IMPACT 2023

CORNERSTONE ANNUAL MEETING & COUNCIL FORUM

Pillars of Safe AI-Powered Lending by Zest AI

Pillars of Safe AI-Powered Lending

April 2023

Introduction

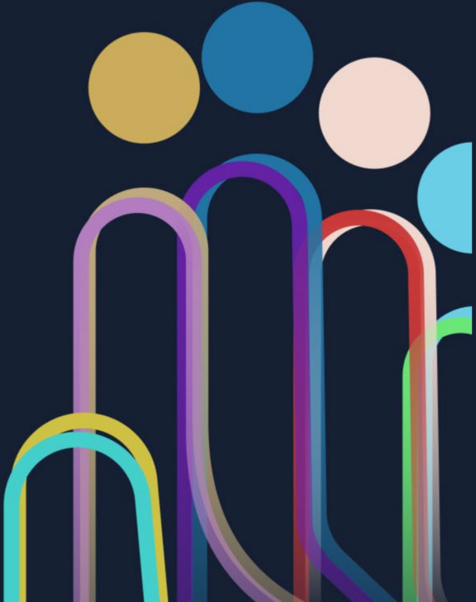
ZEST AI

SMART

MAKE BETTER CREDIT DECISIONS

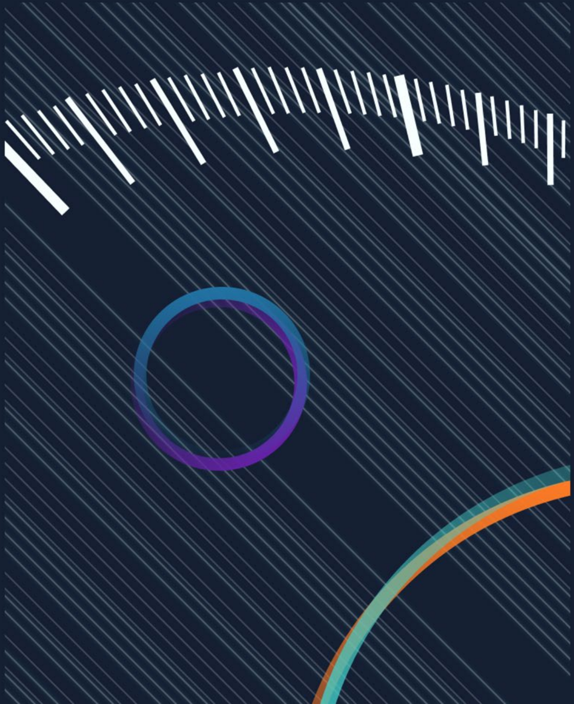
INCLUSIVE

ENSURE ALL YOUR MEMBERS GET A FAIR SHOT



DELIVER MORE DECISIONS FASTER

EFFICIENT



Meet Zest AI

Company Overview

- We're a tech company based in Burbank, CA with 110+ employees
- Founded in 2009 with a mission to make fair and transparent credit available to all
- Help lending institutions make better lending decisions that are smart, efficient, and inclusive
- We do this by leveraging AI-automated credit underwriting technology to provide lending insights that boost accuracy 2x over the leading industry score

Why we exist...

40%

of Americans are difficult for lenders to score accurately

Source: CFPB and Experian

20-40

year old scoring methods are out of date

83%

of current lending tools return decisions in over 30 minutes

Source: Zest survey

How we help...

- Increase lending volumes
- Reduce losses
- Greater automation in decisioning
- Improve fairness and inclusivity
- Lower operating costs

Industries Served

CU's | Banks | Specialty | Fintech

Stress-tested by the largest, most regulated financial institutions

DISCOVER
citi
First National Bank
TRUIST
FIFTH THIRD BANK
Freddie Mac

‘21 Finovate winner; best use of AI/ML

Accessible to and aligned with the credit union movement

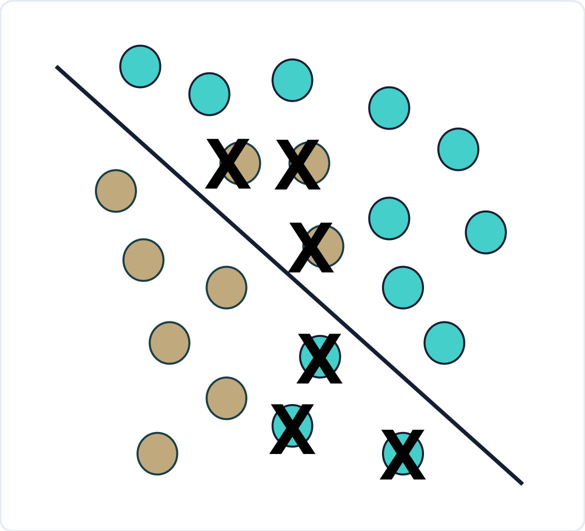
3RIVERS
5Point Credit Union
Addition Financial
All in Credit Union
APCU Atlanta Postal Credit Union
Altra Federal Credit Union
blue
COASTAL federal credit union
Commonwealth Credit Union
Credit Union West
FirstService Credit Union
Greater Texas Credit Union
GreenState Credit Union
HawaiiUSA
ih Credit Union
Redwood Credit Union
Suncoast Credit Union
Truliant Federal Credit Union
VyStar Credit Union
WSECU

