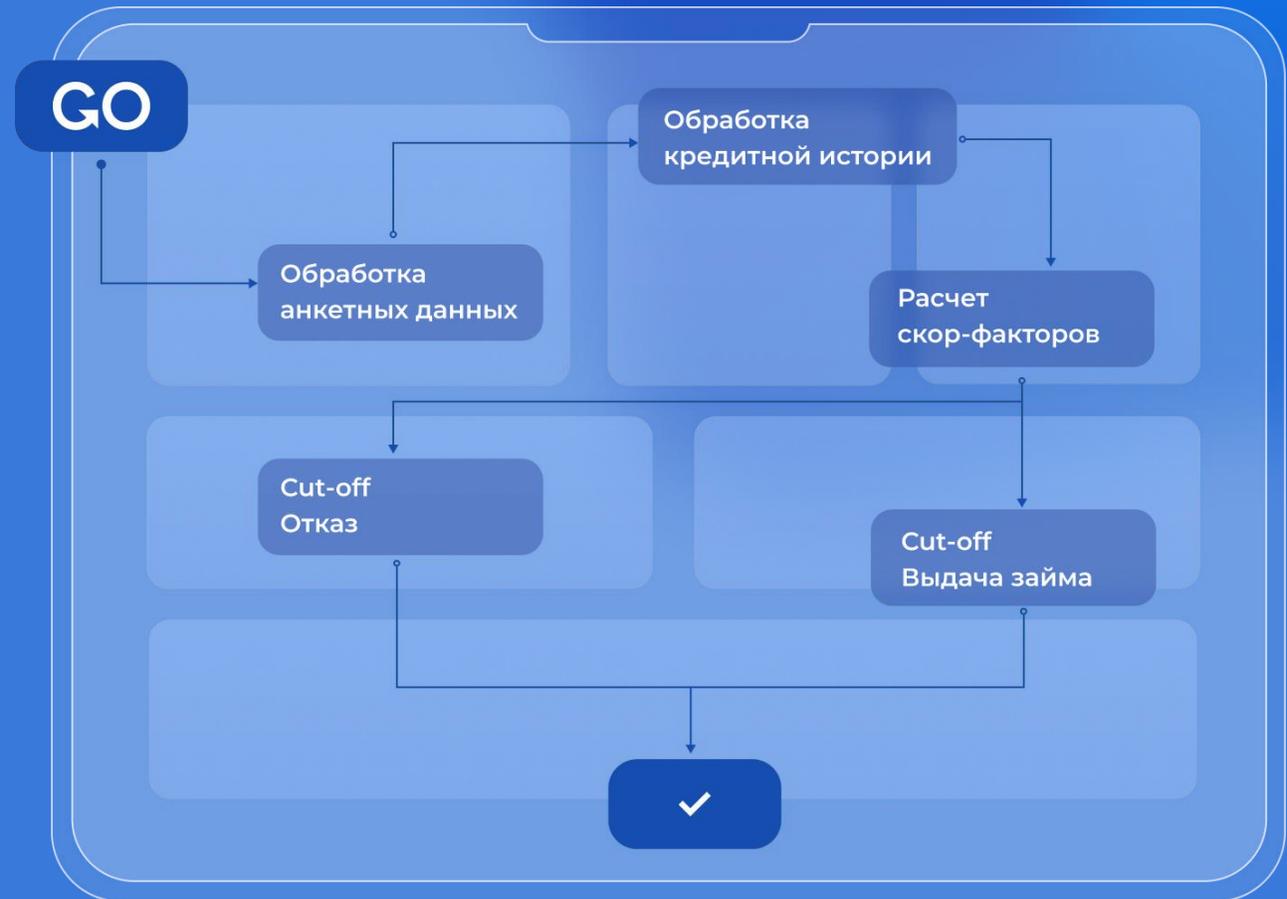


SCORINGO,

Инструмент оценки
заёмщиков для
банков и МФО



Почему Scoringo это важная часть в деятельности МФО



До 58 млн. рублей в месяц можно экономить, оптимизируя только риск!

Оптимизируют риск специалисты и технологии. Причем без технологий специалисты будут безоружными.

	Новые клиенты	Повторные клиенты	Итого
Выдано	100 000 000	400 000 000	500 000 000
Уровень финальных потерь min	22%	15%	
Потери min	22 000 000	60 000 000	82 000 000
Уровень финальных потерь max	40%	25%	
Потери max	40 000 000	100 000 000	140 000 000

Зависимые бизнес метрики: главная – потери; вторичные – data expenses, marketing expenses.
 Размер финансового результата за месяц, который оптимизирует модуль на примере МФО, которое выдает 500 млн. рублей в месяц с пропорцией new/old 20:80.

Scoringo – отдельный модуль в ИС (информационной системе) МФО

Осуществляет три основные функции:

1

«Окно»

Инструмент для взаимодействия с поставщиками данных о благонадёжности потенциальных клиентов;

2

«Данные для DM»

До 80% (или вся «не творческая» часть) процесса data mining. Оставшиеся 20% совершаются в любой удобной для data scientist среде и максимально легко имплементируются в СПР;

3

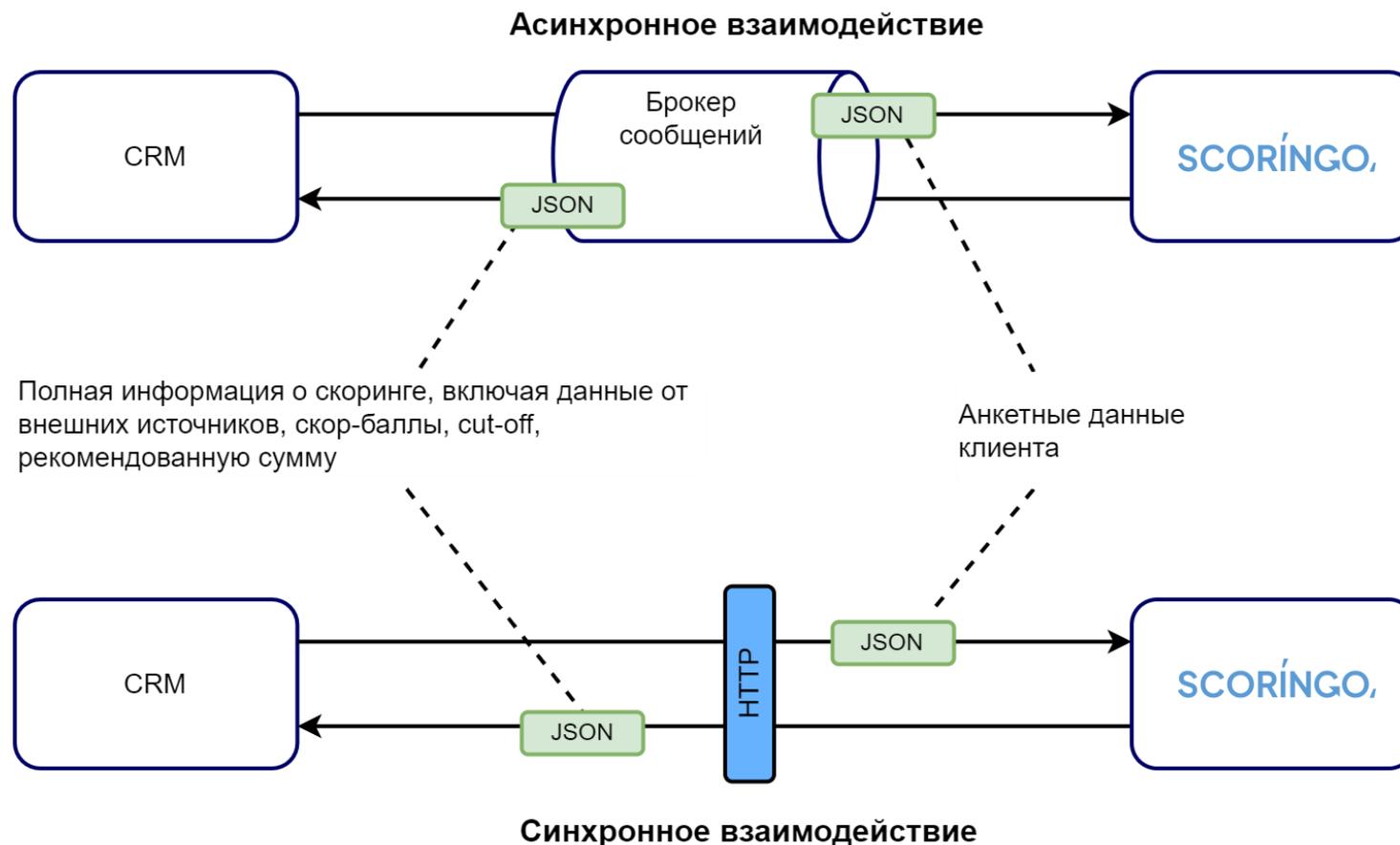
«СПР»

Система принятия решения. Управление логикой принятия решения по кредитной заявке.

