



НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ  
УНИВЕРСИТЕТ

# Интегрированное управление риском и задачи машинного обучения

Масютин А.А.

Факультет компьютерных наук,

Департамент анализа данных и искусственного интеллекта

Руководитель: д.ф.-м.н. Кузнецов С.О.

# Различные виды риска в банковском деле

Под **риском** понимают **возможность наступления некоторого неблагоприятного события**, влекущего за собой различного рода потери (например обесценение активов, получение доходов ниже ожидаемого уровня)

Выступая финансовыми посредниками между агентами в экономике, **банки сталкиваются с различными видами риска:**

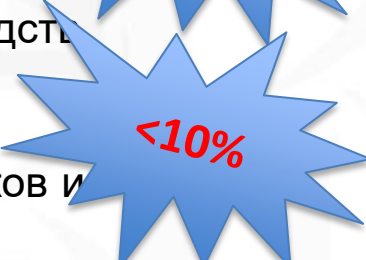
- **Кредитный риск** – возможность потерь банком финансового актива в результате неспособности контрагентов (заемщиков) исполнить свои обязательства по выплате процентов и основной суммы долга в соответствии с условиями договора.
- **Рыночный риск** – возможность потерь при неблагоприятном для банка движении макроэкономических факторов.
- **Риск ликвидности** – возможность недостатка свободных денежных средств для исполнения обязательств со стороны банка.
- **Операционный риск** – частота и вероятный масштаб ошибок сотрудников и сбоев систем.



70-80%



10-15%

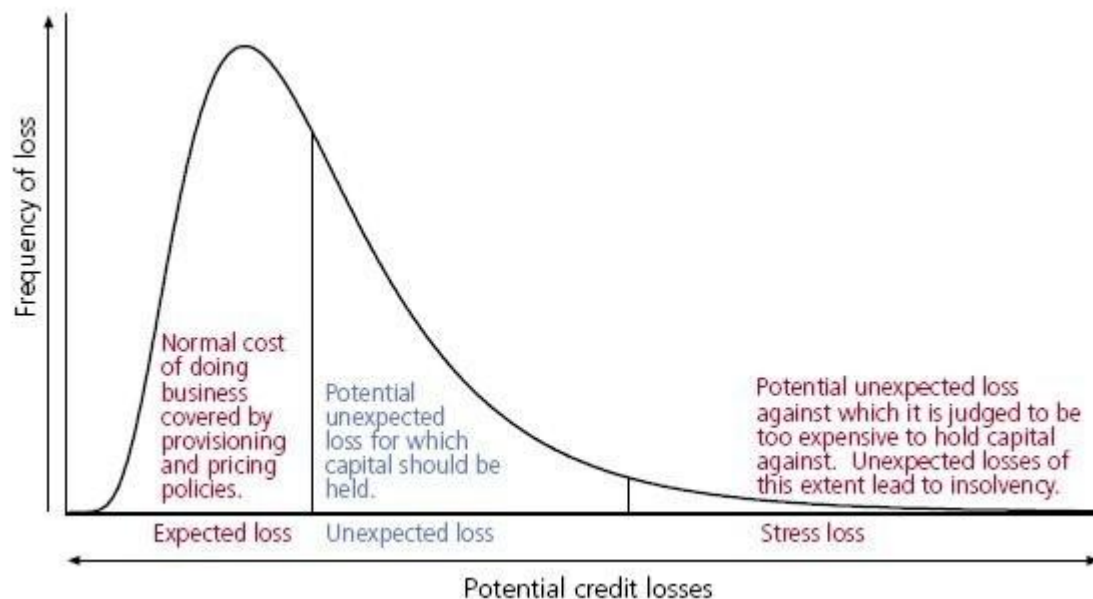


<10%

# Концепция управления кредитным риском

Этапы управления **кредитным риском**:

- Выбор **кредитной политики** (кого кредитуем)
- Определение профиля **риск – доходность** (риск-аппетит)
- **Отбор контрагентов** (**скоринг**)
- Проведение мероприятий **взыскания просроченной задолженности**/реструктуризация
- Оценка ожидаемых и экстремальных потерь (расчет **распределения потерь**)
- Создание резервов под ожидаемые убытки и непредвиденные убытки.



# Расчет распределения потерь по кредитному риску (1)

## Определения:

Пусть  $G$  – множество объектов (заемщики, контрагенты),  $|G| = N$ ,

Пусть  $M$  – множество вещественных атрибутов,  $|M| = K$ ,

В качестве  $m_{ij}$  будем обозначать значение атрибута  $m_j \in M$  для  $i$ -ого заемщика  $g_i \in G$ ,  $\forall j = \overline{1, K}$ ,  $\forall i = \overline{1, N}$ ,  $m_{ij} \in \mathbf{R}$ ,

Отдельно из множества  $M$  выделим атрибут  $e \in M$ , который представляет сумму долга заемщика, а также целевой атрибут  $d$ :

$$d_i = \begin{cases} 1, & \text{в случае если } g_i \text{ перешел в состояние дефолта} \\ 0, & \text{в случае если } g_i \text{ не перешел в состояние дефолта} \end{cases}$$

*Определение состояния дефолта обычно определяется следующим образом: пропуск трех платежей на горизонте 12 месяцев с текущего момента времени.*

Пусть  $D$  – множество описаний вида  $([a_1; b_1], [a_2; b_2], \dots, [a_K; b_K])$ , причем  $a_j, b_j \in \mathbf{R}$ ,  $a_j \leq b_j$ ,  $\forall j = \overline{1, K}$ .

Пусть  $\delta: G \rightarrow D$ , возвращающее описание объекта(-ов)  $g_i \in G$ , а  $\Pi$  - операция пересечения:

$$\forall x, y \in D, x \Pi y = ([\min(a_1^x; a_1^y), \max(b_1^x; b_1^y)], \dots, [\min(a_K^x; a_K^y), \max(b_K^x; b_K^y)])$$

Также определим свойство вложения описаний:  $x \sqsubseteq y \Leftrightarrow x \Pi y = x$

И операторы Галуа:  $A^\circ = \Pi_{g \in A} \delta(g)$ , где  $A \subset G$ ,  $d^\circ = \{g \in G | d \sqsubseteq \delta(g)\}$ , где  $d \in D$

# Расчет распределения потерь по кредитному риску (2)

Традиционно задача расчета распределения потерь по кредитному риску основывается на оценке двух параметров для каждого заемщика:

- Вероятность дефолта заемщика  $PD_i$
- Величина потерь в случае дефолта  $LGD_i$  (выраженная как доля от суммы долга  $e_i$ )

Тогда целевой атрибут  $d$  может быть представлен как случайная величина:

$$d_i = \begin{cases} 1, & \text{с вероятностью } PD_i \\ 0, & \text{с вероятностью } 1 - PD_i \end{cases}$$

Предполагается, что  $PD_i = f(m_{i1}, m_{i2}, \dots, m_{iK})$  так же, как и  $LGD_i = \varphi(m_{i1}, m_{i2}, \dots, m_{iK})$ , где  $f(\cdot)$  и  $\varphi(\cdot)$  – неизвестные отображения, поиск которых и является основной задачей в оценке кредитного риска.

