

# Методы интеллектуального анализа данных

*Мартин Браун (Martin C. Brown) Опубликовано 01.02.2013*

Кабдуллин Азат  
Кабдуллин Максат

Интеллектуальный анализ данных — это?

# Введение

Бизнес-требования привели от простого поиска и статистического анализа данных к более сложному интеллектуальному анализу данных. Для решения бизнес-задач требуется такой анализ данных, который позволяет построить модель для описания информации и в конечном итоге приводит к созданию результирующего отчета.



# Ассоциация

## Association rule

Transaction	Кола	Пиво	Подгузники
0	1	1	1
1	2	0	0
2	3	1	1
3	4	1	1
4	5	0	1

$$supp = \frac{\text{Транзакции с пивом и подгузниками}}{\text{Все транзакции}} = P(\text{Пиво} \cap \text{Подгузники})$$

$$supp = \frac{3}{5} = 60\%$$

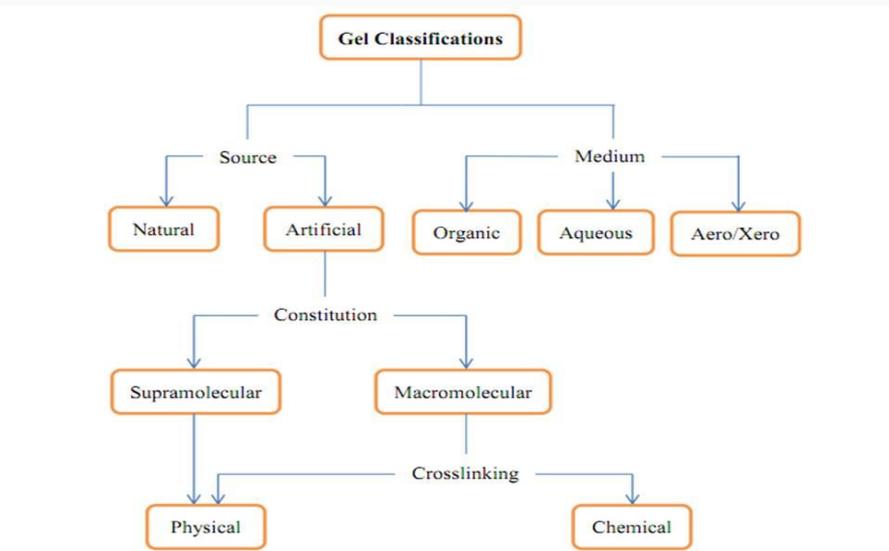
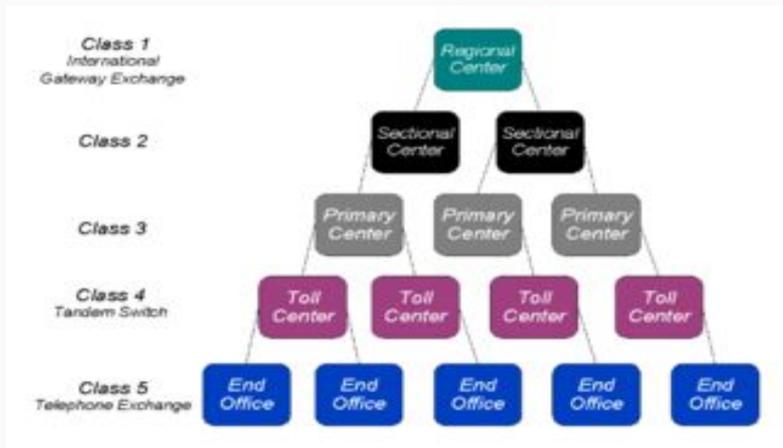
# Деревья решений

Дерево решений начинают с простого вопроса, который имеет два ответа (иногда больше). Каждый ответ приводит к следующему вопросу, помогая классифицировать и идентифицировать данные или делать прогнозы.



