

# Fundamentos de Redes Neurais Profundas

**Palavras-Chave:** REDES NEURAIIS, REDES PROFUNDAS, APRENDIZADO DE MÁQUINA

**Autores(as):**

**GABRIEL DIAS VASCONCELOS, FEEC-UNICAMP**

**Prof. Dr. ROMIS RIBEIRO DE FAISSOL ATTUX (orientador) FEEC-UNICAMP**

---

## INTRODUÇÃO:

O aprendizado profundo com redes neurais é um tema bastante discutido e difundido atualmente, devido à sua aplicabilidade em problemas de reconhecimento, com seu fundamento que se inspira no funcionamento de neurônios animais. Em muitos problemas de áreas como visão computacional, reconhecimento de fala e processamento de linguagem natural [Goodfellow et al., 2016], redes profundas em arquiteturas como as de redes convolucionais, hoje, apresentam resultados de ponta.

O campo de redes neurais teve início com o modelo booleano de neurônio idealizado por McCulloch e Pitts [McCulloch e Pitts, 1943], e ganhou enorme impulso com a proposição do algoritmo de retropropagação de erro (backpropagation, BP) [Haykin, 2008] e tem sido um dos pilares modernos do campo de aprendizado de máquina. Do ponto de vista técnico, essas redes são algoritmos não lineares adaptativos, com memória, capazes de atuar com ou sem realimentação.

Na década passada houve, o avanço na área com a popularização dos modelos de aprendizado profundo (deep learning, DL) permitiu processar grandes massas de dados em camadas correspondentes a diferentes níveis de abstração [LeCun et al., 2015]. Neste projeto, propomos a realização de um estudo sobre essas redes, partindo dos fundamentos de aprendizado de máquina e chegando à implementação prática de uma rede neural profunda convolucional em Python para o reconhecimento de imagens.

## METODOLOGIA:

Na primeira etapa do estudo, foi proposta a análise de um conjunto de dados do Kaggle, uma simulação de espalhamento elétron-próton inelástico para a classificação entre 4 tipos de partículas: positron; pion; kaon e proton, baseadas em leituras de 6 detectores de sinais que forneciam os parâmetros: momentum; angulos beta e teta; número de fotoelétrons; energia interna e energia externa. A análise e tratamento dos dados são aplicações do conteúdo visto nas notas de aula da disciplina de pós-graduação "IA048 – Aprendizado de Máquina" [Bocato e Attux, 2020] e [Géron, A. (2022)].

O conjunto de dados é estruturado conforme na figura 1 e sua descrição está posta na figura 2. Foram plotados histogramas das características dos dados do conjunto para analisar graficamente as proporções dos elementos presentes conforme na Figura 3

	id	p	theta	beta	nphe	ein	eout
0	211	0.780041	1.081480	0.989962	0	0.000000	0.000000
1	211	0.260929	0.778892	0.902450	0	0.000000	0.000000
2	2212	0.773022	0.185953	0.642428	4	0.101900	0.000000
3	211	0.476997	0.445561	0.951471	0	0.000000	0.000000
4	2212	2.123290	0.337332	0.908652	0	0.034379	0.049256

Figura 1: cabeçalho do conjunto de dados analisado

	id	p	theta	beta	nphe	ein	eout
count	5.000000e+06	5.000000e+06	5.000000e+06	5.000000e+06	5.000000e+06	5.000000e+06	5.000000e+06
mean	9.941839e+02	1.230053e+00	4.889321e-01	8.995796e-01	3.575243e+00	5.635883e-02	7.033195e-02
std	9.724034e+02	7.414895e-01	2.942161e-01	1.132321e-01	1.779236e+01	7.882675e-02	1.020604e-01
min	-1.100000e+01	2.101170e-01	9.222140e-02	5.500010e-01	0.000000e+00	0.000000e+00	0.000000e+00
25%	2.110000e+02	6.657510e-01	2.817420e-01	8.378760e-01	0.000000e+00	0.000000e+00	0.000000e+00
50%	2.110000e+02	1.066150e+00	4.048670e-01	9.461940e-01	0.000000e+00	3.442440e-02	4.417490e-02
75%	2.212000e+03	1.649650e+00	6.117800e-01	9.855910e-01	0.000000e+00	6.834520e-02	9.365990e-02
max	2.212000e+03	5.286950e+00	2.219270e+00	1.499590e+00	3.490000e+02	8.993820e-01	1.097410e+00

