

Visual Data Mining

Chidroop Madhavarapu
CSE 591:Visual Analytics

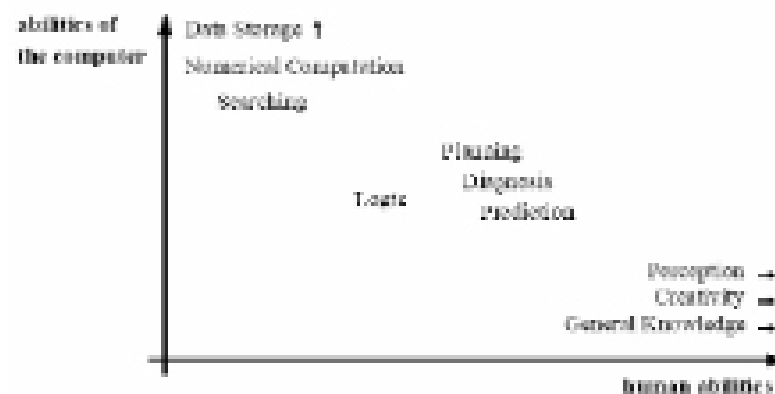
Motivation

Visualization for Data Mining

- Huge amounts of information
- Limited display capacity of output devices

Visual Data Mining (VDM) is a new approach for exploring very large data sets, combining traditional mining methods and information visualization techniques.

Why Visual Data Mining



Integration of visualization and data mining :

Visual Data Mining approaches fall under 3 categories:

- Data Mining process visualization.
- Data Mining result visualization.
- Interactive Visual Data Mining.

Data Mining process visualization

Visualization techniques are used to support Data Mining.

Ex: When required to handle large amount of multidimensional data in the format of Data Tables or relational databases. (Parallel Coordinates, scatter plots etc.)

Data Mining result visualization

To visually convey the results of Mining tasks, such as clustering or classification, to enhance user interpretation.

Examples include Scatter plots, Box plots, BLOB and H-BLOB clustering algorithms, Decision trees, Association Rules,

Interactive Visual Data Mining

- Rather than using Visual Data exploration and analytical mining algorithms as separate tools, a stronger DM strategy would be to tightly couple the visualizations and analytical processes into one DM tool.
- Using visualization tools in the data mining process to help users make smart data mining decisions.
- Examples include the Control project, OptiGrid, PBC (Perception Based Classification).

V-Miner: Using Enhanced Parallel Coordinates to Mine Product Design and Test Data

INTRODUCTION

- V-Miner : Multivariable visualization tool.
- Designed for Mining product design and test data.
- New technique based on Parallel coordinate visualization.
- Goal is to discover useful knowledge from mobile phone testing data that can be used to provide feedback to the design engineers.

Design Process for consumer electronics.

- Engineers design specific sections of phone based on previous successful designs, new product specs, design simulations etc.
- Prototypes are built
- Functional tests are performed on prototypes. If the requirements are not met, start with next design cycle (from step 1).

Above steps are repeated until design meets the specification. Then the phone is released to the NPI team for volume manufacturing.

- For a new product, number of iterations of design revisions should be coordinated.
- 100's of variables involved which are changed/tested in the different revisions.
- V-Miner is used to reduce engineering costs, design defects by mining useful knowledge from the test data .

THE DATA

After each design change, all test variables are measured. Each variable takes numerical values and has the following properties:

- Has an *upper limit* and *lower limit*. If a value does not fall in this range, its unacceptable.
- Has an ideal value called the *target value*.