

o amor é um veneno que me encontrou





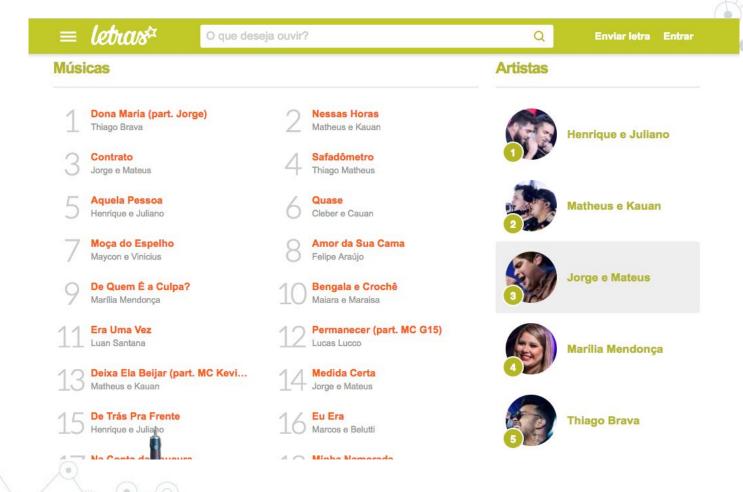
E a madrugada é feita com o cara de carnaval

Objetivos

- Extrair a essência da mais pura poesia brasileira a partir de 5MB de texto.
- Elaborar um script para obter a base de dados (i.e, letras de música) de um estilo específico. Este script entrará em um site de letras, buscará todos os artistas de um determinado estilo musical, buscará todas as letras de cada artista e as copiará para essa base de dados
- Modelar uma rede neural profunda de arquitetura recorrente e configurar seus hiperparâmetros para resultar em um bom gerador de letras de música.
- Treinar este modelo em Hardware de alto desempenho com a base de dados.
- © Gerar letras de música a partir do modelo treinado bem como entender e avaliar o processo de treino visando melhorar o modelo.

Eu sou a mão no joelhinho safado do meu coração, music_201.txt

Web Scrapping

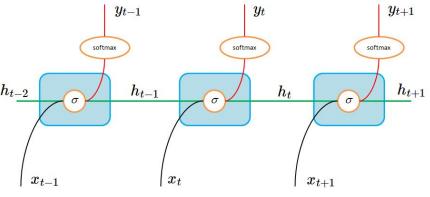


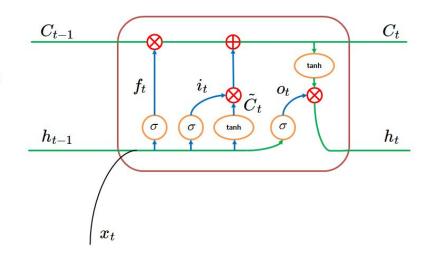
Eu sou o Menor pro mundão Sou do cheiro de peixe, music_192.txt

Modelagem da Rede Neural

Layer (type)	Output Shape	Param #	
Istm_1 (LSTM)	(None, None, 700)	2326800	
dropout_1 (Dropout)	(None, None, 700)	0	
lstm_2 (LSTM)	(None, None, 700)	3922800	
lstm_3 (LSTM)	(None, None, 700)	3922800	
time_distributed_1 (T	imeDist (None, None, 13	0) 91130	
activation 1 (Activation	on) (None, None, 130)	0	

Total params: 10,263,530 Trainable params: 10,263,530 Non-trainable params: 0





Eu sou pobre, mas eu sou gordinho, music_230.txt

Modelo em treinamento

•
"ãÊnnnn@@@@&&&&&fffffffóó
óóóóóóîîîîccccîùùùù
(((((!!!!!!ttttttttUUUUUUUyy33h3
355BBBB
âââââ/////ÊÊÊÊÊ_**nnnnn
nnÇÇÇÇxxxxÒÒÒ###\êêêÊ
_rrrrnnnnnnnnnnnÇÇÇÇxxxxa
aaaaaaaaZaaaa;;;;;;;;
òòòòòòòòòüüüüüüüüüü8888
8"""3}}";;;;;;
òòòòòòòòòüüüüüüüüüü88888"
""}
òòòòòòòòòüüüüüüüüüü8888
8"""}
òòòòòòòòòüüüüüüüüüü8888
8"""}
òòòòòòòòòüüüüüüüüüü8888
8"""}
òòòòòòòòòòüüüüüüüüüü8888

Epoch#0

Epoch#12

la que eu sou funkeira Vai ter que aquentar O baile tá uma uva Se eu te pego Se eu te pego

Epoch#34

inha meu marido dela Aí eu lembro do corpinho dela A mulherada se amarra ela guer o homem pra beijar Então mentaliza na batida dP Graças a Deus e amenizo A minha mulher é feiticeira A minha sogra é fofoqueira A minha sogra é macumbeira A minha sogra é fofoqueira A minha sogra é macumbeira A minha sogra é fofoqueira A minha sogra é macumbeira A minha sogra é fofoqueira A minha sogra é macumbeira A minha sogra é fofoqueira A minha sogra é macumbeira

MC Maique do Marquinho

Modelo gerador de letras de funk brasileiro "w é uma linda mulher

Me pega com jeitinho que eu to cheia de cachorro

Acha que tá com a galera

Vaza!

