Problem 1

Suppose 90% samples are positive (t=1) and 10% are negative (t=0). Compute P, R, and F1 scores of majority guess (always predicting t=1). If in your derivation the denominator is 0, please compute the limit.

Explain the deficiency of the F1-score in this case, and discuss one possible treatment to resurrect P, R, F1 scores for this problem.

Problem 2.

Prove that the sigmoid function $\sigma(z) = \frac{1}{1+\exp\{-z\}}$ is symmetric about the point (0,0.5), in other words, $\sigma(-z) = 1 - \sigma(z)$.

Problem 3.

Prove that minimizing the loss $J = - t\log y - (1 - t)\log(1 - y)$ is equivalent to minimize the Kullback--Leibler (KL) divergence between $\boldsymbol{t}$ and $\boldsymbol{y}$, denoted by KL($\boldsymbol{t}$ || $\boldsymbol{y}$), where $\boldsymbol{t}$=($1 - t, t$) and $\boldsymbol{y}$=($1 - y, y$) are two Bernoulli distributions.

For two discrete distributions $\boldsymbol{P}$= $(p_1, \cdots, p_K)$and $\boldsymbol{Q}$= $(q_1, \cdots, q_K)$, the KL divergence is defined as

KL($\boldsymbol{P}$ || $\boldsymbol{Q}$)= $\sum_{k=1}^{K} p_k \log\frac{p_k}{q_k}$

Note: KL divergence is not symmetric between $\boldsymbol{P}$ and $\boldsymbol{Q}$. To minimize the KL, $\boldsymbol{Q}$ must cover all the support of $\boldsymbol{P}$. Thus, the learned distribution may be smoother than it should be.

END OF W6