



emotion detection in
song lyrics



— introduction —

1 the problem

2 datasets

— approach —

3 neural networks/RNN

4 embeddings

5 LSTM/GRU

— conclusion —

6 results

7 next steps



THE PROBLEM

emotion detection is the subfield of sentiment analysis aimed at extracting subjective feelings and thoughts from text.

music is related to emotions.

the scope of this project is then to train a model that, analyzing the lyrics, extracts emotions carried by songs.



train a model to predict emotions
analyzing text



study the correlation between song
genres and emotions



create playlists based on user's
genre preference and mood

three objectives

TRAINING AND TEST SET

[WASSA-2017](#) tweets dataset.
it's already divided in training and test splits; both parts consist of one file for each of the four basic emotions (anger, fear, joy, sadness); all the files have the same structure (look at the table).

3613 tweets for training and 3142 for testing.

	id	text	sentiment	intensity
0	20,000	How the fu*k! Who the heck! moved my fridge!... should I knock the landlord door. #angry #mad ##	anger	0.938
1	20,001	So my Indian Uber driver just called someone the N word. If I wasn't in a moving vehicle I'd have jumped out #disgusted	anger	0.896
2	20,002	@DPD_UK I asked for my parcel to be delivered to a pick up store not my address #fuming #poorcustomerservice	anger	0.896
3	20,003	so ef whichever butt wipie pulled the fire alarm in davis bc I was sound asleep #pissed #angry #upset #tired #sad #tired #hangry #####	anger	0.896
4	20,004	Don't join @BTCare they put the phone down on you, talk over you and are rude. Taking money out of my acc willynilly! #fuming	anger	0.896

