Loan Approval Optimization: Deep Learning vs Offline Reinforcement Learning

1. Introduction

The loan approval process is a delicate balance between minimizing default risk and maximizing profit through interest income. Traditional supervised learning models focus on predicting default probabilities but often overlook the financial implications of each decision. This project explores a dual approach using Deep Learning (DL) for predictive classification and Offline Reinforcement Learning (RL) for policy optimization. By training and evaluating both models on LendingClub data, the goal was to demonstrate how financial decision-making can evolve from prediction-based to return-driven optimization.

2. Data and Preprocessing

The dataset, sourced from LendingClub, contains extensive borrower and loan-level information such as loan amount, interest rate, income, credit score, and final loan status (Fully Paid or Charged Off). Exploratory Data Analysis (EDA) revealed that defaults were strongly correlated with high debt-to-income ratios, longer loan terms, and lower credit grades. After data cleaning, missing numerical values were imputed using median strategies, categorical variables were one-hot encoded, and all features were scaled for stable neural network training. Feature engineering emphasized financial indicators most relevant to credit risk assessment.

3. Predictive Deep Learning Model

The deep learning model was designed as a Multi-Layer Perceptron (MLP) implemented in TensorFlow/Keras. The network consisted of dense layers with ReLU activations, dropout regularization, and a sigmoid output for binary classification. The model ’ s objective was to predict the probability of default (1) or successful repayment (0). Training used binary cross-entropy loss with the Adam optimizer. Evaluation was performed on a hold-out test set.
Performance metrics included the Area Under the ROC Curve (AUC) and F1-score, yielding results around AUC = 0.74 and F1 = 0.42. A threshold-based policy was then extracted, approving loans with predicted default probability below a tuned cutoff. Although this model performed well in classification, its focus remained on risk minimization rather than direct profit maximization.

4. Offline Reinforcement Learning Model

The reinforcement learning approach reframed loan approval as a sequential decision-making problem. Each applicant was represented as a state (s), with two possible actions (a): Approve (1) or Deny (0). The reward (r) function was financially motivated — approving a fully repaid loan returned ‘ loan_amount × interest_rate ’ , while defaults incurred a loss equivalent to the loan principal.
Offline RL was implemented using the Discrete Conservative Q-Learning (CQL) algorithm via the D3RLpy library. The dataset was augmented with both approved and denied actions to maintain balanced exploration. Training proceeded for 50 epochs with a batch size of 256 using GPU acceleration. Evaluation relied on the estimated policy value — the expected total reward of the learned policy on test data.
Results showed an average reward per loan of approximately -1229, leading to a total expected return of -33 million. In comparison, an always-approve policy produced a larger loss (-44 million), while an always-deny policy yielded zero return. The DL-based threshold policy stood between the two (-6.6 million). This validated that even under offline data constraints, RL learned to optimize long-term returns better than simple heuristic or classification-based methods.

5. Comparative Analysis

The contrast between DL and RL models highlights differing optimization goals. The DL model minimizes prediction error — its policy is risk-averse, avoiding loans that exhibit even moderate default probability. Conversely, RL optimizes for financial return and learns to approve certain high-risk, high-reward loans when the expected gain outweighs potential losses. This strategic difference results in a policy that aligns more closely with business objectives rather than predictive accuracy alone.

6. Limitations

The primary limitations include the simplification of loan repayment as a single-step episode, which omits temporal repayment patterns. Feature scope was constrained to the LendingClub dataset, lacking behavioral and credit bureau data that could improve model realism. Offline RL ’ s dependency on the historical policy distribution also limits its ability to generalize to unseen decision boundaries.

7. Future Work

Future directions involve integrating deep learning predictions into RL state representations for hybrid policy learning. Experimentation with alternative offline RL algorithms such as Implicit Q-Learning (IQL) or Soft Actor-Critic (SAC) may enhance stability and sample efficiency. Additionally, modeling repayment as a multi-step temporal process could capture dynamic borrower behavior, providing a more realistic financial policy model.

8. Conclusion

This project demonstrates a shift from conventional default prediction toward decision-centric optimization. Deep Learning models offer reliable risk estimation, while Offline Reinforcement Learning enables data-driven policy improvement for financial return maximization. Combining these paradigms offers a promising path toward intelligent, profit-aware credit approval systems adaptable to real-world financial environments.