

REDES NEURAIS

Configuração dos hiperparâmetros de uma rede MLP

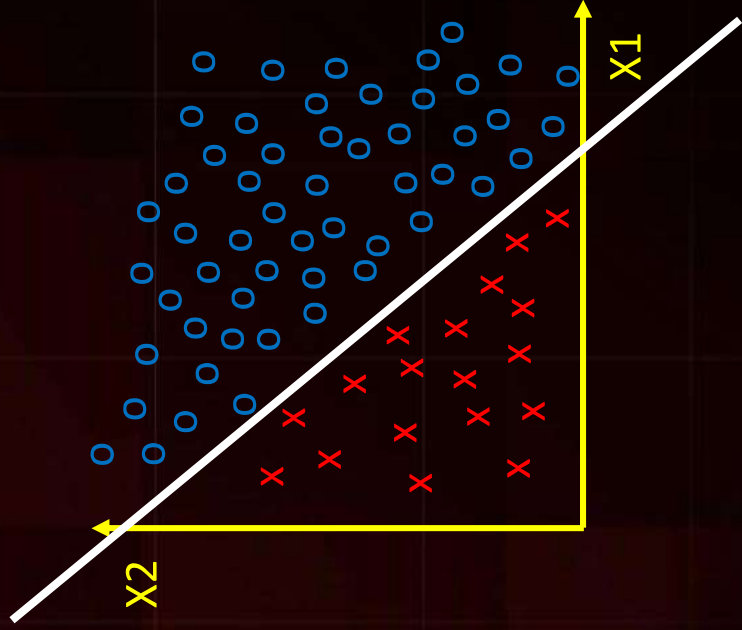


TÓPICOS

1. Configuração da Topologia de uma rede Multilayer Perceptron
2. Aceleração do treinamento da rede

ILUSTRAÇÃO 01

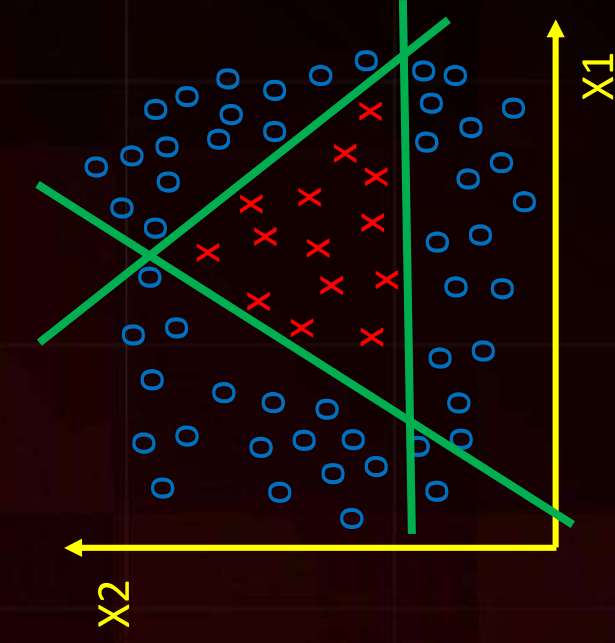
Qual topologia de rede MLP resolve este problema?



Resposta: Um
único neurônio

ILUSTRAÇÃO 02

Qual topologia de rede MLP resolve este problema?



