

Question:

The Binary Cross-Entropy (BCE) loss function is given by

$$L_{\text{BCE}}(y, y_{\text{hat}}) = - [y \log(y_{\text{hat}}) + (1 - y) \log(1 - y_{\text{hat}})],$$

where $y \in \{0, 1\}$ is the true label and $y_{\text{hat}} \in (0, 1)$ is the predicted probability.

1. Explain the importance of the logarithmic function in the Binary Cross-Entropy loss. Discuss its role in optimization, gradient behavior, and probabilistic interpretation.

2. Analyze what happens if the logarithmic function in the loss is replaced by $\sin(x)$, i.e.,

$$L_{\text{sin}}(y, y_{\text{hat}}) = - [y \sin(y_{\text{hat}}) + (1 - y) \sin(1 - y_{\text{hat}})].$$

Comment on the implications for loss behavior and learning.

3. Similarly, analyze the effect of replacing the logarithmic function with the exponential function, i.e.,

$$L_{\text{exp}}(y, y_{\text{hat}}) = - [y e^{y_{\text{hat}}} + (1 - y) e^{1 - y_{\text{hat}}}].$$

Discuss whether such a loss function is suitable for binary classification and justify your answer.