Fundamentos de Deep Learning

Dsc. Nauber Gois
Fametro
Serviço Federal de Processamento de Dados

Apresentação

Francisco Nauber Bernardo Gois

Analista aprendizado de máquina no Serviço Federal de Processamento de Dados

Doutor em Informática Aplicada

Mestre em Informática Aplicada

Especialista em desenvolvimento WEB

Dúvidas: naubergois@gmail.com

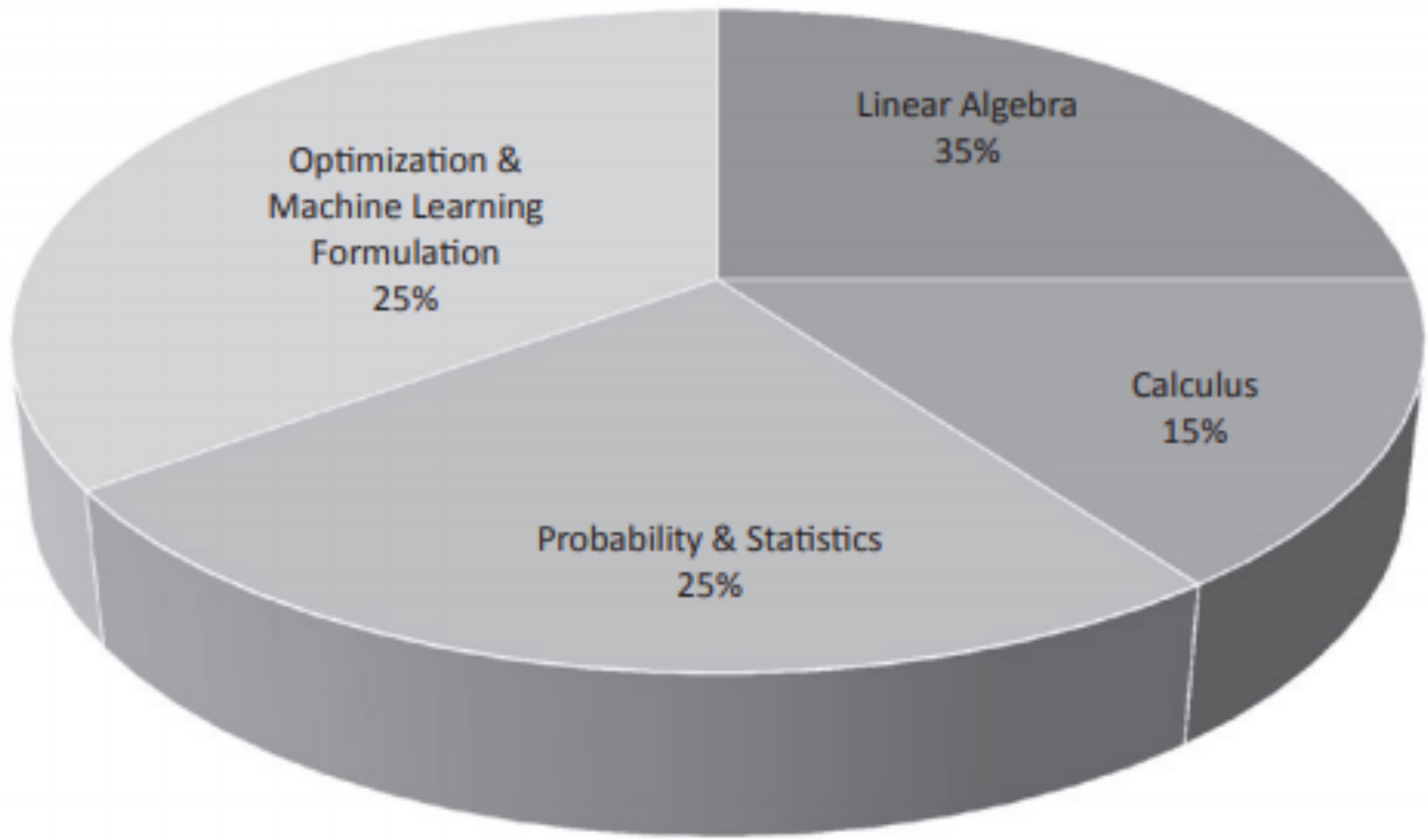

Figure 1-1. Importance of mathematics topics for machine learning and data science

Addition of Two Matrices

The addition of two matrices A and B implies their element-wise addition. We can only add two matrices, provided their dimensions match. If C is the sum of matrices A and B, then

$$c_{ij} = a_{ij} + b_{ij} \quad \forall\, i \in \{1,2,..m\}, \forall\, j \in \{1,2,..n\}$$

$$\textit{where}\; a_{ij} \in A, b_{ij} \in B, c_{ij} \in C$$

Example: $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} B = \begin{bmatrix} 5 & 6 \\ 7 & 8 \end{bmatrix}$ then $A + B = \begin{bmatrix} 1+5 & 2+6 \\ 3+7 & 4+8 \end{bmatrix} = \begin{bmatrix} 6 & 8 \\ 10 & 12 \end{bmatrix}$

Subtraction of Two Matrices

The subtraction of two matrices A and B implies their element-wise subtraction. We can only subtract two matrices provided their dimensions match.

