



GERAÇÃO DE LETRAS DE RAP A PARTIR DE UMA REDE NEURAL

Júlio César de Sousa





Motivação

- Importância do RAP para a cultura nacional;
- Relevância do Racionais MCs;
- Afinidade com o tema;
- Desafio na geração de letras coesas;
- Inserção do underground no meio acadêmico.



Objetivos

- Utilizar uma rede LSTM para gerar texto longo;
- Treinar um modelo para compor músicas;
- Enviar este modelo para o universo Racionais MCs;
- Gerar novas letras de RAP;
- Participar de algum evento relacionado ao tema.

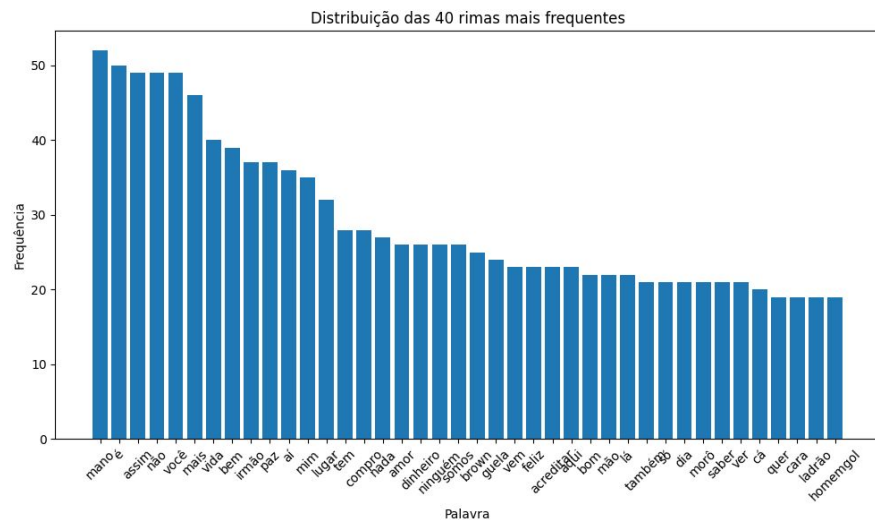


Base de dados

- 104 letras de música do Racionais disponíveis em [Letras.mus.br];
- As letras foram disponibilizadas em [JotaChina];
- O arquivo possui 9010 linhas;
- É baixado para o ambiente do notebook através da biblioteca `urllib.request`.

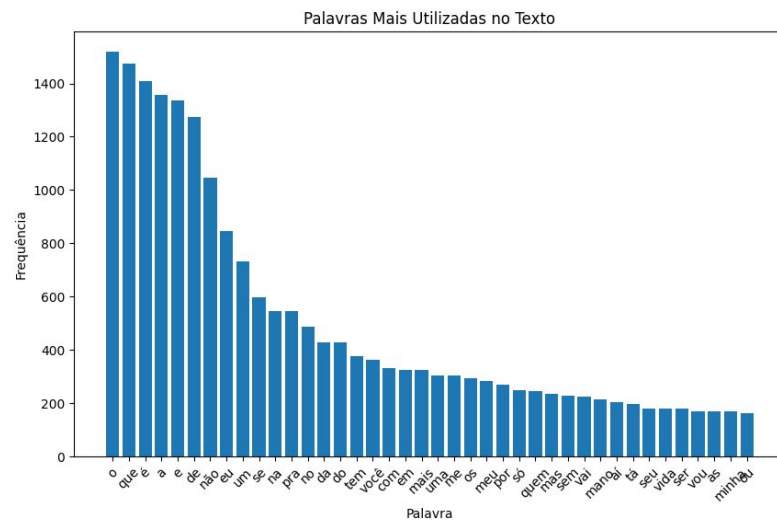
Base de dados

Figura 1 - 40 rimas mais frequentes



Fonte: O autor. (2024)

