# Open Banking current account information has the power to transform credit scoring and lending decisions

A Case Study

# Agenda

- What is Open Banking?

- Overview of the case example and available data

- Summary of the challenges

- Our approach to developing an Open Banking credit score

- Summary of the results

- Implications for the industry

cantab
predictive intelligence

TRENT
Advisory Services

# What is Open Banking?

Open Banking is the secure way to give providers access to your financial information *

- **The EU Revised Payment Services Directive (PSD2)** and national regulations such as the UK's Open Banking are designed to drive greater transparency, security, innovation and market competition

- It enables bank customers to use third-party service providers to carry out various activities based on their financial data, with their explicit consent

- This requires banks to provide access to customer account data and/or initiate payments to these providers.

Consumer lending is one of the main use cases for Open Banking, where customers can gain better access to credit when sharing their data.

* Source: UK Open Banking Implementation entity

cantab
predictive intelligence

TRENT
Advisory Services

# Open Banking provide a major opportunity to enhance credit scoring significantly

Our project has shown that credit scorecards using using Open Banking data alone have Ginis above 70%

Access to banking transaction data now puts all lenders on a similar level: all lenders now know as much about the customer as his/her own bank does

Based on our findings, we argue that most lending organisations operating in an Open Banking environment have a tremendous opportunity to enhance their credit decisioning capabilities significantly

cantab
predictive intelligence

TRENT
Advisory Services

# The project background and result

**The challenge**

- To develop a predictive behavioural credit scoring model, by using only data which will be realistically available through Open Banking banking API

- To take into account current level of uncertainty related to Open Banking implementation in specific markets
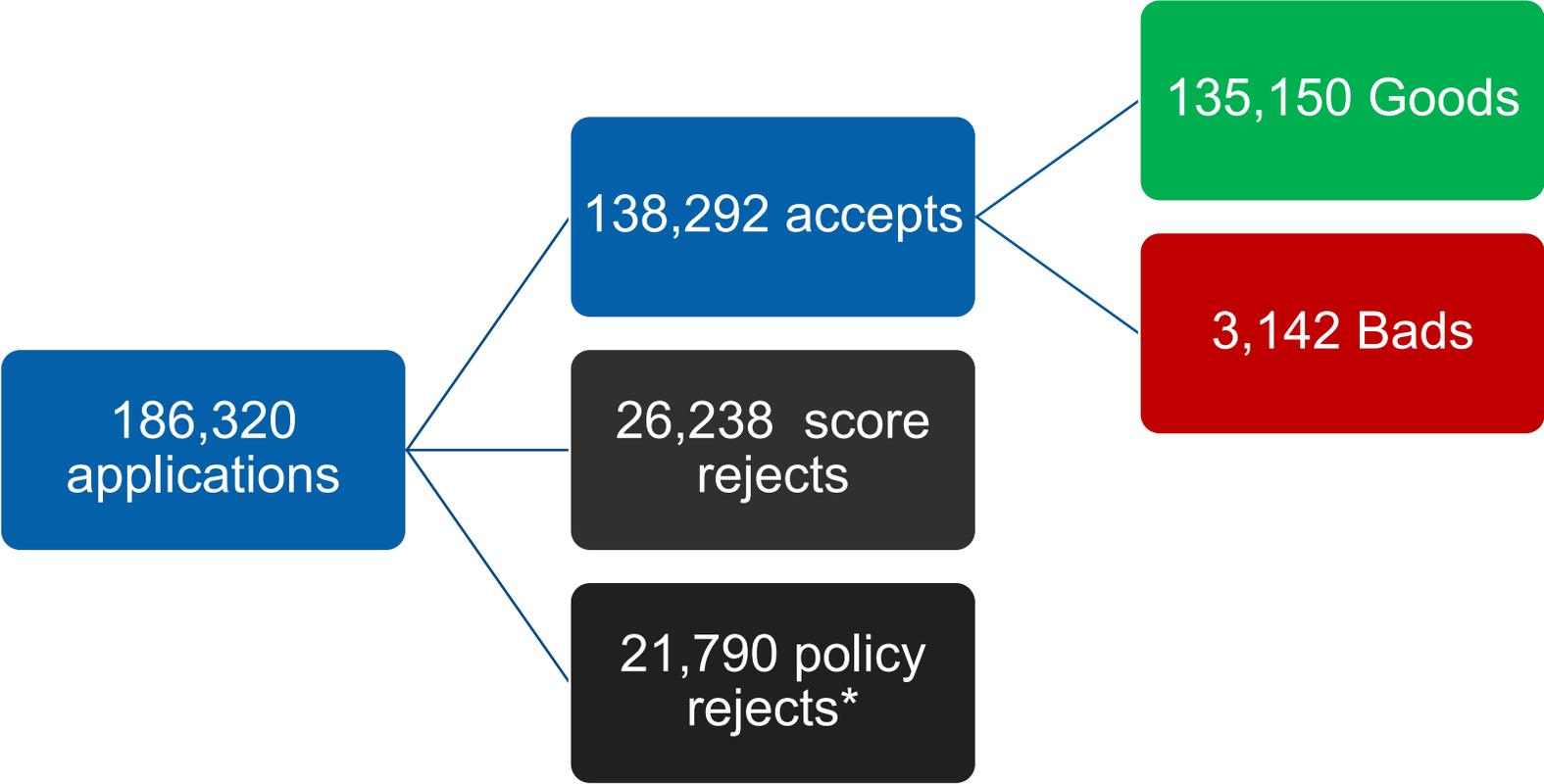
**The outcome**

