



EXPERIAN. learning DAY 2022

Accelerate and maximize your data potential

EXPERIAN **learning**DAY₂₀₂₂

Accelerating your strategy using data and analytics



Richard Adamski

Senior Analytical Consultant
Experian





Sigma Bureau Scoring Suite **Experian's latest consumer bureau scores**

Suite of next-generation scores and our non-traditional credit score for thin files

Sigma Suite of Scores



Sigma Standard

A general-purpose score used to give an overall risk view of an applicant or customer



Sigma Unsecured Credit

Tailored for the Unsecured credit market



Sigma Banking Finance

An acquisition risk prediction score specifically modelled for the banking industry.



Sigma Retail Credit

Sigma Retail Credit provides risk prediction unique to the retail space



Sigma Customer Management

Developed to identifying any early signs of distress



Sigma Transcend

Non-Traditional thin file score developed using latest machine learning techniques

Why Sigma Scores?

**“Best
of both”
data**

**New
diverse
variables**

**Simplify
the solution**

**Minimise
score
inconsistenc
ies**

**Suite of
Scores**

**New BIQ
platform and
BI Scores**



Sigma Banking Finance & Retail Credit Results



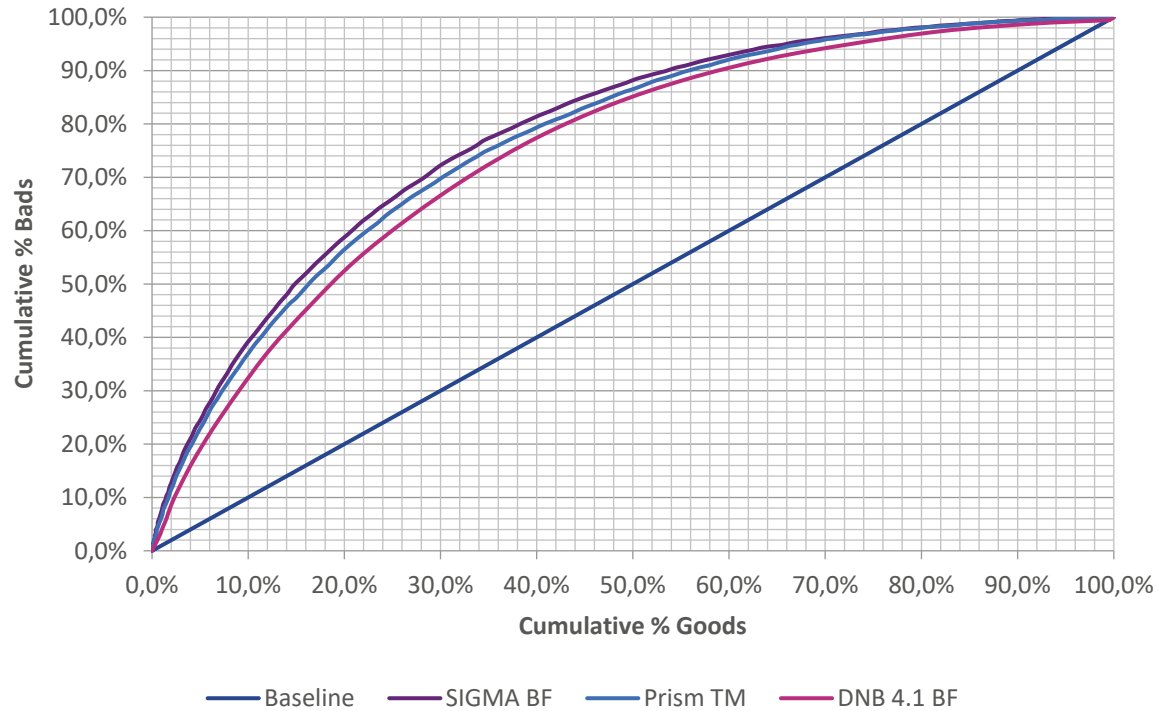
Consumer Sigma Results – Gini

Gini Strength: Sigma significantly outperforms Delphi and Prism

Sigma BF Gini = 56.35
Delphi BF 4.1 Gini = 49.38
Prism TM Gini = 53.85

Gini Chart
Sigma BF Model

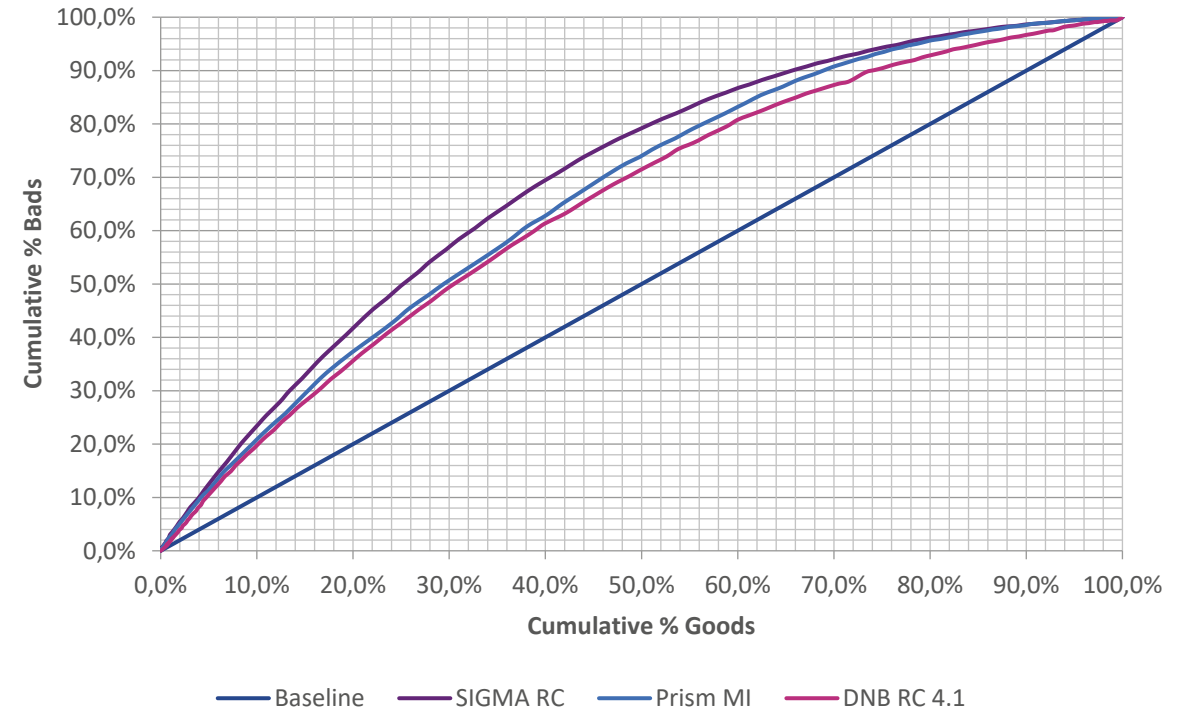
Uplift = 7%



Sigma RC Gini = 39.31
Delphi RC 4.1 Gini = 29.28
Prism MI Gini = 33.18

Gini Chart
Sigma RC Model

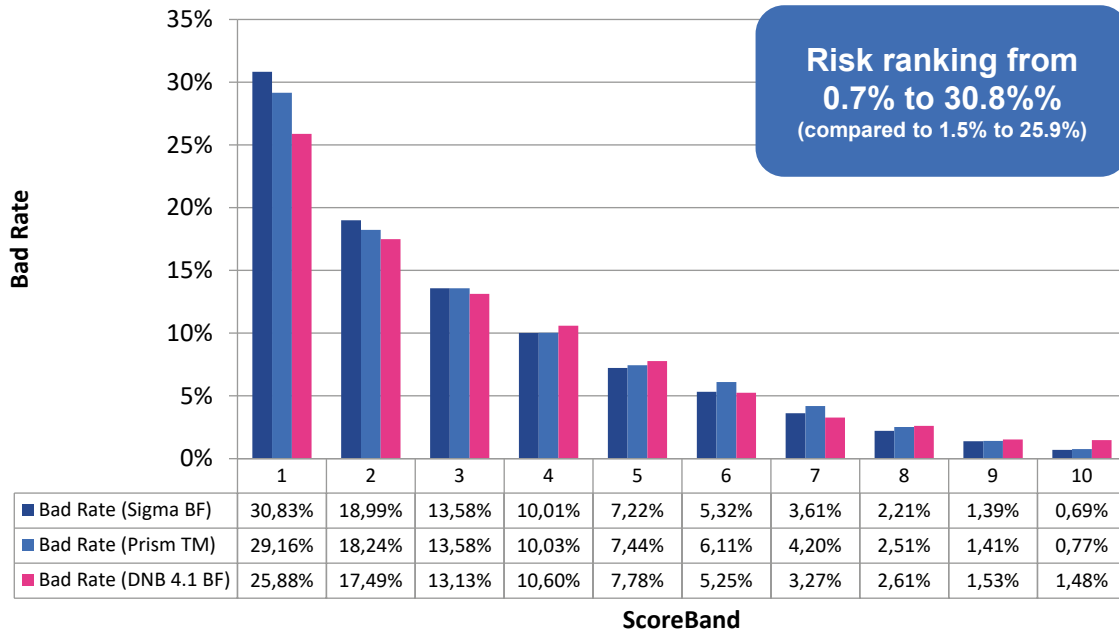
Uplift = 10%



Consumer Sigma Results – Ranking

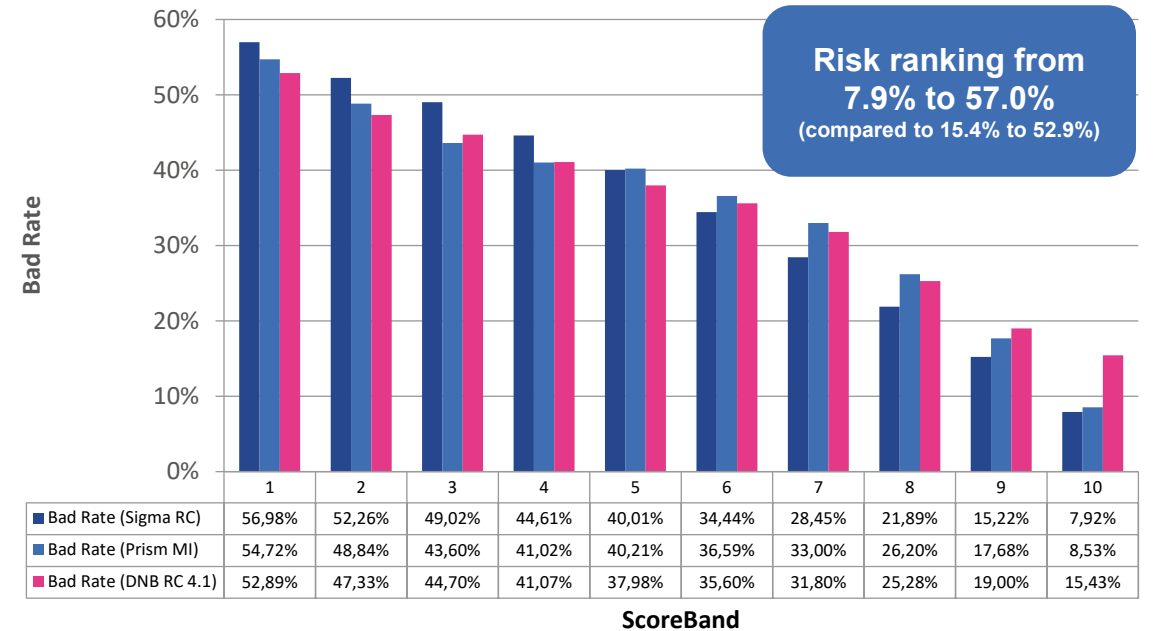
Ranking: Sigma significantly outperforms Delphi and Prism

