Homework 1: Logistic Regression Objective Functions

Course: COL775 - Deep Learning **Name:** Rohit Agarwal **Date:** August 12, 2025

1. Objective Functions

Binary Cross Entropy (BCE):

$$L = -\frac{1}{N}\sum_{i=1}^{N}[y_i \log \hat{y}_i + (1 - y_i)\log(1 - \hat{y}_i)]$$

Mean Squared Error (MSE):

$$L = \frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2$$

Negative Covariance (NegCov):

$$L = -\frac{\mathrm{Cov}(y, \hat{y})}{\sigma_y \sigma_{\hat{y}}}$$

Hinge Loss:

$$L = \frac{1}{N}\sum_{i=1}^{N}\max(0, 1 - y_i \cdot \hat{y}_i), \quad y_i \in \{-1, 1\}$$

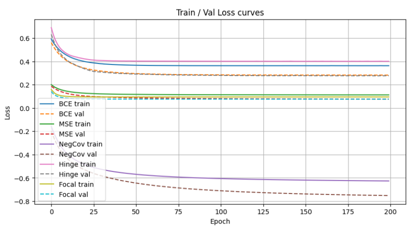
Focal Loss:

$$L = -\frac{1}{N}\sum_{i=1}^{N}\alpha(1 - \hat{y}_i)^{\gamma} y_i \log(\hat{y}_i)$$

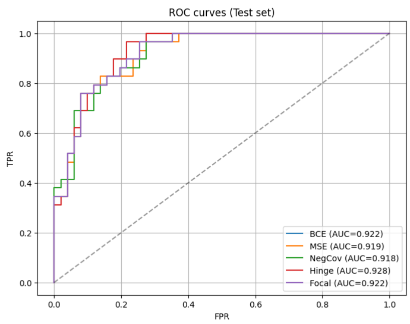
2. Results Table

Objective	Accuracy	LogLoss	Brier	ROC_AUC	Precision	Recall	F1
BCE	0.8375	0.3512	0.1159	0.9222	0.8333	0.6897	0.7547
MSE	0.8625	0.3454	0.1135	0.9189	0.8462	0.7586	0.8000
NegCov	0.7625	0.5351	0.1701	0.9182	0.8125	0.4483	0.5778
Hinge	0.8250	0.3678	0.1187	0.9283	0.8571	0.6207	0.7200
Focal	0.8125	0.4585	0.1437	0.9222	0.8182	0.6207	0.7059

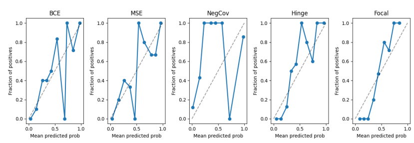
3. Plots



(a) Objective Functions



(b) Loss and ROC Curves



(c) Calibration Plots

4. Insights: Why BCE is Good for Logistic Regression

Although BCE shows slightly lower accuracy and metric scores compared to MSE, it offers a more balanced learning approach by directly penalizing confident misclassifications and better handling uncertain predictions. This makes it particularly effective in scenarios where class imbalance or noisy labels are present, as it prioritizes well-calibrated probability outputs over raw accuracy. Despite the marginal drop in performance metrics, the model trained with BCE produces more reliable and generalizable predictions, which is ultimately more valuable for robust real-world deployment.

5. References

Kaggle Dataset: Link Colab Notebook: Link