

Submetido 31/05/2024. Aprovado 12/04/2025

Avaliação: revisão duplo-anônimo

# Redes Perceptron Multicamadas para modelar efeitos de distorção em sinais de guitarra elétrica

MULTILAYER PERCEPTRON NETWORKS TO MODEL DISTORTION EFFECTS IN ELECTRIC GUITAR SIGNALS

REDES DE PERCEPTRONES MULTICAPA PARA MODELAR EFECTOS DE DISTORSIÓN EN SEÑALES DE GUITARRA ELÉCTRICA

**Alisson Assis Carvalho**

Universidade Federal de Goiás (UFG)  
alsnac@ufg.br

**Murilo Guimarães Correia**

Universidade Federal de Goiás (UFG)  
murilogcorreia@discente.ufg.br

**Ricardo Augusto Pereira Franco**

Universidade Federal de Goiás (UFG)  
ricardo@inf.ufg.br

**Samuel Carvalho de Almeida**

Universidade Federal de Goiás (UFG)  
carvalho.carvalho@discente.ufg.br

## Resumo

Este trabalho investiga a modelagem de efeitos de distorção em sinais de guitarra elétrica utilizando redes perceptron multicamadas (MLPs). Na proposta, a MLP é responsável por transformar o sinal elétrico original da guitarra em um sinal distorcido. Para o treinamento da rede, foram gerados/coletados dados sintéticos de efeito de distorção a partir de sinais elétricos reais aplicados em um simulador SPICE de circuitos eletrônicos. Os resultados do estudo são apresentados em termos de Erro Quadrático Médio (EQM) e testes estatísticos de aderência de Kolmogorov-Smirnov. As análises demonstram que a proposta apresentada para modelar o sistema dinâmico de um pedal elétrico de distorção com o uso da MLP tem bom desempenho, representando os sinais distorcidos com fidelidade.

**Palavras-chave:** modelagem de distorção de áudio; redes MLP; processamento de sinal; sistemas dinâmicos.

## Abstract

This paper investigates the modeling of distortion effects in electric guitar signals using multilayer perceptron (MLP) networks. In the proposed approach, the MLP is employed to transform the original guitar signal into a distorted output. For network training, synthetic distortion data were generated from real electrical signals processed through a SPICE electronic circuit simulator. The results are evaluated using Mean Square Error (MSE) and Kolmogorov–Smirnov statistical adherence tests. The findings

indicate that the proposed MLP-based model effectively captures the dynamics of an electric guitar distortion pedal, providing a faithful representation of the distorted signals.

**Keywords:** audio distortion modeling; MLP neural networks; guitar signal processing; dynamic systems.

### Resumen

Este trabajo presenta una investigación del modelado de efectos de distorsión en señales de guitarra eléctrica mediante el uso de redes de perceptrones multicapa (MLP). Para el entrenamiento de la red, se generaron/recogieron datos sintéticos del efecto de distorsión a partir de señales eléctricas reales, estas aplicadas a un simulador de circuito electrónico SPICE. Los resultados del estudio se presentan en términos de error cuadrático medio (MSE) y pruebas de adherencia estadística de Kolmogorov-Smirnov. Los análisis demostraron que la propuesta presentada para modelar el sistema dinámico de un pedal de distorsión eléctrico con el uso de MLP tiene buen desempeño, representando fielmente las señales distorsionadas.

**Palabras clave:** modelado de distorsión de áudio; redes neuronales MLP; procesamiento de señales de guitarra; sistemas dinámicos.

## Introdução

As guitarras elétricas e os pedais são elementos fundamentais no contexto do rock como gênero musical. Os pedais desempenham um papel crucial na manipulação e na criação de uma ampla gama de timbres, variando de tons agressivos e excêntricos a nuances mais suaves e delicadas. Essa diversidade sonora não apenas permite que os músicos desenvolvam sua própria identidade musical, mas também contribui significativamente para a identidade sonora de uma banda (Reiss; Mcpherson, 2014).

Um modelo icônico de pedal de distorção é o “MXR Distortion+”, conhecido também como “Distortion Plus” ou “D+”, classificado como um Fuzz de distorção mais suave. Lançado pela empresa MXR entre 1978 e 1979, utiliza amplificador operacional 741, potenciômetros de volume e ganho, além de diodos germânicos para recortar tensões acima de um limiar definido, resultando no efeito de distorção característico (Self, 2023).

Compreender o funcionamento dos circuitos eletrônicos é essencial, dada sua natureza não linear. Essa não linearidade origina-se do emprego de componentes não lineares em sua construção, os quais são indispensáveis à formação da sonoridade característica proporcionada pela distorção. Esse desafio se intensifica pela necessidade das tecnologias de distorção digital de reproduzir com fidelidade os sons produzidos pelos circuitos analógicos, que estão intrinsecamente atrelados à sonoridade histórica de diversas bandas e artistas.

Uma alternativa prática e motivadora para lidar com essa complexidade reside na utilização de inteligências artificiais para modelar sistemas não lineares. A ideia central consiste em fornecer à inteligência artificial sinais de entrada de guitarras sem distorção e aplicá-los a um processo de distorção. Empregando técnicas de aprendizado de máquina, a inteligência artificial é capaz de gerar um modelo do efeito de distorção aplicado ao sinal de guitarra elétrica. Entre essas técnicas destaca-se o emprego de redes neurais como solução promissora, conforme proposto em Purwins *et al.* (2019).

As redes neurais, inspiradas no funcionamento dos neurônios do cérebro humano, configuram-se como uma ferramenta poderosa para o processamento de dados

e a execução de tarefas complexas (Brunetto; Schmidt; Dalla'rosa, 2023). A unidade fundamental de uma rede neural, o neurônio artificial, estabelece conexões com outros neurônios por meio de sinapses artificiais, análogas aos dendritos biológicos. Os valores recebidos por cada neurônio são ponderados e combinados por meio de uma função matemática, gerando a saída do neurônio. Esses neurônios se organizam em camadas interconectadas, em que a informação é processada e refinada sucessivamente, culminando na saída final da rede neural (Faceli *et al.*, 2011).

Este trabalho tem como objetivo investigar a capacidade de uma rede MLP (*Multilayer Perceptron*) em modelar efeitos de distorção em sinais de guitarra elétrica. Além disso, busca-se avaliar a fidelidade das modelagens geradas pela MLP em comparação com os sinais de distorção originais. Os resultados obtidos são apresentados e discutidos ao longo do texto, evidenciando o potencial das redes MLP na reprodução de efeitos de áudio complexos.

## Referencial teórico

Na música, a distorção é um efeito frequentemente utilizado com a finalidade de alterar a forma do áudio e gerar um som saturado e expressivo na guitarra elétrica. Modelar e controlar essa distorção com precisão é fundamental para músicos e engenheiros de áudio. Redes Neurais Artificiais (RNAs), especialmente Perceptrons Multicamadas (MLPs), apresentam-se como uma abordagem promissora para essa modelagem. Este referencial teórico explora os fundamentos da distorção de áudio, a aplicação de RNAs no processamento de áudio e a relevância das MLPs nesse contexto.

### Distorção de áudio

A distorção de áudio pode ser classificada em dois tipos principais: harmônica e não harmônica. A distorção harmônica ocorre quando os harmônicos de um tom fundamental são produzidos de maneira controlada, resultando em um som mais rico e musical. Em contraste, a distorção não harmônica produz componentes de frequência não relacionados ao som original, muitas vezes resultando em uma cacofonia áspera (Reiss; Mcpherson, 2014).

### Pedal e simulação

- Elementos para obtenção dos dados

A escolha arbitrária do pedal não afeta no resultado da pesquisa, bem como a forma que o sinal de saída do pedal é coletado. Como o sinal de entrada e o sinal de saída do pedal são fornecidos à rede neural, o que importa é se a rede neural consegue uma aproximação da função de transferência, mas qual é a função de transferência em si não importa para este escopo.

