

A night city street scene with a tram and a digital overlay of binary code. The scene is illuminated by blue and yellow lights, with a tram visible in the distance. The binary code is overlaid on the right side of the image, creating a sense of digital data flow.

**Программное обеспечение учетных и
статистических операций**

Введение в бизнес-анализ

- Методика извлечения знаний (KDD)
- Data Mining
- Программное обеспечение в области анализа данных
- Аналитические платформы
- Языки визуального моделирования

Технологии KDD и DM

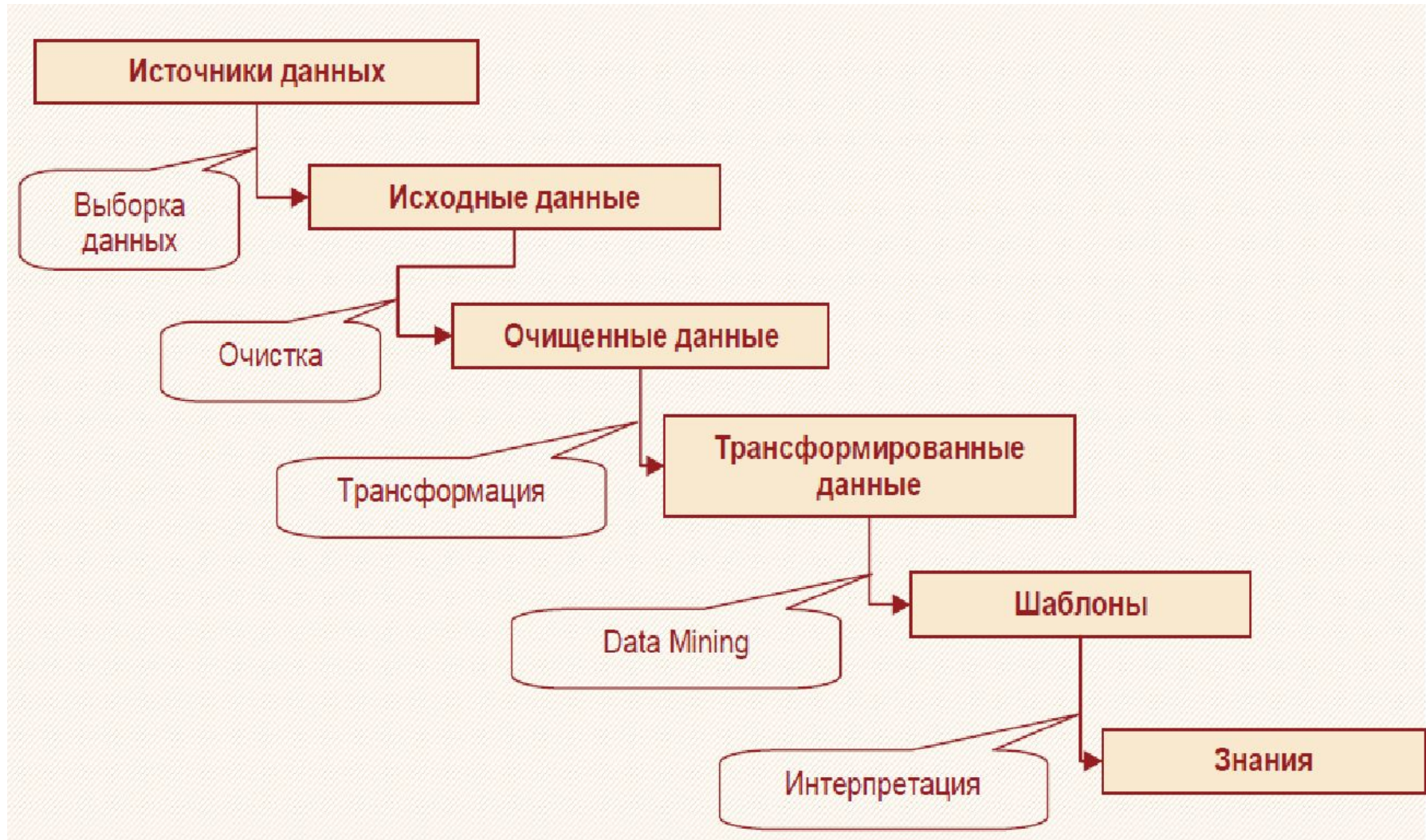
Методика извлечения знаний

Knowledge Discovery in Databases («извлечение знаний из баз данных») - последовательность действий, которую необходимо выполнить для обнаружения полезного знания.