‘22 Credit Union Service Org of the Year

The Move to AI in Lending

AI Models Classify Risk Better

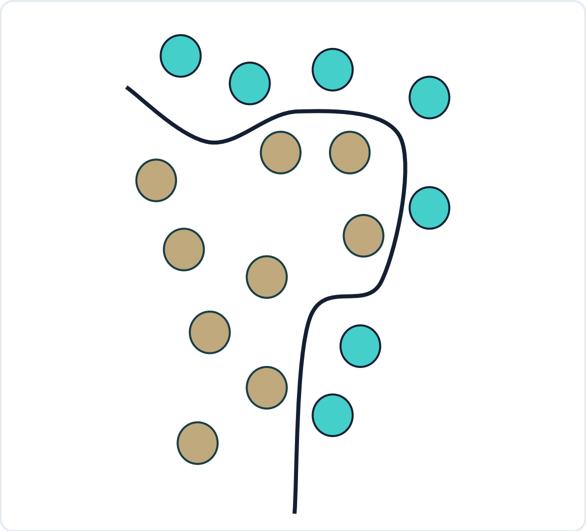
Linear Model



Linear Models Can't Fit Complex Data

This Linear Model Makes **Six Classification Mistakes**

Machine Learning Model



Machine Learning Models Successfully Fit Complex Data

This Machine Learning Model Makes **No Classification Mistakes**

AI Models are More Accurate

Huge lift over the Benchmark

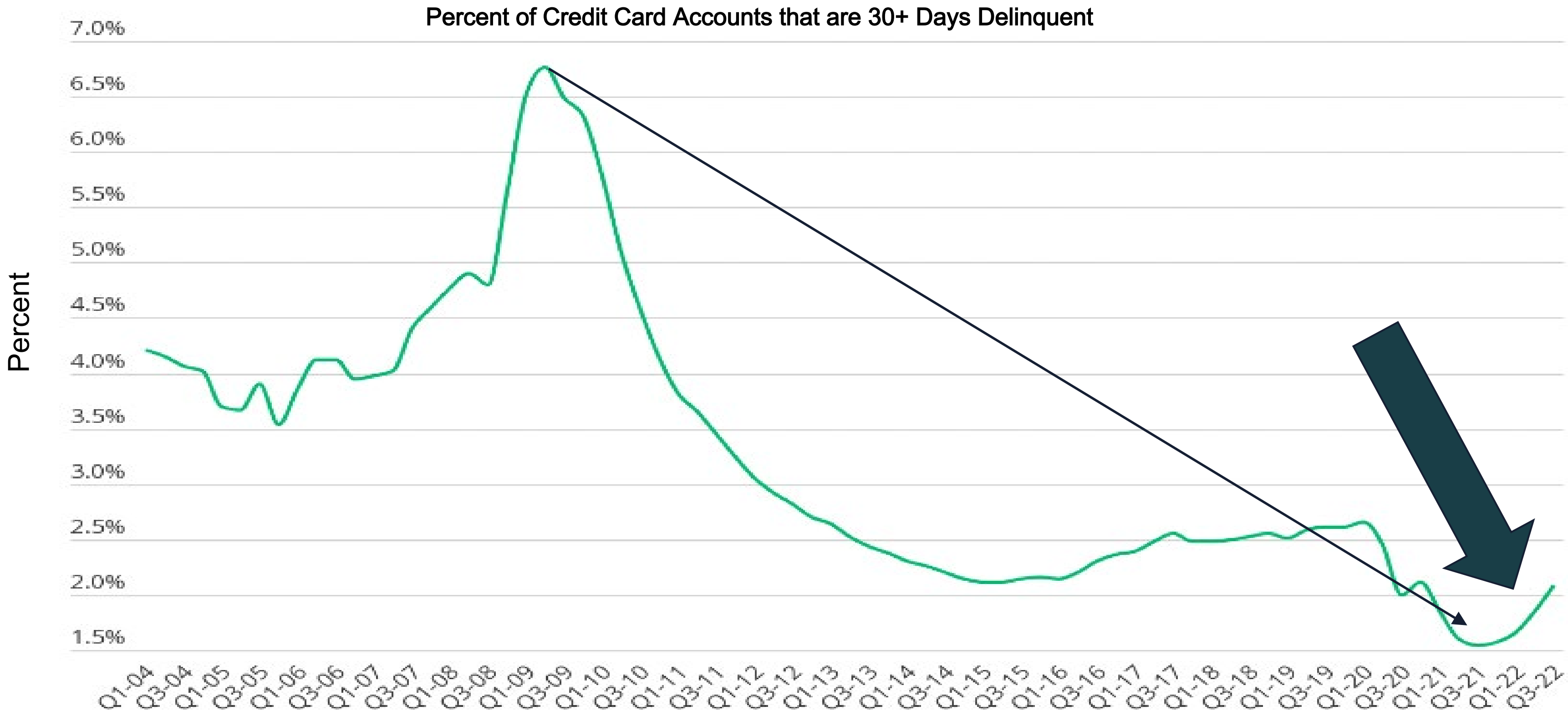
What are we showing?
Zest vs. National Credit Score comparisons of AUC, a measure of statistical accuracy. The AUC statistic assesses model performance by measuring the model's ability to discern defaults from non - defaults.

What are we looking for?
The higher the AUC, the better.

AUC COMPARISON:
Zest vs. Benchmark Model

Takeaway:
The biggest AUC lift is in the middle tiers, where Zest does a significantly better job of identifying risky borrowers.

(Imprecise Decisions Won't Continue to Work)



AI Models Are More Stable

When market conditions change, traditional algorithms can become unstable . . .

Bloomberg

Credit Scores 'May Lose Some Power' After Covid, Fed Warns

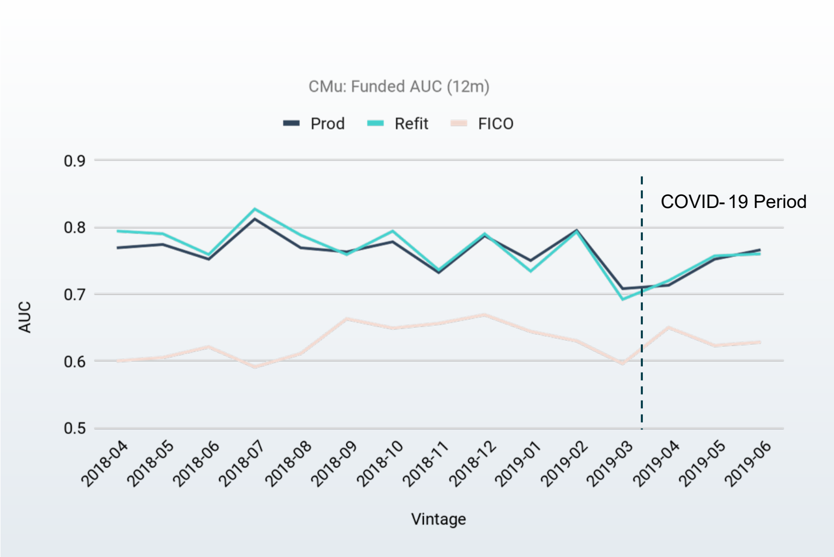
May 19, 2021

Federal Advisory Council
Board of Governors, Federal Reserve

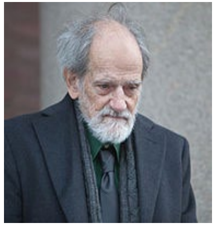
Given market changes during pandemic, "FICO scores do not provide meaningful insights. . . ."

Dec 3, 2020

Machine learning algorithms have proven to be resilient . . .



AI Models Are More Transparent



Nobel Laureate Lloyd Shapley

Value ϕ of a player i in a game f (with S: coalitions, N: players)

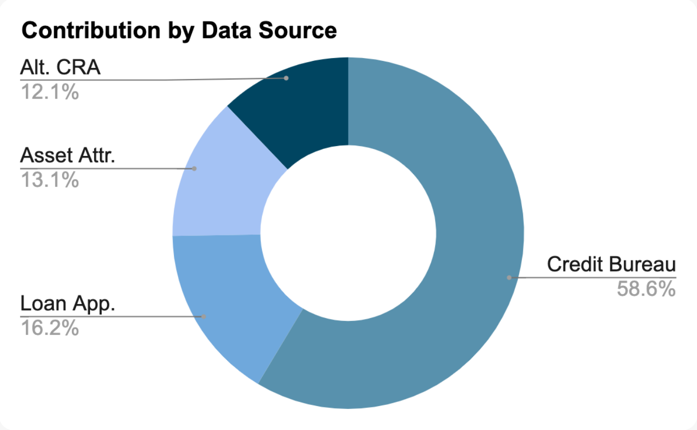
$$\phi_i(f) = \sum_{S \subseteq N \setminus \{i\}} \frac{|S|!(|N|-|S|-1)!}{|N|!} \underbrace{(f(S \cup \{i\}) - f(S))}$$

What happens when a player is missing?

Shapley proved that this formula gives the only attribution that satisfies fundamental mathematical properties of completeness, sensitivity, linearity, and symmetry

Attributes used in the model

Data Source	Attribute Count
Credit Bureau	294
Loan Application	20
Asset Attributes	12
Alternative CRA	50

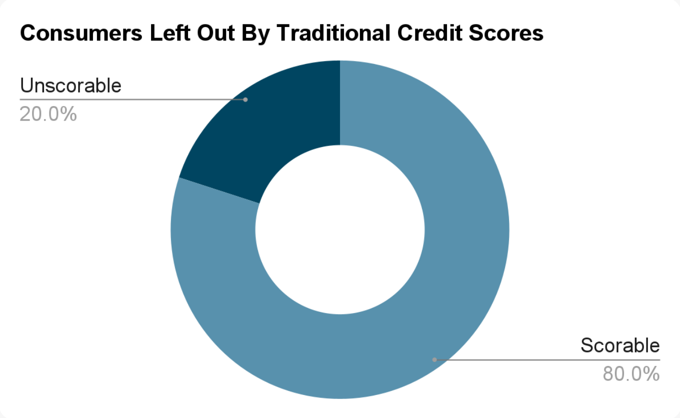


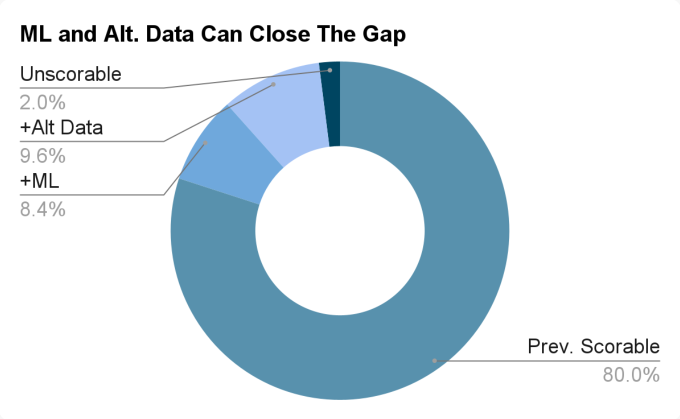
AI Models Can Score Thin File Borrowers

1 in 5 US adults are unscorable by popular credit scoring algorithms. The antiquated math used by credit scoring companies needlessly excludes over 45M Americans (approx 20%).