Bad Rate by Decile Sigma BF Model



■ Bad Rate (Sigma BF) ■ Bad Rate (Prism TM) ■ Bad Rate (DNB 4.1 BF)

Bad Rate by Decile Sigma RC Model



■ Bad Rate (Sigma RC) ■ Bad Rate (Prism MI) ■ Bad Rate (DNB RC 4.1)

- Sigma Scorecards rank risk well
- Sigma has higher risk in the lower deciles and lower risk in the upper deciles

EXPERIAN
learningDAY₂₀₂₂

Sigma Customer Management Results



Consumer Sigma Results – Gini

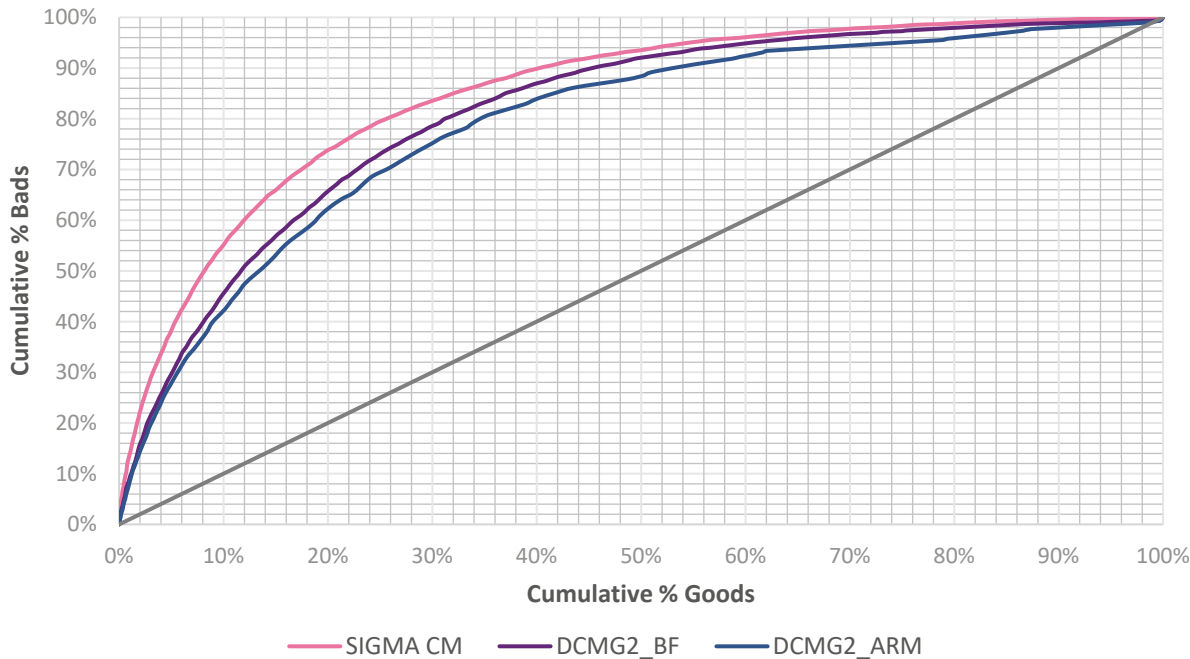
Gini Strength: Sigma Customer Management, with no segmentation, outperforms Delphi on both Banking & Finance and Retail Credit Portfolios