- Developed behavioral credit scoring model with accuracy of Gini of more than 70%

- Multiple models developed for most realistic Open Banking scenarios in the specific market, with good results in all the cases

cantab
predictive intelligence

TRENT
Advisory Services

# The scoring model is based on 6 months of C/A transactions

| | Data source | Description |
|---|---|---|
| **Conservative scenario** | • **Current account transactions** | • 6 months history prior to application for cash loan, amounts and balance only |
| | • Default (yes/no) | • Info if the customer defaulted during the observation time window, according to the definition used by the bank |
| | • Application data (for reject inference) | • Data on all the applicants: info on policy vs. scoring reject rationale; PD calculated by the legacy scoring model |
| **Realistic / Optimistic Scenarios*** | A combination of the following*: | |
| | • Demographic data | • Age |
| | • Approved overdraft | • Amount of approved overdraft at the day of applying for the cash loan |
| | • Length of banking relationship | • Year of opening of the current account |

*\* This data was used only in for testing purposes. The deployed model is using data from the Conservative scenario*

cantab
predictive intelligence

TRENT
Advisory Services

# Our data set was large enough to allow detailed feature analyses

186,320 applications

138,292 accepts

135,150 Goods

3,142 Bads

26,238 score rejects

21,790 policy rejects*

*Policy rejects were excluded from scorecard development*

cantab
predictive intelligence

TRENT
Advisory Services

# To develop the model, our research combined best practice data science analytics and credit scoring analytics

Created over 3,000 features, ie composite variables created from the raw transaction data

20 different scenarios were evaluated in details

For each scenario, we trained 40 models with different test/train splits to determine stability of error (bootstrapping)

Developed several hundreds of codes, routines and visualisations in Tableau, Python and R

6 different machine learning techniques used, including logistic, linear regression, gradient boosting and neural networks

cantab
predictive intelligence

TRENT
Advisory Services

# We created 3000+ of „features" (composite variables) from the transactions dataset

## Challenge

- The data differs substantially from the ones used in traditional credit scoring

- Data contains only low-level features, such as individual transactions

- This is comparable to the challenge that machine learning algorithms are now successfully addressing when trained to recognize the content of an image: their input consists of low-level features – in this case, values of pixels of an image

## Approach

Created 3000+ aggregate features

- **Simple variables**: minimal, maximal, average transaction,...

