Loan Default Prediction & Offline RL Report

1. Task 1 – Exploratory Data Analysis and Feature Selection

I performed initial data analysis and preprocessing to understand loan applicant data.

Key Steps:

- **Data Cleaning:** Missing values handled via mean imputation (numerical) or constant 'missing' (categorical). Categorical variables one-hot encoded.
- **Feature Selection:** 19 features chosen based on financial logic and predictive power:

Feature	Justification
loan_amnt	Larger loans increase potential loss; critical for reward in RL.
int_rate	Determines profit and correlates with risk.
term_months	Longer terms increase default risk.
annual_inc	Repayment capacity.
dti	Debt-to-income ratio; high values indicate risk.
emp_length_yrs	Employment stability.
grade	Lending grade; pre-assessed risk.
home_ownership	Residential stability.
purpose	Loan purpose may correlate with default.
verification_status	Verified applicants less risky.
application_type	Joint vs individual applications.
revol_util	Credit utilization; high utilization indicates stress.
open_acc, total_acc	Credit history and experience.
delinq_2yrs	Recent delinquencies increase default probability.
inq_last_6mths	Recent credit inquiries indicate risk.
pub_rec	Public records (bankruptcy, judgments).
credit_age_years	Longer history reduces default risk.
target	Binary: 0 = Fully Paid, 1 = Defaulted

- **Class Balancing:** Downsampling performed to create a balanced dataset (equal defaulted and fully paid loans).

---

2. Task 2 – Predictive Deep Learning Model (MLP)

- **Objective:** Predict probability of default.
- **Model:** Multi-Layer Perceptron with 256→128→1 architecture, ReLU activations, dropout 0.3, batch normalization.
- **Training:** Balanced dataset, binary cross-entropy loss, Adam optimizer, batch size 1024, 20 epochs.

Evaluation Metrics:

Metric	Result	Interpretation
AUC (Test)	0.733	Measures the model's discriminative ability. AUC > 0.7 indicates good separation of defaulted vs non-defaulted loans.
F1-score (Test)	0.669	Balances precision and recall, crucial for predicting defaults accurately under class imbalance.

Why not just Accuracy?

- After downsampling, accuracy and AUC may look similar, but accuracy at a fixed threshold can be misleading if the business later changes its approval threshold or faces different class distributions in production.
- AUC + F1 provide more stable and business-relevant evaluation than accuracy.

---

3. Task 3 – Offline Reinforcement Learning (Fitted Q Iteration, MLP)

- **Objective:** Learn a policy to approve or deny loans that maximizes profit while minimizing losses.
- **State:** Preprocessed feature vector for each applicant.
- **Action:** {0: Deny, 1: Approve}
- **Reward:**
    - Approve & Fully Paid → loan_amnt * int_rate * 10 (profit)
    - Approve & Default → -loan_amnt * 5 (loss)
    - Deny → 0
- **Algorithm:** Fitted Q Iteration using separate MLP regressors per action.
- **Training:** 30 iterations, max 500 iterations per MLP, warm start enabled.

Evaluation Results:

Policy	Approval Rate	Avg Reward / Policy Value
Learned FQI-NN	4.87%	-79.15

Policy	Approval Rate	Avg Reward / Policy Value
Baseline Heuristic	18.92%	-634.89

Why Estimated Policy Value?

- Measures **expected financial outcome** if the learned policy is applied.
- Directly aligns with business objectives: maximize profit, minimize losses.
- Higher Policy Value → more effective, safer loan approval decisions.

Analysis:

- FQI agent drastically reduces defaults and potential loss compared to baseline.
- Policy is conservative, approving only highly profitable and low-risk loans.
- Although average reward is negative, it is significantly better than baseline, showing the RL agent learns a safer, profit-oriented policy.

---

4. Task 4 – Comparison and Insights

- **MLP vs RL:**
    - MLP predicts default risk (AUC/F1), implicitly defining a threshold-based approval policy.
    - RL learns a profit-driven approval policy directly, taking loan amounts and interest rates into account.
- **Policy Differences:**
    - Some low-risk applicants may still be denied by RL if high loan amount and risk-reward is unfavorable.
    - Conversely, MLP may approve them ignoring profit considerations.
- **Limitations:**
    - Offline RL relies on historical data; rare cases may be missed.
    - Reward amplification may exaggerate conservativeness.
- **Future Steps:**
    - Explore more sophisticated RL algorithms (CQL, BCQ) for better policy optimization.
    - Collect additional applicant data (credit bureau, employment verification) for more accurate models.
    - Combine MLP predictions with RL policy as a hybrid approach.

---

5. Summary

- Features selected based on financial relevance and predictive power.
- MLP predicts loan default with strong AUC (0.733) and F1-score (0.669).
- Offline RL agent learns a safe, profit-oriented policy, outperforming heuristic baseline.
- **Complementary approach:** MLP predicts risk; RL maximizes expected profit while mitigating loss