Компонентная диаграмма решения:



Интеграция Scoringo в экосистему МФО:



Функция «Окно»: интеграции с поставщиками

НБКИ

- Кредитная история
- Скоринговый балл

EQUIFAX®

- Кредитная история
- Скоринговый балл



- Скоринговый балл на основе карточных транзакций

билайн

- Аналитика использования номера, ФИО

JuicyScore

- Анти-фрод инструмент для финансового сектора на основании неперсональных данных

@mail

- Оценка заемщика по профилю
- Скоринг

(ONE)FACTOR

- Данные Мегафон +

QIWI

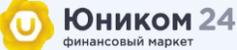
- Скоринг
- Интересы
- Время жизни номера
- Количество платежей
- География платежей
- Сумма платежей за период

Какие еще бывают поставщики:

Поставщик	Тип предоставляемой информации
 ОКБ	Кредитная история, скоринговый балл, прочее
 РУССКИЙ СТАНДАРТ	Кредитная история, скоринговый балл
 croinform.ru Кронос (МБКИ)	Анализ благонадежности физлица
 DOUBLE DATA	Атрибуты социального профиля анализ текстов, интересов заемщика предотвращение мошенничества
 SocioHub FACHINFORMATIONSDIENST SOZIOLOGIE	Анализ данных соцсетей
 МТС	Скоринг бюро, аналитика пользования номера, ФИО
 МЕГАФОН	Скоринг бюро, аналитика пользования номера, ФИО
 ФССП	Наличие либо отсутствие задолженности, арестованное имущество

Поставщик	Тип предоставляемой информации
 ФМС	Действительность документа по данным ФМС
 ФС по финансовому мониторингу	Перечень организаций и физических лиц, в отношении которых имеются сведения об их причастности к экстремистской деятельности или терроризму
 ФНС Госналог	Выписки ЮГРЮЛ и ЕГРИП Сведения: ликвидация, банкротство, недобросовестность. Сведения о задолженности по налогам.
 ЕФ реестр сведений о банкротстве	Сведения о банкротстве физических и юридических лиц
 CyberPlat®	Оценка заемщика
Mandarin	Данные по карточным транзакциям

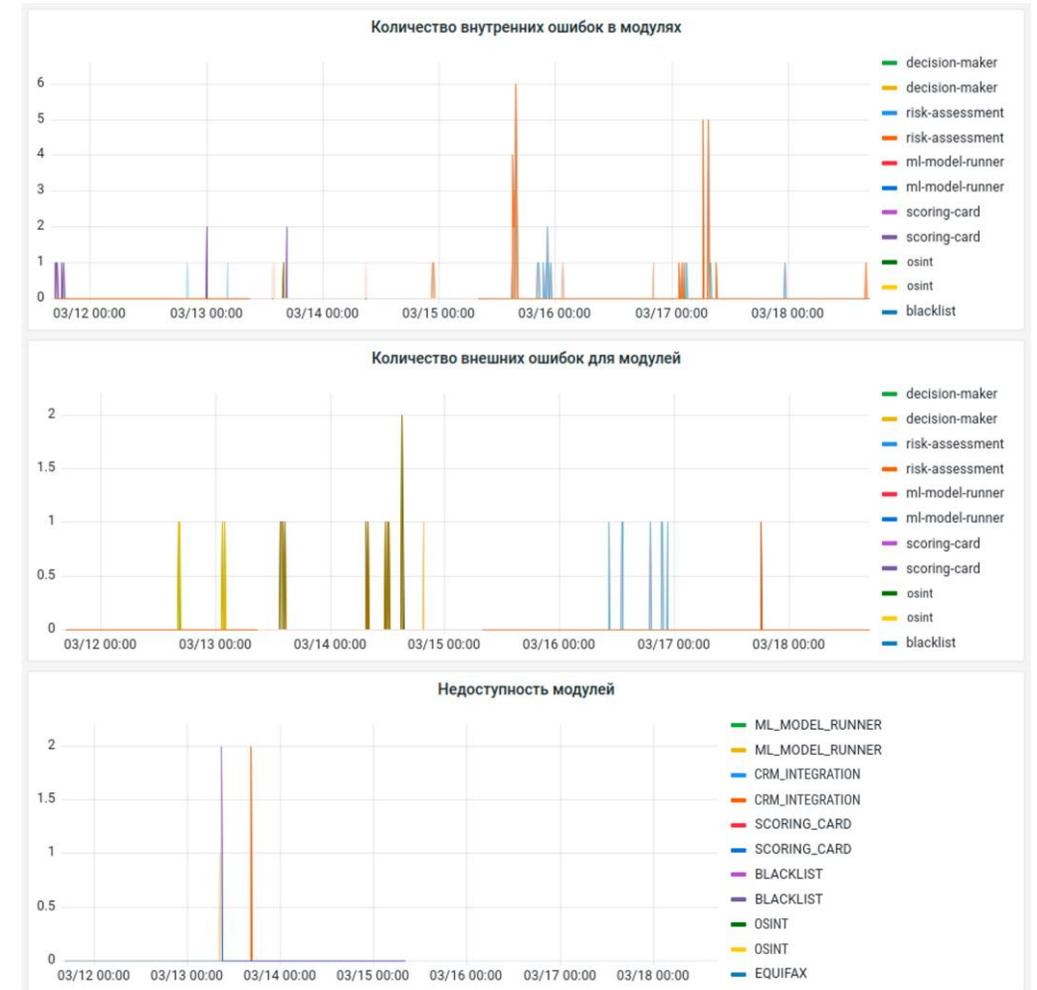
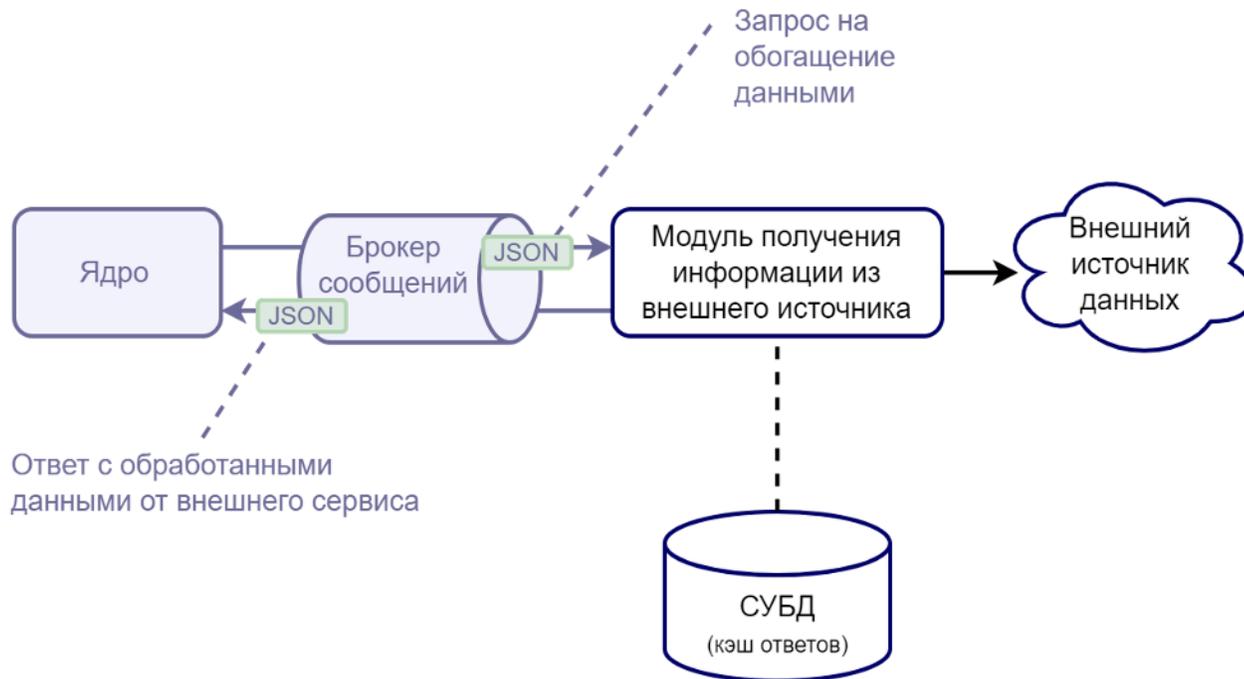
Какие еще бывают поставщики:

