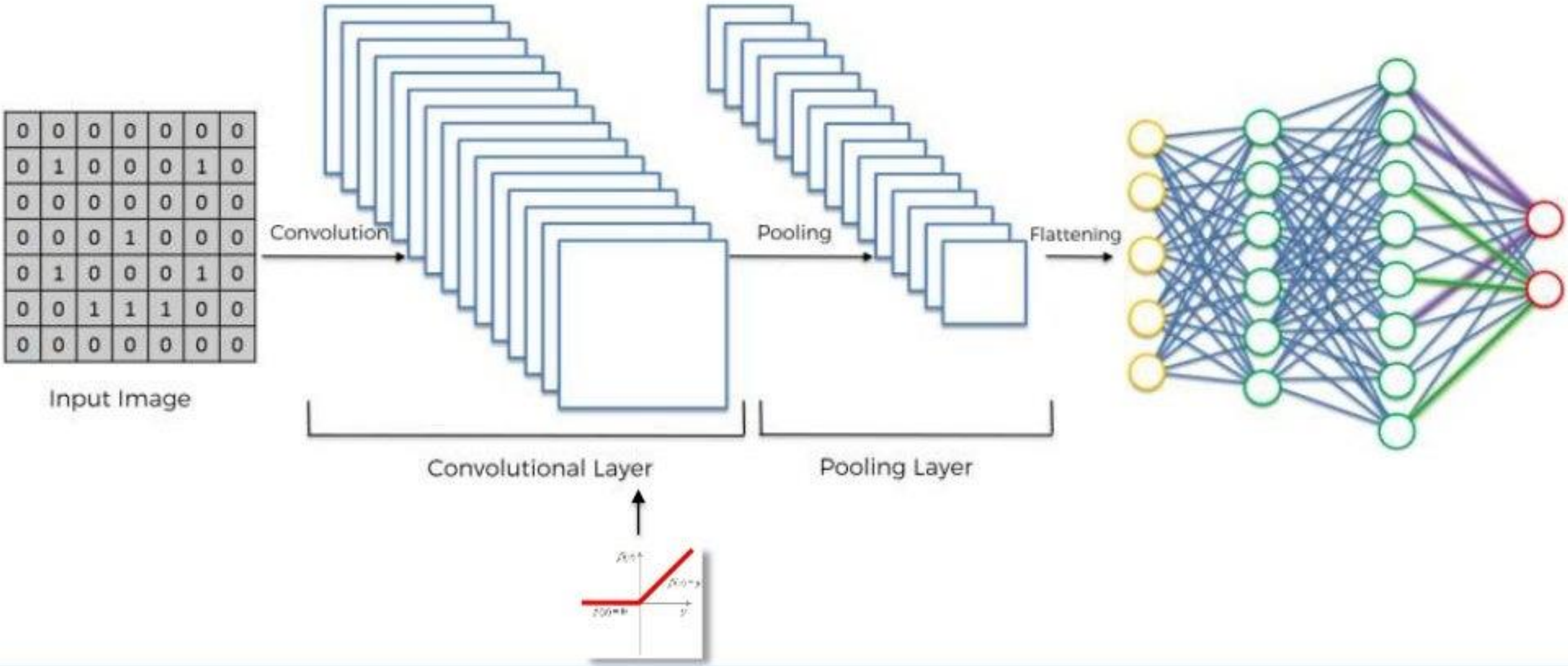
Summary

Summary



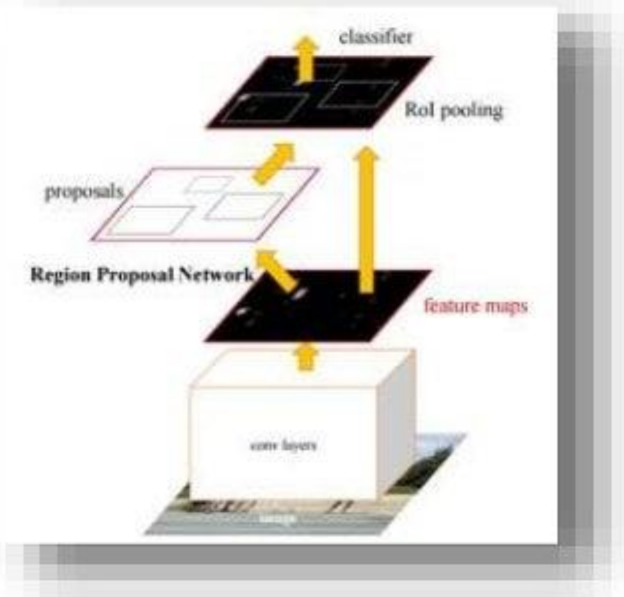
0	0	0	0	0	0	0
0	1	0	0	0	1	0
0	0	0	0	0	0	0
0	0	0	1	0	0	0
0	1	0	0	0	1	0
0	0	1	1	1	0	0
0	0	0	0	0	0	0