Banking and Finance market

Sigma CM Gini = 69.7
 Delphi CM2 BF Gini = 63.7
 Delphi CM2 ARM Gini = 58.6

Gini chart
 Sigma CM

Uplift = 11%

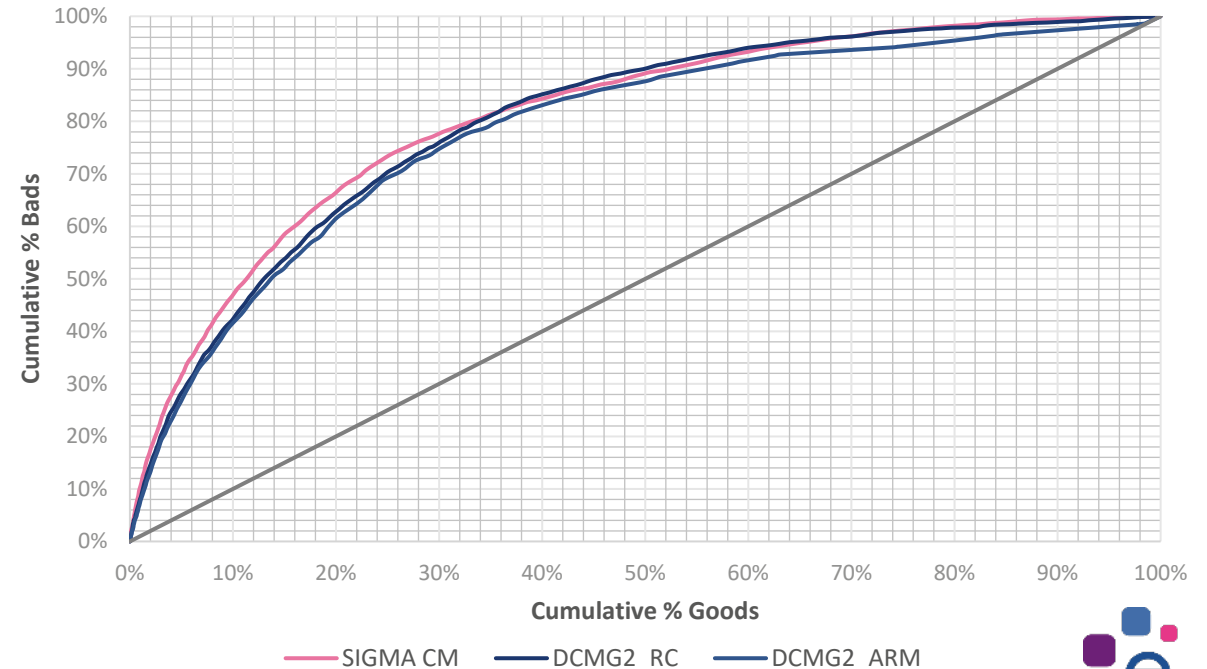


Retail Credit market

Sigma CM Gini = 62.0
 Delphi CM2 RC Gini = 61.0
 Delphi CM2 ARM Gini = 58.0

Gini Chart
 Sigma CM

Uplift = 4%



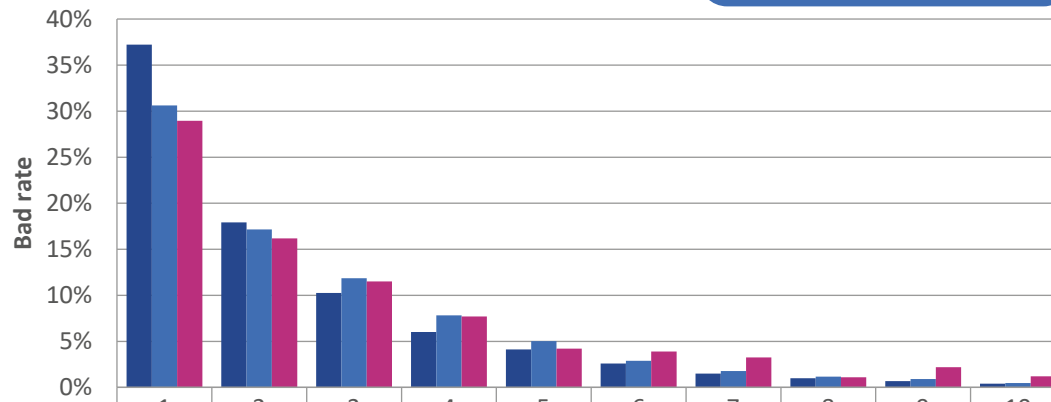
Consumer Sigma Results – Ranking

Ranking: Sigma Customer Management outperforms Delphi

Banking and Finance market

Bad rate by decile Sigma CM

Risk ranking from 0.4% to 37.23% (compared to 0.47% to 30.63%)



	1	2	3	4	5	6	7	8	9	10
SIGMA CM	37,23%	17,93%	10,25%	6,01%	4,12%	2,59%	1,51%	1,00%	0,69%	0,40%
DCMG2_BF	30,63%	17,15%	11,86%	7,83%	5,02%	2,88%	1,78%	1,16%	0,92%	0,47%
DCMG2_ARM	28,96%	16,19%	11,51%	7,70%	4,21%	3,90%	3,26%	1,10%	2,19%	1,21%

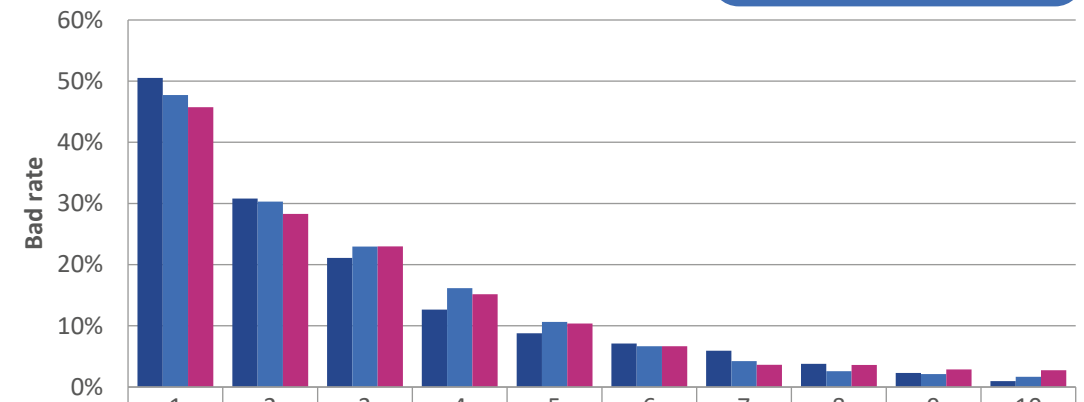
Score band

■ SIGMA CM ■ DCMG2_BF ■ DCMG2_ARM

Retail Credit market

Bad rate by decile Sigma CM

Risk ranking from 0.95% to 50.52% (compared to 1.67% to 47.72%)



	1	2	3	4	5	6	7	8	9	10
SIGMA CM	50,52%	30,79%	21,10%	12,66%	8,78%	7,12%	5,93%	3,79%	2,30%	0,95%
DCMG2_RC	47,72%	30,30%	22,96%	16,15%	10,64%	6,66%	4,24%	2,58%	2,11%	1,67%
DCMG2_ARM	45,74%	28,29%	22,97%	15,16%	10,37%	6,66%	3,63%	3,60%	2,88%	2,75%

Score band

■ SIGMA CM ■ DCMG2_RC ■ DCMG2_ARM

- Sigma Scorecards rank risk well
- Sigma has higher risk in the lower deciles and lower risk in the upper deciles