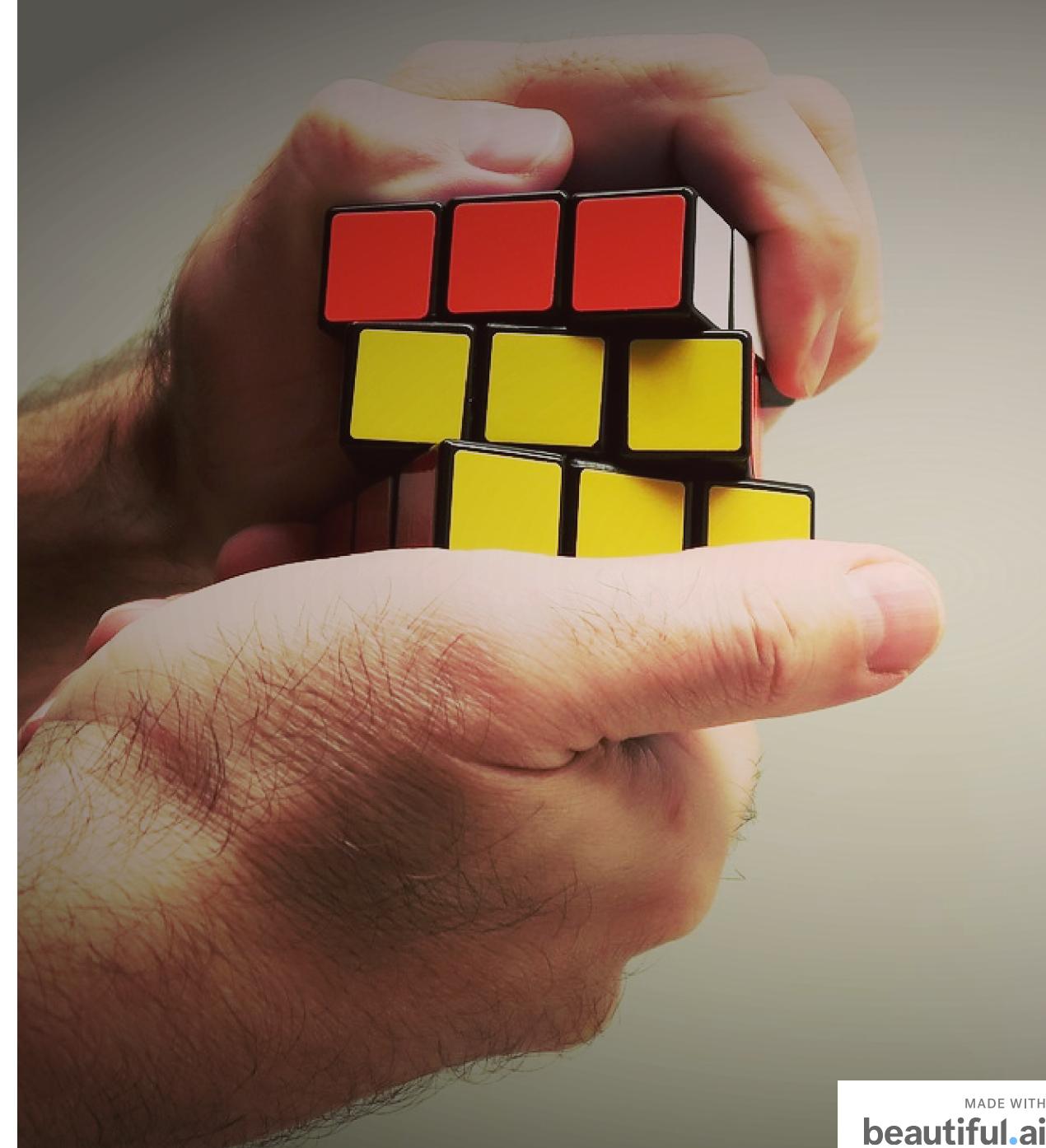
SONG LYRICS DATASET

scraping the [Billboard Year-End Hot 100 singles of 2020](#) page I retrieved songs rank, titles, artists and genres.

through the Genius API I was then able to retrieve song lyrics.

	rank	song	artist	genre	lyrics
0	1	Blinding Lights	The Weeknd	Synthwave	Yeah\n\nI've been tryna call\nI've been on my ...
1	2	Circles	Post Malone	Pop rock	Oh, oh, oh\nOh, oh, oh\nOh, oh, oh, oh, oh\n...
2	3	The Box	Roddy Ricch	Hip hop	Pullin' out the coupe at the lot\nTold 'em fuc...
3	4	Don't Start Now	Dua Lipa	Nu-disco	If you don't wanna see me\nDid a full one-ei...
4	5	Rockstar	DaBaby	Hip hop	Woo, woo\nI pull up like\nHow you pull up, Bab...

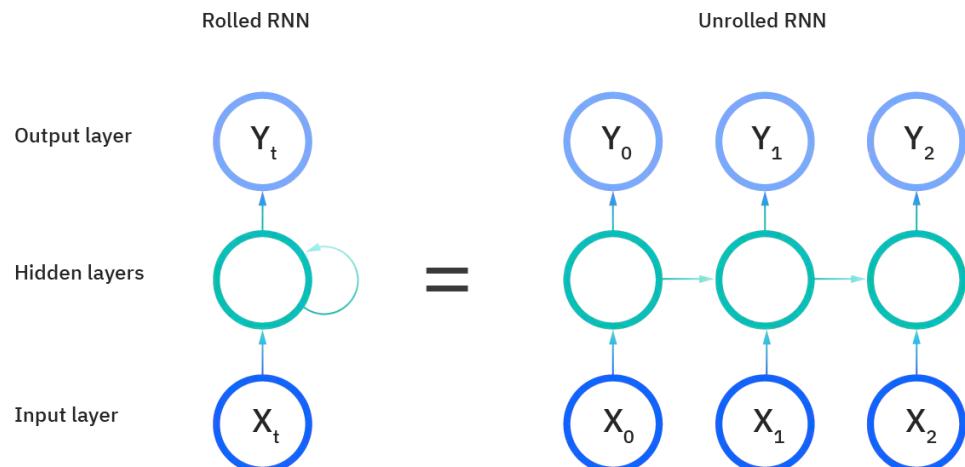
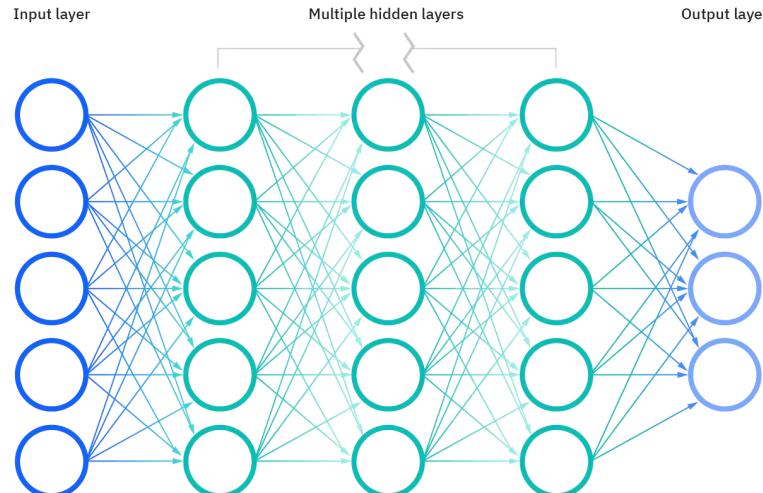
APPROACH



