

Redes Neurais Artificiais

Mariano Faustino Martins A. Sobrinho

Universidade Federal do Amazonas

12 de junho de 2023



Sumário

- 1 Introdução
- 2 Tipos de Redes Neurais
- 3 Topologia de Redes Neurais
- 4 Discussões
- 5 Referências

O que são Redes Neurais?

Redes neurais são modelos computacionais inspirados no cérebro humano, usados para tarefas de aprendizado de máquina e inteligência artificial.

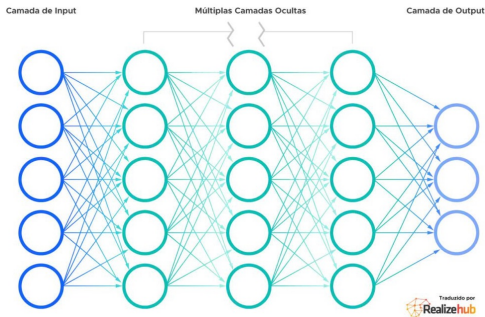
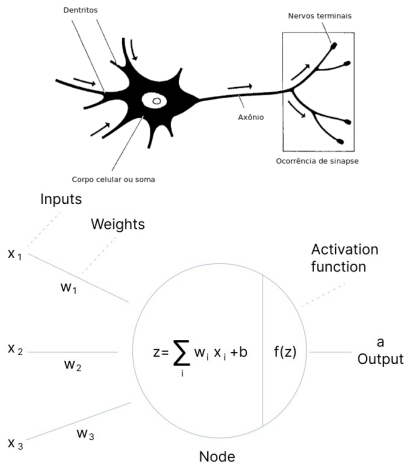


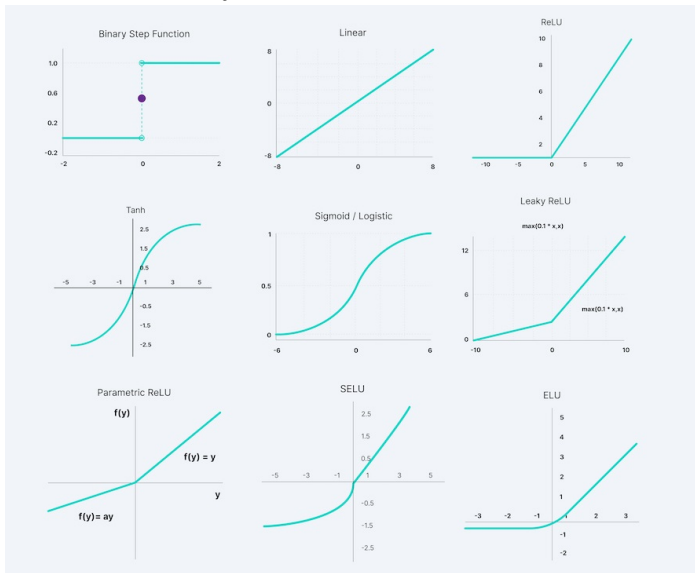
Figura: Uma ilustração de uma rede neural.

Componentes de uma Rede Neural

Uma rede neural é composta por unidades de processamento (neurônios) conectadas por links ponderados (sinapses).



Funções de ativação



Como as Redes Neurais Aprendem?

A rede aprende um mapeamento entrada-saída

$\mathbf{f} : \mathbb{R}^d \mapsto \mathbb{R}^m$ que minimiza uma função de perda genérica $\mathcal{L}(\mathbf{f}(x), y)$, com (x, y) cada par de entrada-saída.

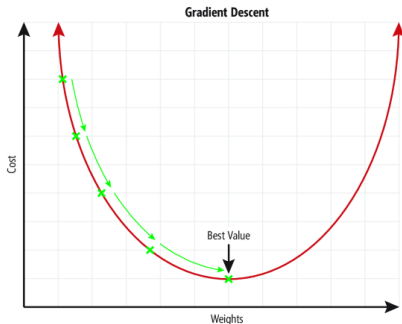
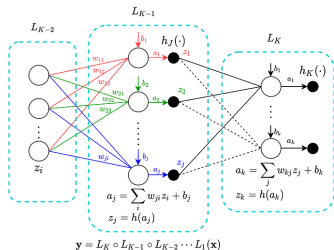


Figura: O processo de aprendizado em uma rede neural.

Aprendizado supervisionado e não supervisionado

No aprendizado não supervisionado, os modelos são treinados usando um conjunto de dados não rotulado. O objetivo é encontrar padrões ou estruturas ocultas nos dados.

