

Федеральное агентство по образованию
СЫКТЫВКАРСКИЙ ЛЕСНОЙ ИНСТИТУТ – филиал государственного образовательного
учреждения высшего профессионального образования
«Санкт-Петербургской государственной лесотехнической академии имени С. М. Кирова»

Доклад по дисциплине: Информационные системы в экономике

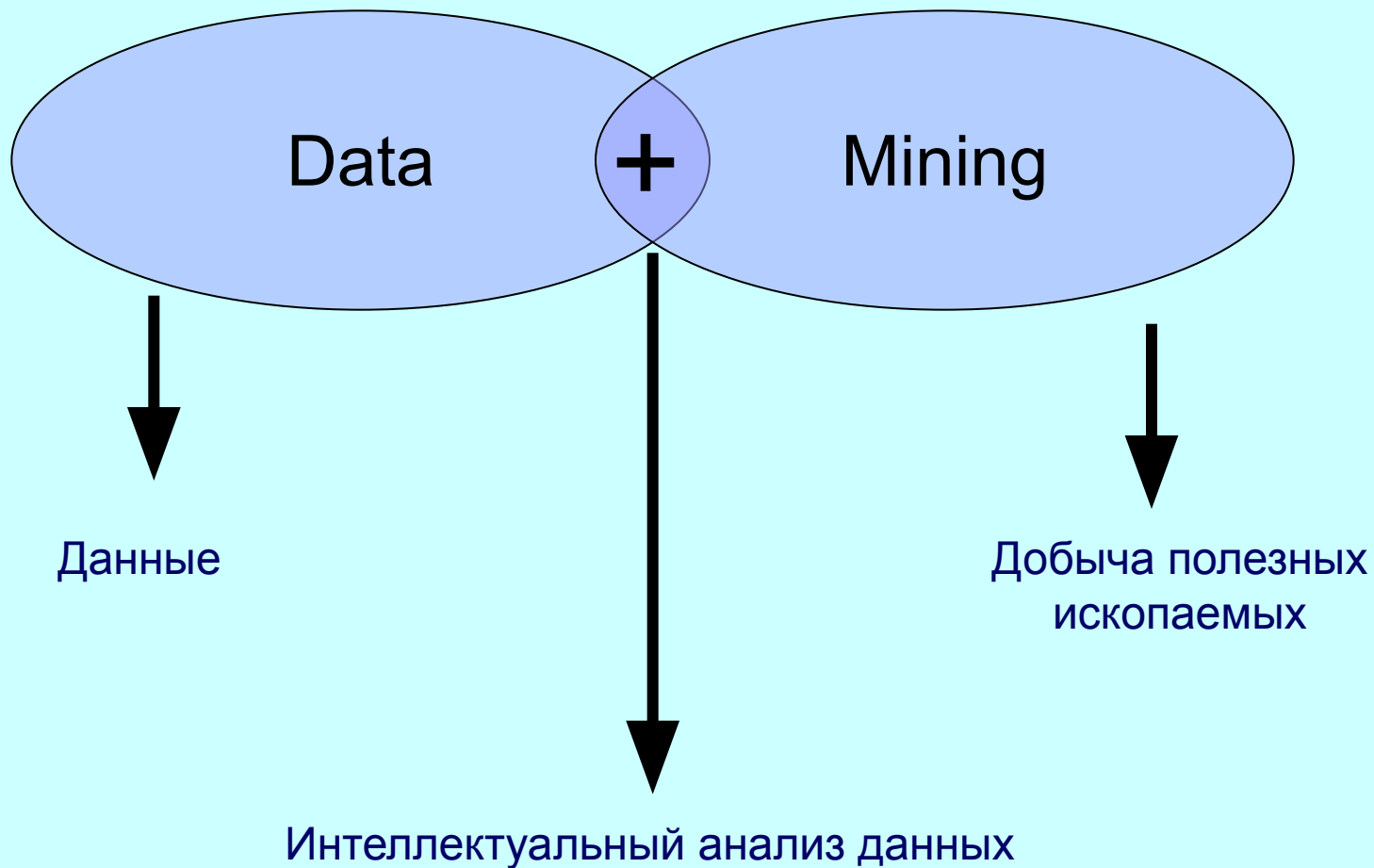
На тему:

Data Mining



Выполнила: студентка 3 курса ФЭиУ,
спец. БУАиА, гр. № 237,
Филиппова Ю.В.
Преподаватель: Ключева Е.А.

Сыктывкар 2010



Data Mining — это процесс обнаружения в сырых данных ранее неизвестных, нетривиальных, практически полезных и доступных интерпретации знаний, необходимых для принятия решений в различных сферах человеческой деятельности



Уровни знаний извлекаемые из данных:



Генераторы запросов;
инфо - поисковая
система OLTP

Аналитическая
обработка
информации OLAP;
DSS-система
поддержки решений

Интеллектуальный
анализ данных Data
Mining

1978 г.