MADE WITH

beautiful.ai

NEURAL NETWORKS



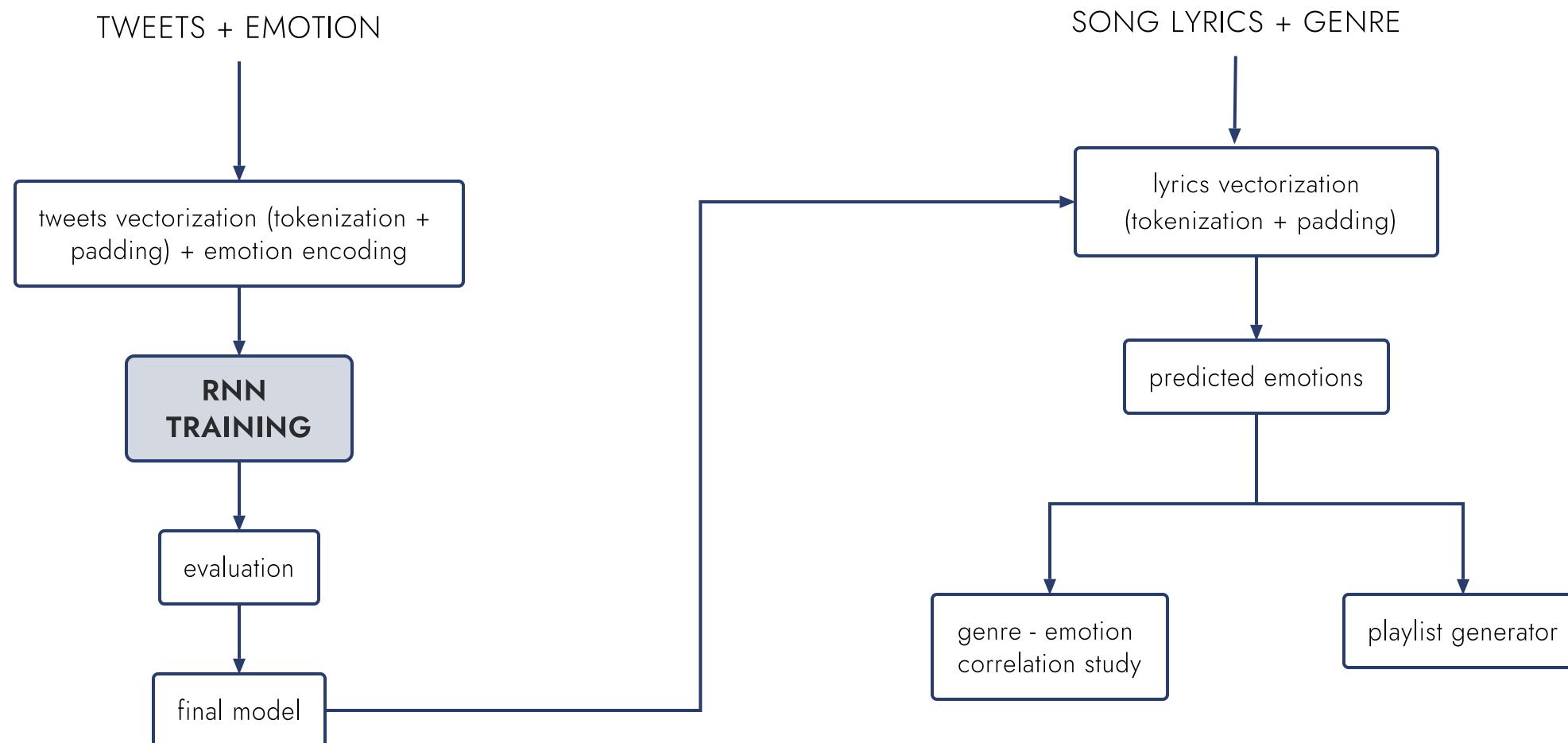
FEED-FORWARD NEURAL NETWORK

issue: inputs and outputs are independent of each other; this structure doesn't take into account information coming from previous predictions.

RECURRENT NEURAL NETWORK

ad-hoc for sequence-to-sequence learning; it has a looping mechanism that acts as an highway allowing information to flow from one step to the next.

