

Lyrics generation

Data Mining 2020/21 project

A photograph of a hand holding a magnifying glass, focusing on a blurred sunset or sunrise over water. The magnifying glass is held in the lower right quadrant, with its handle pointing towards the bottom left.

Lyrics generations

WHAT HAVE WE DONE?

- Goal
- Research
- Dataset
- Preprocessing
- Models
- Results

Goal

WHAT IS OUR OBJECTIVE?

Inputs:

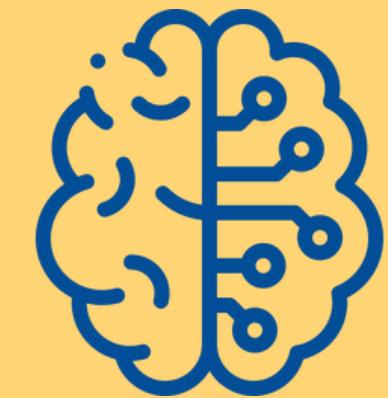
- single word or sequence of words
- musical genre

Output:

- Lyrics starting with the input word(s) whose generation is influenced by the input genre

Research

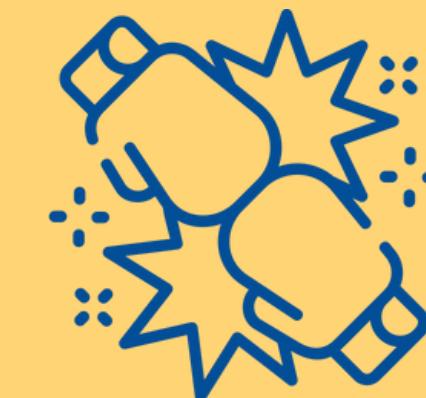
HOW IS THE PROJECT DESIGNED?



