

GERAÇÃO DE LETRAS DE RAP A PARTIR DE UMA REDE NEURAL

Júlio César de Sousa¹

¹ Departamento de Ciência da Computação
Universidade Federal de São João del-Rei, Minas Gerais, Brasil
sousacj1@gmail.com

Abstract. *Musical production is a complex task that has increasingly benefited from and been supported by auxiliary technologies. Crafting song lyrics involves both logical exercises and the expression of essence. The underground¹ scene brings with it originality and a critical perspective on everyday situations, often marginalized due to its vocabulary or sensitive themes. This study explores the training of a Long Short-Term Memory (LSTM)² for automated generation of RAP³ lyrics, investigating the challenges and opportunities associated with this innovative approach. To train this neural network, lyrics from the Racionais MC's group were used, as a tribute to the significant symbol the group represents for national RAP.*

Resumo. *A produção musical é uma tarefa complexa e vem sendo cada vez mais aprimorada e amparada por tecnologias auxiliares. A composição de letras de músicas é um exercício lógico, além da expressividade e essência transcrita. A vertente underground⁴ traz consigo a originalidade e um ponto de vista crítico sobre situações cotidianas e, geralmente, é marginalizado pelo vocabulário ou com temas sensíveis. Este estudo explora o treinamento de uma LSTM⁵ para geração automatizada de letras de RAP⁶, investigando os desafios e as oportunidades associadas a essa abordagem inovadora. Para alimentar essa rede neural foram utilizadas letras de música do grupo Racionais MC's, como forma de homenagem ao símbolo marcante que o grupo representa para o RAP nacional.*

1. Introdução

O RAP, como manifestação cultural, desempenha um papel fundamental na expressão e na narrativa das comunidades urbanas, transmitindo histórias e experiências profundamente enraizadas em questões sociais, políticas e pessoais. Grupos icônicos como o

¹In this context, referring to the marginal musical style, outside the mainstream and the standards imposed by the major music and recording industry. Genres like RAP, Funk, Drill, Grime, etc., fall into this category.

²A type of recurrent neural network architecture used in various Natural Language Processing scenarios.

³Rhythm and Poetry, the verbal musical manifestation of the underground music scene.

⁴Neste contexto diz respeito ao estilo musical marginal, a margem do mainstream e dos padrões impostos pela grande indústria fonográfica e musical. Gêneros como o RAP, Funk, Drill, Grime, etc.

⁵Memória de curto longo prazo, é uma arquitetura de rede neural recorrente que é usada em diversos cenários de Processamento de Linguagem Natural.

⁶Ritmo e poesia, manifestação verbal musical da vertente musical underground.

Racionais MC's se destacam não apenas pela maestria musical, mas também pela capacidade de traduzir em versos a realidade das periferias e das minorias, oferecendo voz e representação para aqueles que muitas vezes são marginalizados pela sociedade. No entanto, o processo criativo por trás da composição de letras de RAP pode ser desafiador e exigente. Os artistas enfrentam a pressão de criar versos que sejam autênticos, coesos e relevantes para o público, ao mesmo tempo em que mantêm a essência e a identidade do gênero musical. Nesse contexto, a automação da geração de letras de RAP por meio de redes neurais surge como uma abordagem inovadora que tem o potencial de simplificar o processo criativo e democratizar a produção musical.

A implementação de uma rede neural capaz de gerar letras de RAP não apenas oferece uma ferramenta valiosa para artistas e produtores musicais, mas também levanta questões importantes sobre a autenticidade e a essência do gênero. Como garantir que as letras geradas mantenham a integridade lírica e o contexto cultural do RAP? Como incorporar as nuances e o estilo único de grupos como os Racionais MC's na geração automatizada de versos? É possível extrair alguns insights valiosos em letras existentes por meio de técnicas de aprendizado de máquina. Ao analisar padrões de rimas, temas recorrentes e sentimentos encapsulados nos versos, podemos não apenas gerar novas composições, mas também compreender melhor a natureza e a evolução do RAP como forma de expressão artística e cultural.

2. Referencial teórico

É importante observar estudos que já foram feitos para finalidades similares, pois o entendimento sobre as problemáticas pode ser facilitado. Deve-se levar em consideração que o intuito desta abordagem é tentar simular a geração de letras de música, com a finalidade de investigar a capacidade da rede neural em inferir na geração de novas composições.

2.1. Racionais MC's

Racionais MC's é um dos grupos mais influentes da história do RAP nacional brasileiro. Formado por Mano Brown, Ice Blue, Edi Rock e KL Jay na década de 1980 em São Paulo, o grupo ganhou destaque por suas letras profundas e críticas, que abordam questões sociais, raciais e políticas do Brasil urbano. A importância do grupo para o cenário underground e o RAP nacional é inegável. Eles foram pioneiros em trazer uma narrativa contundente sobre a realidade das periferias brasileiras para o mainstream musical, ampliando o alcance e a relevância do RAP como forma de expressão cultural e social.

Com discos como "Sobrevivendo no Inferno" (1997), "Nada Como Um Dia Após o Outro Dia" (2002), o grupo não só consolidou-se como ícone do RAP nacional, mas também abriu caminho para uma nova geração de artistas que continuam a explorar temas de marginalização, desigualdade e resistência por meio da música. Além disso, o grupo teve um papel crucial na conscientização e mobilização de jovens das periferias, oferecendo uma voz poderosa para aqueles que frequentemente são marginalizados pela sociedade brasileira.

