Transfer Learning for Credit Recovery Rates: Evaluating Performance under Feature and Label Shifts

Abstract

Accurate forecasting of credit recovery rates is essential for effective credit risk management. A significant challenge in this context is data scarcity, particularly when predicting recovery rates for specialized loan segments or portfolios. Transfer learning (TL) emerges as a valuable approach, enabling banks to leverage insights from related domains with richer datasets, thereby improving prediction accuracy in risk management.

However, TL faces substantial obstacles due to inevitable data shifts occurring when transferring predictive insights between different credit portfolios. Such shifts typically arise from variations in credit products, borrower characteristics, and market conditions, complicating direct knowledge transfer.

Our research systematically evaluates the robustness of predictive models to various data shifts, such as changes in features, labels, or their combinations, using advanced TL methods like attention-based transformer architectures and tree-based ensemble models. We develop a simulation framework to generate synthetic loan data that mirrors real-world distributions derived from industry-standard datasets such as Global Credit Data (GCD). Using automated shifts within a Monte Carlo simulation framework, we rigorously analyze model performance under realistic conditions.

Extensive ablation studies show how different data shifts influence TL effectiveness, quantified using metrics such as the Kullback-Leibler divergence. Additionally, we explore how the size of the target dataset affects TL effectiveness, providing insights into the minimum data required to leverage pretrained models from related credit portfolios. While the empirical analysis is ongoing, preliminary findings indicate substantial performance variability based on shift complexity, data volume, and transfer strategy, highlighting the challenges of effectively applying TL to recovery rate modeling. For instance, even minor changes in feature distributions between source and target datasets can significantly impact the corresponding LGD distributions, jeopardizing the utility of TL. By clarifying the dependencies between data shift types and TL methods, our study provides actionable insights for practitioners on when and how to use TL to enhance recovery rate forecasting.

Authors & Affiliations

Christopher Gerling[1], Hanqiu Peng[2], Prof. Dr. Stefan Lessmann[1], Prof. Dr. Ying Chen[2]
[1]Humboldt University of Berlin, Berlin, Germany. [2]National University of Singapore, Singapore, Singapore