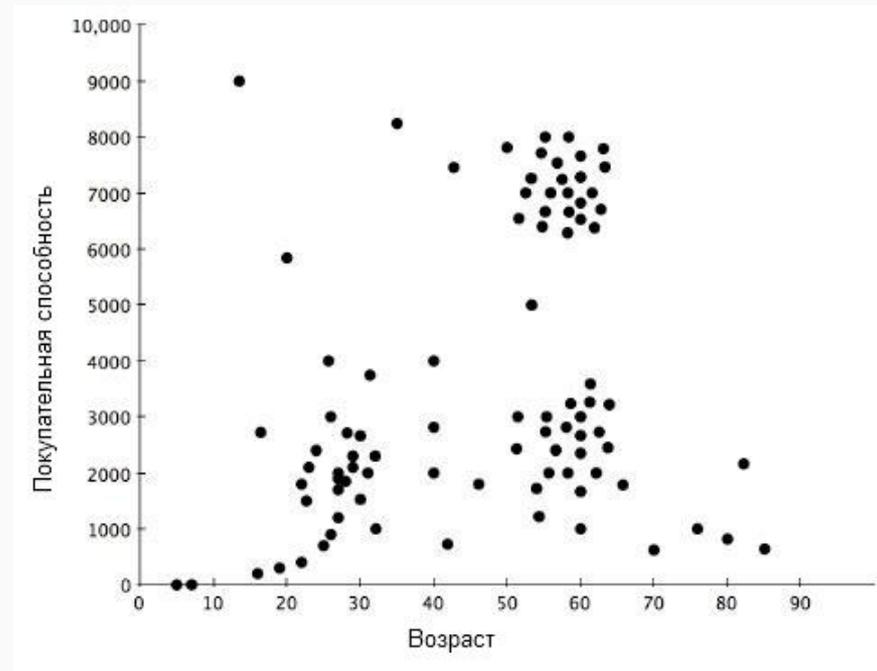
# Классификация

Классификацию можно использовать для получения представления о типе покупателей, товаров или объектов, описывая несколько атрибутов для идентификации определенного класса.



# Кластеризация

На простом уровне при кластеризации используется один или несколько атрибутов в качестве основы для определения кластера сходных результатов. Кластеризация полезна при определении различной информации, потому что она коррелируется с другими примерами, так что можно увидеть, где подобию и диапазоны согласуются между собой.

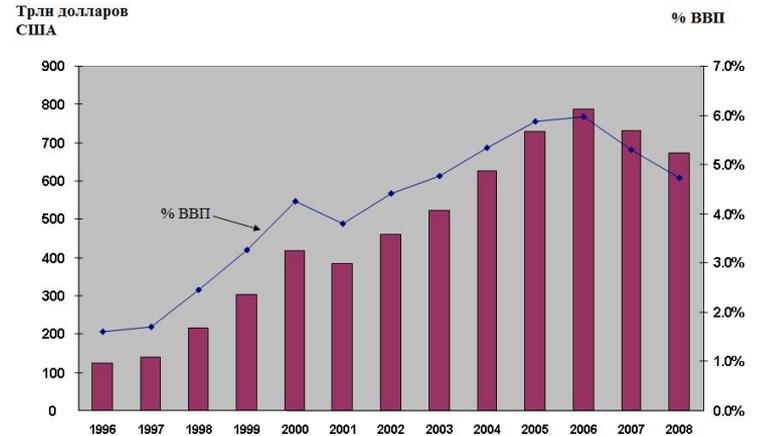


# Прогнозирование

В сочетании с другими методами интеллектуального анализа данных прогнозирование предполагает анализ тенденций, классификацию, сопоставление с моделью и отношения.

Анализируя прошлые события или экземпляры, можно предсказывать будущее.

Дефицит текущего платёжного баланса США: доллары и проценты от ВВП



Источник данных: Бюро экономического анализа США

# Комбинации

---

на практике очень редко используется только один из этих методов. Классификация и кластеризация — подобные методы. Используя кластеризацию для определения ближайших соседей, можно дополнительно уточнить классификацию. Деревья решений часто используются для построения и выявления классификаций, которые можно проследивать на исторических периодах для определения последовательностей и моделей.



# Спасибо!

Контактная информация:

Название компании

ул. Улица, д. 1

Ваш город, Страна, 123456

no\_reply@example.com

[www.example.com](http://www.example.com)