EXPERIAN
learningDAY₂₀₂₂

Tackling the Credit Invisibles

Experian's Sigma Transcend Score

Non-traditional credit score for thin files



Identifying the underserved

Understanding the need for a thin file score



Credit Invisibles



Stricter Credit Policies



Thin-file Score



Machine Learning

EXPERIAN
learningDAY₂₀₂₂

Data Sources



Alternative Data Sources



Geocoded Address Data

- 25M Consumers 18-65yrs
- Industry leading data cleaning and enrichment
- Accuracy to house level for most addresses
- 20-year history



Telephone Number Data

- 28M Consumers 18-65yrs
- Home, work and cell data
- Industry leading data cleaning and enrichment
- 20-year history



Current Pinpoint Data

- Deeds
- CIPC
- Demographics



Experian D2C

- Location
- Handset type
- Data usage
- App installations

Current Model (v1)

Medium-term (v2)

EXPERIAN
learningDAY₂₀₂₂

Thin File Results



Thin File Results

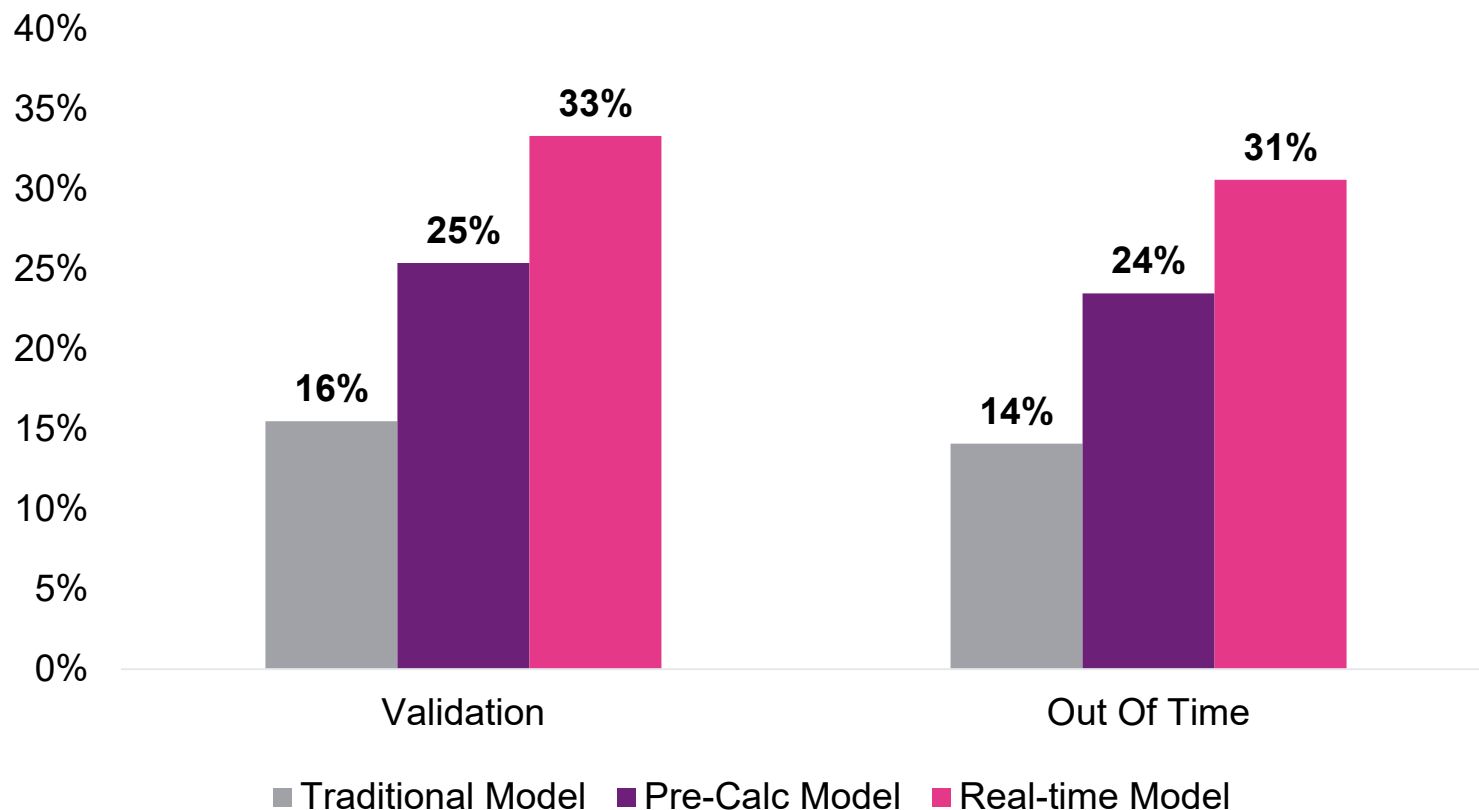
Sigma Transcend provides significant lift over a Traditional Thin File score

'Pre-calc mode' refers to running the model on already reported (historical) data and does not require consumer's contact data from the credit provider at the point of application.

'Real-time mode' requires the credit provider to submit the consumer's address and contact details before calling the scoring API.

This approach provides the most up-to-date information and boost the model performance.

