NPL portfolios pricing based on Machine Learning techniques

FSRM Moscow

May 2019

EY

Building a better working world

The better the question. The better the answer.
The better the world works.

FSRM Group
Quantitative credit risk assessment

Our team

We are part of Advisory department in CIS and Eastern Europe (over 90 professionals experienced in quantitative, regulatory and technological aspects of financial risk management)

All aspects of development and implementation of quantitative credit risk models

Over 40 implemented projects for development / validation / verification of credit risk models for various business processes of financial institutions

Over 100 different clients in financial sector around the world that allows to engage international experts on any issue
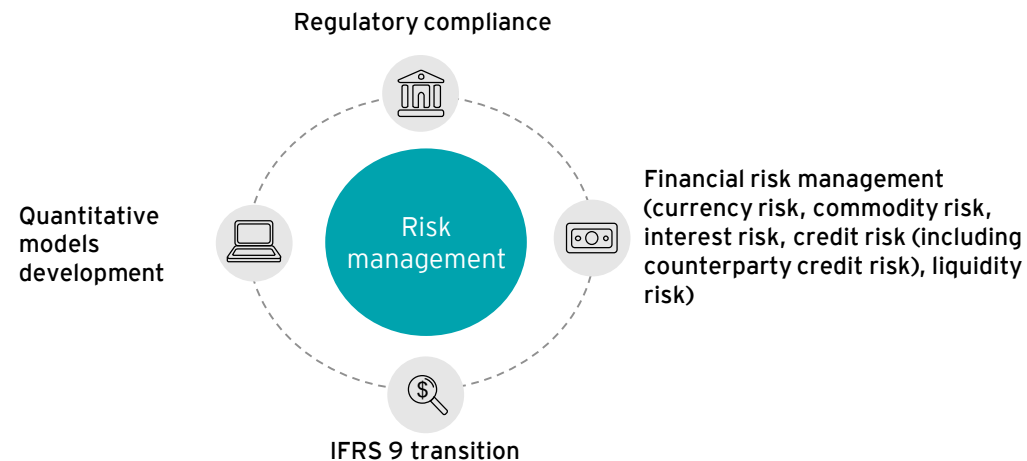
Michail Tsibulevsky, FRM
Partner

Head of FSRM in Russia and CIS

Oleg Chernyshev,
Senior Manager

Head of modeling and quantitative risk assessment

Main competencies

Regulatory compliance

Quantitative models development

Risk management

Financial risk management (currency risk, commodity risk, interest risk, credit risk (including counterparty credit risk), liquidity risk)

IFRS 9 transition

Our clients

СБЕРБАНК   ВТБ   Банк России   ГАЗПРОМБАНК   UniCredit

ВЭБ   ПКБ   ДОМ.РФ   А   citibank

ВНЕШЭКОНОМБАНК   Nordea   Альфа-Банк   INTERNATIONAL INVESTMENT BANK

Kapital Bank BİRİNCİ BANK   unibank   AMERICAN EXPRESS   J.P.Morgan

PASHA Holding

Key service areas

Area	Key topics		
1	Services related to the risk management system, the requirements of the ICAAP and the Recovery Resolution Plan	ICAAP / Recovery plan	RMF / MIS	KPI / Remuneration
2	Cybersecurity and Operational Risk Services	Cyber security / RCSA		KRI / Controls
3	Services related to the optimization of business processes of the bank	Three lines of defense / Process mining		EWS / Collection
4	Services related to IFRS 9	Provisions / Optimization		Back-testing
5	Services related to predictive analytics, machine learning methods and quantitative risk assessment methodology	Models / Validation	AML	Machine learning / PD LGD
6	Services related to the assessment of market risks, liquidity risks and approaches to fair value estimation	CVA DVA / ALM	FTP	VaR / Fair value
7	Services related to IT, data quality and risk management automation	Business requirements / Test scenarios		Data

Project overview

Current situation

- The company's business is based on the brands of two collection agencies with total portfolio of overdue debts of 850 bn rubles

- Company has logistic regression model which Is used for portfolio management and making strategic decisions when purchasing overdue debts portfolios