If C is the matrix representing $A - B$, then

$$c_{ij} = a_{ij} - b_{ij} \quad \forall\, i \in \{1,2,..m\}, \forall\, j \in \{1,2,..n\}$$

$$\textit{where}\; a_{ij} \in A, b_{ij} \in B, c_{ij} \in C$$

Example: $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} B = \begin{bmatrix} 5 & 6 \\ 7 & 8 \end{bmatrix}$ then $A - B = \begin{bmatrix} 1-5 & 2-6 \\ 3-7 & 4-8 \end{bmatrix} = \begin{bmatrix} -4 & -4 \\ -4 & -4 \end{bmatrix}$

Product of Two Matrices

For two matrices $A \in \mathbb{R}^{m \times n}$ and $B \in \mathbb{R}^{p \times q}$ to be multipliable, n should be equal to p. The resulting matrix is $C \in \mathbb{R}^{m \times q}$. The elements of C can be expressed as

$$c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj} \qquad \forall\, i \in \{1,2,..m\}, \forall\, j \in \{1,2,..q\}$$

For example, the matrix multiplication of the two matrices $A, B \in \mathbb{R}^{2 \times 2}$ can be computed as seen here:

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} B = \begin{bmatrix} 5 & 6 \\ 7 & 8 \end{bmatrix}$$

$$c_{11} = \begin{bmatrix} 1 & 2 \end{bmatrix} \begin{bmatrix} 5 \\ 7 \end{bmatrix} = 1 \times 5 + 2 \times 7 = 19 \quad c_{12} = \begin{bmatrix} 1 & 2 \end{bmatrix} \begin{bmatrix} 6 \\ 8 \end{bmatrix} = 1 \times 6 + 2 \times 8 = 22$$

$$c_{21} = \begin{bmatrix} 3 & 4 \end{bmatrix} \begin{bmatrix} 5 \\ 7 \end{bmatrix} = 3 \times 5 + 4 \times 7 = 43 \quad c_{22} = \begin{bmatrix} 3 & 4 \end{bmatrix} \begin{bmatrix} 6 \\ 8 \end{bmatrix} = 3 \times 6 + 4 \times 8 = 50$$

$$C = \begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{bmatrix} = \begin{bmatrix} 19 & 22 \\ 43 & 50 \end{bmatrix}$$

Dot Product of Two Vectors

Any vector of dimension n can be represented as a matrix $v \in \mathbb{R}^{n \times 1}$. Let us denote two n dimensional vectors $v_1 \in \mathbb{R}^{n \times 1}$ and $v_2 \in \mathbb{R}^{n \times 1}$.

$$v_1 = \begin{bmatrix} v_{11} \\ v_{12} \\ . \\ . \\ . \\ v_{1n} \end{bmatrix} v_2 = \begin{bmatrix} v_{21} \\ v_{22} \\ . \\ . \\ . \\ v_{2n} \end{bmatrix}$$

The dot product of two vectors is the sum of the product of corresponding components—i.e., components along the same dimension—and can be expressed as

$$v_1.v_2 = v_1^T v_2 = v_2^T v_1 = v_{11}v_{21} + v_{12}v_{22} + .. + v_{1n}v_{2n} = \sum_{k=1}^{n} v_{1k}v_{2k}$$

Matrix Working on a Vector

When a matrix is multiplied by a vector, the result is another vector. Let's say $A \in \mathbb{R}^{m \times n}$ is multiplied by the vector $x \in \mathbb{R}^{n \times 1}$. The result would produce a vector $b \in \mathbb{R}^{m \times 1}$

$$A = \begin{bmatrix} c_1^{(1)} c_1^{(2)} \dots . c_1^{(n)} \\ c_2^{(1)} c_2^{(2)} \dots . c_2^{(n)} \\ . \\ . \\ . \\ c_m^{(1)} c_m^{(2)} \dots . c_m^{(n)} \end{bmatrix} x = \begin{bmatrix} x_1 \\ x_2 \\ . \\ . \\ . \\ x_n \end{bmatrix}$$

A consists of n column vectors $c^{(i)} \in \mathbb{R}^{m \times 1} \quad \forall i \in \{1,2,3,\dots,n\}$.

$$A = \left[c^{(1)} c^{(2)} c^{(3)} \dots . c^{(n)} \right]$$

$$b = Ax = \left[c^{(1)} c^{(2)} c^{(3)} \dots . c^{(n)} \right] \begin{bmatrix} x_1 \\ x_2 \\ . \\ . \\ . \\ x_n \end{bmatrix} = x_1 c^{(1)} + x_2 c^{(2)} + \dots . + x_n c^{(n)}$$