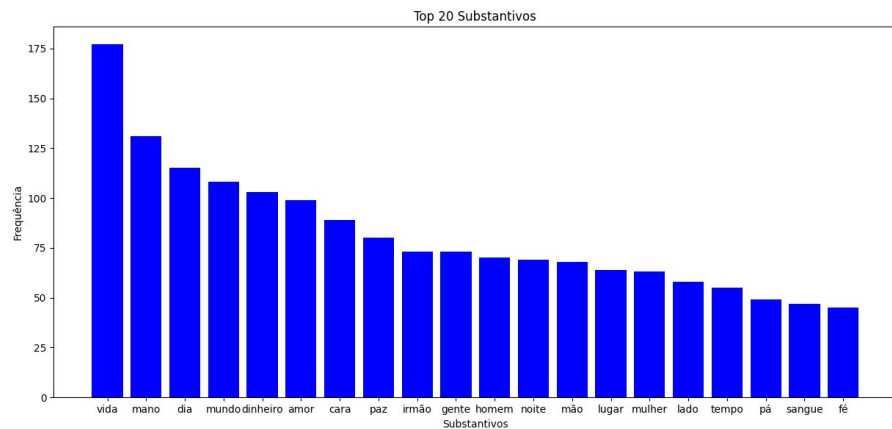
Figura 2 - Palavras mais utilizadas



Fonte: O autor. (2024)

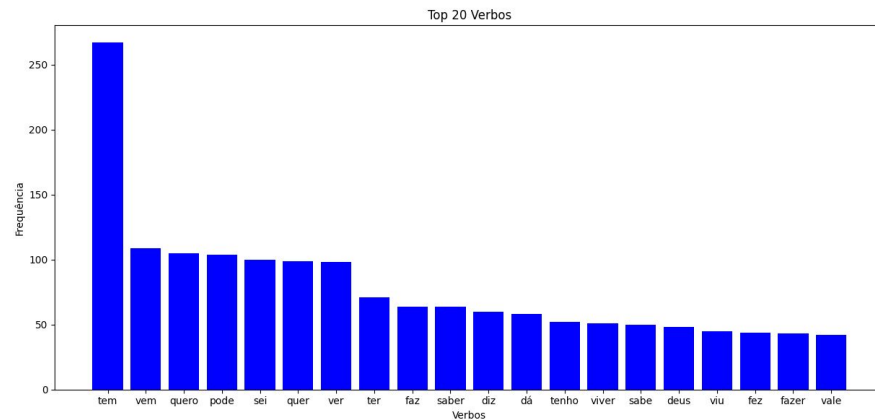
Base de dados

Figura 3 - Top 20 substantivos



Fonte: O autor. (2024)

Figura 4 - Top 20 verbos



Fonte: O autor. (2024)



Pré-processamento

- Remoção de caracteres não alfanuméricos e não pertencentes ao utf-8;
- Unificação em um único corpus;
- Tokenização com [TENSORFLOW];
- Criação das sequências de entrada utilizando o método n-grams
 - começando com tamanho 2, até o comprimento total da linha.
- Pad à esquerda para garantir o mesmo tamanho em todas as sequências.



Abordagens anteriores

- Ajustes finos em modelos pré-treinados;
- Enviesando o modelo para o universo Racionais MC's;
- Geração de novas letras de música
 - Mais coesão pelo fato do modelo pré-treinado; e
 - Geração de texto mais robusto.
- Limitações
 - Hardware e ambiente de execução.



Modelos pré-treinados

- Podem facilitar para obter resultados mais robustos;
- Já possuem uma arquitetura pronta para resolver o problema;
- Motivos cruciais da não utilização desta abordagem
 - Limitação de hardware; e
 - Índices de acurácia abaixo do esperado.



Modelos pré-treinados

Os modelos abaixo foram utilizados em tentativas anteriores

- [BERTimbau]
 - Predição de labels.
- [GPTuguese-2]
 - Geração de texto.