- Current model for portfolio valuation is constantly being improved, but it has some limitations:
  - It can be applied only to historical data
  - It cannot take into account operational changes ang impact of business initiatives
  - It does not allow to use expert judgments when evaluating portfolios
  - It requires improved accuracy of estimates

- Application of existing model for valuation of portfolios with poor quality data and, as a result, for making decisions about business expansion is questionable

Goals and objectives

Goals:
- Improve current model accuracy

Objectives:
- Internal data analysis and advanced model development using machine learning tools
- Development of the model satisfying the following requirements:
  - Confidence interval estimation
  - Consideration of all available data, macroeconomic indicators and operational improvements, interpolation of missing data
  - Possibility for expert adjustments
  - External data consideration
- Back-testing of the model results on the overdue debts portfolios purchased in 2017
- Preparation of technical report on developed model, model documentation and presentation for users and company owners

Results

- Model estimates absolute collection sum

- Model estimates absolute costs

- Model determines 95% confidence interval for collection sum

- Model meets the following:
  - More accurate estimates compared to the current model
  - Larger portion of portfolios with low and medium risk
  - Estimated confidence interval covers 95% of cases

5

Model description

Bank provides information about the portfolio of "bad" loans

Data cleaning and enrichment

Portfolio data is enriched with information about borrowers and Bank from multiple sources

Collection from the borrowers and profit gain

The Bureau collects the loans using soft, hard and legal collection. If the collection costs are lower than the collected sum the bureau gets the profit

ПКБ

Collection bureau business process

Portfolio pricing

Participating in auction

The portfolio is bought by the collection bureau which offers the highest price for the portfolio

- How much should we pay for such a portfolio?
- How many borrowers will pay at least anything?
- How much will we collect from the borrowers?
- How much will we spend on the collection?
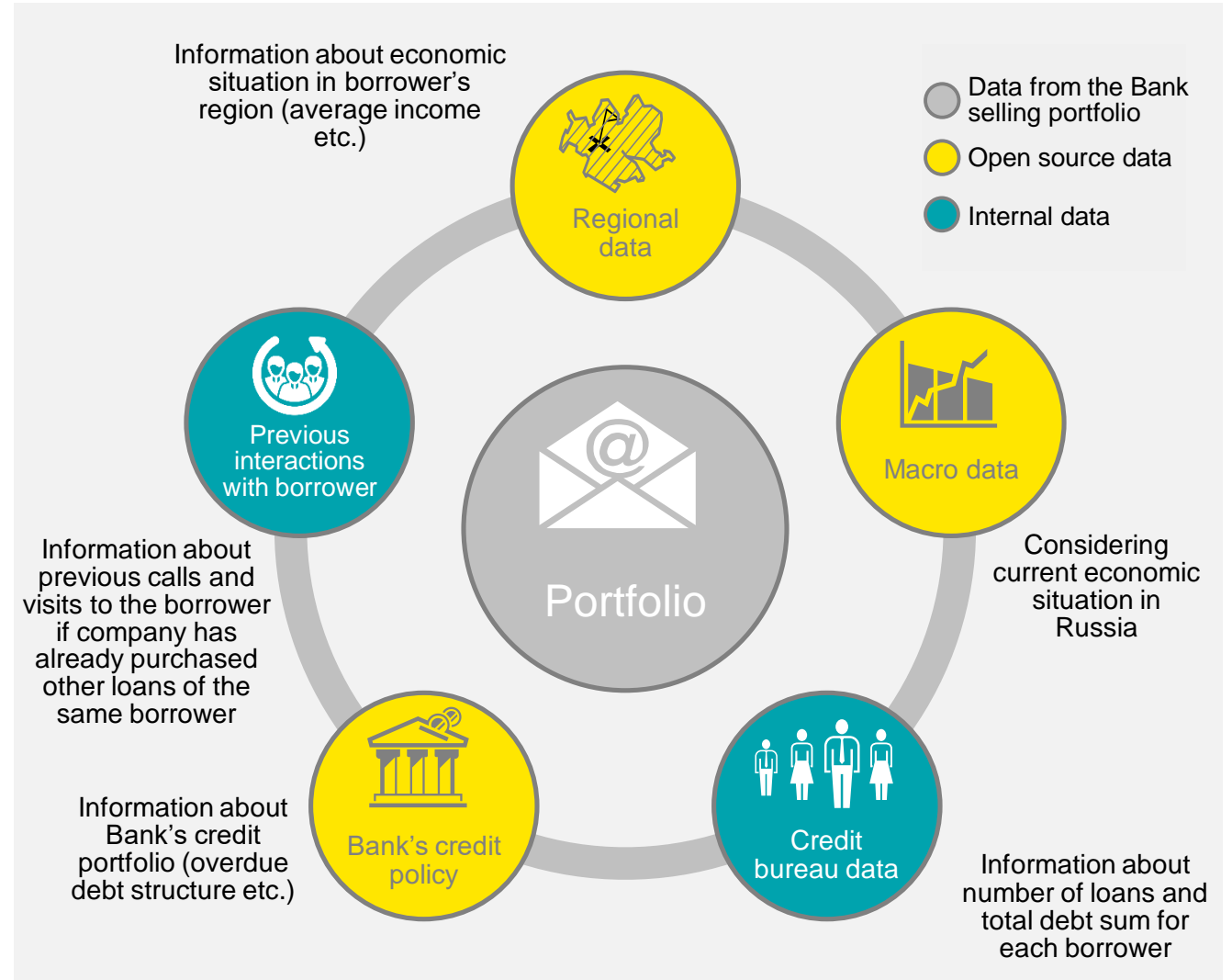- Which is the best way to collect from particular borrower?

