

Aprendizagem de máquina e aprendizado profundo

Lux.AI

INSTITUIÇÃO EXECUTORA



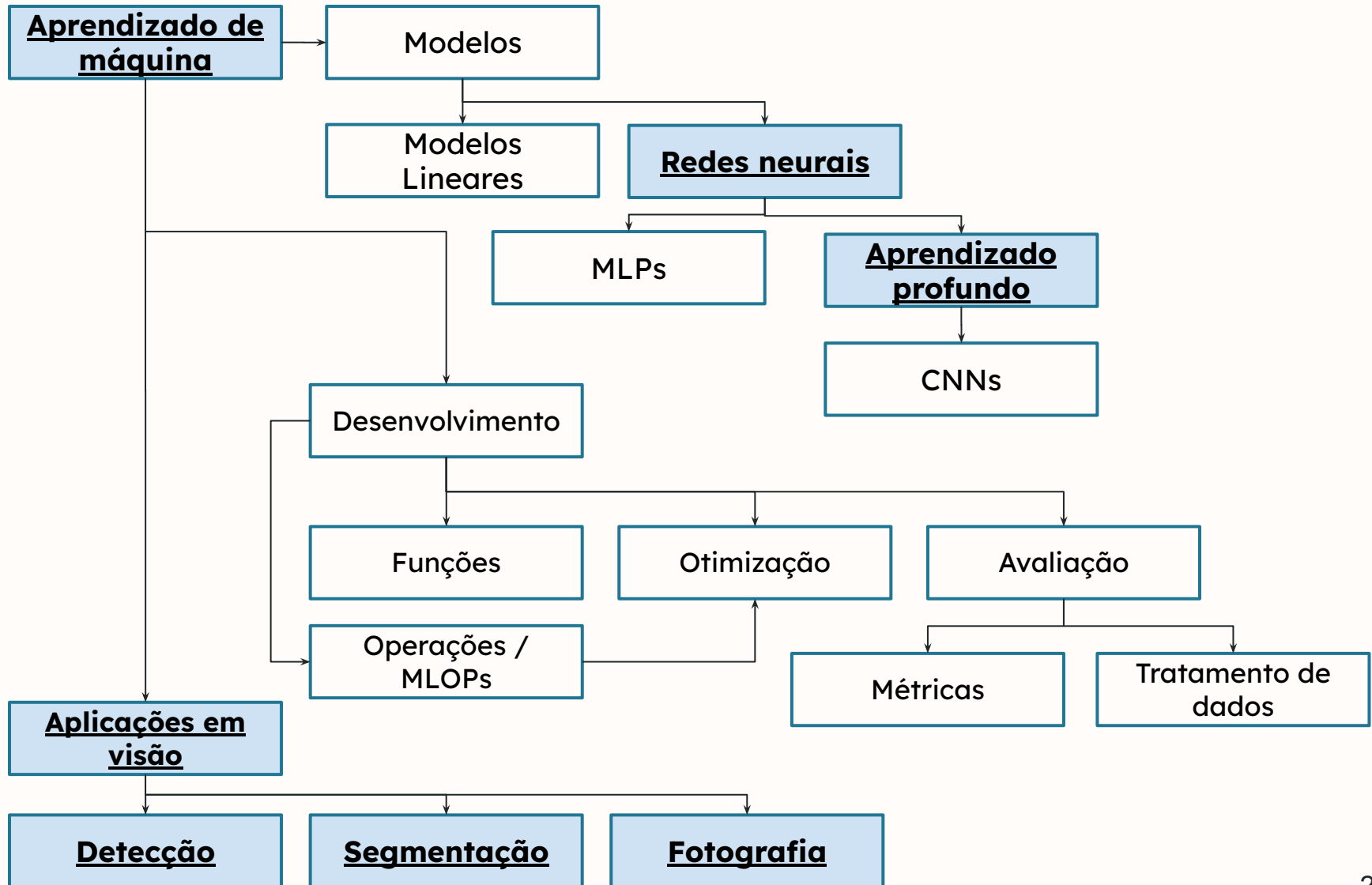
COORDENADORA



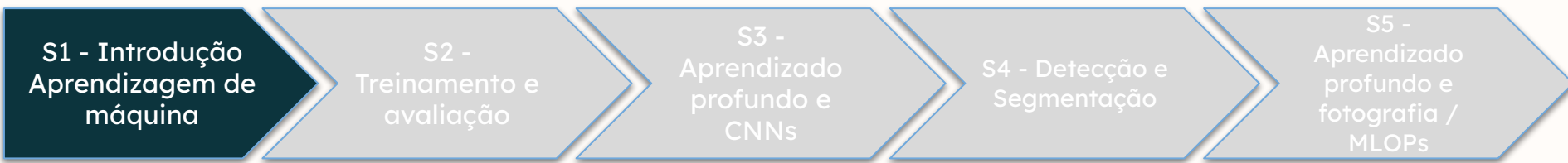
APOIO



Aprendizado de máquina e aprendizado profundo



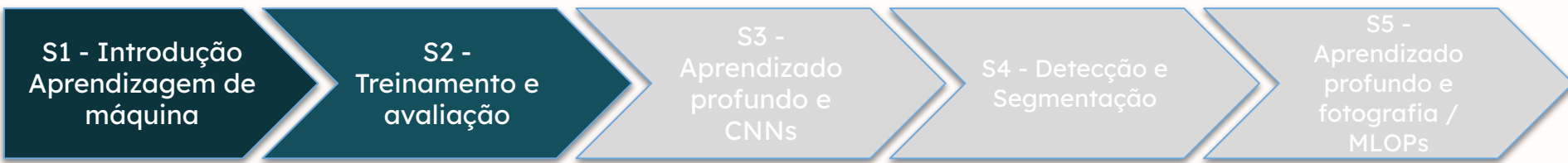
Tópicos



→ Semana 01

- ◆ Introdução ML
- ◆ Modelos lineares
- ◆ Redes Neurais
- ◆ MLPs
- ◆ Backpropagation
- ◆ Otimização

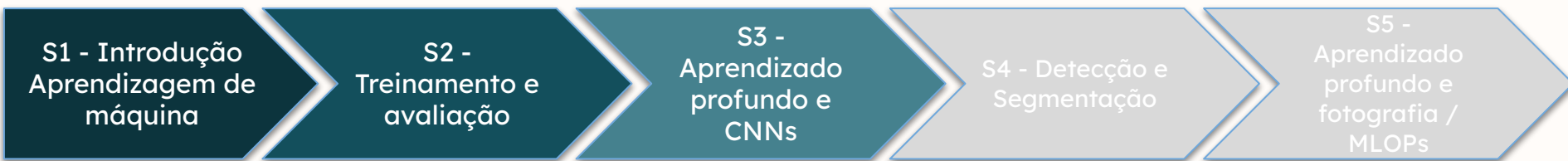
Tópicos