Ambos os modelos suportam o idioma Português brasileiro.

Testes de predição com BERTimbau

Figura 5 - Teste de predição com o label “A [MASK] cheia, clareia as ruas do capão”

```
[{'score': 0.22971616685390472,
  'token': 13943,
  'token_str': 'lua',
  'sequence': 'A lua cheia, clareia as ruas do Capão'},
{'score': 0.21356083452701569,
  'token': 1991,
  'token_str': 'água',
  'sequence': 'A água cheia, clareia as ruas do Capão'},
{'score': 0.06903950870037079,
  'token': 2954,
  'token_str': 'noite',
  'sequence': 'A noite cheia, clareia as ruas do Capão'},
{'score': 0.05121883377432823,
  'token': 596,
  'token_str': 'tempo',
  'sequence': 'A tempo cheia, clareia as ruas do Capão'},
{'score': 0.04503905773162842,
  'token': 7856,
  'token_str': 'Lua',
  'sequence': 'A Lua cheia, clareia as ruas do Capão'}]
```

Fonte: O autor. (2024)

Figura 6 - Teste de predição com o label “hoje eu sou ladrão, [MASK] 157”

```
[{'score': 0.06270167976617813,
  'token': 15212,
  'token_str': 'tenho',
  'sequence': 'Hoje eu sou ladrão, tenho 157'},
{'score': 0.05100703611969948,
  'token': 123,
  'token_str': 'a',
  'sequence': 'Hoje eu sou ladrão, a 157'},
{'score': 0.049017492681741714,
  'token': 712,
  'token_str': 'aos',
  'sequence': 'Hoje eu sou ladrão, aos 157'},
{'score': 0.04188200458884239,
  'token': 4319,
  'token_str': 'artigo',
  'sequence': 'Hoje eu sou ladrão, artigo 157'},
{'score': 0.023356173187494278,
  'token': 1328,
  'token_str': 'art',
  'sequence': 'Hoje eu sou ladrão, art 157'}]
```

Fonte: O autor. (2024)



Testes de geração de texto com GPortuguese-2

Figura 7 - Prompt “A lua cheia, clareia as ruas do Capão”

>> Generated text 1

A lua cheia, clareia as ruas do Capão, que é a primeira cidade de Minas Gerais, a ter suas terras e o maior porto do país (sendo três das mais importantes do mundo). A capital das Minas foi em 1843, por

>> Generated text 2

A lua cheia, clareia as ruas do Capão da Canoa, em meio a muitos tiros a que seus rivais gritaram. Durante o jogo seguinte os clubes se enfrentaram com as torcidas rivais, se tornando: Galondo é uma

Fonte: O autor. (2024)

Figura 8 - Prompt “Tem que ser, tem que pá”

>> Generated text 1

Tem que ser, tem que pá de sua própria, se contar em um lado e se aproxime de dila o outro, quando o morta se vê na cidade ser a sua al

>> Generated text 2

Tem que ser, tem que pá! e se chamar “A.L.A.” Não se sabe onde o clipe foi filmado. As cópias lançadas foram lançadas pela Sony em 7 de abril de 2016, e foram oficialmente lançadas em 5

Fonte: O autor. (2024)