Величину потерь по заемщику  $g_i$  можно представить, как случайную величину  $L_i$ :

$$L_i = \begin{cases} LGD_i \cdot e_i, & \text{с вероятностью } PD_i \\ 0, & \text{с вероятностью } 1 - PD_i \end{cases}$$

Пусть  $\widehat{PD}_i = \hat{f}(\delta(g_i))$ ,  $\widehat{LGD}_i = \hat{\varphi}(\delta(g_i))$  – некоторые оценки вероятности дефолта заемщика и величины потерь.

Тогда, в простейшем случае, для оценки распределения потерь по всему множеству объектов  $G$  (по всему портфелю) можно прибегнуть к методу симуляций.

# Расчет распределения потерь по кредитному риску (3)

Метод симуляций состоит в сценарном анализе наблюдаемого явления, а именно в генерации распределения случайной величины.

Пусть  $L$  – случайная величина, выражающая суммарные потери по портфелю состоящим из заемщиков множества  $G$ .

$$L = \sum_{i=1}^N L_i$$

Задача: для произвольного  $l$  оценить вероятность того, что совокупные потери по портфелю превысят  $l$ :  $\Pr(L > l) = 1 - \Pr(L \leq l) = ?$

Метод Монте-Карло состоит из следующих шагов.

```
for(z in 1:number_of_simulations){#Задается число симуляций
  1. random.vector = runif(N) #Вектор псевдо-случайных равномерно
                             #распределенных чисел от 0 до 1
  2. d = ifelse(PD > random.vector, 1, 0) #Генерируем реализацию вектора
                                           #дефолтов
  3. L = sum(d*e*LGD) #Считаем величину суммарных потерь для данной
                     #реализации
  4. Сохранить L
}
```



# Скоринговые системы (1)

Скоринговой системой называется отображение  $s: \mathbf{R}^K \rightarrow \mathbf{R}$

- задающее разбиение  $m_j^0 < m_j^1 < m_j^2 < \dots < m_j^{s_j-1} < m_j^{s_j}$  на множестве значений каждого атрибута  $m_j$  из  $M$ , где  $m_j^0 = -\infty$ ,  $m_j^{s_j} = +\infty$ ,
- сопоставляющее каждому элементу разбиения  $(m_j^{t-1}; m_j^t]$ ,  $t = \overline{1, s_j}$  скоринговый балл  $scr_j^t \in \mathbf{R}$ ,
- возвращающая для заданных атрибутов заемщика сумму скоринговых баллов по каждому атрибуту:

$$s(m_{i1}, m_{i2}, \dots, m_{iK}) = \sum_{j=1}^K \sum_{t=1}^{s_j} I(m_{j1} \in (m_j^{t-1}; m_j^t]) \cdot scr_j^t$$

- где  $I(\cdot)$  – функция-индикатор.

*Скоринговый балл, таким образом, является мерой, разделяющих «хороших» и «плохих» заемщиков. Предполагается, что скоринговая система должна присваивать «хорошим» заемщикам более высокие баллы относительно «плохих».*

Калибровкой скоринговой системы  $s$  называется отображение  $c: \mathbf{R} \rightarrow [0; 1]$

- ставящее в соответствие скоринговому баллу вероятность дефолта (вероятность того, что  $d_i = 1$ ).

# Скоринговые системы (2)

Заметим, что скоринговая система также может быть представлена как множество описаний  $\{d \in D \mid \Pi d = \mathbf{R}^K\}$  («микросегменты»). Действительно, в качестве такого множества описаний рассмотрим Декартово произведение разбиений атрибутов:

$$(m_1^{t_1-1}; m_1^{t_1}] \times (m_2^{t_2-1}; m_2^{t_2}] \times \dots \times (m_K^{t_K-1}; m_K^{t_K}], \forall j = \overline{1, K}, \forall t_j = \overline{1, s_j}$$

Изначальное разбиение атрибутов производится на основе Weight-of\_evidence transformation<sup>1</sup>. Категоризация атрибутов происходит по следующему алгоритму.

- атрибут разбивается на основе равных перцентилей на несколько групп. В каждой группе считается показатель веса категории WOE:

$$WOE_{(m_j^{t-1}; m_j^t]} = \ln \left( \frac{\sum_{i=1}^N (1-d_i) \cdot I(m_j \in (m_j^{t-1}; m_j^t])}{\sum_{i=1}^N d_i \cdot I(m_j \in (m_j^{t-1}; m_j^t])} \right) / \frac{\sum_{i=1}^N (1-d_i)}{\sum_{i=1}^N d_i}$$