Machine learning and alternative data allows lenders to assess the "credit invisible" and cover 98% of American adults:

- Machine learning models can scores borrowers with thin credit bureau files, covering 42% of unscorables
- Adding alternative data allows Zest models to cover 90%+ of unscorables





ZEST AI

AI Models are More Inclusive

With

Models optimized for both fairness and accuracy

You can

Ensure you are being consistent and equitable

And

Help the underserved

AI Models Produce Meaningful Results

Expanding credit access

VyStar Credit Union

22%	$40M
Increase in approvals; holding risk constant	New credit issued per year

"What really excites us is the opportunity to offer instant decisions, better pricing, and personalized service to our members through our partnership with Zest"

Jenny Vipperman,
Chief Lending Officer

Accelerating change

First National Bank

28%	$37M
Reduced risk; holding approval rate constant	Savings identified by management

"Zest really focused on the things that matter to us: fairness, compliance, documentation, and rigorous ML monitoring. Zest's technology helps us optimize, ensuring we do right by the communities we serve."

Mihaela Kobjerowski,
Chief Credit Officer

Lending more inclusively

Freddie Mac

25,000	$9.4B
Additional approvals of Black and Latinx families	Value of additional home loans to families of color

"Zest allows us to do our machine learning modeling work more efficiently and with less operational risk."

Michael Bradley,
SVP Single Family Modeling & Analytics

Compliant AI Lending

Pillar 1: Advanced Explainability

ML models require advanced math to "open the black box"

Adverse action and fair lending analysis require more rigorous, game-theoretic methods to calculate the drivers of model-based decisions for ML models; older math doesn't cut it

X

Drop One

For each applicant, drop (delete) each variable and see what it does to the model's prediction. If deleting the value changes the score, it is deemed "important" and the code associated with that feature goes in the NOAA.

X

Impute Median

For each applicant, for each variable, substitute the median value for approved borrowers. If replacing the value changes the score, the variable is deemed "important" and the code associated with that feature goes in the NOAA.

✓

Game Theoretic Approaches*

For each applicant, compare the applicant's score to an approved score and analyze the model to determine which variables drove the difference in score on their own and in combination with other variables.

* See Aumann and Shapley (1974), *Values of Non-Atomic Games.*

Drop One and Impute Median are almost always wrong, when run on ML models

- During a CFPB tech sprint on adverse action notices, we generated denial reasons for a machine learning model built to approve loans for a mid-sized auto lender using various methods.
- We compared Drop One and Impute Median to the Shapley game-theoretic baseline to assess their accuracy.
- Drop One was wrong ~90% of the time; Impute Median was almost always incorrect.

Percentage of time the reasons given by each method correctly matched any of the top 3.

TECHNIQUE	1ST REASON	2ND REASON	3RD REASON
Drop One	11%	11%	13%
Impute Median	~0%	~0%	1%

More rigorous adverse action reason methodology should be required for ML models

Game theory rigorously quantifies the impact of each player in a collaborative game

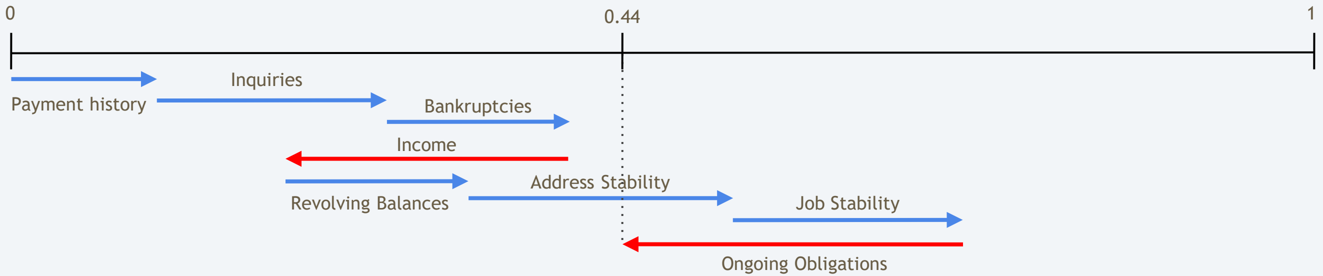
Value ϕ of a player i in a game f (with S: coalitions, N: players)

$$\phi_i(f) = \sum_{S \subseteq N \setminus \{i\}} \frac{|S|!(|N|-|S|-1)!}{|N|!} \underbrace{(f(S \cup \{i\}) - f(S))}$$

What happens when a player is missing?

Shapley proved that this formula gives the only attribution that satisfies fundamental mathematical properties of completeness, sensitivity, linearity, and symmetry

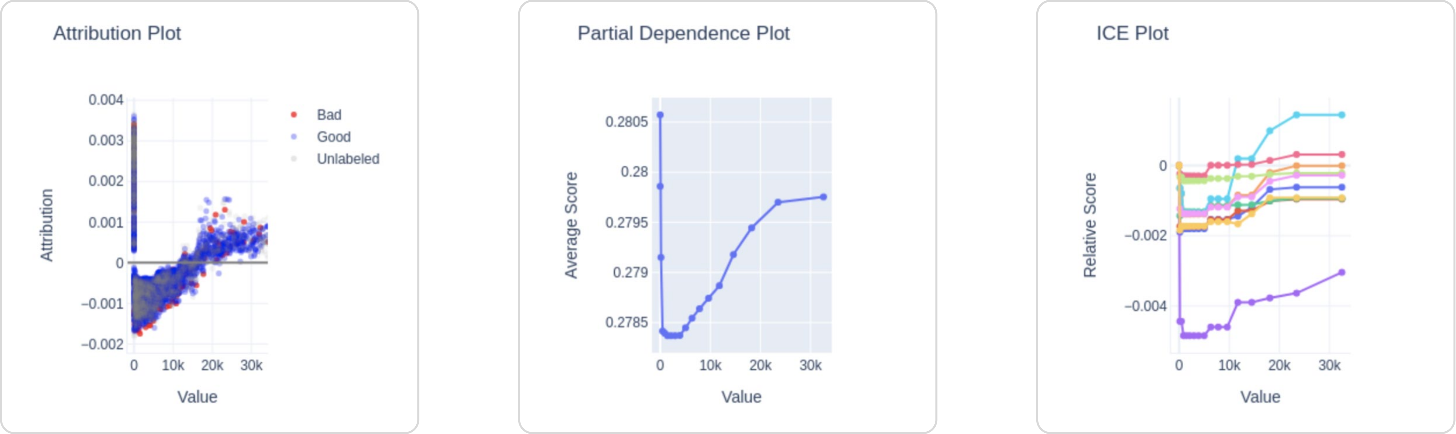
Example of **Shapley decomposition** of applicant with score of 0.44.