A história do grupo é rodeada de capítulos de resistência e manifestações em prol de igualdade, mas também tem uma força cultural que desafiou e continua desafiando as normas estabelecidas, inspirando não apenas músicos, mas também toda uma comunidade a se expressar, lutar por seus direitos e dignidade. Construindo um legado único

e fazendo uma escola onde a massiva maioria dos artistas underground brasileiros tiram suas referências.

2.2. O RAP como manifestação cultural e musical

Expressão verbal da vertente Hip hop, emergiu como um componente da cultura underground para narrar a marginalização, preconceito e repressões impostas aos adeptos do estilo de vida underground. Na costa oeste dos Estados Unidos⁷ que o RAP ganhou força e representatividade para se tornar um dos gêneros musicais mais ouvidos do mundo. O desenvolvimento tecnológico desempenhou um papel crucial nessa evolução. Em Compton, na década de 1980, o gangsta RAP começou a florescer como uma vertente do RAP que explorava as realidades duras e muitas vezes violentas da vida nas comunidades urbanas.

Figuras emblemáticas como Dr. Dre, Eazy-E e Ice Cube, foram pioneiros desse subgênero, destacando questões sociais e experiências pessoais, onde contavam histórias de preconceito e opressão que eram submetidos. Dr. Dre, um dos principais produtores e artistas do RAP, teve um impacto significativo na evolução sonora do gênero. Sua utilização inovadora de tecnologias de produção, como a mesa de mixagem SSL (Solid State Logic) e samplers, contribuiu para criar batidas distintas e produções de alta qualidade. O G-Funk, subgênero derivado do gangsta RAP, popularizou o uso de sintetizadores e samples de Funk. Dr. Dre, mais uma vez, foi fundamental na introdução de sintetizadores pesados e batidas groove, como evidenciado no álbum "The Chronic" de 1992, estabelecendo uma sonoridade característica do RAP da Costa Oeste.

A ascensão da MTV⁸ e a disseminação de videocliques na década de 1980 também desempenharam um papel crucial na popularização do RAP. Artistas como N.W.A, Wu Tang Clan, Racionais MC's, Facção Central, 509-E e outros tantos mais, aproveitaram os videocliques para narrar visualmente suas histórias e expressar a realidade das comunidades urbanas. Com a popularização do acesso a outras formas de streaming e a redução em adesão da TV como forma de consumo, a disseminação de músicas online através de plataformas de streaming e redes sociais permitiu que artistas de RAP atingissem audiências globais, independentemente de sua localização geográfica ou apoio das grandes gravadoras da indústria fonográfica.

2.3. Desafios na composição de letras de RAP

Os desafios na composição de letras de RAP a partir de redes neurais consistem na contextualização do ponto de vista marginalizado e as mensagens que são passadas nas entre-linhas das letras. As histórias que são contadas nas composições geralmente tem uma narrativa coesa e realmente conta uma história⁹. Principalmente as letras do Racionais, que é o universo de pesquisa deste estudo, que sempre tenta trazer uma reflexão sobre a vida do crime e seus malefícios, em busca da conscientização dos ouvintes e apaziguar os conflitos entre os envolvidos neste cenário.

Além do desafio de ter coesão textual e ser uma história narrada, uma problemática é a geração de rimas para que possa haver fluidez lírica e, desta forma, ser mais similar a

⁷Particularmente em Compton, Califórnia.

⁸MTV é um canal de televisão norte-americano básico por cabo e satélite, de propriedade da Paramount.

⁹Que tenha início, meio e fim. Narre um acontecimento, de fato.

uma composição de uma letra de RAP. O uso de uma LSTM pode contribuir para tentar resolver este problema, visando mapear o estilo de escrita original, para uma composição mais próxima possível da original.

2.4. Automação da geração de letras de RAP por meio de redes neurais

A tarefa de automatizar a geração de novas linhas de RAP surgiram a partir da vivência do autor, além disso, a inserção no meio underground. Entretanto, a vontade de homenagear um dos maiores grupos de RAP da história, foi uma grande motivação para utilizar o grupo como tema central deste estudo. Esta tarefa não é simples, pois foi preciso uma sequência de etapas que serão descritas na seção 4 para baixar a base de dados, escolha da arquitetura da rede neural, dentre outras escolhas que foram tomadas.

Particularmente, destaca-se o trabalho [Barrat 2018], cuja abordagem está diretamente relacionada ao campo de pesquisa. Além disso, os conceitos trazidos por [Juliani 2019], que exploram a geração de texto a partir de letras musicais utilizando redes neurais, onde a utilização de uma LSTM auxilia neste processo, de forma que motivou o desenvolvimento deste estudo. Outros estudos que são interessantes e serviram de referências foram [Anand et al.], [Satija 2017], [Draca 2017]. Por fim, [Potash et al. 2015] que é uma referência em comum dentre os estudos voltados a geração de letras musicais utilizando técnicas de aprendizado de máquina.