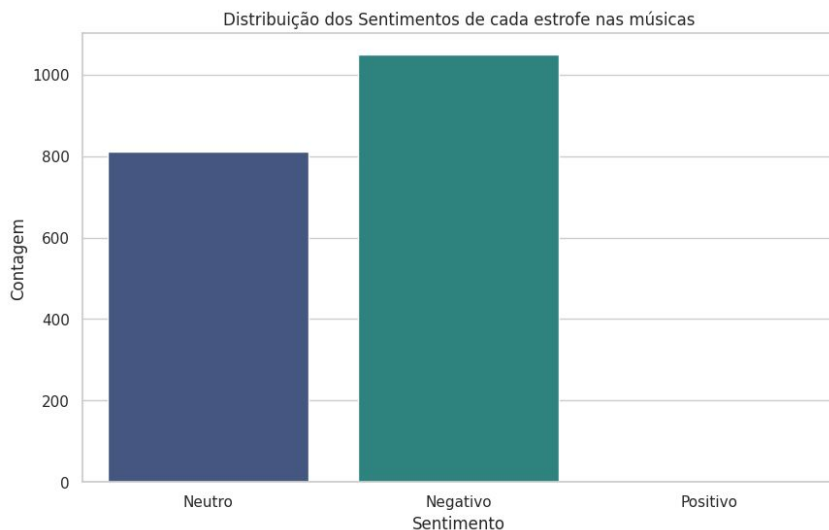


Análise de sentimento

- Descobrir o sentimento preponderante no corpus;
- Utilizá-lo para influenciar na geração das novas letras;
- Feito utilizando [LEME];
- Trouxe informações sobre a base, mas não teve efetividade na aplicação.

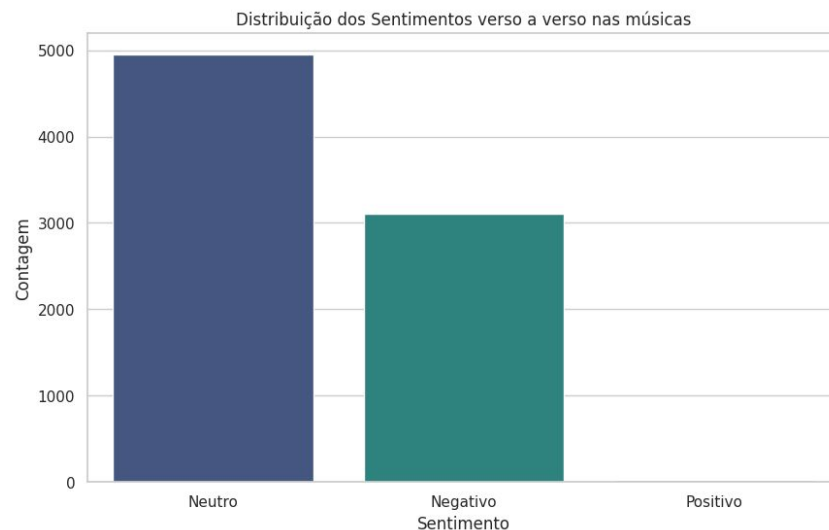
Análise de sentimento

Figura 9 - Sentimentos em blocos de 4 versos



Fonte: O autor. (2024)

Figura 10 - Sentimentos verso a verso



Fonte: O autor. (2024)



Abordagem atual, LSTM

- Camada de embedding, onde a saída é um tensor tridimensional.
 - Tamanho 250.
- Camada LSTM, que retorna sequências completas, também tem como saída um tensor tridimensional
 - 200 neurônios.
- Camada LSTM, retorna apenas a saída do último passo de tempo, resultando em um tensor bidimensional
 - 100 neurônios.
- Camada densa, com função de ativação softmax para calcular as probabilidades da próxima palavra gerada na sequência
 - Número total de palavras únicas (7799) de neurônios.



Métricas de treinamento

- O treino foi feito em 100 épocas;
- Utilizou-se um batch size de 128;
- A acurácia das predições foi uma métrica de avaliação;
- Assim como a perda.