Появление понятия Data Mining.

1990-е г.г.

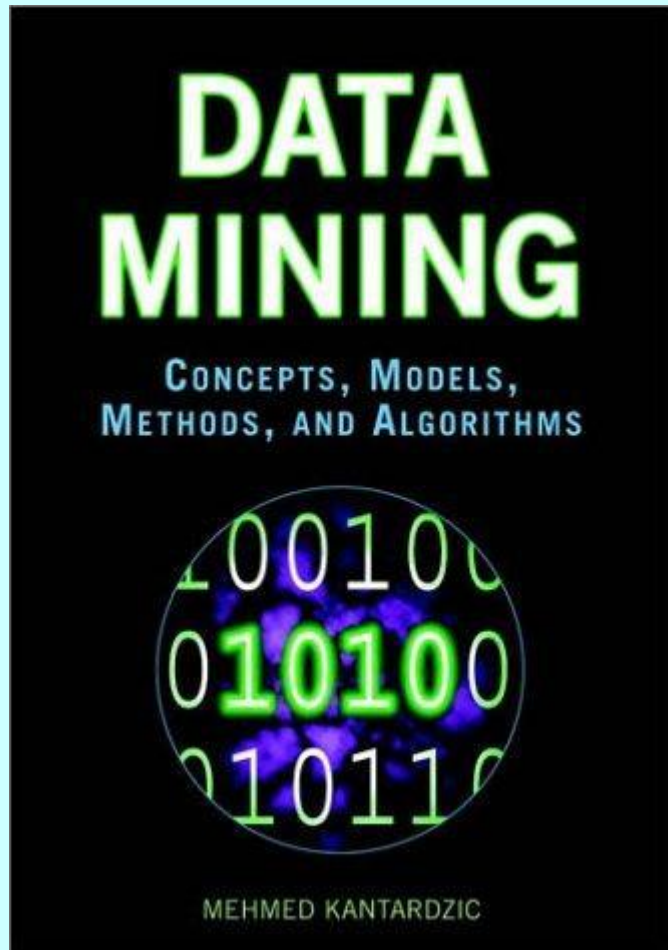
Понятие Data Mining, приобрело высокую популярность в современной трактовке.



Возникновение и развитие Data Mining обусловлено следующими факторами:

- совершенствование аппаратного и программного обеспечения;
- совершенствование технологий хранения и записи данных;
- накопление большого количества данных;
- совершенствование алгоритмов обработки информации.

Задачи (закономерности, техники) Data Mining:



- ассоциация;
- последовательность;
- классификация;
- кластеризация;
- временные закономерности.

Методы Data Mining:

Статистические

- дескриптивный анализ,
- корреляционный и регрессионный анализ,
- факторный анализ,
- дисперсионный анализ,
- компонентный анализ,
- дискриминантный анализ,
- анализ временных рядов

Кибернетические

- искусственные нейронные сети,
- эволюционное программирование,
- генетические алгоритмы,
- ассоциативная память,
- деревья решений,
- системы обработки экспертных знаний



Недостатки технологии Data Mining:

- Data Mining не может заменить аналитика;
- Сложность разработки и эксплуатации приложения Data Mining;
- Требуется определенной квалификации пользователя;
- Сложность подготовки данных;
- Большой процент ложных, недостоверных или бессмысленных результатов;
- Высокая стоимость.



Особенности областей, где применение технологии Data Mining будет успешным:

- требуют решений, основанных на знаниях;
- имеют изменяющуюся окружающую среду;
- имеют доступные, достаточные и значимые данные;
- обеспечивают высокие дивиденды от правильных решений.

Примеры применения Data Mining

Финансы

- Прогноз оттока клиентов
- Выявление мошенничества
- Потенциальные банкротства
- Возможные продажи

Маркетинг

- Предложение купить конкретный продукт
- Более нацеленные кампании
- Выявление возможностей для кросс-продаж

Телеком

- Выявление оттока клиентов
- Выявление возможностей для кросс-продаж

Страхование

- Выявление аномалий
- Уменьшение стоимости расследований подозрительных случаев и фальшивых заявок

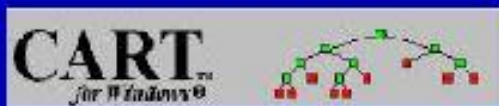
Розница

- Программы лояльностей
- Кросс-продажи
- Анализ корзины
- Мошенничества

Медицина

- Выявление факторов, влияющих на здоровье пациентов
- Генетический анализ
- Новые лекарства

Популярные продукты для Data Mining





Спасибо за внимание!