LSTM



WORD2VEC

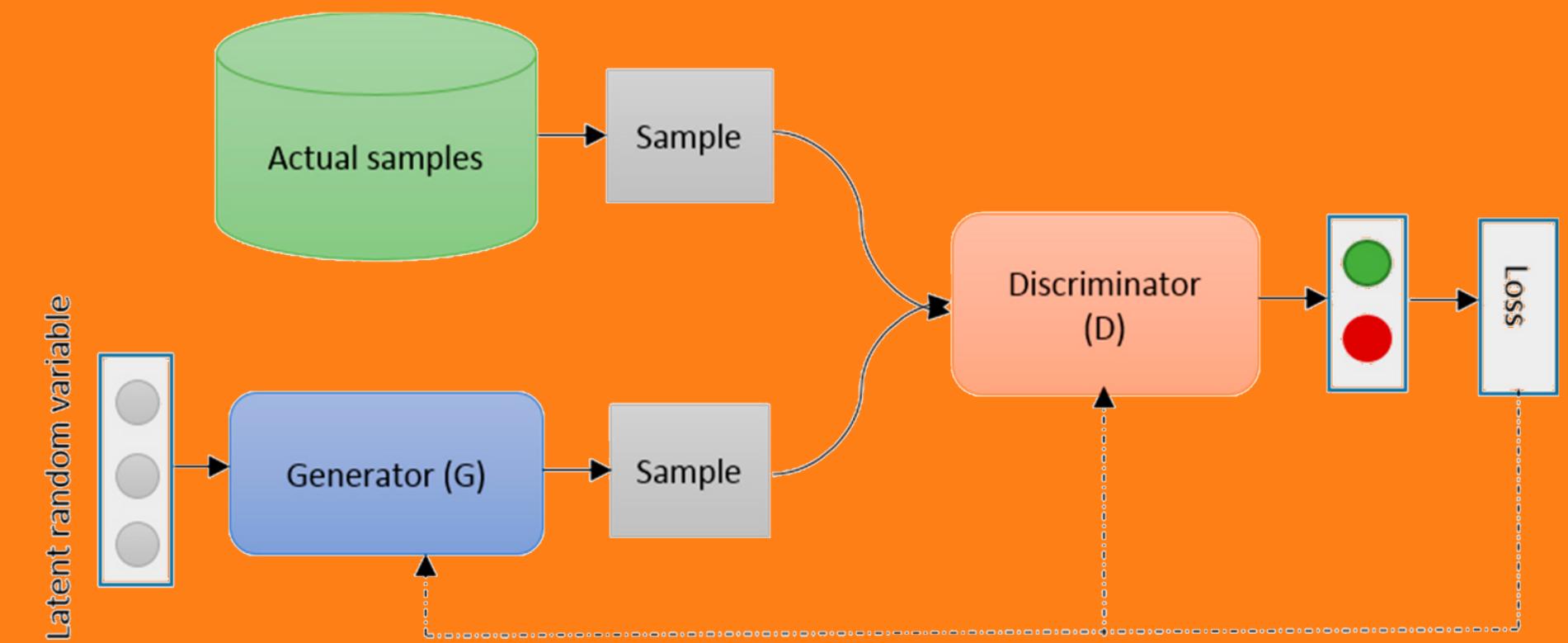


GAN

Generative Adversarial Networks

HOW ARE THEY MADE?

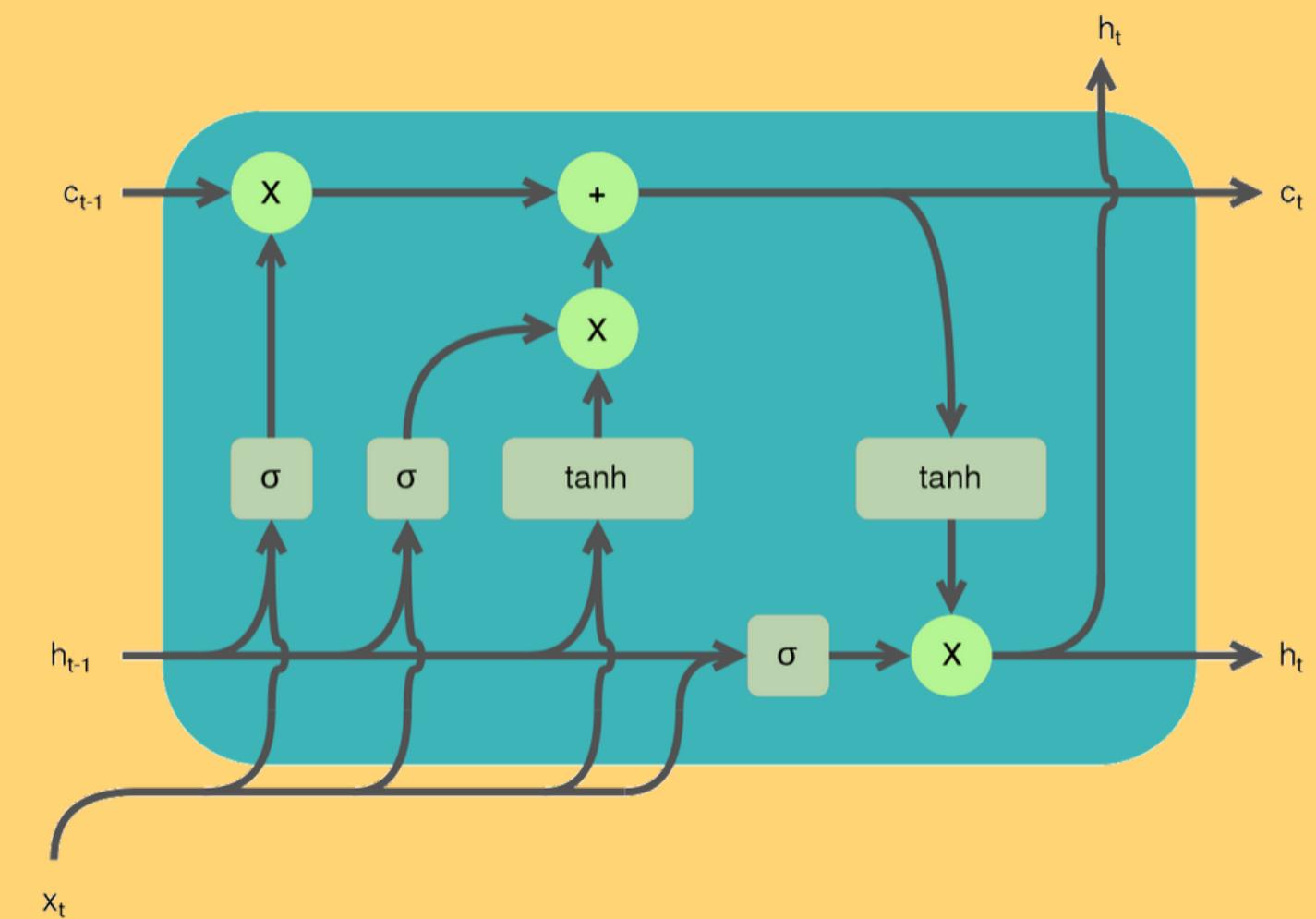
- Generator model
- Discriminator model
- Maximize similarity between fake and real data
- Minimize discrimination loss
- Convergence failures