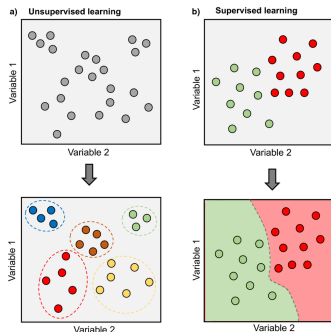


Figura: Uma ilustração do aprendizado supervisionado e não supervisionado

Redes Multilayer Perceptrons (MLP)

As MLPs são uma classe de redes neurais feedforward. Elas consistem em múltiplas camadas de nós em uma direção completa do input para o output.

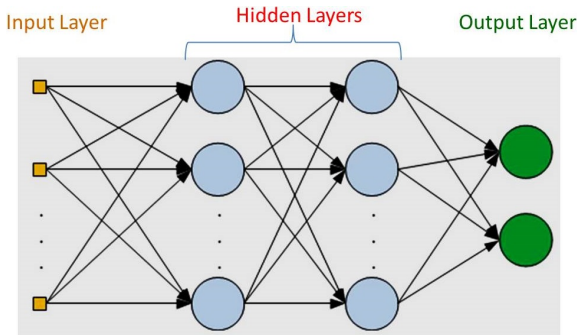


Figura: Uma ilustração de uma rede neural MLP.

Redes Neurais Convolucionais (CNN)

As CNNs são especialmente úteis para tarefas de processamento de imagem. Elas têm camadas convolucionais que podem processar sub-regiões de seus inputs, permitindo-lhes reconhecer características visuais locais.

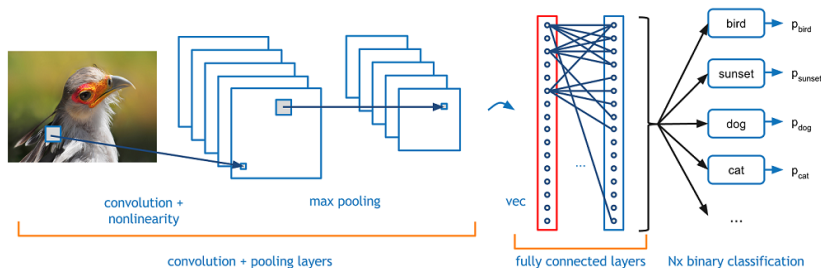


Figura: Uma ilustração de uma rede neural CNN.

Redes Neurais Recorrentes (RNN)

As RNNs têm conexões que formam um ciclo. Isso lhes permite manter informações ao longo do tempo, tornando-as úteis para tarefas que envolvem sequências, como processamento de linguagem natural.

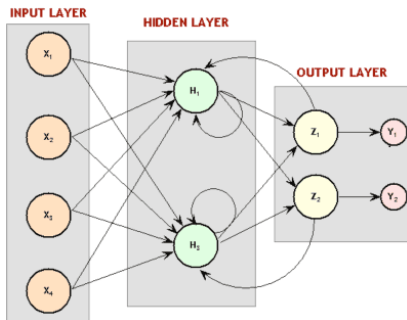
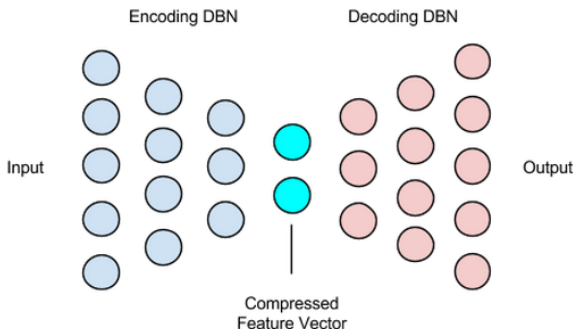


Figura: Uma ilustração de uma rede neural RNN.

Deep Auto-Encoders

Os Auto-Encoders são redes neurais usadas para aprendizado não supervisionado de codificações de dados eficientes. A ideia principal é aprender uma representação (codificação) para um conjunto de dados, geralmente para a finalidade de redução de dimensionalidade.