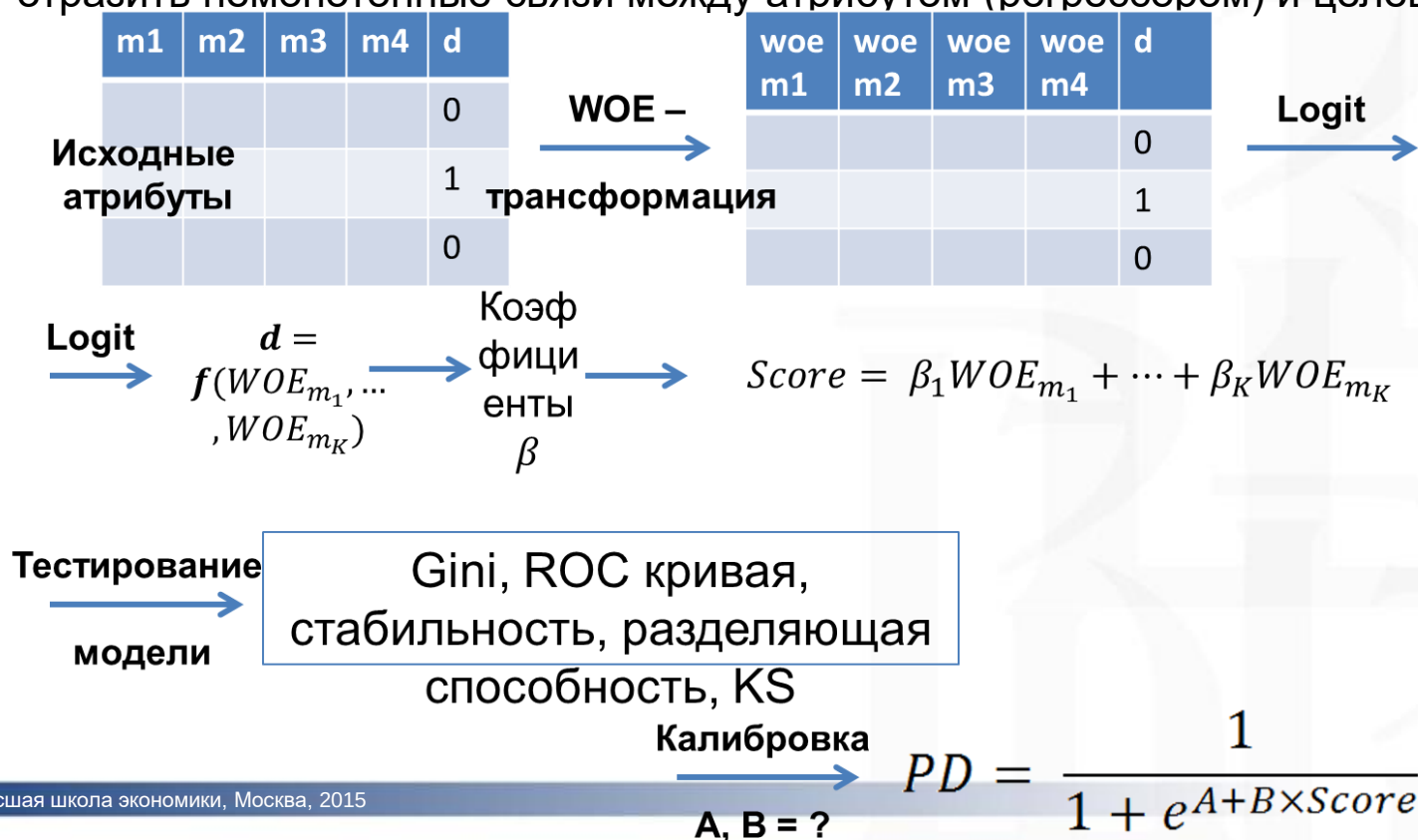
- Итеративно перцентили объединяются в более крупные группы, на основе близости значений WOE между категориями.
- После категоризации переменной каждому диапазону значения ставится в соответствие величина WOE.
- Далее оценивается логистическая регрессия на основе преобразованных атрибутов.



# Скоринговые системы (3)

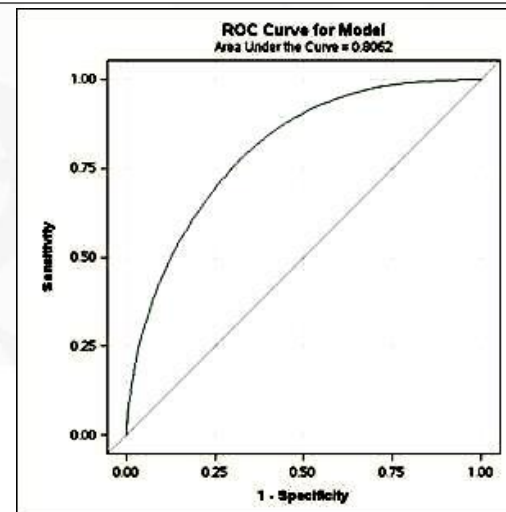
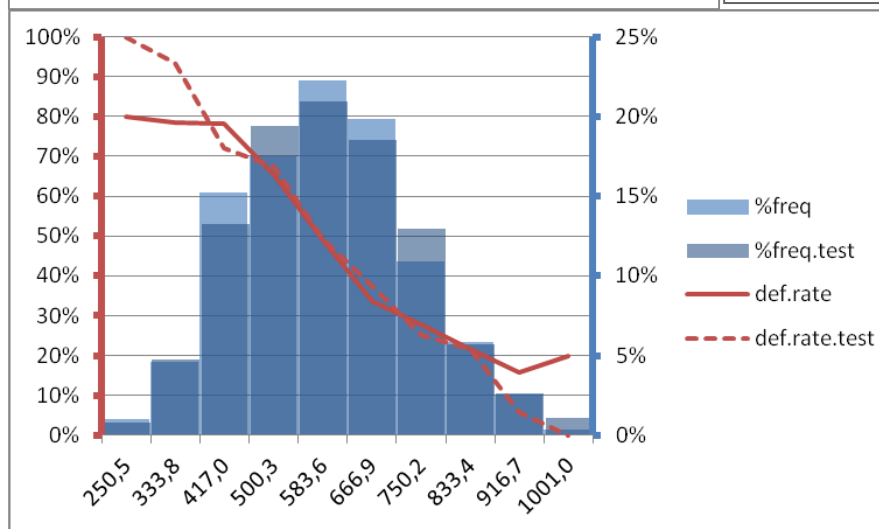
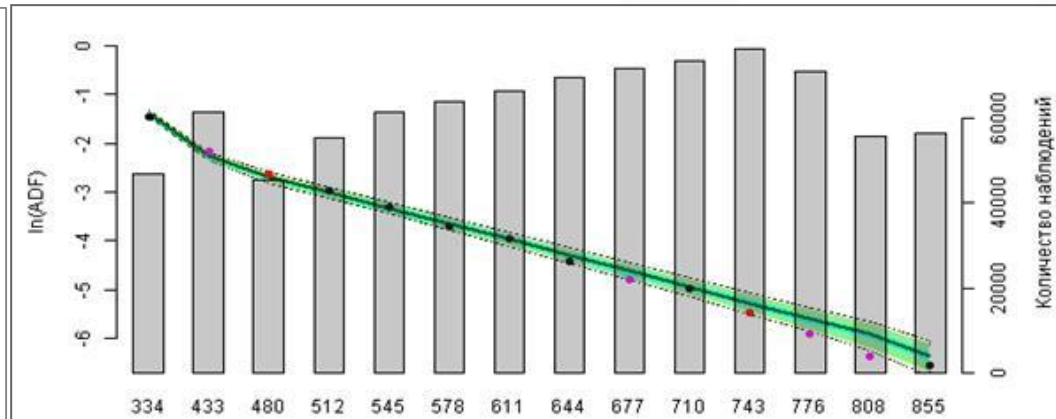
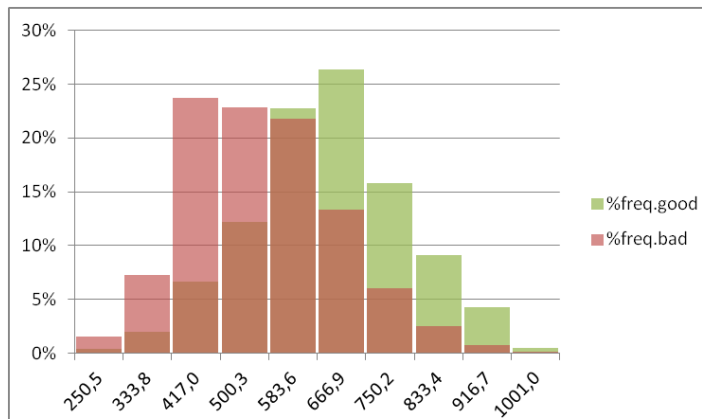
Категоризация количественных переменных позволят добиться следующих основных преимуществ при построении скоринговой карты:

- облегчить обработку выбросов и экстремальных значений количественных атрибутов;
- упростить интерпретацию скоринговой карты;
- отразить немонотонные связи между атрибутом (регрессором) и целевым атрибутом



# Скоринговые системы (4)

- Плюс скоринговых карт состоит в том, что сформированные сегменты имеют понятную интерпретацию, по ним удобно осуществлять мониторинг кредитного портфеля:



# Скоринговые системы (5)

## Итак, текущая ситуация:

- Скоринговые карты дают прозрачный инструмент управлением риска, позволяют проводить сценарный анализ для сегментов.
- Банки остаются «глухими» к «немым» black-box методам ML.
- Однако, по сути базируются на однофакторном анализе, не учитывая всю структуру описания объектов.
- Используют инструмент логистической регрессии предпосылки которой могут не выполняться

## Таким образом, существует потребность в инструменте, который:

- рассматривает атрибуты не по отдельности, а учитывает всю структуру описания объекта
- производит классификацию объектов на основе интерпретируемых и прозрачных правил
- возвращает не только предсказанное значение целевого атрибута, но и предоставляет численную меру ранжирования объектов (ранжирования заемщиков по риску)

Возможным решением такой потребности является использование таких алгоритмов анализа формальных понятий и узорных структур (FCA).

# Алгоритм ленивой классификации

Пусть множество объектов  $G$  делится на положительный контекст  $G^+$  (там где  $d_i = 1$ ) и отрицательный контекст  $G^-$  ( $d_i = 0$ ).

Пусть  $g_t$ - тестовый объект с неизвестным значением атрибута  $d_t$ .

Алгоритм ленивой классификации состоит из следующих шагов

1. Для каждого  $g \in G$  найти  $(\delta(g) \sqcap \delta(g_t))^\circ$
2. Если для некоторого  $g$  все объекты из  $(\delta(g) \sqcap \delta(g_t))^\circ$  принадлежат множеству  $G^+$ , то  $\widehat{d}_t = 1$ . Если  $(\delta(g) \sqcap \delta(g_t))^\circ$  принадлежат  $G^-$ , то  $\widehat{d}_t = 0$ .

Для непрерывных вещественных атрибутов такой подход может вызывать следующие трудности:

- При больших  $K$  пересечения описаний тестового объекта и объекта из контекста  $G$  представляют собой очень специфичные описания («узкие интервалы»), в образ которых входят только они сами.
- Хотелось бы получать описания с большей поддержкой, поэтому появилась идея пересекать тестовый объект не с одним, а сразу с несколькими.
- В статье<sup>1</sup> было предложено использовать рандомизацию с настраиваемыми параметрами:
  - Мощность множества объектов из контекста  $G$
  - Допустимый уровень фальсификации «гипотезы» (“ $\alpha$ -weak premise”)
  - Число сэмплингов

# Модифицированный алгоритм ленивой классификации (MLCA)

## Определения:

$\alpha$ -слабым положительным описанием называется  $h_+ \in D$  если и только если:

$$\frac{\|h_+ \cap G^-\|}{\|G^-\|} \leq \alpha \text{ and } \exists A \subseteq G^+: h_+ \sqsubseteq A^\circ$$

Пусть *subsample.size* – параметр, который определяет число объектов (как долю от общего числа), с которыми будет пересекаться тестовый объект.

Пусть *n.iter* – параметр, определяющий число случайных подвыборок объектов размером  $\text{subsample.size} \times G^q$ ,  $q = \{+, -\}$

Пусть  $\alpha$  – параметр пороговое значение («*anti-support*»)

Тогда последовательность шагов алгоритма представляется так:

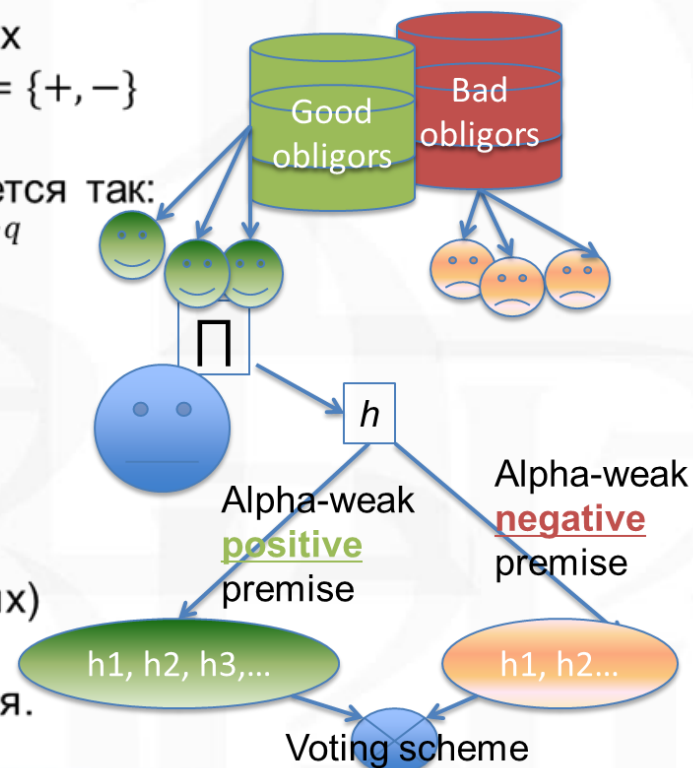
**Шаг 1.** Произвести случайную подвыборку объектов из  $G^q$  размером  $\text{subsample.size} \times G^q$ .

**Шаг 2.** Найти образ пересечения  $\delta(g_1) \sqcap \dots \sqcap \delta(g_t)$

**Шаг 3.** Если  $\delta(g_1) \sqcap \dots \sqcap \delta(g_t)$  является  $\alpha$ -слабым  $q$ -ым описанием, то добавить его в коллекцию.

