

DATA MINING

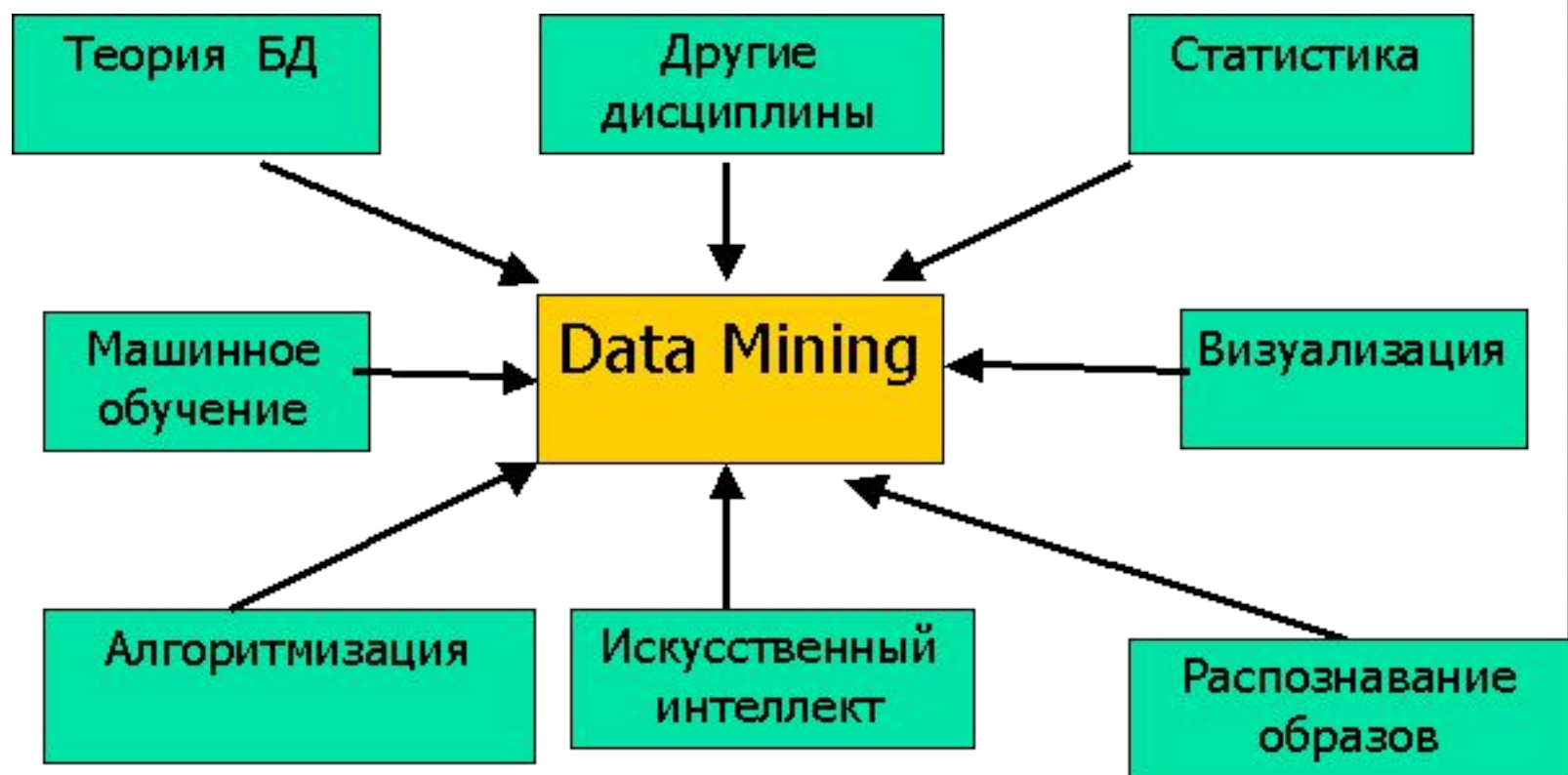
Выполнили студентки группы У04-01
Фоминых Татьяна Игоревна
Сурнина Елизавета Алексеевна

ПОНЯТИЕ DATA MINING

- Получило распространение в 1989 году
- Основоположник - Григорий Пиатецкий - Шапиро
- Data Mining - это технология для поиска в больших объемах данных неочевидных, объективных и полезных закономерностей.



DATA MINING КАК МУЛЬТИДИСЦИПЛИНАРНАЯ ОБЛАСТЬ



СРАВНЕНИЕ СТАТИСТИКИ, МАШИННОГО ОБУЧЕНИЯ И DATA MINING

Data Mining

- ⦿ Интеграция теории и эвристик.
- ⦿ Сконцентрирована на едином процессе анализа данных, включает очистку данных, обучение, интеграцию и визуализацию результатов.

