

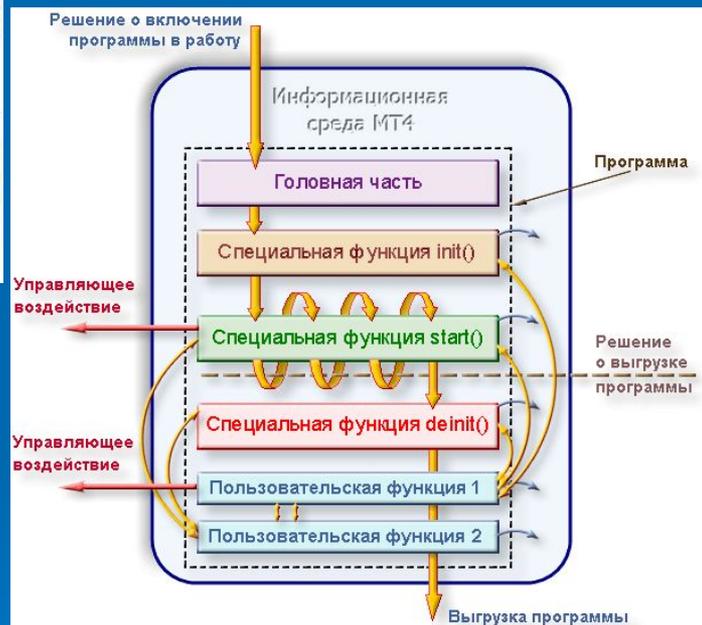
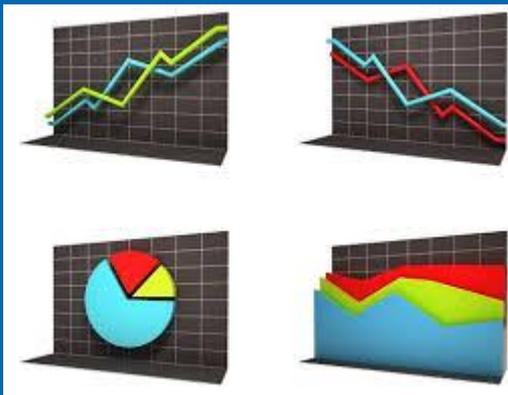
Лекция 11

Визуализация данных



Визуализация данных

К способам визуального или графического представления данных относят графики, диаграммы, таблицы, отчеты, списки, структурные схемы, карты и



ОЧЕНЬ ОПАСНЫЕ	E123	E510	E519E	E527				
ОПАСНЫЕ	E102	E110	E120	E124	E127	E129	E155	E180
	E201	E220	E222	E223	E224	E228	E233	E242
	E400	E401	E402	E403	E404	E405	E501	E502
	E503	E620	E636	E637				
КАНЦЕРОГЕННЫЕ	E131	E142	E153	E210	E212	E213	E214	E215
	E216	E219	E230	E240	E249	E280	E281	E282
	E283	E310	E354					
РАССТРОЙСТВО ЖЕЛУДКА	E338	E339	E340	E341	E343	E450	E461	E462
	E463	E465	E466					
КОЖНЫЕ ЗАБОЛЕВАНИЯ	E151	E160	E231	E232	E239	E311	E312	E320
	E907	E951	E1105					
РАССТРОЙСТВО КИШЕЧНИКА	E154	E626	E627	E628	E629	E630	E631	E632
	E633	E634	E635					
ДАВЛЕНИЕ	E154	E250	E252					
ОПАСНЫЕ ДЛЯ ДЕТЕЙ	E270							
ЗАПРЕЩЕННЫЕ	E103	E105	E111	E121	E123	E125	E126	E130
	E152	E211	E952					
ПОДОЗРИТЕЛЬНЫЕ	E104	E122	E161	E171	E173	E241	E477	

Визуализация традиционно рассматривалась как вспомогательное средство при анализе данных, однако сейчас все больше исследований говорит о ее самостоятельной роли.

Визуализация данных

Традиционные методы визуализации могут находить следующее применение:

- представлять пользователю информацию в наглядном виде;
- компактно описывать закономерности, присущие исходному набору данных;
- снижать размерность или сжимать информацию;
- восстанавливать пробелы в наборе данных;
- находить шумы и выбросы в наборе данных.

Методы визуализации

Методы визуализации, в зависимости от количества используемых измерений, принято классифицировать на две группы:

- представление данных в одном, двух и трех измерениях;
- представление данных в четырех и более измерениях.

Представление данных в одном, двух и трех измерениях

К этой группе методов относятся хорошо известные способы отображения информации, которые доступны для восприятия человеческим воображением. Практически любой современный инструмент Data Mining включает способы визуального представления из этой группы.