Resposta: 3 neurônios ocultos + 1 neurônio de saída

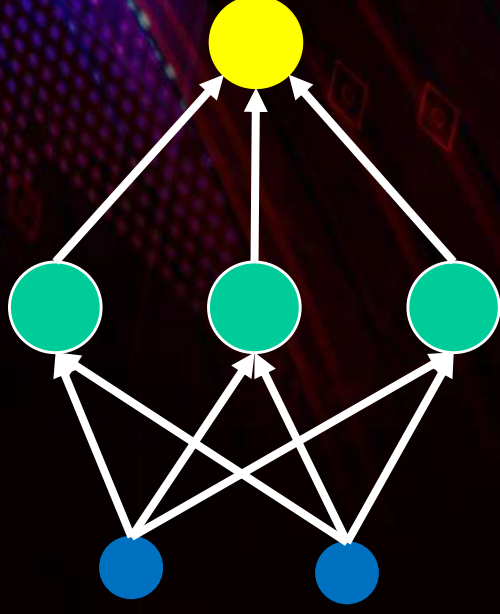
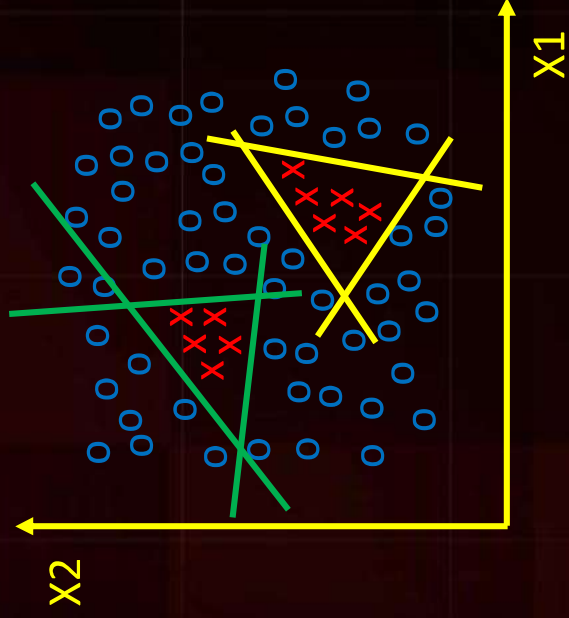
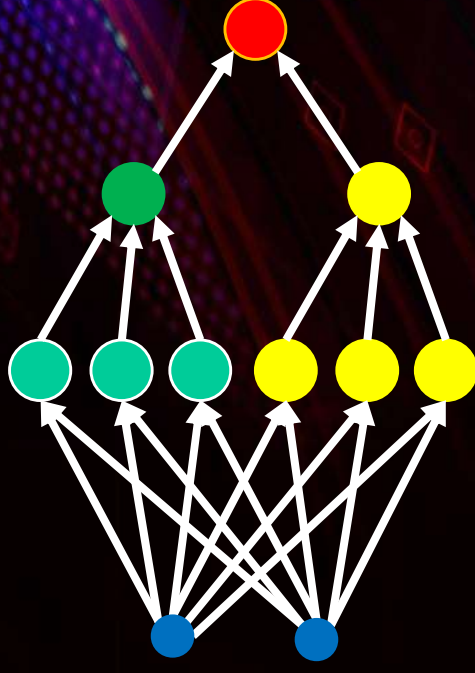


ILUSTRAÇÃO 03

Qual topologia de rede MLP resolve este problema?



Resposta:
6 neurônios primeira camada oculta
2 neurônios segunda camada oculta
1 neurônio de saída



CONFIGURAÇÃO DA TOPOLOGIA

1. QUAL É O PROBLEMA DESSA ABORDAGEM?

R: Difícilmente conseguiremos visualizar a distribuição de um problema real em R^n

Questão: Como configurar a topologia?

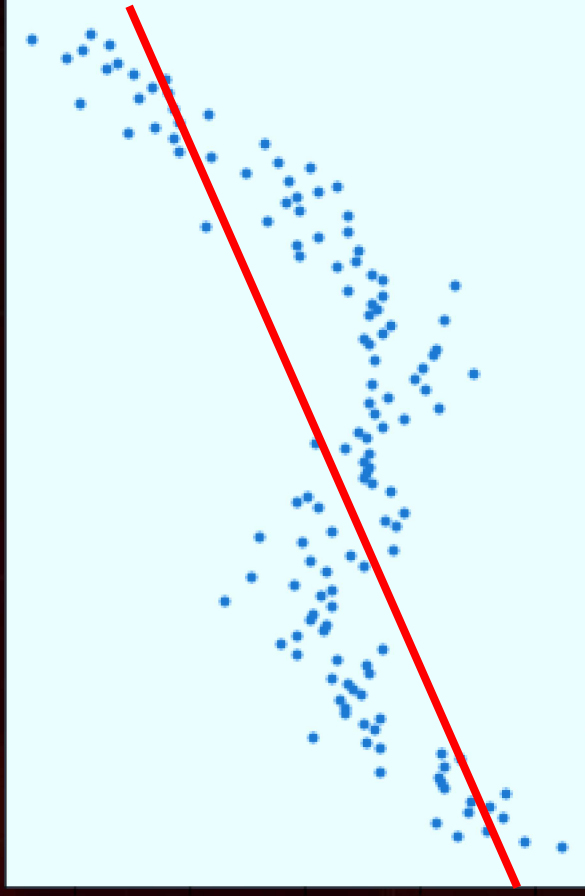
ALGUNS MÉTODOS

1. Força bruta
2. Busca em grade
3. Métodos de otimização
 - a. Algoritmos genéticos
 - b. Simulated Annealing
 - c. Etc.
4. Configuração empírica

