Final Report — Loan Policy Project

1. Overview

This report summarizes the EDA, supervised predictive model, and offline RL experiments.

2. Supervised DL model (MLP)

- Best MLP config: {'hidden_dims': [512, 256], 'dropout': 0.2, 'lr': 0.001, 'batch_size': 512, 'epochs': 10, 'patience': 3}

- Test AUC: 0.7367

- Best Test F1: 0.4539 at threshold 0.26

3. Offline RL agent (Discrete CQL)

- Estimated policy value (mean reward per case on test): -0.15966403481283487

4. Why these metrics?

AUC & F1 for DL model:

- AUC measures the model discrimination across thresholds, useful for ranking risk.

- F1 measures balance between precision and recall at a chosen threshold; important when class imbalance exists and when we care about correct default detection.

Estimated Policy Value for RL agent:

- In offline RL we aim to learn a policy that maximizes expected return. Estimated policy value (average monetary reward on a held-out test set) directly measures the downstream business objective (profit minus losses).

5. Policy comparison and example differences

The supervised DL model defines a thresholding policy (approve when predicted default probability < t). The RL agent learns a policy that maximizes expected reward and may approve some high-risk applicants if the expected reward (interest) outweighs expected losses. Examples with differing decisions can be inspected by comparing model predictions and RL action for the same applicant.

6. Limitations & Future Steps

- Behavior policy was simulated because raw logged approve/deny decisions were unavailable. Replace with real logged propensities if available.

- Use OPE (IS/DR) for robust policy evaluation, and tune CQL/IQL for better performance.

- Consider reward normalization (e.g., divide monetary values by loan amount) and more complex reward that includes time-discounted cashflows.

7. Reproducibility

Follow the instructions in `README.md`. Key scripts:

- `train_mlp_grid.py` — grid search for MLPs

- `evaluate_best.py` — evaluate and save final supervised metrics

- `build_rl_dataset.py` and `train_offline_rl.py` — offline RL pipeline