deeplearning.ai

Learning from multiple tasks

Multi-task learning

Simplified autonomous driving example

$x^{(i)}$

$y^{(i)}$ $(4,1)$

Pedestrians 0

Cars 1

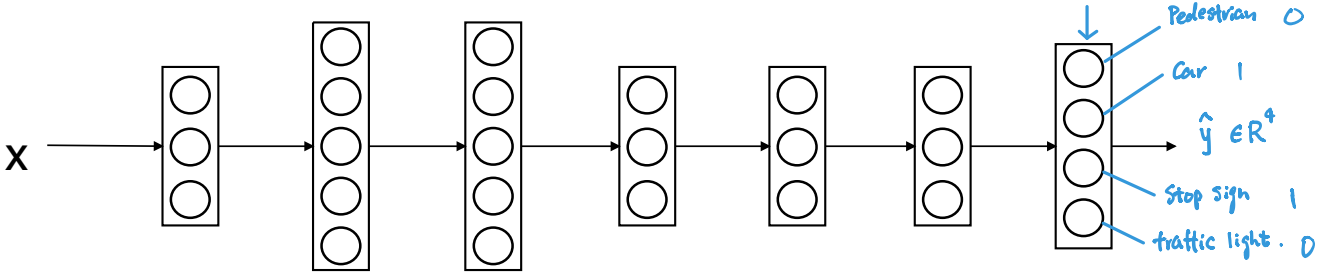
Stop signs 1

Traffic lights 0

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

$4 \times m$

Neural network architecture



Loss: $y^{(i)}$ $(4,1)$

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

Sum only over value of j with 0/1 label

Usual logistic loss

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

Unlike softmax regression:
One image can have multiple labels

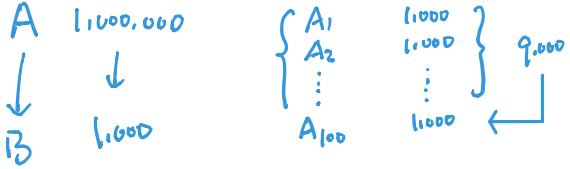
Multi-task learning

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

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.



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