Long Short-Term Memory layers

HOW ARE THEY MADE?

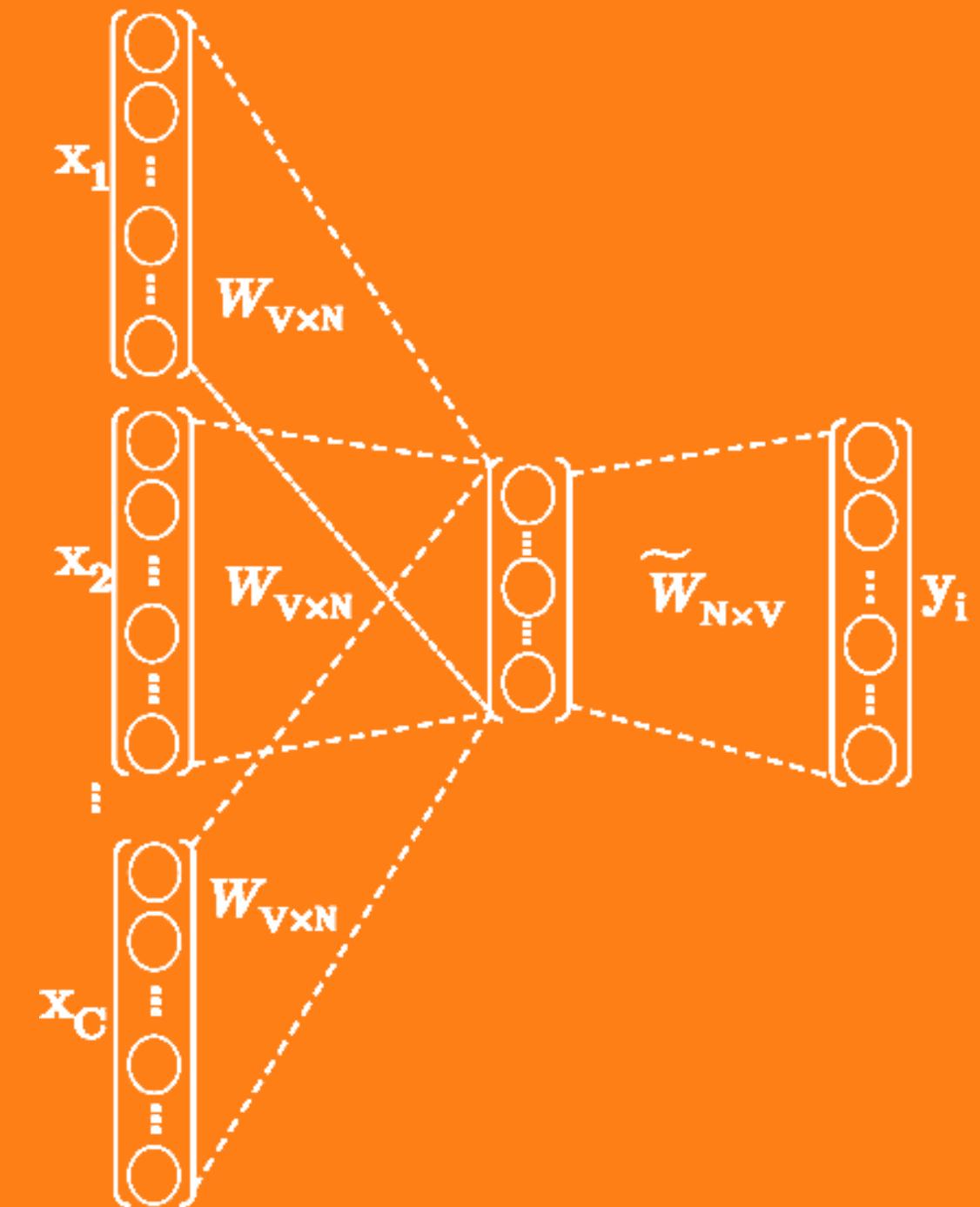
- Can process entire sequences of data
- Composed by cell, input gate, output gate and forget gate
- Partially deal with the vanishing gradient problem of traditional RNNs
- Can be layered or bi-directed



Pretrained word vector representation

WHAT IS IT?

