



deeplearning.ai

Learning from
multiple tasks

Multi-task
learning

Simplified autonomous driving example



$x^{(i)}$

Pedestrians

Cars

Stop signs

Traffic lights

\vdots

$y^{(i)}$

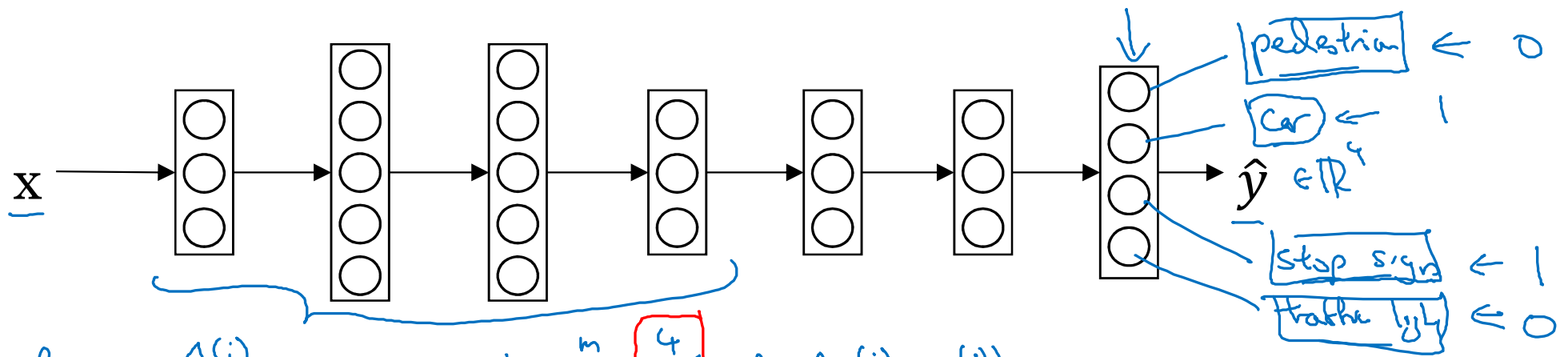
$(4, 1)$

$\begin{matrix} 0 \\ 1 \\ 1 \\ 0 \\ \vdots \end{matrix}$

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

$(4, m)$

Neural network architecture



Loss: $\mathcal{L}(\hat{y}^{(i)}, y^{(i)})$

$$\rightarrow \frac{1}{m} \sum_{i=1}^m \sum_{j=1}^4$$

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

$\mathcal{L}(\hat{y}^{(i)}, y^{(i)})$

Usual logistic loss

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

Multi-task learning

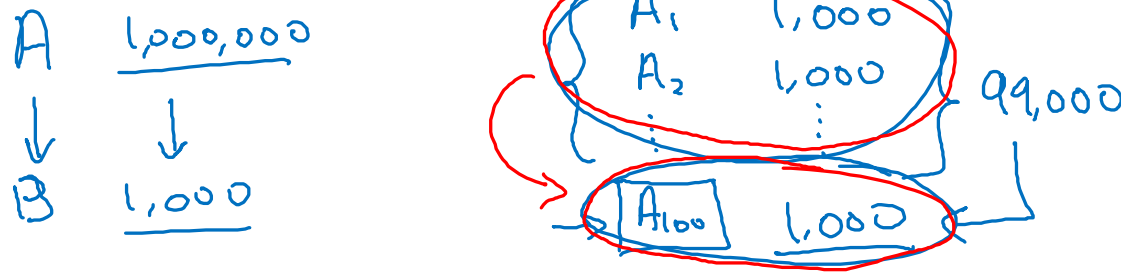
$$Y = \begin{bmatrix} 1 & 1 & 0 & ? \\ 0 & 1 & 1 & ? \\ ? & ? & 1 & ? \\ ? & ? & 0 & ? \end{bmatrix}$$

Unlike softmax regression:

One image can have multiple labels

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.



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