3. Sobre a base de dados

A base de dados utilizada para este projeto consiste em todas as letras de músicas do Racionais MC's que foram disponibilizadas em [Letras.mus.br 2024], até o momento deste estudo, um total de 104 músicas. Para facilitar sua incorporação ao projeto, este arquivo é baixado, por meio de funções da biblioteca `urllib.request`, de um arquivo RAW no repositório GitHub deste projeto, disponível em [de Sousa 2024a]. O arquivo possui 9010 linhas, após ser pré-processado totalizam-se 8087 linhas, onde cada uma destas é um verso de uma música do grupo. Optou-se por manter interlúdios e introduções, para dar mais dados e tentar auxiliar no ganho de contexto e coesão do modelo para o universo de interesse. A escolha inicial das letras do Racionais MC's foi motivada pela extensão e pela diversidade de temas abordados, tornando-as adequadas para análise. No entanto, a necessidade de compreensão do contexto das letras pelo algoritmo apresenta desafios adicionais.

3.1. Pré-processamento da base de dados

O pré-processamento da base de dados consiste em baixar o arquivo supracitado, fazer a remoção de qualquer caractere que não seja alfanumérico ou que faça parte da codificação UTF-8 que represente a língua portuguesa.

Após esta etapa, foi elaborada uma forma de dividir os dados entre conjuntos de treino e de validação para que possam ser utilizados no treinamento da rede LSTM. Este momento é crucial para que o modelo possa compreender e fazer as conexões entre os dados dispostos. Sendo assim, tendo a finalidade de geração de novas letras de música, dividiu-se o texto pré-processado em dois conjuntos, onde um contém as linhas quase completas, deixando de fora apenas o último token, o outro conjunto consequentemente possui o último token de cada linha. A finalidade é fazer com que o modelo aprenda sobre

a composição e criação de contexto e tente sempre fazer a predição correta da próxima palavra.

Para a manipulação destes conjuntos foi utilizada a biblioteca do Keras¹⁰ do TensorFlow¹¹ para tokenizar cada palavra da base de dados. Subsequente, têm-se a criação de entradas n-grams. Cada sequência dessas possui, no máximo, 8 tokens. É feito um padding à esquerda, completando com 0 onde não houverem tokens e garantindo que todas as sequências tenham o mesmo tamanho. As linhas são divididas entre o conjunto de treinamento e validação utilizando arrays da biblioteca NumPy¹².

3.2. Explorando a base de dados

É importante entender o conteúdo da base que está sendo utilizada, para isso é necessário tentar traduzir os dados em informações. Desta forma, em um estudo voltado a geração de novas letras de RAP, a base de dados deve conter letras deste gênero musical. Conforme descrito nesta seção, a base de dados é composta de diversas letras do grupo Racionais MC's. Porém, deve-se tentar extrair informações desta base.

Ao explorar a seção 6 têm-se alguns gráficos que ilustram algumas informações relevantes sobre a base de dados e sobre o universo de interesse. Sendo assim, na figura 1 podemos ver as 40 rimas mais frequentes nas músicas contidas na base. Isso traz algumas informações relevantes sobre a forma em que as letras são compostas, tendo em vista que são consideradas rimas as últimas palavras de cada linha do texto. Cada rima é feita a partir da harmonia sonora entre a palavra da linha atual e da linha subsequente¹³.

Outro ponto interessante de observar, referente à base de dados, é a frequência que as palavras aparecem, na figura 2 pode-se visualizar a frequência das palavras mais utilizadas nas letras de música. Este é um outro termômetro para verificar a qualidade da geração das novas letras de RAP, uma vez que é possível observar a frequência das palavras mais utilizadas nas letras que foram geradas.

Percorrendo outras características da base de dados, outro fator interessante é a frequência de substantivos e versos presentes nas composições originais. Podem ser vistos alguns exemplos na figura 3, onde têm-se os 20 substantivos mais frequentes no texto e na figura 4, onde são exibidos os 20 verbos mais frequentes nas composições.

Foi feita uma análise de sentimentos no texto para que possa ser observado o sentimento mais frequente. Essa análise foi feita utilizando o modelo pré-treinado disponibilizado em [Leme 2024].

Esta tática foi utilizada em duas abordagens diferentes, onde na figura 5 é vista a frequência de cada sentimento no texto com a análise feita verso a verso. Na figura 6 pode-se observar a frequência dos sentimentos no texto, onde a análise foi feita em blocos de 4 em 4 linhas do texto. Isso trouxe uma reflexão sobre a diferença em que o

¹⁰Site oficial da biblioteca Keras. Disponível em: <https://keras.io/>. Acesso em jun de 2024.

¹¹Site oficial da biblioteca TensorFlow. Disponível em: <https://www.tensorflow.org/?hl=pt-br>. Acesso em jun de 2024.

¹²Site oficial da biblioteca NumPy. Disponível em: <https://numpy.org/doc/stable/>. Acesso em 20 jun de 2024.

¹³Neste caso, não levou-se em consideração jogos de palavras e rimas que estejam sendo feitas entre as linhas ou em posições diferentes que não a final.

sentimento foi exibido no texto, pois quando houve mais contexto¹⁴ o sentimento teve uma diferença notória em distribuição e frequência. No caso da abordagem de verso a verso, o sentimento neutro era preponderante. A outra abordagem obteve o sentimento negativo como o sentimento mais presente no texto, o que já era esperado na hipótese inicial. Isso serve para uma próxima abordagem, onde o desenvolvimento pode ser mais complexo, contendo ajustes finos e novas abordagens, após os resultados deste protótipo, percorridos na seção 4.4.

4. Metodologia

É importante discorrer sobre como foi feito o processo de treinamento e validação do modelo. Uma vez que já foi explicitada a base de dados e o interesse final deste estudo.