Поставщик	Тип предоставляемой информации
 ФИНКАРТА	История предоставления займов и попыток погашения займов на указанные при подаче заявки пластиковые карты
 CleverData	Сервис позволяет потребителям самостоятельно выбрать нужные данные на Бирже поставщика
 Юникон 24 финансовый маркет	Скоринг, кредитный отчет
 CASHOFF	Облачный сервис, собирающий финансовую информацию из разных источников
 ЦИАН	Анализ объектов недвижимости
Casebook	Оценка факторов риска и исковой нагрузки потенциальных контрагентов
ТРАНСДЕКРА	Расшифровка параметров ТС, стоимость, регистрация

Поставщик	Тип предоставляемой информации
AudaVIN	Кол-во расчётов, сумма ремонта
 SEON	Предотвращение мошенничества
mfms	Сервис по предоставлению данных анализа входящих смс отправленных на абонентский номер заявителя финансовыми компаниями
 spectrum data	Платформа решений для бизнеса на основе больших данных о физических лицах, организациях, автомобилях и объектах недвижимости.
Контур Фокус	Проверка данных по юрлицам и ИП

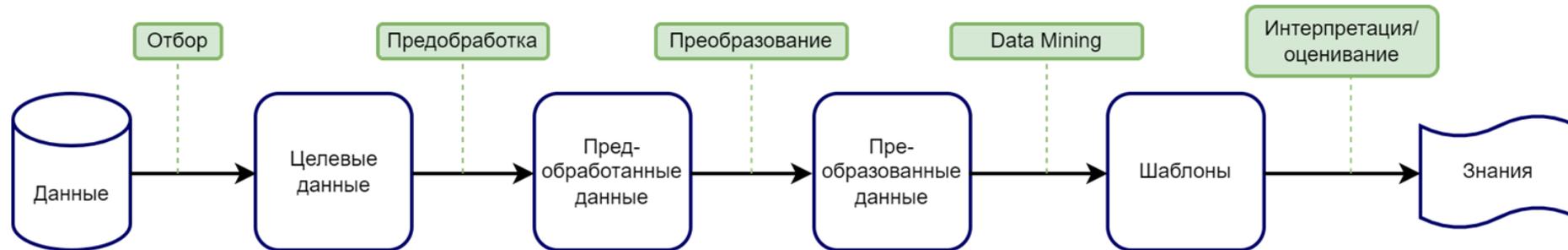
«Окно» Архитектура и технологии

Интеграция с внешним источником:



Функция «Данные для DM»

Этапы Data mining:



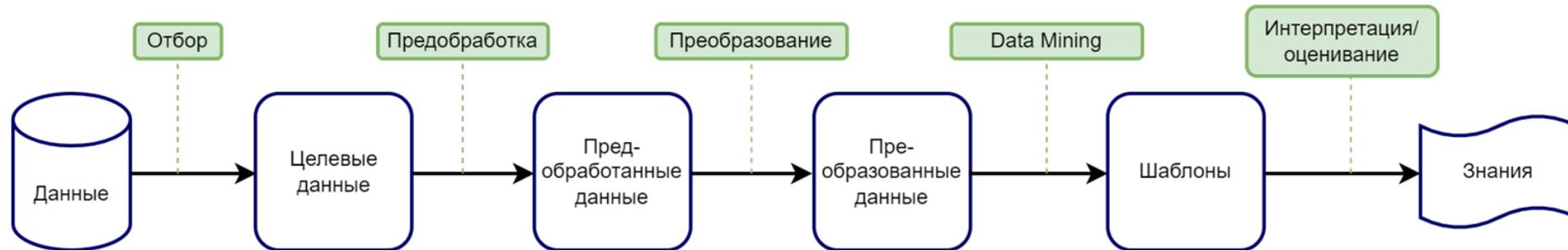
При использовании Scoringo, упрощаются или полностью реализуются все этапы, кроме непосредственно Data mining.



Отбор, и предобработка данных производится в Scoringo один раз для каждого нового источника данных.

Функция «Данные для DM»

Этапы Data mining:



Преобразование данных

– это этап генерации фич

Признак (фича, feature) – это переменная, которая описывает отдельную характеристику объекта. Бывают фичи с бинарными, номинальными (список значений), количественными (числовыми) значениями.

В системе Scoringo фичи создаются при помощи java кода «программистами», либо при помощи SQL-кода «риск-аналитиками».

Большинство полезных фич из самых популярных источников для Scoringo уже создано на java



Интерпретация/оценивание

– это этап заключающийся в применении к собранному вектору данных модели, разработанной на этапе Data mining во внешней среде.

Модель переносится в Scoringo максимально просто, на тип модели и внешнюю среду разработки практически нет ограничений. На выходе Scoringo позволяет получать не только вероятность дефолта, но и назначать в зависимости от нее лимиты, ставки, продукты благодаря отдельному функционалу.

Функция «Данные для DM»



Какие данные идут в комплекте:

В коробке идут 200+ рассчитанных скор-факторов с описанием логики

Правила расчета скор-факторов по КИ НБКИ

Код	Название	Описание	Тип данных	Описание расчета
NBCH_1	BAD_LOAN_COUNT	Количество проблемных кредитов в КИ	int64	Количество всех неодинаковых accountType, в которых accountRating = 52, 61 или 14.
NBCH_2	ACTIVE_MICRO_LOAN_WITHOUT_PMNT_PCT	Процент активных микрозаймов без платежа среди всех кредитов	float64	Отношение займов в которых accountRating != 13, currentBalanceAmount = 0 и accountType = 16 к общему количеству, в процентах.
NBCH_3	CLOSED_MICRO_LOAN_COUNT	Количество закрытых микрозаймов	int64	Количество всех AccountReply, где accountType = 16 и accountRating = 13.
NBCH_4	LAST_MICRO_LOAN_STATUS	Статус последнего микрозайма	float64	Значение accountRating, где accountType = 16, serialNum = max(serialNum) при max(openedDt). Ищем кредиты с максимальной последней датой. Если их несколько на одну дату, выбираем из них тот, у которого serialNum больше.
NBCH_5	LAST_LOAN_DAYS	Количество дней с открытия последнего кредита до обращения	int64	Разница между датой создания заявки и максимальной датой кредита (openedDt).
NBCH_6	REQ_AMOUNT_TO_MAX_PCT	Процентное отношение запрашиваемой суммы к максимальному выплаченному кредиту	float64	Доля loanAmountRequested (из анкетных данных) к максимальному значению creditLimit в процентах.
NBCH_7	SUM_OVERDUE	Общая сумма просроченной задолженности (без ипотеки)	float64	Сумма по всем amountPastDue с accountType не равно 6. В случае amountPastDue = NULL возвращает 0.
NBCH_8	DELAY_LOAN_COUNT	Кол-во просроченных	float64	Количество кредитов с accountType = 16 и accountRating = 52.

Функция «Данные для DM». Как создать новый фактор.

Есть 2 сценария - первый этап риск-аналитик, второй программист.

1 Создание скор-фактора в Java:

Java - надёжность и скорость работы системы.

```

/**
 * Возраст заёмщика - количество лет между датой рождения и датой подачи заявки на займ
 */
@Component
public class Age extends AbstractScoringRule {

    public Age() {
        super(ScoringRuleType.AGE);
    }

    @Override
    public Long calculate(ScoringData scoringData) {

        LocalDateTime loanApplicationApplied = scoringData
            .getRegistrationData()
            .getApplicationCreated();

        LocalDate registrationFormClientBirthDate = scoringData
            .getRegistrationData()
            .getPerson()
            .getBirthDate();

        return ChronoUnit.YEARS.between(registrationFormClientBirthDate, loanApplicationApplied.toLocalDate());
    }
}

```

2 Создание скор-фактора из графического интерфейса:

SQL - скорость и меньше надёжность.