Compliant AI tools must provide detailed insight into model behavior

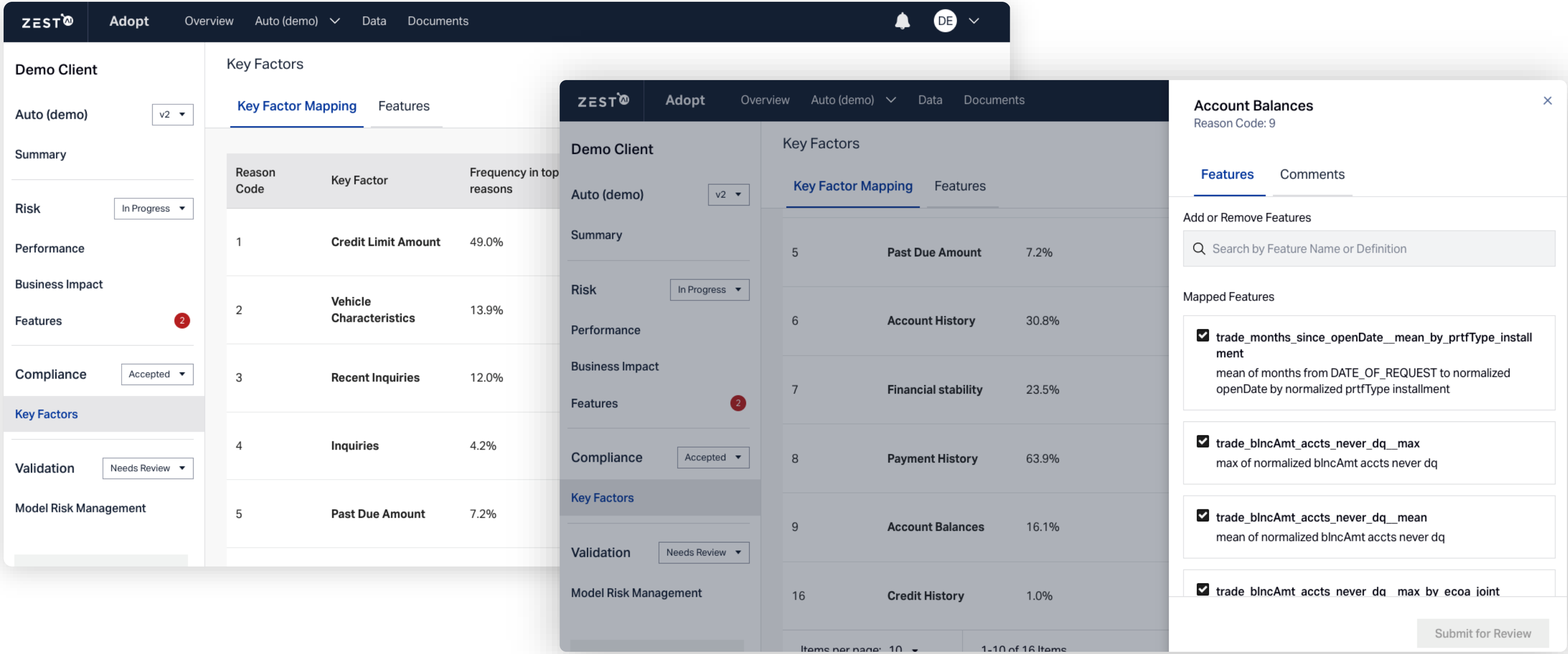
Attribution, partial dependence and ICE plots deliver a detailed view on how a model behaves

In the example below, we are looking at how the model's risk assessment is impacted by one variable: **Sum of current balances on revolving accounts** (displayed along the x-axis)



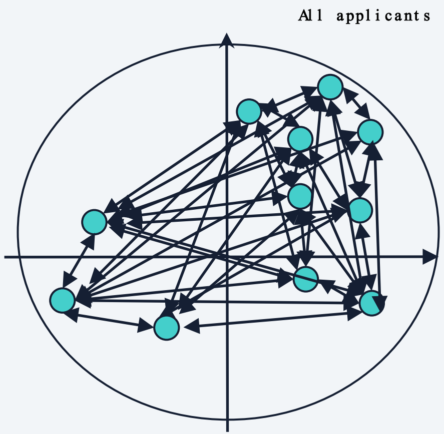
The model's risk assessment is non-monotonic with respect to this variable; this can easily be corrected

Compliant AI tools should make it easy to map model variables to ECOA adverse action reasons



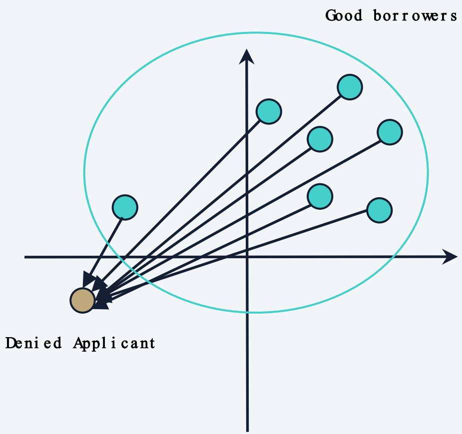
Compliant explainability tools must be able to explain more than just adverse action reasons

Feature Importance



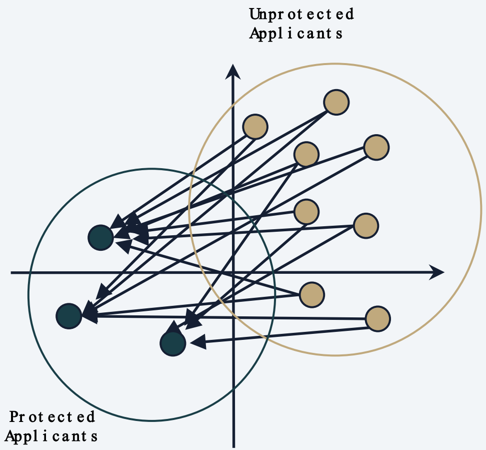
Compare all applicants with each other to compute the average marginal contribution

Adverse Action



Compare the denied applicant with approved applicants to compute the marginal contribution of each variable to the denial

Disparate Impact



Compare protected applicants with their unprotected counterparts to compute the marginal contribution of each variable to a difference in score

Compliant AI Lending

Pillar 2: Advanced Fair Lending

Fair lending review has 5 components

Disparate Treatment Assessment

Evaluate whether demographic characteristics are used directly when assessing an applicant's creditworthiness

Disparate Impact Analysis

Assess the degree to which disparity exists between protected and non -protected classes

Feature Attribution

If disparate impact exists, identify and quantify the features most responsible for driving that disparity

LDA Search

Evaluate whether demographic characteristics are used directly when assessing an applicant's creditworthiness

Fair Lending Documentation

Evaluate whether demographic characteristics are used directly when assessing an applicant's creditworthiness

Software can automate this process for any kind of origination model

Step 1: Disparate treatment analysis

Software identifies whether a feature proxies for a protected class by building a univariate model that predicts the protecte d status; AUC = 1 indicates a perfect proxy

Hispanic:
Top 5 most impactful features

Impact Rank	Feature Name	AUC
1	Avg. credit limit amount on all revolving accounts	0.61
2	Avg. credit limit amount on revolving accounts with high credit to credit limit > 0.25	0.61
3	Avg. credit limit amount on never-delinquent revolving accounts	0.6
4	Avg. months since open date	0.6
5	Avg. credit limit amount on active revolving accounts	0.6

Gender:
Top 5 most impactful features

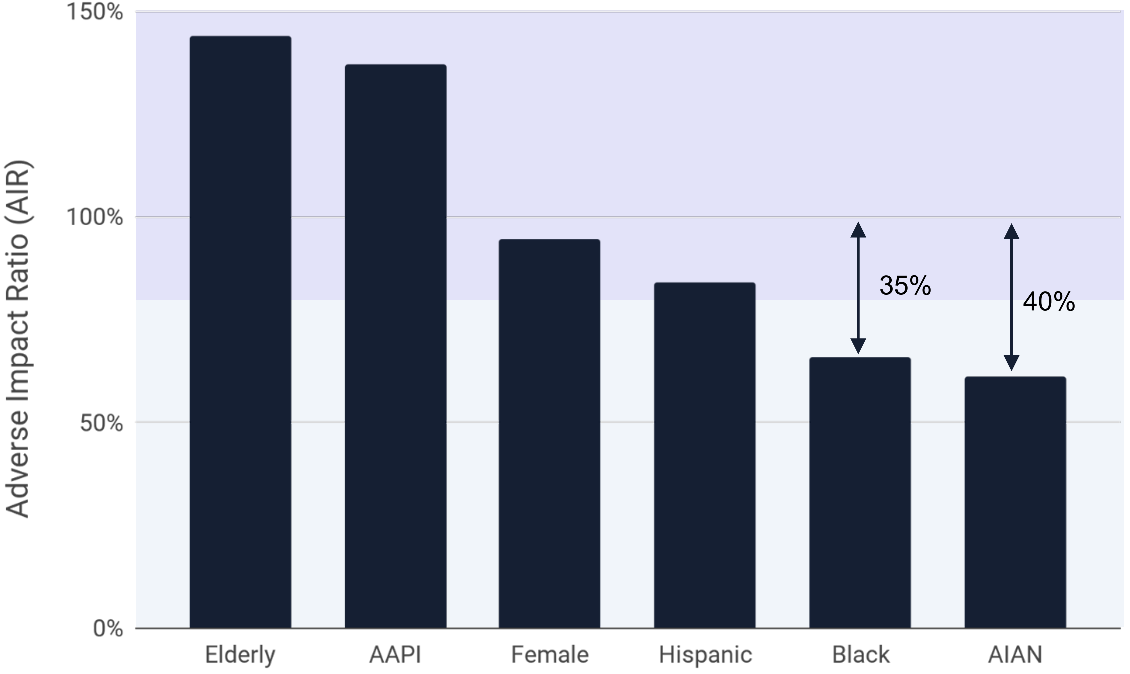
Impact Rank	Feature Name	AUC
1	Min. credit limit amount on never-delinquent revolving accounts	0.59
2	Max high credit amount on individual accounts with recent payment	0.57
3	Avg. credit limit amount on all revolving accounts	0.57
4	Avg. credit limit amount on active revolving accounts	0.57
5	Avg. credit limit amount on revolving accounts with high credit to credit limit greater than 0.25	0.56