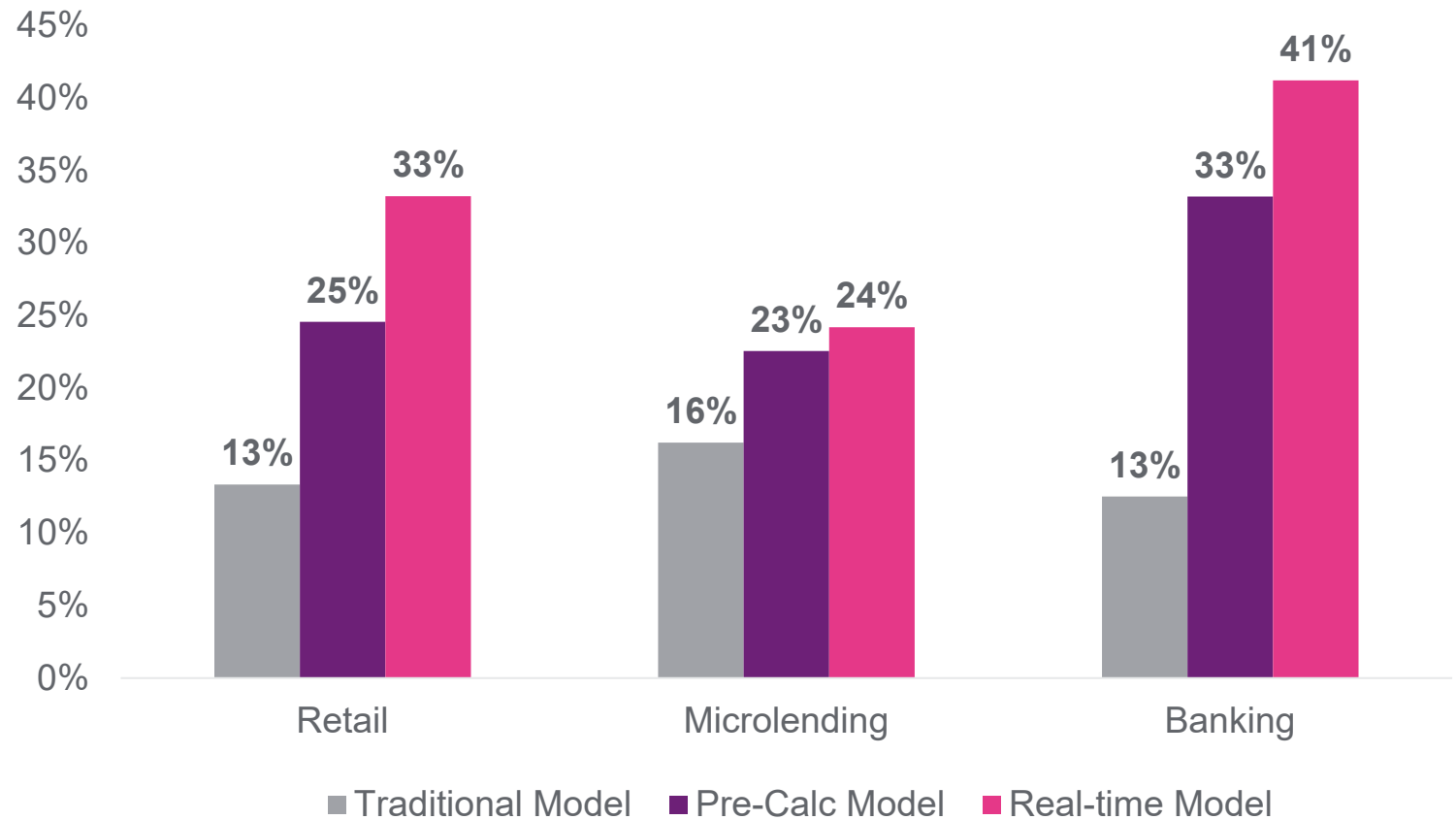
Gini Comparison



Thin File Results by Industry

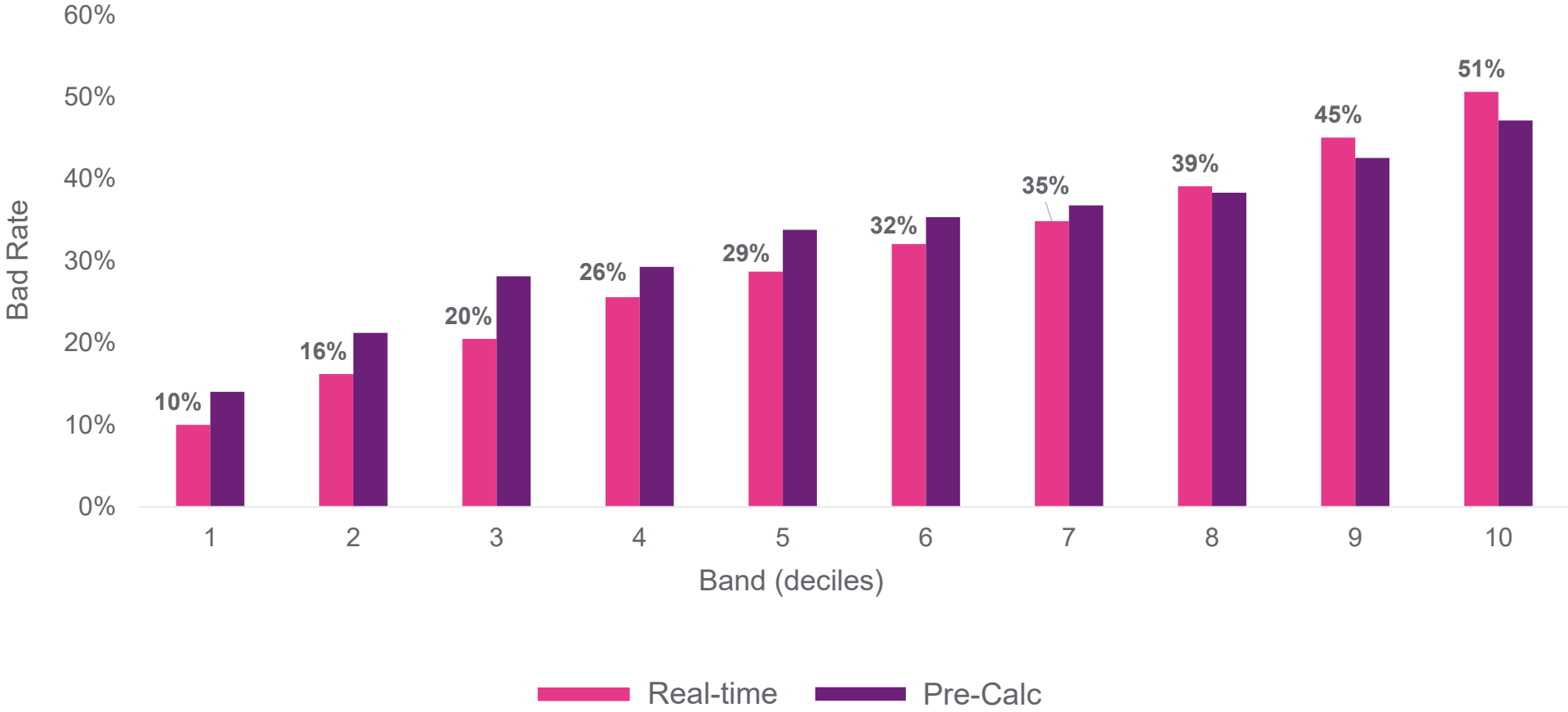
Sigma Transcend provides significant lift for all segments of the market

