



deeplearning.ai

Learning from
multiple tasks

Multi-task
learning

Simplified autonomous driving example



$x^{(i)}$

several different things를 detect해야 한다면...

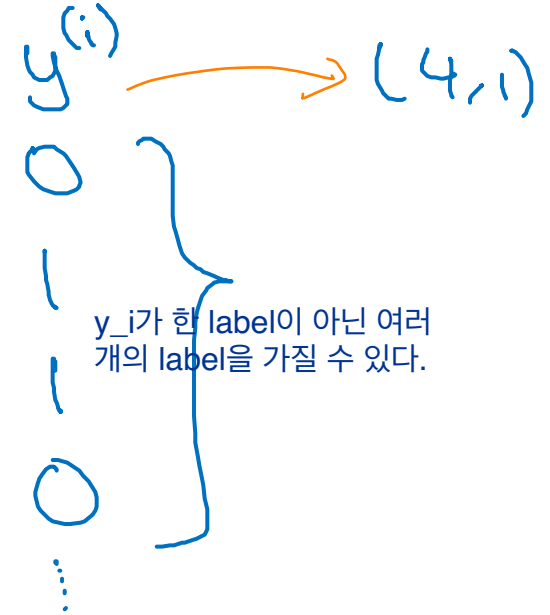
Pedestrians

Cars

Stop signs

Traffic lights

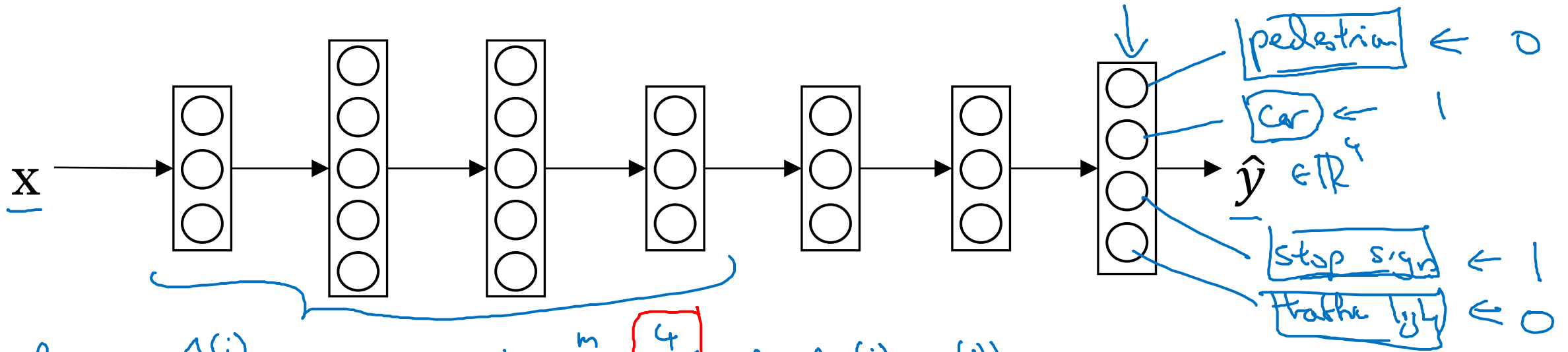
⋮



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

→ (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.

여기서 앞에서 공부한 binary classification example과 다른 점은 이 부분

Usual logistic loss \mathcal{L} 은 일반 logistic loss와 동일하고,
 $-y_j^{(i)} \log \hat{y}_j^{(i)} - (1-y_j^{(i)}) \log (1-\hat{y}_j^{(i)})$

Unlike softmax regression:
one image can have multiple labels

고로 위 cost function을 최소화하도록 NN을 학습시키면 multi-task learning을 수행하는 것이다. 왜냐하면 각 이미지에 대해 4 가지 오브젝트가 있는지에 대한 문제 4개를 풀어야 하기 때문이다.

Multi-task learning

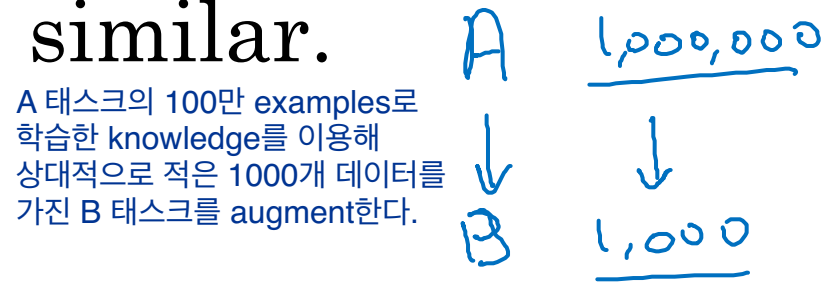
이때 4개의 개별적인 net을 학습시키는 대신에 하나의 net으로 학습하면 earlier features를 공유하게 되어 더 좋은 성능을 보이게 된다.

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

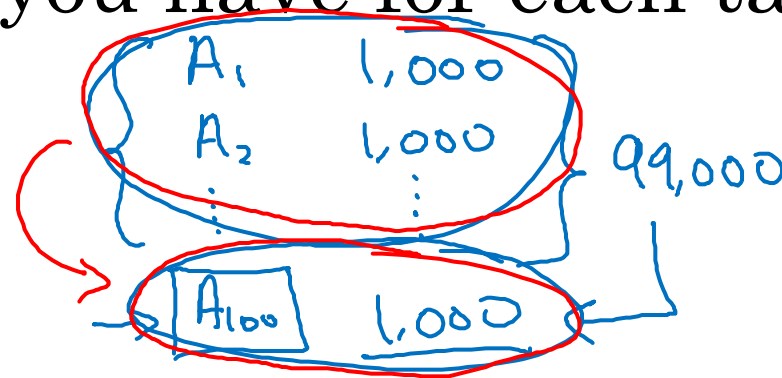
multi-task learning의 또 다른 장점은 이렇게 어떤 example들이 subset of labels만 갖고 있다고 해도 학습이 가능하다는 점이다.

When multi-task learning makes sense

- Training on a set of tasks that could benefit from having shared lower-level features. 예를 들어 앞의 예제에서는 모두 도로 이미지이기 때문에 reasonable하다.
- Usually: Amount of data you have for each task is quite similar.



transfer learning



multi-task learning

각 problem이 1000개의 examples를 가진다고 할 때, 각각을 분리시켜 생각하면 A100 태스크는 1000 개의 데이터만 가지지만 나머지 A1~99의 데이터 99,000 개 example이 big boost를 주고 knowledge to

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

네트워크가 충분히 크지 않으면 각각을 별개의 네트워크로 따로 학습시키는 것에 비해 multi-task learning이 성능이 떨어진다는 연구 결과가 있다.