Редактирование скор-фактора

← Назад

Наименование*
AGE x Пользовательский

Комментарий
Возраст заёмщика x

Источник данных*
MODULE

SQL запрос*
SELECT DATE_PART('year', AGE(cast(model -> 'model' -> 'registrationData' -> 'person' -> 'birthDate' as varchar), cast(model -> 'model' -> 'registrationData' -> 'applicationCreated' as varchar))) FROM datasource_interactions WHERE uuid = uuid ORDER BY id DESC LIMIT 1

Наименование параметра	Наименование поля в скоринговой модели
uuid	uuid

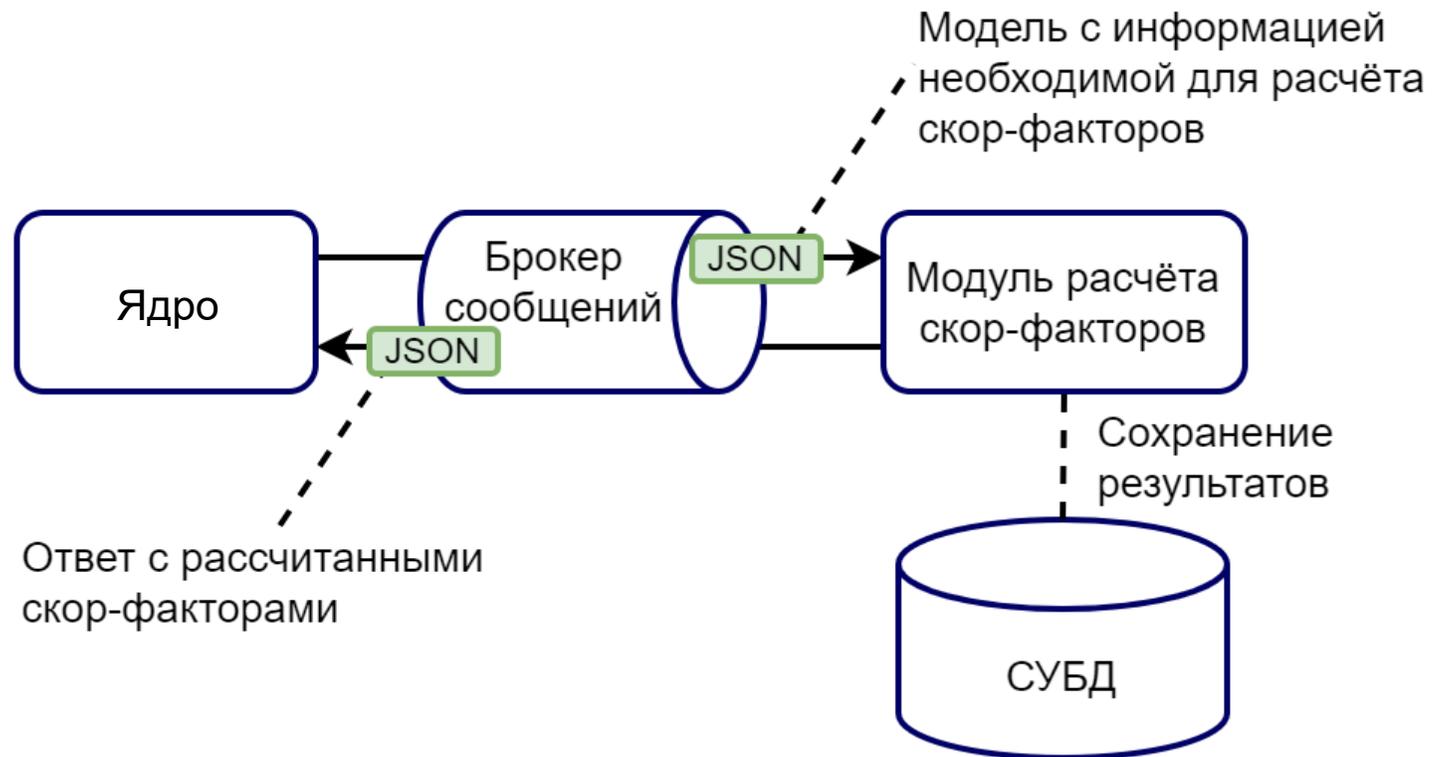
+ Добавить параметр

Удалить Сохранить

История изменений: 2023-09-17 17:18:58
Скоринговые карты: FIRST_RBL

Функция «Данные для DM». Архитектура и технологии.

Отдельный микро-сервис для расчёта скор-факторов:



Функция «СПР». Логика принятия решения в зрелой МФО.

Тренды:

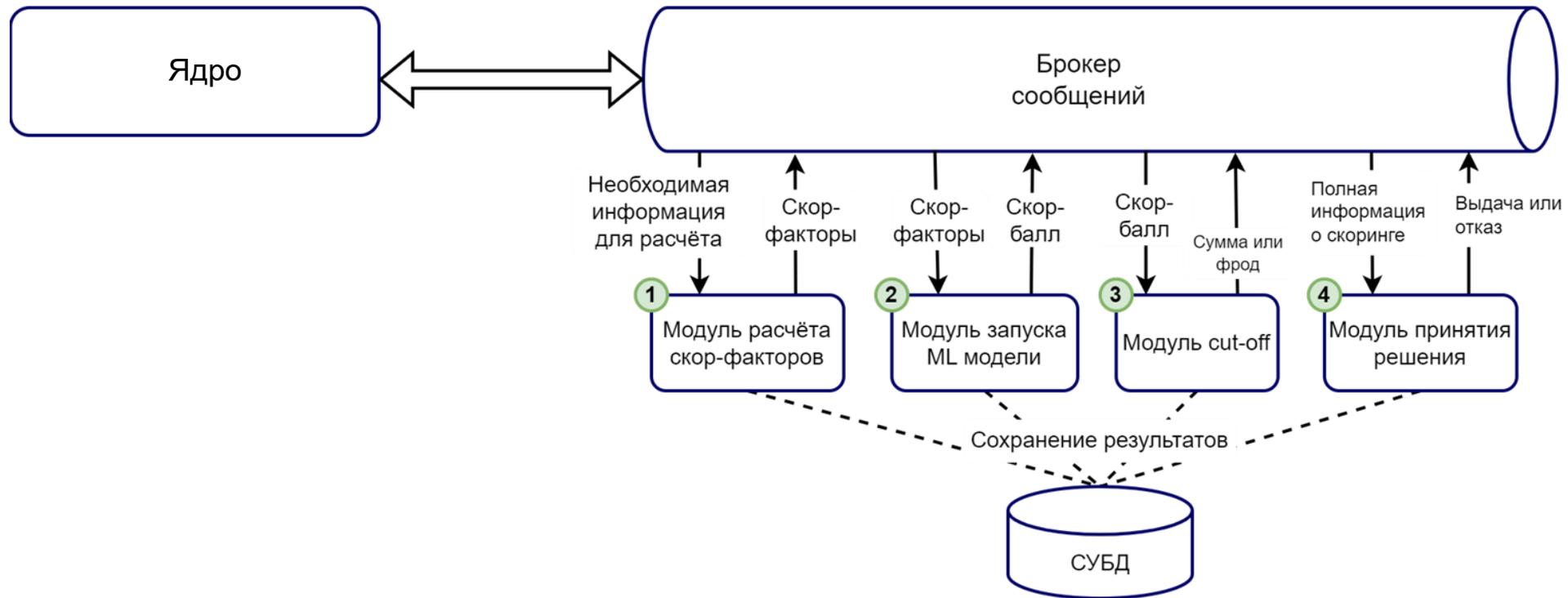
- 1 Разделение моделей для 1,2,3,4+ займов;
- 2 Вынос вперед бесплатных проверок;
- 3 Использование “чистого ручейка” для улучшения моделей;
- 4 Отдельные антифрод модели, вынесенные до моделей оценки кредитного риска;
- 5 Антифрод отсечение правилами на факторах, правилами на группах, скорингом;
- 6 Последовательное использование кредитных бюро;
- 7 Отсечение по данным бюро правилами до того, как отсекается картой совместно с другими данными;
- 8 Использование моделей на нейросетях, а не на логистической регрессии;
- 9 Выделение отдельных моделей под сегменты (кластеры);
- 10 Правила для назначения Risk-based limits и Risk-based pricing;
- 11 Внедрение изменения через champion/challenger по принципу раскатывания 10%/50%/100%.



