Cut your credit losses without losing borrowers

ZAML™ allows you to improve your underwriting by leveraging machine learning

Why traditional underwriting makes it hard to grow without increasing credit losses

In today's low interest rate environments, growing lending is the primary path to revenue growth—and minimizing credit losses is key to growing lending businesses. But to grow lending requires solving one particularly hard problem: Underwriting more consumers, particularly thin- and no-file borrowers.

Most underwriting technology in use today does a good job of identifying creditworthy borrowers with an easily accessible credit history. But traditional underwriting has not changed in 50 years. This lack of innovation makes it challenging to identify millions of creditworthy borrowers.

Fifty years ago, most people with a credit bureau file had neither missing data nor errors in their file. This is no longer true. As many as 40% of Americans—including tens of millions of millennials—now have thin credit files, or no credit file. These applicants—whether they will be good credit risks or not—are neglected because they haven't amassed the extensive credit histories needed to fuel traditional underwriting models.

This problem is even worse in many emerging markets because the data needed for traditional underwriting doesn't exist in those markets. The result: Businesses needing to grow their lending in these markets often struggle to control credit losses.

ZestFinance's new approach to underwriting enables lenders to minimize credit losses while lending to thin-file and no-file applicants in the US and internationally.

How the Zest Automated Machine Learning (ZAML) platform can help

Scoring thin-file applicants effectively requires, not surprisingly, adding more data than that found in the credit bureau files. In fact, there are thousands of pieces of information on applicants both on the internet and in company internal databases. However, traditional underwriting is unable to process more than about 50 data points. This contradiction begs the question: What's next? How can I use that mass of data to help provide fair and transparent credit to applicants?

Machine learning (ML) is the answer. ML can help lenders improve their risk performance across these previously hard-to-score populations by using all that data instead of the 50 or fewer data points traditional models use.

But ML is not magic. It's quite difficult to move from traditional underwriting methods. Upfront costs—in time and money—can be prohibitive for acquiring and preparing the necessary data and building the supporting ML infrastructure.

And even given the data, there is a dearth of data scientists who know the math and computer science that underpins ML. As a result, it's extremely hard to hire great, experienced talent.

In addition, machine learning models often function as "black boxes." One can see the model's output but can't explain what drove that output. This affects the lender and its regulators. The lack of transparency makes it hard for modelers to understand how to iterate and improve their models.

Even more importantly, the black box nature of ML models makes it impossible for lenders to provide legally required information to applicants—like adverse action—and to regulators—like disparate impact reports.

The ZAML platform was built to overcome these obstacles, in addition to providing world-class data handling and modeling environments.

First, ZAML makes it easy for data scientists to learn the math underpinning machine learning: ZAML makes good modelers into great ones and novice modelers into experienced ones quickly.

Second, ZAML provides extremely powerful explanation tools that make any black-box ML model transparent. The platform automatically generates dynamic reports to support modeler iteration. ZAML tells modelers what to focus on to improve and does much of the grunt work for them.

Finally, these same explanation tools allow the lender to provide adverse action in whatever format they choose to applicants. Disparate impact reports, using the regulators' approved techniques, are automatically generated.

ZestFinance developed ZAML—an end-to-end underwriting platform—over almost a person-century of experience lending to and scoring diverse customers segments. ZAML's ***data assimilation*** tools allow lenders to acquire, onboard, and prepare massive amounts of disparate data for modeling. This data can come from external sources. However, it often starts with additional internal data that the lender has but can't use in underwriting. ZAML's ***modeling environment*** makes it easy for data scientists to train, ensemble and productionalize models extremely efficiently. Together these tools drastically lower the time and financial cost of adopting machine learning. And ZAML's ***explainability tools*** solve black box concerns, providing model insights to executives and tools to support analyses needed for compliance.

ZAML in action: Two examples of lenders who are improving risk performance

ZAML will let your modelers find nontraditional, complex indicators of applicant risk, even thin-file applicants. ZAML models have both found surprising high-risk populations and identified denied applications that were likely good borrowers. It's possible to reduce aggregate portfolio risk by denying future no-longer surprising applicants. Alternatively, one could approve the newly found good borrowers, increasing the total book size with no increase in portfolio risk. Let's focus on the reduction of portfolio risk.

*** A top-five U.S. credit card issuer** is using ML models built on the ZAML platform to decrease loss rates by more than 15%, saving more than $100 million in annual losses.

The ZAML model built with this customer was able to identify distinct customer segments other models missed and then recognize which variables drove behavior in each respective segment. For example, a large proportion of approved borrowers used their card only one time and then paid it off. Normal underwriting often would categorize these customers as great credit risks — but they are wildly unprofitable. The ZAML model was able, automatically, to recognize this customer segment and underwrite it separately from the longer-term, more profitable — and higher risk — card users. Additionally, the ZAML model noticed that the application channel — online versus at a gas station, for example — had big risk impacts.

The team used the ZAML explanation module to identify and understand these factors. Both of these factors are intuitive in retrospect, which is exactly what one wants in an ML model. To the extent that factors seem totally surprising, one's model is potentially broken. However, these intuitive factors are very hard to include in a traditional model and are almost impossible to add automatically: If a model doesn't know to expect them, the model won't find them for her. And seemingly random factors are unlikely to survive regulatory scrutiny. The ZAML explanation module ensures both that models don't include such random factors and that all factors are human-understandable.

*** A major U.S. auto lender** is using ML models built with ZAML to cut losses from low- and no-FICO borrowers by more than 20%.

Using its data assimilation and modeling tools, the ZAML model built with this customer spots risk indicators masked in high-level data. For example, while a traditional model might focus on metrics like loan-to-value and payment-to-income ratios, the ZAML model looks deeper into the data, spotting hidden indicators that a customer has taken out an additional loan to cover their down payment. While this is also an intuitive indicator of credit risk, it's one that traditional underwriting models and their limited supporting data might miss.

Neither of these companies relied heavily on ML before working with ZAML. But they had capable and experienced data scientists and analysts who were able to harness the value of machine learning because ZAML endowed them with the right tools.

ZAML's data assimilation tools allowed the clients' data scientists and analysts to use a variety of disparate data sources and produce a clean, robust data set for modeling, with hundreds or thousands of variables for each applicant. The ZAML modeling tools enabled them to develop submodels and ensemble them into a single, integrated underwriting model. And finally, ZAML's explainability tools allowed them to unpack the model results in each case to understand what was driving the credit scoring. With these explainability tools, both clients were able to ensure that they continued to meet regulations for adverse action and disparate impact.

Grow your lending business with new math and more data

Machine learning requires both new math and more data. Often, that data already exists within an organization. ZAML allows lenders to use ML and new data—either from internal sources or a third party—to improve the performance of their underwriting. And it allows them to do so efficiently, cost-effectively, and compliantly.

To learn more about how ZAML can help your lending business cut losses without sacrificing borrowers, contact us at partner@zestfinance.com or visit www.zestfinance.com.