- Word2Vec pre-trained vectors
- Necessary to produce word embeddings
- Words with similar meanings are closer in space representation
- Capture semantics & syntax
- Boost performance in NLP tasks



Dataset

WHAT KIND OF DATA TO USE?

Different datasets can produce very different results:

- MetroLyrics (directly available)
- Kaggle (directly available)
- Genius (through API)
- Musixmatch (through API)

METROLYRICS

METROLYRICS

- +380000 lyrics
- Publicly available, already preprocessed data
- No punctuation
- No newlines
- Typing errors

GENIUS

- +190000 lyrics
- No publicly available, self-made
- Punctuation
- Newlines
- Rare typing errors



Preprocessing

HOW TO REWORK LYRICS
IN ORDER TO PRODUCE
GOOD RESULTS?

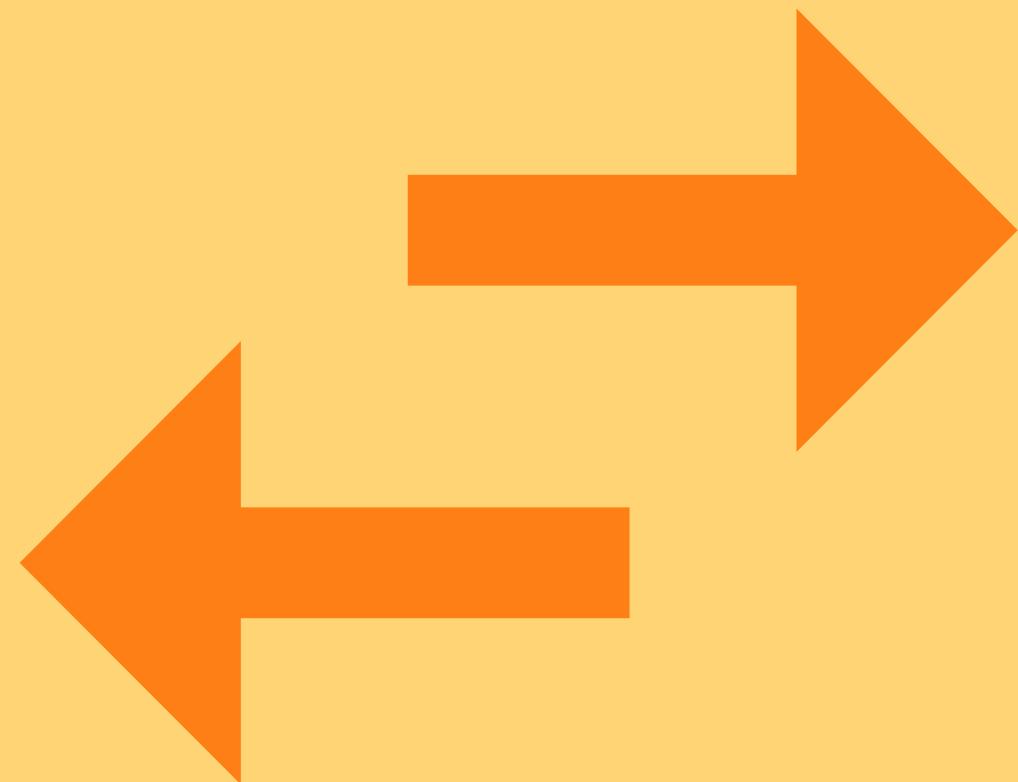
- Working with words or characters?
- A priori vocabulary or lyrics-based?
- Lowercased vocabulary?
- Remove non-english lyrics
- Remove stop words
- Handle punctuation and newlines to be tokenization-ready
- Make use of pre-trained word vectorization vocabulary*
- Rework out-of-vocabulary* words



Characters or words?

