Computational Graphs

Đồ thị tính toán

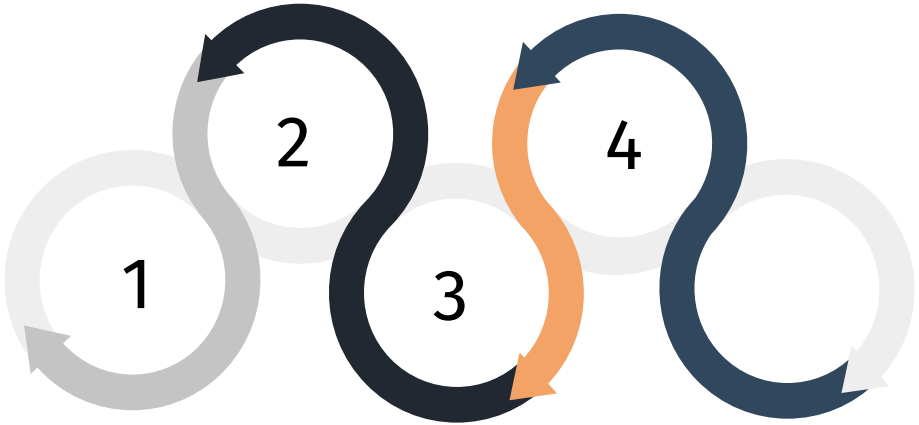
Giảng viên: Nguyễn Thanh Sơn

Nhóm 13

Lê Chí Thành - Vương Vĩnh Thuận

Nội dung

1. Dẫn nhập
2. Computational Graphs
3. Ứng dụng
4. Demo



1 Dẫn nhập

Dẫn nhập

- Đạo hàm là gì?
- Các thuật toán Machine Learning?

Dẫn nhập

- **Các thuật toán Machine Learning?**

  - Supervised Learning: regression, classification
  - Unsupervised Learning: Clustering (phân nhóm), association (sự liên kết)
  - Reinforcement Learning (Học củng cố)

LINEAR REGRESSION

Computational Graphs

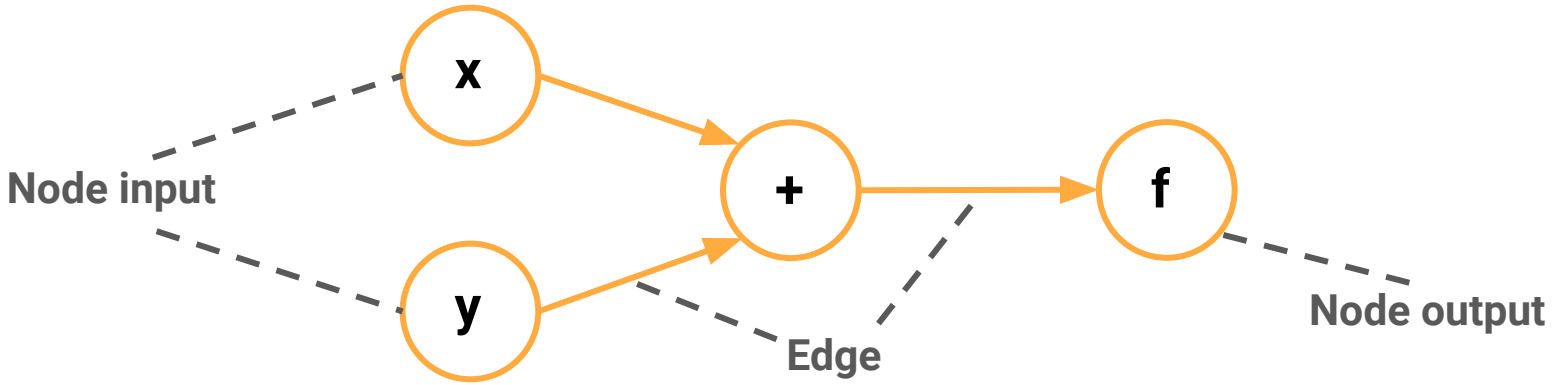
Đồ thị tính toán

- Là một loại đồ thị **có hướng** dùng để biểu diễn các biểu thức toán học.
- Là ngôn ngữ mô tả mô hình Deep Learning.

Đặc điểm

- Node (nút):
  + Đầu vào (input): vector, ma trận.
  + Phép toán (operation): cộng, trừ, nhân và chia.

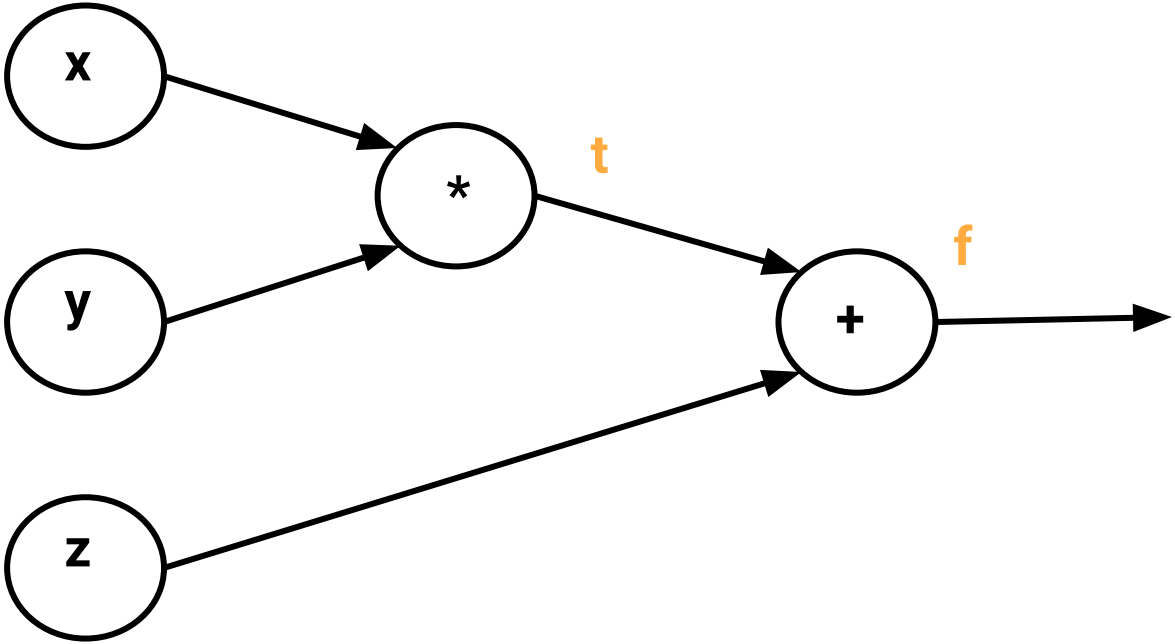
- Cạnh (edge): một đối số (dữ liệu phụ thuộc), giống một con trỏ đến node.