Повторить шаги 1 - 3 *n.iter* число раз для  $q = \{+, -\}$ .

Если коллекции  $\alpha$ -слабых положительных (отрицательных) описаний не пусты, произвести классификацию тестового объекта на основе некоторой схемы голосования.





# Схемы голосования (1)

Схемой голосования называется отображение

$$F: (g_t, h_+^1, \dots, h_+^p, h_-^1, \dots, h_-^n) \rightarrow \{0, 1, \emptyset\}$$

где  $h_+^j$  –  $\alpha$ -слабое положительное описание  $\forall j = \overline{1, p}$ , а  $\{0, 1, \emptyset\}$  – значения целевого атрибута  $d_t$ ,  $\emptyset$  – отказ от классификации.

Чтобы конструировать разные схемы голосования, нужно определить взвешивающую функцию  $w(\cdot)$ , агрегирующий оператор  $A(\cdot)$  и сравнивающий оператор  $\otimes$ :

$$F(g_t, h_+^1, \dots, h_+^p, h_-^1, \dots, h_-^n) = (A_{i=1}^p [w(h_+^i)]) \otimes (A_{j=1}^n [w(h_-^j)])$$

Некоторые примеры схем голосования:

JSM:

$$\omega(h) = \begin{cases} 1, & \text{if } \delta(g_{test}) \sqsubseteq h \\ 0, & \text{otherwise} \end{cases}$$

$$A(h) = \bigvee h$$

$$\otimes = \wedge \neg$$

Число гипотез:

$$A(h) = \sum h$$

$$\omega(h) = \begin{cases} 1, & \text{if } \delta(g_{test}) \sqsubseteq h \\ 0, & \text{otherwise} \end{cases}$$

$$a \otimes b = \begin{cases} \text{sign}(b - a), & \text{if } a \neq b \\ \emptyset, & a = b \end{cases}$$

Суммарное число  
объектов в коллекции:

$$A(h) = \sum h$$

$$\omega(h) = ||h^\diamond||$$

$$a \otimes b = \begin{cases} \text{sign}(b - a), & \text{if } a \neq b \\ \emptyset, & a = b \end{cases}$$

# Схемы голосования (2)

Некоторые примеры схем голосования:

По ширине интервалов:

Нормализация атрибутов:

$$X_{norm} = \frac{X - \min(X)}{\max(X) - \min(X)}$$

$$A(h) = \sum h$$

$$\omega(h) = \frac{M - \sum_{k=1}^M (b_k^h - a_k^h)}{M}$$

$$a \otimes b = \begin{cases} \text{sign}(a - b), & \text{if } a \neq b \\ \emptyset, & a = b \end{cases}$$

По ширине интервалов

и с числом объектов:

$$\omega(h) = \|h\| \frac{M - \sum_{k=1}^M (b_k^h - a_k^h)}{M}$$

$$\begin{aligned} F(g_{test}, h_1^+, \dots, h_p^+, h_1^-, \dots, h_n^-) = \\ = \left( \sum_{i=1}^p \|h_i^+\| \frac{M - \sum_{k=1}^M (b_k^{h_i^+} - a_k^{h_i^+})}{M} \right) \otimes \\ \otimes \left( \sum_{j=1}^n \|h_j^-\| \frac{M - \sum_{k=1}^M (b_k^{h_j^-} - a_k^{h_j^-})}{M} \right) \end{aligned}$$

# Эксперименты и использованные данные

- Данные представляют собой данные по заявкам на **потребительские кредиты** крупного банка на горизонте фев.2011 по май.2014. **Факт дефолта** определен как наличие **просрочки более 90 дней** по регулярным платежам **на горизонте 12 месяцев** с даты оценки заемщика.
- Выборки формировались репрезентативными по отношению ко всему портфелю розничных кредитов банка. **Объем выборки** дефолтных заемщиков после очистки данных составил **492 320 наблюдений**. Объем недефолтных был скорректирован так, чтобы соответствовать объему выборки дефолтных.
- Тестовое множество представляло 100 000 отобранных за тот же период наблюдений
- Состав атрибутов отражает **социально-демографическое** положение заемщика, **параметры кредитного продукта**, **данные БКИ** на дату оценки:

| Фактор                                                                            | Фактор                                                                               | Фактор                                                                                                    |
|-----------------------------------------------------------------------------------|--------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------|
| Возраст                                                                           | Наличие недвижимости                                                                 | Просроченная задолженность по основному долгу и процентам на предыдущий месяц                             |
| Длительность последней просрочки по основному долгу                               | Наличие ТС                                                                           | Просроченная задолженность по основному долгу и процентам на предыдущий месяц по потребительским кредитам |
| Длительность последней просрочки по процентам                                     | Обязательства заемщика                                                               | Просроченная задолженность по основному долгу и процентам по потребительским кредитам                     |
| Доход семьи                                                                       | Отношение лимита к доходу семьи                                                      | Расчетный срок кредита                                                                                    |
| Категория занимаемой должности                                                    | Отношение основного долга к лимиту                                                   | Семейное положение                                                                                        |
| Количество действующих потребительских кредитов                                   | Отношение остатка срока к сроку кредита                                              | Стаж на последнем месте работы                                                                            |
| Количество членов семьи                                                           | Отношение подтвержденного дохода к сумме подтвержденного и неподтвержденного доходов | Сумма выдачи/лимита                                                                                       |
| Максимальное количество дней просрочки за предыдущие 3 месяца, не включая текущий | Пол                                                                                  | Сумма номеров типов клиентов                                                                              |
| Наличие детей                                                                     | Просроченная задолженность по основному долгу по потребительским кредитам            | Сумма просроченного основного долга и процентов за предыдущие 6 месяцев, не включая текущий               |
| Наличие заграничного паспорта                                                     | Просроченная задолженность по основному долгу и процентам                            | Фактический срок кредита                                                                                  |

# Результаты

- Был произведен grid-search по параметрам алгоритма с оценкой качества точности на кросс-валидационной выборке.
- Расчеты производились на Fujitsu TX300 S7 , 2 процессора Xeon E5-2670 по 8 ядер, параметры, 64 Гб, 2.6ГГц
- Модели сравнивались с методами, которые используется в банке.