Dessa forma, também se justifica a utilização de um software de simulação de circuitos para simular o pedal, pois não cabe a esta pesquisa avaliar a qualidade de simulação do software, mas sim a capacidade da rede neural de relacionar dois sinais. Esse fato ressalta a aplicabilidade do processo em todos os sistemas dinâmicos não lineares, como os pedais de distorção. Assim, um pedal de distorção genérico foi escolhido por sua relativa simplicidade e ampla utilização no gênero do rock.

Semelhantemente, o software para simulação, o LTSpice (“Analog Devices”, 2023), foi escolhido pela simplicidade e pelo fato de ser gratuito (LTSpice, 2023).

- Elementos para obtenção dos dados

O pedal está organizado em 4 partes que realizam funções individuais no processamento do sinal. A primeira dessas partes é o estágio de entrada de sinal (do inglês, *input signal stage*). O sinal escolhido para a simulação estava com uma tensão alta, mais alta do que costuma ser fornecido de entrada por uma guitarra. Por esse motivo, foi colocado uma fonte dependente com ganho de 0.2, valor esse que trouxe o sinal de entrada para um valor mais próximo da realidade.

Na montagem do circuito analógico, o sinal da guitarra deve ser aplicado no nó sinalizado pela palavra *input*. A segunda parte do pedal é o estágio de alimentação de energia (do inglês, *power supply stage*). O uso dos altos valores de resistência contribui para uma alta impedância de entrada do circuito, além de uma alta impedância para o terra do curto-circuito virtual. O capacitor tem o papel de eliminar a variação residual conhecida como ondulação (do inglês, *ripple*), que é um sinal AC indesejado. Na Figura 1, pode-se visualizar o diagrama elétrico do pedal de efeito de distorção.

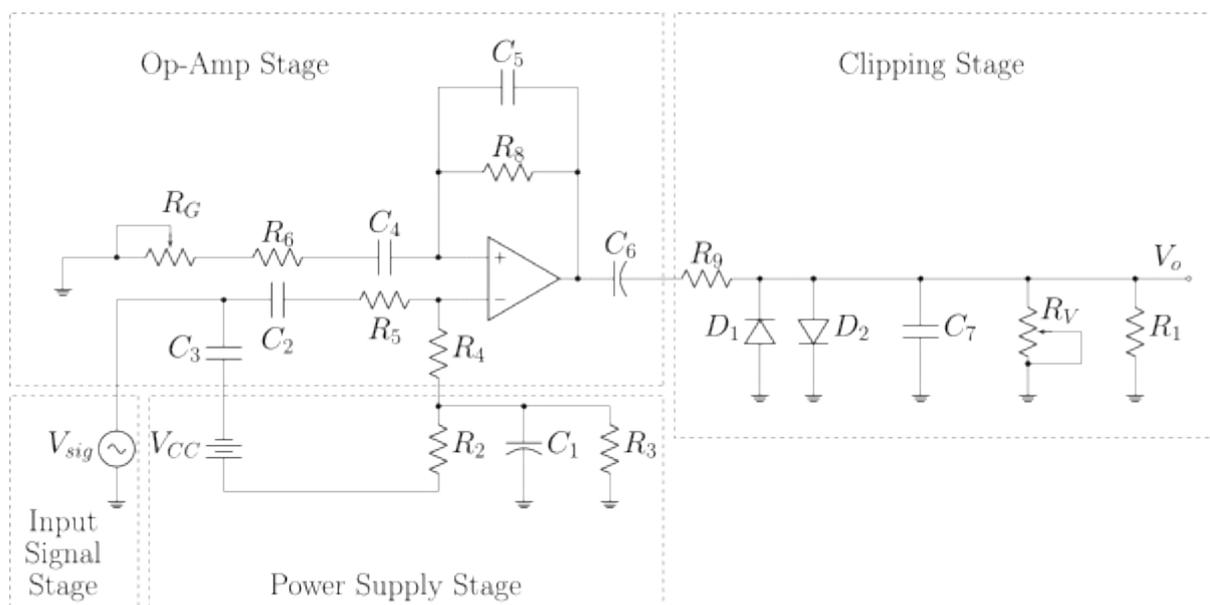


Figura 1 – Diagrama elétrico do pedal

Fonte: elaborado pelo(as) autores(as).

Ademais, a terceira parte do pedal é o estágio de amplificação do Amp Op (do inglês, *Op Amp Amplifier Stage*, ou somente *Op Amp Stage* como referido no esquemático). O amplificador se encontra na configuração não inversora, o que gera um sinal de saída como relação dos resistores  $R_8$ ,  $R_6$  e  $R_G$ , sendo este último um potenciômetro que permite ao usuário controlar o ganho. Além disso, nesse estágio, há o capacitor  $C_3$ , cuja função é evitar ruído de rádio frequência, bem como ajudar em descargas eletrostáticas e oscilações. O capacitor  $C_6$  tem a função de remover a tensão DC para o próximo estágio. Os outros capacitores contribuem para a resposta em frequência com pico em torno de 1,5 KHz, característica comum em outros pedais que ajuda a guitarra a ter destaque em relação a outros instrumentos na música ao ter maior ganho na faixa de frequência audível para os seres humanos.

O último estágio é o de recorte (do inglês, *Clipping Stage*). Nessa etapa, o resistor R9 limita a corrente que chega aos diodos, protegendo-os. Os diodos possuem uma tensão de saturação e, pela disposição deles, saturam com tensão aproximadamente menor que -0,6V e valores maiores que 0,7V. Com isso, o sinal é cortado/limitado além desses limites, dando a sonoridade característica da distorção. Quanto mais abrupto o corte, mais distorcido e com excessos de harmônicos o som fica (Reiss *et al.*, 2014). Finalmente, há um potenciômetro para a regulação do volume, bem como a impedância de saída do pedal – esta última não está presente na construção do pedal, mas é importante para a simulação por representar uma carga para ser realizada a máxima transferência de potência entre o pedal e a carga.

## Redes neurais artificiais em processamento de áudio

Redes neurais perceptron multicamadas (MLPs) são um tipo de RNA comumente usado em problemas de aprendizagem supervisionada. Eles consistem em camadas de neurônios (perceptrons) organizadas em uma estrutura de camadas, incluindo uma camada de entrada, uma ou mais camadas ocultas e uma camada de saída. A MLP é conhecida por sua capacidade de modelar funções não lineares complexas (Zhang *et al.*, 2023).

A capacidade de modelar funções não lineares complexas da MLP é explicada ao entender que uma rede com uma camada intermediária pode implementar qualquer função contínua. A utilização de duas camadas intermediárias permite a aproximação de qualquer função (Faceli *et al.*, 2011). Sendo a função que se deseja modelar não linear, usar o método com múltiplas camadas permite uma aproximação dessa função, por isso a MLP foi escolhida como método de aprendizado de máquina aplicado.

## Métodos de avaliação dos resultados

A utilização das técnicas de teste de erro quadrático médio (EQM) e Kolmogorov-Smirnov (KS) é muito importante na avaliação de resultados de projetos de processamento de áudio, como a modelagem dos efeitos de distorção de um sinal de guitarra elétrica utilizando uma rede MLP.

- Erro Quadrático Médio (EQM)

O erro quadrático médio é uma métrica comumente usada nos campos de processamento de sinais e aprendizado de máquina para quantificar a diferença entre um sinal estimado e um sinal de referência (Bishop, 2006). Este projeto usa EQM para medir a fidelidade de uma rede MLP a um sinal de áudio de uma guitarra elétrica com som distorcido como referência. Quanto menor o valor do EQM, mais próxima a saída MLP está do sinal de referência, indicando melhor capacidade do modelo em reproduzir o efeito de distorção desejado.

