

Driving Financial Inclusion Through Advanced Scoring with Alternative Data: How Sophisticated Modelling leveraging Open Banking Data can Increase Lending, Decrease Bad Rates and improve Responsible Lending

Abstract

In this study, we apply the standard credit scoring methodology to open banking data, incorporating reject inference through multi-dimensional parcelling and advanced feature engineering for transaction data. The increasing availability of open banking data introduces unique opportunities and challenges in credit risk modelling, particularly in the treatment of rejected applicants. Unlike traditional bureau data, open banking data provides granular insights into consumer behaviour, including affordability measures that can enhance the assessment of creditworthiness. However, it also poses challenges in constructing swap sets post-reject inference, which must be carefully addressed to maintain model performance and fairness.

We demonstrate the effectiveness of our approach using real-world open banking data. Our results show that logistic regression, combined with a weights-of-evidence (WoE) transformation, remains a robust modelling technique capable of outperforming bureau credit scores for a subprime population. The incorporation of multi-dimensional parcelling for reject inference plays a crucial role in refining the model's predictive power, particularly when leveraging transaction data. This data provides additional insights into spending patterns, income stability, and financial behaviour, which are essential for feature construction in the rejected population.

A critical aspect of reject inference is the recalibration of score cutoffs to ensure that newly inferred applicants are assessed fairly while maintaining an optimal balance between risk and approval rates. We explore the challenges associated with setting new score thresholds and propose strategies to mitigate potential pitfalls, such as shifts in bad rate distributions and model drift over time.

Our findings indicate that this approach yields significant benefits, including:

- A reduction in bad rate, leading to more effective risk management.
- An increase in overall lending, allowing financial institutions to expand their portfolios with greater confidence.
- A boost in financial inclusion, as certain underrepresented populations gain improved access to credit based on alternative data sources.

By leveraging open banking data and refining reject inference techniques, this study highlights a path toward more accurate, fair, and inclusive credit decisioning.

Authors & Affiliations

Jacobus Johannes Eksteen^{1,2}

¹Matogen Applied Insights (Pty) Ltd, Cape Town, South Africa; ²Atto, Scotland, United Kingdom