Gini Comparison by Industry



Bad Rates by Decile

Bad Rate By Decile - Overall



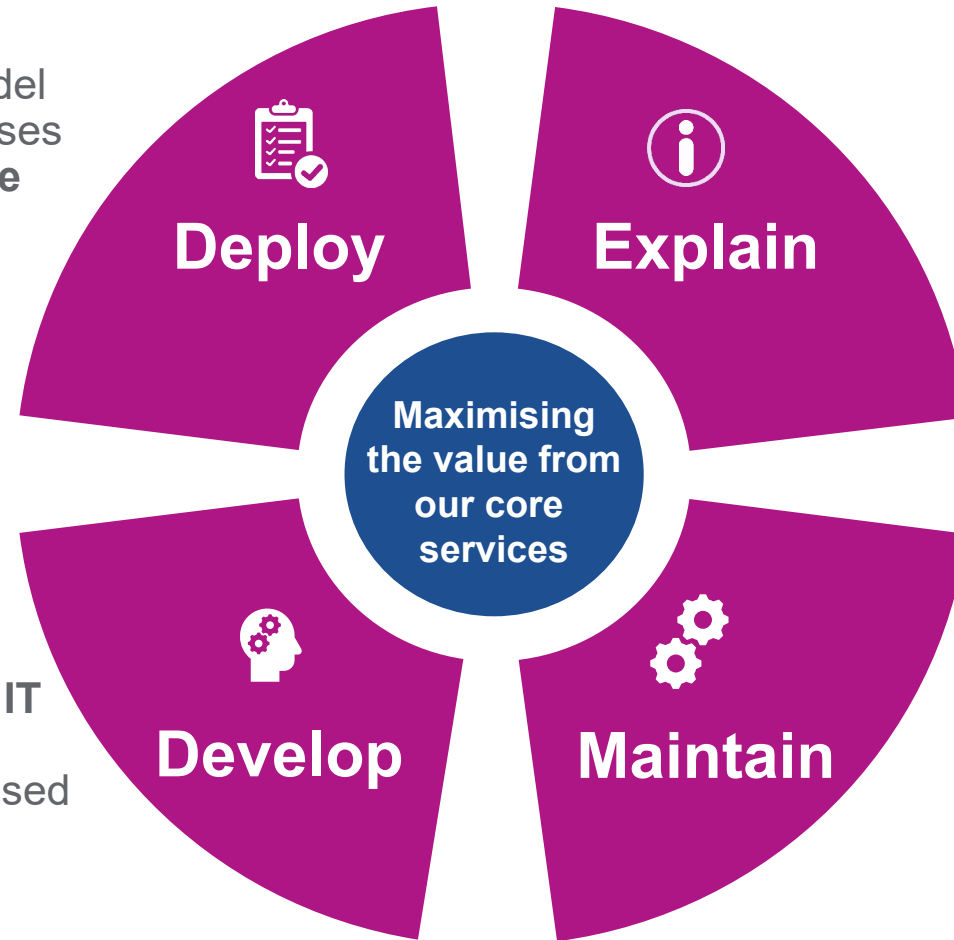
EXPERIAN
learningDAY₂₀₂₂

Machine Learning Explainability



Four Core Principles to Focus on

Develop complex ML model operationalisation processes to **maximise ROI, reduce cost and risk.**



Ensure explainability, data traceability, fairness and avoidance of bias capabilities are **part of solution.**

Data access, quality, analytics expertise and IT availability core to the development of data focused solutions.

Make tracking and monitoring models, ongoing validation and maintaining ongoing stability **part of core value proposition.**