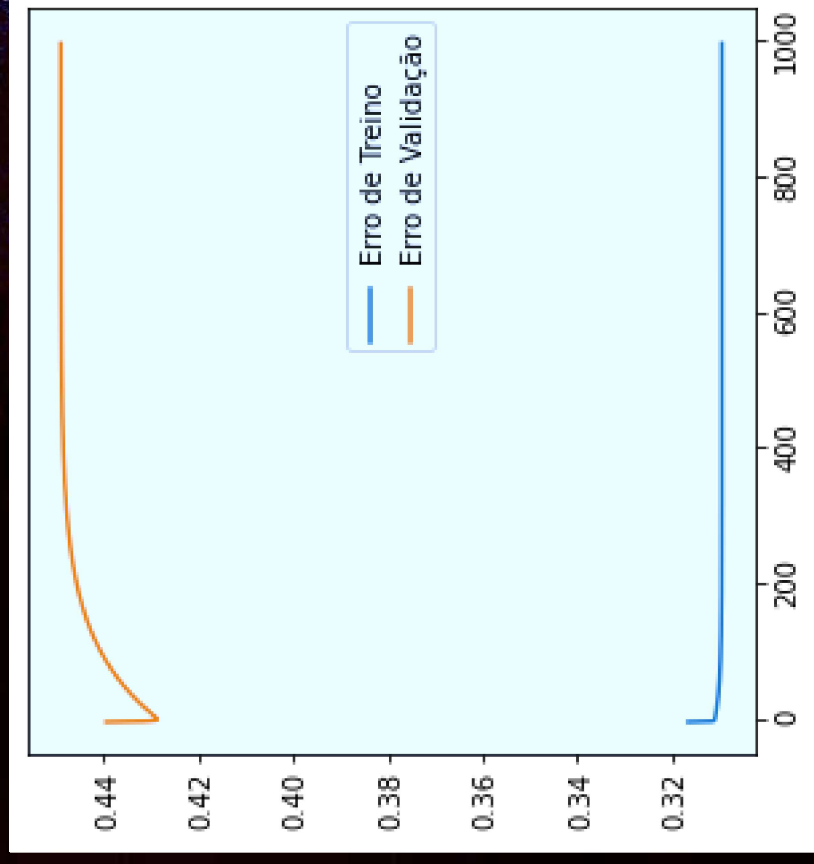
ALGUNS MÉTODOS

1. Divisão do conjunto treino/validação/teste
2. Validação Cruzada

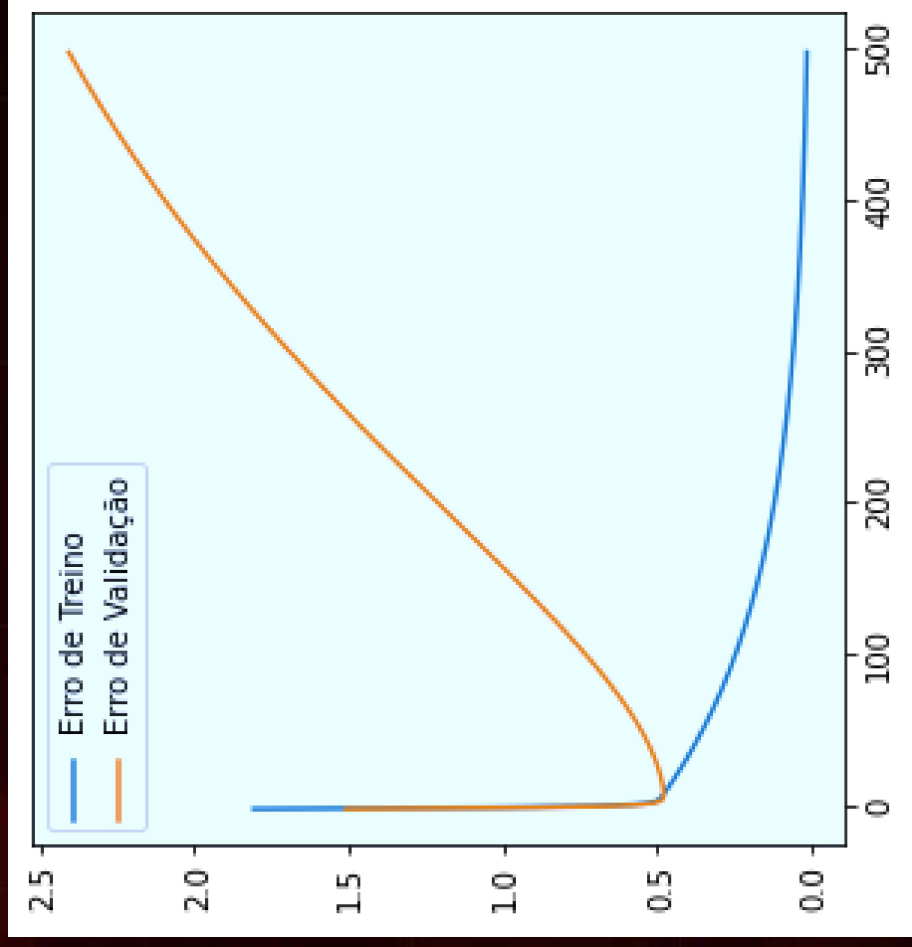
CENÁRIO 01 - UNDERFITTING



Modelo Linear – 1 camada