Методы визуализации

В соответствии с количеством измерений представления это могут быть следующие способы:
одномерное (univariate) измерение, или 1-D;
двумерное (bivariate) измерение, или 2-D;
трехмерное или проекционное (projection) измерение, или 3-D.

Следует заметить, что наиболее естественно человеческий глаз воспринимает двухмерные представления информации.



Методы визуализации

При использовании двух- и трехмерного представления информации пользователь имеет возможность увидеть закономерности набора данных:

- его кластерную структуру и распределение объектов на классы (например, на диаграмме рассеивания);
- топологические особенности;
- наличие трендов;
- информацию о взаимном расположении данных;
- существование других зависимостей, присущих исследуемому набору данных.

Методы визуализации

Если набор данных имеет более трех измерений, то возможны такие варианты:

- использование многомерных методов представления информации (они рассмотрены ниже);
- снижение размерности до одно-, двух- или трехмерного представления. Существуют различные способы снижения размерности, один из них - факторный анализ.

Методы визуализации

Представления информации в четырехмерном и более измерениях недоступны для человеческого восприятия. Однако разработаны специальные методы для возможности отображения и восприятия человеком такой информации.

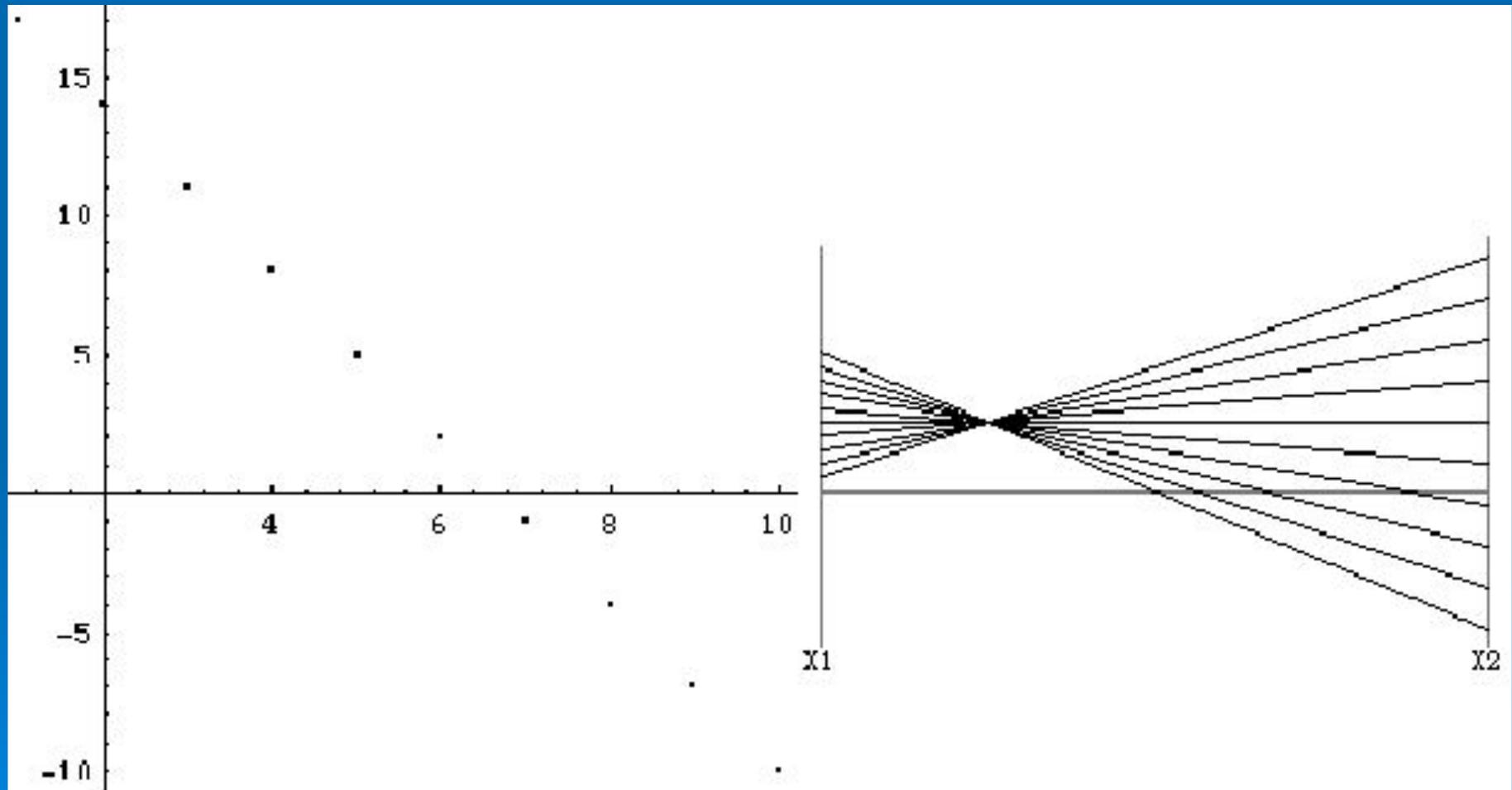
Наиболее известные способы многомерного представления информации:

- параллельные координаты;
- "лица Чернова";
- лепестковые диаграммы.