- **Aggregate variables**: e.g. max number of days without income, etc.

- **Statistical measures**: median, average, standard deviation of transactions of a certain type

- **Complex variables**, generated by specialized packages (e.g. Fourier transforms of the time series)

cantab
predictive intelligence

TRENT
Advisory Services

# The most important variables in the model are often related to variance / deviation of the observed behaviour

**The most important variables for one of the best performing models (Example)**

- **6months_Current_balance_std_min** - Minimum of monthly standard deviations of current balance in last 6 months prior to application

- **6months_Current_balance_min_max** - Maximum of monthly minimums of current balance in last 6 months prior to application

- **6months_I_mean_min** - Minimum of monthly average outgoing payment in last 6 months prior to application

- **4week_size_of_transactions2week** – Number of transactions in 2 weeks prior to application

- **6months_avg_CTO_minus_DTO_ratio** - Difference between sum of all incoming and outgoing transactions by month: avg of last 3 months / avg of 3 months before

- **pilot_below_11_net_mean** - Average of transactions below xx Eur in the last 6 months prior to application

- **I_max3** – Maximum of outgoing payments in last 3 months prior to application

# To deal with the lack of standards, we modelled a number of different scenarios

## Challenge

- At the time of modelling, no agreed standard accross EU on data to be shared, as well as in multiple markets

- For e.g. Berlin Group standards, a large variability of possible interpretations

- The credit scoring models built for any specific market need to be flexible enough to cope with multiple future scenarios

### Base and Composite scenarios

| Variables used | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| Base | ● | ● | ● | ● | ● | ● | ● | ● | ● | ● |
| Tsfresh* | | | | | | ● | ● | ● | ● | ● |
| Age | | ● | | | | | ● | ● | ● | ● |
| City | | | ● | | | | | | | |
| Overdraft* | | | | ● | | | | | ● | ● |
| Gdays* | | | | | ● | | ● | | ● | |

cantab
predictive intelligence

TRENT
Advisory Services

# To deal with the noise and correlations, we compared multiple machine learning techniques and chose the best

**Challenge**

- Irregularity in current account transactions over time

- It is thus difficult to identify specific information such as salary

- In addition, information extracted from current account transactions in the form of specific features or variables is extremely correlated
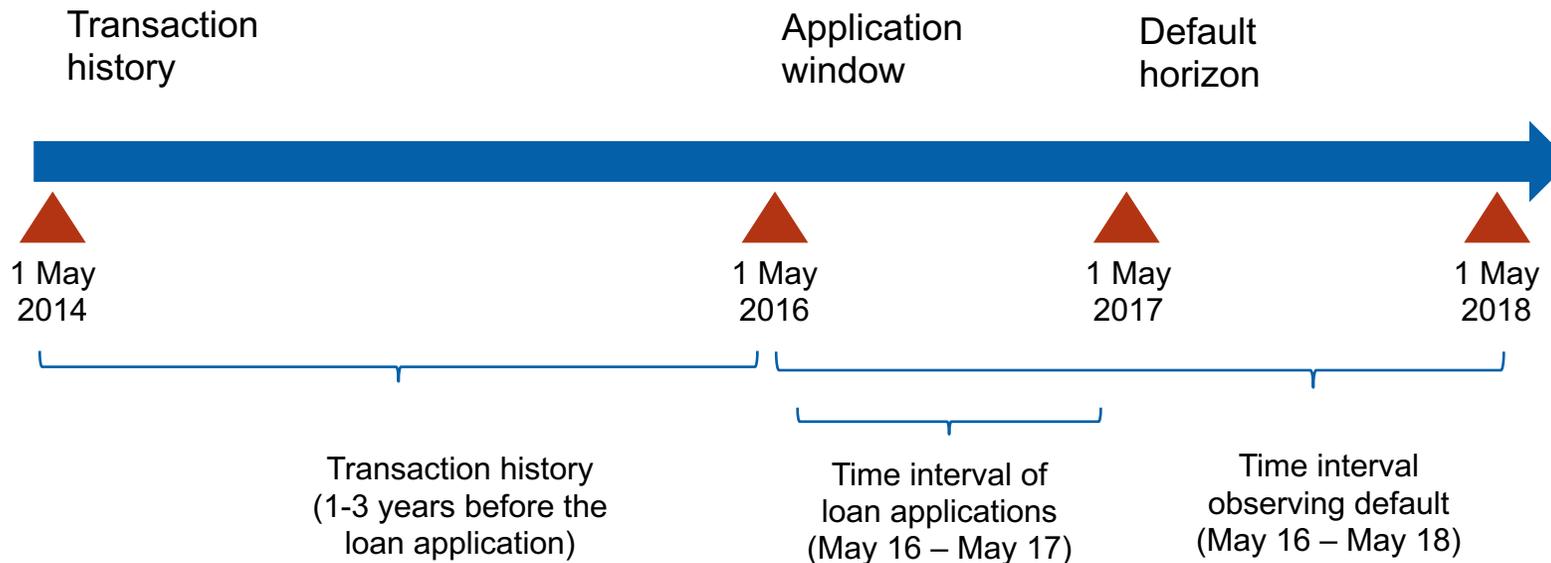
**Compared machine learning models**