Knowledge Discovery in Databases — процесс получения из данных знаний в виде зависимостей, правил, моделей, обычно состоящий из таких этапов, как выборка данных, их очистка и трансформация, моделирование и интерпретация полученных результатов.

Технологии KDD и DM

Этапы KDD



Технологии KDD и DM

Этапы KDD

- запросы
- фильтрация данных
- сэмплинг

- заполнение пропусков
- подавление аномальных значений
- сглаживание
- исключение дубликатов и противоречий

- преобразование временного ряда
- приведение типов
- выделение временных интервалов
- квантование
- сортировка
- группировка

- моделирование

- формальные методы
- знания аналитика

Технологии KDD и DM

Data Mining

(«добыча данных», «раскопка данных») – совокупность большого числа различных методов обнаружения знаний.

Data Mining — обнаружение в «сырых» данных ранее неизвестных, нетривиальных, практически полезных и доступных интерпретации знаний, необходимых для принятия решений в различных сферах человеческой деятельности.

Технологии KDD и DM

Условная классификация задач DM

- установление зависимости дискретной выходной переменной от входных переменных

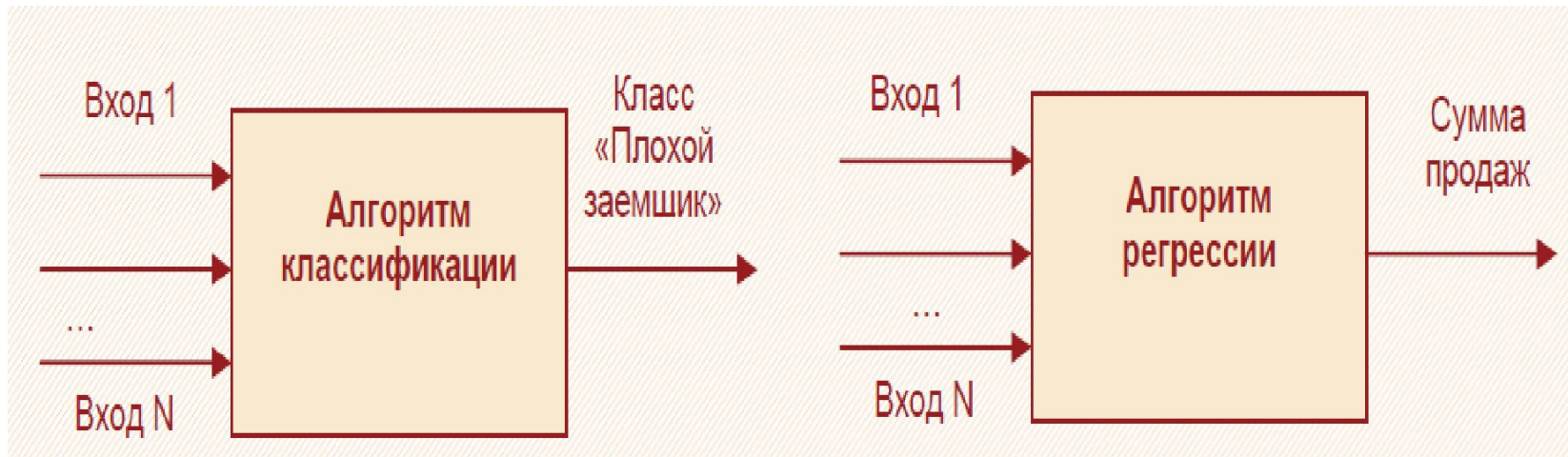
- установление зависимости непрерывной выходной переменной от входных переменных

- группировка объектов (наблюдений, событий) на основе данных, описывающих свойства объектов