Age:
Top 5 most impactful features

Impact Rank	Feature Name	AUC
1	Avg. months since open date	0.74
2	Avg. months since open date on all revolving accounts	0.72
3	Avg. months since open date on all individual accounts	0.71
4	Avg. since open date on all credit card accounts	0.68
5	Max credit limit amount	0.64

In this case, no disparate treatment was found (all AUCs < 1)

Step 2: Disparate impact assessment



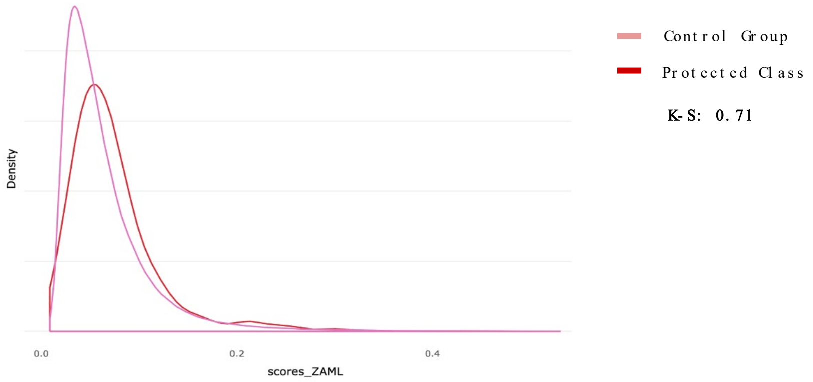
Significant Disparate Impact:

AIR below 80% indicates significant disparity and comes with increased fair lending enforcement risk

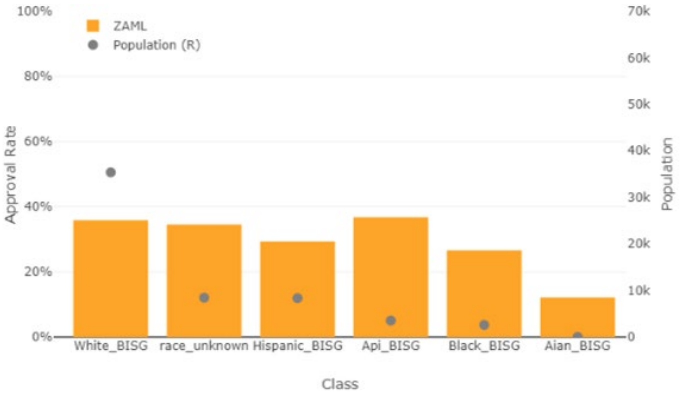
Step 2: Disparate impact assessment

Quantify any differences in score distribution and outcomes for protected borrowers

First, examine the score distributions of the protected class and the control group



Second, examine how this distribution affects approval rates, AIR, pricing, etc

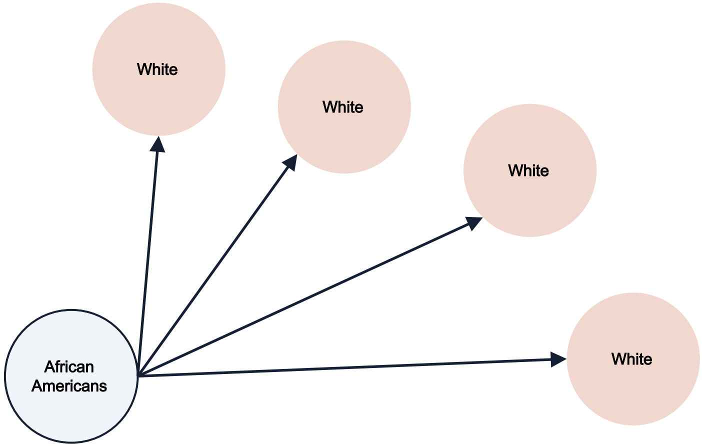


Zest software identifies disparate impact by computing metrics like max K-S

Step 3: Feature Attribution

Identify which features are driving the disparate impact

Each protected class borrower is explained in reference to the control group



Sample of top 5 impactful features for protected class borrowers in a test model

RANK	FEATURE	IMPACT
1	Ratio of satisfactory trades to total	2.5%
2	Total high credit	2.4%
3	Total mortgage balances	1.6%
4	Sum of avg. balances for unclassified trades	1.5%
5	Total inquiries last 6 months	1.4%

Shapley values quantify the extent to which a variable causes a difference in score

Step 3: Compare how each variable contributes to predictive accuracy and disparate impact

Average marginal contribution of each variable

MODEL VARIABLE	CONTRIBUTION TO MODEL PREDICTIONS	CONTRIBUTION TO DISPARATE IMPACT
Credit Score	32%	28%
Loan To Value	21%	17%
Down Payment Amount	11%	14%
Monthly Income	8%	12%
Count of Bankruptcies	6%	2%
Delinquencies	4%	2%
Length of Credit History	4%	1%

* For simplicity, the contribution to disparate impact is shown on an aggregated basis, these statistics are typically disaggregated by protected group. Bootstrap sampling enables us to put confidence intervals on all statistics.

Feature selection is critical

Some of the best predictors of credit risk treat protected classes unfairly, but which signals should you cut? If you drop one, the model doesn't work.

Step 4: LDA search

Determine whether there is a practical less discriminatory alternative. If so, the lender may have fair lending enforcement risk.

Goal of this step

LDA search establishes whether a change to the model is required – The search process may establish there is no less discriminatory alternative model. If so, the documented search process is the lender's business justification for disparate impact under ECOA.

LDA Search Methods

"Drop one" – Model variables that contribute most to disparate impact can be dropped / neutralized. This can be useful to explore scenarios but often results in a model that performs worse and therefore doesn't get adopted. Considering one variable at a time isn't as thorough as considering all the variables and how they interact.

Adversarial training – A more thorough algorithmic search can identify more practical alternatives. The optimizer is instructed to consider model accuracy and fairness simultaneously by defining a new objective function that combines losses from a system of iteratively trained models.