Erro de treinamento alto



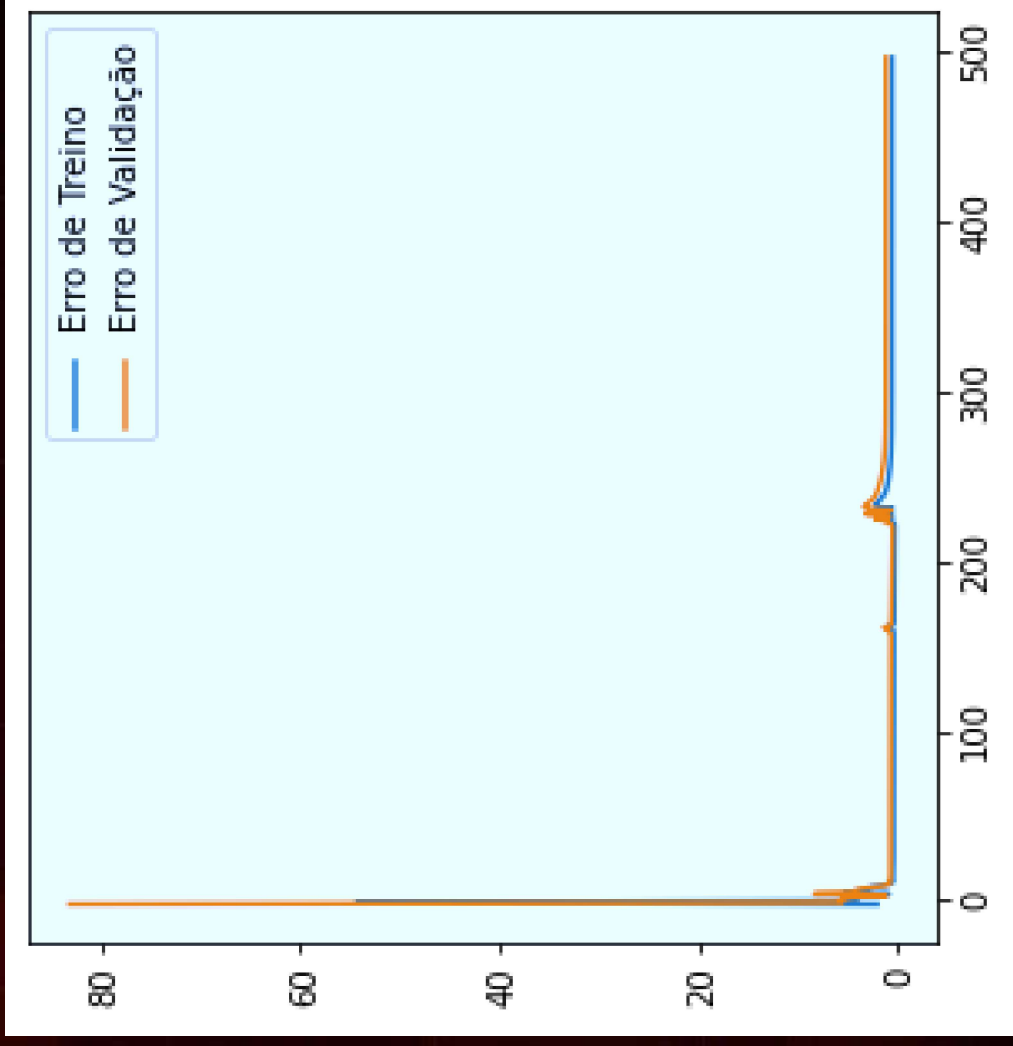
CENÁRIO 02 - OVERFITTING



- Rede com múltiplas camadas
- Erro de treinamento baixo
- Erro de validação alto
- Típico cenário de overfitting (excesso de parâmetros e ajuste)