- Logistic regression

- Random forests

- Support-vector machine

- Extreme gradient boosting (XGB)

- Neural networks

The best performing on the test set, deals well with correlated variables

cantab
predictive intelligence

TRENT
Advisory Services

# We implemented the approach in line with the practices of the bank

**Considered time windows**

**Definition of "bad"**

Transaction history

Application window

Default horizon

1 May 2014

1 May 2016

1 May 2017

1 May 2018

Transaction history (1-3 years before the loan application)

Time interval of loan applications (May 16 – May 17)

Time interval observing default (May 16 – May 18)

- At lest Euro 250 of overdue payment for more than 90 consecutive days

cantab
predictive intelligence

TRENT
Advisory Services

# Summary of the approach – reject inference

We carried out reject inference using a combination of:

- Previous score (PD) used for decisioning

- Accepts performance by score range

- Behaviour on rejects who had another credit product

**186,320 applications**

**138,292 accepts**
- 135,150 Goods
- 3,142 Bads

**26,238 score rejects**
- 20,944 inferred goods
- 5,294 inferred bads

**21,790 policy rejects***
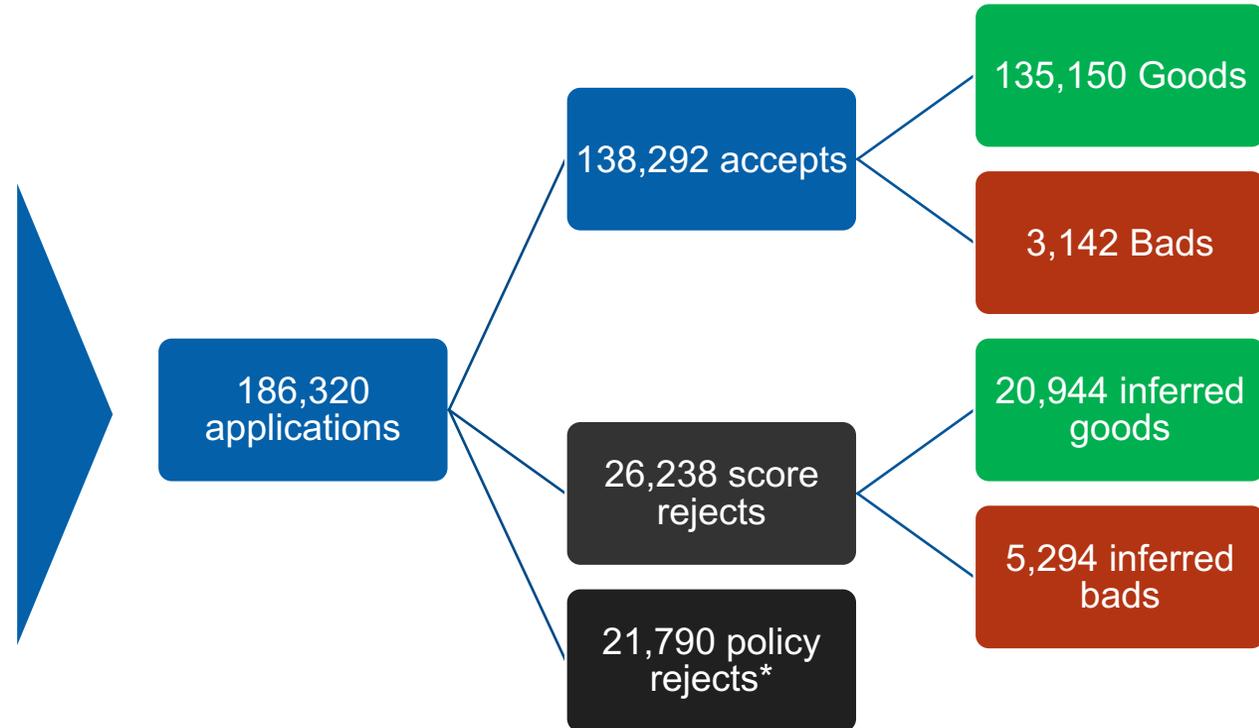
*\* Policy rejects were excluded from scorecard development*