- выявление закономерностей между связанными событиями

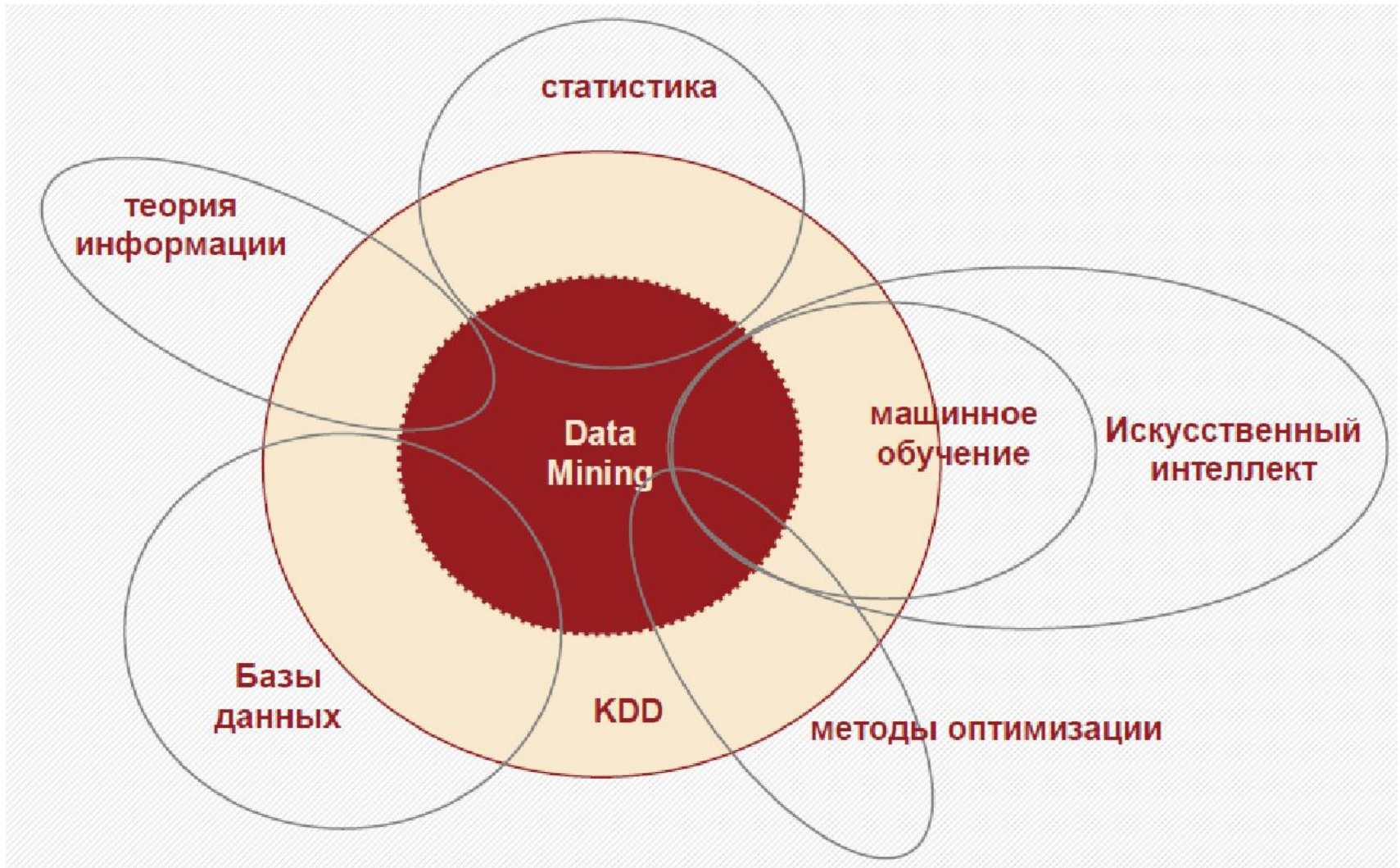
Технологии KDD и DM

Алгоритмы DM



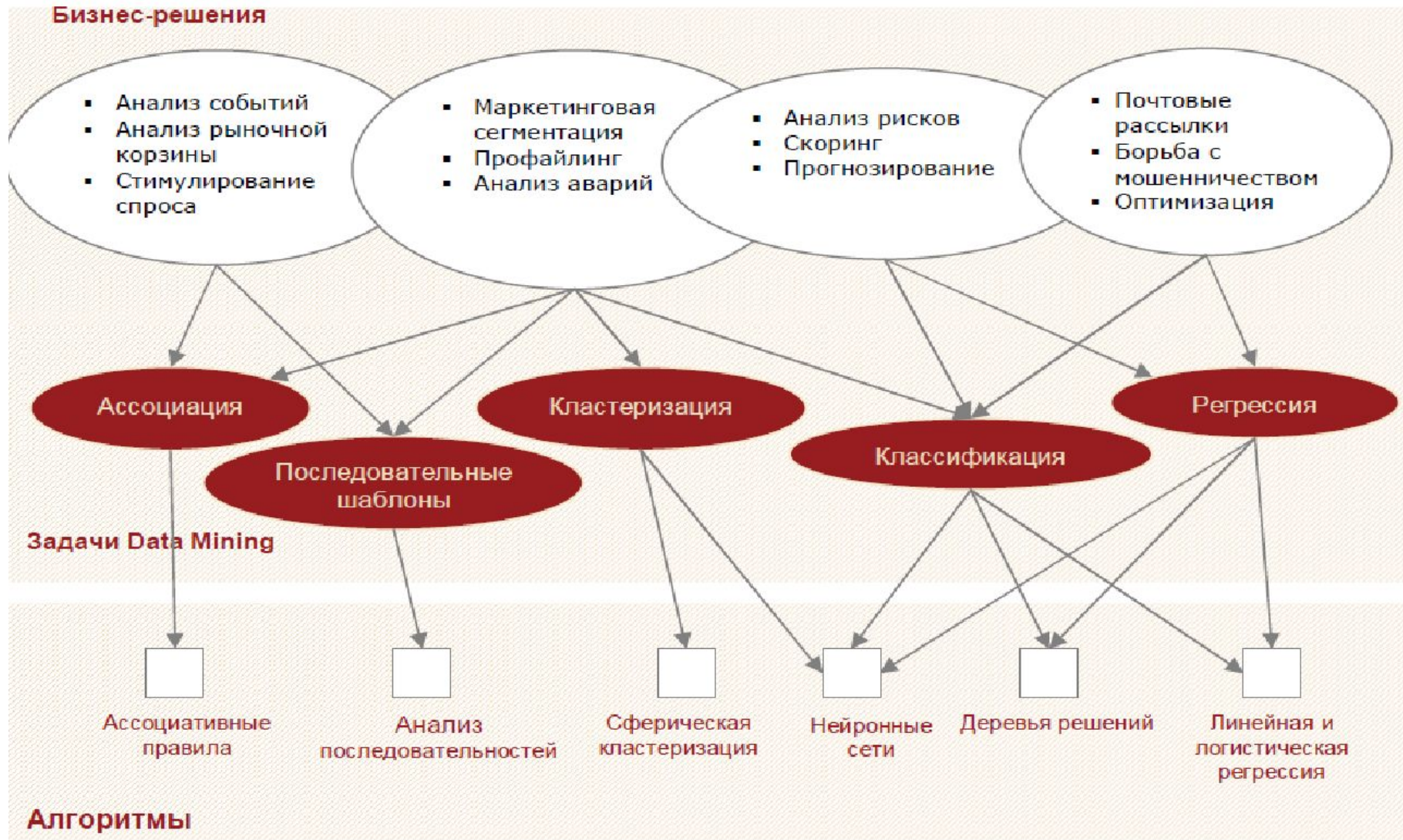
Технологии KDD и DM

Алгоритмы и методы DM



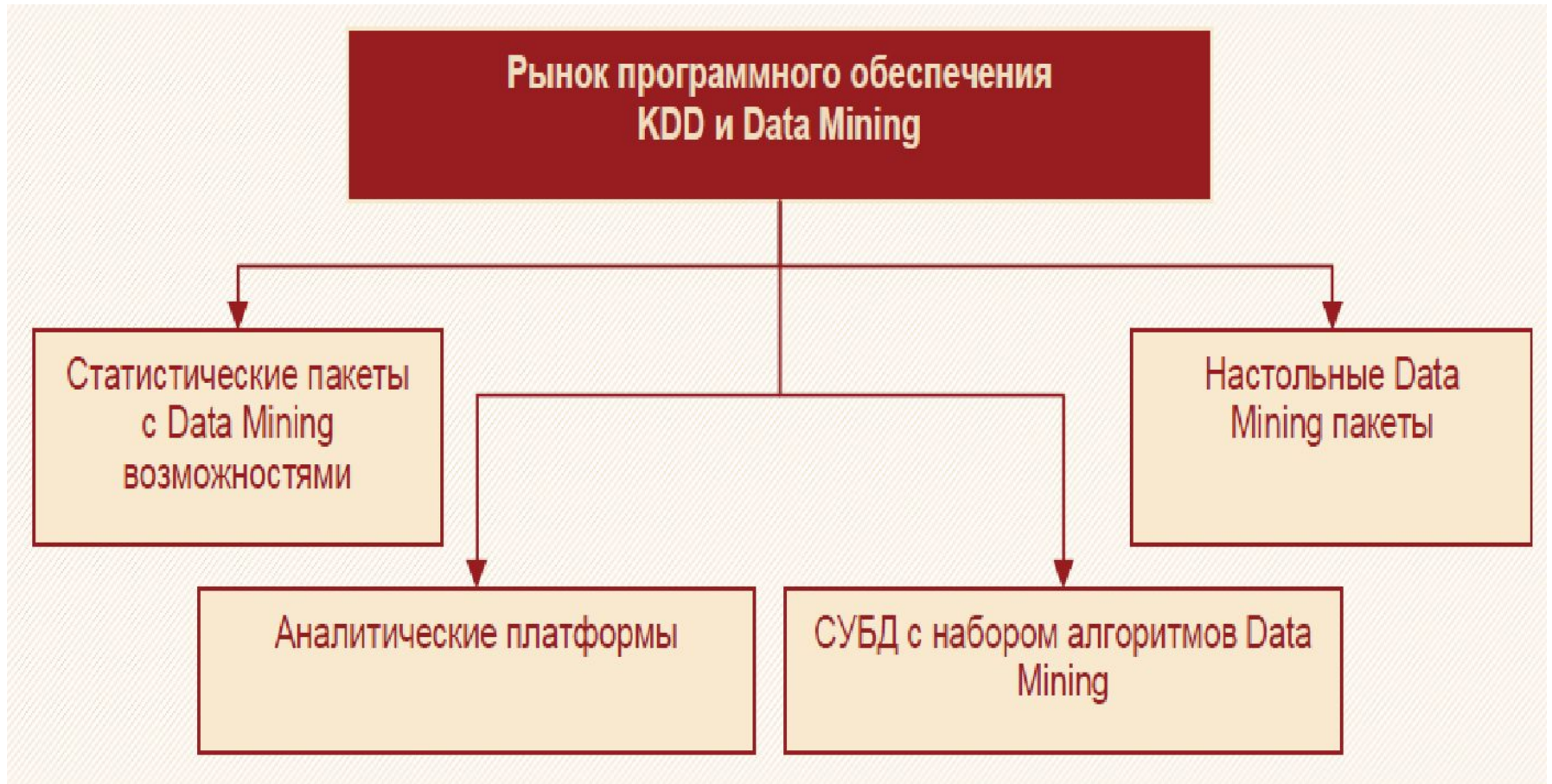
