Digital Footprint Credit Scoring Modeling

Abstract

Digital footprint data, gathered from users' interactions and behavior during online processes, is becoming a valuable resource in assessing credit risk. This study looks at how behavioral signals—like device usage, time spent on screens, and how applicants fill in online forms—can help predict creditworthiness, particularly for customers without existing credit histories.

In total, more than 600,000 variables were collected from the digital application journey. Out of these, 20 key parameters were selected and analyzed, including the type and number of devices used for registration, session activity, time spent on specific fields, and email structure. While most of these variables, when looked at individually, showed weak to medium predictive power, they performed much better when combined into a full model.

The resulting model, built only on digital footprint data, achieved a Gini coefficient of 32.3%, which is acceptable for application scorecards—especially in cases where credit bureau data is unavailable. When credit bureau data was added to the model, the Gini increased to 51%, showing the benefit of combining behavioral and traditional sources. Importantly, digital footprint data was available for about 95% of applicants, making it a highly scalable and practical approach.

This work shows that digital footprint variables can play an important role in credit scoring, especially in markets or segments where traditional data is limited. For financial institutions looking to serve underbanked or new-to-credit customers, incorporating behavioral data into risk models can improve both the accuracy of decisions and the reach of credit products.

Authors & Affiliations

Dr. Dmytro Kolechko[1]
[1]Wing Bank, Phnom Penh, Cambodia