Figura 2: descrição do conjunto de dados

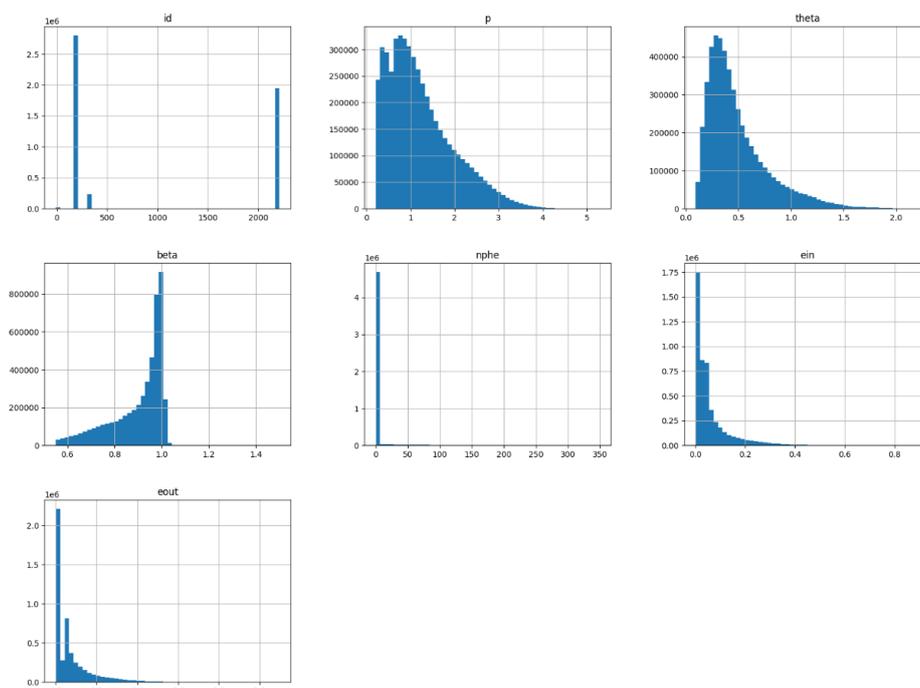


Figura 3: Histogramas dos parâmetros

Após esse processo, uma matriz de correlação foi gerada, para analisar correlações lineares entre os parâmetros, como mostra a Figura 4