Eu gosto de você

Meu Deus que boquinha

Chama as gatas no rasante

De santa de Cidade de Deus

E o Buiú esta com Deus

E o Buiú esta com Deus

E o Buiú esta com a gente

Se eu te pego de jeito gata eu ti do o meu piru de boca

Para com a bala pra voar

Vida loka paga seu sublino

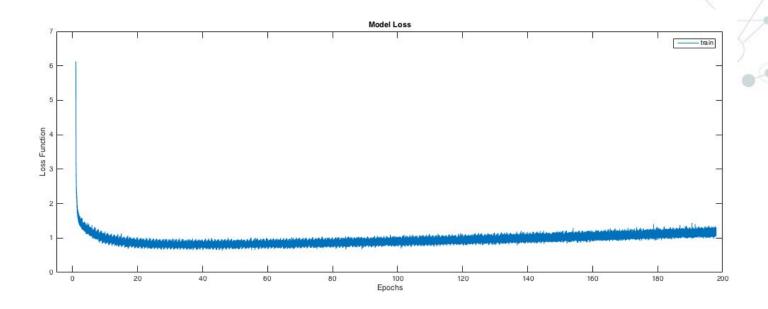
E a novinha tá muito louca

Eu sou MC Maique do Marquinho

Vou confessar a verdade

Sempre de olho grande nos olhos e comerar os m"

E a solidão e a minha mente forma de espuma, music_185.txt



- 200 epochs
- Batch Size = 50
- Otimizador = RMSprop
- Epoch = 3000 segundos; aprox. 6 dias de treinamento
- GPU: GeForce GTX 1080 Ti

Zé Qualquer do Mato

Modelo gerador de letras de sertanejo

"Zé Qualquer do mato

Do mal luero tirano, pensando que já cansada de me atender

O tempo passa e ela desenhava

O mal se paga com o seu poder

Mais eu não importa o que vai acontecer

Você é o ar mas que ele traz no peito e um cara pra chamar

de seu

Eu te avisei meu amor,

quem sabe um dia eu perdi a correr

E sempre procurando outra emoção

Meu coração sempre esteve aqui no camaro amarelo

E agora você vem, né?

Agora você quer?

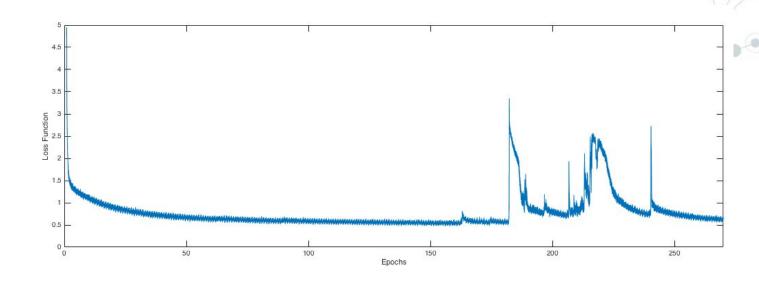
Só que agora vou escolher

Tá sobrando mulher

Agora eu faço tudo que você di"



O céu está escrito nas madrugadas, music_260.txt



- 270 epochs
- Batch Size = 128
- Otimizador = Adam
- Epoch = 1000 segundos; aprox. 2 dias de treinamento para 200 epochs
- GPU: GeForce GTX 1080 Ti

E a minha sogra tinha um lança pra gente se arrumar, music_179.txt

Conclusões e Trabalhos Futuros

- Os modelos aqui apresentados foram capazes de aprender o idioma, de forma a cometer poucos erros ortográficos (e os cometidos se devem apenas aos erros contidos no próprio dataset de treino), extrair padrões de versificação e rimas. Por outro lado, os modelos ainda sofreram com o problema de overfit - o que pode gerar sérios problemas de plágio.
- Trabalhos Futuros:
 - Regularização + Fine-tuning dos hiperparâmetros
 - Classificador de palavras (OBS: curse of dimensionality e a solução por embeddings)
 - Aprendizado semi-supervisionado: GANs e Variational Autoencoders

Um beijo na boca E desce um combo de Red no delegado Ele que estava mais apaixonado Pois eu não desisti de te amar Não tenho que achar ninguém pra ver novamente Mas uma noite foi pouco Cê tá se achando, é? Aproveita que ela vai ter fim. Ela entrou em minha vida E me entreguei

<u>Final de semestre</u> que chorava já perdeu a linha, Você não pode para me perdoar

Eu te peço perdão
Vai me perdoando, mas tem linha e cheiro de explosão
Eu senti naquela hora
Achei que era saudade
Por mim faltaram rumo amigo