METHODOLOGY



WORD EMBEDDINGS

when working with text, the input for neural networks needs to be a word vector.

word embeddings allow us to represent words in the form of a real-valued vector that encodes the meaning of the word such that words that are closer in the vector space are expected to have a similar meaning.

word embeddings can be obtained using a set of language modeling and feature learning techniques where words are mapped to vectors of real numbers passing from a space with many dimensions per word to a vector space with a much lower dimension.

these word vectors can be used either by training a specific word embedding model (Word2vec, GloVe...) for a corpus (assuming to have enough contents) or by exploiting a pre-trained model (as in the case of this project) to embed the corpus words.

GloVe is based on ratios of probabilities from the word-word co-occurrence matrix, these probabilities have the potential for encoding some form of meaning.

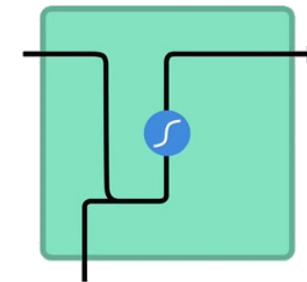
LSTM AND GRU

short term memory issue in RNNs:

- "Alice is allergic to nuts.
She can't eat *peanut butter*."
- "Alice is allergic to nuts. [...]
She can't eat *peanut butter*."

as the gap grows, RNNs become unable to learn to connect information.

two proposed solutions: long short term memory networks and gated recurrent units.



