

Tema: **IA nita: um estudo de inteligência artificial para geração de músicas do gênero mais influente da cultura brasileira atual - o funk**

Contexto/Motivação

A área da engenharia, em especial da engenharia da computação, é intrinsecamente interdisciplinar. Ainda assim, engenheiros têm pouco contato com movimentos culturais durante a graduação. Isso tende a criar um isolamento da profissão, que perde potencial de construir uma realidade considerando muito mais do que apenas métodos e técnicas. Conhecer todo o contexto social envolvido é essencial para o exercício responsável da profissão.

Objetivo

O objetivo do projeto foi estudar técnicas e modelos de aprendizagem profunda (“Deep Learning”) que representassem o estado da arte para geração de música a partir de áudio cru (.mp3, .wav, etc.) e aplicá-las ao gênero funk.

Metodologia

Foi utilizada uma metodologia de pesquisa organizada de forma iterativa. Compõe-se uma base de dados e realiza-se processo de treinamento/validação de uma arquitetura. Com esses resultados, geram-se novas ideias para o aprimoramento das amostras geradas, cria-se uma nova base de dados e treina-se uma nova rede com melhores super-parâmetros.

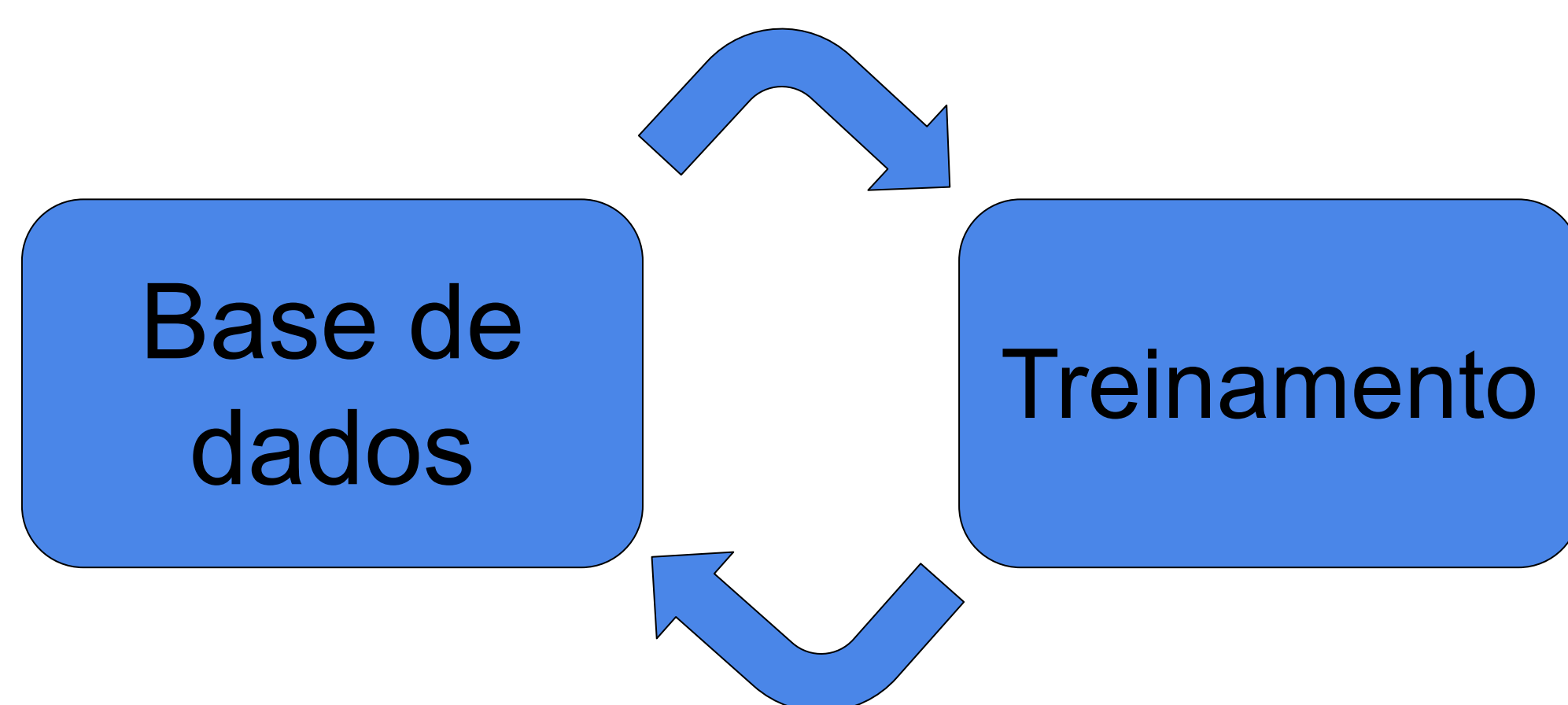


Figura 1: Representação da metodologia empregada

O projeto se desenvolveu em um ambiente virtual chamado Google Colab e utilizou conceitos como redução de dimensionalidade, aprendizagem não supervisionada e supervisionada.

