TA Sample Solution

Problem 1

Self-supervised learning does not need labeled data. Rather, it creates indicators for training the model from the data itself. In the area of NLP, self-supervised learning can also be used to fill in the blank of sentence based on the context. This enables any sentence to serve as the training sample without further labeling, and the model can understand the structure of the semantics. Then, the model will be helpful in multiple applications including text analysis, synthesis, and others. Lots of the LLMs in use today also use this technique a lot.

Problem 4

Let the superscript denote the data index and the subscript denote the index of a vector component. Then, we know that $\forall n$, $x_n^0 = 1$.

Let $T = T_+ + T_-$ denote the total number of updates and w_{PLA} is the final weights,

$$
\begin{aligned}
w_{\mathrm{PLA}} &= w_T \\
&= w_{T-1} + y_{n(T)} x_{n(T)} \\
&\vdots \\
&= w_0 + \sum_{t=1}^{T} y_{n(t)} x_{n(t)}, \\
\implies w_{\mathrm{PLA}}^0 &= 0 + \sum_{t=1}^{T} y_{n(t)} \cdot 1 \\
&= \sum_{t=1}^{T_+} 1 + \sum_{t=1}^{T_-} (-1) \\
&= T_+ - T_-.
\end{aligned}
$$