| Модель                  | Джини на тестовом сете | Модель | Схема взвешивания                             | Джини на тестовом сете |
|-------------------------|------------------------|--------|-----------------------------------------------|------------------------|
| Логистическая регрессия | 48,08%                 | MCLA   | <u>Число гипотез</u>                          | 56,12%                 |
| Скоринговая карта       | 52,02%                 |        | <u>Суммарное число объектов</u>               | 51,40%                 |
| CART                    | 54,81%                 |        | <u>По ширине интервалов</u>                   | 58,12%                 |
|                         |                        |        | <u>По ширине интервалов с числом объектов</u> | 61,50%                 |

| Модель                  | Время<br>расчета<br>одной<br>настройки<br>модели | Модель | Схема взвешивания                             | Время<br>расчета<br>одной<br>настройки<br>модели |
|-------------------------|--------------------------------------------------|--------|-----------------------------------------------|--------------------------------------------------|
| Логистическая регрессия | ~ 2 мин 30 сек                                   | MCLA   | <u>Число гипотез</u>                          | ~10 часов                                        |
| Скоринговая карта       | ~ 5 мин                                          |        | <u>Суммарное число объектов</u>               | ~11 часов                                        |
| CART                    | ~ 2 мин                                          |        | <u>По ширине интервалов</u>                   | ~10 часов                                        |
|                         |                                                  |        | <u>По ширине интервалов с числом объектов</u> | ~12,5 часов                                      |



НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ  
УНИВЕРСИТЕТ

# Спасибо за внимание!

101000, Россия, Москва, Мясницкая ул., д. 20

Тел.: (495) 621-7983, факс: (495) 628-7931

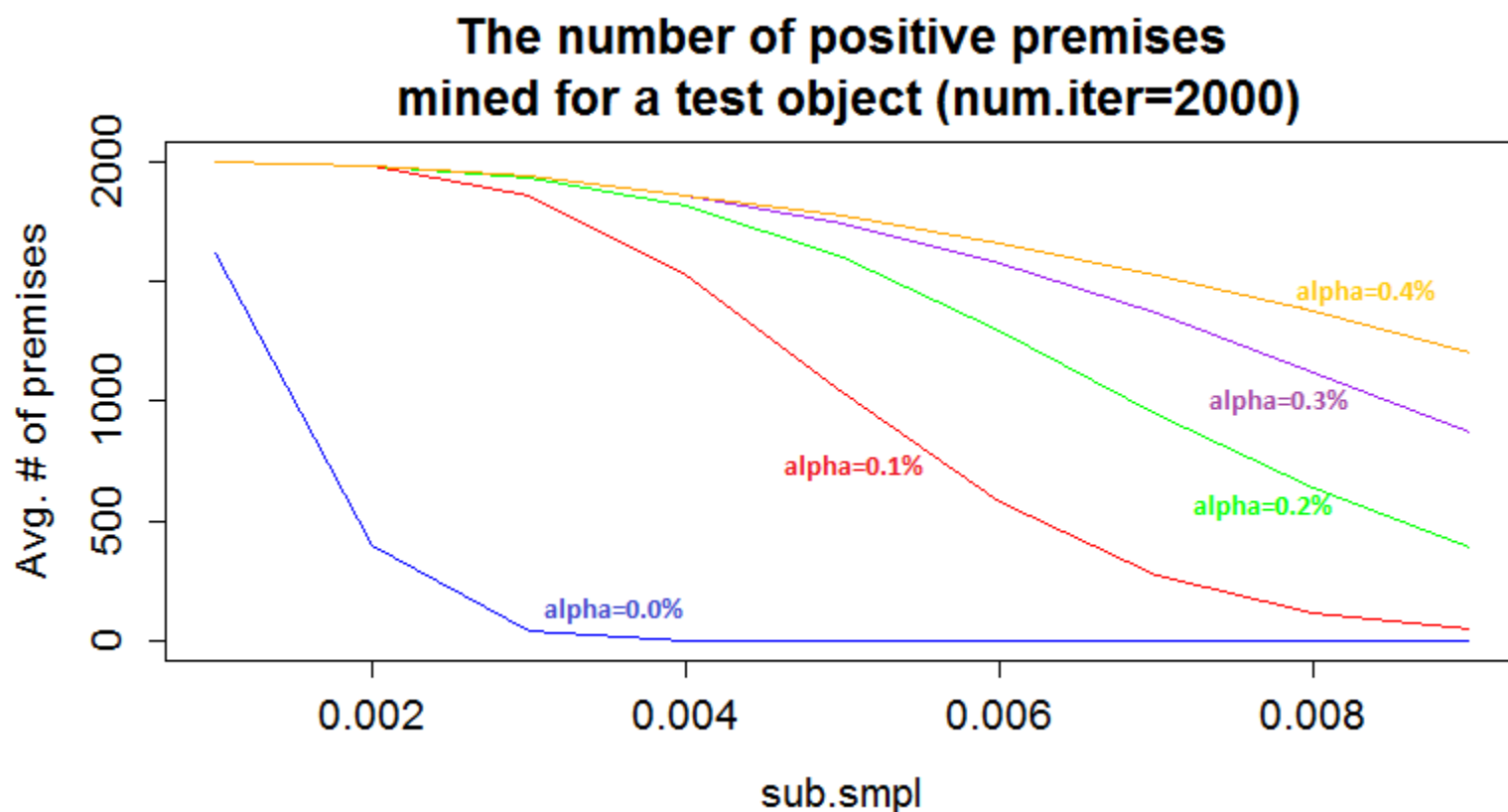
[www.hse.ru](http://www.hse.ru)