4.1. Abordagens anteriores

Antes de chegar ao protótipo final deste estudo, foram feitas algumas abordagens utilizando modelos pré-treinados para geração das novas letras de RAP. A arquitetura pronta destes modelos é um diferencial para obter resultados mais robustos e refinados. A problemática desta solução, que fez com que ela fosse desencorajada, é o alto custo computacional para a execução de treinamento nas etapas de ajuste fino para o contexto desejado.

Utilizando o modelo [Souza et al. 2020], a intenção era fazer a predição de labels¹⁵ para compor as novas letras geradas. Ao fazer testes iniciais o modelo alcançava resultados satisfatórios, conseguindo fazer a predição de letras já existentes do Racionais MC's. Porém, ao tentar treiná-lo e utilizá-lo para a geração de novas letras de RAP, não obteve um grande êxito em acurácia, tendo índices de acurácia em torno de 0.30 e perda maior do que 5. Também na seção 6, é possível ver alguns resultados dos testes de predição nas figuras 10 e 11.

Na abordagem em que foi experimentado o modelo [Guillou 2020] para geração de novas letras, pela característica deste, com intuito de gerar um texto fluido e coeso. Porém, as mesmas limitações anteriormente citadas também limitaram um pouco nas tentativas de treinamento. Pois foi necessário utilizar um corpus muito extenso e, também, a inserção do corpus contendo as letras de música do Racionais MC's. Porém, foram feitos testes iniciais que não foram tão animadores quanto os do [Souza et al. 2020], pois não foram gerados textos coesos. Nem mesmo faziam sentido com o que foi enviado de prompt para a geração dos resultados. Na seção 6, é possível observar nas figuras 12, 13 alguns exemplos da geração do texto. Porém, teve uma acurácia similar ao modelo anterior, por volta de 0.40 e perda por volta de 6.

4.2. Arquitetura da rede LSTM

Após algumas tentativas com outras abordagens, uma solução que esteve presente nos estudos relacionados, que foram percorridos na seção 2, tomou-se a iniciativa de utilizar uma rede neural sequencial com arquitetura LSTM para a geração das novas letras de RAP, além de ser uma abordagem que foi implementada e vastamente utilizada durante as pesquisas para este estudo e foram supracitadas na seção 4. Busca-se com a memória

¹⁴No caso, na abordagem de blocos de 4 em 4 linhas.

¹⁵Neste estudo é considerado o label [MASK].

de curto longo prazo, um facilitador para geração de novas letras de RAP. Esta arquitetura foi escolhida com base no êxito desta para geração de texto longo.

Esta rede é composta de 4 camadas, onde a primeira delas é a camada de embedding, responsável por transformar os números inteiros (representando palavras) em vetores densos de tamanho 250, onde a saída desta camada será um tensor tridimensional. A segunda camada é LSTM e possui 200 neurônios e retorna sequências completas em vez de apenas a saída do último passo de tempo e também tem como saída um tensor tridimensional. A terceira camada é uma outra camada LSTM, porém com 100 neurônios, esta camada retorna apenas a saída do último passo de tempo, resultando em um tensor bidimensional. Por fim, a quarta camada é uma camada densa, possuindo a quantidade de neurônios relativa à quantidade de palavras únicas disponíveis no texto e usa a função de ativação softmax para calcular as probabilidades de cada palavra no vocabulário ser a próxima palavra na sequência gerada.

4.3. Processo de treinamento e validação

Optou-se por um treinamento composto de 100 épocas, com batch size de 128. Os conjuntos de treinamento e validação são compostos pela letras de música pré-processadas, conforme supracitado. Onde cada palavra, ou símbolo, é tokenizado utilizando o tokenizer do Keras (TensorFlow). Após isso são divididos em X e y, conforme descrito na seção 3 e são usados para treinamento e validação da rede neural.

O modelo foi compilado com a função de perda 'categorical_crossentropy', que pensa-se ser adequada para problemas de classificação com várias classes, visando observar como a perda se comporta no decorrer do treinamento. Esse é o motivo de utilizar a categorização no conjunto y. O otimizador escolhido foi o 'adam', que é conhecido por seu desempenho eficiente. Escolheu-se a métrica 'accuracy', para monitorar a precisão da predição das novas letras durante o treinamento e avaliação.

4.4. Resultados

Espera-se que as letras de música sejam autênticas e envolventes para cativar a atenção do público. O RAP traz consigo o símbolo da resistência contracultural e expressão de pontos de vista marginais sobre a sociedade, fazendo críticas sobre situações de opressão, descaso e preconceito, oriundas do estado ou seus representantes legais. Histórias são contadas e mensagens¹⁶ são deixadas nas entre linhas, fazendo com que os ouvintes possam refletir sobre a realidade do mundo.

O contexto na geração das novas linhas a partir de inferências no modelo carregam um leve contexto e tentou imitar a fluidez das palavras em sua predição. Contudo, na seção 6 têm-se alguns exemplos de letras geradas a partir de um texto semente nas figuras 7, 8 e 9. Optou-se por sequências de palavras que estão contidas na base de dados que alimentou o treinamento da LSTM, alguns exemplos são, "A lua cheia, clareia as ruas do capão", "Tem que ser, tem que pá", "Cotidiano na periferia e as dificuldades". A acurácia de treinamento que foi obtida no protótipo deste estudo foi de 0.9928 e a acurácia de validação foi de 0.9947, o modelo teve uma perda, no treinamento de 0.0336 e em validação de 0.0269. O arquivo está disponível em [de Sousa 2024b].

