

Data Mining Techniques

Chapter 1: Why and What is Data Mining?

Analytic customer relationship management	2
Information systems	3
Data warehousing	4
Customer relationship management	5
Data mining I	6
Data mining II.	7
Data mining III	8

Analytic customer relationship management

- IS → DW → DM → CRM (headings on board)
- IS (information systems) notice customer behavior
- DW (data warehousing) remember behavior over time
- DM (data mining) learn from behavior (focus of this course)
- CRM (customer relationship management) act on learning

© Iain Pardoe, 2006

2 / 8

Information systems

- Operational, but can sometimes be used for DM, e.g.:
 - phone companies' call records to find residential numbers used like businesses;
 - catalog companies' order histories to identify customers for future mailings;
 - Fedex change in shipping patterns during UPS strike;
 - supermarket POS data to decide which coupons to print;
 - web retailers' past purchases to determine what to display on return visits.

© Iain Pardoe, 2006

3 / 8

Data warehousing

- Gather operational data together and organize it in a consistent and useful way over time.
- DSC 444/544.

© Iain Pardoe, 2006

4 / 8

Customer relationship management

- Understand each customer individually.
- Use that understanding to make it easier for the customer to do business with you rather than competitors.
- Transform from a product-focused organization into a customer-centric one.
- Organization must be able to change its behavior as a result of what it learns through DM.
- Need to know both how DM tools work and also how they will be used.

© Iain Pardoe, 2006

5 / 8

Data mining I

- Analytical techniques that can be used to turn customer data into customer knowledge:
 - over time learn more about customers, use that knowledge to serve them better.
- Exploration and analysis of large datasets to discover meaningful patterns and decision rules:
 - knowledge discovery in databases (KDD);
 - knowledge creation.
- Business decisions should be based on learning:
 - informed decisions are better than uninformed;
 - measuring results is beneficial.
- Add intelligence to the data warehouse.
- Data from the past contains information that will be useful in the future (if customer/business behavior is not completely random):
 - need to find systematic signal and distinguish it from random noise.

© Iain Pardoe, 2006

6 / 8

Data mining II

- Directed DM:
 - explain or categorize a *target* field/variable such as income or response (to an offer, say);
 - build models (algorithms/rules/formulas) to connect inputs to target or outcome;
 - e.g., regression, neural networks, decision trees, nearest neighbors;
 - models produce *scores* (fitted or predicted values) used to *rank* customers.
- Undirected DM:
 - find patterns/similarities among group of records/observations;
 - e.g., affinity grouping (association rules, market basket analysis), clustering, self-organizing maps.
- Description/profiling:
 - e.g., decision trees, association rules, clustering.

© Iain Pardoe, 2006

7 / 8

Data mining III

- Why now?
 - data is being produced and warehoused;
 - computing power is affordable;
 - interest in CRM is strong (focus on services, information as a product);
 - software is readily available.

© Iain Pardoe, 2006

8 / 8