Технологии KDD и DM

Алгоритмы DM для бизнес-решений



Технологии KDD и DM

Классификация ПО в области Data Mining и KDD



Технологии KDD и DM

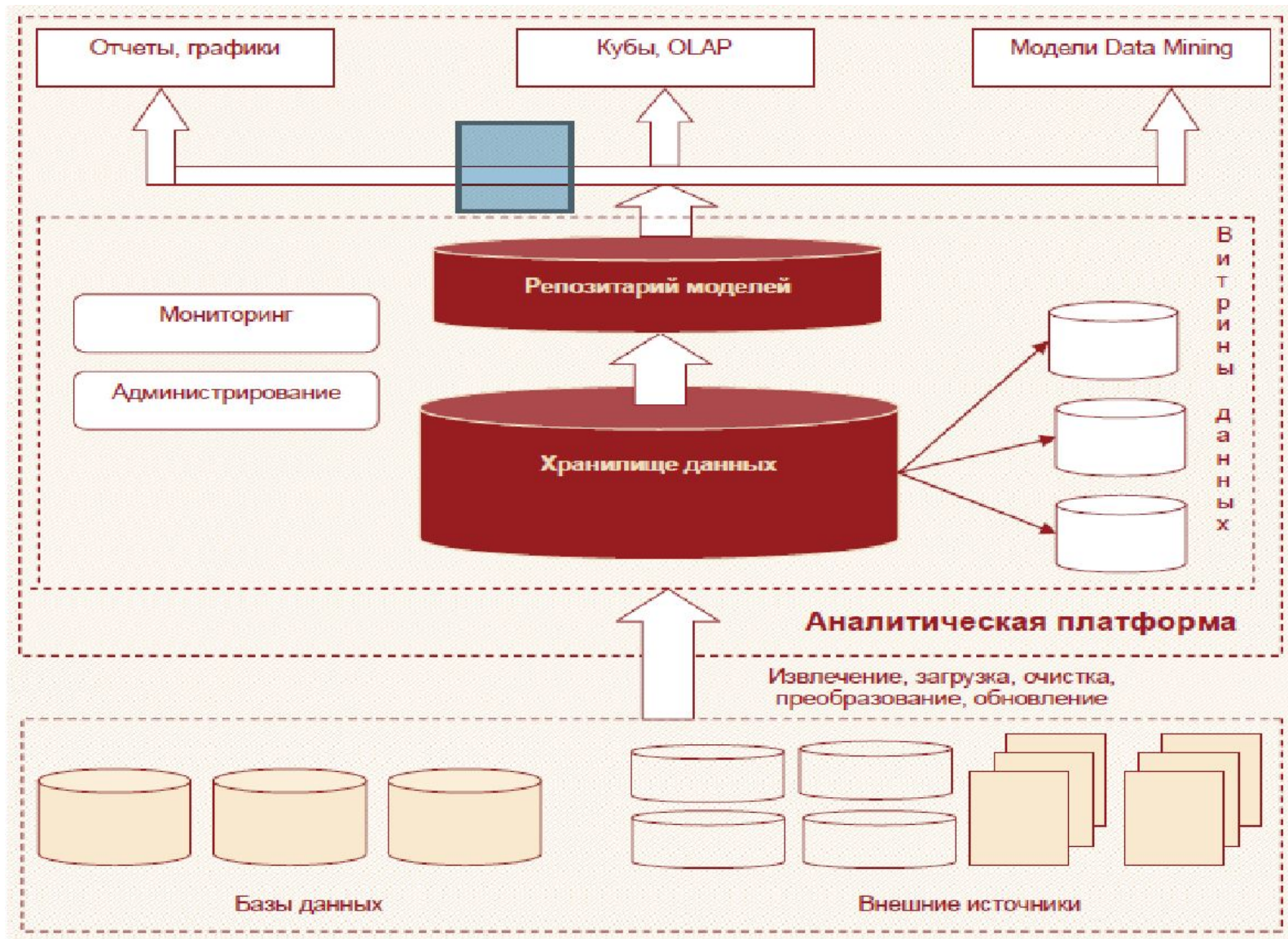
Аналитические платформы –

ориентированы на анализ данных и предназначены для создания готовых решений

Аналитическая платформа — специализированное программное решение (или набор решений), которое содержит в себе все инструменты для извлечения закономерностей из «сырых» данных: средства консолидации информации в едином источнике (хранилище данных), извлечения, преобразования, трансформации данных, алгоритмы Data Mining, средства визуализации и распространения результатов среди пользователей, а также возможности «конвейерной» обработки новых данных.