- Teste de Kolmogorov-Smirnov (KS)

O teste Kolmogorov-Smirnov é um método estatístico usado para comparar duas distribuições de probabilidade e avaliar se elas vêm da mesma população (Stephens, 1974). Neste projeto, o KS foi aplicado para verificar se a saída MLP segue a mesma distribuição estatística do sinal distorcido da guitarra elétrica. Um valor alto para a estatística KS ou um valor menor que o valor significativo de 0.05 indica que as duas

distribuições são diferentes, sugerindo que a saída da MLP não reproduz adequadamente o efeito de viés desejado.

O Teste de Kolmogorov-Smirnov (KS) é usado para verificar a semelhança entre duas distribuições de dados. A estatística de teste KS é calculada como:

onde  $F_n(x)$  é a função de distribuição acumulada da primeira distribuição,  $G(x)$  é a função de distribuição acumulada da segunda distribuição e  $\sup_x$  denota o supremo sobre todos os valores possíveis de  $x$ . O valor-associado ao KS é usado para determinar a significância estatística da diferença entre as duas distribuições.

- Função de Distribuição Acumulada (FDA)

Além das métricas EQM e KS, a avaliação dos resultados incorpora a análise visual das Funções de Distribuição Acumulada (FDA), derivadas das saídas da MLP e dos sinais de referência. As Figuras 5 e 6 ilustram as FDA para os ganhos de 10k e 30k, respectivamente. Esses gráficos proporcionam uma representação visual das distribuições acumuladas das amplitudes, permitindo uma comparação intuitiva entre a saída da MLP e os sinais de referência.

A função de distribuição acumulada (FDA) é uma ferramenta fundamental em estatística e probabilidade. Ela é definida como a integral da função de densidade de probabilidade (PDF). No contexto deste estudo, a FDA é aplicada às amplitudes dos sinais, proporcionando uma visão acumulativa das probabilidades associadas a diferentes valores de amplitude. A sobreposição próxima entre as curvas FDA da saída da MLP e dos sinais de referência indica consistência estatística, validando a capacidade da MLP em reproduzir os efeitos de distorção esperados nos sinais de guitarra elétrica (Ross, 2020; Wasserman, 2004; Zhang *et al.*, 2023).

### Trabalhos relacionados

A área de processamento de áudio e inteligência artificial conta com diversas contribuições notáveis, algumas das quais são apresentadas no Quadro 1. Esses trabalhos anteriores exploram o uso de redes neurais em processamento de áudio, abordando técnicas e aplicações em síntese musical.

Título	Referência	Correlação	Diferencial
"Deep Learning for Audio Signal Processing"	(Purwins <i>et al.</i> , 2019)	Explora o uso de redes neurais em processamento de áudio	Destaca-se pela revisão abrangente de técnicas e aplicações em processamento de áudio
"Neural Audio Synthesis of Musical Notes with WaveNet Autoencoders"	(Engel <i>et al.</i> , 2017)	Modelagem de síntese neural de áudio, incluindo efeitos	Aplica autoencoders WaveNet na síntese de áudio musical, destacando-se pela aplicação específica em notas musicais
"Audio Super Resolution using Neural Networks"	(Kuleshov <i>et al.</i> , 2017)	Apresenta uma abordagem de super-resolução de áudio utilizando redes neurais profundas	Aplicação de <i>deep learning</i> para melhorar a qualidade de áudio de baixa resolução

Quadro 1 - Trabalhos relacionados  
 Fonte: elaborado pelos(as) autores(as).

## Diferencial deste trabalho

Este trabalho diferencia-se ao concentrar-se na análise específica dos efeitos de distorção em sinais de guitarra elétrica. Enquanto estudos anteriores podem ter uma visão mais abrangente do processamento de áudio, esta pesquisa aprofunda-se em uma aplicação particular, buscando compreender e reproduzir com fidelidade os efeitos de distorção desejados.

### CONTRIBUIÇÕES ESPECÍFICAS

A análise detalhada dos efeitos de distorção em sinais de guitarra elétrica, usando uma abordagem baseada em redes neurais, mais precisamente, uma Multilayer Perceptron (MLP), é uma contribuição única deste trabalho. Enquanto estudos anteriores podem oferecer uma visão mais ampla, nossa pesquisa proporciona uma compreensão aprofundada e aplicada em um contexto musical específico.

### POTENCIAIS IMPACTOS

As implicações deste trabalho não se limitam à esfera acadêmica, apresentando também potenciais impactos práticos. A capacidade de modelar de forma eficaz os efeitos de distorção em sinais de guitarra elétrica pode ter aplicações diretas na indústria musical, contribuindo para o desenvolvimento de processadores de efeitos mais avançados e realistas. Além disso, essa pesquisa pode servir como base para explorações mais aprofundadas em áreas relacionadas, como síntese de áudio, processamento de sinais musicais e até mesmo em campos mais amplos, como reconhecimento de padrões sonoros.

### FUTURAS DIREÇÕES DE PESQUISA

O enfoque específico deste trabalho abre portas para futuras pesquisas que podem explorar diferentes instrumentos musicais, ampliando a aplicabilidade da abordagem proposta. Além disso, a incorporação de técnicas mais avançadas de aprendizado de máquina, como redes neurais recorrentes ou redes generativas, pode enriquecer ainda mais a modelagem de efeitos de distorção em contextos musicais diversos.

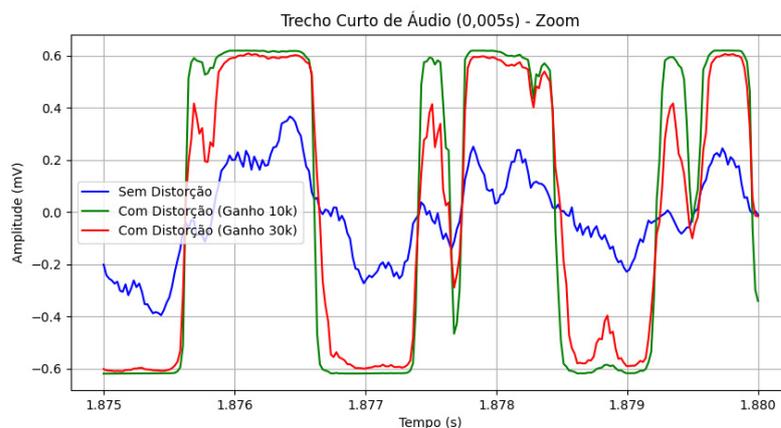
## Metodologia

### Obtenção e preparação de dados

Na fase inicial de obtenção e preparação dos dados, foi adotada uma abordagem meticulosa para garantir a qualidade das informações coletadas. Foi empregado um software de simulação em conjunto com um circuito eletrônico modelado para realizar simulações de áudio. O sinal de uma guitarra genérica, sem efeitos, foi selecionado como entrada principal. A plataforma de simulação, munida da funcionalidade *waveform*, possibilitou a incorporação desse sinal como entrada no circuito simulado. Após a definição dos parâmetros apropriados, o software gerou o sinal de saída em formato de áudio, fornecendo, assim, os dados cruciais para as próximas etapas da pesquisa.

Os dados passaram por uma fase de preparação cuidadosa, visando assegurar a qualidade das entradas fornecidas à MLP. Os áudios da guitarra elétrica, tanto com quanto sem distorção, foram registrados a uma taxa de amostragem de 48kHz, com duração de 30 segundos, totalizando 1.440.000 amostras e capturando nuances sutis. A distorção foi introduzida por meio da aplicação de ganhos específicos, criando pares de áudios correspondentes: um com distorção gerada por um resistor de 10 k $\Omega$  e outro com distorção devido a um resistor de 30k $\Omega$ . Esse procedimento possibilitou a comparação direta dos efeitos da distorção na saída da MLP em relação ao áudio não distorcido.