CHARACTERS

- ✓ Reduced domain
- ✓ Reduced training time
- ✓ No out-of-vocabulary handling
- ✗ Word misspellings*
- ✗ Non-existent words*
- ✗ Neither semantics nor syntax



WORDS

- Extended domain ✗
- Extended training time ✗
- Out-of-vocabulary handling ✗
- No typing errors ✓
- Word vectorization ✓
- Semantics & Syntax ✓



Word vectorization and vocabulary building

HOW PREPROCESSING IMPACTS RESULTS?

Working on lyrics can significantly improve the results as preprocessing affects:

- words domain
- word miss rate
- lyrics quality

WHAT WE OBTAINED

**A priori Word2Vec vocabulary and lyrics lowercased
(dataset without punctuation and newlines)**

RESULT:

Saturday was silent surely it moulding wrest Tollway Engineered monica
MUMMumbai Rade 1416 rests Pre-owned byes erases nhs nonfinancial
inDon x13 dumbest Everlane participatory Devore tortious Aaa PS Wiggin
Lubrication Injecting wack progressively msa anti-Semitism heart-
[...]

WHAT WE OBTAINED

A priori filtered Word2Vec vocabulary and lyrics lowercased
(dataset without punctuation and newlines)

RESULT:

Saturday was silent surely it geostationary downstate servitude flowchart
mauling ascertain thyristor exasperation constituents credentialed
shouldve word viewers psychomotor pseudonym crappy visor
antagonizing punish bodyguard steelers emphases columella sexton
[...]

WHAT WE OBTAINED

**Lyrics-based vocabulary and lyrics lowercased & reworked
(dataset without punctuation and newlines)**

RESULT:

Saturday was silent surely it had my sentence there is no joy but if only
fools wanted this mend this world with me i'm so is still ease for good if
not fair till same girl likes will be nice as the one across my fellow heart
her been waitin i said that hell can't help him the west is proud as blue as
[...]

WHAT WE OBTAINED

Lyrics-based vocabulary, lyrics partially lowercased & reworked (dataset with punctuation and newlines)

RESULT:

Saturday was silent surely I had just too focused to all is taking in
Heaven, burning of my day
Now I just wanna understand
Everything that makes
[...]

Genre

HOW IS GENRE ENCODED IN THE MODEL?

- One-hot-encoded genres
- Genre bypass embedding and LSTM layers
- The output of LSTM is concatenated with one-hot-encoding vectors

In this way we managed to produce a unique model for different genres, without producing specific genre-trained models.