Data enrichment

Both open-source and company's internal information on borrowers can be used for dataset enrichment

Data enrichment of the dataset allows to optimize the following processes

Borrowers prioritization (by the probability of response), determination of the most risky borrowers

Development of individual collection strategy for each borrower

Information about economic situation in borrower's region (average income etc.)

Regional data

Data from the Bank selling portfolio

Open source data

Internal data

Previous interactions with borrower

Macro data

Portfolio

Considering current economic situation in Russia

Information about previous calls and visits to the borrower if company has already purchased other loans of the same borrower

Bank's credit policy

Credit bureau data

Information about Bank's credit portfolio (overdue debt structure etc.)

Information about number of loans and total debt sum for each borrower

Key challenges

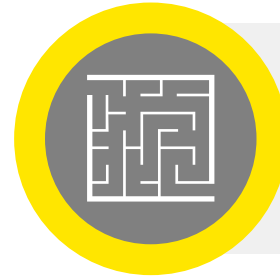**Characteristics of portfolios purchased from different banks are heterogeneous**

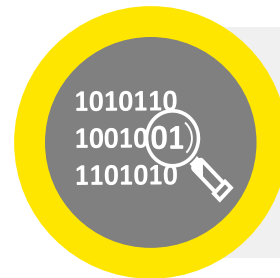**Data in portfolios are different for each bank**.