A Figura 2 ilustra um trecho breve (0,005 segundos) comparando os áudios com distorção, considerando ganhos dos resistores de 10k $\Omega$  e 30k $\Omega$ , com o áudio sem distorção.

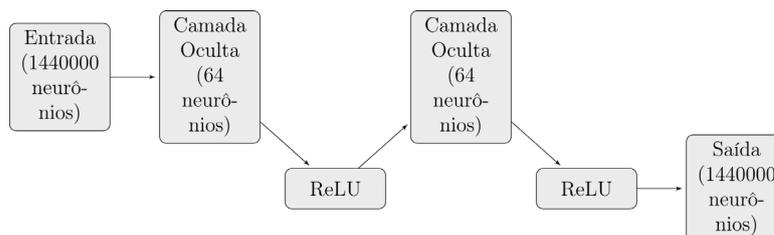


**Figura 2 – Comparação entre os áudios com distorção (ganho de 10k $\Omega$  e 30k $\Omega$ ) e o áudio sem distorção com tempo de 0,005 segundos**  
Fonte: elaborado pelo(as) autores(as).

## Arquitetura da Perceptron Multicamadas (MLP)

A arquitetura da MLP desempenha um papel fundamental na modelagem de efeitos de distorção em sinais de guitarra elétrica. A perceptron multicamadas foi projetada para aprender representações complexas de sinais de áudio, permitindo a reprodução de efeitos de distorção realistas. Esta seção detalha a arquitetura MLP usada neste estudo.

A MLP utilizada neste projeto consiste em três camadas principais: uma camada de entrada, duas camadas ocultas e uma camada de saída. A camada de entrada é responsável por receber o sinal de áudio bruto, e as camadas ocultas aprendem uma representação intermediária do sinal. Já a camada de saída produz uma representação estimada do sinal de áudio distorcido. A arquitetura da MLP é representada na Figura 3.



**Figura 3 – Diagrama da arquitetura da MLP.**  
Fonte: elaborado pelos(as) autores(as).

Pode-se observar as seguintes informações sobre a arquitetura dessa MLP.

- **Camada de entrada:** a camada de entrada possui 1.440.000 neurônios, correspondendo ao número de amostras dos áudios em cada entrada. Isso ocorre devido à frequência de amostragem de 48kHz e ao uso dos áudios de entrada de 30 segundos. A frequência de amostragem determina o número de amostras por segundo, resultando em 48.000 amostras/segundos vezes 30 segundos, o que é igual a 1.440.000 amostras no total por áudio com cada neurônio recebendo uma amostra de cada áudio, ou seja, duas amostras por neurônio.
- **Camadas ocultas:** a MLP possui duas camadas ocultas, cada uma com 64 neurônios. As camadas ocultas são responsáveis por aprender representações intermediárias dos sinais de áudio. A função de ativação da unidade de retificação unificada, traduzida de *rectified linear unit* (ReLU), é aplicada após cada camada oculta para introduzir não linearidades na rede.
- **Camada de saída:** a camada de saída possui 1.440.000 neurônios, correspondendo ao número de amostras do áudio na saída. Ela produz a representação estimada do sinal de áudio distorcido.

A função de ativação ReLU desempenha um papel fundamental nas camadas ocultas da MLP. O comportamento da função ReLU consiste em mapear qualquer valor negativo para zero, enquanto mantém os valores positivos inalterados. Esse processo introduz não linearidades nas camadas ocultas da MLP, possibilitando a aprendizagem de relações não lineares complexas nos dados de entrada (Faceli *et al.*, 2011). A presença da função ReLU é crucial para capacitar a rede a discernir padrões e características mais intrincadas nos sinais de áudio.

A derivada da ReLU apresenta uma característica peculiar: o gradiente é zero para valores negativos e igual a um para valores não negativos, sendo crucial durante o treinamento da MLP. A propriedade de permitir a passagem irrestrita de gradientes positivos (durante a retropropagação do erro) é fundamental para evitar o problema do desaparecimento do gradiente, contribuindo para uma convergência mais eficaz durante o processo de otimização. Em contraste com funções de ativação que introduzem saturação, a ReLU facilita o treinamento de redes mais profundas, tornando-a uma escolha prevalente em arquiteturas de MLP (Faceli *et al.*, 2011).

## Treinamento MLP

O treinamento da MLP (Perceptron Multicamadas) desempenhou um papel crucial neste projeto, possibilitando que a rede neural modele de maneira eficaz os efeitos da distorção em sinais de guitarra elétrica. Detalhamos a seguir os principais aspectos desse processo.

O otimizador Adam (Kingma; Ba, 2017; Zhang *et al.*, 2023) foi escolhido para ajustar os parâmetros da MLP durante o treinamento. O algoritmo do otimizador Adam é apresentado a seguir.

**Entrada:**

- $\theta$  (parâmetros da rede neural)
- $\alpha$  (taxa de aprendizado)
- $\beta_1, \beta_2$  (hiperparâmetros de momento)
- $\varepsilon$  (termo de estabilização numérica)
- $f(\theta)$  (função de perda)

**Inicialização:**

- $m \leftarrow 0$  (vetor inicial de primeiras ordens do momento)
- $v \leftarrow 0$  (vetor inicial de segundas ordens do momento)
- $t \leftarrow 0$  (contador de iterações)

**Passo a passo:****1. Para cada iteração:**

1.  $t \leftarrow t + 1$
2.  $g_t \leftarrow \nabla_{\theta} f(\theta_t)$  (gradiente da função de perda)
3.  $m_t \leftarrow \beta_1 * m_{t-1} + (1 - \beta_1) * g_t$  (primeiro momento do gradiente)
4.  $v_t \leftarrow \beta_2 * v_{t-1} + (1 - \beta_2) * g_t^2$  (segundo momento do gradiente)
5.  $\hat{m}_t \leftarrow m_t / (1 - \beta_1^t)$  (correção do viés no primeiro momento)
6.  $\hat{v}_t \leftarrow v_t / (1 - \beta_2^t)$  (correção do viés no segundo momento)
7.  $\theta_t \leftarrow \theta_{t-1} - \alpha * \hat{m}_t / (\sqrt{\hat{v}_t} + \varepsilon)$  (atualização dos parâmetros)

**2. Retorne  $\theta$  (parâmetros otimizados)****Figura 4 – Algoritmo de treinamento utilizando para obter os melhores pesos possíveis da MLP**

Fonte: elaborado pelos(as) autores(as).

A função de perda escolhida foi a função de perda do Erro Quadrático Médio (EQM), adequada para problemas de regressão, como é o caso deste projeto. O objetivo é minimizar a discrepância entre a saída da MLP e o áudio de referência distorcido por um pedal de guitarra.

Além desses elementos fundamentais, foram aplicadas técnicas adicionais para aprimorar o treinamento da MLP. Uma dessas técnicas consistiu na redução da taxa de aprendizado a cada 10 épocas. A taxa de aprendizado é um hiperparâmetro crítico que determina a magnitude dos ajustes dos pesos durante o treinamento. A redução periódica da taxa de aprendizado pode ajudar a estabilizar o treinamento e a evitar oscilações indesejadas.

O treinamento da MLP foi conduzido ao longo de 500 épocas, conforme apresentado nas Tabelas 1 e 2. Embora os resultados apresentados nas figuras da seção 4 mostrem resultados até a época 300, é importante destacar que o treinamento foi estendido até a época 500. Após a análise do comportamento da rede durante o treinamento, observou-se que, a partir da época 300, não há mudanças visíveis nas curvas, exceto com um zoom muito alto (intervalo de tempo muito pequeno).