Word

0, 0, 1, 0, 0, 0, 0, 0

Models

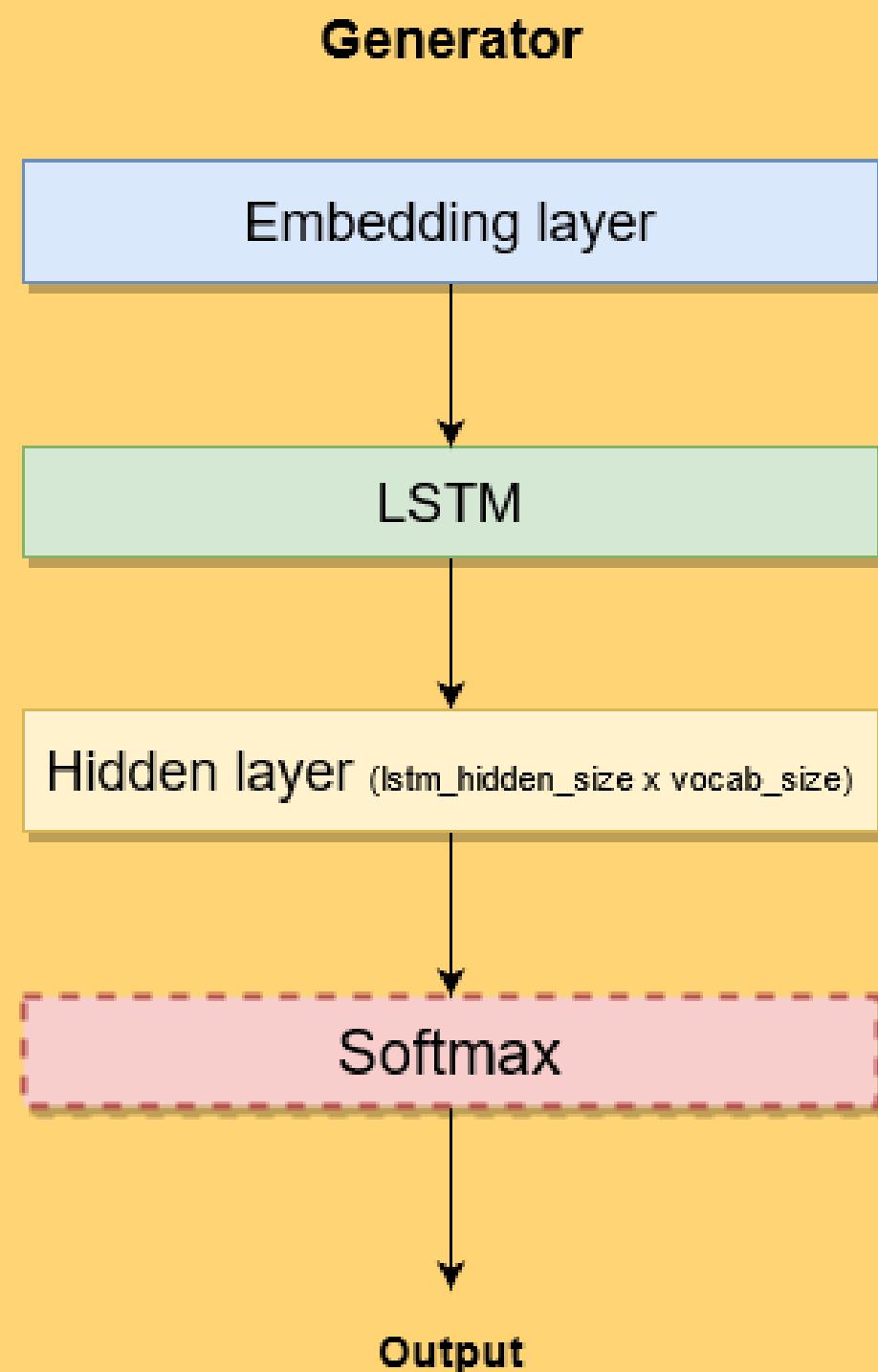
GENERATION AND DISCRIMINATION

- **Generator:** generates new realistic lyrics. We used this to produce our best lyrics
- **Discriminator:** tries to understand if passed lyrics are real or fake

Generative Adversarial Networks arise from the attempt to put together the two models.