Summary

Additional Reading:

The 9 Deep Learning Papers You Need To Know About (Understanding CNNs Part 3)

Adit Deshpande (2016)

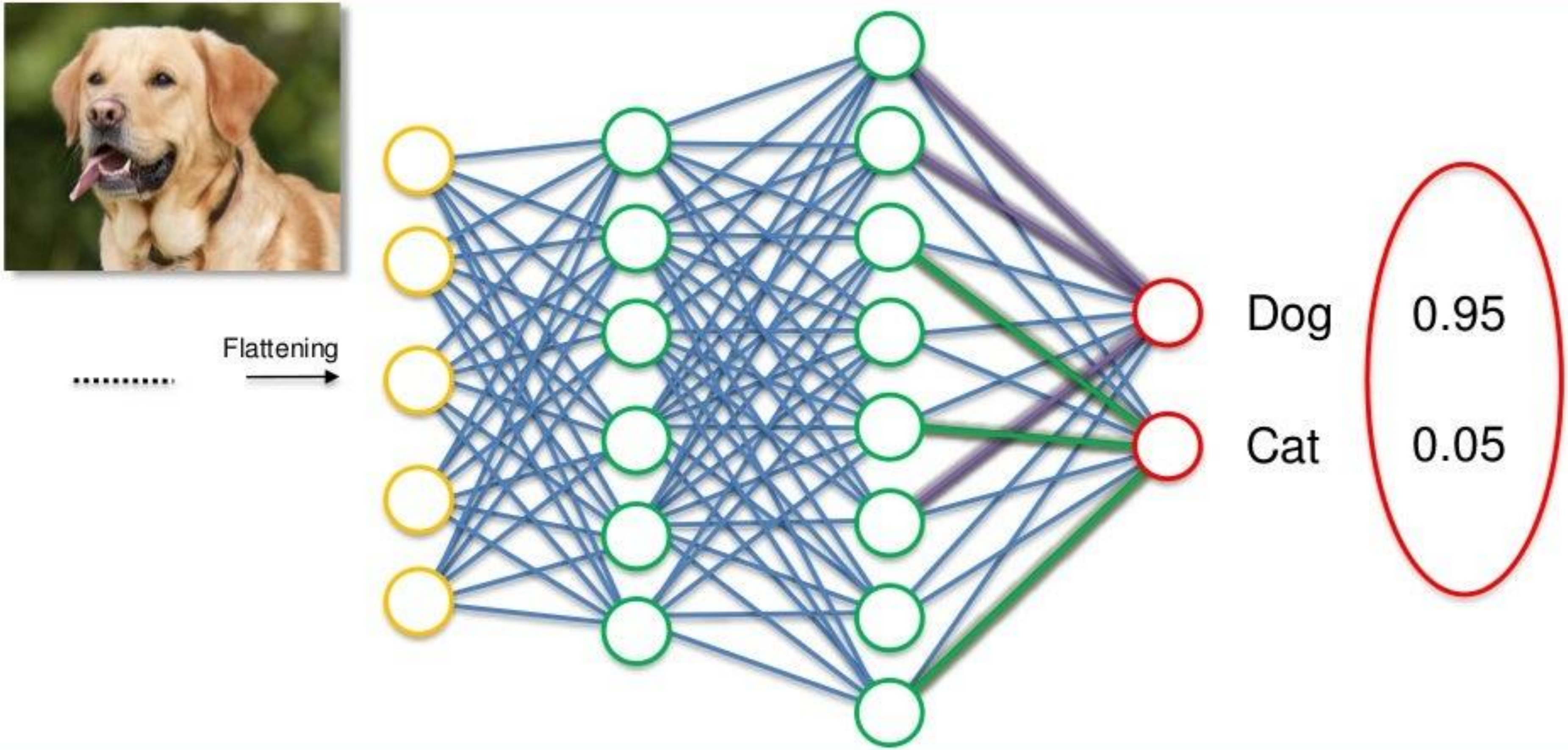


Link:

https://adeshpande3.github.io/adeshpande3.github.io/The-9-Deep-Learning-Papers-You-Need-To-Know-About.html