Essa escolha foi feita para otimizar a apresentação dos resultados, eliminando informações redundantes e concentrando a atenção nas fases iniciais do treinamento, onde as mudanças são mais proeminentes. A Figura 2 mostra um trecho curto (0,005 segundos) comparando o áudio com distorção com ganho do resistor de (10 k $\Omega$ ), com distorção com ganho do resistor de (30 k $\Omega$ ) e o áudio sem distorção.

## Avaliação dos resultados

A avaliação dos resultados neste projeto foi conduzida por meio de métricas quantitativas, visando mensurar a qualidade das previsões da MLP e sua similaridade com a fala de referência. As métricas adotadas foram o Erro Quadrático Médio (EQM),

teste de Kolmogorov-Smirnov (KS) e a análise da Função de Distribuição Acumulada (FDA).

O EQM foi empregado como uma medida de proximidade entre as saídas previstas pela MLP e os áudios de referência distorcidos. Especificamente, quanto menor o EQM, mais próximas as previsões da MLP estão dos áudios de referência, indicando melhor desempenho da rede.

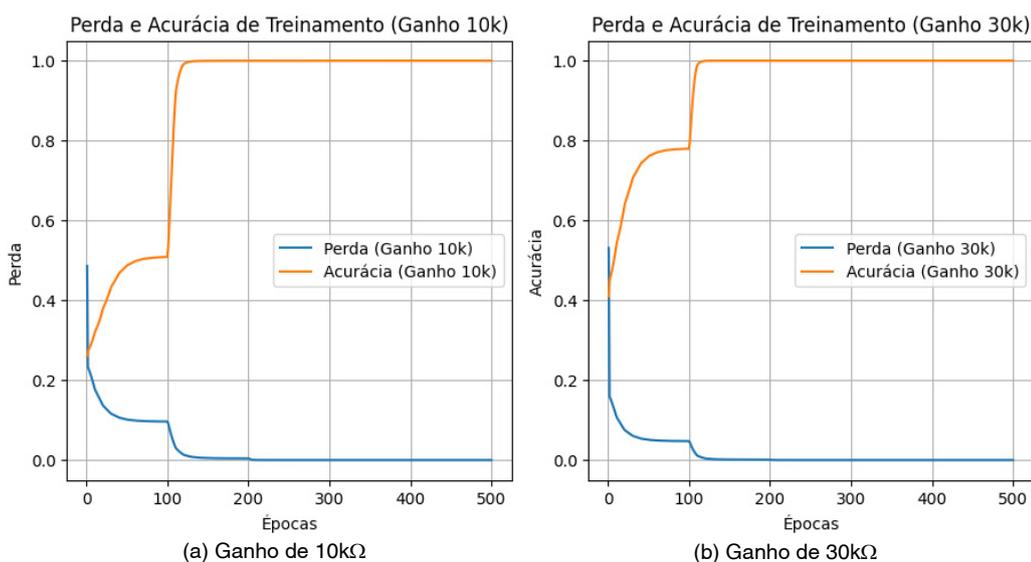
O teste de Kolmogorov-Smirnov (KS) foi aplicado para avaliar a similaridade entre as distribuições acumuladas das amplitudes dos áudios de referência e das saídas geradas pela MLP. Uma estatística KS próxima de zero e um valor- associado (*p-value*) próximo de 1 indicam uma boa concordância entre as distribuições, enquanto valores mais elevados sugerem divergências significativas.

Além disso, para uma compreensão mais aprofundada, comparamos diretamente as saídas das MLPs com o áudio de referência distorcido pelo pedal, como mostrado nas Figuras 8 e 9. Essas figuras destacam visualmente a similaridade entre as previsões da MLP e os áudios de referência.

Para uma análise estatística adicional, examinamos as Funções de Distribuição Acumulada (FDA), geradas a partir das saídas da MLP e dos sinais de referência. As Figuras 5 e 6 apresentam visualmente as FDA para os ganhos de 10k $\Omega$  e 30k $\Omega$ , respectivamente. Esses gráficos proporcionam uma representação visual das distribuições acumuladas das amplitudes, permitindo uma comparação intuitiva entre a saída da MLP e os sinais de referência.

Os resultados dessas métricas foram resumidos nas Tabelas 1 e 2 para os diferentes períodos de treinamento (100, 200, 300, 400 e 500 épocas), considerando os dois diferentes ganhos do resistor (10k $\Omega$  e 30k $\Omega$ ).

Essa abordagem permitiu uma análise abrangente da performance da MLP em diferentes condições de treinamento, proporcionando insights sobre o impacto do ganho do resistor na qualidade das previsões da rede.

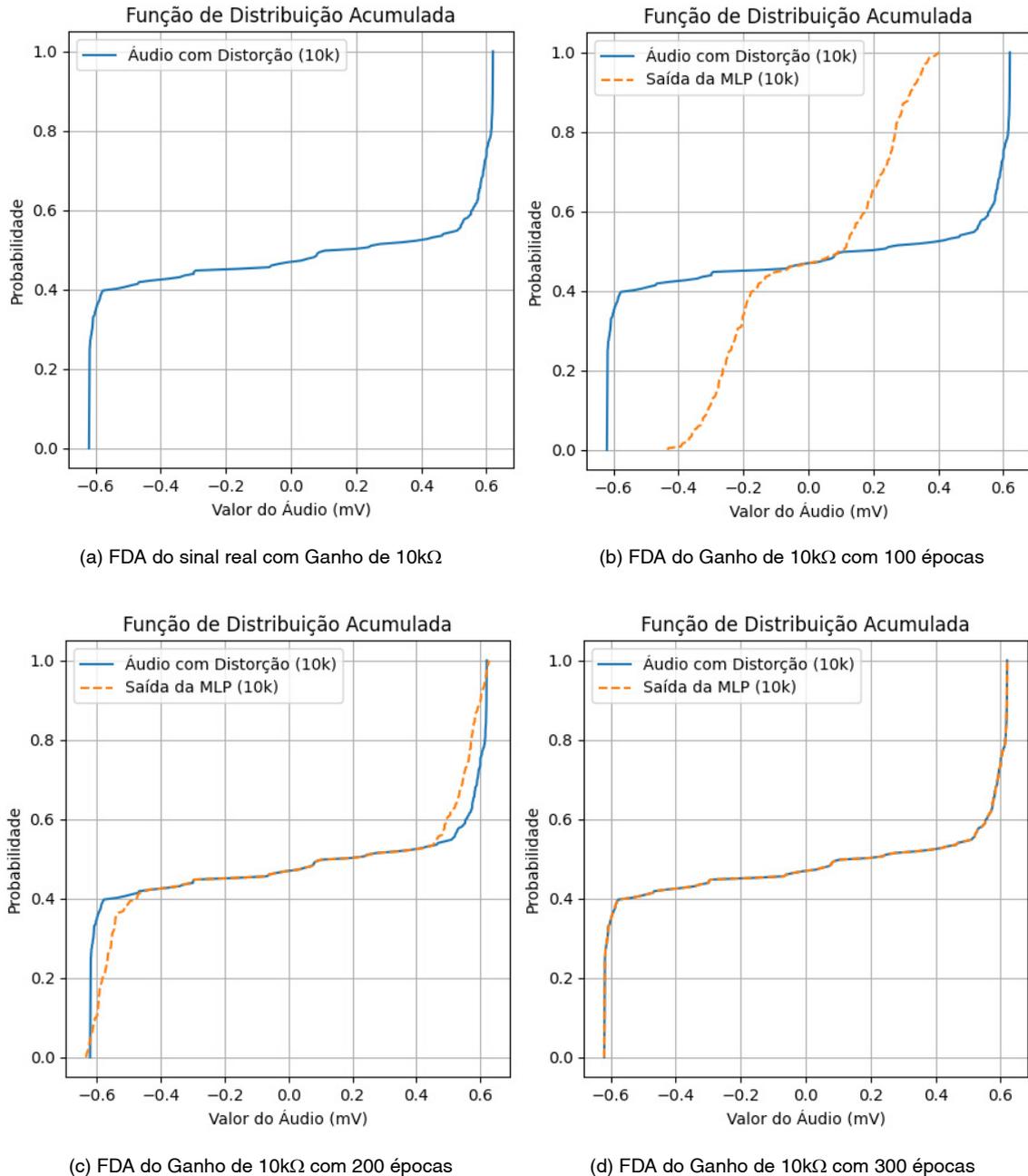


**Figura 5 – Variação da perda e acurácia durante o treinamento**

Fonte: elaborado pelo(as) autores(as).