Aplicações de Redes Neurais na Física

- **Física de Partículas**

- Análise de dados de colisões de partículas

- **Astrofísica**

- Classificação de galáxias
- Identificação de supernovas
- Detecção de ondas gravitacionais

- **Física do Estado Sólido**

- Previsão das propriedades dos materiais

- **Física Nuclear**

- Modelagem da estrutura dos núcleos atômicos

- **Física Quântica**

- Simulação de sistemas quânticos

- **Física Estatística**

- Modelagem de sistemas complexos e não lineares

- **Física Computacional**

- Aceleração de simulações computacionais

Topologia da Aprendizagem em Redes Neurais Artificiais

Topology of Learning in Artificial Neural Networks

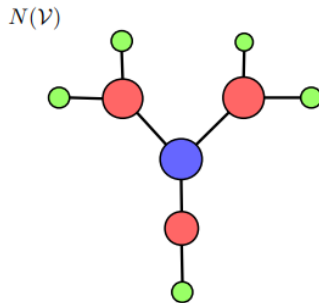
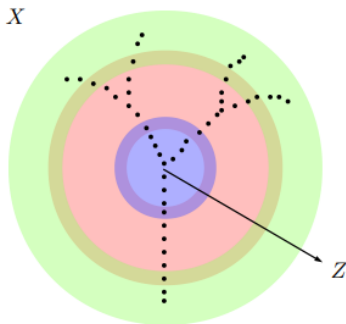
Maxime Gabella
MAGMA Learning
Lausanne, Switzerland
maxime@magmalearning.com

Abstract

Understanding how neural networks learn remains one of the central challenges in machine learning research. From random at the start of training, the weights of a neural network evolve in such a way as to be able to perform a variety of tasks, like classifying images. Here we study the emergence of structure in the weights by applying methods from topological data analysis. We train simple feedforward neural networks on the MNIST dataset and monitor the evolution of the weights.

Inicialização Simétrica e Árvores de Ramificação

Construindo um Mapper graphs que representam o aprendizado dos pesos para cada camada de uma rede neural feedforward totalmente conectada. Para uma dada camada i com neurônios N_i , rastreamos a evolução dos pesos de entrada para cada neurônio."



Dataset MNIST

label = 5



label = 0



label = 4



label = 1



label = 9



label = 2



label = 1



label = 3



label = 1



label = 4



label = 3



label = 5



label = 3



label = 6



label = 1



label = 7



label = 2



label = 8



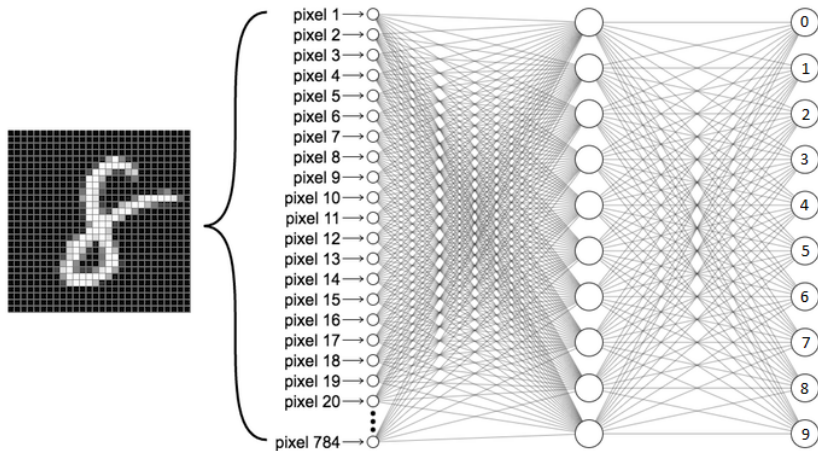
label = 6



label = 9



Modelo de Rede Neural



Inicialização em zero

- Foi utilizado o dataset MNIST de dígitos manuscritos.
- Uma camada oculta de 100 neurônios.
- Os pesos são inicializados com 0.

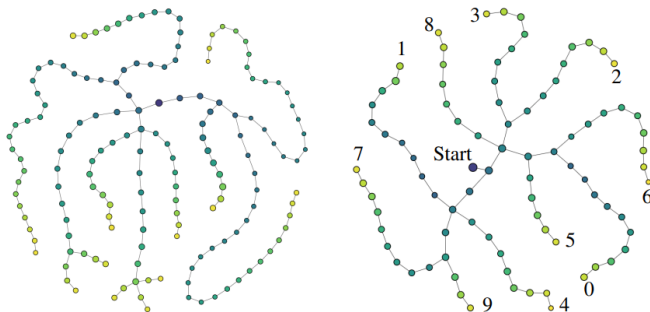


Figura: Evolução dos pesos para uma rede neural com uma camada oculta, inicializada em zero. Esquerda : Camada oculta. Direita : Camada de saída (os dígitos são indicados).

Inicialização com pequenos valores aleatórios

- Duas camadas ocultas de 100 neurônios.
- Os pesos são inicializados com uma distribuição normal com média 0 e desvio padrão $\sigma = 10^{-6}$.