→ Semana 02

- ◆ Treinamento e avaliação
- ◆ Análise de erro
- ◆ Bias e variância
- ◆ Comparação de performance
- ◆ Treino x inferência
- ◆ Técnicas de otimização
- ◆ Regularização
- ◆ Métricas de avaliação

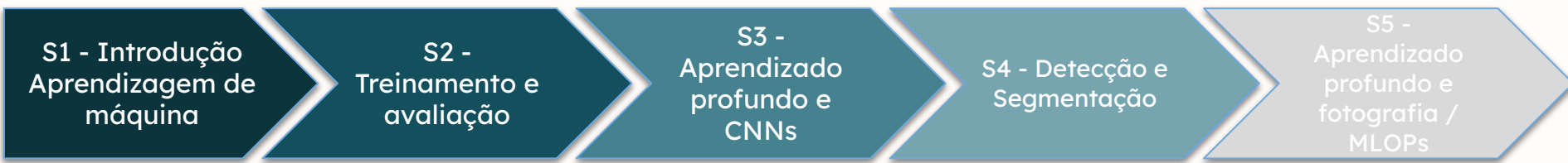
Tópicos



→ Semana 03

- ◆ Redes neurais para imagens
- ◆ Aprendizado profundo
- ◆ Redes convolucionais
- ◆ Transformações e augmentation
- ◆ Transfer Learning
- ◆ Evolução das CNNs
- ◆ Modelos generativos