sigmoid



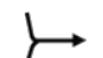
tanh



pointwise multiplication

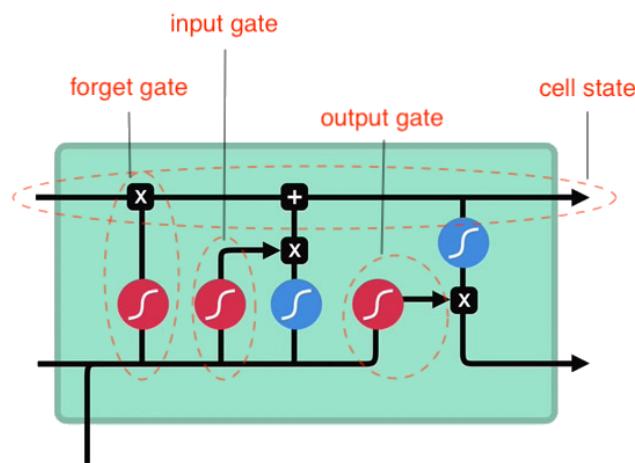


pointwise addition

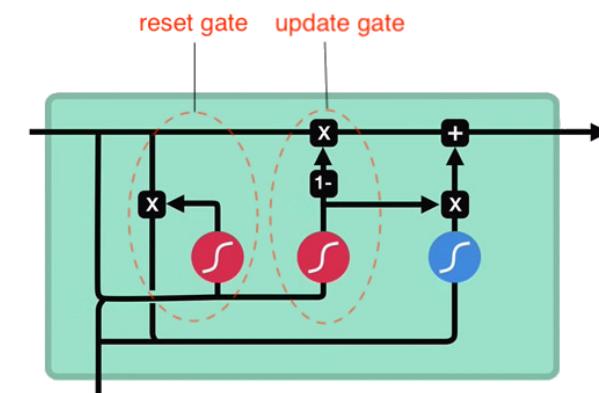


vector concatenation

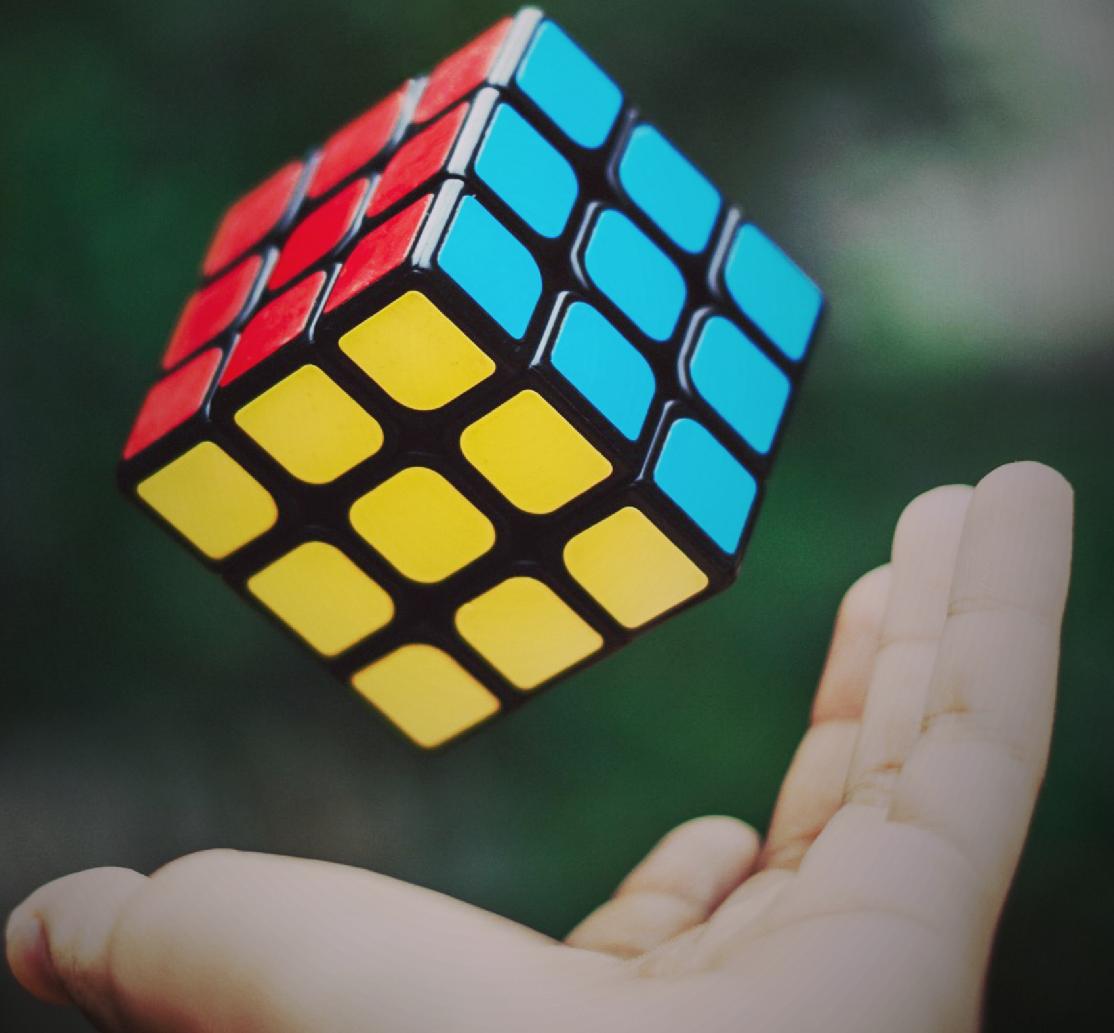
Vanilla Recurrent Neural Networks



Long Short Term Memory (LSTM)



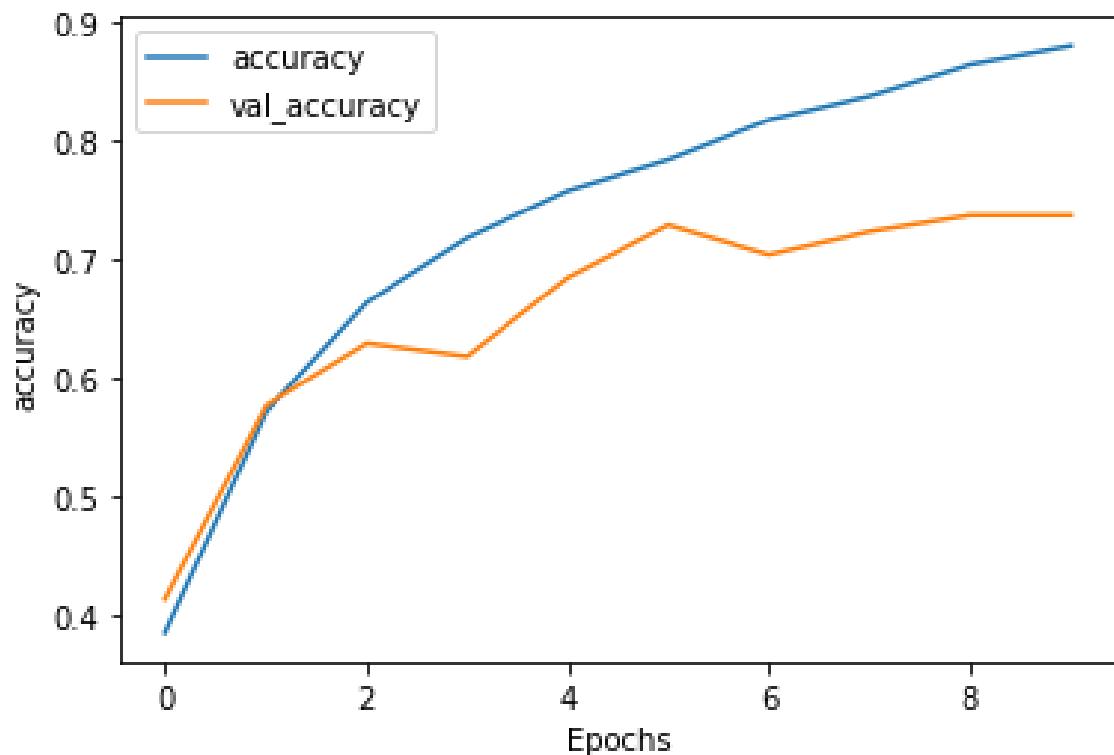
Gated Recurrent Units (GRU)