Đồ thị tính toán

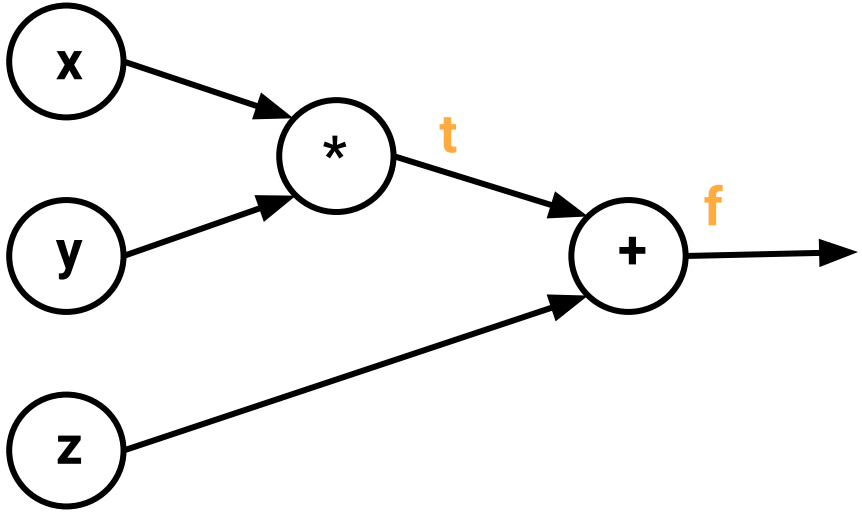
Xây dựng Computational graph cho hàm f sau:

f(x, y, z) = (x * y) + z



Đồ thị tính toán

f(x, y, z) = (x * y) + z Đặt t = x * y, f (t, z) = t + z



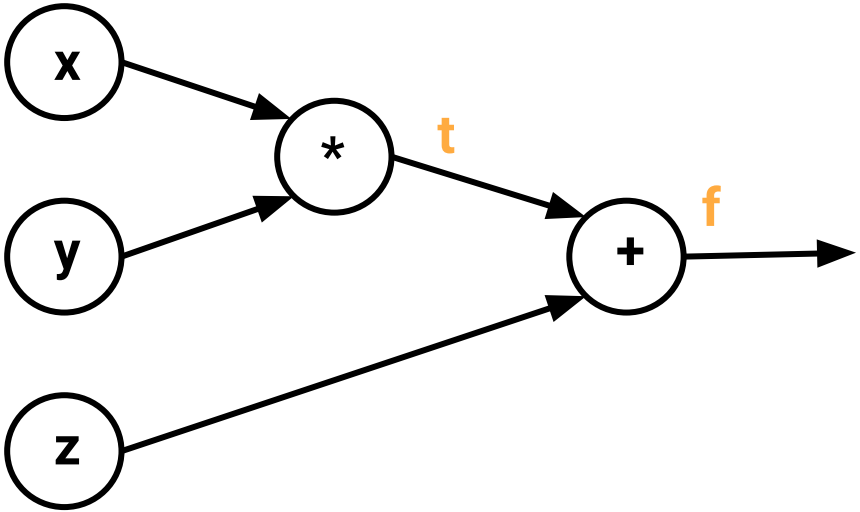
Đạo hàm riêng - Partial Derivative

$$\frac{\partial t}{\partial x} = y \qquad \frac{\partial f}{\partial z} = 1 \qquad \frac{\partial f}{\partial x} = y$$

$$\frac{\partial t}{\partial y} = x \qquad \frac{\partial f}{\partial t} = 1 \qquad \frac{\partial f}{\partial y} = x$$

Đồ thị tính toán

Tính chất



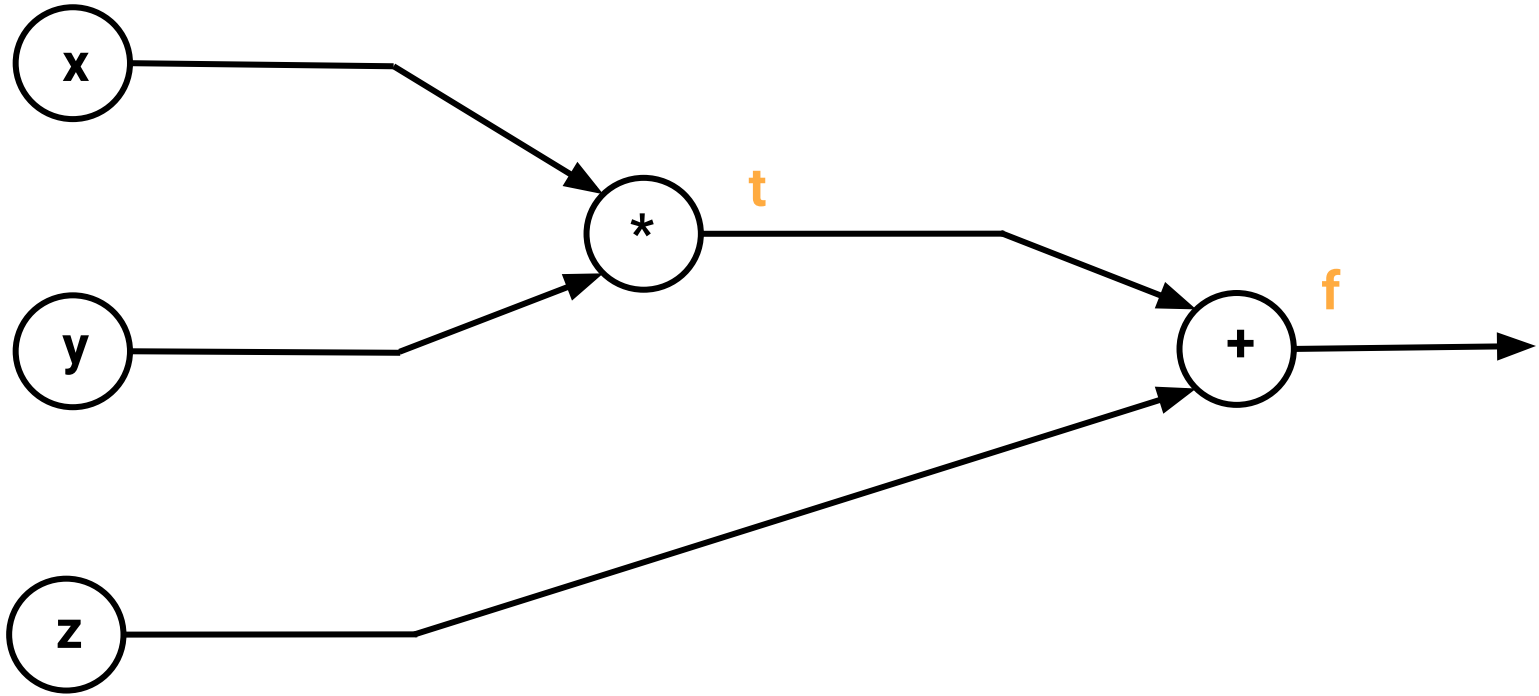
Đạo hàm riêng - Partial Derivative

$$\frac{\partial t}{\partial x} = y \qquad \frac{\partial f}{\partial z} = 1 \qquad \frac{\partial f}{\partial x} = y$$

$$\frac{\partial t}{\partial y} = x \qquad \frac{\partial f}{\partial t} = 1 \qquad \frac{\partial f}{\partial y} = x$$