# For regulatory purposes, we developed models to estimate monthly salary and debt ratio

**The requirement**

For regulatory purposes,

- Estimate algorithmically monthly salary

- Estimate algorithmically debt ratio

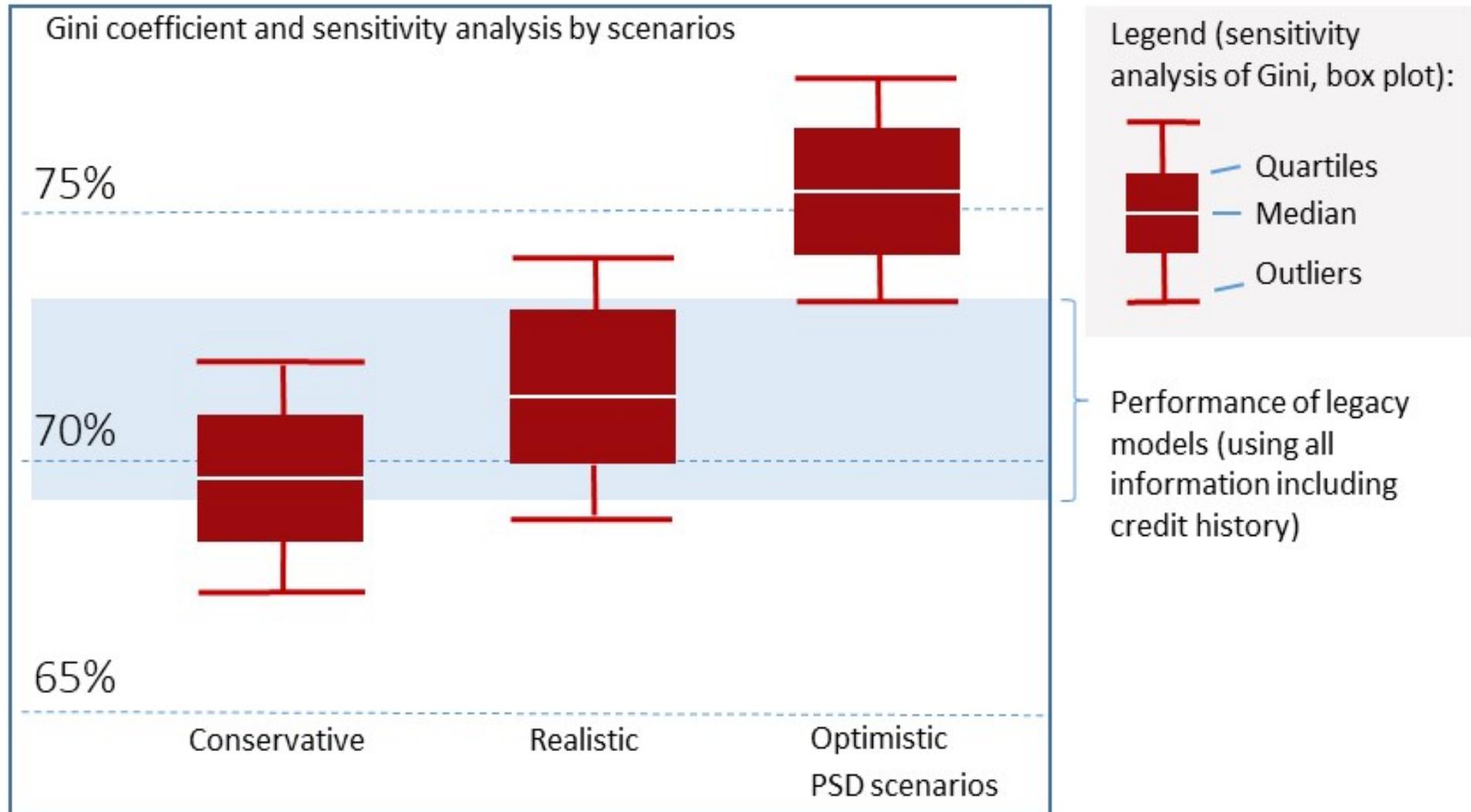- Establish accuracy of these estimates / predictions