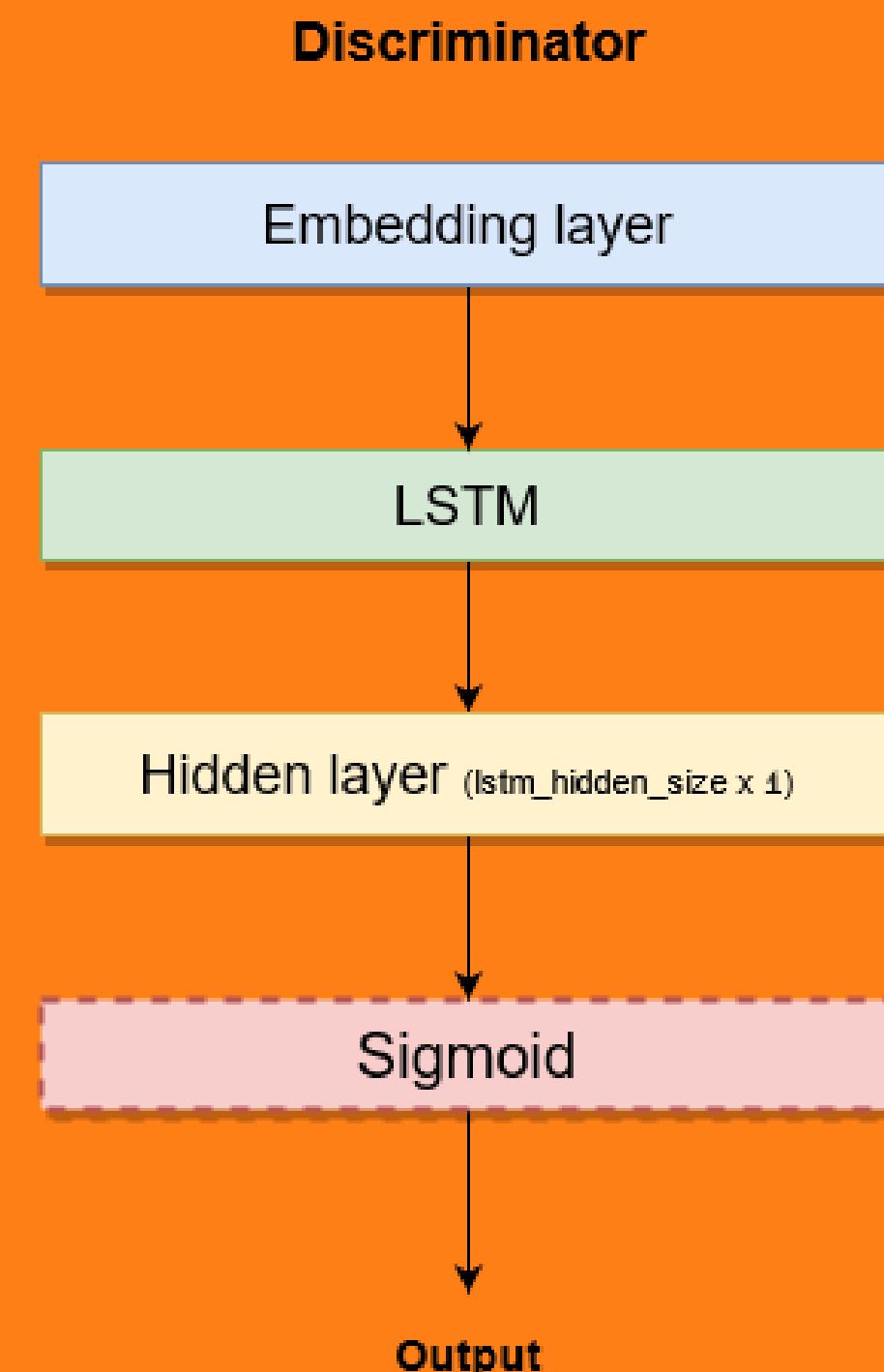
Generator

- Inputs are sequences of six terms
- Embedding layer convert word indices to word vectors
- LSTM extends dimensions, gaining informations from terms sequence
- LSTM output and genre one-hot-encoding are concatenated
- Hidden layer output size is vocabulary size
- Softmax activation function converts outputs to probability distribution in [0,1]
- Output is sampled from multinomial distribution in order to make the lyrics different



Discriminator

- Input size is the same as the one of the generator
- Three-layered LSTM
- LSTM output is concatenated with genre one-hot-encoding
- Hidden layer output size is one because the discriminator is a binary classifier
- Sigmoid activation function
- Output is a boolean value which tells if sequence is real or fake