¹⁶Têm-se como conselhos, aprendizados de vida, regras implícitas em algumas situações. Palavras de aconselhamento.

5. Conclusões e trabalhos futuros

Ao gerar letras de RAP a partir de uma rede neural, busca-se coesão textual e fluidez lírica nos resultados. Porém, a limitação de hardware¹⁷ para efetuar o treinamento restringiriam o potencial máximo deste estudo, pois não foi possível utilizar um modelo pré-treinado para partir de uma arquitetura já modelada e utilizada para geração de texto longo, como a família GPT¹⁸ ou Gemini¹⁹ para auxiliar no treinamento do modelo.

Os resultados iniciais são animadores e trazem a ânsia de aplicar uma nova abordagem, talvez utilizar ambientes pagos para fazer a execução do treinamento e tentar inferir com um modelo mais fino. A inserção de modelos pré-treinados supracitados neste estudo é vista com entusiasmo para aproveitar o ambiente mais parrudo e apropriado para fazer as simulações. Sendo assim, outras técnicas de ajuste fino e regularização do modelo podem ser aplicadas, em busca de melhorar a acurácia das predições e, consequentemente, diminuir a perda durante as épocas de treinamento.

O desenvolvimento de uma rede neural para geração de novas letras de RAP é um desafio que traz a tona a correlação entre tecnologia e arte, em busca de aproximar-se do pensamento humano e suas formas de expressão. Além de ser uma homenagem a um dos maiores grupos de RAP da história, que são referências para tantas pessoas.

6. Imagens

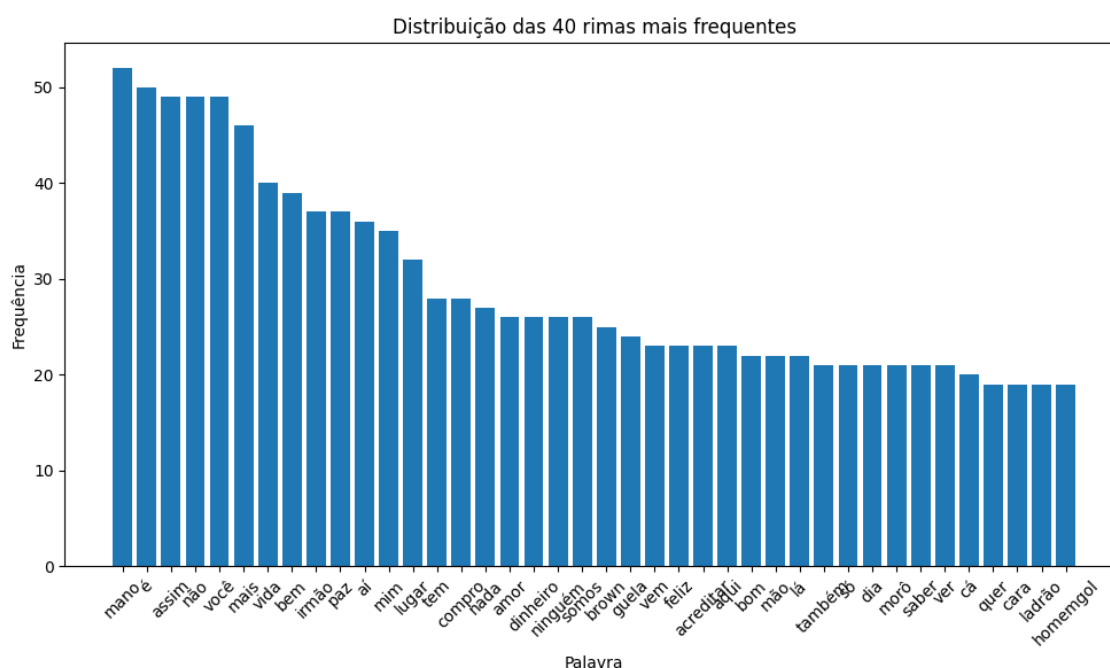


Figure 1. Distribuição das 40 rimas mais frequentes

¹⁷Foi utilizado o computador pessoal do autor e o ambiente gratuito do Google Colab.

¹⁸Documentação do GPT. Disponível em: <https://platform.openai.com/docs/models>. Acesso em 26 de jun de 2024.

¹⁹Google DeepMind. Disponível em: <https://deepmind.google/technologies/gemini/>. Acesso em 26 de jun de 2024.