## Simulações e resultados

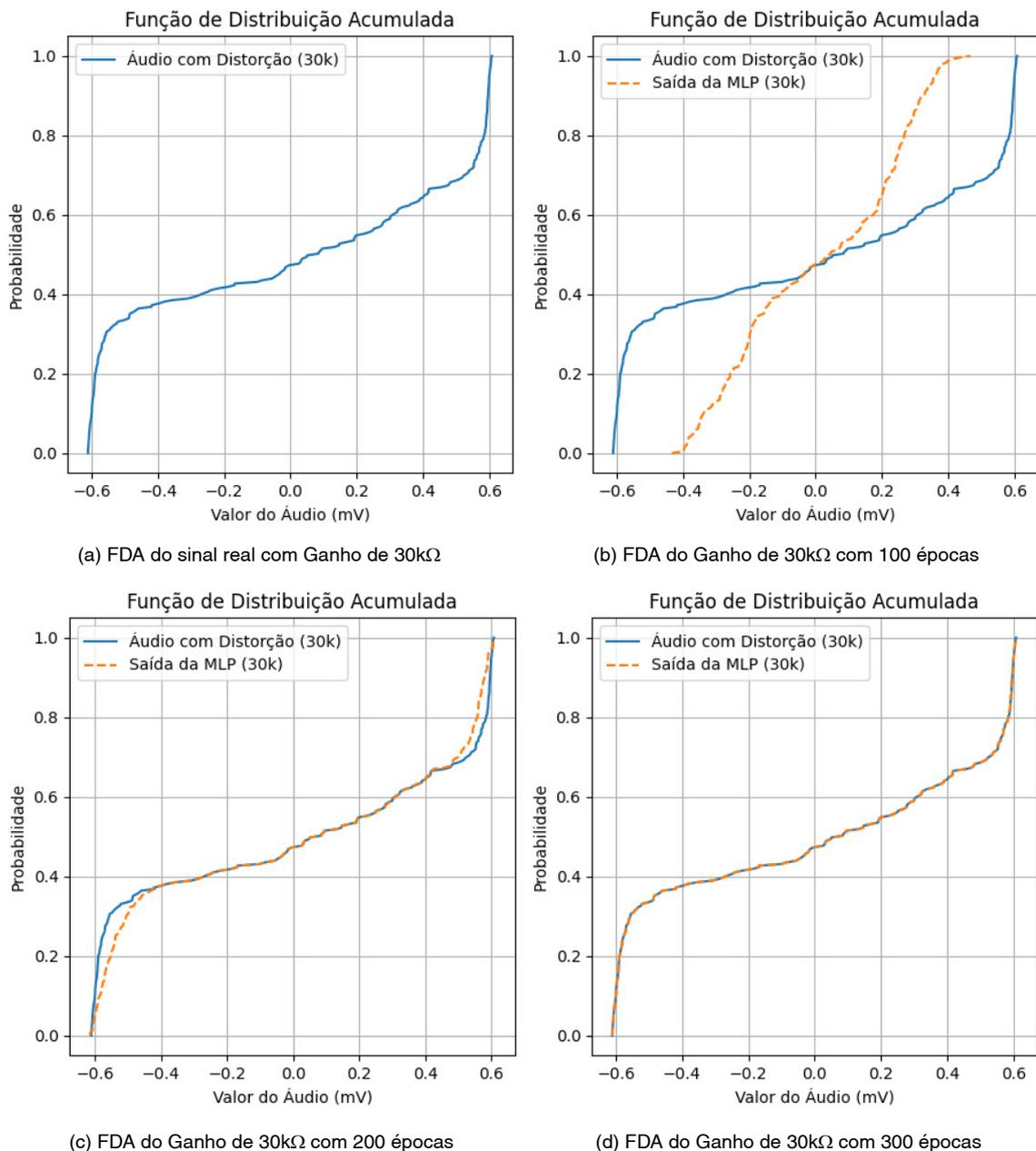
Nesta seção, são apresentados os resultados das simulações conduzidas para avaliar a eficácia do modelo MLP em reproduzir a distorção de áudio em um sinal de guitarra elétrica. Inicialmente, o processo de treinamento do modelo é detalhado, seguido pela análise do desempenho do modelo treinado e pela comparação entre a saída da MLP e os áudios distorcidos.



**Figura 6 – Função de distribuição acumulada para o sinal de áudio com ganho de 10kΩ real e a saída da MLP com 100, 200 e 300 épocas**  
Fonte: elaborado pelo(as) autores(as).

## Treinamento do modelo

O modelo MLP foi treinado utilizando o otimizador Adam (Kingma; Ba, 2017; Zhang *et al.*, 2023) por 500 épocas, seguindo um procedimento de treinamento robusto. É importante salientar que o treinamento do modelo foi realizado executando o código cinco vezes, cada uma com duração de 100 épocas. Em cada execução, os valores da taxa de aprendizado foram redefinidos, garantindo uma abordagem mais ampla na exploração do espaço de parâmetros e promovendo a convergência do modelo em diferentes cenários.



**Figura 7 – Função de distribuição Acumulado para o sinal de áudio com ganho de 30kΩ real e a saída da MLP com 100, 200 e 300 épocas**

Fonte: elaborado pelo(as) autores(as).

Durante o treinamento, a função de perda escolhida foi o Erro Quadrático Médio (EQM), a qual foi continuamente minimizada para ajustar os parâmetros do modelo. Na Figura 5, pode-se visualizar os valores de perda e de acurácia no decorrer das épocas de treinamento para os ganhos de 10 k $\Omega$  e 30 k $\Omega$ . Observa-se uma convergência gradual do modelo, indicada pela estabilização da perda e pelo aumento consistente da acurácia com o progresso das épocas.