CENÁRIO 03 - IDEAL

- Rede com múltiplas camadas
- Modelo consegue generalizar
- Baixo erro de treino e de validação



HIPERPARÂMETROS

- Iniciar com a topologia mínima (linear) e avaliar os resultados
- Se o erro de treino é alto: ampliar o número de parâmetros (número de camadas / neurônios)
- Se o erro de treino é baixo: o modelo atual consegue aprender e é um bom modelo inicial

HIPERPARÂMETROS

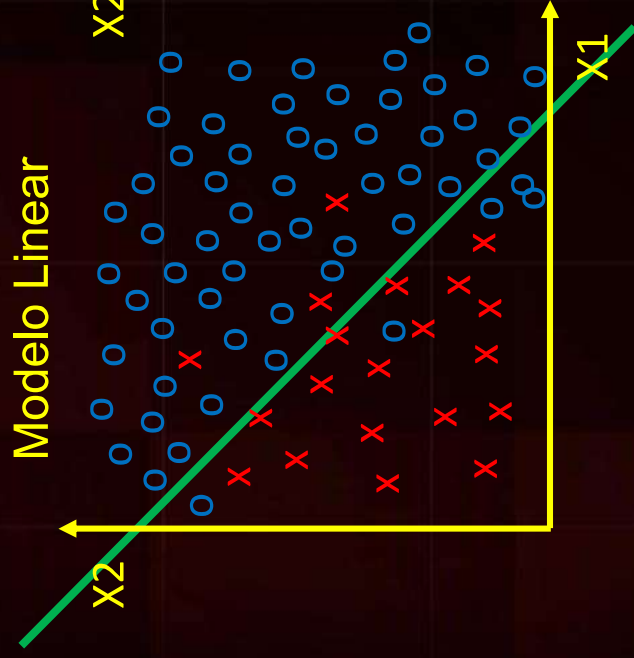
- Quando atingimos um erro baixo de treino: avaliar o erro de **validação**
 - Se o erro é baixo, o modelo está ajustado e consegue generalizar
 - Se o erro de validação é alto, temos a presença de overfitting
 - Reduzir a quantidade de parâmetros
 - Aplicar regularização