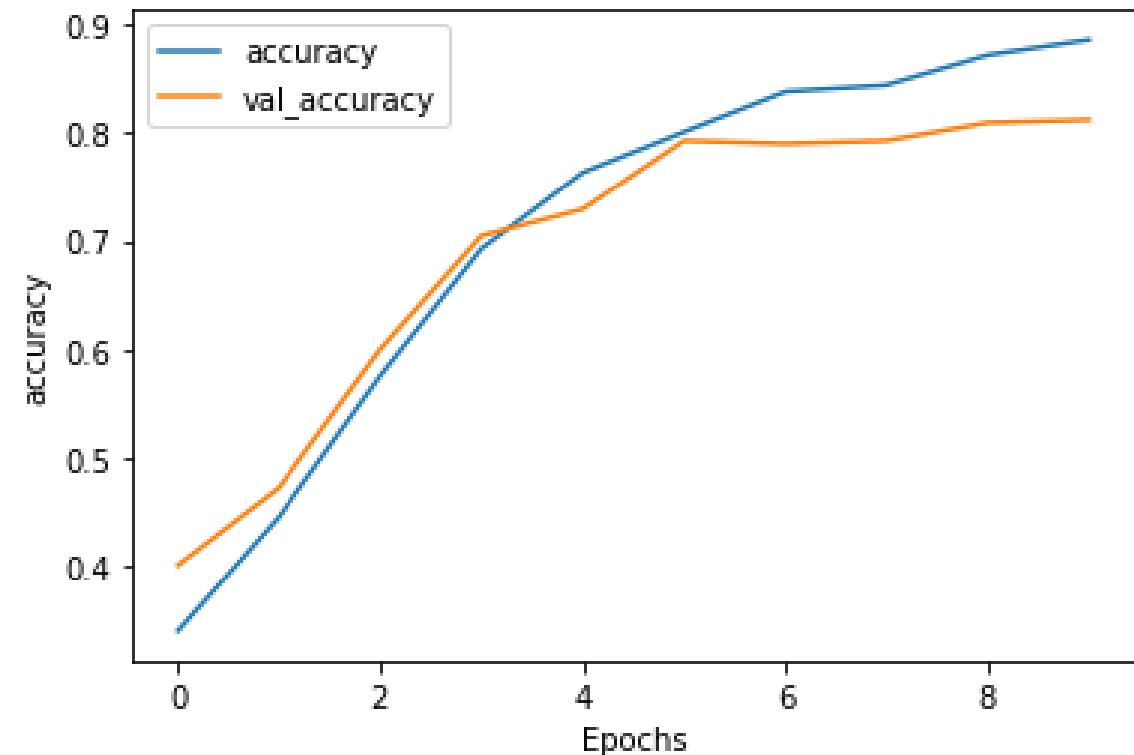
CONCLUSION

LSTM VS GRU:

training - validation accuracy



LSTM



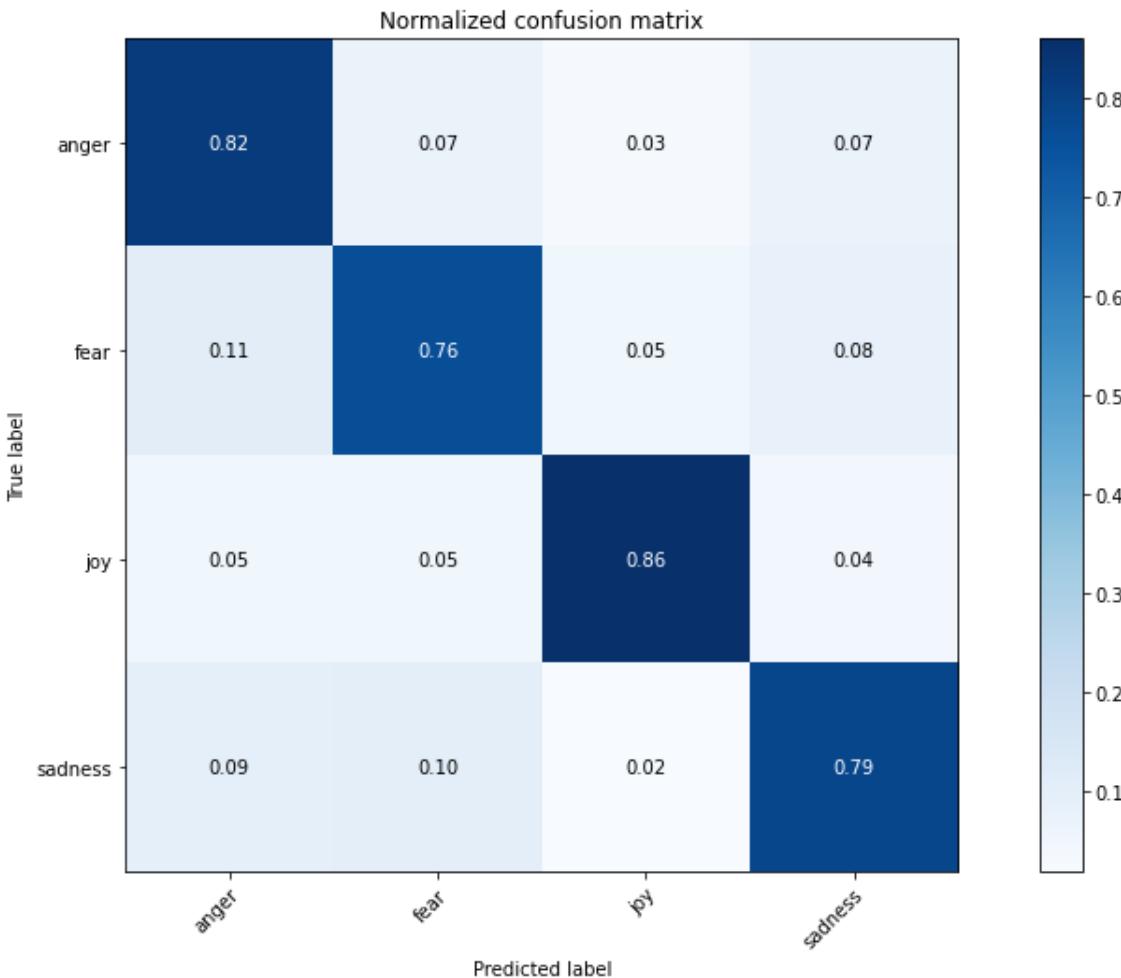
GRU

LSTM VS GRU

	training time	test accuracy
LSTM	13m 39s	74.79%
GRU	11m 01s	80.52%

GRU FTW

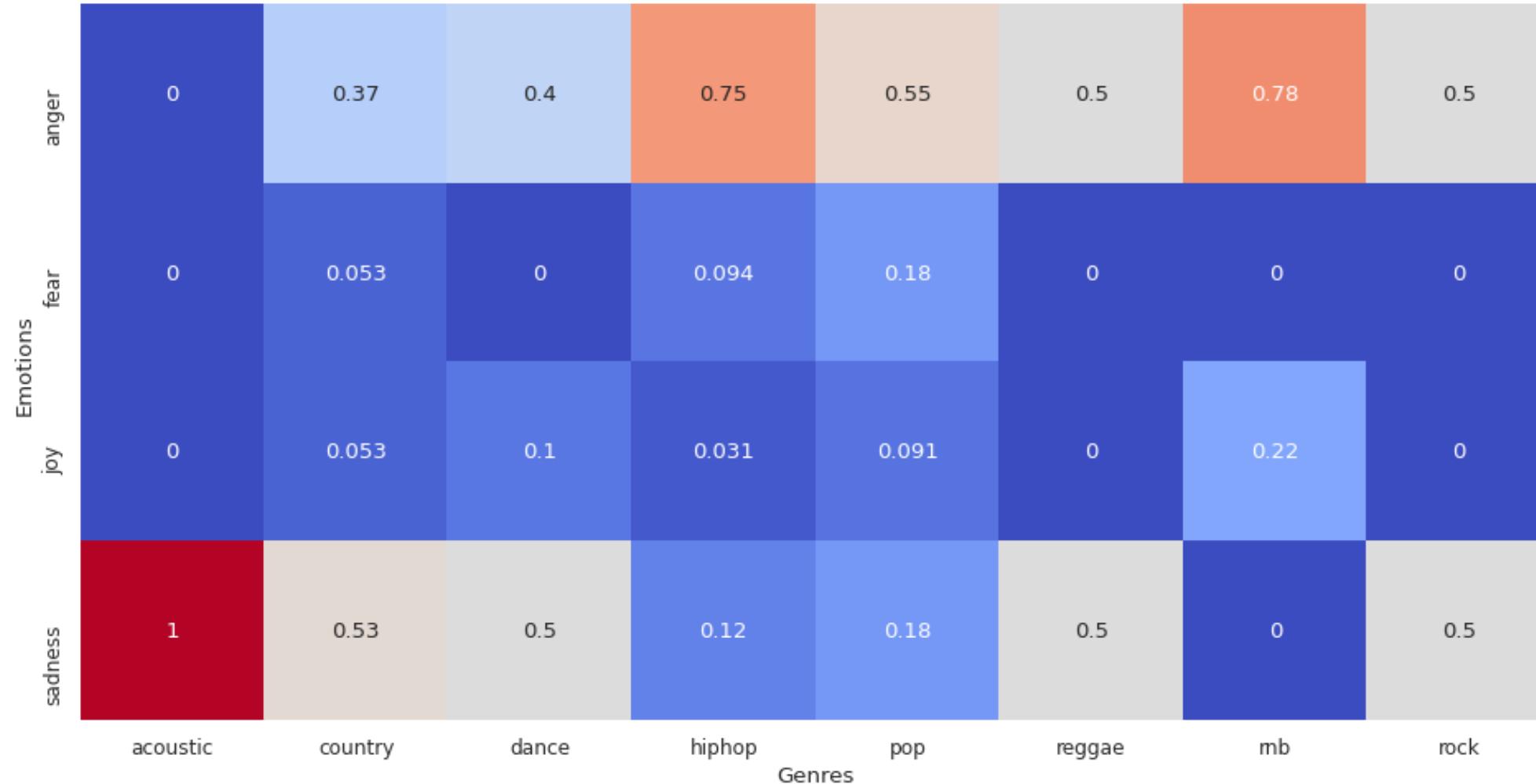
confusion matrix





PREDICTIONS WITH
SONG LYRICS

GENRE - EMOTION CORRELATION MATRIX

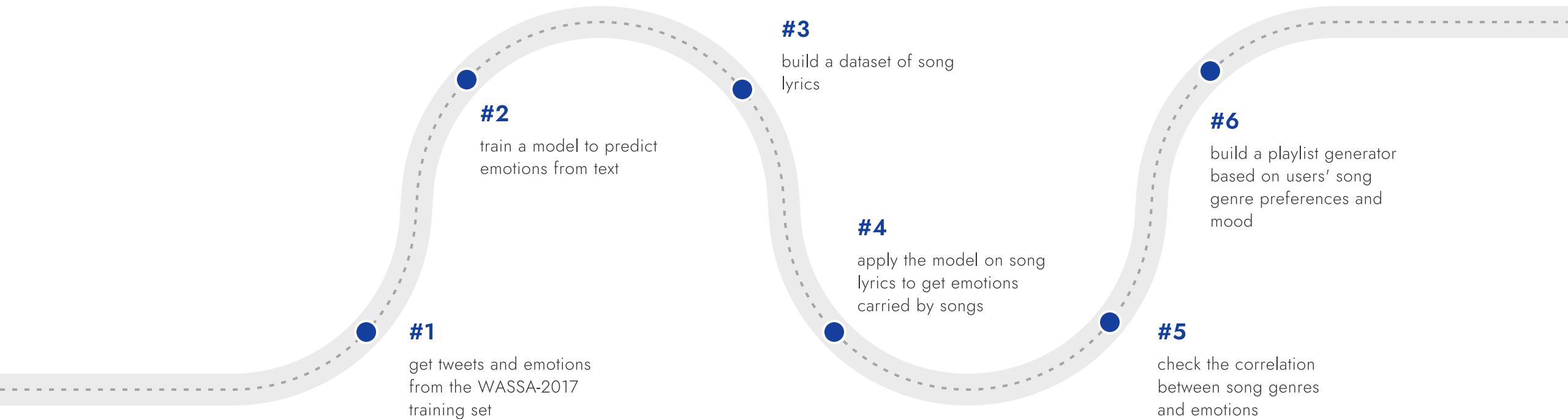


PLAYLIST GENERATOR

What genre you'd like to listen to? hiphop
What's your mood? sadness

Go listen these songs:
Toosie Slide - Drake
Death Bed - Powfu
Heart on Ice - Rod Wave
Suicidal - YNW Melly

RECAP



ISSUES AND NEXT STEPS



unbalanced distribution of emotions in training set



enlarge training dataset



homogeneous song genres



enlarge song lyrics dataset



emotion prediction based only on lyrics



different embeddings/CNNs

THANK YOU