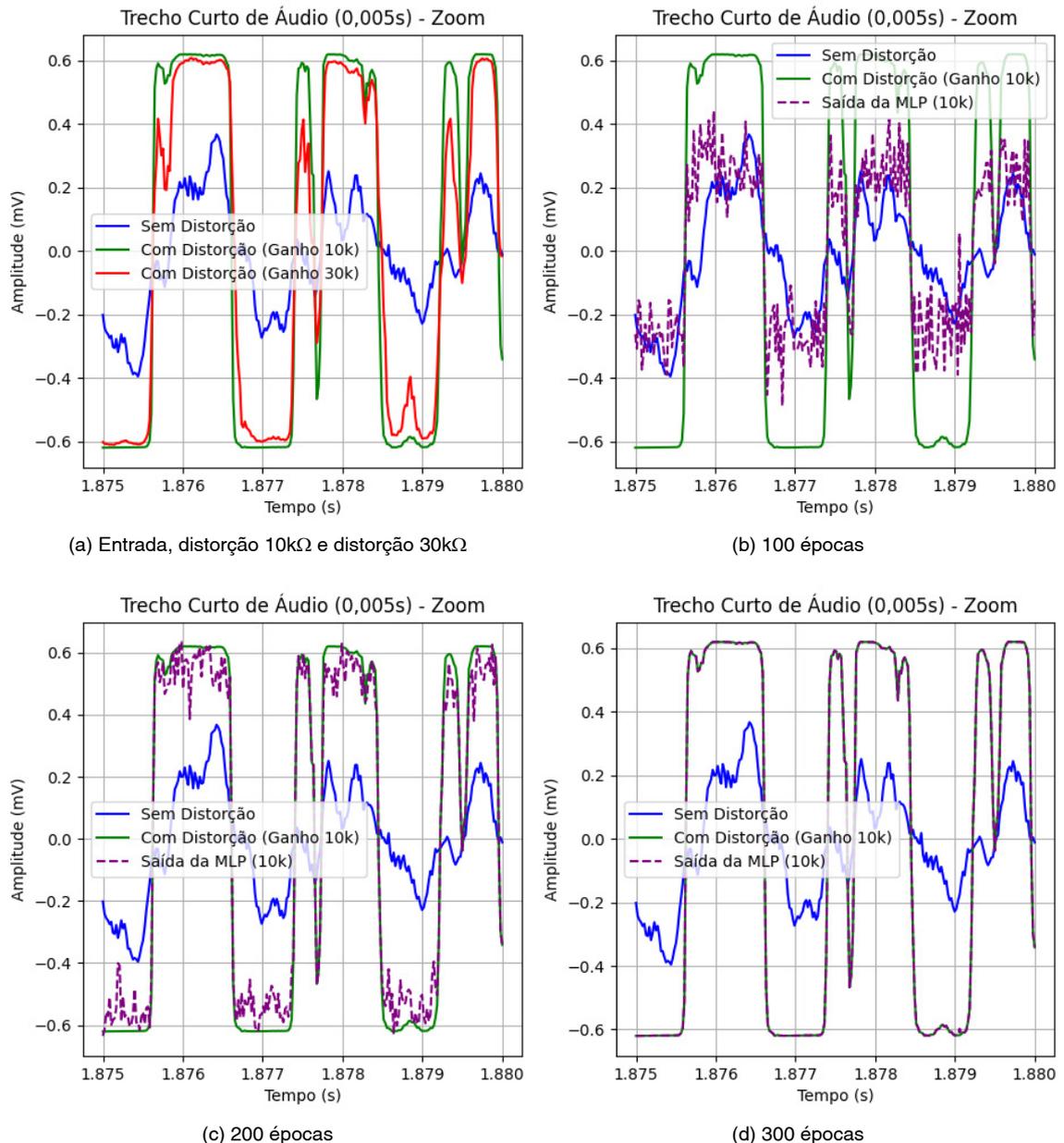


Figura 8 – Comparação entre as saídas da MLP, o áudio com distorção e o áudio sem distorção, para o ganho de 10k $\Omega$

Fonte: elaborado pelo(as) autores(as).

## Avaliação do modelo

A avaliação do desempenho do modelo treinado é um passo crítico para garantir sua capacidade de generalização e aplicabilidade em cenários do mundo real. Neste contexto, foram empregadas métricas objetivas para quantificar a qualidade da saída da MLP em relação ao áudio de entrada distorcido.

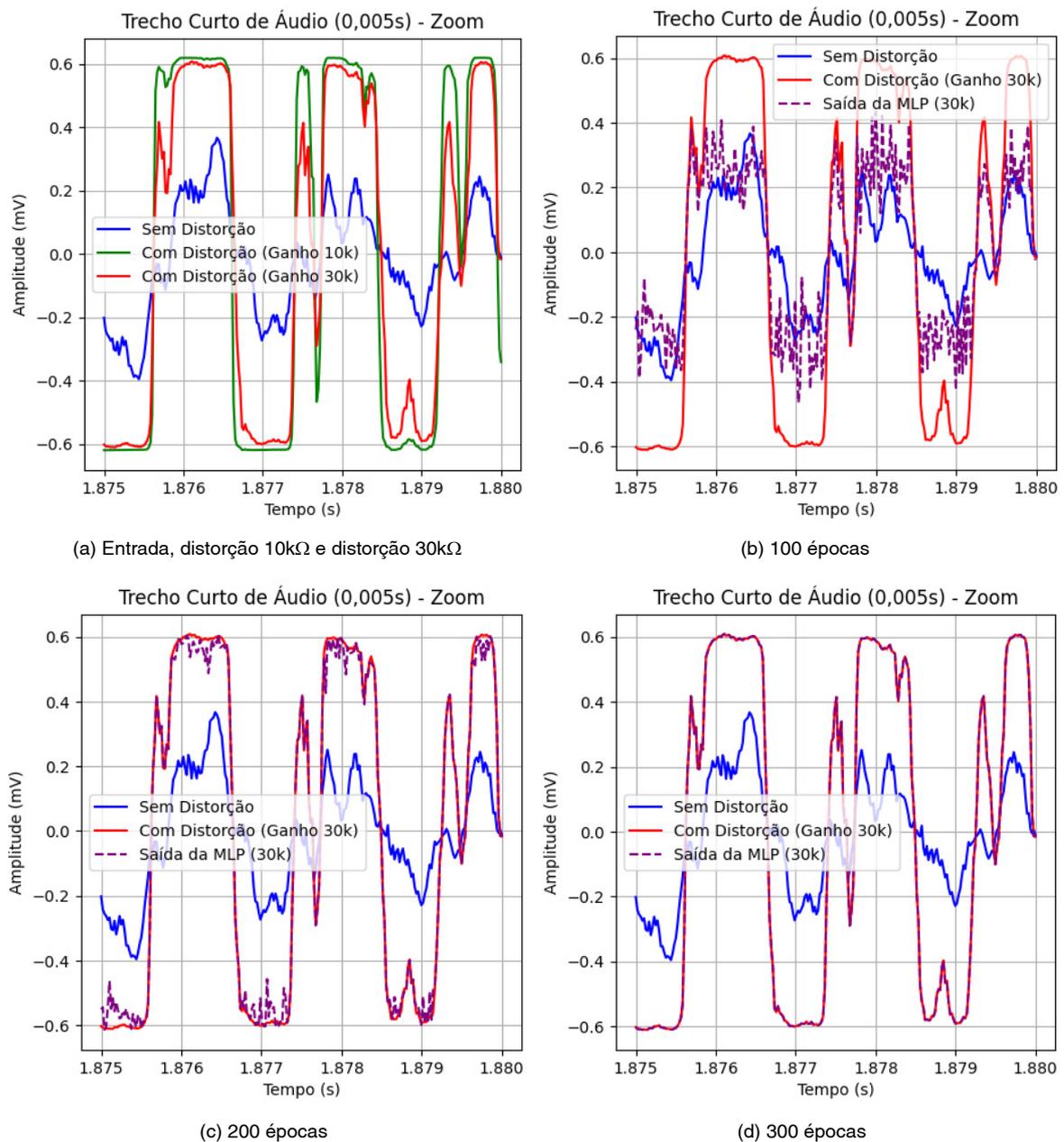


Figura 9 – Comparação entre as saídas da MLP, o áudio com distorção e o áudio sem distorção, para o ganho de 30kΩ

Fonte: elaborado pelo(as) autores(as).

Primeiramente, calculamos a função de perda EQM (Erro Quadrático Médio) entre a saída da MLP e o áudio distorcido. Essa métrica fornece uma medida direta da diferença entre os sinais, permitindo uma avaliação quantitativa do desempenho do modelo em minimizar a distorção.

Além disso, realizamos o Teste Kolmogorov-Smirnov (KS), uma técnica estatística amplamente utilizada para comparar duas distribuições de dados. Neste caso, comparamos a distribuição da saída da MLP com a distribuição do áudio distorcido. O teste KS fornece uma medida objetiva da semelhança entre as distribuições, permitindo avaliar a fidelidade da saída do modelo em replicar as características do áudio original.