Достаточно большие деревья.
Пример: 60 элементов в процессе для оценки только первичных клиентов.

Функция «СПР». Архитектура и технологии этой части модуля.

Система принятия решения:



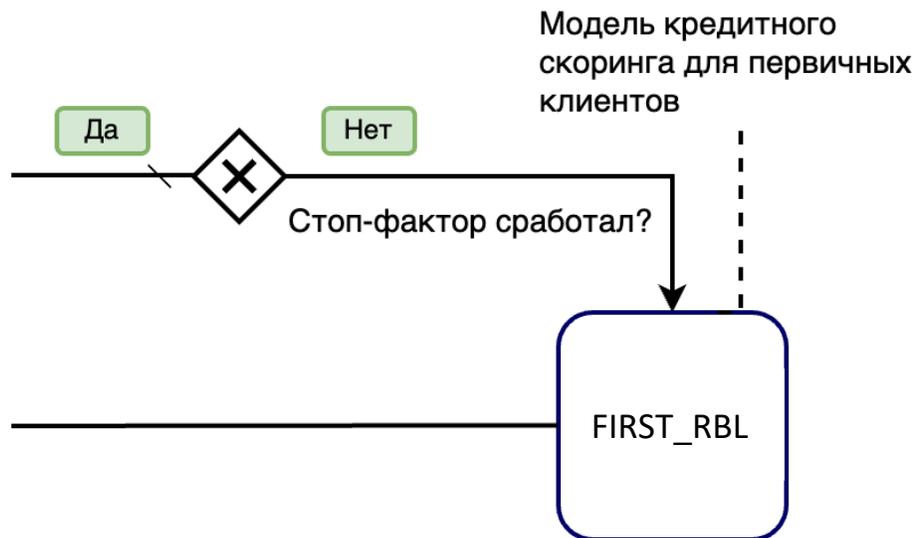
Функция «СПР». Какие типы узлов поддерживаются в диаграмме процесса.

- 1 Запрос анкетных данных из CRM;
- 2 Рандомизация для выбора альтернативной стратегии, расчёт дополнительных атрибутов;
- 3 Отдельные типы узлов для каждого внешнего источника данных;
- 4 Расчёт стоп-факторов;
- 5 Расчёт скор-карт;
- 6 Запуск ML модели;
- 7 Применение балла отсеечения;
- 8 Принятие решения;
- 9 Определение лимита суммы к выдаче;
- 10 Определение продукта (процентной ставки);
- 11 Модуль поддержки тестирования (ретро и прочие тесты).

Функция «СПР»



Заглядываем детально в узел с обчетом скоркарты и разбираемся как все работает от картинки на схеме до получения оценки вероятности дефолта по заявке.



Change value for "In parameters"

Source	Source express...	Target	Target expressi...
	SCORING_CARD	dataSource	
	FIRST_RBL	scoringCards	

Source: Enter a source

Source expression: FIRST_RBL

Target: scoringCards

Target expression: Enter a target expression

↑ ↓ + -

Cancel Save

Функция «СПР». Редактирование скоринговой карты.

☰ Редактирование скоринговой карты

← Назад

Наименование*

FIRST_RBL X

Комментарий

RBL для первичных клиентов X

Тип*

Скоринг

№ Наименование скор-ф

5	BEELINE
6	COUNT_ACTIVE_MFO
7	AGE
7	LIFETIME_SMALL_CR
7	SUM_ACTIVE

Удалить Сохранить

📘 Информация

Дата последнего обновления:
2023-09-17 17:19:14

№	Наименование скор-фактора	Тип скор-фактора	Импортировать	Добавить
5	BEELINE	Встроенный		X
6	COUNT_ACTIVE_MFO	Встроенный		X
7	AGE	Встроенный		X
7	LIFETIME_SMALL_CRED	Встроенный		X
7	SUM_ACTIVE	Встроенный		X

Функция «Данные для DM». Интерфейсы, мониторинги, отчеты и т. д.

☰ Просмотр скоринга

← Назад Атрибуты Детали работы

Детали работы процесса скоринга

№	Источник данных	Параметры источника	Дата и время запроса	Дата и время ответа	
EQUIFAX		SCORE	2023-09-17 16:03:46	2023-09-17 16:03:46	🔗
QIWI			2023-09-17 16:03:46	2023-09-17 16:03:46	🔗
MAILRU		SCORE	2023-09-17 16:03:46	2023-09-17 16:03:46	🔗
BEELINE			2023-09-17 16:03:46	2023-09-17 16:03:46	🔗
NPCS			2023-09-17 16:03:47	2023-09-17 16:03:47	🔗
ONE_FACTOR			2023-09-17 16:03:47	2023-09-17 16:03:47	🔗
SCORING_CARD		FIRST_RBL	2023-09-17 16:03:47	2023-09-17 16:03:47	🔗
ML_MODEL_RUNNER		FIRST_RBL	2023-09-17 16:03:47	2023-09-17 16:03:48	🔗
RISK_ASSESSMENT		FIRST_RBL	2023-09-17 16:03:49	2023-09-17 16:03:49	🔗
DECISION_MAKER		RISK_ASSESSMENT_CUTOFF FIRST_RBL	2023-09-17 16:03:49	2023-09-17 16:03:53	🔗

```
    },
    "scoringCardReports": [
      {
        "code": "FIRST_RBL",
        "totalScore": 435.673,
        "cutOffValue": 7000,
        "cutOffType": "RBL",
        "cutOffGridCode": "FIRST_RBL",
        "scoringRuleDetails": [
          { ... },
          { ... },
          { ... },
          { ... },
          { ... },
          { ... },
          {
            "ruleCode": "AGE",
            "value": "41"
          },
          { ... },
        ]
      }
    ]
  }
```

Функция «СПР». Интерфейсы, мониторинги, отчеты и т. д.

The screenshot displays a web interface titled "Cut-off". It features a table with four columns: "№", "Название сетки", "Дата последнего обновления", and "Тип сетки". The table contains four rows of data. Below the table is a blue button labeled "Создать сетку".

№	Название сетки	Дата последнего обновления	Тип сетки
1	FIRST_AF_1	15.09.2023 05:22	Бинарная
2	FIRST_AF_2	15.09.2023 05:22	Бинарная
3	FIRST_RBL	15.09.2023 15:44	RBL
4	REPEAT_RBL	15.09.2023 05:22	RBL

[Создать сетку](#)

Функция «СПР». Матрица RBL.

Интерфейс пользователя для регулирования сумм выдачи от скор балла.

Редактирование скор-фактора

← Назад Копировать сетку Сгенерировать сетку с шагом

Тип сетки* RBL × Название сетки* FIRST_RBL ×

Комментарий
Введите комментарий

Информация
Дата последнего обновления:
2023-09-15 16:23:55

Нижний диапазон	Верхний диапазон	Значение	
0	100	0	×
100	200	2000	×
200	300	3000	×
300	400	5000	×
400	500	7000	×
500	600	10000	×

Добавить запись

Удалить Сохранить

Работа не ИТ сотрудников с системой: Риск-стратег

Создает логику принятия решения:

- 1 Какие поставщики и вообще источники используются – те, что показали свою предсказательную силу на тестах и стоят адекватных денег;
- 2 Какие именно проверки осуществляются и в каком порядке – бесплатные разумно выносить вперед, самые дорогие в конце.
- 3 Сколько есть моделей для целевой переменной «дефолт» и для каких они сегментов – всегда разделяют модели для новых и повторных клиентов. Часто создают модели отдельно для вторых и третьих займов, для 4+ обычно единая. Если хорошо выделяются кластеры, то будут ли модели под них.;
- 4 Будут ли использоваться модели с другими целевыми переменными – часто дополнительно делают модель на целевую «фрод», реже на «дефолт, но коллектится»;
- 5 Будет ли использоваться «чистый ручеек»;
- 6 Настраивает результат оценки риска – отсечение, лимитная политика, risk-based pricing.