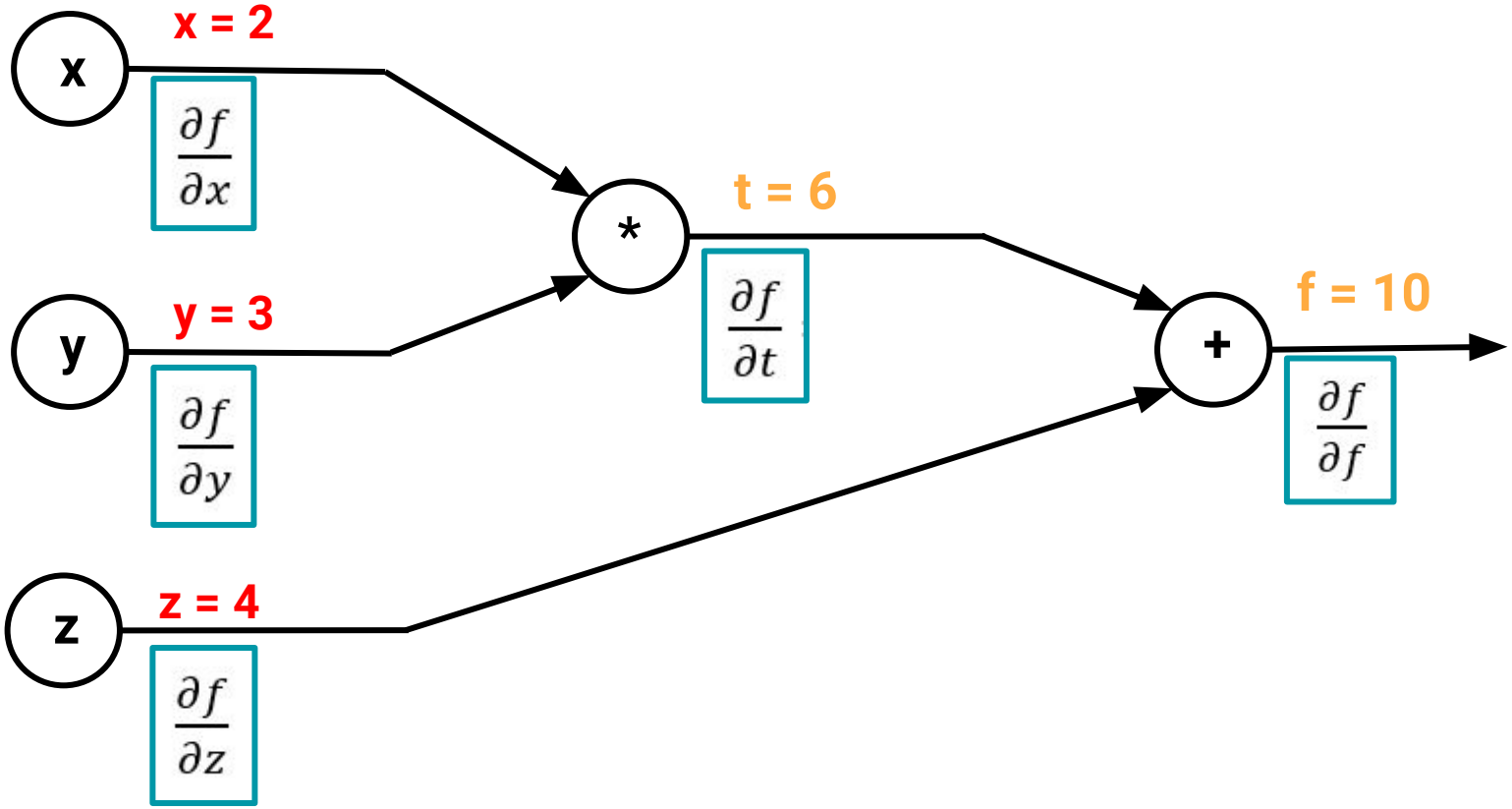
- ***Đạo hàm của biểu thức cộng tại 1 biến = 1***
- ***Đạo hàm của biểu thức nhân theo biến này = giá trị của biến còn lại***