Технологии KDD и DM

Архитектура аналитической платформы



Технологии KDD и DM

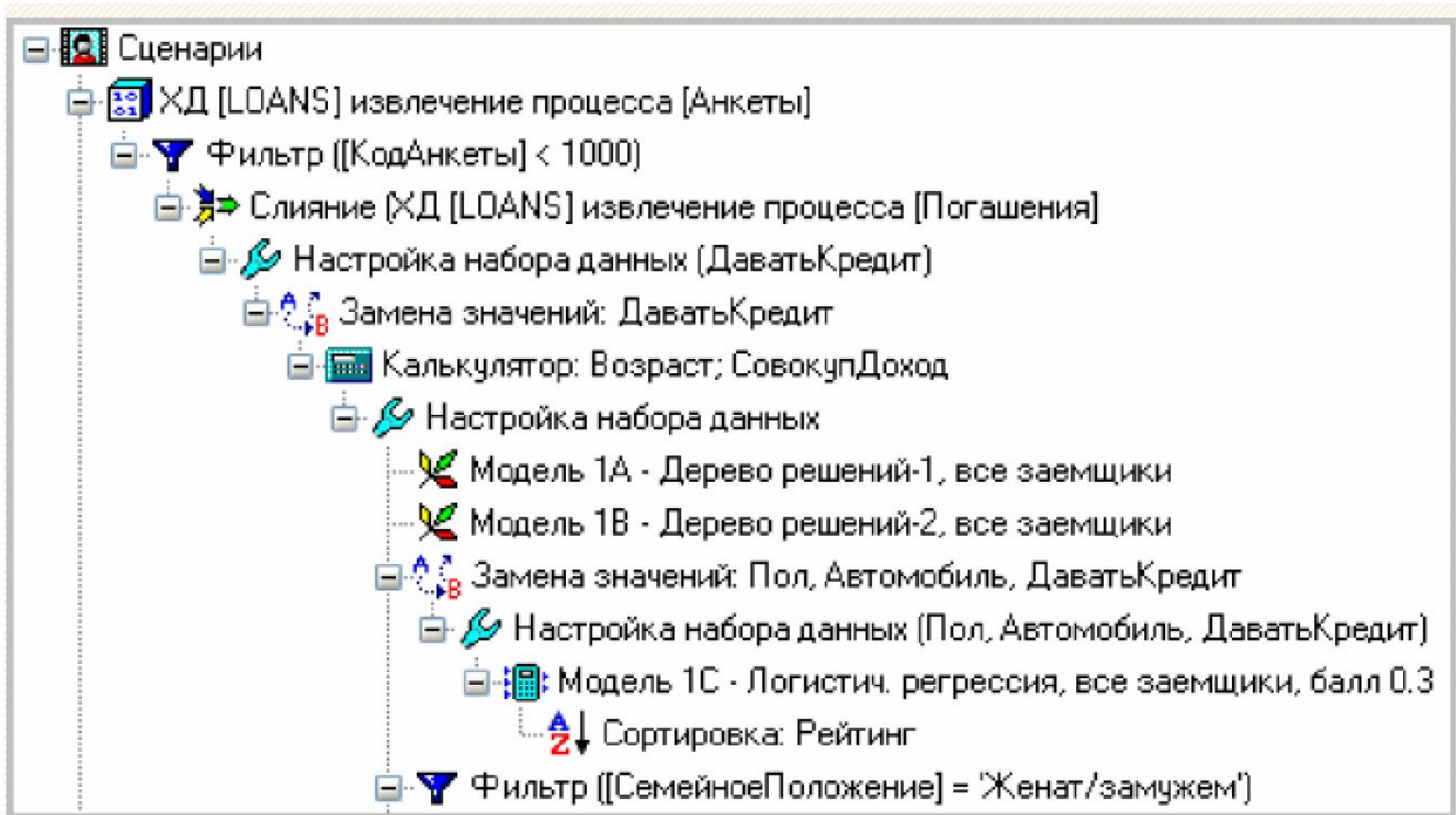
Языки визуального моделирования -

- освобождают аналитика от необходимости углубленного понимания сложных математических алгоритмов;
- позволяют аналитику в визуальной среде строить последовательности шагов по обработке данных от получения «сырых» данных до конечного результата.

