Cost-sensitive Reinforcement Learning for Credit Risk

Source: ScienceDirect, Jorge et al., 2025

Link: https://www.sciencedirect.com/science/article/pii/S0957417425003306

1. **Goal/Problem Statement**

   To maximize long-term business benefit in dynamic credit risk, where loss/costs depend on loan amounts. Aim: adapt reinforcement learning (RL) and bandit algorithms for cost-sensitive settings, overcoming bias due to data limited to approved transactions.

2. **Strategies/Tools/Methods Used**

- Reinforcement learning (RL) and cost-sensitive classification algorithms.
- Cost-sensitive passive-aggressive online learning.
- Cost-sensitive logistic bandit using Thompson Sampling.
- Simulations and experiments on credit risk datasets.

3. **Key Results and Findings**

- Cost-sensitive RL models outperform traditional ones when decisions directly impact monetary loss.
- Bandit algorithms can balance exploitation/exploration to minimize opportunity cost & bias.
- Instance-dependent loss functions lead to greater business reward than global accuracy.

4. **Critical Learnings/Takeaways**

- RL and bandit methods, tailored to cost-sensitive credit decisions, significantly improve aggregate financial outcomes in real-world, dynamic financial environments.
- Addressing only observed approved transactions prevents unfair bias.

5. **Other Noteworthy Points**

- Applicability to insurance and fraud detection by adjusting cost matrices.
- Data privacy restricts full reproduction; code or pseudocode referenced for implementation.

Example-dependent Cost Sensitive Learning Based Selective Deep Ensemble Model for Customer Credit Scoring

Source: Nature, Jin Xiao et al., 2025

Link: https://www.nature.com/articles/s41598-025-89880-7

1. **Goal/Problem Statement**

   Optimize customer credit scoring by minimizing business loss with example-dependent cost-sensitive (ECS) deep learning ensembles, handling imbalanced and varied-loss scenarios.

2. **Strategies/Tools/Methods Used**

- Development of ECS-TabNet for interpretable tabular data.
- GMDH (self-organizing inductive modeling) for selective deep ensemble.
- Use of bespoke ECS loss functions tailored for individual sample costs.
- Comparison with CCS and deep ensemble methods.

3. **Key Results and Findings**

- ECS-SDE model outperforms six cost-sensitive and five deep ensemble credit scoring models.
- Achieves top performance on savings, AUC, and "cost of misclassification" metrics.
- Enhanced feature interpretability with attention mechanisms.

4. **Critical Learnings/Takeaways**

- ECS methods better reflect true business costs than uniform-cost algorithms, critical for scalable, explainable credit/fraud management.
- Selective deep ensembling reduces redundancy and increases predictive accuracy.

5. **Other Noteworthy Points**

- TabNet's interpretability aids regulatory and audit compliance.
- GMDH enables efficient model selection within large parameter spaces.

Cost-sensitive Thresholding over a Two-Dimensional Decision Region for Fraud Detection

Source: ScienceDirect, Jorge & Cao et al., 2024

Link: https://www.sciencedirect.com/science/article/pii/S0020025523015414

1. Goal/Problem Statement

   Reduce aggregate financial loss in fraud detection by considering both transaction probabilities and amounts in classification thresholds.

2. Strategies/Tools/Methods Used

- Two-dimensional thresholding (fraud probability + transaction amount).
- Novel algorithm for decision region optimization.
- Comparative study with standard classifiers.

3. Key Results and Findings

- Two-dimensional approach consistently improves loss reduction over traditional one-dimensional thresholding.
- More effective assignment of fraud investigation resources.

4. Critical Learnings/Takeaways

- Incorporating transaction amount is essential for scalable, cost-sensitive fraud operations.
- Strategies can be adapted to insurance/fraud integration projects.

5. Other Noteworthy Points

- Demonstrates generalization to various classifiers; data privacy limits reproduction.

Reinforcement Learning Applied to Insurance Portfolio Pursuit

Source: Alan Turing Institute, Young et al., 2024

Link: https://www.turing.ac.uk/sites/default/files/2024-08/2408.00713v2.pdf

1. **Goal/Problem Statement**

   Address portfolio pursuit in insurance using RL—modulate customer offers to achieve targeted portfolio distributions and maximize aggregate profit, not just one-off wins.

2. **Strategies/Tools/Methods Used**

- Formulation of portfolio pursuit as a Markov Decision Process (MDP).
- RL-based sequential decision methodology.
- Synthetic agent-based simulation of insurance market interactions.
- Baseline industry comparisons.

3. **Key Results and Findings**

- RL method yields more profit (7% increase) and comparable portfolio quality to baseline over hundreds of simulated customer interactions.
- Method enables firm to prioritize long-term strategic portfolio goals over short-term pricing.

4. **Critical Learnings/Takeaways**

- RL-based pricing/pursuit outperforms static industry methods for insurance portfolio management.
- Method is flexible and capable of integration with existing cost-estimation models.

5. **Other Noteworthy Points**

- Provides open-source codebase for industry benchmarking.
- Identifies future research directions: handling customer exits, longer-term horizons, real-world deployment risks.