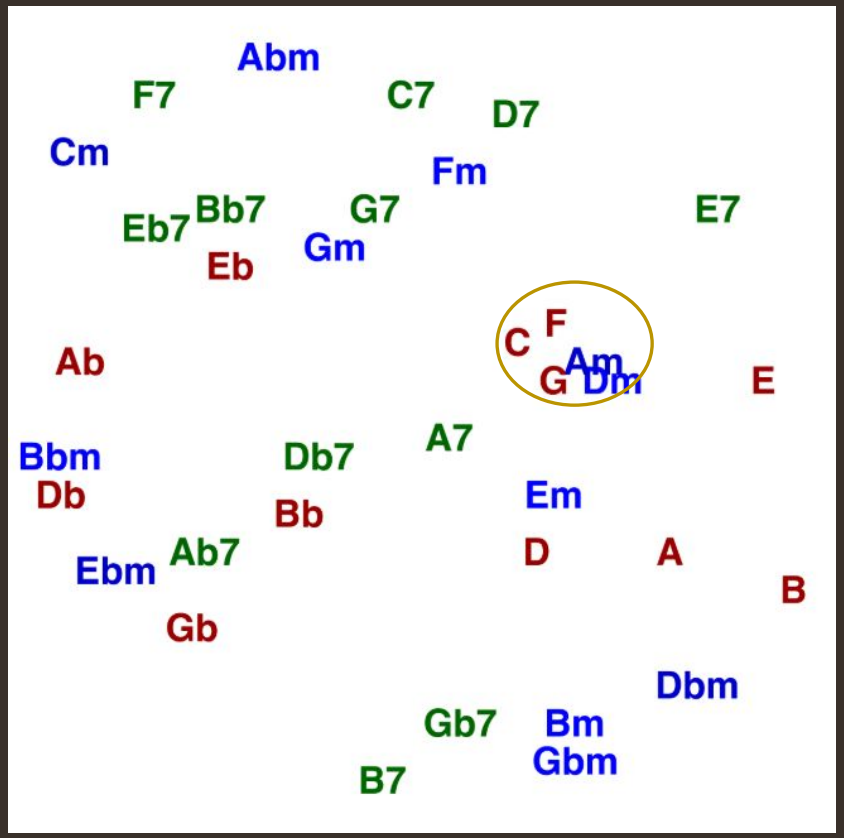
Bluesy chords are clustered in space

Observations from Embeddings

Embedding with Chords Only



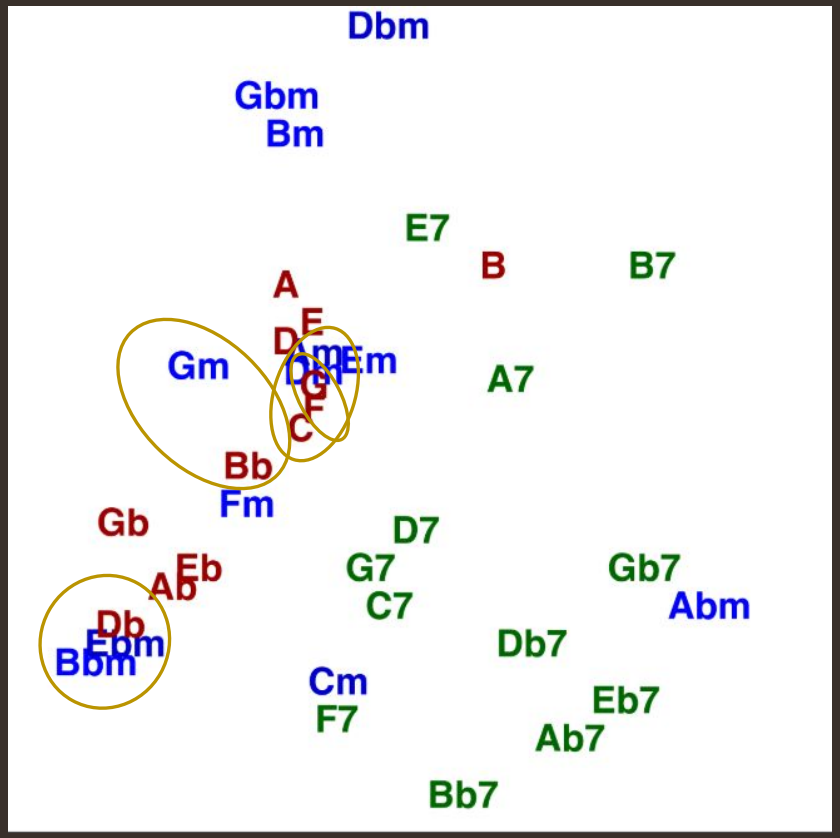
Embedding with Chords and Lyrics



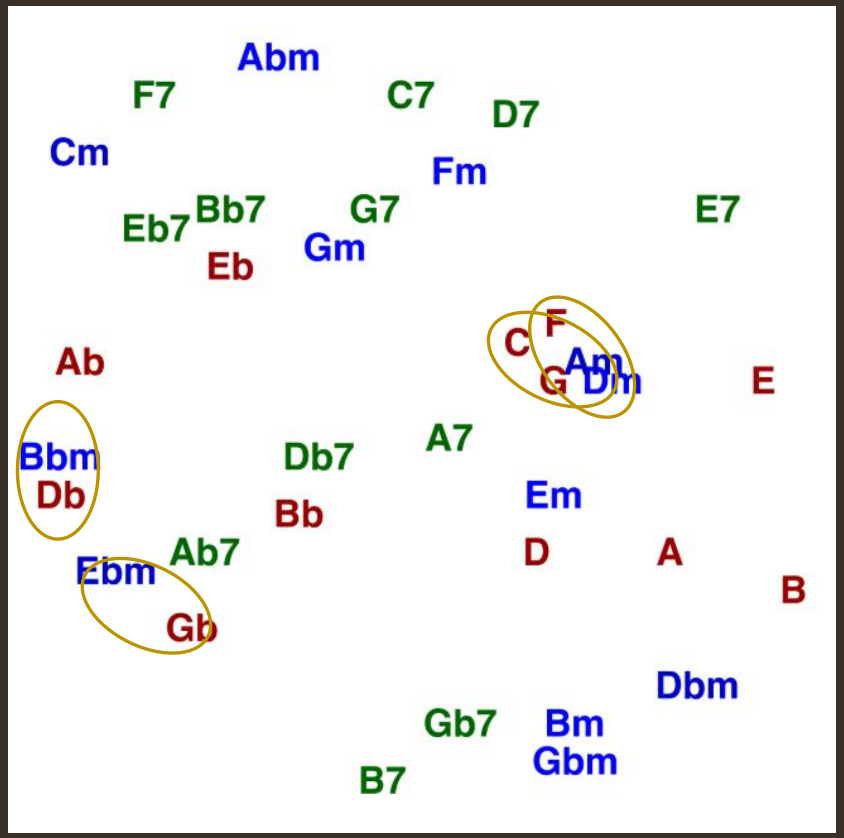
Enharmonic chords are clustered in space

Observations from Embeddings

Embedding with Chords Only



Embedding with Chords and Lyrics



Chord substitutions are clustered in space

Conclusions/Future Work

- Studying chords and lyrics together may be helpful for music-related studies
- Music theory concepts are supported by our approach
- Is this joint model useful in other areas?
- Can performance be improved with other models?

Thank You

- Special thanks to Morteza Dehghani at USC for encouraging this work

- References:

  - [1] Johnny Cash, “Hurt” *American IV: The Man Comes Around*, American, Lost Highway. *Spotify*, https://open.spotify.com/track/28cnXtME493VX9NOw9cIUh?si=lj4D0y8aQz2Rmw-eNW6OBQ.
  - [2] ABBA, “Mamma Mia” *ABBA*, Polar. *Spotify,* https://open.spotify.com/track/22NN4BS1AlqVbyKIWExgON?si=9Fsgknt2RIOGtNpg_fSbLQ
  - [3] Luong, H Pham, and C D Manning, “Bilingual word representations with monolingual quality in mind,” in Proc. of the 1st Workshop on Vector Space Modeling for Natural Language Processing, 2015, pp. 151–159

The Dataset

- Ukutabs.com
- Chords and lyrics are aligned

```
C       D       Am          C       D       Am
I hurt myself today    to see if I still feel
C       D       Am          C       D       Am
I focus on the pain    the only thing that's real
```