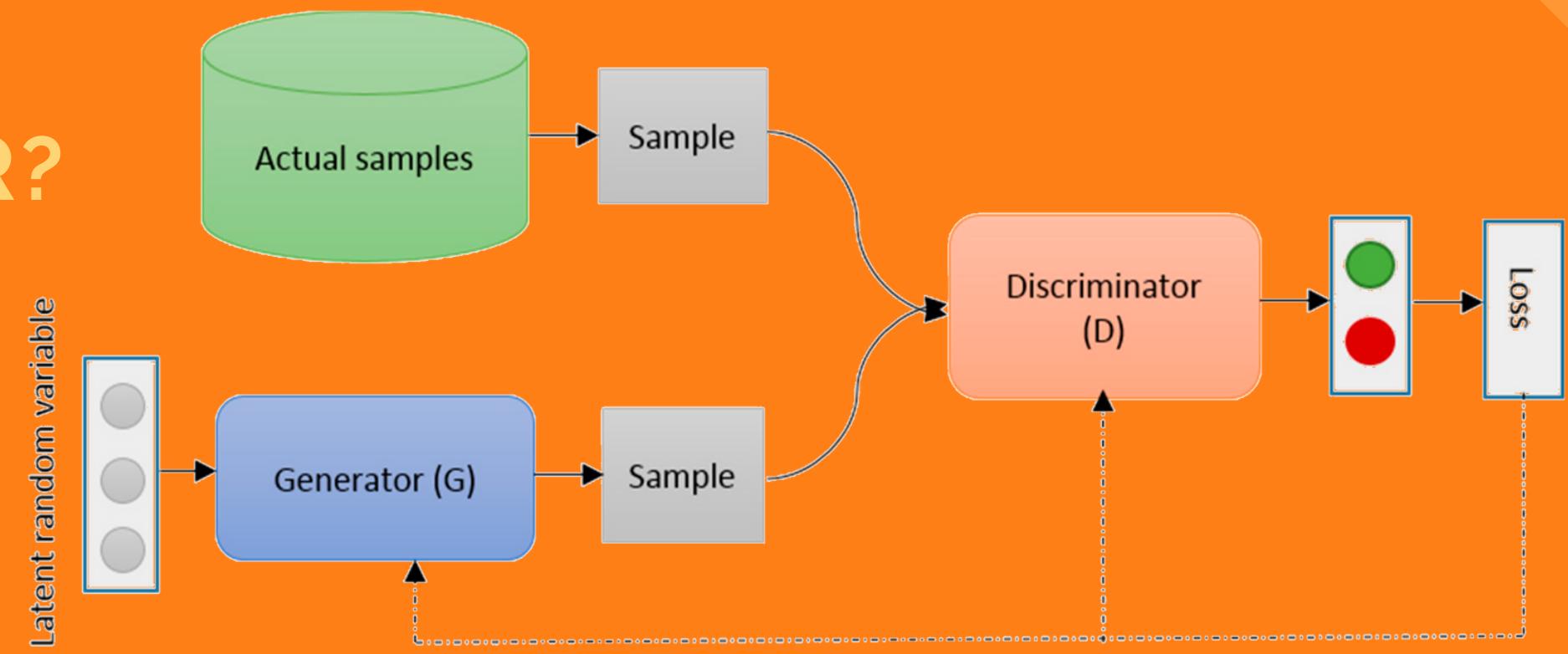


GAN Model

HOW WE PUT ALL TOGETHER?

We tried different approaches:

- Classic GAN
- Wasserstein GAN



Generator gradient is obtained from discriminator's loss
fed with fake samples and true targets, in the first case.

In the second case, we made use of the Wasserstein
Loss which tries to avoid problems with vanishing
gradients.

WHAT WE OBTAINED

Results

INPUT WORD: SAY

GENRE: ROCK

RESULT:

Say

I have started all the time keep it every times when I see your face,
slipping in front of my eye

Closer to your eye

Just falls away

Caught another escape

But I now, can come to your anger

[...]

WHAT WE OBTAINED

Results

INPUT WORD: SAY

GENRE: COUNTRY

RESULT:

Say that you again again

Even though you used love when it's true

Grudge, did not have

I don't need the

Man, they can never be found

It was her wrong though that it's work for a fella

But I was always busy enough

In the night with the I said, but I was angry

[...]

WHAT WE OBTAINED

Results

INPUT WORD: MAN

GENRE: HIP-HOP

RESULT:

Man

Enough to block up and they are too hard for the drugs and that body gonna work in

But you haven't get my stash all my life

Keep it, keep it life, too girl, ya, no niggas

You gotta get home, hah hah

I wasn't right, I was walking with with no cooling

[...]

Future developments

WHAT CAN WE DO IN THE FUTURE?

BALANCED DATASET

Currently, the dataset is not balanced by genre. A good move could be to extend it, balancing the lyrics

OUR OWN WORD2VEC

Currently, we are using a pre-trained word vector representation. Why not trying to construct it using our own dataset?

PREPROCESSING AND INPUT IMPROVEMENTS

Currently, sequences are truncated during generation to conclude a lyrics, windows are of fixed size and the input is the first word of the sequence. Varying this parameters could be a good idea

TRANSFER LEARNING

Models can be improved using transfer learning and fine tuning, maybe with advanced models like BERT

THE TEAM



IVAN FARDIN



CLAUDIU GABRIEL IVAN

A wide-angle photograph of a massive outdoor concert or event held in a large amphitheater or stadium. The foreground and middle ground are filled with thousands of people, their faces mostly obscured by the low light. The stage is visible in the center, with bright lights illuminating it from below. In the background, a massive firework display erupts from the center of the arena, sending numerous bright sparks and streaks of light into the dark sky. The overall atmosphere is one of a major, celebratory event.

Thank you for the attention!