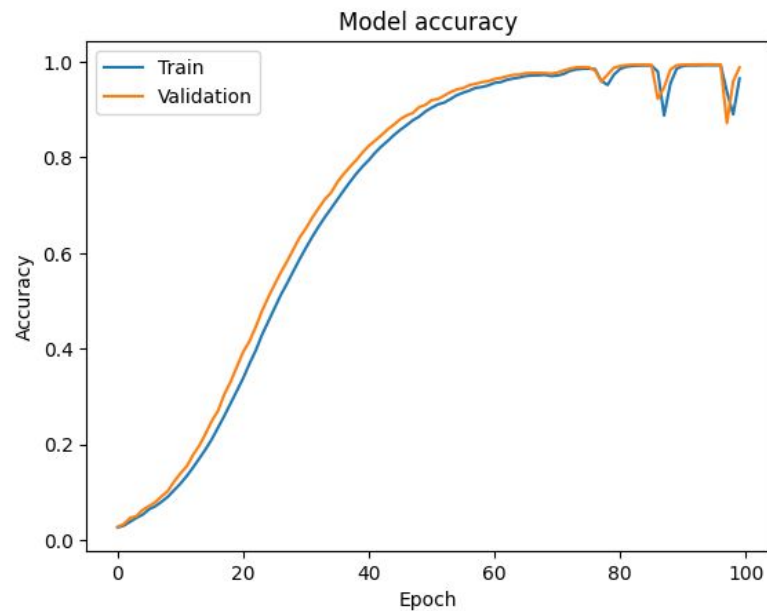


Resultados

- Acurácia do protótipo final
 - Treino: 0.9928; e
 - Validação: 0.9947.
- Perda do protótipo final
 - Treino: 0.0336; e
 - Validação: 0.0269.

Resultados

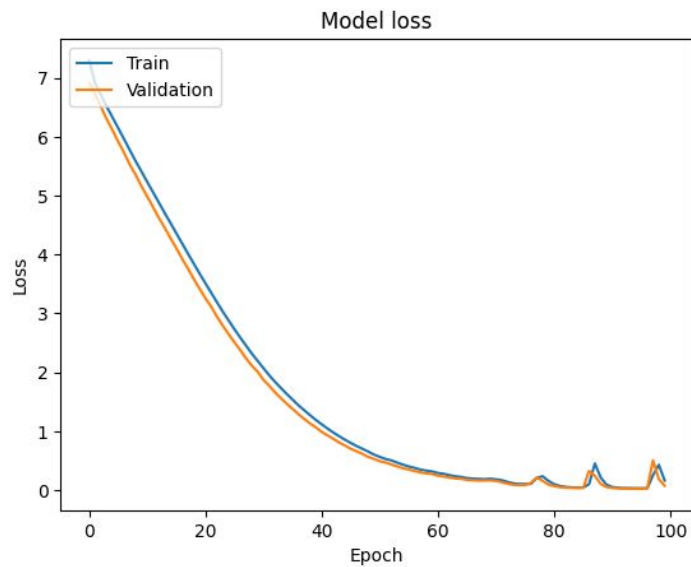
Figura 11 - Gráfico de acurácia



Fonte: O autor. (2024)

Resultados

Figura 12 - Gráfico de perda



Fonte: O autor. (2024)



Resultados

Figura 13 - Letras geradas a partir do prompt “Tem que ser, tem que pá”

```
Tem que ser, tem que pa  
me deixou feliz refrão pelas olhar vice favela  
favela outro vergonha que se hoje aqui tem uns  
uns direitos em mais na gente favela tem shopping  
  
shopping o gente veio graças a paz vai o  
o povo para o sistema é deus certo meu  
meu maluco tá na solidão que eu conheço a  
a vida todo dia igual até no final violento  
  
violento é quente nem fica com muita treta e  
e você é o lugar então ontem o que  
que pensa um de levantar só acredito de mim  
mim um irmão na minha liberdade morreu registrar registrar  
  
registra dá o rap não tira o que tem  
tem 4 pano é claro de me salvar morra  
morra câmara câmara câmara câmara câmara câmara câmara câmara  
câmera mano mó raça sonha para eu já não  
  
panão entrando por amor carcerário dinheiro já é o  
o dinheiro que você não entra quando confia e  
e uma oportunidade tudo de acudi acudi acudi acudi  
audi sem cola 4 pessoas várias cor de madrugada
```