Tópicos



→ Semana 04

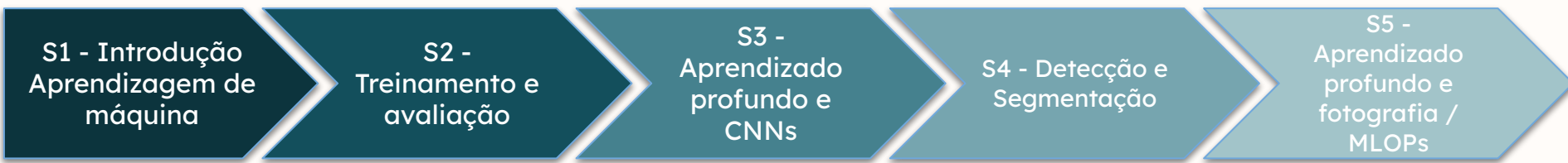
◆ Detecção

- Detecção singleshot x two-stage
- RCNN
- YOLO
- Evolução da arquitetura YOLO

◆ Segmentação

- Arquiteturas de segmentação
- U-Net
- Restauração de imagens

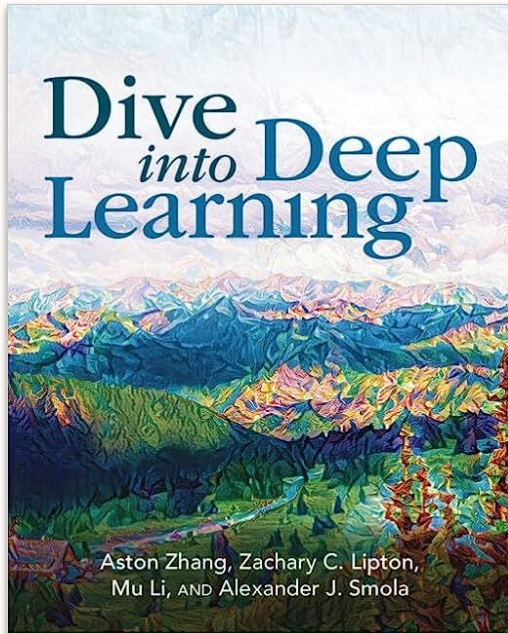
Tópicos



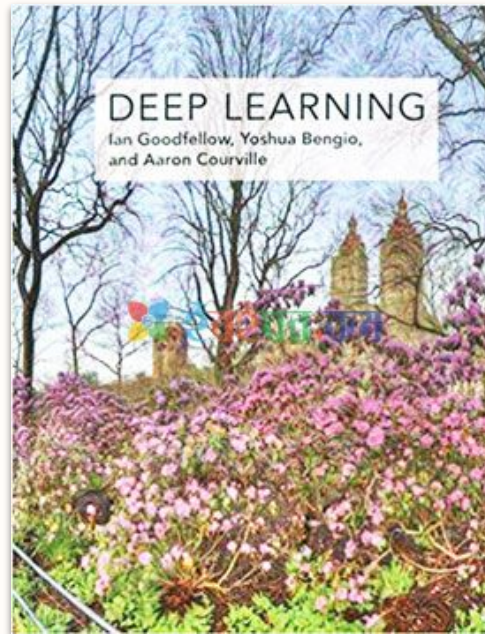
→ Semana 05

- ◆ Aplicações de aprendizado em fotografia
 - Photo-enhancement
 - Deblurring
 - Super resolution
 - Neural Style Transfer
- ◆ Introdução MLOPs

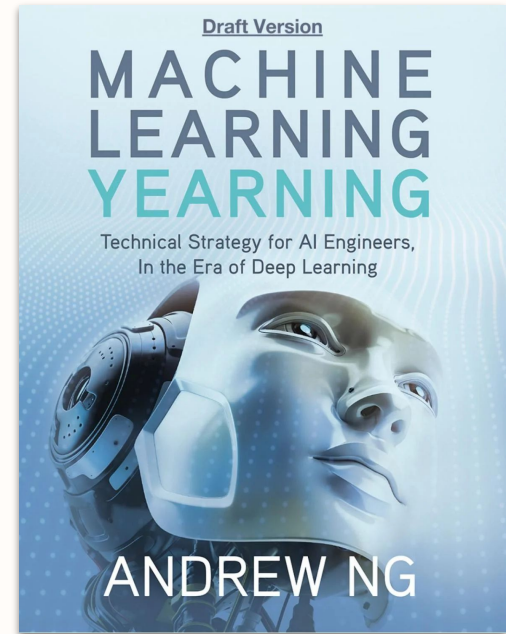
Bibliografia



d2l.ai



deeplearningbook.org



info.deeplearning.ai/machine-learning-yearning-book


Pré-requisitos

- Álgebra Linear
 - Escalares, vetores, matrizes e tensores. Operações com matrizes, normas, determinantes, svd
- Cálculo
 - Funções, derivadas, integrais
- Computação numérica
 - Otimização, overflow, underflow
- Estatística e Probabilidade
 - variáveis aleatórias, probabilidades (marginal, condicional), distribuições de probabilidade, regra de Bayes, variância e covariância, modelos probabilísticos

