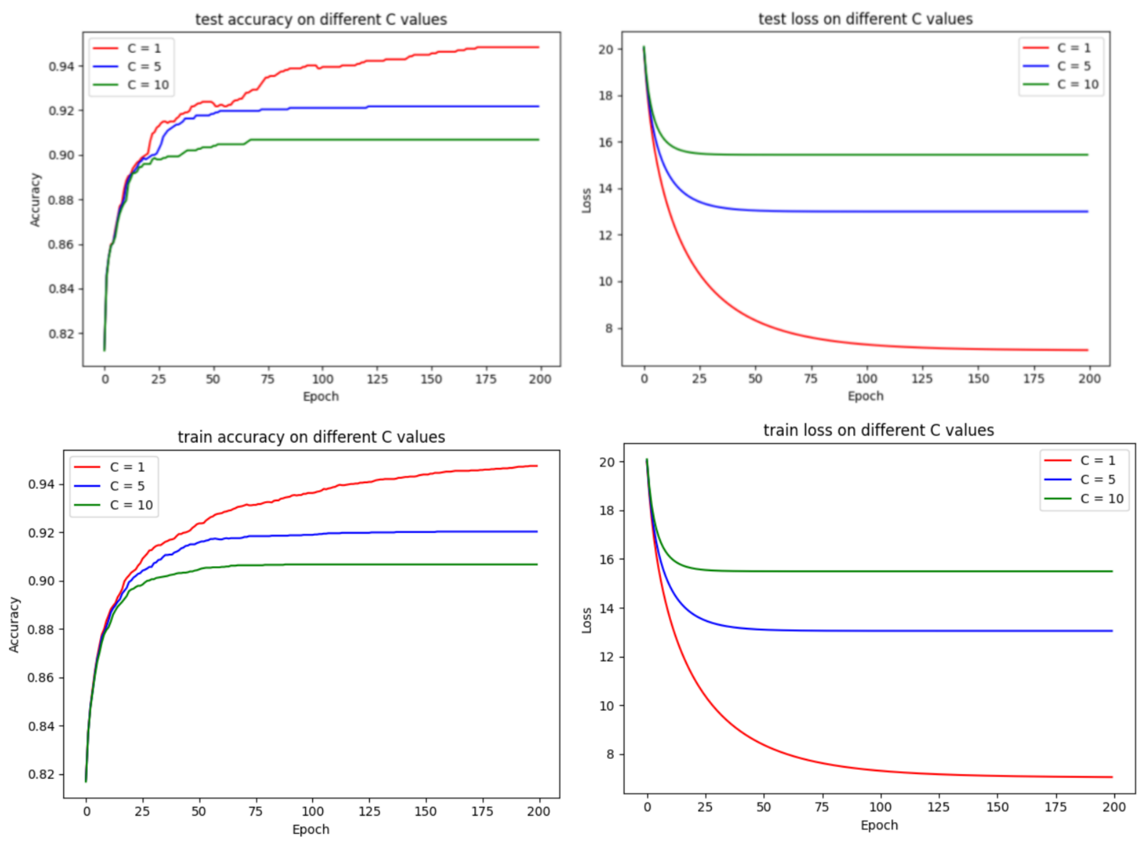
HW2

Practical Report

By Faezeh PouyaMehr and Rachel Ebrahimi

Q4-



As depicted in the figures above, the model performs almost the same on the test set as it did on the train set and the decreasing flow of the both loss plot in iterations demonstrate that overfitting is not an issue for this model and this dataset.

Subject. Practical Report _ HW 2

Date.

1)

$$R = \frac{c}{2} \sum_{j'=1}^{m} \| w^{j'} \|_2^2 = \frac{c}{2} \sum_{j'=1}^{m} \left(\sqrt{\sum_{k=1}^{p} (w_k^{j'})^2} \right)^2 = \frac{c}{2} \sum_{j=1}^{m} \sum_{k=1}^{p} {w_k^{j}}^2$$

$$\Rightarrow \frac{\partial R}{\partial w_k^j} = \frac{c}{2} \sum_{j=1}^{m} \sum_{k=1}^{p} \frac{\partial {w_k^j}^2}{\partial w_k^j} = \frac{c}{2} \sum_{j=1}^{m} \sum_{k=1}^{p} 2 w_k^j =$$

$$c \sum_{j=1}^{m} \sum_{k=1}^{p} w_k^j \xrightarrow{\text{for class } j} = c \sum_{k=1}^{p} w_k^j \checkmark$$

2)

$$H = \frac{1}{n} \sum_{(x_i, y_i) \in S} \sum_{j'=1}^{m} L(w^{j'}; (x_i, y_i))$$

$$L(w^{j'}; (x_i, y_i)) = \boxed{\overbrace{\left(\max\{0, 2 - (\langle w^{j'}, x_i \rangle) 1\{y_i = j'\}\} \right)}^{=M}}^{2}$$

$$\frac{\partial H}{\partial w_k^j} = \frac{1}{n} \sum_{(x_i, y_i) \in S} \sum_{j'=1}^{m} \frac{\partial M^2}{\partial w_k^j} \rightarrow \text{we assume } \frac{\partial \max\{0, a\}}{\partial a} = \begin{cases} 1 & \text{if } a > 0 \\ 0 & \text{if } a \leq 0 \end{cases}$$

$$\Rightarrow \frac{\partial H}{\partial w_k^j} = \frac{1}{n} \sum_{(x_i, y_i) \in S} \sum_{j'=1}^{m} 2M \frac{\partial M}{\partial w_k^j}$$. Now we calculate $\frac{\partial M}{\partial w_k^j}$

$$= \begin{cases} -x_{i,k} 1\{y_i = j\} & \text{if } 2 - (\langle w^j, x_i \rangle) 1\{y_i = j\} > 0 \\ 0 & \text{otherwise} \end{cases}$$

if $2 - (\langle w^j, x_i \rangle) 1\{y_i = j\} \leq 0$, then the max function in M will return 0 (the same as $\frac{\partial M}{\partial w_k^j}$). so we can just remove the if condition of our $\frac{\partial M}{\partial w_k^j}$ and write it as: $-x_{i,k} 1\{y_i = j\}$

$$\Rightarrow \frac{\partial H}{\partial w_k^j} = \frac{-2}{n} \sum_{(x_i, y_i) \in S} M \times (-x_{i,k} 1\{y_i = j\}) \checkmark$$

PAPCO