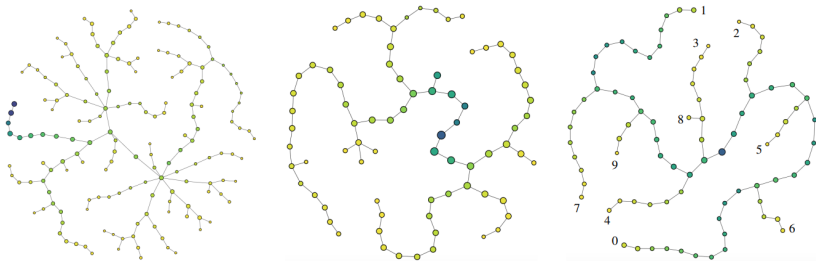


Figura: Evolução dos pesos de uma rede neural com duas camadas ocultas, inicializadas em pequenos valores aleatórios.

Inicialização com pequenos valores aleatórios

Com função de filtro dada pelas 3 primeiras direções do PCA (Análise de Componentes Principais).

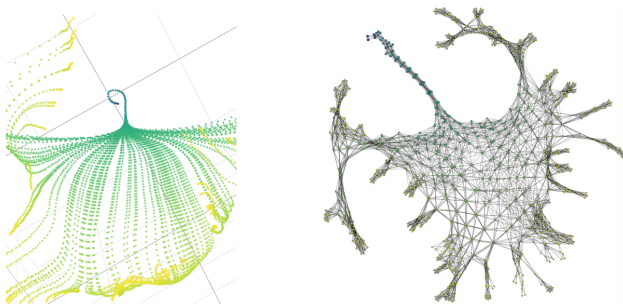
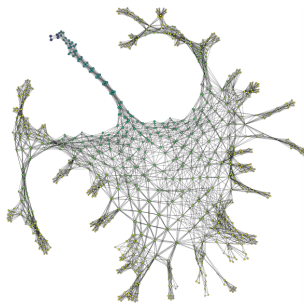
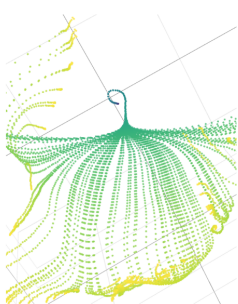


Figura: Esquerda: Detalhe da projeção PCA da evolução dos pesos para a primeira camada oculta. Várias fases são visíveis: evolução uniforme (azul), evolução paralela ao longo de uma superfície (verde), evolução caótica (amarelo). Direita: gráfico de aprendizado representando a superfície como uma grade densamente conectada.

Inicialização com pequenos valores aleatórios

Embora se esperasse que a forma da superfície fosse determinada pelos valores aleatórios exatos dos pesos iniciais, ela parece, de fato, ser independente da semente aleatória. Além disso, a superfície varia continuamente sob pequenas mudanças no número de neurônios, taxa de aprendizado, tamanho do lote ou média e desvio padrão da inicialização normal.



Discussões

Em resumo, a aplicação da teoria das redes complexas no estudo de redes neurais não apenas nos proporciona insights valiosos sobre o processo de aprendizado dessas redes, mas também abre caminhos para aplicações práticas significativas. Por exemplo, essas descobertas podem orientar a seleção de arquitetura e hiperparâmetros em futuros projetos de redes neurais. Além disso, as propriedades quantitativas dos gráficos de aprendizado, como graus ou centralidades têm o potencial de fornecer métricas robustas para aprimorar a eficácia do aprendizado.

Bibliografia

- GABRIELSSON, R. B.; SCHMIDT, M. N.; WINDEKILDE, I. Learning Graphs from Data: A Topological Figure of Merit. 2019. Disponível em: <https://arxiv.org/abs/1902.08160>.
- La Malfa, E.; La Malfa, G.; Caprioli, C.; Nicosia, G.; Latora, V. Deep Neural Networks as Complex Networks. 2022.
- La Malfa, E.; La Malfa, G.; Nicosia, G.; Latora, V. Characterizing Learning Dynamics of Deep Neural Networks via Complex Networks. 2021.
- Testolin, A.; Piccolini, M.; Suweis, S. Deep learning systems as complex networks. 2018.
- Boccaletti, S.; Latora, V.; Moreno, Y.; Chavez, M.; Hwang, D.-U. Complex networks: Structure and dynamics. Physics Reports, v. 424, n. 4–5, p. 175-308, 2006.

END OF PRESENTATION

SLIDE:23

THANK YOU FOR YOUR ATTENTION