



deeplearning.ai

Learning from
multiple tasks

Multi-task
learning

Simplified autonomous driving example



$x^{(i)}$

Pedestrians

Cars

Stop signs

Traffic lights

⋮

$y^{(i)}$

0

1

1

0

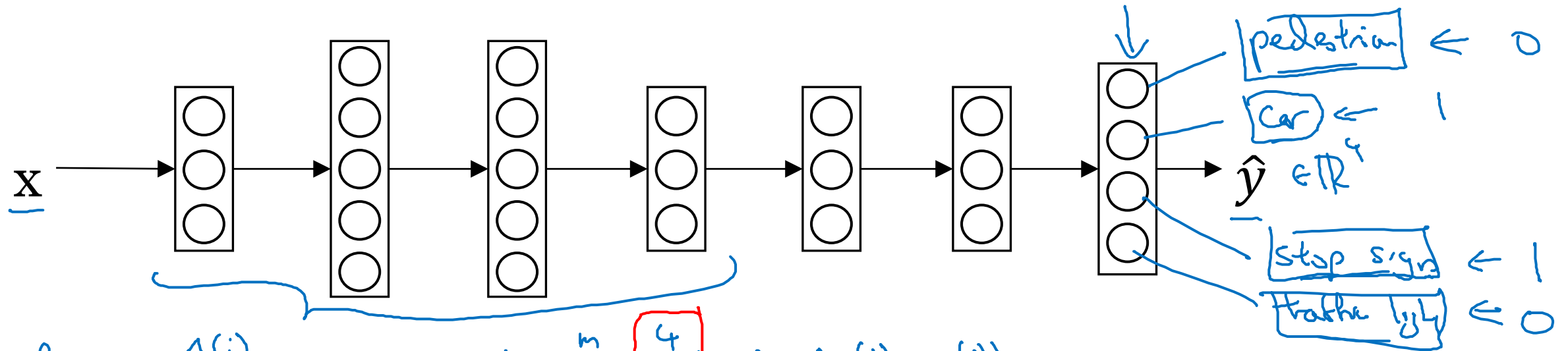
⋮

(4, 1)

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

(4, m)

Neural network architecture



Loss: $\frac{1}{m} \sum_{i=1}^m \sum_{j=1}^4 \mathcal{L}(\hat{y}_j^{(i)}, y_j^{(i)})$

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

Unlike softmax regression:
One image can have multiple labels

Usual logistic loss

$$-y_j^{(i)} \log \hat{y}_j^{(i)} - (1 - y_j^{(i)}) \log (1 - \hat{y}_j^{(i)})$$

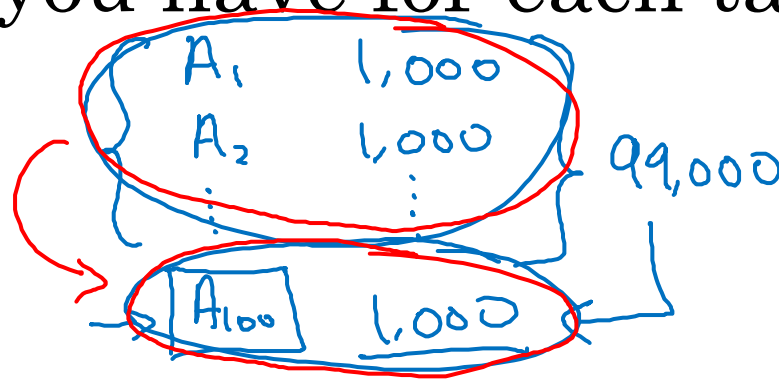
Multi-task learning \leftarrow

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

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 1,000,000
↓ ↓
B 1,000



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