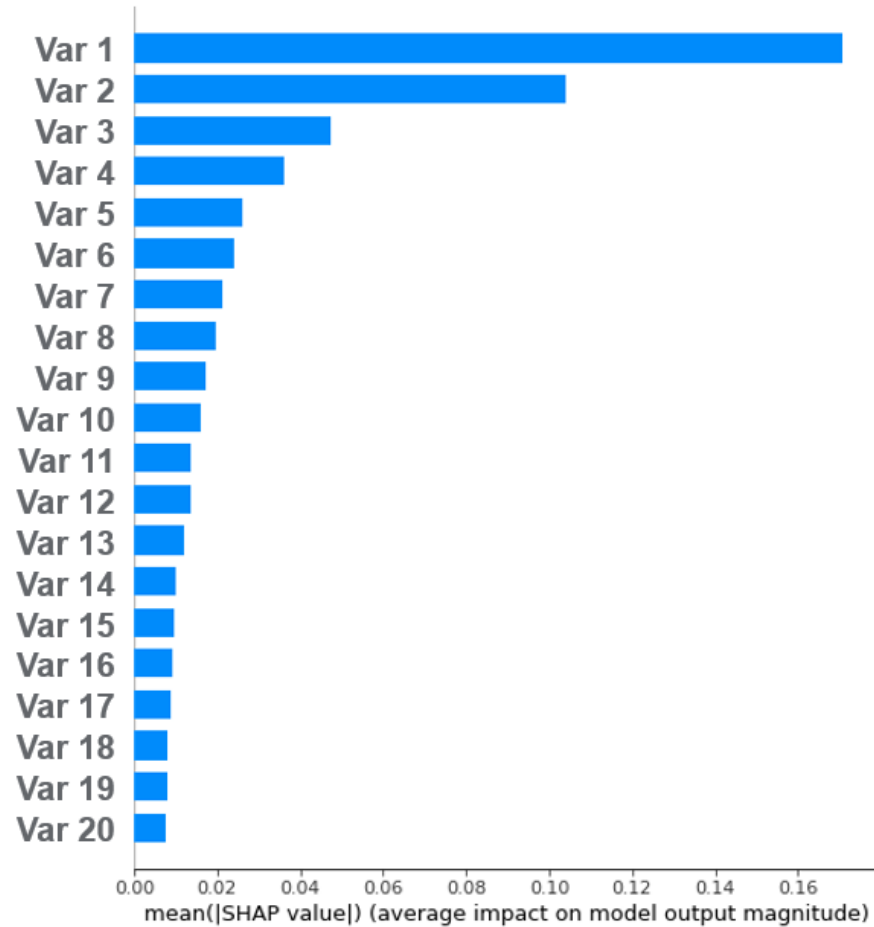
EXPERIAN
learningDAY₂₀₂₂

Model Transparency



Variable Contribution

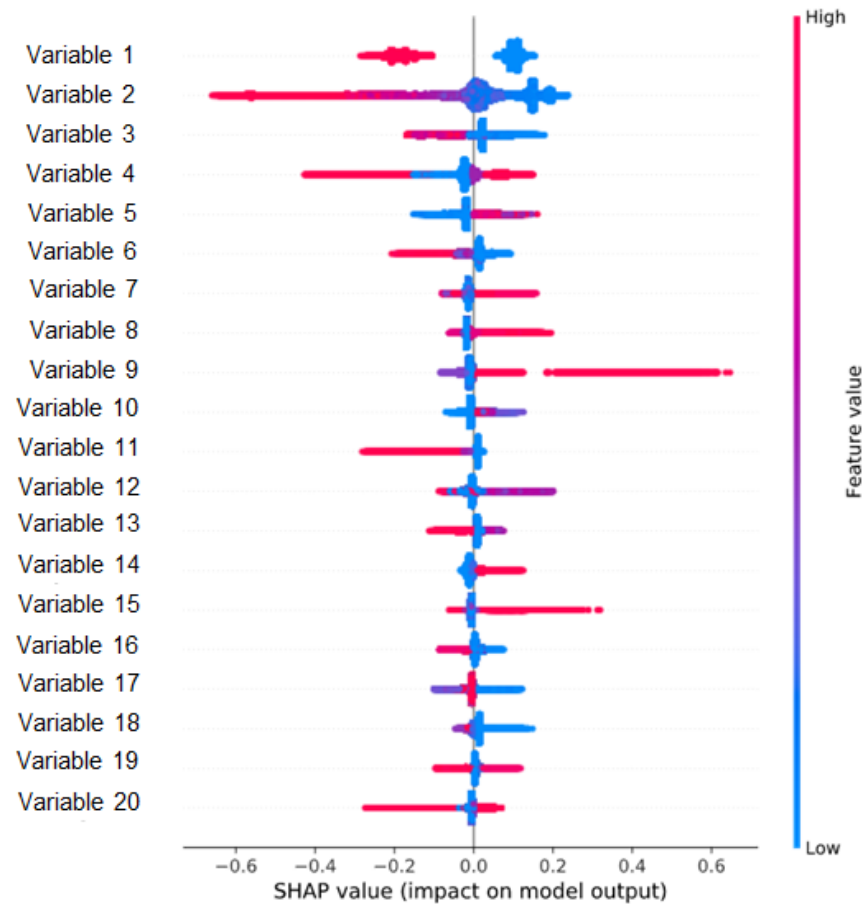
The top 20 contributing variables are shown below:



SHAP values give us a complete transparency and a good understanding of the workings of the model. One can assess the variable contribution to the model (similar to marginal contributions in Logistic Regression)

Variable Discrimination

The top 20 contributing variables are shown below:



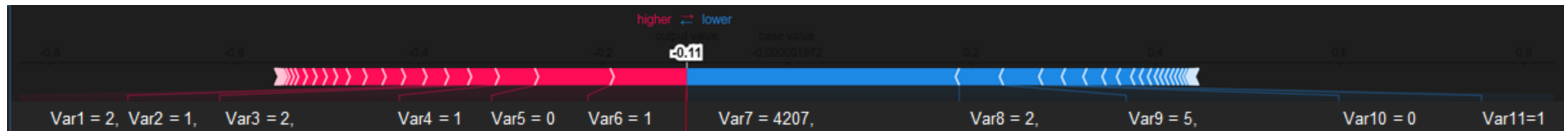
SHAP values show us the discriminatory power of each variable and allows us to sense check the logic of the trend against business rationale

Customer Level Explainability

An example of a scenario where a customer is higher risk (red indicates high risk contributors, while blue is low risk)



An example of a scenario where a customer is lower risk (red indicates high risk contributors, while blue is low risk)



Questions?

Contact Us



Marketingafrica@experian.com



www.experian.co.za