Resultados

Figura 14 - Letras geradas a partir do prompt “Cotidiano na periferia e as dificuldades”

```
Cotidiano na periferia e as dificuldades
a cadeia e alguém hã nego te deixa
deixa porque que e ontem eu quero esse xá
x2 igual o frio ele era os manos na

na lado do outro for ele vai em onde
onde para um monstro e da lei da selva
selva é necessário aqui não sabe disso se não
não vivo qual lá é hoje cê tem mais

mais que ninguém quer ter mais dólar e sair
sair os sério qualquer pior do capão cinco na
na esquina no remetente e ele morre o minha
minha mão não quer mais as famílias dinheiro eu

eu conseguir um corre por você e os manos
manos mas sem gamba gamba gamba gamba gamba
gambé vem no chão tem continuar se vale a
a cara é o sangue dele perde a madrugada

madrugada na distância nível a alma me mostra a
a ingrata vela hoje a noite era hoje que
que não cachimbar de lutar complexo polícia pobres essa
essa final tio e mulher entre uma solução na
```



Resultados

Figura 15 - Letras geradas a partir do prompt “A lua cheia, clareia as ruas do Capão”

```
A lua cheia, clareia as ruas do capão
acima de nós só deus humilde é nós
nós fica perdida mulheres ligado vive se têm a
a vida subiu a boca eu tive meu corpo

corpo a testemunha é calada de asfalto te imita
imita desprezo então o seu osso eu sou pôr
pôr na sarjeta no barril barril barril barril barril
brasil aí é se erreí para você gosta de

de meu pai tá as flores da ponte para
pra ser não vou encontrar o cá agora precisa
precisa não vou dizer ninguém ninguém tá na rua
rua não foi embalado embalado embalado embalado embalado embalado

embaçado por ceguinho ceguinho ceguinho ceguinho ceguinho ceguinho
neguinho na selva mais o que vinho acha que
que é bem ganharam aqui você vai sabe o
o tempo será de lá vai ser a fita

fita ao problema será de justiça irmão meu rei
rei a lua escura no homem para to no
no coração uma próprias lágrimas la uma viela em
em seu todas de novo do barril barril barril
...

mó vela sem querer ele vai vai mesmo foi
```



Conclusões

- A geração de letras de música é uma tarefa árdua;
- A utilização de uma rede neural para essa finalidade é um desafio peculiar;
- A forma em que a letra é gerada pode interferir no resultado;
- LSTM x modelos pré-treinados
 - Limitações; e
 - Vantagens.
- Campo de pesquisa interessante, onde há um enorme leque de possibilidades.



Referências

- Letras.mus.br. Racionais mc's. Disponível em: <https://www.letras.mus.br/racionais-mcs/>. [Acesso em 21 jun. 2024];
- JotaChina. Computer intelligence. Disponível em: <https://raw.githubusercontent.com/JotaChina/Computer-intelligence/update%231/TP2/RacionaisLyrics.raw>. [Acesso em 21 jun. 2024];
- LEME, Lucas. FinBERT-PT-BR. Hugging Face, 2024. Disponível em: <https://huggingface.co/lucas-leme/FinBERT-PT-BR>. Acesso em: 22 jun. 2024;
- TENSORFLOW. Tokenizer. TensorFlow, 2024. Disponível em: https://www.tensorflow.org/api_docs/python/tf/keras/preprocessing/text/Tokenizer. Acesso em: 22 jun. 2024;
- Souza, F., Nogueira, R., and Lotufo, R. (2020). BERTimbau: pretrained BERT models for Brazilian Portuguese. In 9th Brazilian Conference on Intelligent Systems, BRACIS, Rio Grande do Sul, Brazil, October 20-23 (to appear). Acesso em: 28 jun. 2024;
- Guillou, P. (2020). Gportuguese-2 (portuguese gpt-2 small): a language model for portuguese text generation (and more nlp tasks...). Acesso em: 28 jun. 2024.