As Figuras 6 e 7 apresentam as funções de distribuição acumulada para os ganhos de 10 k $\Omega$  e 30 k $\Omega$ , respectivamente. Esses gráficos oferecem insights sobre a correspondência entre as distribuições, auxiliando na análise da fidelidade da saída da MLP em relação ao áudio distorcido.

### Comparação entre saída da MLP e áudios com distorção

A comparação entre a saída da MLP e os áudios com distorção é essencial para verificar a capacidade do modelo em replicar com precisão as características do sinal de entrada. Com base nos resultados obtidos no teste KS, observa-se uma diferença estatisticamente significativa entre a distribuição da saída da MLP e a distribuição do áudio distorcido. No entanto, uma análise visual mais aprofundada das Figuras 8 e 9 revela que a saída da MLP é altamente semelhante ao áudio distorcido.

Essa semelhança visual entre a saída da MLP e o áudio distorcido sugere que o modelo foi eficaz na captura e reprodução das distorções presentes no sinal de entrada. As figuras apresentam as comparações entre as saídas da MLP, os áudios com distorção e os áudios sem distorção para os ganhos de 10 k $\Omega$  e 30 k $\Omega$ , respectivamente.

As análises visuais dessas figuras oferecem uma perspectiva valiosa sobre a capacidade do modelo em preservar as características essenciais do sinal de entrada, mesmo após o processo de distorção. Esses resultados corroboram a eficácia do modelo MLP na tarefa de modelagem de distorção de áudio.

### Resultados de avaliação

Os resultados obtidos na avaliação do modelo para os ganhos de 10k $\Omega$  e 30k $\Omega$  são apresentados nas Tabelas 1 e 2, respectivamente. Esses resultados são fundamentais para compreender o desempenho do modelo em diferentes condições de treinamento.

Épocas	Teste KS	Valor-p	MSE
100	0.322905	0	0.080315
200	0.116924	0	0.002267
300	0.003869	0	$3.169582 \cdot 10^{-7}$
400	0.003216	0.0000006	$6.741544 \cdot 10^{-9}$
500	0.003034	0.0000034	$6.235293 \cdot 10^{-9}$

Tabela 1 - Resultados de Avaliação para Ganho de 10 k $\Omega$

Fonte: elaborado pelos(as) autores(as).

Épocas	Teste KS	Valor-p	MSE
100	0.2045111	0	0.0313119
200	0.0211680	0	0.0001642
300	0.0004034	0.99980123	$6.259833 \cdot 10^{-9}$
400	0.0004291	0.99936664	$6.068709 \cdot 10^{-9}$
500	0.0004416	0.99896655	$6.044989 \cdot 10^{-9}$

Tabela 2 - Resultados de Avaliação para Ganho de 30 k $\Omega$ 

Fonte: elaborado pelos(as) autores(as).

As Tabelas fornecem uma análise detalhada dos valores de teste KS, valor- e EQM para diferentes números de épocas de treinamento. Observa-se que, à medida que o número de épocas aumenta, o desempenho do modelo melhora, como evidenciado pela diminuição dos valores de teste KS e EQM.

Esses resultados destacam a importância do treinamento adequado do modelo, bem como a necessidade de monitorar seu desempenho ao longo das épocas para garantir resultados confiáveis e precisos.

## Considerações finais

Este estudo investigou a modelagem de efeitos de distorção em sinais de guitarra elétrica por meio de uma rede MLP (*Multilayer Perceptron*), desde a conceituação teórica da distorção de áudio até a implementação prática da arquitetura MLP e a análise dos resultados obtidos.

Os experimentos realizados demonstraram que a MLP é capaz de reproduzir o efeito de distorção em sinais de guitarra elétrica com um bom desempenho. Essa constatação foi corroborada pela comparação direta entre os resultados gerados pela MLP e os sinais de referência. A análise do erro quadrático médio (EQM) também corrobora esse bom desempenho, com valores próximos de zero, indicando boa correspondência entre os sinais gerados pela rede e os originais.

Os resultados obtidos no teste KS revelaram diferenças estatisticamente significativas entre as distribuições dos sinais de saída da MLP e os áudios distorcidos. Contudo, os valores de KS foram extremamente baixos, sugerindo uma alta semelhança entre essas distribuições e, conseqüentemente, indicando que o modelo foi capaz de capturar efetivamente as características das distorções presentes nos sinais de entrada.

Adicionalmente, os valores de EQM obtidos para ambos os conjuntos de dados de treinamento foram bastante reduzidos, indicando precisão considerável na reprodução dos sinais de saída desejados pela MLP. A análise das Funções de Distribuição Acumulada (FDAs) também corroborou esses resultados, revelando alta similaridade entre as distribuições dos sinais de saída da MLP e dos áudios distorcidos originais.

Assim, este trabalho representa um avanço significativo no campo da modelagem de efeitos de distorção em áudio, com potenciais aplicações que abrangem desde a indústria musical até a produção de áudio. Possíveis direções futuras de pesquisa incluem explorar as capacidades de outras arquiteturas de redes neurais na modelagem de efeitos de áudio complexos e avaliar a metodologia proposta em outros contextos que envolvam sistemas dinâmicos não lineares.



## Referências

- BISHOP, C. M. *Pattern recognition and machine learning*. Cambridge: Springer, 2006. v. 4.
- BRUNETTO, L. F. M.; SCHMIDT, C. E.; DALLA'ROSA, A. *Seleção de hiperparâmetros para uma rede multi-layer perceptron aplicada na predição do preço da soja*. Revista Tecnia, Goiânia, v. 8, n. 2, p. 1-15, 2023.
- LTSPICE. [S. l.]: Analog, 2023. Disponível em <https://www.analog.com/en/design-center/design-tools-and-calculators/ltspice-simulator.html>. Acesso em: 31 maio 2023.
- ENGEL, J.; RESNICK, C.; ROBERTS, A.; DIELEMAN, S.; NOROUZI, M.; ECK, D.; SIMONYAN, K. *Neural audio synthesis of musical notes with wavenet autoencoders*. PMLR. International Conference on Machine Learning, p. 1068-1077, 2017.
- FACELI, K.; LORENA, A. C.; GAMA, J.; CARVALHO, A. C. P. L. F. *Inteligência artificial: uma abordagem de aprendizado de máquina*. Rio de Janeiro: LTC, 2011.
- KINGMA, D. P.; BA, J. Adam: A method for stochastic optimization. Arxiv, [s. l.], v. 1, p. 1-15, 2017.
- KULESHOV, V.; ENAM, S. Z.; ERMON, S. *Audio super resolution using neural networks*. Arxiv, [s. l.], p. 1-8, 2017. Disponível em: <https://arxiv.org/abs/1708.00853>. Acesso em: 5 set. 2025.
- PURWINS, H.; LI, B.; VIRTANEN, T.; SCHLÜTER, J.; CHANG, S.-Y.; SAINATH, T. Deep learning for audio signal processing. *IEEE Journal of Selected Topics in Signal Processing*, [s. l.], v. 13, n. 2, p. 206-219, 2019.
- REISS, J. D.; MCPHERSON, A. *Audio effects: theory, implementation and application*. Florida: CRC Press, 2014.
- ROSS, S. M. *Introduction to probability and statistics for engineers and scientists*. [Estados Unidos]: Academic Press, 2020.
- SELF, D. *Small signal audio design*. 4 th. Waltham: Focal Press, 2023.
- STEPHENS, M. A. *Edf statistics for goodness of fit and some comparisons*. Journal of the American statistical Association, Taylor & Francis, v. 69, n. 347, p. 730-737, 1974.
- WASSERMAN, L. *All of statistics: a concise course in statistical inference*. Berlim: Springer Science & Business Media, 2004.
- ZHANG, A.; LIPTON, Z. C.; LI, M.; SMOLA, A. J. *Dive into Deep Learning*. Cambridge: Cambridge University Press, 2023.