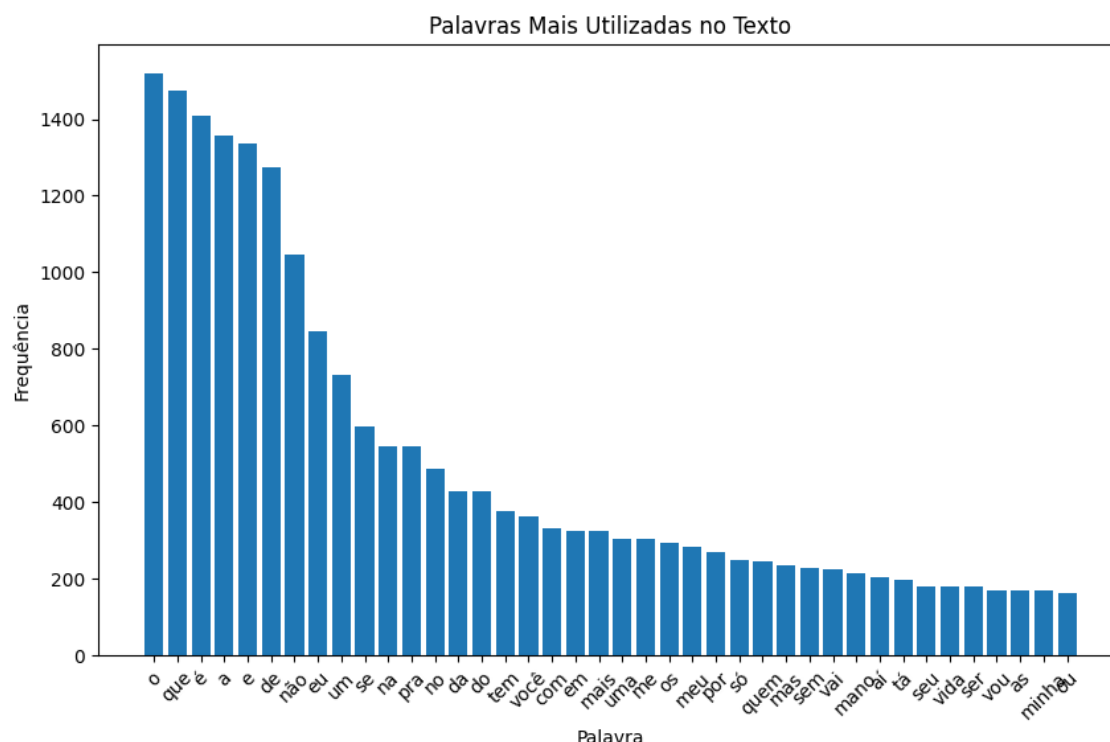


Figure 2. Distribuição das palavras mais frequentes

References

- Anand, A., Anand, A., Maiyuran, J., and Shum, M. Rap lyric generation: A phoneme-based lstm approach. [Acesso em 21 jun. 2024].
- Barrat, R. (2018). Rapping-neural-network. <https://github.com/robbiebarrat/rapping-neural-network>. [Acesso em 21 jun. 2024].
- de Sousa, J. C. (2024a). Computer intelligence. <https://raw.githubusercontent.com/JotaChina/Computer-intelligence/update%231/TP2/RacionaisLyrics.raw>. [Acesso em 21 jun. 2024].
- de Sousa, J. C. (2024b). Computer intelligence. <https://github.com/JotaChina/Computer-intelligence/tree/main/TP4>. [Acesso em 28 jun. 2024].
- Draca, N. (2017). rddt. <https://github.com/nikodraca/rddt>. [Acesso em 21 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].
- Juliani, J. d. S. (2019). Gerando letras musicais utilizado uma rede neural recorrente lstm-long short-term memory. [Acesso em 21 jun. 2024].
- Leme, L. (2024). Finbert-pt-br. <https://huggingface.co/lucas-leme/FinBERT-PT-BR>. [Acesso em: 25 jun. 2024].
- Letras.mus.br (2024). Racionais mc's. <https://www.lettras.mus.br/racionais-mcs/>. [Acesso em 21 jun. 2024].

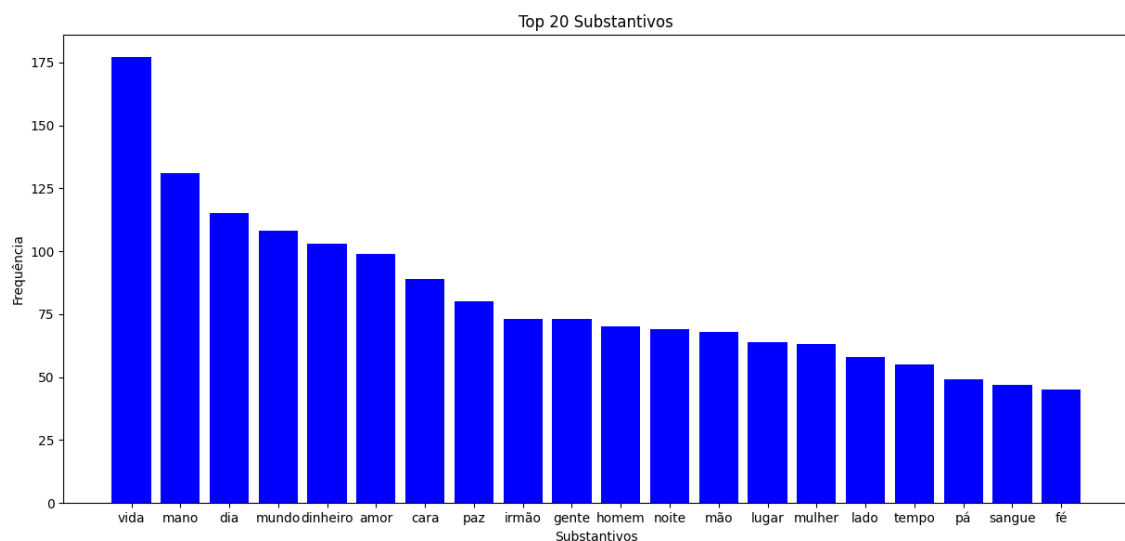


Figure 3. Distribuição dos substantivos mais frequentes

Potash, P., Romanov, A., and Rumshisky, A. (2015). Ghostwriter: Using an lstm for automatic rap lyric generation. In *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing*, pages 1919–1924. [Acesso em 21 jun. 2024].