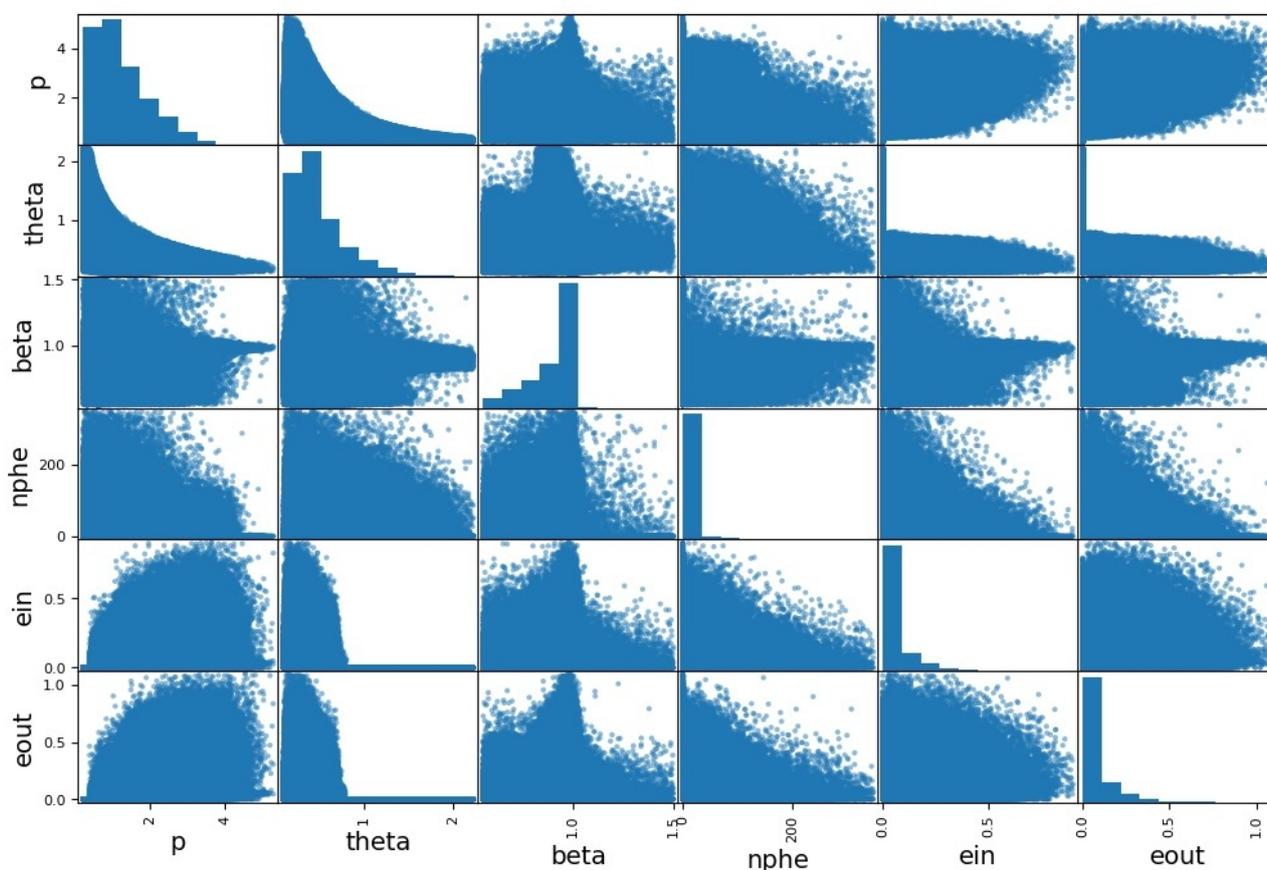


Figura 4: Matriz de Correlação

Foi identificada correlação próxima de meio entre os momentos de partículas e energia interna (41,3%) e externa (54,7%) e próxima de vinte por cento entre momento e ângulos beta (22,7%). Em seguida, dividiu-se a tabela de dados em conjuntos de treinamento e de teste, para evitar o sobreajuste de dados do modelo linear, em uma proporção 80/20, importando da biblioteca “Sklearn” a função “train\_test\_split”, e utilizando a biblioteca “pandas” para remover os dados nulos do conjunto total. No conjunto de dados de treinamento recém tratado prosseguiu-se a utilizar o modelo linear de classificação rotulada com a estratégia de gradiente estocástico descendente, atingindo uma precisão de 92,8%. Em seguida, partimos para a aplicação do modelo não linear de floresta de decisão “Random Tree Regressor”, com três árvores de decisão foi atingida precisão de 99,3% e para verificar que não houve sobreajuste, aplicamos o modelo treinado ao conjunto de testes e obtivemos precisão de 96,7% de resultados.

Na segunda parte do trabalho, foi realizada a aplicação do modelo de classificação linear no conjunto de dados MNIST, que consiste em um banco de dados de setenta mil imagens de dígitos manuscritos por alunos do ensino médio norte americano e funcionários do US Census Bureau como na Figura 5. Pouco tratamento de dados foi necessário no caso desse conjunto por já vir separado em parcelas de treino e teste (60 mil imagens para treino e 10 mil para validação), foi aplicado então o

modelo de classificação linear no conjunto de dados após o escalonamento com a função standard scaler, utilizando novamente o método de gradiente descendente estocástico.



*Figura 5: Dígito do MNIST (5)*

Para avaliar o desempenho do classificador foi utilizado o método de validação cruzada com 3 folds importada da biblioteca “Sklearn” no conjunto de testes e obtivemos valores de precisão de 87% no conjunto de testes. Aplicando o mesmo procedimento com o classificador não linear de floresta aleatória, obtivemos precisão de 99% no conjunto de treino e de 96% no conjunto de teste, levando-nos a conclusão de que houve pouco sobreajuste no processo de implementação.

## **RESULTADOS E DISCUSSÃO:**

Ao aplicar métodos não lineares no conjunto de dados de partículas foi possível atingir scores mais significativos que pelo método linear. Não foi aplicado com sucesso o método não linear com mais árvores de decisão por exigir poder computacional e tempo de computação suficiente para inviabilizar o processo, mas acreditamos que resultados melhores de precisão ainda pudessem ser atingidos sem necessariamente ocasionar em sobreajuste de dados.

Agora, nosso próximo é construir uma Rede Neural Profunda Convolutiva (CNN) para reconhecimento de dígitos MNIST, que é uma aplicação clássica de visão computacional. O objetivo é projetar e implementar uma arquitetura poderosa, capaz de identificar corretamente os dígitos escritos à mão, utilizando camadas convolucionais, de pooling e totalmente conectadas.

## **CONCLUSÕES:**

Durante esse projeto de Iniciação científica exploramos técnicas fundamentais do aprendizado de máquina, como tratamento de dados, regressão e classificação linear, métodos de classificação não linear como florestas aleatórias e redes neurais rasas e profundas, com uma etapa inicial de estudo e aprofundamento nos temas e segunda etapa de aplicações práticas dos conceitos em conjuntos de dados open-source.

O processo viabilizou a capacitação do aluno para o trabalho e pesquisa na área de aprendizado de máquina e forneceu uma visão abrangente acerca da aplicabilidade do machine learning e sua capacidade de tratar com eficácia conjuntos extensos de dados

## **BIBLIOGRAFIA**

[Boccatto e Attux, 2020] L. Boccatto, R. R. F. Attux, Notas de Aula do Curso IA048 – Aprendizado de Máquina, FEEC/UNICAMP, 2020.

[Géron, 2022] A. Géron, Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems, O'Reilly, 2022.

[Goodfellow et al., 2016] I. Goodfellow, Y. Bengio, A. Courville, Deep Learning, MIT Press, 2016.

[Haykin, 2008] Haykin, S. (2008). Neural Networks and Learning Machines. Pearson.

[LeCun et al., 2015] LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep Learning . MIT Press.

[McCulloch e Pitts, 1943] McCulloch, W.S., Pitts, W. A logical calculus of the ideas immanent in nervous activity. Bulletin of Mathematical Biophysics 5, 115–133 (1943).

[Harrison,2018].Particle Identification from Detector Responses,  
<https://www.kaggle.com/datasets/naharrison/particle-identification-from-detector-responses/data> .