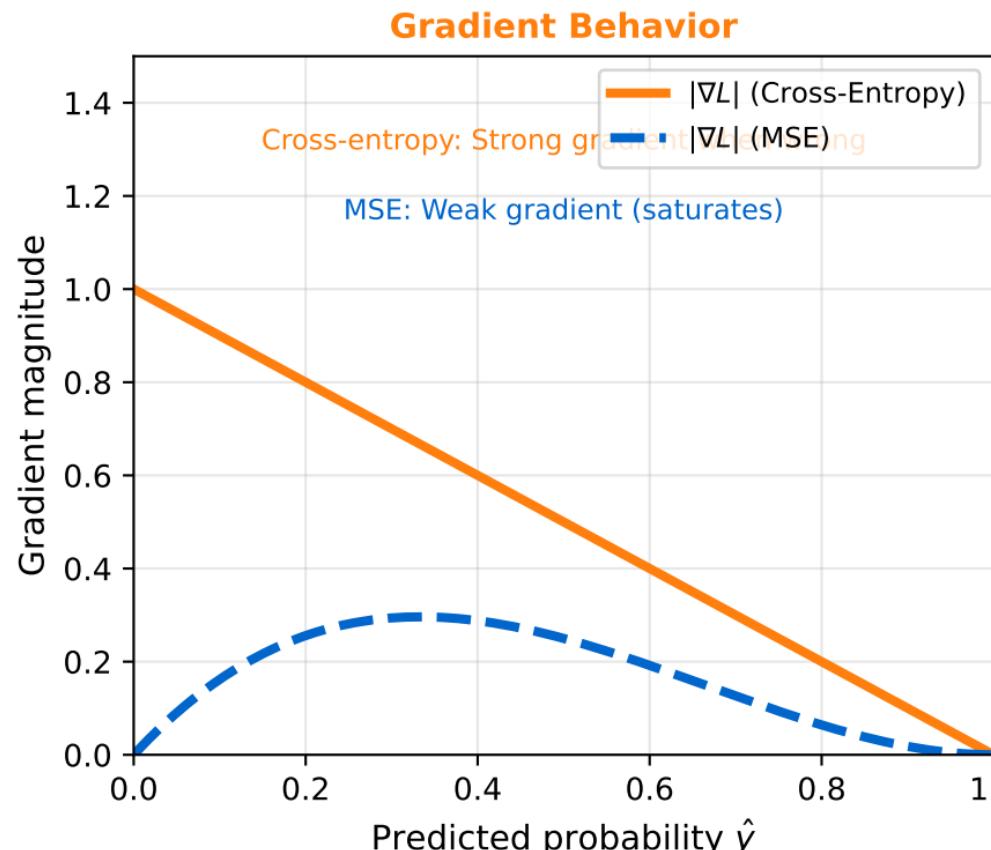
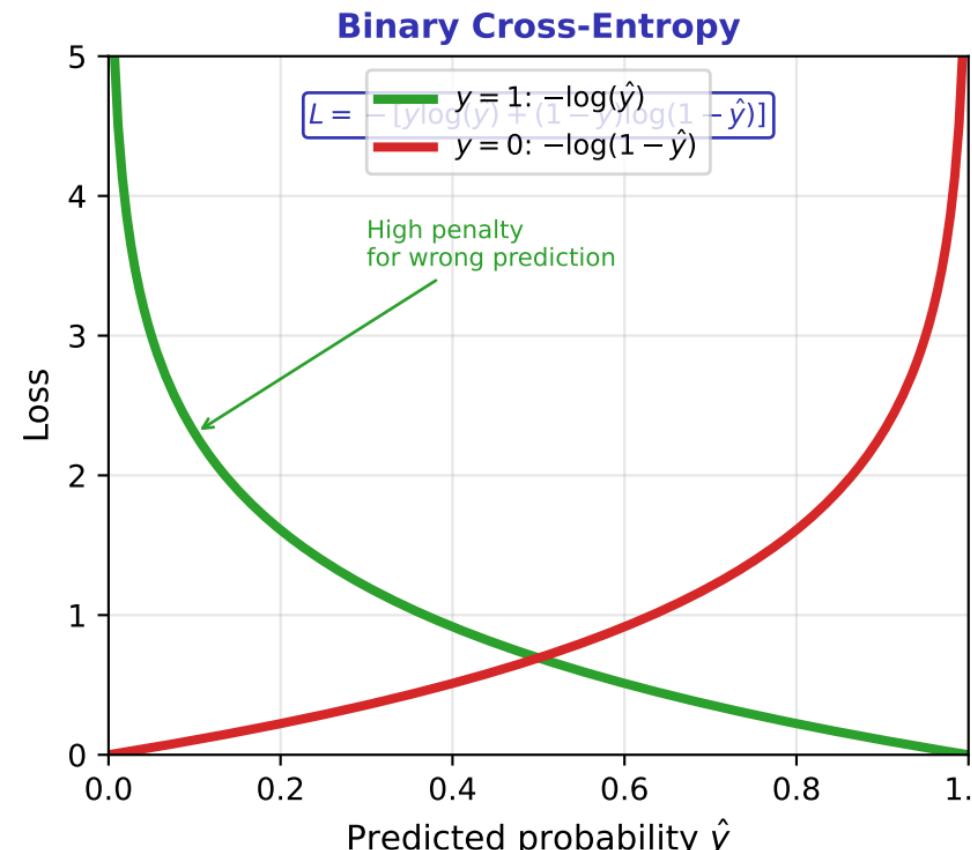


Cross-Entropy: The Standard Classification Loss



Cross-Entropy Properties

Probability-based: Measures information difference

Strong gradients: Fast learning from mistakes

No saturation: Always learns from errors

Natural for classifications: Softmax output

Why use Cross-Entropy over MSE?

- Stronger gradients for confident wrong predictions
- Information-theoretic foundation
- Standard for classification tasks