Satija, V. (2017). Rhymenet. <https://github.com/vidursatija/rhymenet>. [Acesso em 21 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].

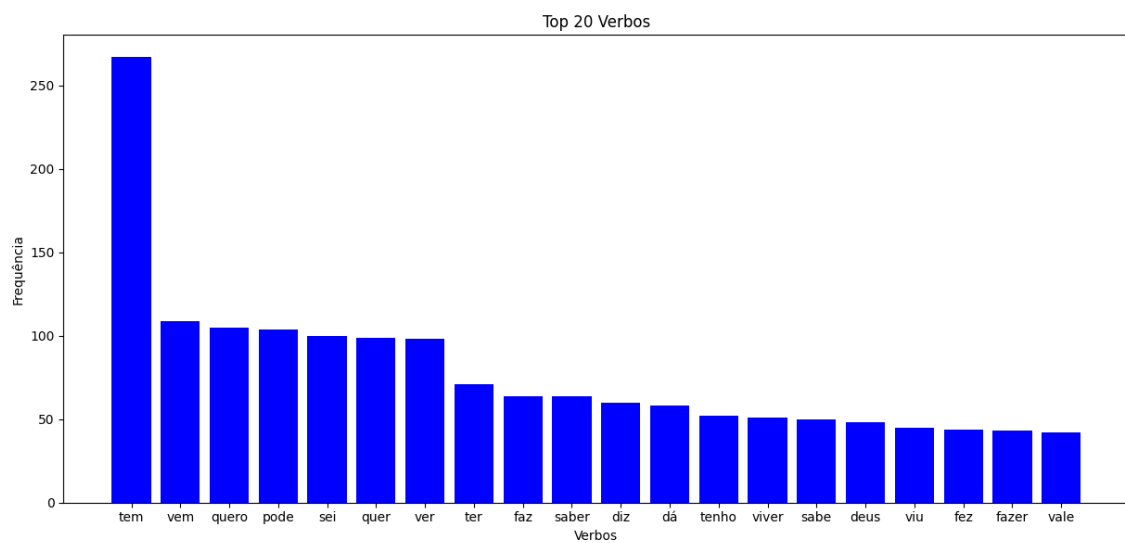


Figure 4. Distribuição dos verbos mais frequentes

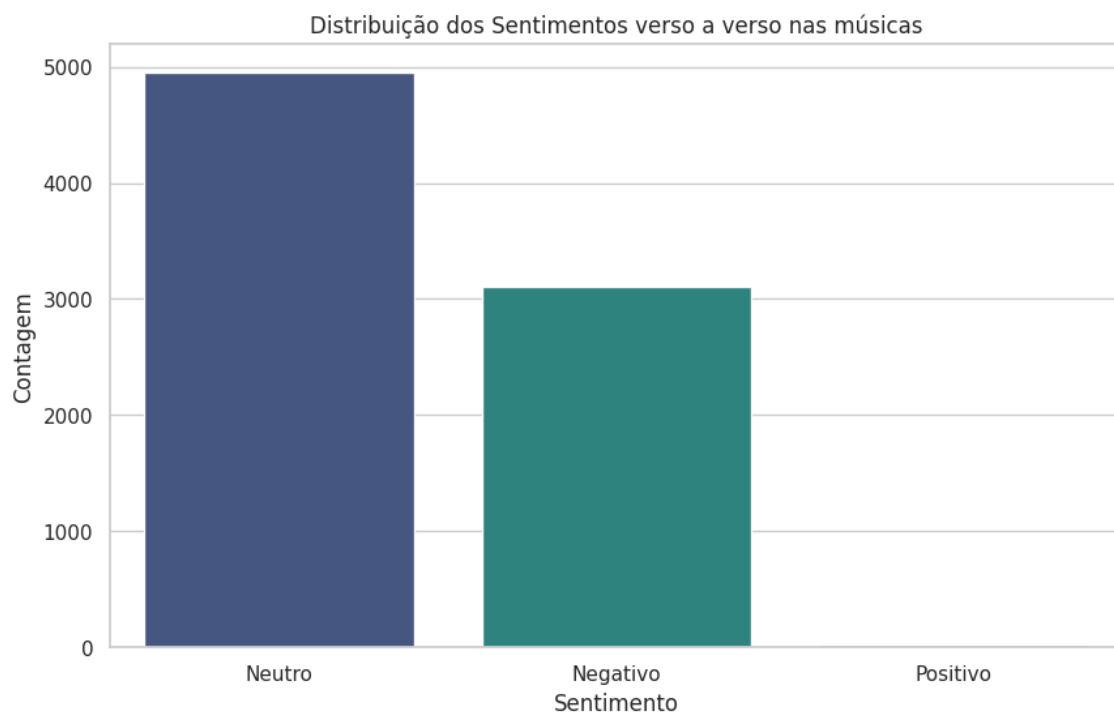


Figure 5. Distribuição dos sentimentos, verso a verso, no texto

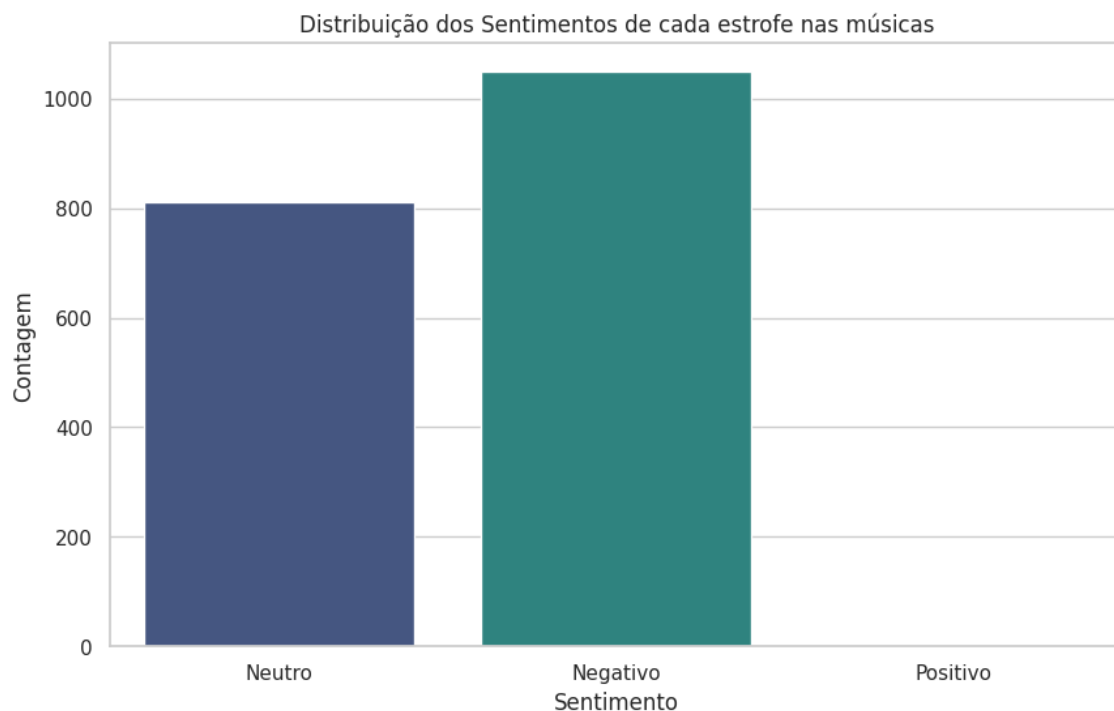


Figure 6. Distribuição dos sentimentos, bloco de 4 versos, no texto

Tem que ser, tem que pa
 é pouco dia periferia é desabado muito mano
 mano então nem já não acho que você tá
 tá tudo se tocar só isso 3 pagando de

de quem não sei responder todo celular não conhece
 conhece todos esse faz mano aqui os aliado foi
 foi grana mas ela tem que você viu nessa
 nessa que diz que só roubar em dinheiro igual

igual de sócio o dia que é mano firmeza
 firmeza vamos sempre vive fica fica nem aí é
 né a favela na área assim se lembre ao
 ao lado cada um dois puro para dinheiro pelo

pelo júri é periferia já diz meu mestre parece
 parece uma doença não é rico em cupão cupão
 cuzão na esquina as deixem em sol local e
 e veja seu som para que as conquistas metralhado

metralhado 4 branca e carente o mundo é má
 má sem dez dos trevas tradição de sina lenta
 lenta uma nossa sangue onde é muito uma vida
 vida sem si não é mulher mas ele ainda

ainda tá entre o bane bane bane bane bane
 bang pique e seis de monte por ódio hã
 hã da bala só do dia de pensar ele

Figure 7. Letras geradas a partir do prompt “Tem que ser, tem que pá”

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 policia pobres essa
essa final tio e mulher entre uma solução na

Figure 8. Letras geradas a partir do prompt “Cotidiano na periferia e as dificuldades”