Статистика

- ⦿ Более, чем Data Mining, базируется на теории.
- ⦿ Более сосредотачивается на проверке гипотез.

Машинное обучение

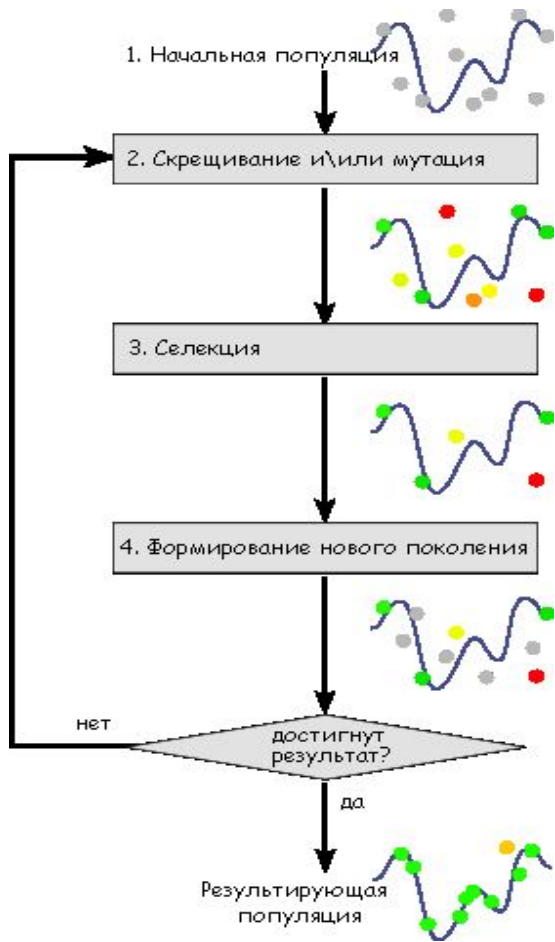
- ⦿ Более эвристично
- ⦿ Концентрируется на улучшении работы агентов обучения.

МЕТОДЫ DATA MINING

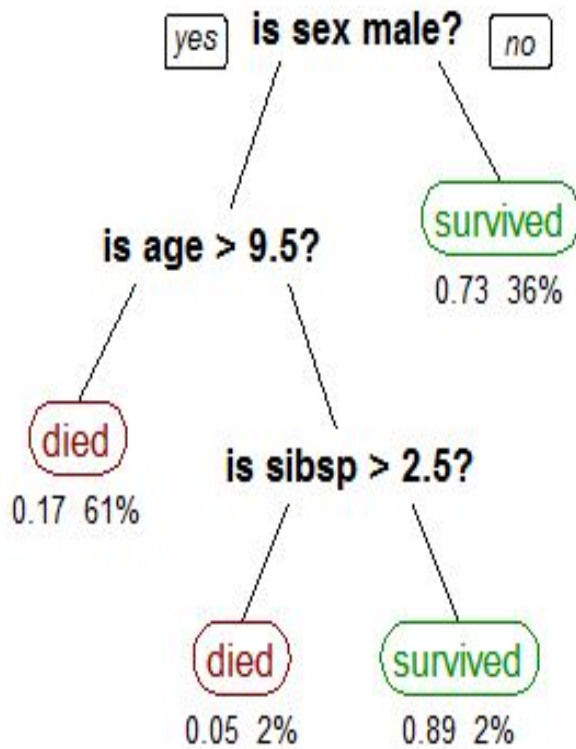
- **Методы**
 - Статистика
 - Современные
 - ИТ
 - Алгоритмизация
 - Прогнозирование
- **Классификация**
- **Моделирование**

МЕТОДЫ DATA MINING

Генетический алгоритм



Дерево принятия решений



СТАДИИ DATA MINING

- Выявление закономерностей (свободный поиск)
- Использование выявленных закономерностей для предсказания неизвестных значений (прогностическое моделирование)
- Анализ исключений

ЗАДАЧИ DATA MINING

- Классификация
- Кластеризация
- Ассоциация
- Последовательная ассоциация
- Прогнозирование
- Определение отклонений
- Оценивание
- Визуализация
- Подведение итогов

СФЕРЫ ПРИМЕНЕНИЯ DATA MINING

- Решение бизнес-задач
- Государственные задачи
- Научные исследования
- WEB - задачи



WEB MINING



TEXT MINING

- Text Mining охватывает новые методы для выполнения семантического анализа текстов
- Программы должны некоторым образом оперировать естественным человеческим языком и при этом понимать семантику анализируемого текста.



Call Mining

Распознавание
речи

Анализ речи

Data Mining



**СПАСИБО ЗА
ВНИМАНИЕ!**