Sofya Aksenyuk, 150284.

IT – Lab 5

Cross-entropy & Kullback-Leibler Divergence

I. Prove that $\sum_i p(i)\log_2 \frac{p(i)}{q(i)} = H_q(p) - H(p)$, where $H_q(p)$ is the cross-entropy of q given the average distribution p, and $H(p)$ is the entropy of p.

$$H_q(p) - H(p) = \sum_x p(x)\cdot \log\frac{1}{q(x)} - \sum_x p(x)\cdot\log\frac{1}{p(x)} =$$
$$= \sum_x p(x)\cdot\left(\log\frac{1}{q(x)} - \log\frac{1}{p(x)}\right) = \sum_x p(x)\cdot\log\frac{p(x)}{q(x)}.$$

II. Neural network returns an output vector: $[-1, 2, 0]$. Use an appropriate function (softmax, sigmoid) for the classification task to produce a vector on which we can use cross-entropy loss measure.

a) Multi-class problem.

$$\left[\frac{e^{-1}}{e^{-1}+e^2+e^0}; \frac{e^2}{e^{-1}+e^2+e^0}; \frac{e^0}{e^{-1}+e^2+e^0}\right] = [0.042; 0.844; 0.114].$$

b) Multi-label problem.

$$\left[\frac{e^{-1}}{1+e^{-1}}; \frac{e^2}{1+e^2}; \frac{e^0}{1+e^0}\right] = [0.268; 0.88; 0.5].$$

III. Use the values above to compute a cross-entropy loss, if a correct output vector for this particular case was $[0, 1, 0]$. Take into account the type of classification problem, so that the error makes sense.

a) Multi-class problem.

from II a) $[-1; 2; 0] \to [0.042; 0.844; 0.114]$.

$$\left.\begin{array}{l} t = [0, 1, 0]; \\ y = [0.042; 0.844; 0.114] \end{array}\right\} CE(t,y) = 0\cdot\log\frac{1}{0.042}$$
$$+ 1\cdot\log\frac{1}{0.844} + 0\cdot\log\frac{1}{0.114} = \log\frac{1}{0.844} = 0.17.$$

b) Multi-label problem.

from IIb) $[-1,2,0] \to [0.268, 0.88, 0.5]$.

$t = [0,1,0]$;

$y = [0.268, 0.88, 0.5]$

$$\left.\begin{array}{l} t=[0,1,0]; \\ y=[0.268,0.88,0.5] \end{array}\right\} CE(t,y) = CE_1 + CE_2 + CE_3 = 0 \cdot \log\frac{1}{1-0.268} + $$

$$+ 1 \cdot \log\frac{1}{0.268} + 1 \cdot \log\frac{1}{0.88} + 0 \cdot \log\frac{1}{1-0.88} + 0 \cdot \log\frac{1}{1-0.5} +$$

$$+ 1 \cdot \log\frac{1}{0.5} = \log\frac{1}{0.268} + \log\frac{1}{0.88} + \log\frac{1}{0.5} = 1.32 + 0.69 +$$

$$+ 0.13 = 2.01 + 0.13 = 2.14.$$

IV. Compute a Kullback-Leibler divergence (relative entropy) between the probability distributions. Compare the result when distributions are swapped.

a) $P_1 = \{0.5, 0.5\}$, $P_2 = \{0.25, 0.75\}$: $d_{KL}(P_1 \parallel P_2)$ and $d_{KL}(P_2 \parallel P_1)$

1) $d_{KL}(P_1 \| P_2) = 0.5 \cdot \log\frac{1}{0.25} + 0.5 \cdot \log\frac{1}{0.75} - 0.5 \cdot \log\frac{1}{0.5} -$
$- 0.5 \cdot \log\frac{1}{0.5} = 0.5 \cdot (1.386 + 0.287 - 0.69 - 0.69) =$
$= 0.1375.$

2) $d_{KL}(P_2 \| P_1) = 0.25 \cdot \log\frac{1}{0.5} + 0.75 \cdot \log\frac{1}{0.5} - 0.25 \cdot \log\frac{1}{0.25}$
$- 0.75 \cdot \log\frac{1}{0.75} = 0.25 \cdot (0.69 - 1.386) + 0.75 \cdot (0.69 - 0.28)$
$= -0.174 + 0.3075 = 0.1335.$

$d_{KL}(P_1 \| P_2) = 0.1375 > d_{KL}(P_2 \| P_1) = 0.1335.$

b) $P_1 = \{0.25, 0.25, 0.5\}$, $P_2 = \{0.1, 0.2, 0.7\}$: $d_{KL}(P_1 \parallel P_2)$ and $d_{KL}(P_2 \parallel P_1)$

1) $d_{KL}(P_1 \| P_2) = 0.25 \cdot \log\frac{1}{0.1} + 0.25 \cdot \log\frac{1}{0.2} + 0.5 \cdot \log\frac{1}{0.7}$
$- 0.25 \cdot \log\frac{1}{0.25} - 0.25 \cdot \log\frac{1}{0.25} - 0.5 \cdot \log\frac{1}{0.5} = 0.25 \cdot (2.3 + 1.6 -$
$- 1.386 - 1.386) + 0.5 \cdot (0.356 - 0.69) = 0.115.$

2) $d_{KL}(P_2 \| P_1) = 0.1 \cdot \log\frac{1}{0.25} + 0.2 \cdot \log\frac{1}{0.25} + 0.7 \cdot \log\frac{1}{0.5}$
$- 0.1 \cdot \log\frac{1}{0.1} - 0.2 \cdot \log\frac{1}{0.2} - 0.7 \cdot \log\frac{1}{0.7} = 0.1 \cdot (1.386 -$
$- 2.3) + 0.2 \cdot (1.386 - 1.6) + 0.7 \cdot (0.69 - 0.356) = 0.0996.$

$d_{KL}(P_1 \| P_2) = 0.115 > d_{KL}(P_2 \| P_1) = 0.0996.$

$$d_{KL}(p \| q) = H(p \| q) - H(p) = E_p\left(\log\frac{1}{q}\right) - E_p\left(\log\frac{1}{p}\right) =$$

$$= \sum_x p(x) \cdot \log\frac{1}{q(x)} - \sum_x p(x) \cdot \log\frac{1}{p(x)}.$$