deeplearning.ai

Learning from multiple tasks

Multi-task learning

Simplified autonomous driving example

Pedestrians

Cars

Stop Signs

Traffic lights

$y^{(i)}$

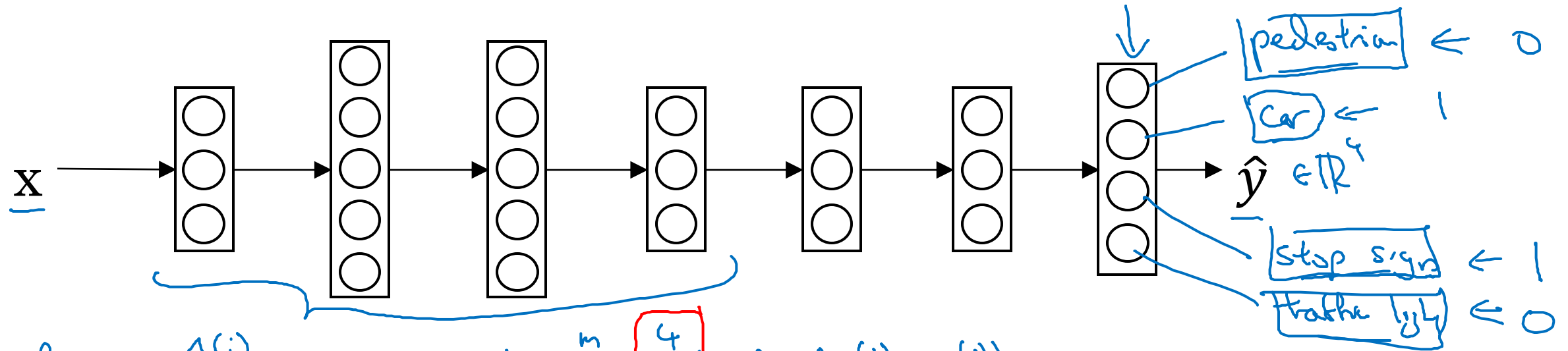$\begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \end{bmatrix}$

$(4, 1)$

$$Y = \begin{bmatrix} y^{(1)} & y^{(2)} & y^{(3)} & \cdots, & y^{(m)} \end{bmatrix}$$

$(4, m)$

$x^{(i)}$

Andrew Ng

Neural network architecture

x →

$\hat{y} \in \mathbb{R}^4$

pedestrian ← 0
Car ← 1
stop sign ← 1
traffic light ← 0

Loss: $\overset{\hat{y}^{(i)}}{\underset{(4,1)}{}}$

$$\to \frac{1}{m} \sum_{i=1}^{m} \boxed{\sum_{j=1}^{4}} L\left(\hat{y}_j^{(i)}, y_j^{(i)}\right)$$

Sum only over value of j with $0/1$ label.

↳ Usual logistic loss
$-y_j^{(i)} \log \hat{y}_j^{(i)} - (1-y_j^{(i)}) \log(1-\hat{y}_j^{(i)})$

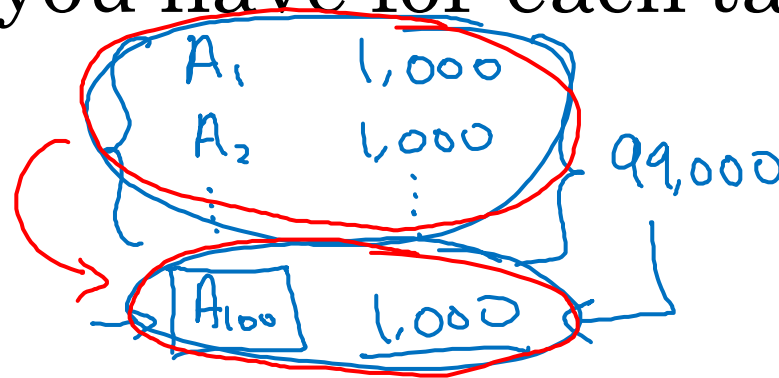Unlike softmax regression:
One image can have multiple labels

Multi-task learning ←

$$Y = \begin{bmatrix} 1 & 1 & 0 & ? \\ 0 & 1 & 1 & ? & \cdots \\ ? & ? & 1 & ? & \cdots \\ ? & ? & 0 & ? \end{bmatrix} \Leftarrow$$

Andrew Ng

When multi-task learning makes sense

- Training on a set of tasks that could benefit from having shared lower-level features.
- Usually: Amount of data you have for each task is quite similar.

$$A \quad 1,000,000$$

$$\downarrow \qquad \downarrow$$

$$B \quad 1,000$$

$$A_1 \quad 1,000$$
$$A_2 \quad 1,000$$
$$\vdots \qquad \vdots \quad \Big\} \; 99,000$$

$$A_{100} \quad 1,000$$

- Can train a big enough neural network to do well on all the tasks.