



Universidade Federal de Pernambuco - DES

Implementação de Rede Neural com Otimizador RMSProp para uso no banco de dados MNIST

UFPE

11 de julho de 2019



DEPARTAMENTO DE ELETRÔNICA E SISTEMAS

Sumário:

- 1 Motivação
- 2 Projeto
- 3 Conclusão

MNIST

Informações básicas

- Hospedado por Yan LeCun
- Coletado de funcionários da United States Census Bureau e de alunos de ensino médio
- O dataset contém 70046 imagens manuscritas dos algarismos 0 a 9

MNIST

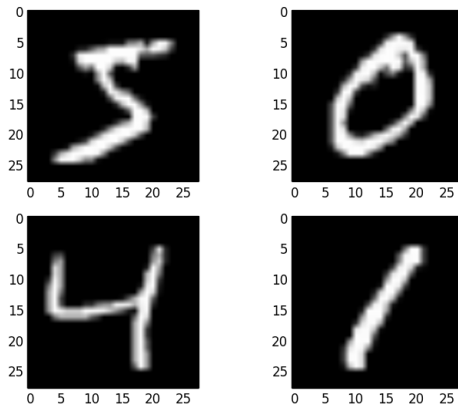


Figura: Exemplos de imagens do MNIST

Tratamento dos dados

- Normalização 0 – 255 para 0 – 1
- *Cross-Validation 10 Folds* : 54024 - 6003 - 10019

Multilayer Perceptron - MLP

- MLP totalmente conectada

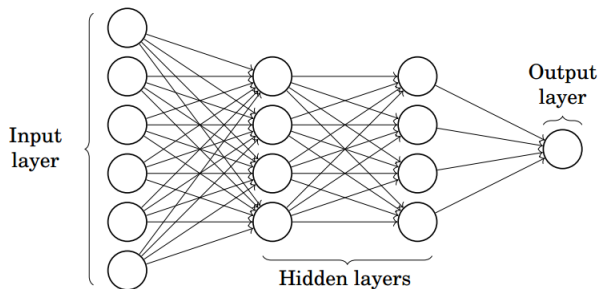


Figura: Arquitetura de uma rede neural MLP

Multilayer Perceptron - MLP

Hiperparâmetros (Parâmetros de Definição do Projetista)

- Quantidade de camadas
- Quantidade de neurônios por camada
- Taxa de aprendizagem
- Função de custo
- Função de ativação
- Modo de treinamento
- Número de épocas

Quantidade de Camadas e Neurônios por Camada

- 2 camadas (1 hidden layer e 1 camada de saída)
- 15 neurônios na primeira camada, 10 na de saída

Taxa de aprendizagem

- Random de 5 valores, entre 10^{-4} e 10^{-6}
- Melhor valor $\alpha = 5 \times 10^{-5}$

Função de custo

$$C = -(y \log p + (1 - y) \log (1 - p))$$

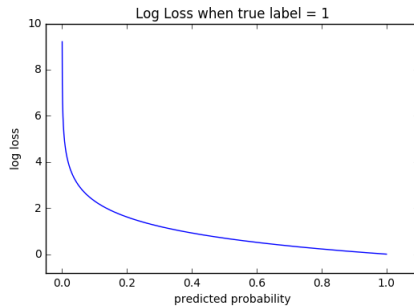


Figura: Função log-loss

Função de ativação

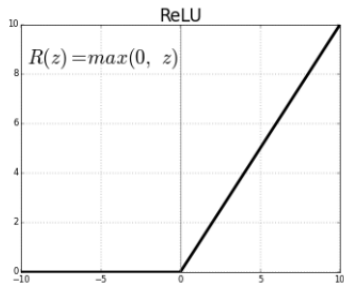
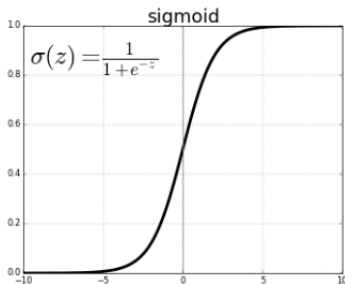


Figura: Funções de ativação usadas

Modo de treinamento

- Gradiente Descendente Estocástico (SGD)

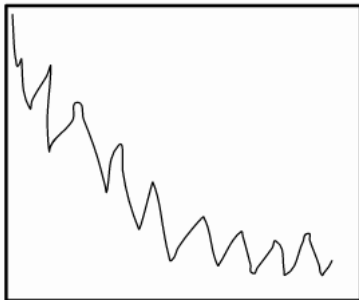


Figura: Gradiente Descendente Estocástico

Número de épocas

- Random de 5 valores, entre 1 e 10
- Melhor valor $e = 10$

RMSPProp

$$E[g^2]_t = E[g^2]_{t-1} + (1 - \beta) \left(\frac{\partial C}{\partial W} \right)^2$$
$$W_t = W_{t-1} - \frac{\eta}{\sqrt{E[g^2]_t}} \frac{\partial C}{\partial W}$$

Cross-Validation 10 Folds

- 6003 dados
- Biblioteca sci-kit learn
- Porcentagem de acerto (conjunto de validação): 82%

Resultados

- Tempo de treinamento total: 500 s com 10 épocas
- Porcentagem de acerto (conjunto de teste): 91%
- Porcentagem de acerto (conjunto de treino): 90,5%

Conclusão

- Bom resultado \neq alta complexidade

O que é um neurônio?

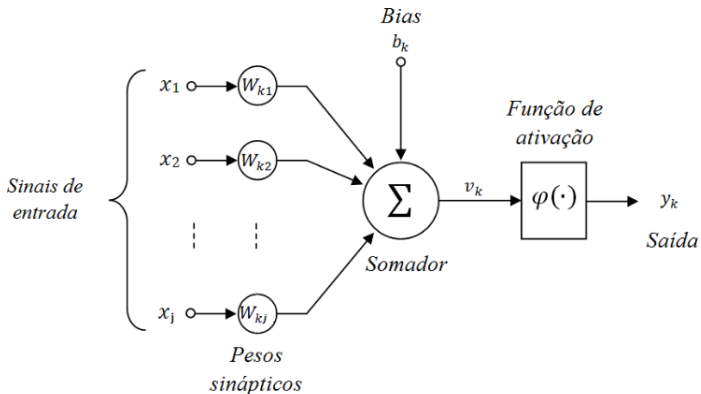


Figura: Um neurônio artificial

Forward Propagation

$$z = W^T x + b$$

$$a = \text{relu}(z)$$

Back Propagation

$$W_t = W_{t-1} - \Delta W$$

$$\Delta W = \eta \frac{\partial C}{\partial W}$$

$$b_t = b_{t-1} - \Delta b$$

$$\Delta b = \eta \frac{\partial C}{\partial b}$$

thank
you!