Đồ thị tính toán



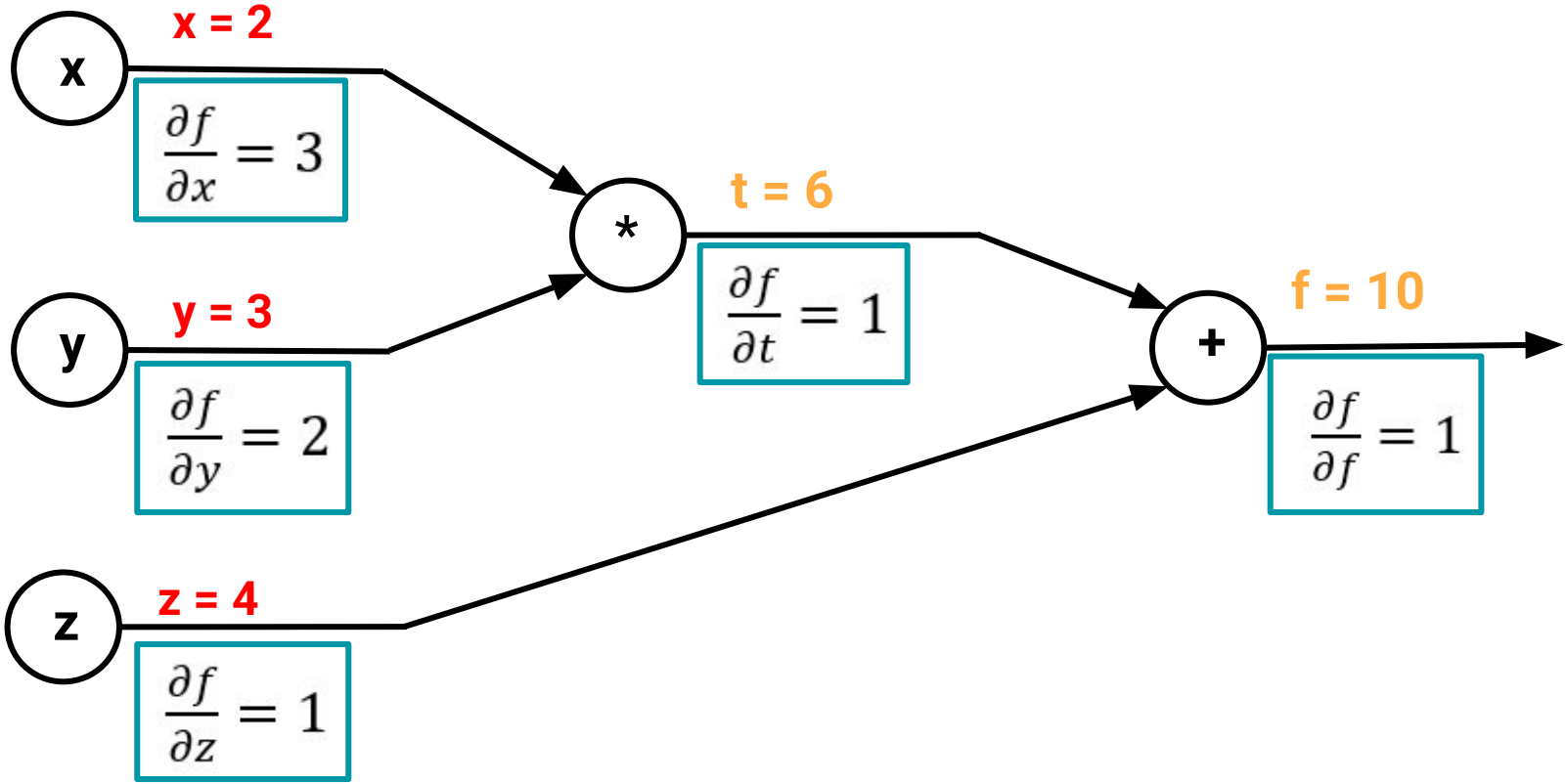
Đồ thị tính toán

Tính f(x, y, z) = (x * y) + z với **x = 2, y = 3, z = 4**

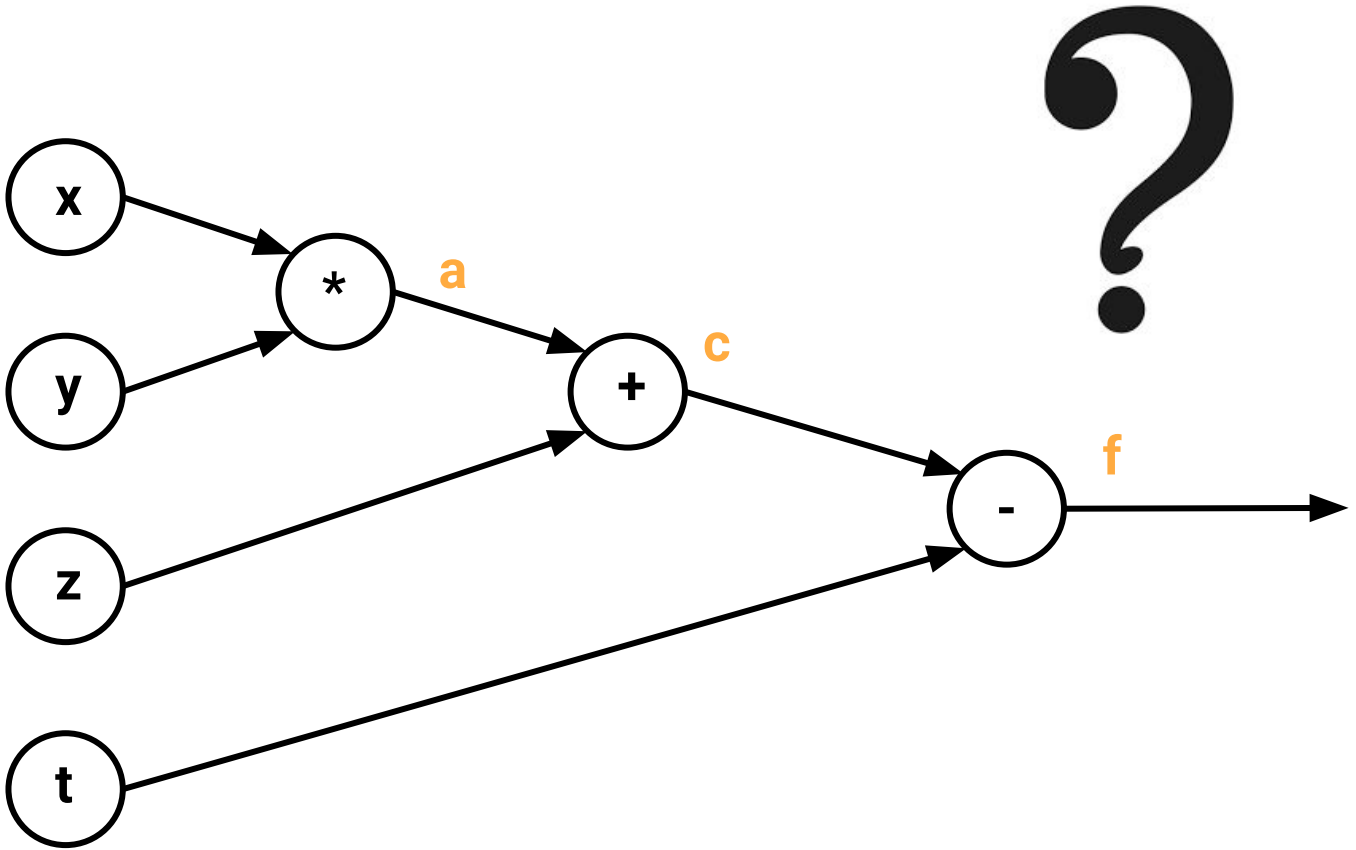


Đồ thị tính toán

Tính f(x, y, z) = (x * y) + z với **x = 2, y = 3, z = 4**



Đồ thị tính toán



Đồ thị tính toán

Quy tắc dây chuyền - Chain rule

$$\frac{\partial f}{\partial x} = \frac{\partial f}{\partial c}\frac{\partial c}{\partial a}\frac{\partial a}{\partial x} = 1.1.y = y$$

$$\frac{\partial f}{\partial y} = \frac{\partial f}{\partial c}\frac{\partial c}{\partial a}\frac{\partial a}{\partial y} = 1.1.x = x$$

$$\frac{\partial f}{\partial z} = \frac{\partial f}{\partial c}\frac{\partial c}{\partial z} = 1.1 = 1$$

$$\frac{\partial f}{\partial t} = -1$$

Computational graphs giúp tính đạo hàm từng vị trí (node)

Đồ thị tính toán

Dự đoán cân nặng

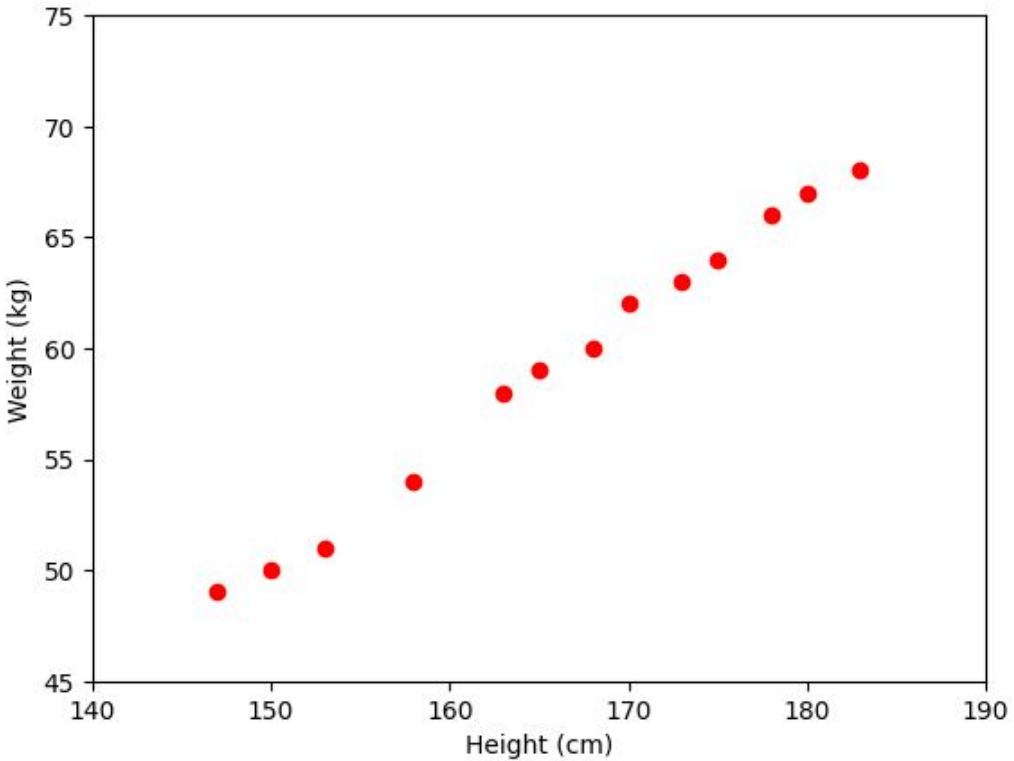
Chiều cao (cm)	Cân nặng (kg)	Chiều cao (cm)	Cân nặng (kg)
147	49	168	60
150	50	170	72
153	51	173	63
155	52	175	64
158	54	178	66
160	56	180	67
163	58	183	68
165	59		

Đồ thị tính toán

Dự đoán cân nặng

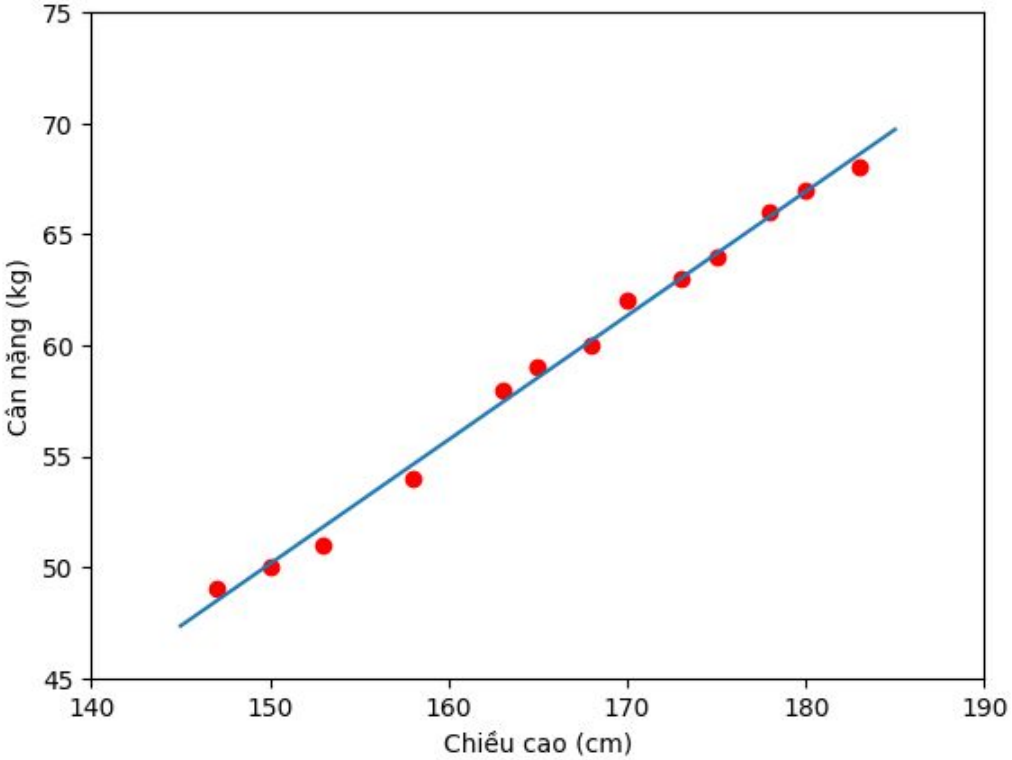
Chiều cao (cm)	Cân nặng (kg)	Chiều cao (cm)	Cân nặng (kg)
147	49	168	60
150	50	170	72
153	51	173	63
155	52	175	64
158	54	178	66
160	56	180	67
163	58	183	68
165	59		