Работа не ИТ сотрудников с системой: Data scientist

1

Создает модели в привычной ему среде;

2

Проводит ретротесты созданных моделей;

3

Генерит новые фичи. Ставит задачи на их разработку в java, либо создает сам в SQL.

Работа не ИТ сотрудников с системой: Anti-Fraud менеджер

1

Использует хранилище Scoringo, как базу для поиска групповых кейсов фрода;

2

Описывает найденный кейс исключаящим SQL запросом и имплементирует его в Scoringo.

3

Аналогичным образом исключается маркетинговый фрод.

Работа не ИТ сотрудников с системой: Прочие сотрудники

Во время работы так или иначе существует запрос на понимание, почему в конкретном случае произошел отказ или одобрение с неожиданным результатом. Для этого в Scoringo существует инструмент, который визуализирует процесс оценки каждой отдельной заявки. Доступ на данный инструмент выделяется нужному специалисту (например, руководителю Call Center), чтобы снять потенциально ненужный поток запросов с ИТ или Рисков.

Просмотр скоринга

← Назад Атрибуты Детали работы

Ключевые атрибуты процесса скоринга

№	Атрибут	Детали
1	ID Заявки	11111
2	Процесс	PresentationScoring
3	ID Клиента	20047942
4	Тип клиента	Первичный
5	Результат	Одобрение
6	Детали	Окончательная сумма: 7000 Сумма по ML модели: 7000

Ключевые атрибуты процесса скоринга

№	Атрибут	Детали
1	ID Заявки	11111
2	Процесс	PresentationScoring
3	ID Клиента	20047942
4	Тип клиента	Первичный
5	Результат	Отказ
6	Детали	BAD_PASSPORT

Дополнительные бизнес-фичи:

- 1 Champion challenger/Тестирование скор-карт;
- 2 Ретро скоринг;
- 3 Чистый ручеек;
- 4 Кэширование ответов поставщиков.

RetroScoring Test

 Таблица Фильтр

Решение

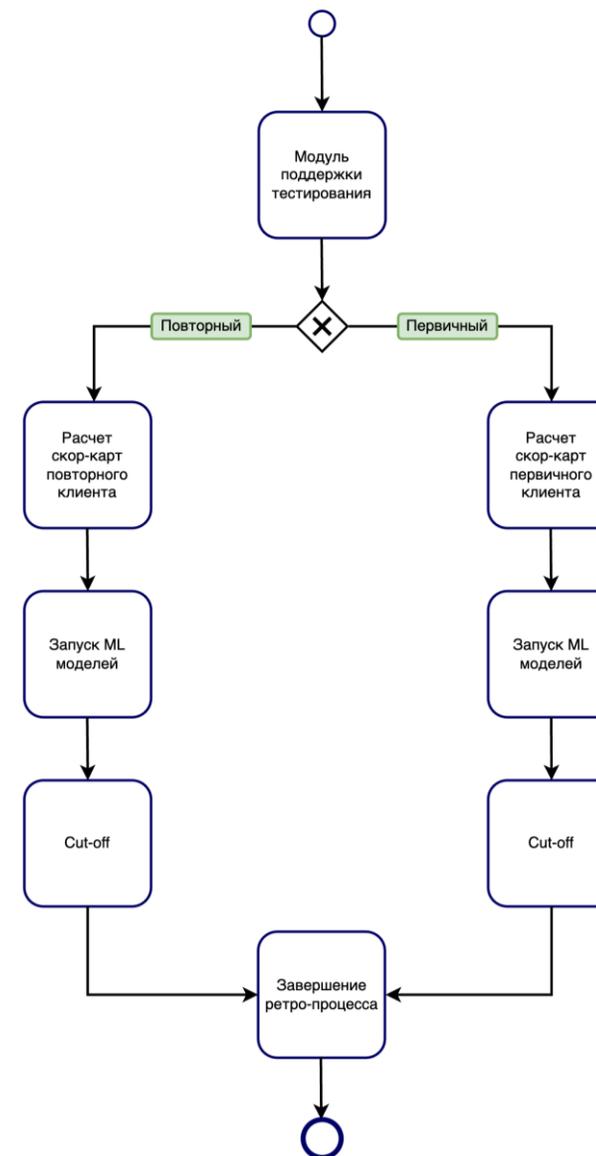
Тип клиента

Дата и время

Поиск

Обработано 100 из 1000 заявок 0 ошибок

Завершить ретро-тест



Технологии и инструменты:



Java 11



SpringBoot 2.6



Maven 3



Docker



Helm



Flowable BPMN



Kubernetes



PostgreSQL



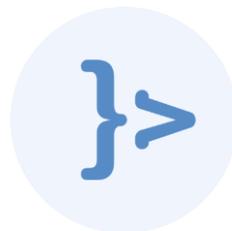
RabbitMQ



Prometheus



Grafana



VAADIN

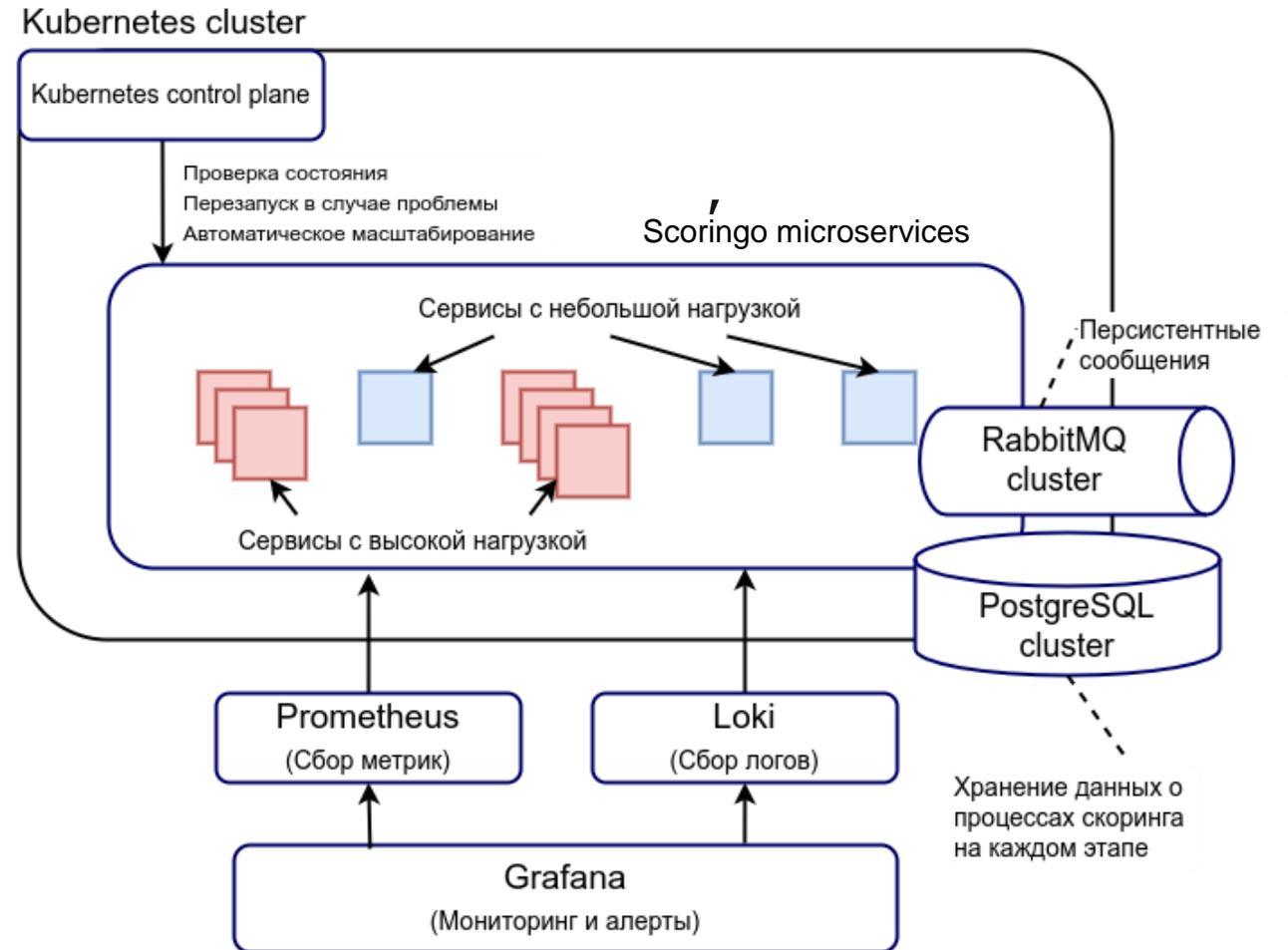
Про отказоустойчивость, мониторинг и масштабирование:

Конфигурация действующего контура:

- 1 сложный процесс, в котором используются 10 источников данных и считается 5 ML моделей
- 2 15 заявок на скоринг (в мин)
- 3 пиковая нагрузка 25-30 (в мин)
- 4 16 сервисов с 1 CPU и 1Gb RAM + 1 сервис с 2 CPU и 2Gb RAM
- 5 + Инфраструктура: PostgreSQL, RabbitMQ, Grafana, Loki, Prometheus, Kubernetes

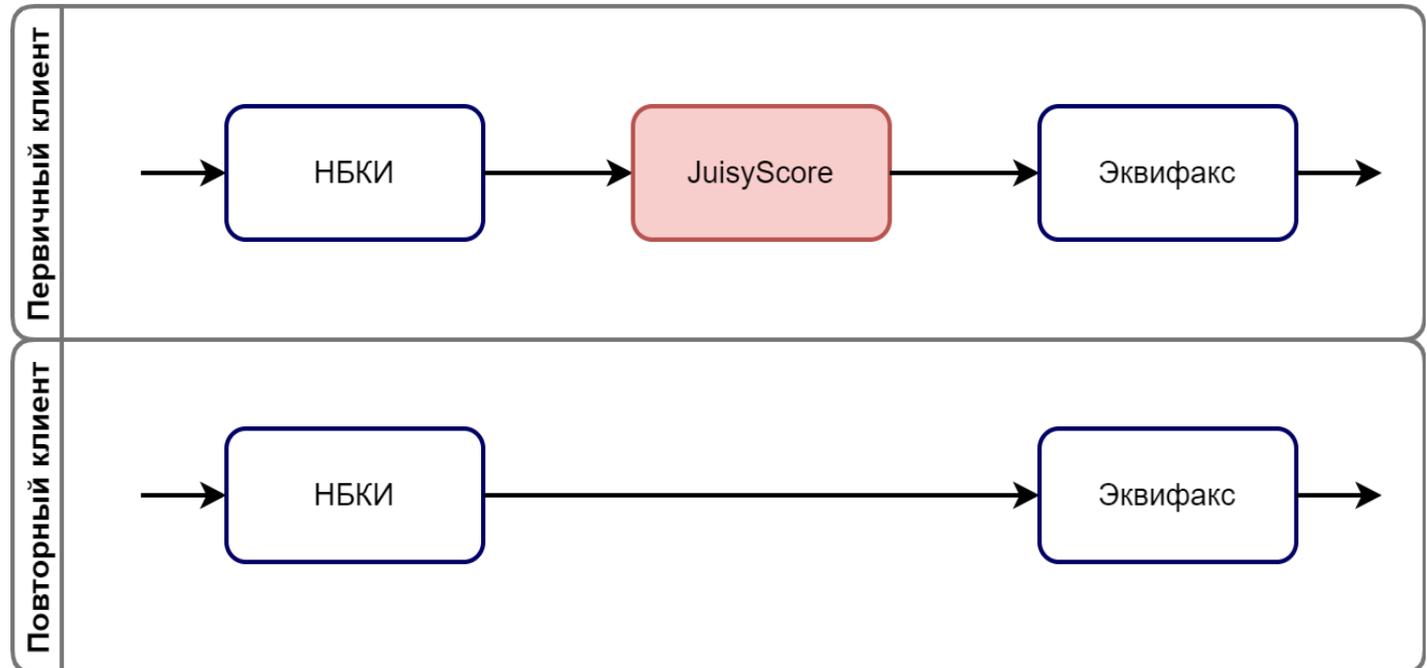
K8S:

- 1 Горизонтальное/вертикальное масштабирование
- 2 Отказоустойчивость 24x7
- 3 CI/CD с использованием Docker образов



Отказоустойчивость бизнес-логики

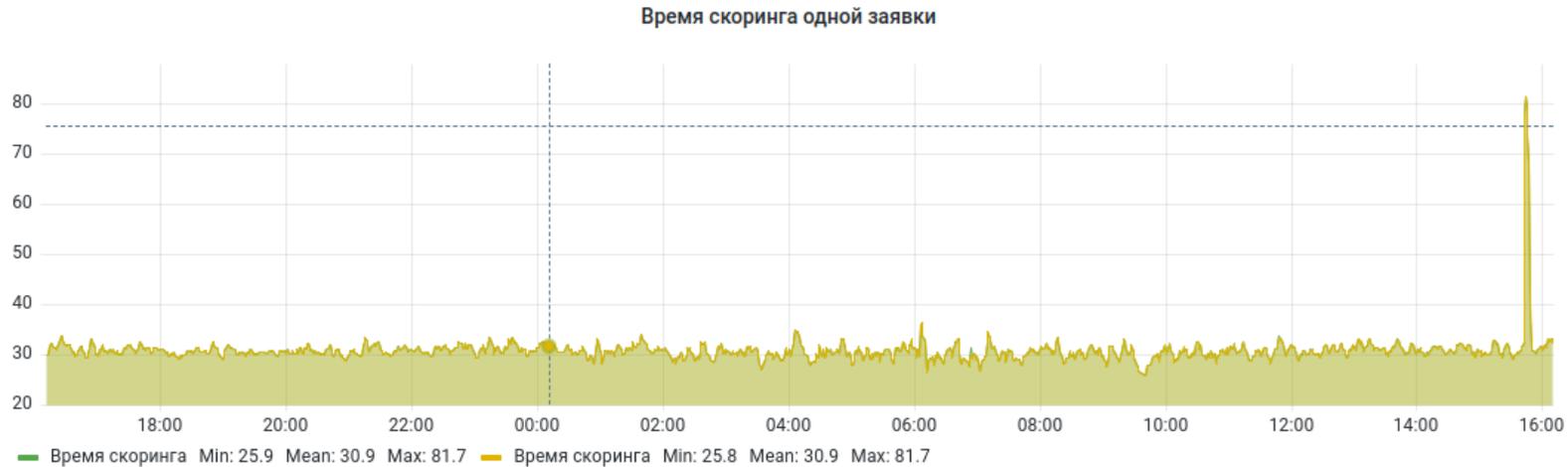
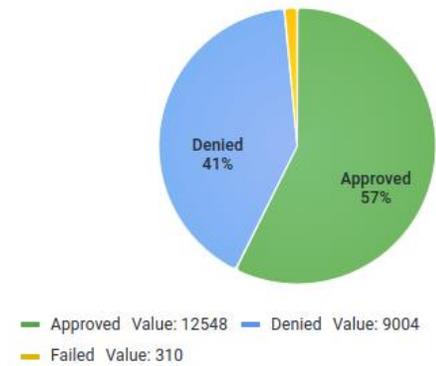
Основные проблемы связаны с получением данных от внешних источников, при этом система от этого не умирает, процесс восстановится после исправления ситуации. Если для заявки не нужен этот внешний источник, то она проскорится без ошибок.