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

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

brasil broca broca broca broca broca broca broca broca

Figure 9. Letras geradas a partir do prompt “A lua 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'}]
```

Figure 10. Predição de label no prompt “A [MASK] cheia, clareia as ruas do Capão”

```
[{'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'}]
```

Figure 11. Predição de label no prompt “Hoje eu sou ladrão, [MASK] 157”

```
Cotidiano na periferia e as dificuldades que se encontravam em relação a esses problemas, mas que ainda se agradavam. Em 1455, o conde Maurício de Nassau (João de Hesse-Darmstadt), filho de Maurício de Nassau,  
>> Generated text 2  
  
Cotidiano na periferia e as dificuldades da vida social em bairros periféricos do país. A luta contra a ditadura foi um grande evento de resistência política. Em 1960, um grupo de jovens estudantes universitários do Brasil foi convocado para lutar contra a ditadura  
>> Generated text 3  
  
Cotidiano na periferia e as dificuldades do trabalho na capital, que ainda se agravavam, e que lhe agradasse. O seu irmão mais velho, o senador João da Silva, era um homem de espírito de grande capacidade e capacidade intelectual,
```

Figure 12. Prompt “Cotidiano na periferia e as dificuldades”

```
>> Generated text 1  
  
Tem que ser, tem que pá como um lugar onde não exista nenhuma locação. Tem que ser o lugar onde o único mundo vivo existe, é que existe todos os seus primos como uma vida que existe  
>> Generated text 2  
  
Tem que ser, tem que pá -- em volta - para o lado, - e a gola, - e o papel. O disco foi lançado no dia 1 de
```

Figure 13. Prompt “Tem que ser, tem que pá”