Nguồn: Bảng dữ liệu về chiều cao và cân năng



Đồ thị tính toán

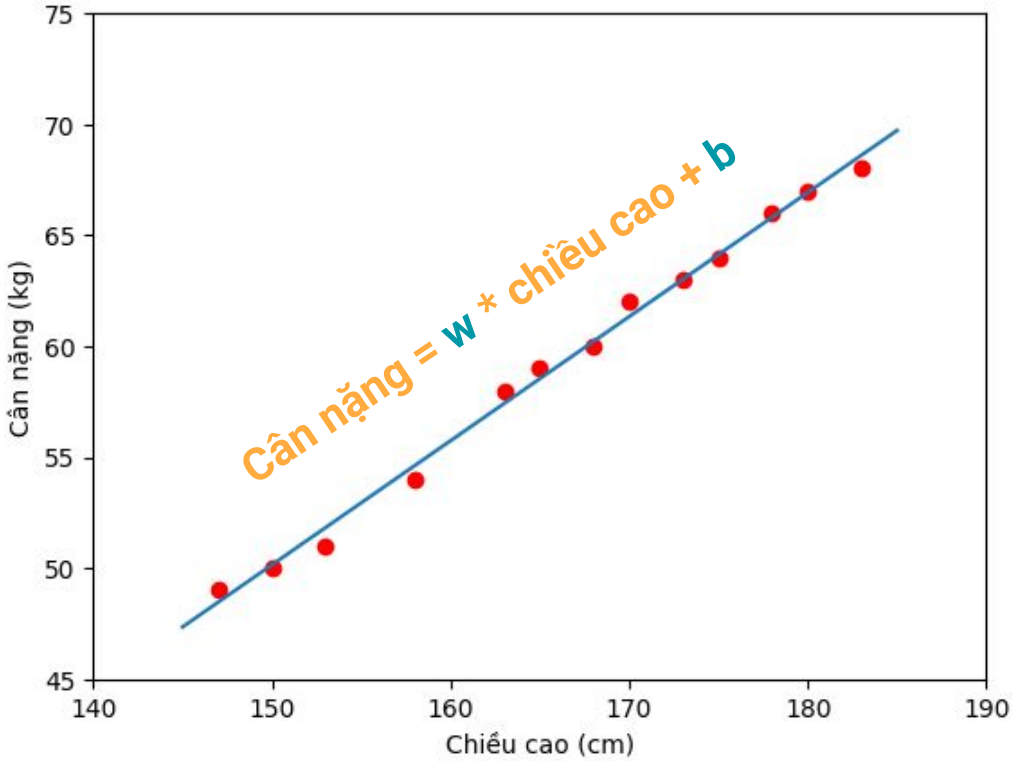
Dự đoán cân nặng



Đồ thị tính toán

Dự đoán cân nặng

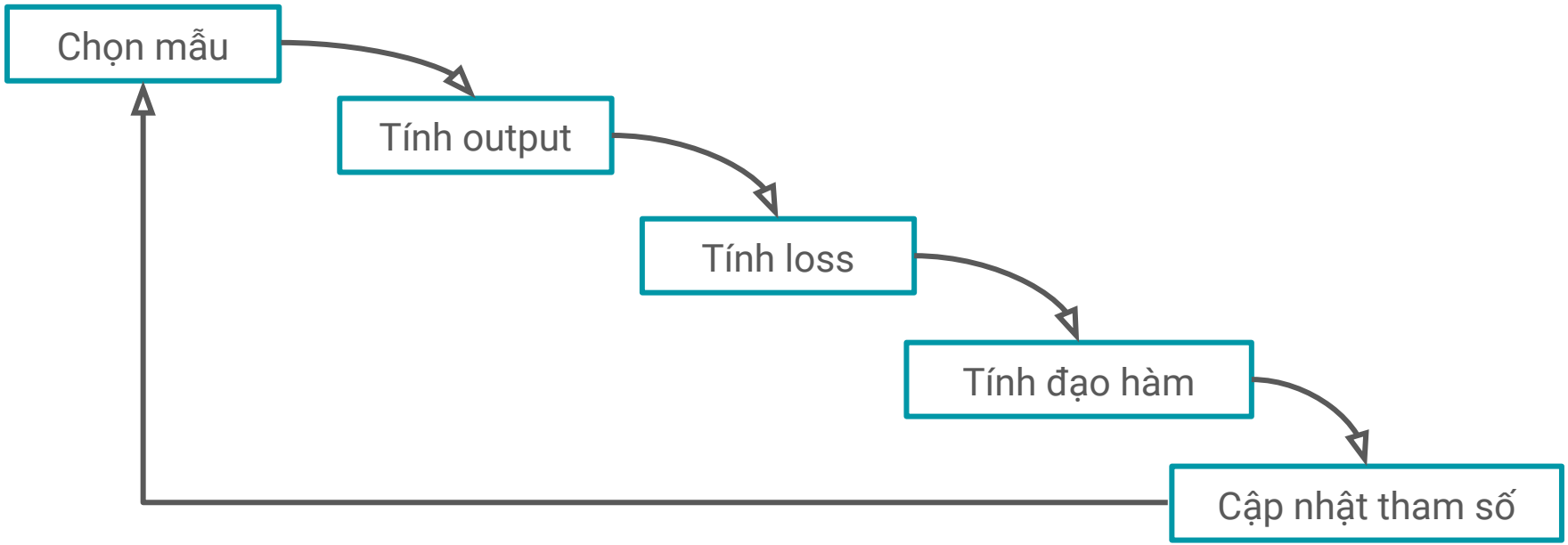
1 người cao **155cm** sẽ nặng bao nhiêu kg?