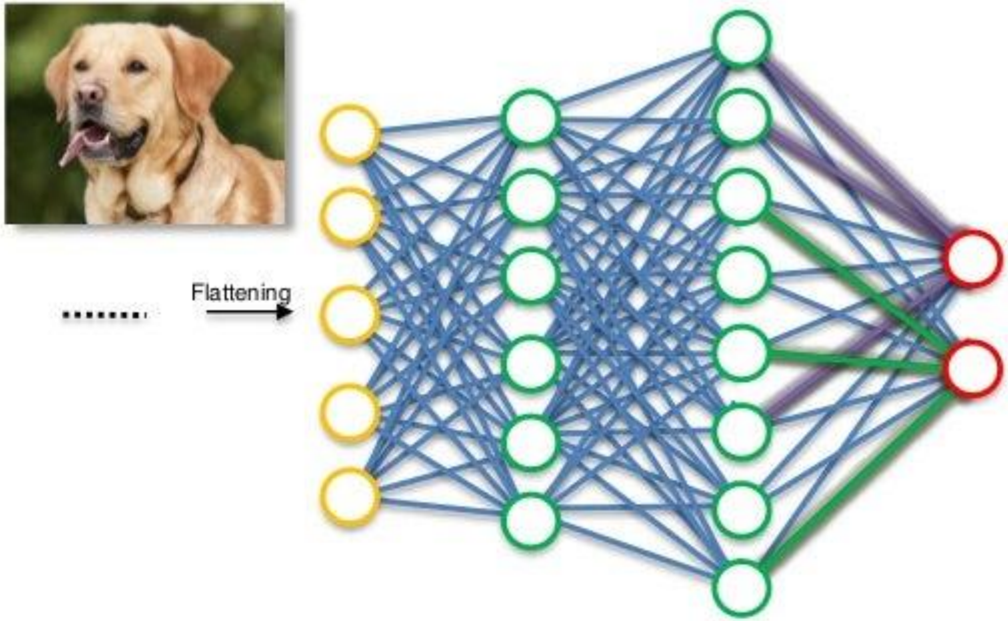
Softmax & Cross-Entropy

Softmax & Cross-Entropy



Softmax & Cross-Entropy

$$f_j(z) = \frac{e^{z_j}}{\sum_k e^{z_k}}$$



Dog $\longrightarrow$ z_1 $\longrightarrow$ 0.95

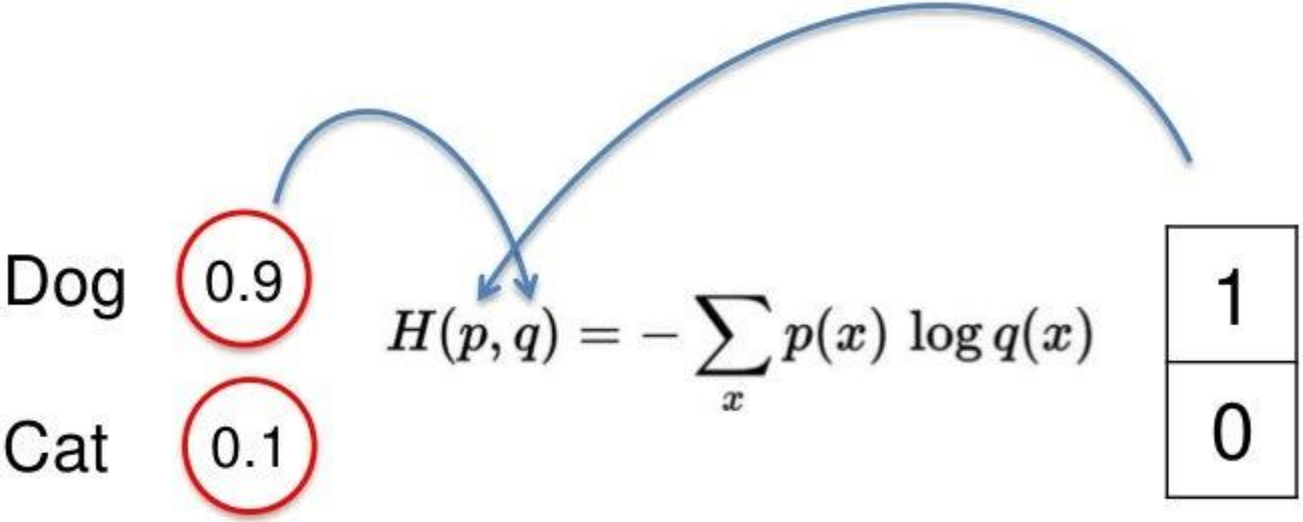
Cat $\longrightarrow$ z_2 $\longrightarrow$ 0.05

Softmax & Cross-Entropy

$$L_i = -\log\left(\frac{e^{f_{y_i}}}{\sum_j e^{f_j}}\right)$$

$$\boldsymbol{H}(\boldsymbol{p},\boldsymbol{q}) = -\sum_{\boldsymbol{x}} \boldsymbol{p}(\boldsymbol{x})\ \mathbf{log}\,\boldsymbol{q}(\boldsymbol{x})$$