BIAS VS VARIÂNCIA

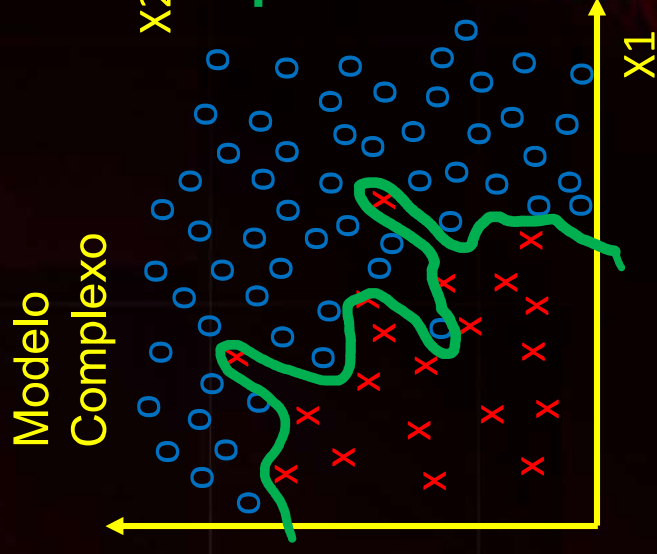
Conjunto de Treino

Conjunto de Validação

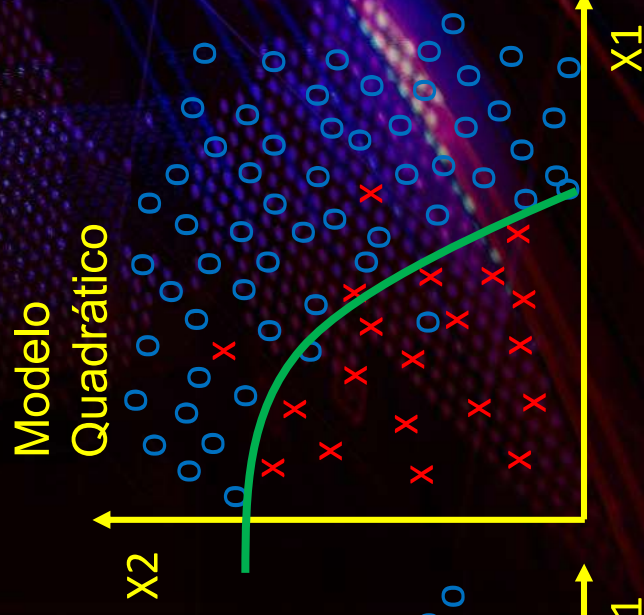
Recordando



Alto viés - Underfitting



Alta variância
Overfitting



Modelo adequado

BIAS VS VARIÂNCIA

- **Ajuste do Bias – Underfitting (treino)**
 - Mais parâmetros
 - Treinar por mais tempo
- **Ajuste da Variância – Overfitting (validação)**
 - Obter mais dados
 - Regularização / ajuste na topologia
- Buscamos pela menor rede capaz de resolver o problema de forma adequada

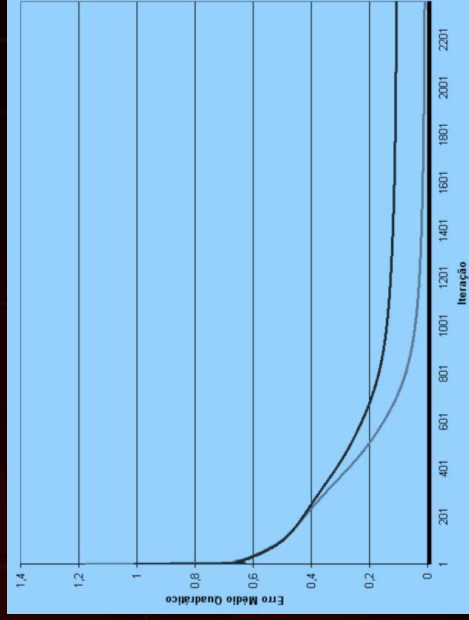
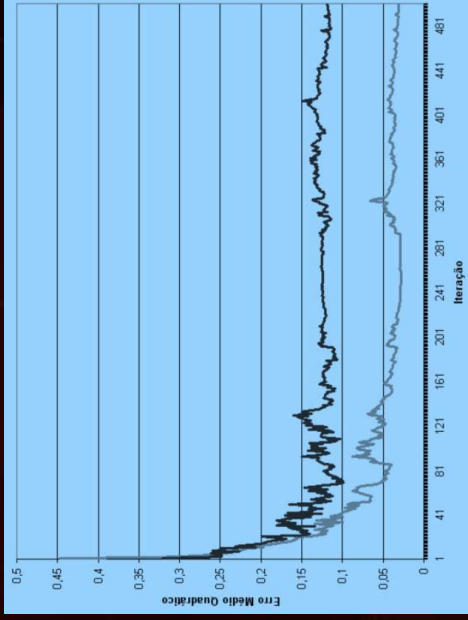
Conjunto de Trei

Conjunto de
Validação

ACELERANDO O TREINAMENTO

- Forma de treinamento (padrão / lote / mini lote)
- Configuração inicial dos pesos
- Taxa de aprendizagem
- Utilizar o termo de momentum
- Adicionar nós intermediários
- Normalização do conjunto de dados

ACELERANDO O TREINAMENTO



- **Modo padrão:**
parâmetros ajustados
para cada exemplo
- **Modo em lote (batch):**
parâmetros ajustados
uma vez por época
- **Mini batch:** Ajuste por
subconjuntos

CONFIGURAÇÃO DOS PESOS

- Iniciar a rede com pesos pequenos e aleatórios
- Podemos ponderar o valor inicial pela quantidade de sinapses do neurônio
- Neurônio com 1000 entradas vs neurônios com 10 entradas