Pré-requisitos

[1] https://www.deeplearningbook.org/contents/part_basics.html

[2] https://d2l.ai/chapter_preliminaries/index.html



DIVE INTO

DEEP LEARNING

2. Preliminaries

Preview Version

PyTorch

MXNet

Notebooks

Preface

Installation

Notation

1. Introduction

2. Preliminaries

3. Linear Neural Networks for Regression

4. Linear Neural Networks for Classification

5. Multilayer Perceptrons

6. Builders' Guide

7. Convolutional Neural Networks

2. Preliminaries

To prepare for your dive into deep learning, you will need a few survival skills: (i) techniques for storing and manipulating data; (ii) libraries for ingesting and preprocessing data from a variety of sources; (iii) techniques for handling high-dimensional data elements; (iv) just enough linear algebra to understand the theory; (v) the ability to automatically check your work. In probability, our primary language for reasoning about uncertainty when you get stuck.

In short, this chapter provides a rapid introduction to the tools you will need to get started.

2.1. Data Manipulation

2.1.1. Getting Started

2.1.2. Indexing and Slicing

2.1.3. Operations

1 Introduction 1

1.1 Who Should Read This Book? 8

1.2 Historical Trends in Deep Learning 12

2 Applied Math and Machine Learning Basics 27

2 Linear Algebra 29

2.1 Scalars, Vectors, Matrices and Tensors 29

2.2 Multiplying Matrices and Vectors 32

2.3 Identity and Inverse Matrices 34

2.4 Linear Dependence and Span 35

2.5 Norms 37

2.6 Special Kinds of Matrices and Vectors 38

2.7 Eigendecomposition 40

2.8 Singular Value Decomposition 42

2.9 The Moore-Penrose Pseudoinverse 43

2.10 The Trace Operator 44

2.11 The Determinant 45

2.12 Example: Principal Components Analysis 45

3 Probability and Information Theory 51

3.1 Why Probability? 52

i

Ferramentas



colab.research.google.com
pytorch.org/
scikit-learn.org/
numpy.org/
pandas.pydata.org/

Aprendizagem de máquina e aprendizado profundo

INSTITUIÇÃO EXECUTORA



COORDENADORA



APOIO



Conceitos Preliminares

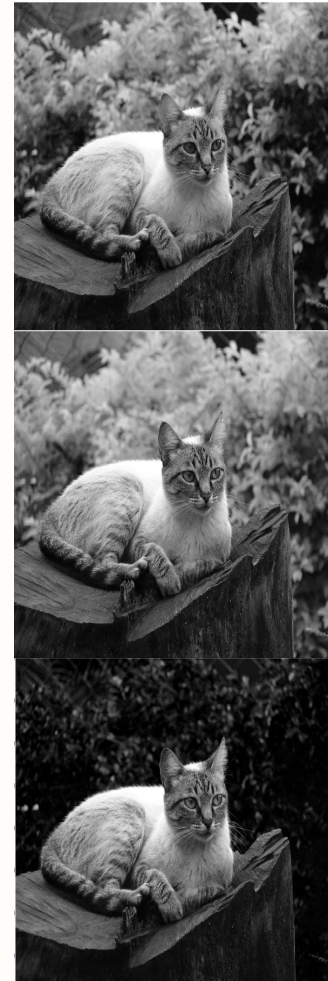
Pré-processamento

- Durante o processo de aquisição de dados, muitas variáveis geralmente não estão na mesma escala. O exemplo a seguir demonstra uma extração de informações de alguns indivíduos.

Indivíduo	Ano de Nasc.	Altura (m)	Peso (kg)
01	1996	1,80	83
02	2001	1,65	80
03	1980	1,70	90
04	1998	1,86	70
05	2005	1,67	68

Pré-processamento

- No caso de imagens, os pixels geralmente estão disponibilizados na escala RGB que estão no intervalo entre 0 e 255.