Softmax & Cross-Entropy



Dog 0.9

Cat 0.1

$$H(p,q) = -\sum_x p(x)\,\log q(x)$$

1
0

Softmax & Cross-Entropy

			NN1	NN2
	Dog	1	0.9	0.6
	Cat	0	0.1	0.4
	Dog	0	0.1	0.3
	Cat	1	0.9	0.7
	Dog	1	0.4	0.1
	Cat	0	0.6	0.9

Softmax & Cross-Entropy

NN1

Row	Dog^	Cat^	Dog	Cat
#1	0.9	0.1	1	0
#2	0.1	0.9	0	1
#3	0.4	0.6	1	0

NN2

Row	Dog^	Cat^	Dog	Cat
#1	0.6	0.4	1	0
#2	0.3	0.7	0	1
#3	0.1	0.9	1	0

Classification Error

1/3 = 0.33

1/3 = 0.33

Mean Squared Error

0.25

0.71

Cross-Entropy

0.38

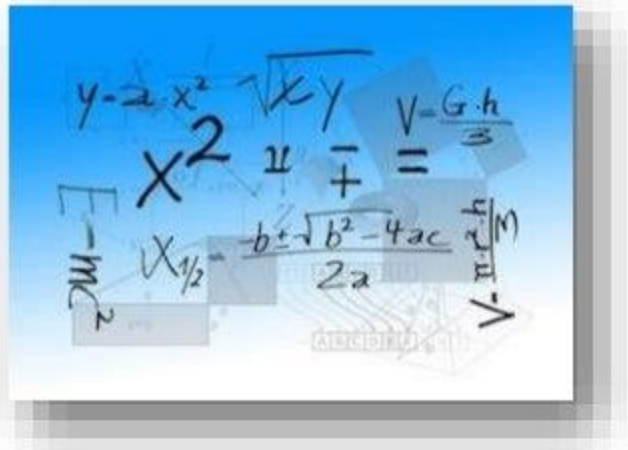
1.06

Softmax & Cross-Entropy

Additional Reading:

A Friendly Introduction to Cross-Entropy Loss

By Rob DiPietro (2016)



Link:

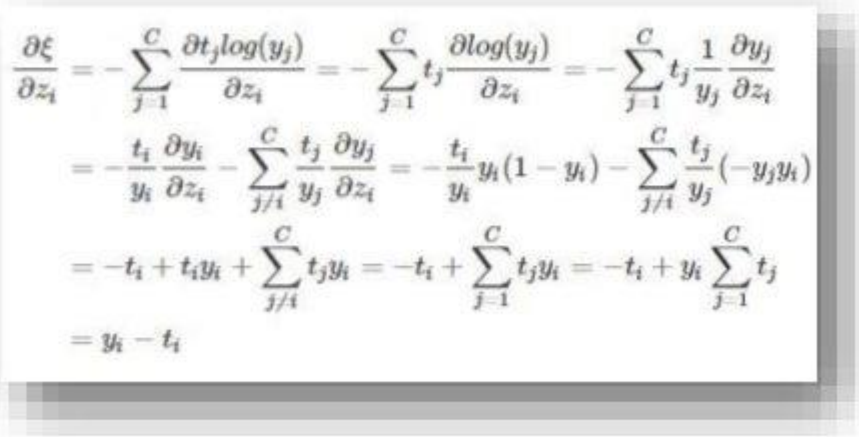
https://rdipietro.github.io/friendly-intro-to-cross-entropy-loss/

Softmax & Cross-Entropy

Additional Reading:

How to implement a neural network Intermezzo 2

By Peter Roelants (2016)



$$\begin{aligned}\frac{\partial \xi}{\partial z_i} &= -\sum_{j=1}^{C} \frac{\partial t_j log(y_j)}{\partial z_i} = -\sum_{j=1}^{C} t_j \frac{\partial log(y_j)}{\partial z_i} = -\sum_{j=1}^{C} t_j \frac{1}{y_j} \frac{\partial y_j}{\partial z_i} \\ &= -\frac{t_i}{y_i} \frac{\partial y_i}{\partial z_i} - \sum_{j \neq i}^{C} \frac{t_j}{y_j} \frac{\partial y_j}{\partial z_i} = -\frac{t_i}{y_i} y_i (1 - y_i) - \sum_{j \neq i}^{C} \frac{t_j}{y_j} (-y_j y_i) \\ &= -t_i + t_i y_i + \sum_{j \neq i}^{C} t_j y_i = -t_i + \sum_{j=1}^{C} t_j y_i = -t_i + y_i \sum_{j=1}^{C} t_j \\ &= y_i - t_i\end{aligned}$$

Link:

http://peterroelants.github.io/posts/neural_network_implementation_intermezzo02/