Đồ thị tính toán

Dự đoán cân nặng

One-sample training



Đồ thị tính toán

Dự đoán cân nặng

One-sample training

1) Chọn mẫu

2) Tính output o

$$o = wx + b$$

3) Tính loss

$$L = (o - y)^2$$

4) Tính đạo hàm

$$L'_w = 2x(o - y)$$
$$L'_b = 2(o - y)$$

5) Cập nhật tham số

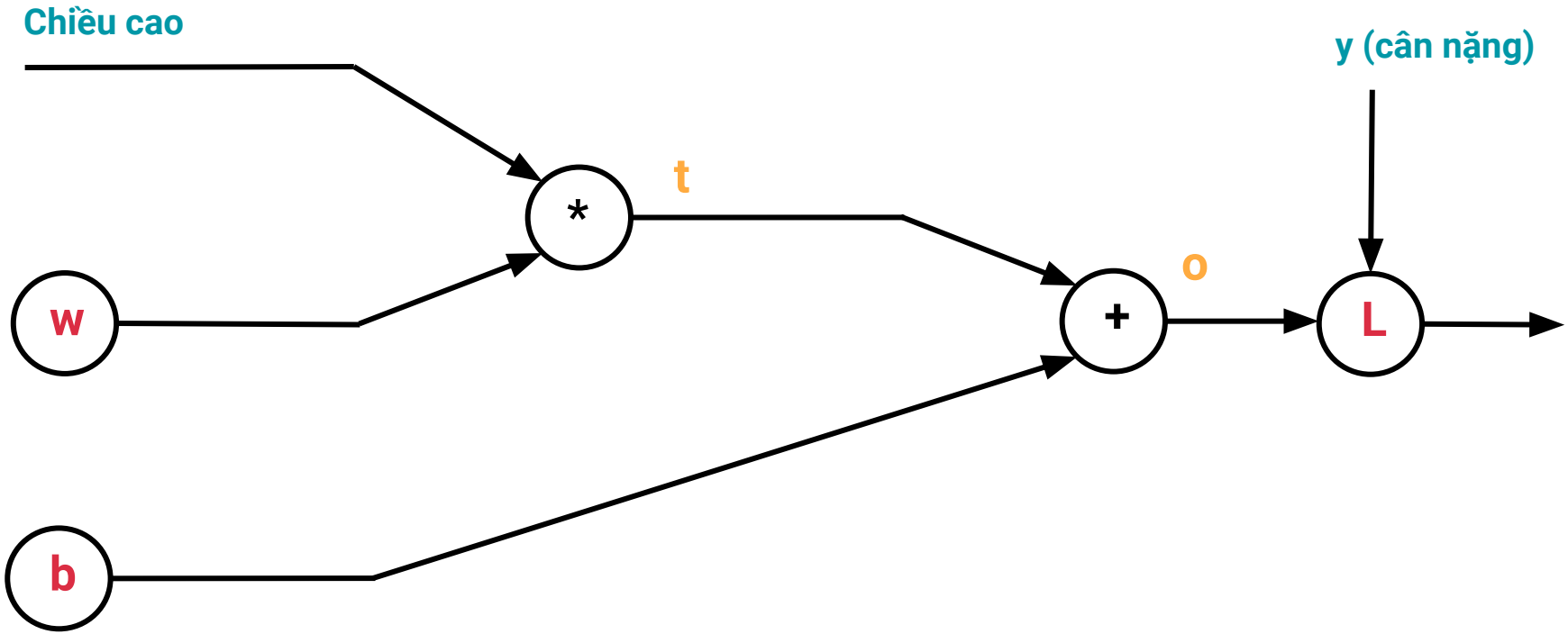
$$w = w - \eta L'_w$$
$$b = b - \eta L'_b$$