In the search for fairer models, “drop one” often leads to a decrease in predictive accuracy

For example:
Compliance flags total mortgage balances for review

RANK	FEATURE	IMPACT
1	Ratio of satisfactory trades to total	2.5%
2	Total high credit	2.4%
~~3~~	~~Total mortgage balances~~	~~1.6%~~
4	Sum of avg. balances for unclassified trades	1.5%
5	Total inquiries last 6 months	1.4%

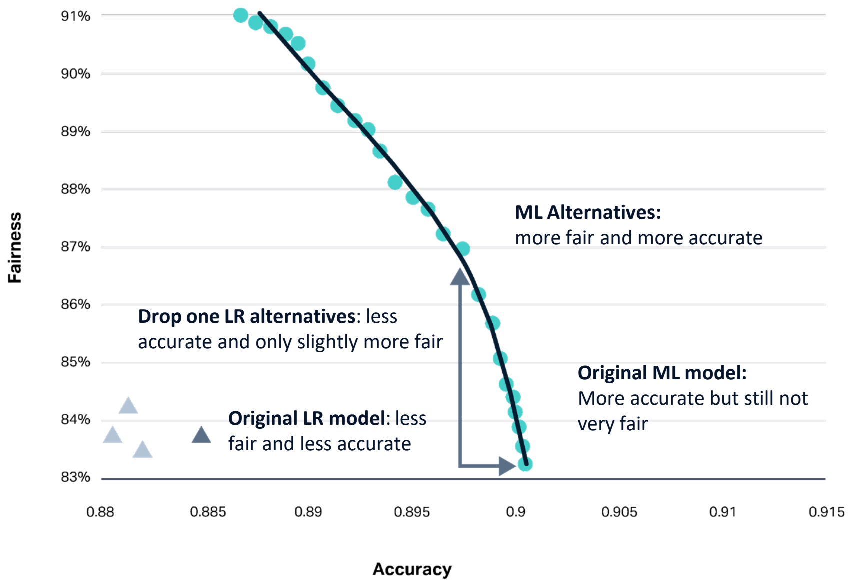
We neutralize the feature, and disparity & accuracy are recalculated

	CLASS	BEFORE	AFTER
Disparity (KS Score)	Hispanic	0.08	0.05
	African American	0.08	0.10
	API	0.08	0.05
	AIAN	0.08	0.22
Accuracy (Bad Rate)	All Borrowers	2.39%	2.48%

Neutralizing the feature reduced disparity but led to a **4% increase in bad rate**

A more thorough search can more effectively mitigate disparate impact and fair lending risk

Instead of omitting predictive variables, we can optimally and automatically adjust the influence of features causing disparate impact to generate a series of more fair models.



How it works

- Users can adjust "how fair" the model should be using a gain knob -- different values can be used to find the efficient frontier
- The efficient frontier provides lender options to manage trade-offs between accuracy and fairness

The status quo (drop -one) is inferior because:

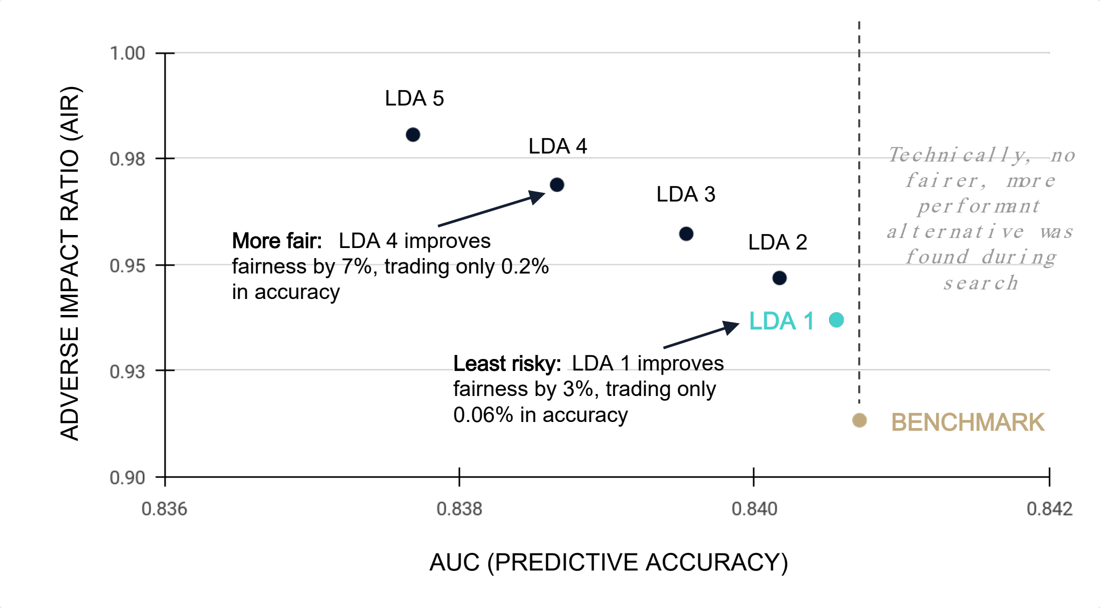
- It uses all-or-nothing approach (features are in-or-out rather than being attenuated)
- It uses a greedy feature-at-a-time process and constrains the search-space.
- Features can only be dropped! Others might need to be increased
- Tedious manual process prone to errors and lengthy timeline.

A more thorough search can more effectively mitigate disparate impact and fair lending risk

Identify fairer models
Identify alternative models that are more fair while maintaining the highest possible accuracy

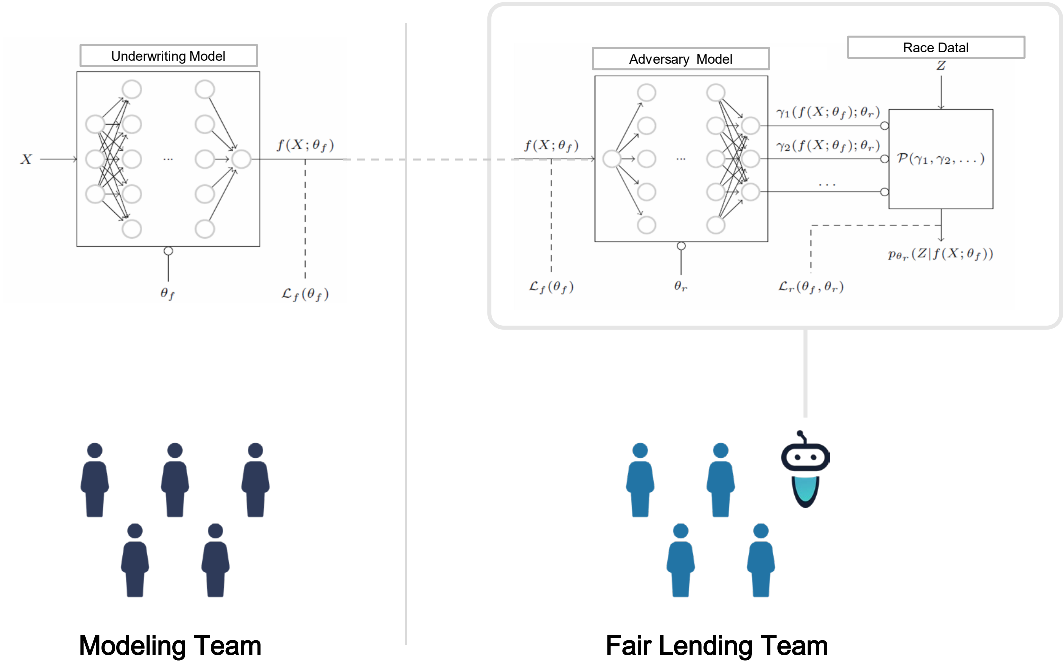
Ensure no less discriminatory alternative model exists
Banks can reduce enforcement risk by more thoroughly searching for less discriminatory alternatives and documenting the search process

LDA SEARCH WITH ADVERSARIAL DEBIASING



All of the LDA models are more fair and more accurate than drop -one alternatives

Adversarial Training is consistent with current practice and procedure in place today at banks



This is NOT using race as a model feature

- *Only* the adversary has access to race data.
- The adversary *never* communicates race data to the underwriting model.
- Instead, the adversary communicates the *correlation* between the model scores and race.

This does NOT weaken the wall between modeling and fair lending

- Fair lending lore requires a strict wall between modeling and fair lending teams.
- Zest's method doesn't weaken the wall. It only changes what is communicated between the two groups.
- Instead of saying "drop X variable," the fair lending team encourages modeling to drop variables OR change their importance.