Pré-processamento

- Para resolver esse problema de escala muito difusa, a aplicação do pré-processamento sobre os dados é de extrema importância, uma vez que os pesos treinados no processo pode ser sensível a essas variações.
- Aplicações de uma normalização como a Standard Scaler é frequentemente utilizada para resolver esses problemas, uma vez que normaliza os valores para a média zero e desvio padrão unitário.

Pré-processamento

- Além do caso de valores de diferentes *features*, algumas informações podem se tornar ausentes durante o processo de aquisição de dados, seja por falha do equipamento ou informação que não existia.
- Nesse caso, os valores são tidos como *not a number* (NAN), no qual nenhuma informação pode ser inferida nesse caso. Por esse motivo, o tratamento desses dados deve ser aplicado antes do treinamento dos modelos.
- Alguns tratamentos podem ser aplicados nesse caso, como a remoção da amostra que apresenta esse problema ou introduzir o valor médio das amostras no dado ausente.

Álgebra Linear

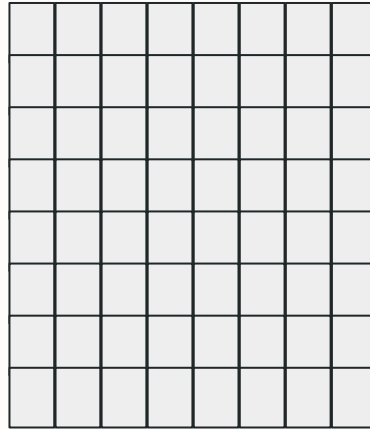
Escalar $x \in \mathbb{R}$



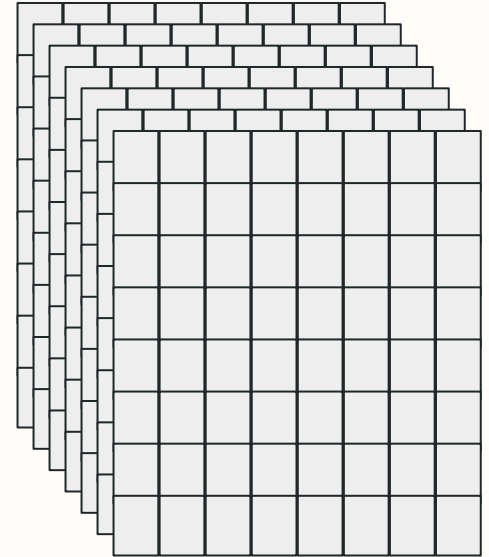
Vetor $x : \mathbb{R} \rightarrow \mathbb{R}$



Matriz $x : \mathbb{R}^2 \rightarrow \mathbb{R}$



Tensor $x : \mathbb{R}^n \rightarrow \mathbb{R}$



Álgebra Linear

- Soma ou multiplicação de um escalar pelo tensor (vetor ou matriz) resulta em uma saída com a mesma dimensão da original.

$$a \rightarrow \text{escalar} ; B = \begin{bmatrix} b_{11} & \dots & b_{1n} \\ \vdots & \ddots & \vdots \\ b_{m1} & \dots & b_{mn} \end{bmatrix}$$

$$a + B = \begin{bmatrix} a + b_{11} & \dots & a + b_{1n} \\ \vdots & \ddots & \vdots \\ a + b_{m1} & \dots & a + b_{mn} \end{bmatrix}$$

$$a \cdot B = \begin{bmatrix} a \cdot b_{11} & \dots & a \cdot b_{1n} \\ \vdots & \ddots & \vdots \\ a \cdot b_{m1} & \dots & a \cdot b_{mn} \end{bmatrix}$$

Álgebra Linear

- A realização de operações elemento por elemento gera uma saída de mesma dimensão dos tensores de entrada.
- Note que o produto elemento a elemento, gera um operador conhecido com **produto da matriz de Hadamard**