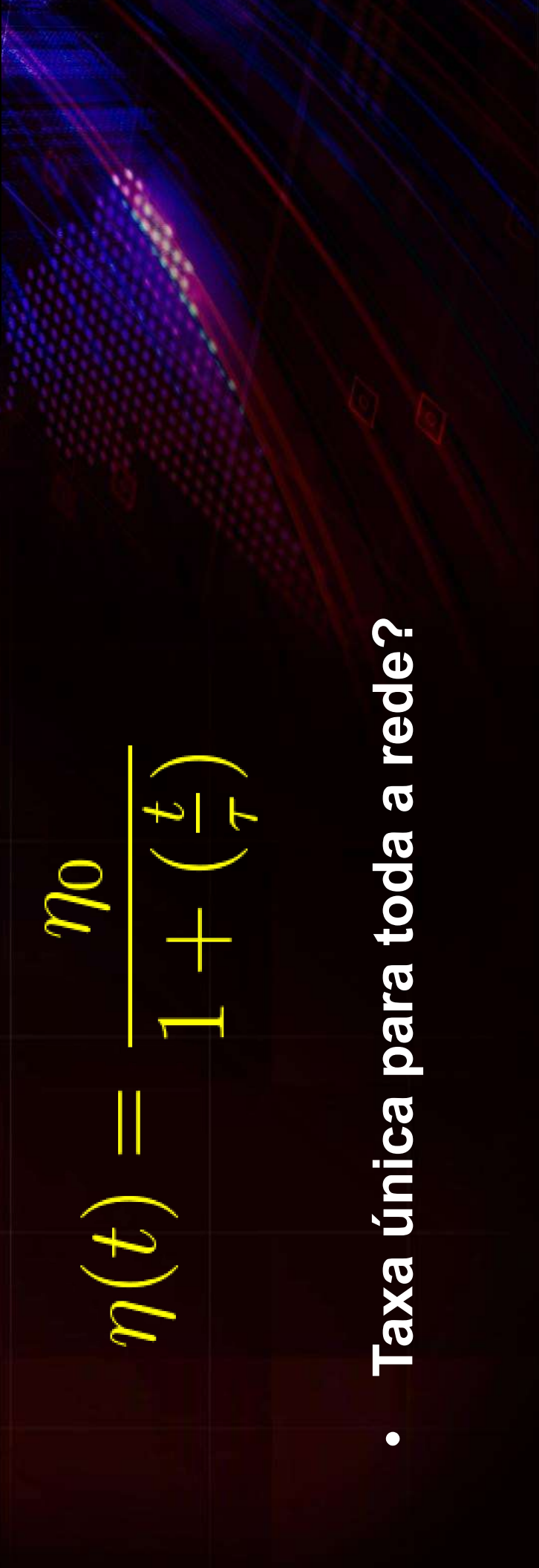
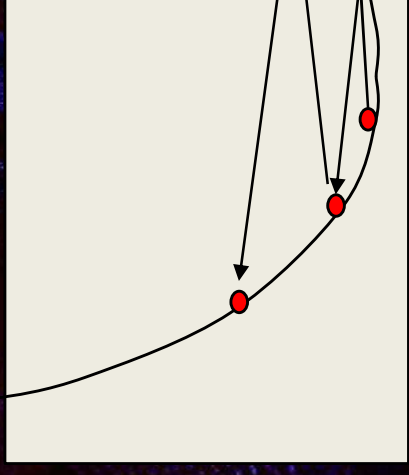
$$v_k = \sum w_{kj} x_j$$

TAXA DE APRENDIZAGEM

- Fixa vs. Decrescente
- Como realizar o decaimento?

$$\eta(t) = \frac{\eta_0}{1 + \left(\frac{t}{\tau}\right)}$$

- Taxa única para toda a rede?



TERMO DE MOMENTUM

- Reduz o perigo de instabilidade
- Gera uma inércia na descida do gradiente

$$\Delta w^t = \eta \delta + \alpha \Delta w^{t-1}$$

$$\Delta w^t = \eta((1 - \alpha)\delta + \alpha \Delta w^{t-1})$$

- Correção do Bias

$$\Delta w^t = \frac{\Delta w^t}{1 - \alpha^t}$$



O QUE VIMOS?

- Aprendemos a configurar a topologia de uma rede MLP utilizando o conjunto de treino e validação
- Conhecemos algumas heurísticas para acelerar o treinamento dos modelos

PRÓXIMA VIDEOAULA

- Conhecer alguns algoritmos de treinamento otimizados
- Entender o processo de regularização da rede (pesos)

ATÉ A PRÓXIMA!!