# Приложения

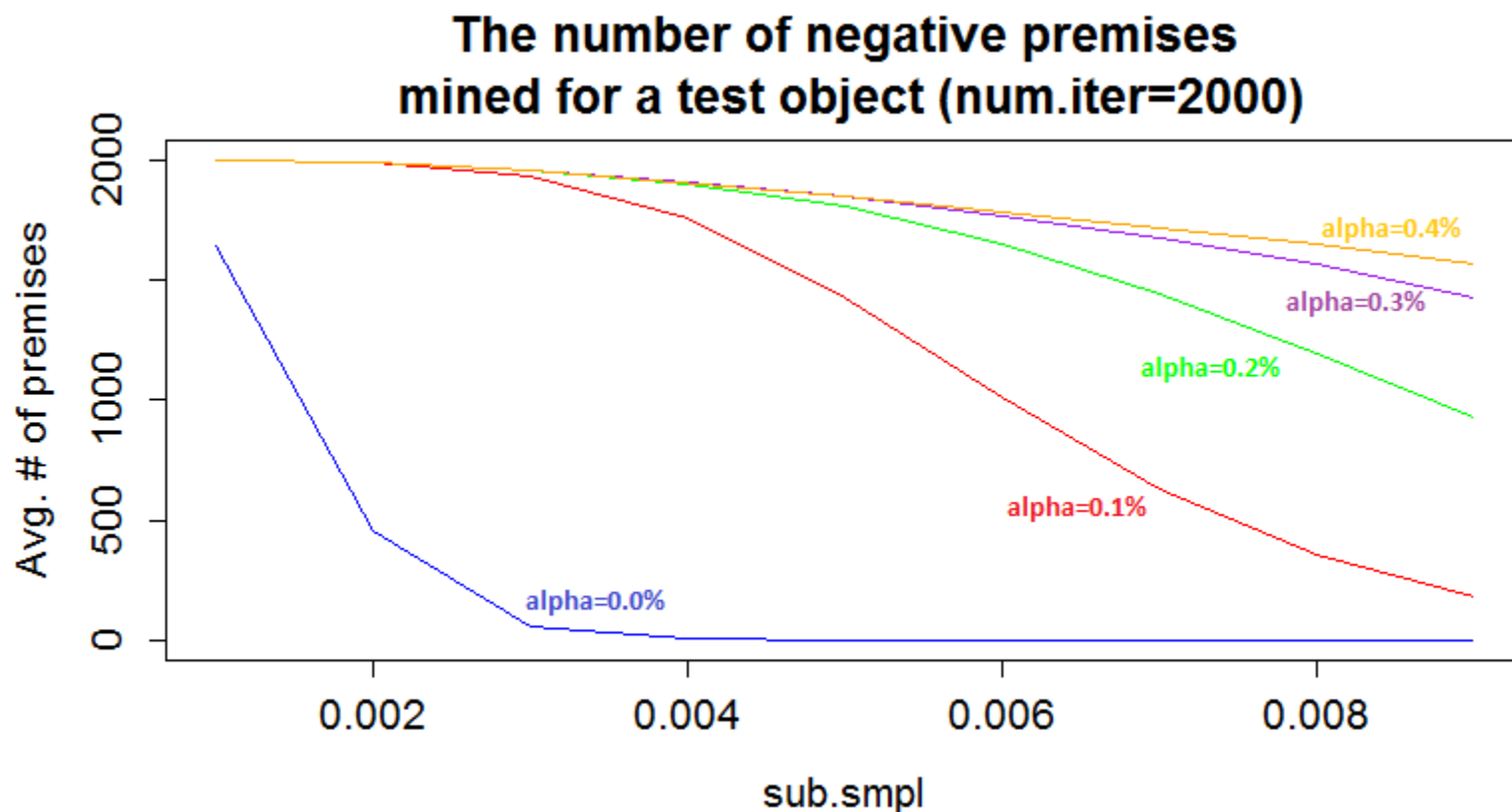
# Data and Experiments Results

- The dynamics of  $\alpha$ -weak positive premises mining with fixed number of iterations parameter:



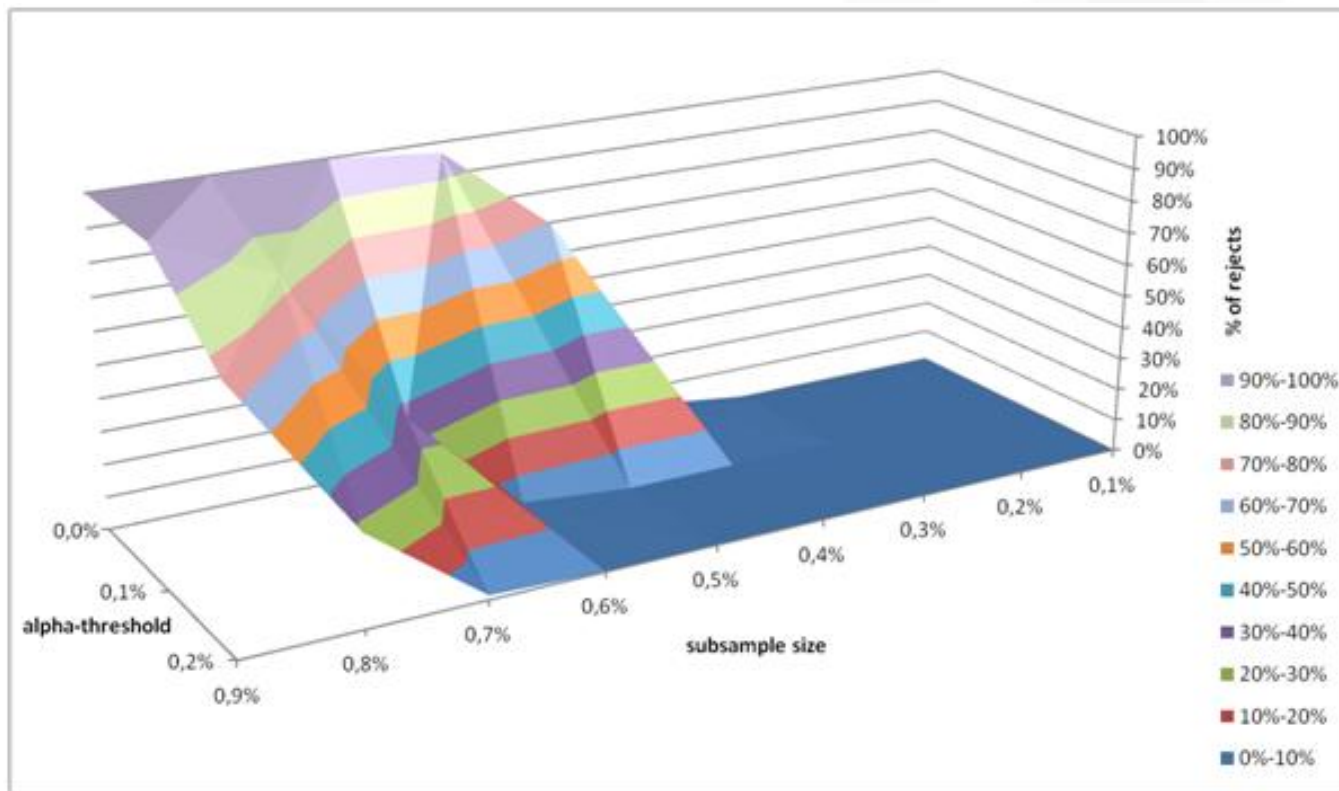
# Data and Experiments Results

- The dynamics of alpha - weak negative premises mining with fixed number of iterations parameter:



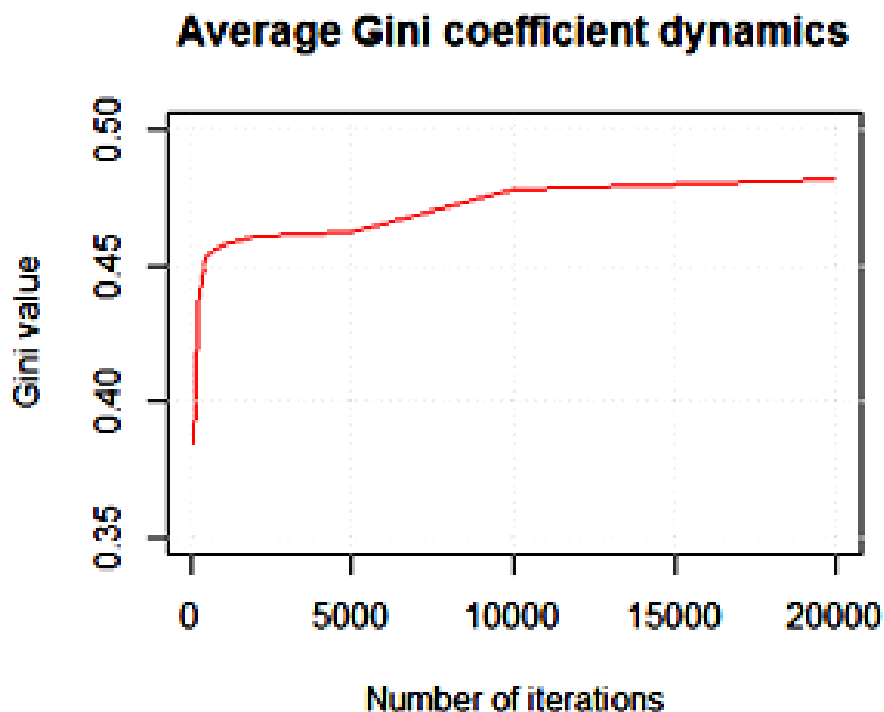
# Data and Experiments Results

- Below, the dynamics of the percentage of rejects from classifications. The abstain from classifications can arise in case if there was no premise found. This can be typical for low number of iterations and low alpha threshold:



# Data and Experiments Results

- Average Gini grouped by the different number of iterations (over all other parameter values):



# Appendix

- Below we present the classification accuracy obtained for different combinations of parameters (grid search).

|                 |                      | Subsample size |      |      |      |      |      |      |      |      |  |
|-----------------|----------------------|----------------|------|------|------|------|------|------|------|------|--|
| Alpha-threshold | Number of iterations | 0.1%           | 0.2% | 0.3% | 0.4% | 0.5% | 0.6% | 0.7% | 0.8% | 0.9% |  |
| 0.0%            | 100                  | 40%            | 44%  | 39%  | 18%  | 1%   | 0%   | 0%   | 0%   | 0%   |  |
|                 | 150                  | 35%            | 46%  | 35%  | 5%   | 0%   | 0%   | 0%   | 0%   | 0%   |  |
|                 | 200                  | 42%            | 37%  | 36%  | 12%  | 5%   | 1%   | 0%   | 0%   | 0%   |  |
|                 | 500                  | 39%            | 44%  | 44%  | 25%  | 6%   | 1%   | 0%   | 0%   | 0%   |  |
|                 | 1000                 | 44%            | 47%  | 44%  | 41%  | 11%  | 3%   | 0%   | 0%   | 0%   |  |
|                 | 2000                 | 44%            | 48%  | 46%  | 36%  | 17%  | 4%   | 0%   | 0%   | 0%   |  |
| 0.1%            | 100                  | 33%            | 37%  | 40%  | 40%  | 44%  | 43%  | 34%  | 32%  | 34%  |  |
|                 | 150                  | 41%            | 34%  | 33%  | 43%  | 41%  | 47%  | 41%  | 37%  | 37%  |  |
|                 | 200                  | 40%            | 40%  | 34%  | 42%  | 51%  | 43%  | 44%  | 41%  | 36%  |  |
|                 | 500                  | 37%            | 42%  | 47%  | 49%  | 51%  | 49%  | 43%  | 41%  | 34%  |  |
|                 | 1000                 | 37%            | 42%  | 46%  | 48%  | 49%  | 48%  | 43%  | 43%  | 37%  |  |
|                 | 2000                 | 39%            | 43%  | 45%  | 49%  | 51%  | 49%  | 46%  | 41%  | 38%  |  |
|                 | 5000                 | 43%            | 40%  | 44%  | 49%  | 46%  | 50%  | 48%  | 38%  | 36%  |  |
| 0.2%            | 100                  | 29%            | 38%  | 42%  | 32%  | 43%  | 37%  | 46%  | 43%  | 37%  |  |
|                 | 150                  | 27%            | 42%  | 41%  | 41%  | 36%  | 47%  | 48%  | 45%  | 41%  |  |
|                 | 200                  | 32%            | 40%  | 43%  | 42%  | 42%  | 49%  | 46%  | 47%  | 48%  |  |
|                 | 500                  | 39%            | 46%  | 46%  | 48%  | 47%  | 48%  | 51%  | 48%  | 51%  |  |
|                 | 1000                 | 41%            | 50%  | 48%  | 47%  | 49%  | 53%  | 52%  | 52%  | 47%  |  |
|                 | 2000                 | 38%            | 48%  | 50%  | 48%  | 47%  | 53%  | 52%  | 53%  | 50%  |  |
| 0.3%            | 100                  | 35%            | 38%  | 39%  | 42%  | 39%  | 45%  | 34%  | 45%  | 39%  |  |
|                 | 150                  | 27%            | 43%  | 44%  | 42%  | 42%  | 39%  | 37%  | 40%  | 46%  |  |
|                 | 200                  | 34%            | 46%  | 47%  | 45%  | 49%  | 47%  | 45%  | 45%  | 52%  |  |
|                 | 500                  | 31%            | 45%  | 49%  | 50%  | 49%  | 46%  | 50%  | 51%  | 47%  |  |
|                 | 1000                 | 37%            | 48%  | 49%  | 49%  | 49%  | 47%  | 52%  | 51%  | 51%  |  |
|                 | 2000                 | 38%            | 46%  | 48%  | 51%  | 51%  | 50%  | 50%  | 52%  | 52%  |  |
|                 | 5000                 | 40%            | 47%  | 46%  | 51%  | 52%  | 51%  | 49%  | 51%  | 53%  |  |
|                 | 10000                | 40%            | 44%  | 43%  | 46%  | 46%  | 48%  | 50%  | 52%  | 54%  |  |
|                 | 20000                | 40%            | 43%  | 42%  | 46%  | 47%  | 49%  | 50%  | 52%  | 53%  |  |