η: *learning rate*

Đồ thị tính toán

Dự đoán cân nặng

One-sample training

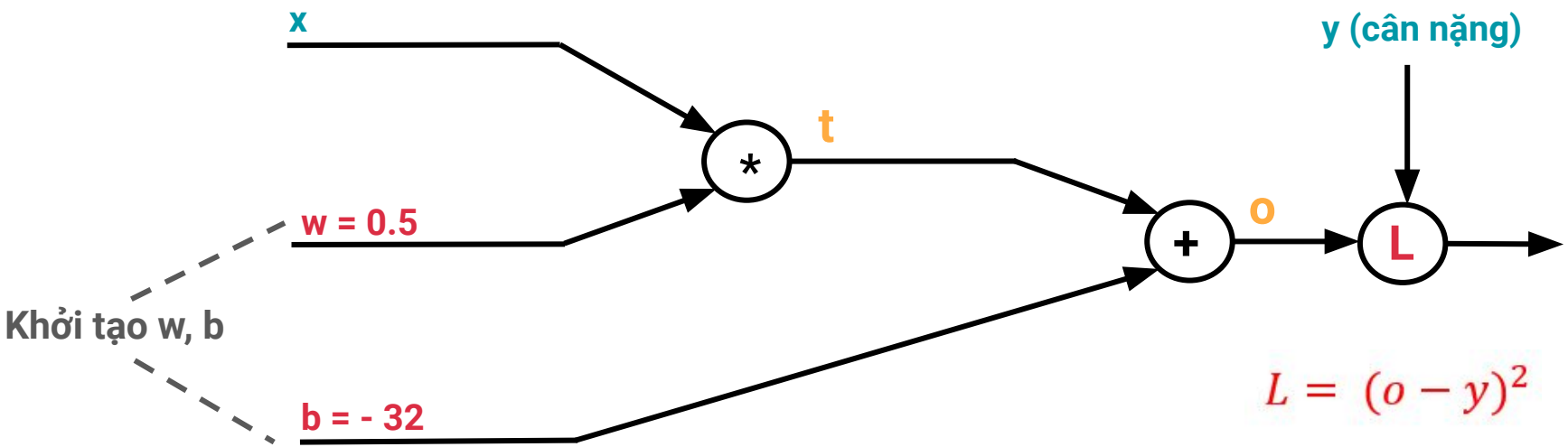


Đồ thị tính toán

Dự đoán cân nặng

One-sample training

$$cân\ nặng = o = wx + b \qquad t = w * chiều\ cao = wx \qquad o = t + b$$

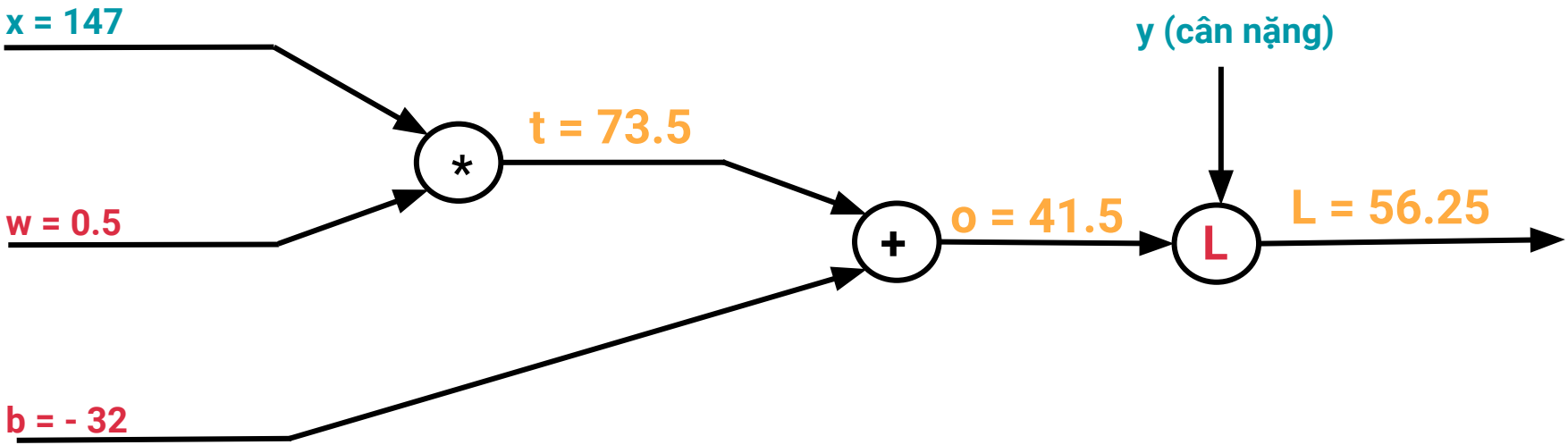


Đồ thị tính toán

Dự đoán cân nặng

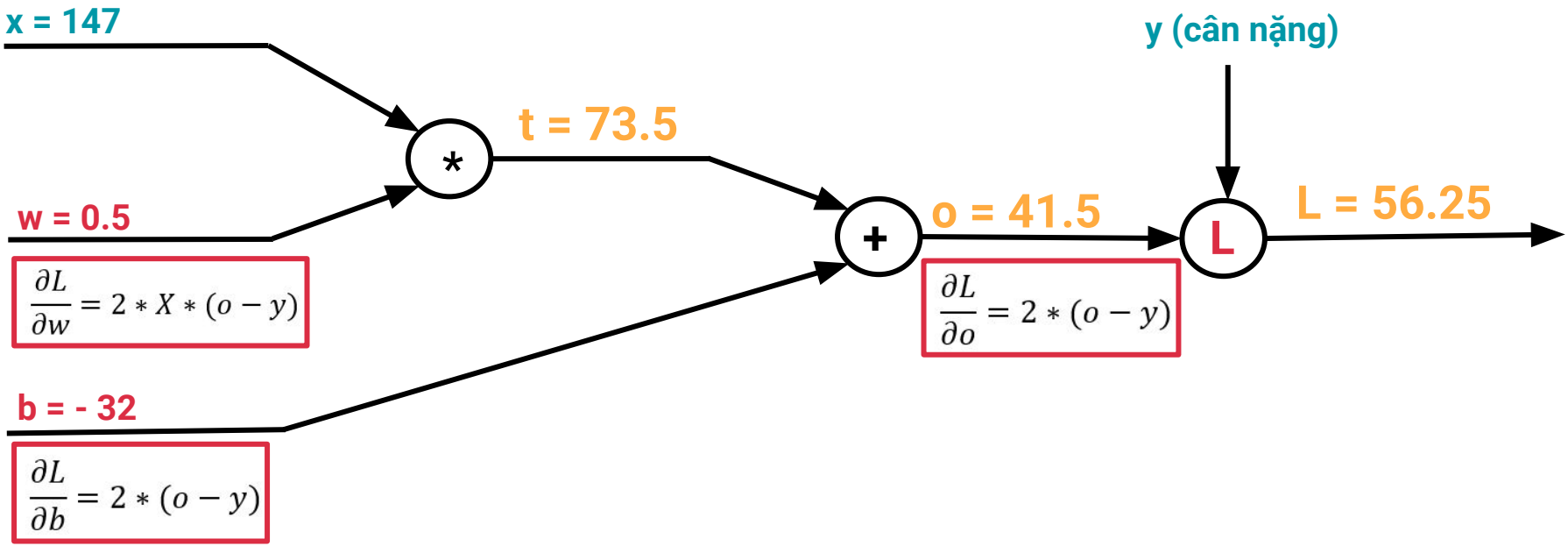
One-sample training

$$cân\ nặng = o = wx + b \qquad t = w * chiều\ cao = wx \qquad o = t + b$$



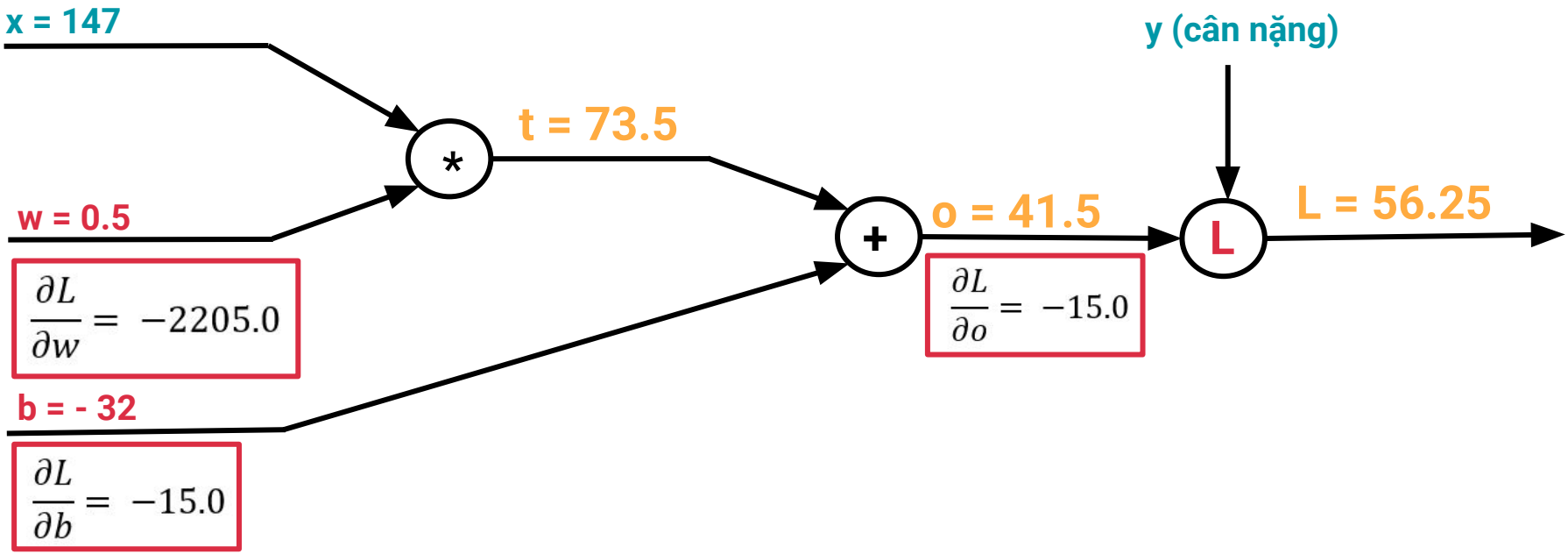
Đồ thị tính toán

Cơ chế đi ngược - Backpropagation



Đồ thị tính toán

Cơ chế đi ngược - Backpropagation



Đồ thị tính toán

Cơ chế đi ngược - Backpropagation

Cập nhật w và b

$$w = w - \eta L'_w$$

$$b = b - \eta L'_b$$

$$\eta = 0.00001$$

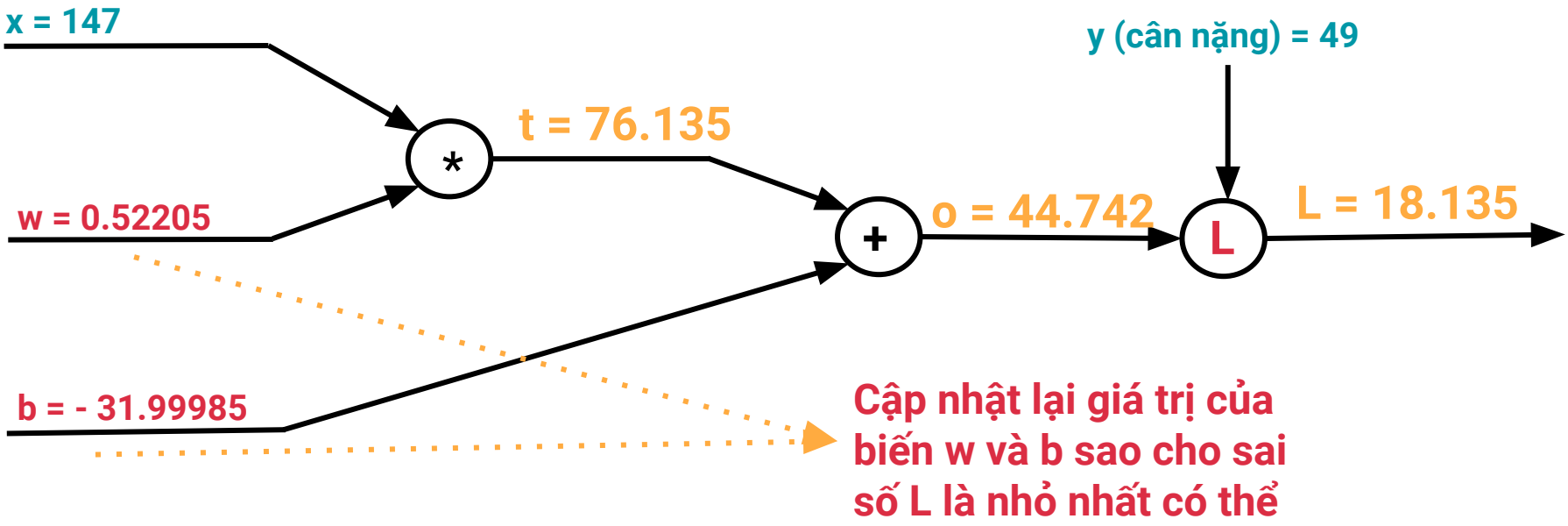
$$w_{cập\ nhật} = w - \eta . L'_w = 0.5 - 0.00001 * (-2205.0) = 0.52205$$

$$b_{cập\ nhật} = b - \eta . L'_b = -32 - 0.00001 * (-15.0) = -31.99985$$

Đồ thị tính toán

Cơ chế đi ngược - Backpropagation

Cập nhật w và b



Đồ thị tính toán

Đồ thị tính toán có phải là đồ thị vô hướng?