Resultados

Foram realizados 3 ciclos de desenvolvimento ao todo e uma prova de conceito utilizando-se uma rede já treinada.

Os resultados das gerações realizadas pela arquitetura SaShiMi, um dos modelos que representa o estado-da-arte na área de geração de áudio, foram decepcionantes. Pouca melhoria foi encontrada entre cada ciclo, ainda que boas reflexões tenham sido produzidas.

A rede jukebox, treinada com muito mais recursos computacionais, se mostrou capaz de compreender parte da musicalidade do funk, e gerou extratos interessantes.

Tal experimento revela a grande necessidade de recursos para se trabalhar com formatos ricos em informação e utilizar técnicas de Deep Learning. A tarefa de gerar áudio a partir de um gênero musical proveniente das periferias do país se mostra de difícil execução. Analogamente, o gênero é marginalizado nas escolhas de material de treinamento, ainda que possivelmente seu emprego gerasse descobertas devido a sua musicalidade peculiar.

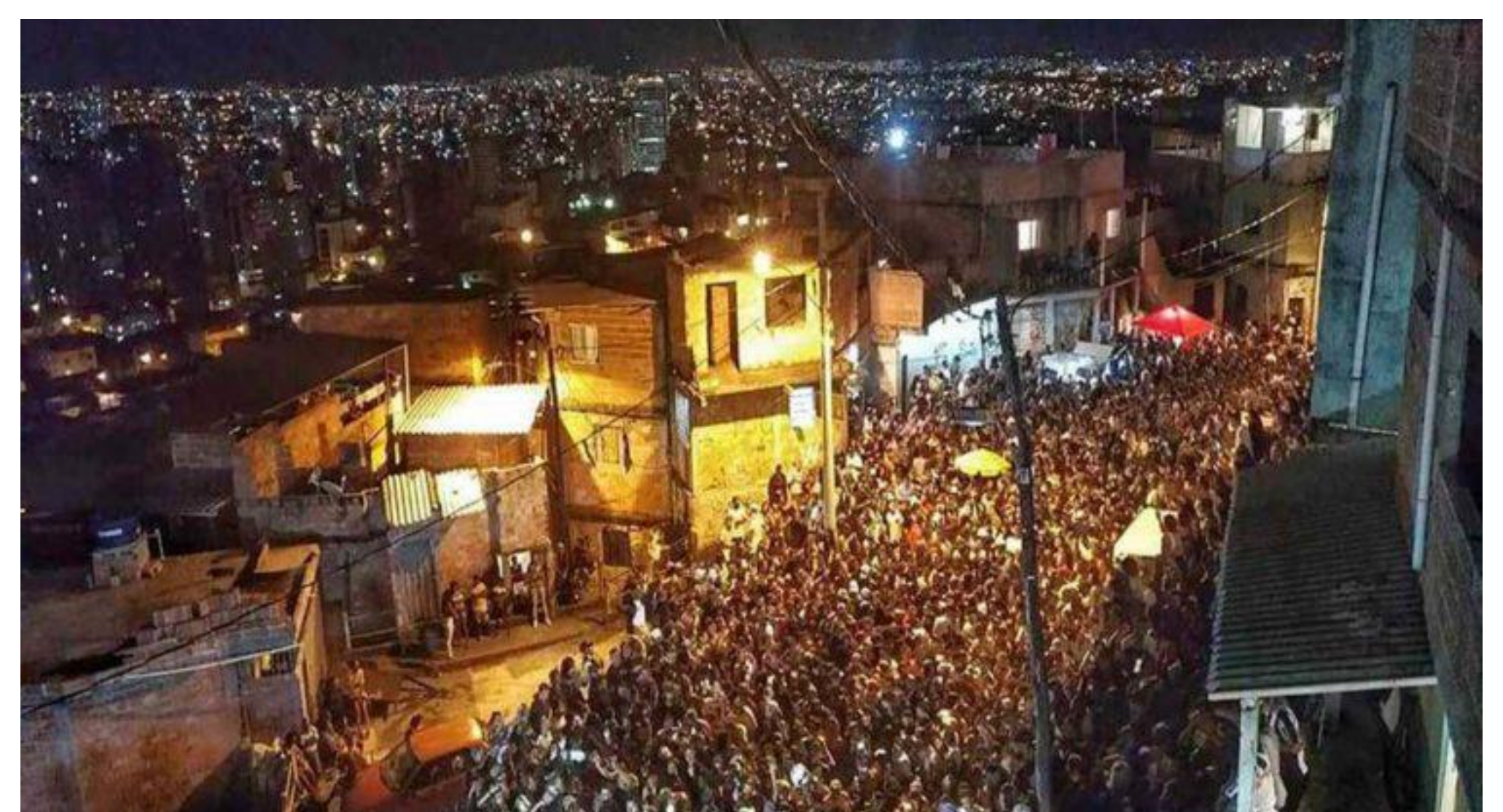


Figura 2: Foto aérea de um baile funk.

Aplicação

As amostras geradas, assim como reflexões a respeito do projeto como um todo, podem ser acessados no website:

<https://huggingface.co/spaces/renatottin/TCCapp>