Методы визуализации

- параллельные координаты; В параллельных координатах переменные кодируются по горизонтали, вертикальная линия определяет значение переменной.



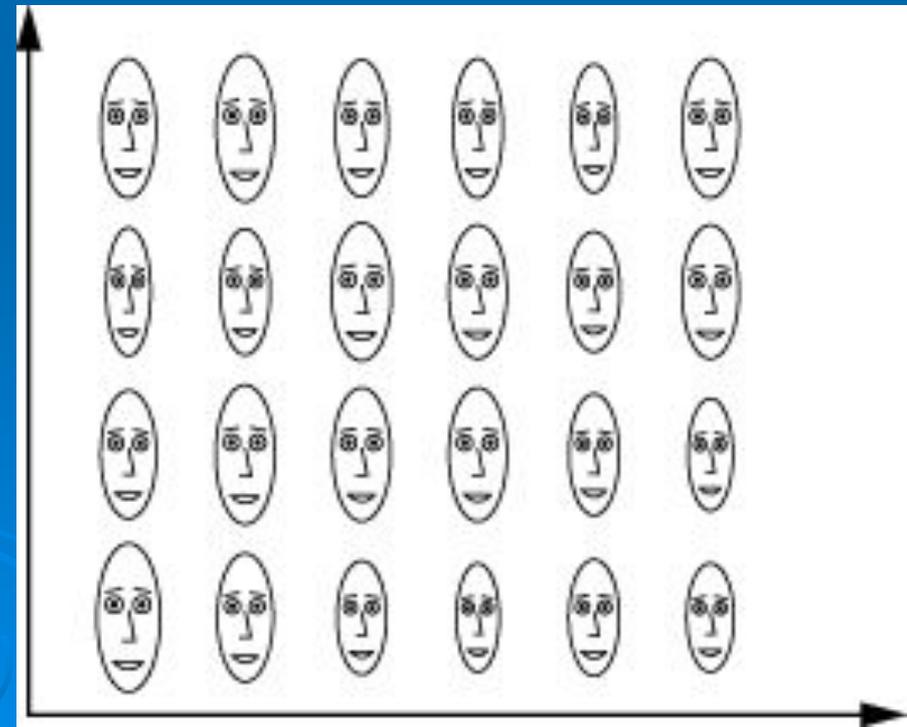
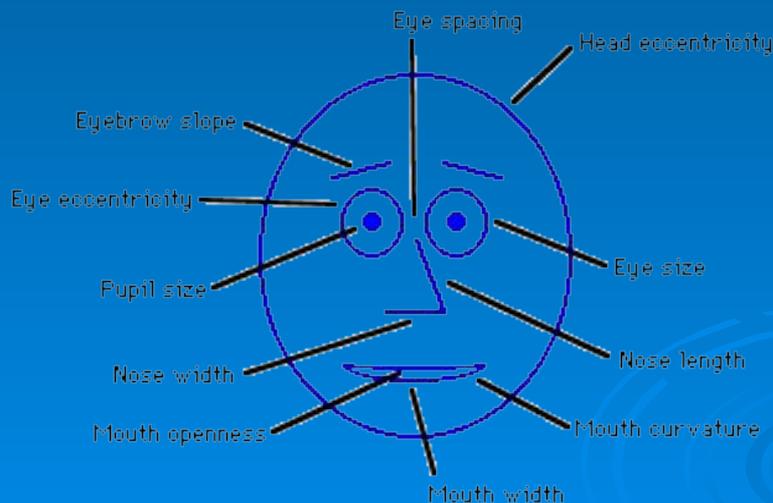
Методы визуализации



"лица Чернова"; Основная идея представления информации в "лицах Чернова" состоит в кодировании значений различных переменных в характеристиках или чертах человеческого лица

Для каждого наблюдения рисуется отдельное "лицо". На каждом "лице" относительные значения переменных представлены как формы и размеры отдельных черт лица (например, длина и ширина носа, размер глаз, размер зрачка, угол между бровями).

Анализ информации при помощи такого способа отображения основан на способности человека интуитивно находить сходства и различия в чертах лица.



Методы визуализации

- лепестковые диаграммы.