**Approach**

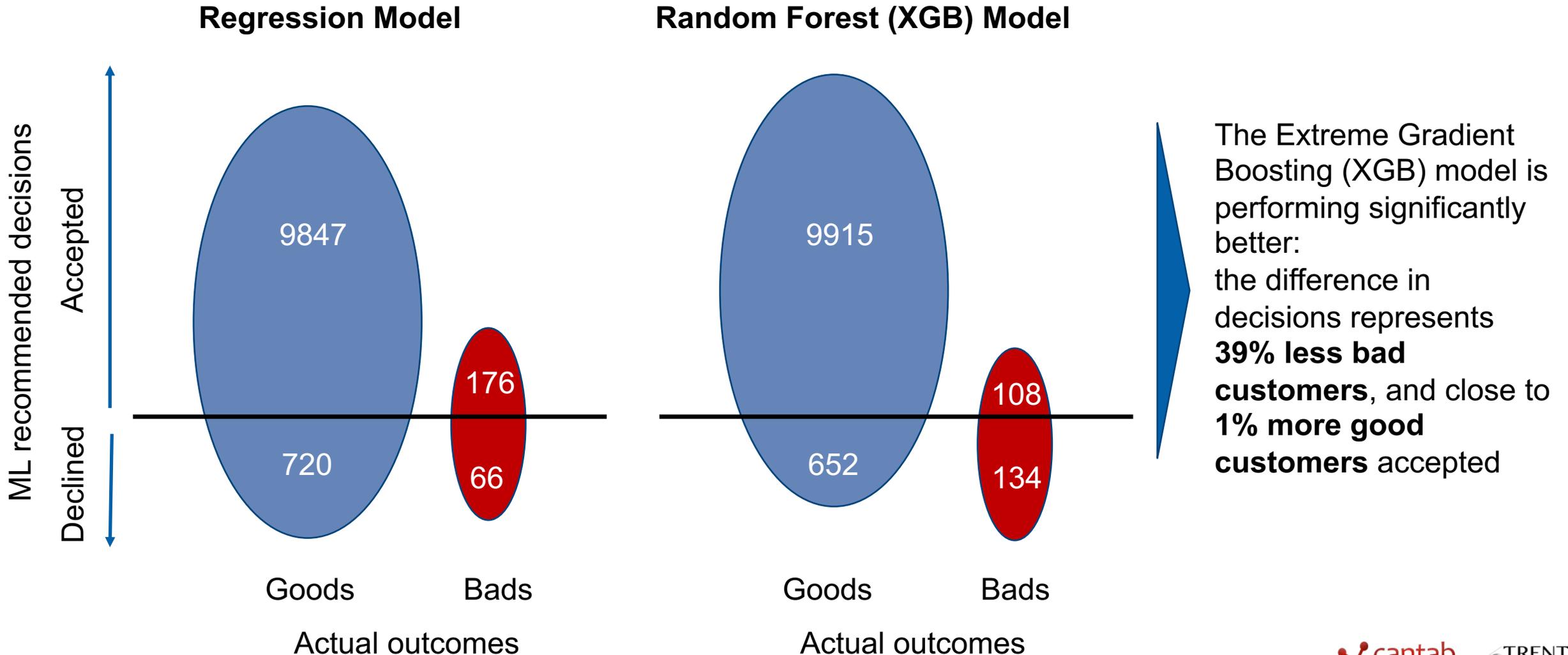Developed separate predictive models:

- To predict average monthly salary from the transactional data

- To predict average monthly loan payments from the transactional data

- Established statistical accuracy of these models, satisfactory for regulatory purposes

Outputs of these models as variables **did not significantly** improve predictivity of the credit scoring model

cantab
predictive intelligence

TRENT
Advisory Services

# The performance of developed models, measured in Gini, is comparable or better than the legacy models.



Gini coefficient and sensitivity analysis by scenarios

75%

70%

65%

Conservative

Realistic

Optimistic
PSD scenarios

Legend (sensitivity analysis of Gini, box plot):

— Quartiles
— Median
— Outliers

Performance of legacy models (using all information including credit history)

# The Extreme Gradient Boosting model is performing significantly better than the regression model

**Regression Model**

**Random Forest (XGB) Model**



ML recommended decisions
Accepted
Declined

Regression Model:
- Goods: 9847 (Accepted), 720 (Declined)
- Bads: 176 (Accepted), 66 (Declined)

Random Forest (XGB) Model:
- Goods: 9915 (Accepted), 652 (Declined)
- Bads: 108 (Accepted), 134 (Declined)

Goods   Bads
Actual outcomes

Goods   Bads
Actual outcomes

The Extreme Gradient Boosting (XGB) model is performing significantly better:
the difference in decisions represents **39% less bad customers**, and close to **1% more good customers** accepted

Note: analysis carried on hold-out test sample

cantab
predictive intelligence

TRENT
Advisory Services

# The better performance of XGB models can be understood by much more precise capturing of „feature interactions"

XGB model, score 0.09, classified as bad, actually became bad



Regression model, score 0.05, classified as good, actually became bad



- Regression underpeforms, as two features typically correlated with good clients in this case drive the score up independently of the other variables

- XGB model is able to „dampen" typically good features, depending on other features, thus resulting with a more accurate score