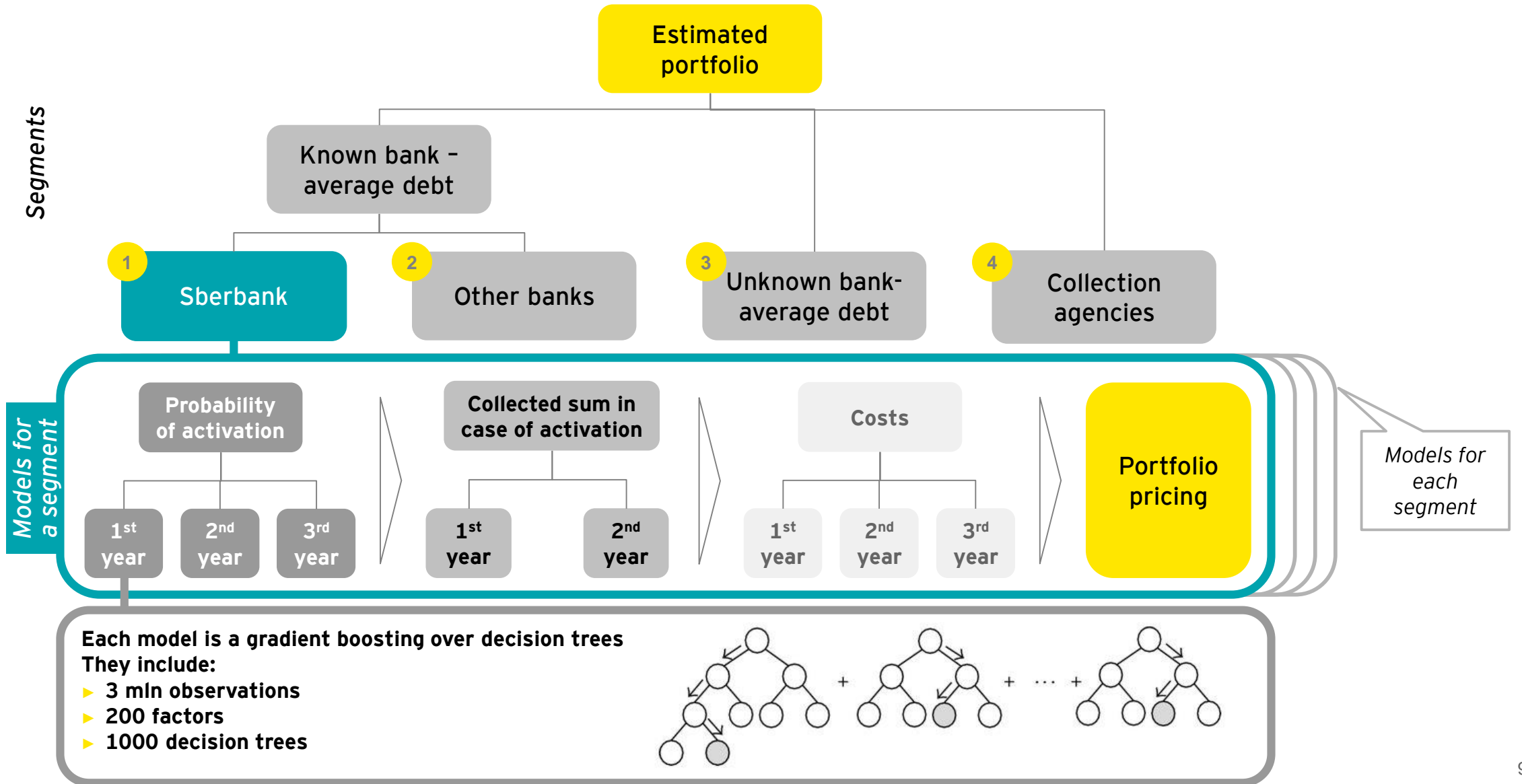**Portfolios from new banks can be purchased**

New collection strategies evolved.

Transitions between strategies changed

1010110
1001001
1101010

Data formats in internal database changed

Model structure

Segments

```
                        ┌──────────────────┐
                        │    Estimated     │
                        │    portfolio     │
                        └──────────────────┘
```

Known bank – average debt

(1) **Sberbank**   (2) **Other banks**   (3) **Unknown bank-average debt**   (4) **Collection agencies**

Models for a segment

Probability of activation	Collected sum in case of activation	Costs	Portfolio pricing
1st year 2nd year 3rd year	1st year 2nd year	1st year 2nd year 3rd year	

Models for each segment

Each model is a gradient boosting over decision trees
They include:
► **3 mln observations**
► **200 factors**
► **1000 decision trees**

Project results (1 / 2)

Processed Data

Train and test sample included

3 mln loans

Number of factor used for the model development counted

200 factors

Model factors

The developed model takes into Account macroeconomic forecasts as well as the external data describing the credit policy of the Bank selling the portfolio

Model Architecture

Model architecture was developed in accordance with the target segments of the collection bureau and it included

26 models

Model Algorithm

The main algorithm for Forecasting of portfolio collection and costs was

Gradient tree boosting

Forecasting Horizon

The developed model builds forecasts for the next

3 years

Project results (2 / 2)

High model accuracy

95% of the test sample is inside the confidence interval

Average **deviation** of the forecasted collection sum from the fact collection is **less than 20%**

Collection strategy determination

Optimal collection strategy

is automatically determined by the model for each loan based on detailed analysis of collection sum and costs

Model integration

Model is fully integrated with internal database and company's business processes

User interface

User-friendly **application** was developed for model estimation and portfolio price evaluation