Đồ thị tính toán

Nếu có nhiều hơn 1 sample thì sao?

Đồ thị tính toán

Công thức đối với dataset có N sample

$$\bar{Y} = \begin{bmatrix} \bar{y}^{(1)} \\ \vdots \\ \bar{y}^{(N)} \end{bmatrix} = (\sum_j w_j X_j) + b = (\sum_j w_j \begin{bmatrix} x^{(1)} \\ \vdots \\ x^{(N)} \end{bmatrix}) + b$$

$$\bar{y}^{(i)} = (\sum_j w_j x^{(i)}) + b$$

$$L = \frac{1}{N} \sum_i (\bar{y}^{(i)} - y^{(i)})^2$$

$$\begin{cases} L'_{w_j} = \frac{1}{N} \sum_i 2x^{(i)} (\bar{y}^{(i)} - y^{(i)}) \\ L'_b = \frac{1}{N} \sum_i 2(\bar{y}^{(i)} - y^{(i)}) \end{cases}$$

$$w_j = w_j - \eta L'_{w_j}, \qquad b = b - \eta L'_b$$

$$for\ 0 < j \leq m, \qquad 0 < i \leq N$$

with m is the number of var w, N is the number of samples

Đồ thị tính toán

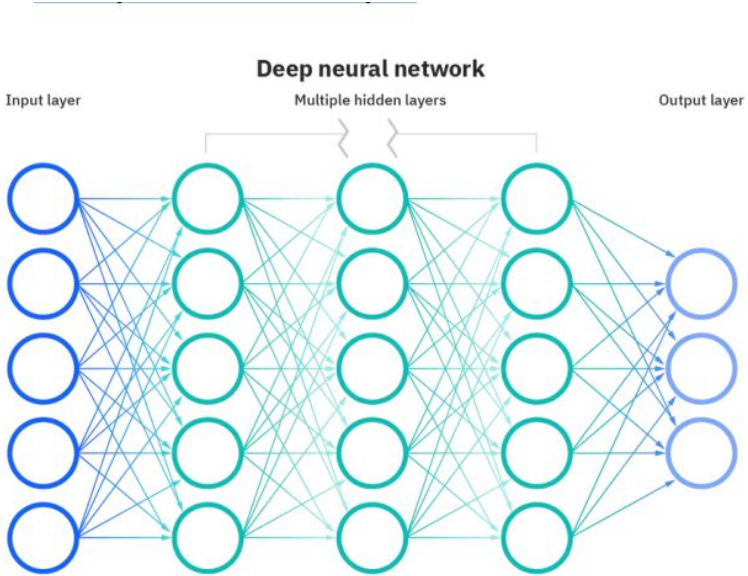
Nếu có nhiều hơn 1 đặc trưng?

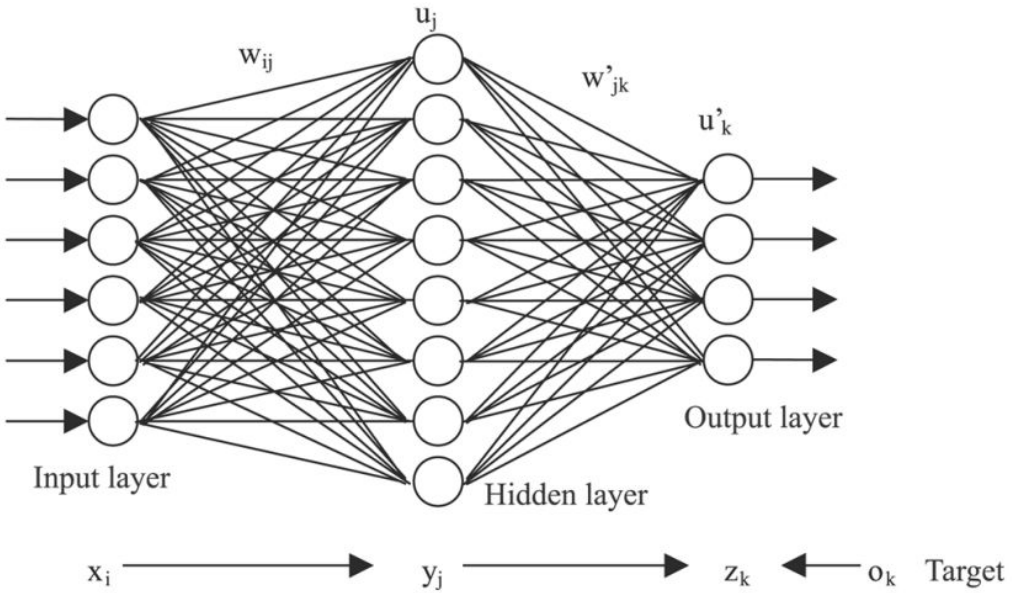
3

Ứng dụng

Ứng dụng

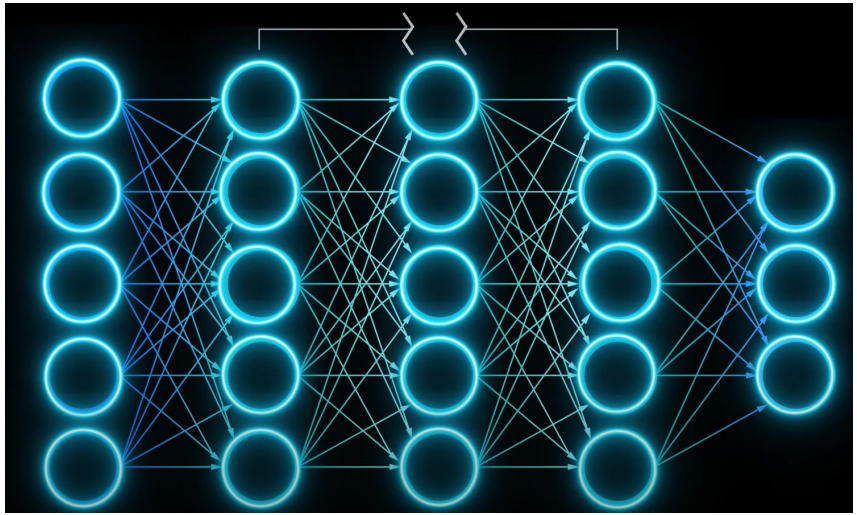
- Giải thích cách tổ chức tính toán trong Neural Network (theo forward pass hay backpropagation để tính toán được output của Network).
- Neural Network là một dạng đặc biệt của Computational Graph.





Ứng dụng

- Là công cụ mạnh để đạo hàm theo từng biến.
- Trong các bài toán network nhiều lớp, cần phải tính đạo hàm theo từng biến. Để làm việc này hiệu quả hơn, người ta sử dụng thuật toán lan truyền ngược (backpropagation).
- **Parallelization (tính toán song song)**



DEMO

Cảm ơn mọi người đã lắng nghe