Case Study: Cost-Sensitive Focal Label Smoothing (CP-FLS)

---

1. Problem Statement & Objectives

Highly imbalanced datasets are common in real-world applications such as fraud detection, network intrusion detection, and medical imaging. Traditional classifiers tend to be biased towards majority classes, leading to poor recall for minority classes. The objective of this case study is to evaluate **Cost-Sensitive Focal Label Smoothing (CP-FLS)**, a novel loss function that integrates focal scaling, label smoothing, and cost-sensitive weighting, and to compare its performance against baseline methods.

---

2. Data Preprocessing

- **Dataset:** CIFAR-10 with artificially induced class imbalance (minority-to-majority ratio = 0.01).
- **Transformations:**
- Random cropping
- Horizontal flipping
- Normalization to dataset mean and standard deviation
- **Splits:**
- Training: imbalanced distribution
- Validation: balanced split from test set
- Test: balanced split from test set

---

3. Model Selection & Development

- **Baseline Models:**
- Cross-Entropy Loss (with class weights)
- Focal Loss
- Label Smoothing
- **Proposed Model:**
- SimpleCNN trained with CP-FLS loss
- **Hyperparameters:**
- Optimizer: Adam (lr=0.001)
- Batch size: 128
- Epochs: 50
- Gamma (focal): 2.0
- Smoothing: 0.1
- Cost-weight exponent (q): 0.5

---

4. Visualizations & Insights

4.1 Training Curves

- Loss, Accuracy, F1-score, and AUC over epochs for both training and validation sets.

4.2 Precision-Recall Curves

- PR curves comparing CE, Focal, LS, and CP-FLS.
- CP-FLS shows superior PR-AUC.

4.3 Calibration Curves

- Reliability diagrams demonstrating better calibration (lower ECE) for CP-FLS.

4.4 Confusion Matrices

- CP-FLS achieves better minority recall compared to baselines.

---

5. Comparative Results (Placeholder)

Loss Type	Accuracy	F1	AUC	PR-AUC	ECE ↓	Brier ↓
CE						
Focal						
LS						
CP-FLS						

---

6. Recommendations

- **When to use CP-FLS:**
- Applications with highly imbalanced datasets.
- Domains requiring both high recall and good calibration (fraud detection, medical imaging).
- **Benefits:**
- Improves minority class recall without sacrificing calibration.
- Reduces expected cost under fixed false-positive budgets.

---

7. Conclusion

This case study demonstrates that CP-FLS outperforms baseline loss functions in terms of recall, calibration, and overall robustness on imbalanced CIFAR-10. The findings suggest CP-FLS as a practical solution for real-world imbalanced classification tasks.

Appendix

- Figures: Training plots, PR curves, Calibration curves, Confusion matrices.
- JSON results file with per-epoch metrics.
- Best model checkpoint (`best_model.pth`).