# To interpret and monitor the model, we are using the „SHAP" technique
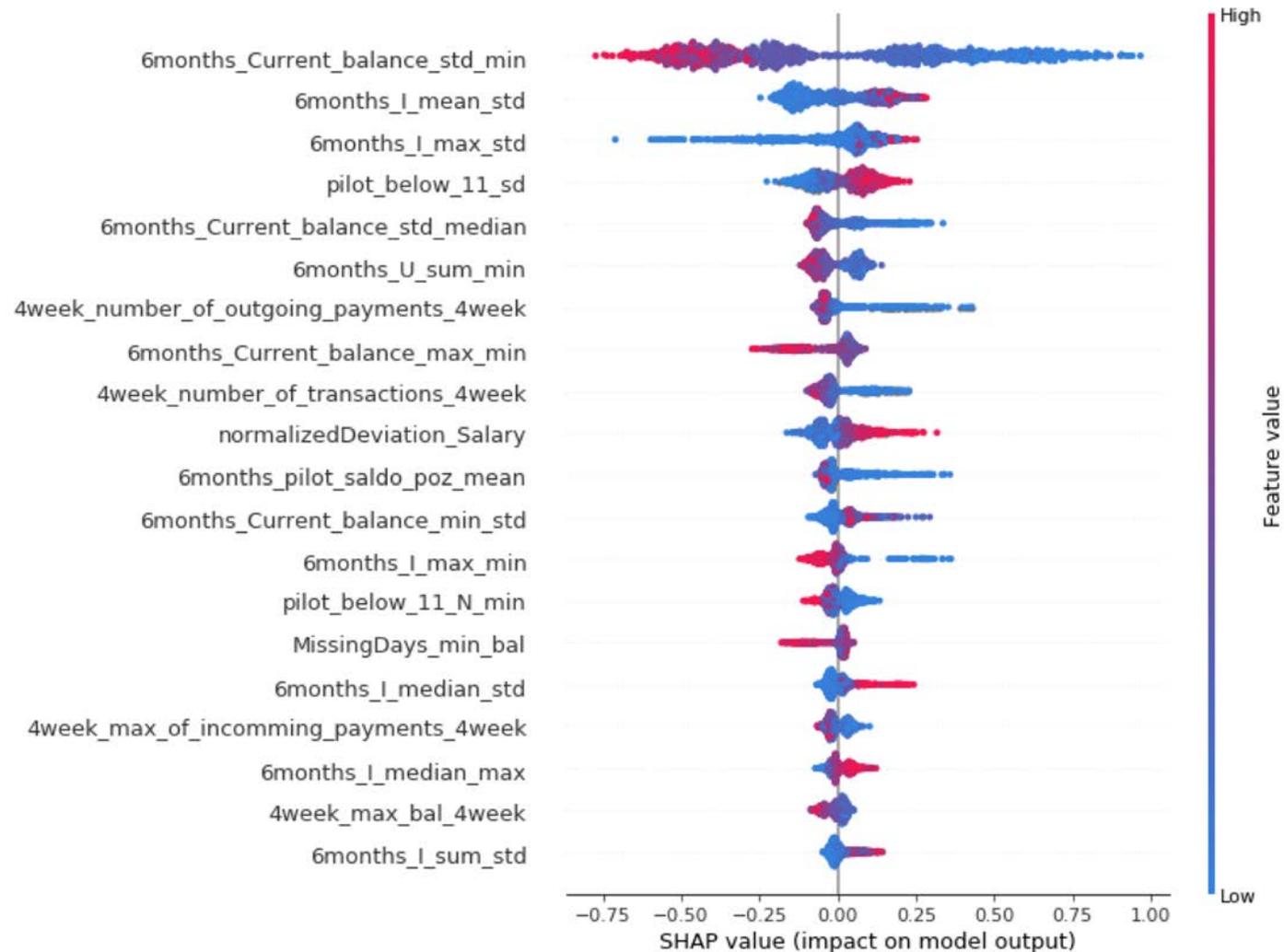
# Implications for the industry

Open Banking enables lenders to have access to a new wealth of information that is extremely predictive of credit risk:

This information is granular, recent, and provides a unique perspective on credit applicants financial behavior

Combined with traditional credit history, lenders now have a comprehensive picture to take automated credit decisions

**This means:**

- Better assessment of credit risk
- Better assessment of affordability
- Faster decisions
- More relevant and adequate credit offers
- Level playing field competition

**And this benefits:**

- Consumers
- Lenders
- Regulators:
  - ✓ Competition
  - ✓ Conduct
  - ✓ Prudential

cantab
predictive intelligence

TRENT
Advisory Services

# Thank you!

Sinisa Slijepcevic: [sinisa@cantbpi.com](mailto:sinisa@cantbpi.com)

Marc Gaudart: mg@trentadvisoryservices.com