Дополнительные материалы

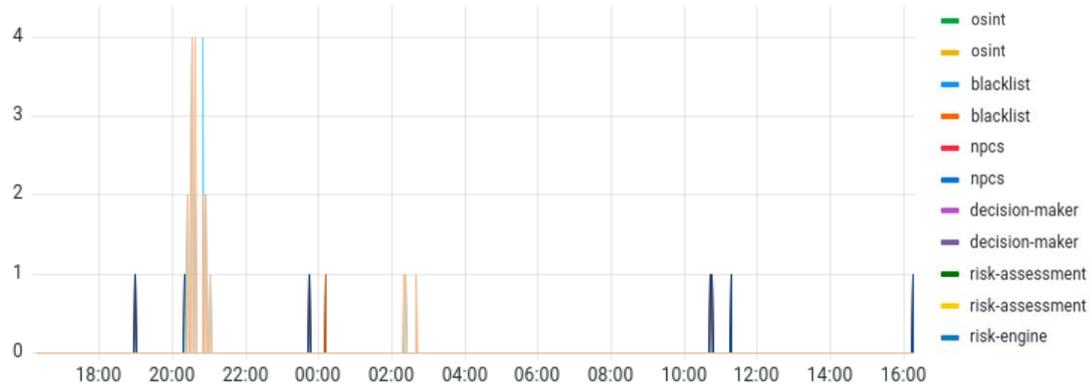


Успешные/неуспешные запросы

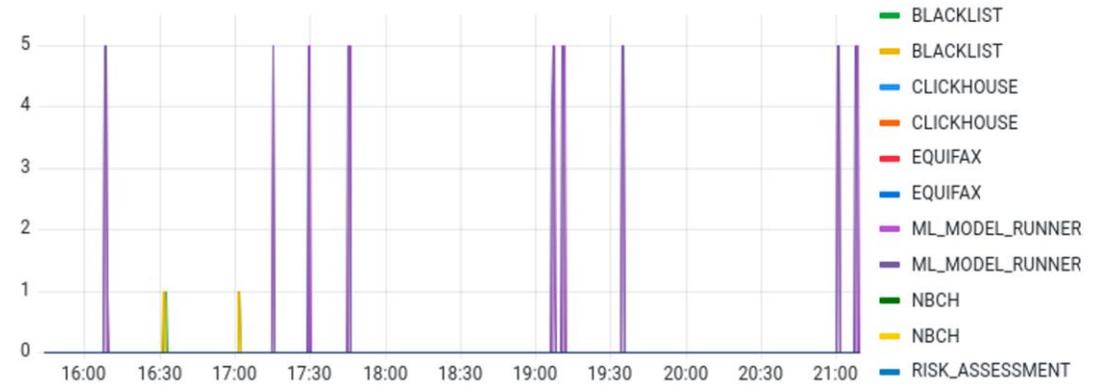


Дополнительные материалы

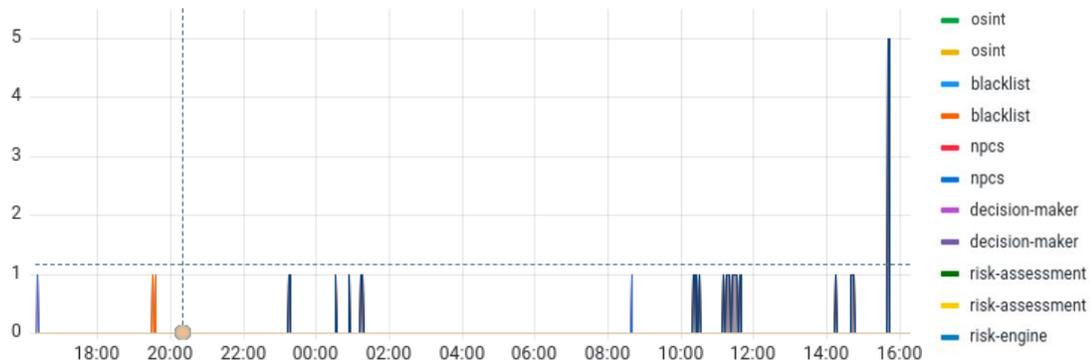
Количество внутренних ошибок в модулях



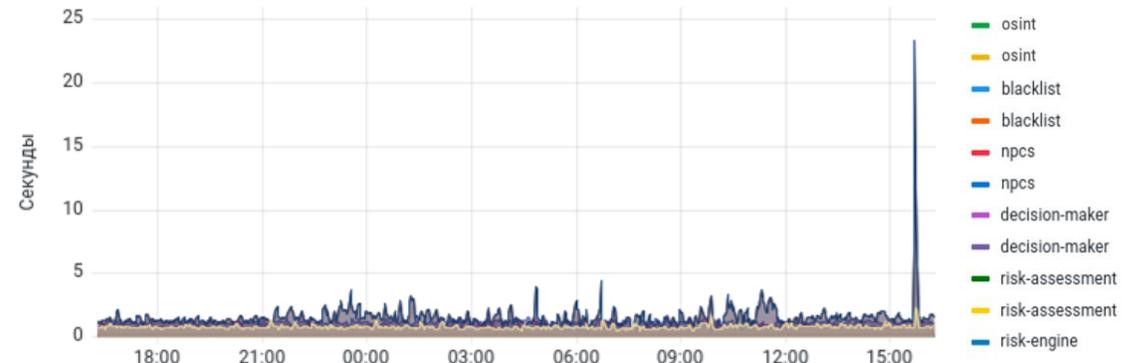
Недоступность модулей



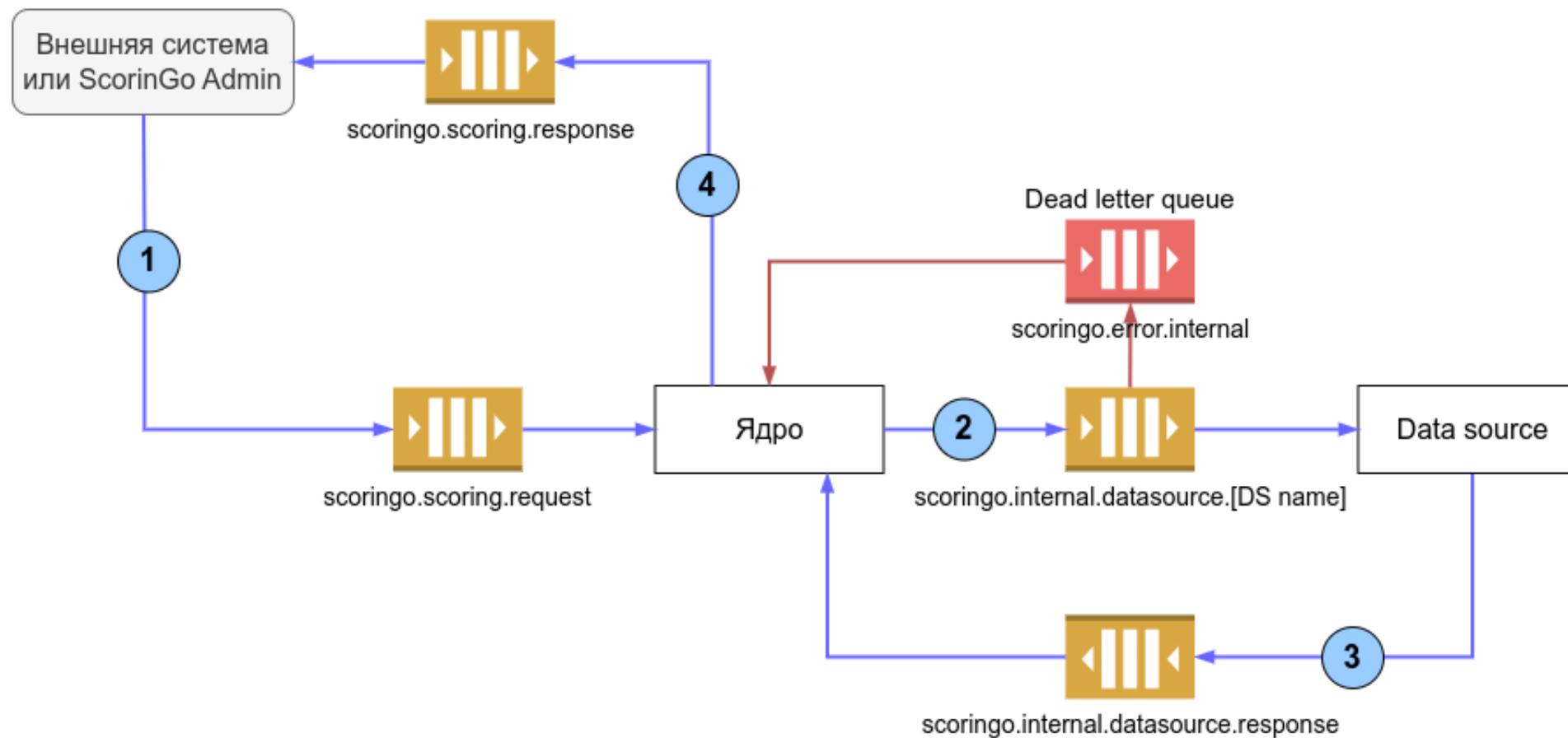
Количество внешних ошибок для модулей



Время обработки запросов модулями



Дополнительные материалы



SCORÍNGO,

Спасибо за внимание!



sales@intabia.ru



<http://www.intabia.ru/>