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}$$

$$A + B = \begin{bmatrix} a_{11} + b_{11} & a_{12} + b_{12} & \dots & a_{1n} + b_{1n} \\ a_{21} + b_{21} & a_{22} + b_{22} & \dots & a_{2n} + b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} + b_{m1} & a_{m2} + b_{m2} & \dots & a_{mn} + b_{mn} \end{bmatrix}$$

$$B = \begin{bmatrix} b_{11} & b_{12} & \dots & b_{1n} \\ b_{21} & b_{22} & \dots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{m1} & b_{m2} & \dots & b_{mn} \end{bmatrix}$$

$$A \odot B = \begin{bmatrix} a_{11} b_{11} & a_{12} b_{12} & \dots & a_{1n} b_{1n} \\ a_{21} b_{21} & a_{22} b_{22} & \dots & a_{2n} b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} b_{m1} & a_{m2} b_{m2} & \dots & a_{mn} b_{mn} \end{bmatrix}$$

Álgebra Linear

- Produto vetorial

$$x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \text{ e } y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$

$$x \cdot y = \sum_{i=1}^n x_i y_i$$

Álgebra Linear

- Multiplicação de Matrizes.

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1k} \\ a_{21} & a_{22} & \dots & a_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nk} \end{bmatrix} = \begin{bmatrix} a_1^\top \\ a_2^\top \\ \vdots \\ a_n^\top \end{bmatrix}$$

$$B = \begin{bmatrix} b_{11} & b_{12} & \dots & b_{1m} \\ b_{21} & b_{22} & \dots & b_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ b_{k1} & b_{n2} & \dots & b_{nm} \end{bmatrix} = [b_1 \quad b_2 \quad \dots \quad b_m]$$

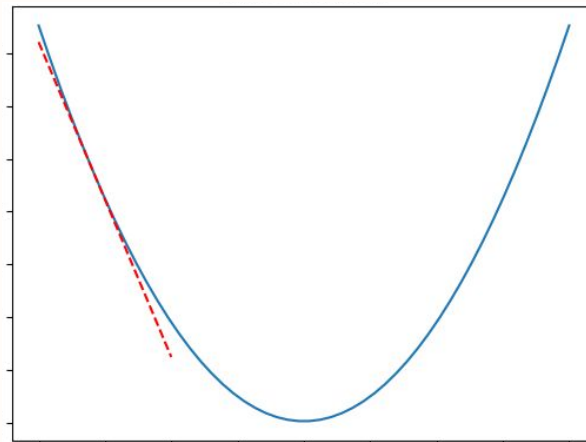
$$C = AB = \begin{bmatrix} a_1^\top \\ a_2^\top \\ \vdots \\ a_n^\top \end{bmatrix} [b_1 \quad b_2 \quad \dots \quad b_m]$$

Cálculo

- Cálculo
 - Derivada

$$f : \mathbb{R} \rightarrow \mathbb{R}$$

$$f'(x) = \lim_{h \rightarrow 0} \frac{f(x+h) - f(x)}{h}$$



Conceitos Preliminares

- Cálculo
 - Derivada parcial e gradientes

$$f : \mathbb{R}^n \rightarrow \mathbb{R}$$

$$\frac{\partial f(x)}{\partial x_i} = \lim_{h \rightarrow 0} \frac{f(x_1, \dots, x_{i-1}, x_i + h, x_{i+1}, \dots, x_n) - f(x_1, \dots, x_{i-1}, x_i, x_{i+1}, \dots, x_n)}{h}$$

$$\nabla_x f(X) = [\partial_{x_1} f(x), \partial_{x_2} f(x), \dots, \partial_{x_n} f(x)]$$

Conceitos Preliminares

- Probabilidade e Estatística

- A probabilidade é um valor que determina a possibilidade de um evento acontecer, quanto maior esse valor, maior as chances de ocorrência desse evento.
- A probabilidade de um evento acontecer é um número que se encontra no intervalo $[0, 1]$, no qual 1 (um) significa certeza do evento acontecer e 0 (zero) o evento é impossível.
- A soma de todas as probabilidade de um espaço amostral é 1 (um).

Conceitos Preliminares

- Probabilidade e Estatística

- Dados um conjunto de variáveis aleatórias independentes P , a probabilidade de ocorrer dois eventos simultâneos é o produto da probabilidade dos eventos separadamente.

$$P(X, Y) = P(X) P(Y)$$

- Teorema de Bayes é a probabilidade de um evento ocorrer, dado um conhecimento *a priori*.

$$P(Y | X) = \frac{P(X | Y) P(Y)}{P(X)}$$

$$P(Y | X) = \frac{P(X | Y) P(Y)}{\sum_a P(X | Y = a) P(Y = a)}$$