Технологии KDD и DM

Языки визуального моделирования

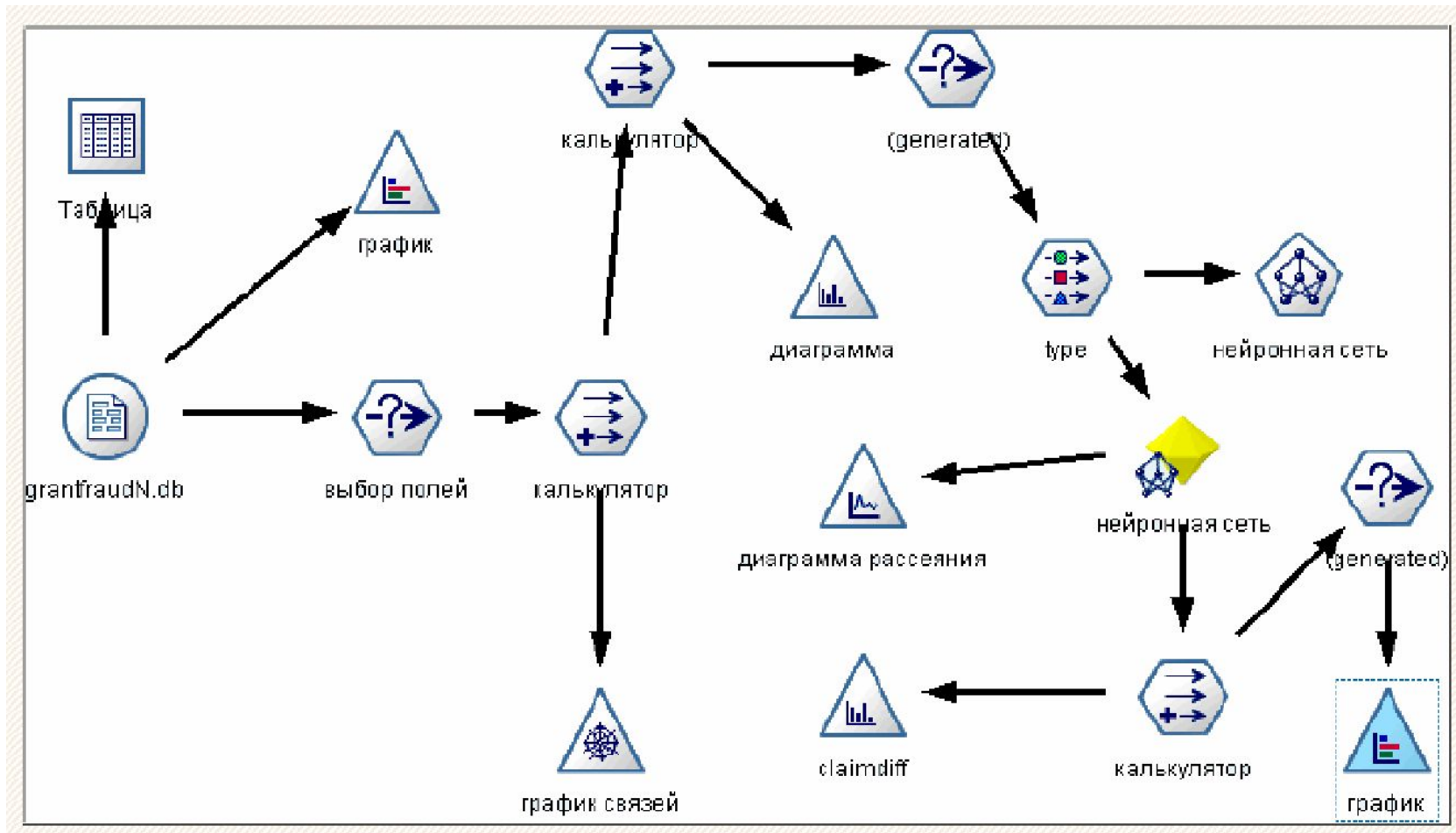
Диаграмма в виде дерева



Технологии KDD и DM

Языки визуального моделирования

Диаграмма в виде графа



Вопросы - ?

ТЕСТ – 15 вопросов

15 минут