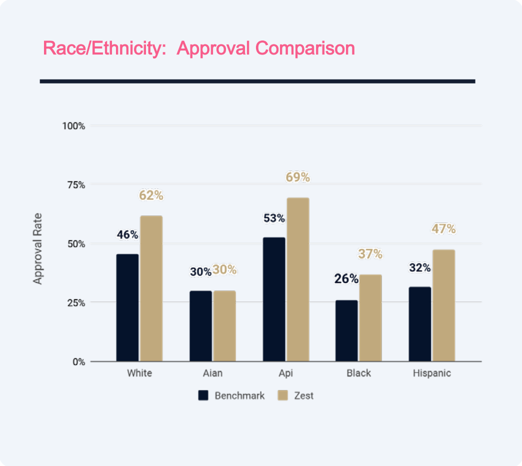
More fair alternatives include many minor adjustments to achieve more equitable outcomes, while still preserving predictive accuracy

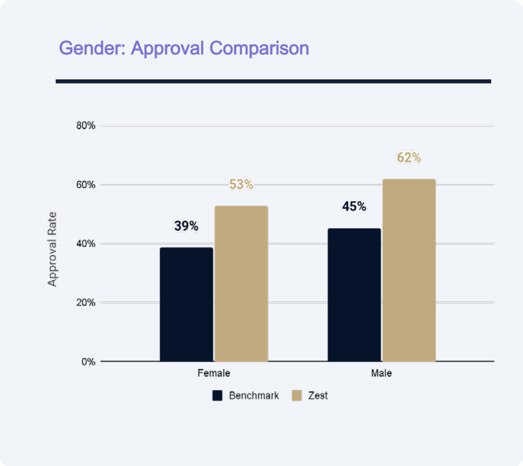
This process is impossible for a human to do manually but is easy for modern mathematics.

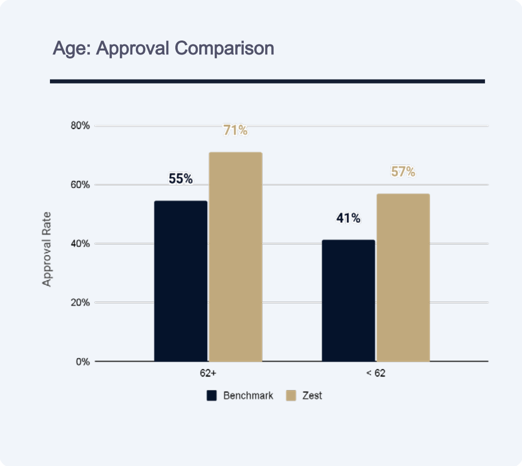
#	Feature Name	Feature Importance (%)	Absolute Difference (%)
1	Average Credit Limit	8.1%	~ 0.00%
2	Parent Listed as Co - Borrower	5.0%	- 0.01%
3	Average Payment Pattern Length	4.0%	- 0.02%
4	Max Number of Months Delinquent	3.3%	+ 0.05%
5	Max Delinquency Length	2.8%	- 0.40%
6	Max Credit Limit on any product	2.4%	- 0.11%
7	Total Credit Limit	2.2%	+ 0.32%
...	Hundreds more	-	-

Assessment Results: Approval rates

The model will increase HIspanic & Black approvals by +40% and Female by +36%







49% Increase in Hispanic approvals

41% Increase in Black approvals

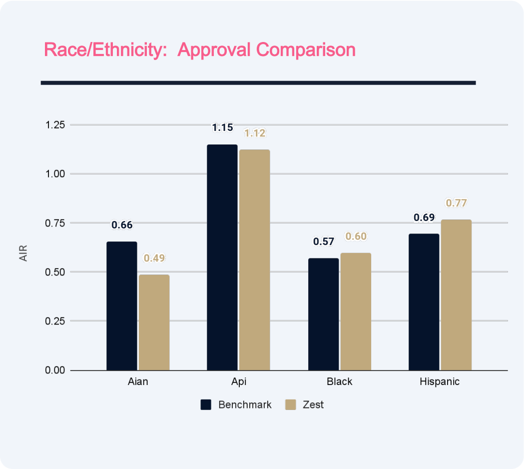
31% Increase in Api approvals

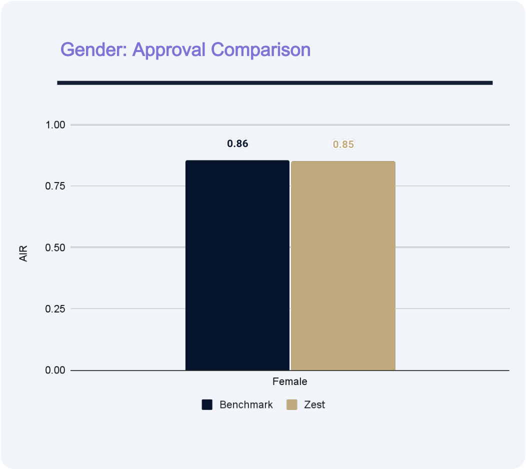
36% Increase in female approvals

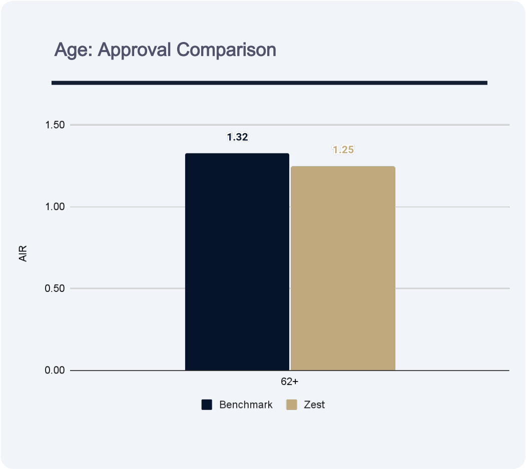
36% Increase in elderly approvals

Assessment Results: Adverse Impact Ratio (AIR)

The model will increase HIspanic & Black approvals by +40% and Female by +36%







11% Increase in AIR for Hispanic borrowers

5% Increase in AIR for Black borrowers

Minor degradation in AIR for Female applicants (within 99% CI)

No adverse impact for age

Compliant AI Lending

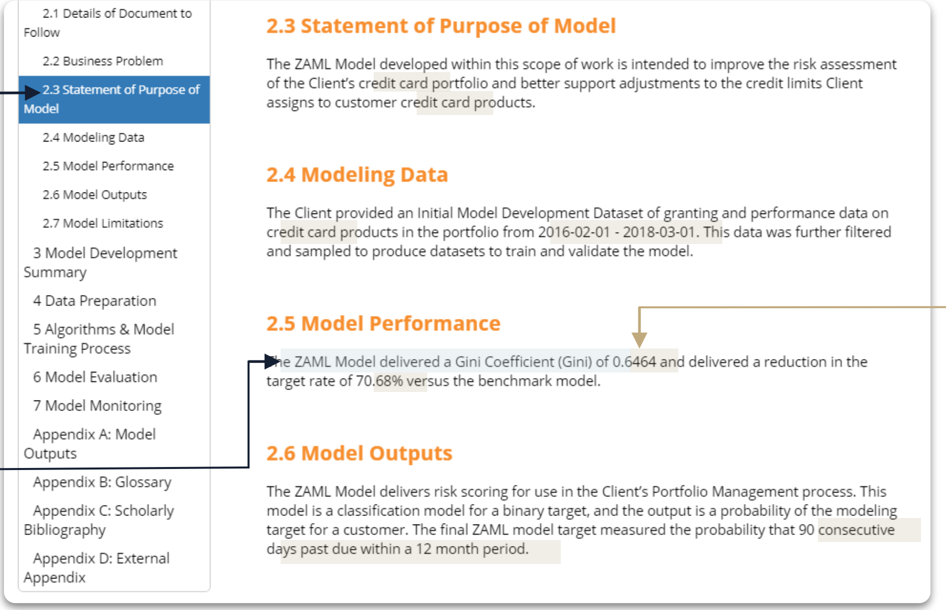
Pillar 3: Proper Documentation

Model Risk Management Documentation

Documentation is a critical component of model risk management. Standardized, fast, and easy -to-update documentation allows users to track every step in the modeling process and incorporate changes instantly.

Configs play many roles, including allowing you to hide / show / reorder different sections to suit your documentation needs.

Static text can be modified and will generally not change if the modeling use case (i.e. portfolio management) remains the same from project to project.



2.3 Statement of Purpose of Model

The ZAML Model developed within this scope of work is intended to improve the risk assessment of the Client's credit card portfolio and better support adjustments to the credit limits Client assigns to customer credit card products.

