Universidade Federal do Rio Grande do Norte - CCET/DIMAp Aprendizado Profundo – DIM0446 – 2024.2

**Título do trabalho** Warlike Richard da Silva Soares

Natal, 29/01/2025.

**Resumo:** este documento tem como objetivo descrever um modelo que utiliza redes recorrentes para transformar um áudio de cantarolado em uma música tocada em piano. Para isso foram gravados os cantarolados a partir da base de dados Maestro v3 de arquivos MIDI.

# Introdução

Neste trabalho eu implementei uma IA que transforma áudios .wav de cantarolados em um outro áudio .wav com um som de piano, que visa corresponder ao cantarolado original. Para isso, o sistema transforma o áudio do cantarolado em um vetor com os embeddings extraídos do áudio, usando o modelo wav2vec. Esses embeddings passam então pelo modelo que utiliza RNN para ser treinado, gerando então um espectrograma que logo após será usado em outro modelo para gerar o áudio. Sendo assim, o sistema criado utiliza duas ferramentas para atingir o objetivo, já que o modelo puro apenas transforma os embeddings em espectrogramas.

# Fundamentação teórica

Classificadores são...

Existem várias técnicas de classificação, dentre elas podemos mencionar...

# O modelo implementado

O modelo consiste em... (descrever com detalhes o tipo de modelo usado: redes neurais, HMM, SVM, quantas camadas).

# Resultados

Descrever o que foi feito na base de dados em termos de pré-processamento. Mostrar como a base de dados foi particionada (treino, validação e teste).

Falar sobre o ajuste de hiperparâmetros (caso tenha sido feito). Mostrar as curvas de aprendizado (no caso de redes neurais).

Mostrar o resultado final: ao final o modelo obteve uma acurácia de XX%. A Figura XX mostra a tabela de confusão...

O modelo em questão foi treinado em uma máquina XX usando a(s) ferramenta(s) YYY... O treinamento durou ZZZ...

# Conclusão

Implementamos um modelo classificador de XXX baseado em YYYY. A base de dados utilizada foi ZZZ que foi particionada da seguinte forma: AA amostras para treinamento, BB amostras para validação e CC amostras para teste.

O modelo alcançou uma acurácia de UUU%.

Aqui estão algumas sugestões para melhoria: XXX, YYY, ZZZ

# Referências bibliográficas

Artigos, revistas e blogs.