2.4 Modeling Data

The Client provided an Initial Model Development Dataset of granting and performance data on credit card products in the portfolio from 2016-02-01 - 2018-03-01. This data was further filtered and sampled to produce datasets to train and validate the model.

2.5 Model Performance

The ZAML Model delivered a Gini Coefficient (Gini) of 0.6464 and delivered a reduction in the target rate of 70.68% versus the benchmark model.

2.6 Model Outputs

The ZAML Model delivers risk scoring for use in the Client's Portfolio Management process. This model is a classification model for a binary target, and the output is a probability of the modeling target for a customer. The final ZAML model target measured the probability that 90 consecutive days past due within a 12 month period.

Dynamic inputs come from the model outputs and artifacts

All dynamic inputs in the screenshot have been highlighted.

Among other things, controllers pull in dynamic content.

INPUTS

- Dataset
- Exclusion List (based on EDA)
- Target variable
- Development sets
- Benchmark values

AUTOMATIC ARTIFACTS

- Model-ready new and excluded features by transformer
- Model pipeline flowchart
- Algorithm type and default parameters
- Updated model parameters
- List of features by importance by algorithm
- Ensemble weights
- List of features by importance for the ensemble
- Statistical performance metrics
- Statistical performance charts
- Economic analysis metrics and projected impact
- Partial Dependence Plots
- Attribution Plots
- Individual score attributions

CUSTOM ARTIFACTS

- Model details: End user, use case, product, objective
- Custom transformer description

Fair Lending Documentation

A fair lending report incorporating the above analysis is essential to document compliance

Fair Lending Review

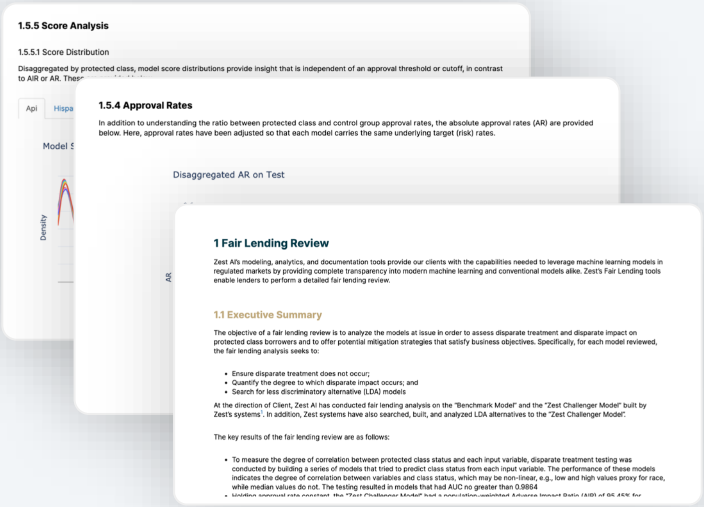


Table of Contents

- ✓ Fair Lending Techniques & Tools
- ☐ Protected Class Information
- ☐ General Statistical Comparison
- ☐ Adverse Impact Ratios
- ☐ Approval Rates
- ☐ Score Analysis
- ☐ Disparate Treatment Evaluation
- ☐ Disparate Impact Analysis
- ☐ LDA Search and business justification

THANK YOU

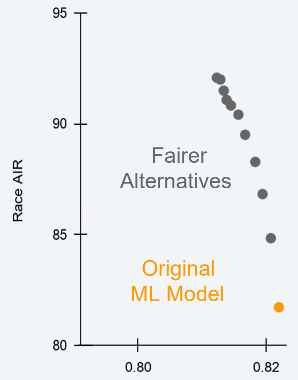
Theodore ("Teddy") Flo
Teddy.Flo@Zest.AI

Thank You

Appendix 1:
Open Questions About AI and Fair Lending

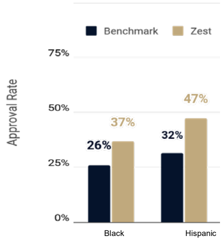
Decisions remain:

Which LDA to pick?

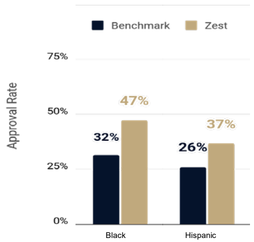


- Trading accuracy for fairness is *not* required today. (But may be soon.)
- It's your decision; be principled about it; consult your own attorneys.
- Consider the following:
  - Existing accuracy standards, and
  - Accuracy loss or trade -offs you accept elsewhere.

What if protected groups "clash"?



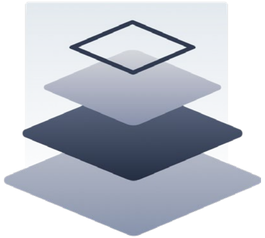
or



- What if some LDAs are better for different groups?
- Only consider LDA models where every group that had an AIR below 100% in the original model is better off (or at least not worse off).
- Then, be principled. Consider the following:
  - Protected group population sizes
  - AIR, true positive, false positive, of the LDAs
  - Margin of error for different protected groups
- The advice of your own counsel.

Appendix 2
Recent Developments

A Few Recent Developments



Fair lending law and policy are in a state of flux right now, with the White House, CFPB, and other regulators driving for increases in fairness and inclusivity in lending, with a focus on Machine Learning and Artificial Intelligence (and a response by fintechs and responsible AI practitioners)

Use of "Unfairness" to Police Fair Lending

- CFPB announced its use of the unfairness UDAAP framework to police fairness in lending
- This leaves numerous open questions for fair lending compliance
- This has been challenged in court. It's application is likely to be delayed

CFPB Circular 2022 - 03

- On May 26, 2022, CFPB released a circular on the importance of providing accurate NOAAs
- Shortly thereafter, Zest AI produced a whitepaper discussing ML explainability compliance

Increased Risk of Relying on Medical Debt in Underwriting

- A March 1, 2022, CFPB report calls into question predictive accuracy of medical debts
- Beginning July 1, 2022, the three main credit bureaus will stop reporting medical collection debt

FinRegLab / Stanford Study Shows AI Lending Can Be Fair, Profitable, and Compliant

- In April 2022, FinRegLab released a report on AI in Lending. Adversarial debiasing found to be the most effective in de-biasing AI underwriting models

Let’s keep moving the conversation forward

May 2018:
First FDIC-chartered bank accepts ML originations model built using Zest tools to run their $7B personal loans business

Aug 2018:
Zest AI discusses model risk management in an ML era with Senior Fed Supervisor David Palmer, a key architect of the interagency guidance on model risk management (Fed SR 11-7/OCC 2011-12)

Sep 2018:
Zest AI releases its FAQs on MRM Guidance, which is circulated within the Fed, OCC, NCUA, and FDIC

Oct 2018:
Zest AI briefs Dilip Patro (FDIC), another key architect of the MRM guidance

Apr 2019:
Zest AI briefs the Fed, OCC, and FDIC supervisors, CFPB office of fair lending enforcement, and FHFA office of fair lending oversight on the responsible use of AI in lending

Jun 2019:
Zest AI invited to testify to the House Financial Services Committee AI Task Force

Aug 2019:
Freddie Mac announces its long-term partnership with Zest AI

Oct 2019:
VyStar Credit Union announces its partnership with Zest AI

Dec 2019:
Zest AI invited to present to FDIC, OCC, and Fed examiners on how to examine ML models; 100s of bank examiners attend the presentation

Oct 2020:
A customer's ML model built using Zest tools passed OCC examination and is being use to underwrite a $6B credit card portfolio

At the invitation of the CFPB, Zest AI participates in the bureau’s first tech sprint on notice of adverse actions, demonstrating how naive methods are almost always wrong when run on ML models to determine adverse action reasons

Nov 2020:
Zest AI briefs the CFPB Office of Fair Lending Enforcement on recent innovations in fair lending analytics

Dec 2020:
Zest AI presents its fair lending technology to OCC Project REACh Homeownership Working Group

Apr 2021:
Zest AI selected by FinRegLab, Stanford GSB for inclusion in its study to document how ML models meet regulatory requirements