Plastic Analysis

Unit - 5

Data Mining Applications

- Data mining is an interdisciplinary field with wide and diverse applications
  - There exist nontrivial gaps between data mining principles and domain-specific applications
- Some application domains
  - Financial data analysis
  - Retail industry
  - Telecommunication industry
  - Biological data analysis

Data Mining for Financial Data Analysis

- Financial data collected in banks and financial institutions are often relatively complete, reliable, and of high quality.

- **Design and construction of data warehouses for multidimensional data analysis and data mining**
  - View the debt and revenue changes by month, by region, by sector, and by other factors
  - Access statistical information such as max, min, total, average, trend, etc.
  - Data warehouses, data cubes, multifeature and discovery-driven data cubes, characterization and class comparisons, and outlier analysis all play important roles in financial data analysis and mining.

- **Loan payment prediction/consumer credit policy analysis**
  - **Attribute selection** and **attribute relevance ranking** - identify important factors and eliminate irrelevant ones.

Risk in Loan payment performance

- loan-to-value ratio, termof the loan, debt ratio (total amount of monthly debt versus the total monthly income), payment to- income ratio, customer income level, education level, residence region, and credit history.

Consumer credit rating

Financial Data Mining

- **Classification and clustering of customers for targeted marketing**
  - multidimensional segmentation by nearest-neighbor, classification, decision trees, etc. to identify customer groups or associate a new customer to an appropriate customer group.

- **Detection of money laundering and other financial crimes**
  - Integration from multiple DBs (e.g., bank transactions, federal/state crime history DBs).
  - Tools: data visualization, linkage analysis, classification, clustering tools, outlier analysis, and sequential pattern analysis tools (find unusual access sequences).

- Data visualization tools - To display transaction activities using graphs by time and by groups of customers.
- Linkage analysis tools - To identify links among different customers and activities.
- Classification tools - to filter unrelated attributes and rank the highly related ones.
- Clustering tools - to group different cases.
- Outlier analysis tools – to detect unusual amounts of fund transfers or other activities.
- Sequential pattern analysis tools - to characterize unusual access sequences.

Data Mining for the Retail Industry

Design and construction of data warehouses based on the benefits of data mining

- Retail data cover a wide spectrum (including sales, customers, employees, goods transportation, consumption, and services), there can be many ways to design a data warehouse for this industry.
- It involves deciding which dimensions and levels to include and what pre processing to perform.

Multidimensional analysis of sales, customers, products, time, and region

- The retail industry requires timely information regarding customer needs, product sales, trends, and fashions, as well as the quality, cost, profit, and service of commodities.

- Powerful multidimensional analysis and visualization tools - construction of sophisticated data cubes.
- ***Multifeature data cube*** - facilitates analysis on aggregates with complex conditions.

Analysis of the effectiveness of sales campaigns

- The retail industry conducts sales campaigns using advertisements, coupons, and various kinds of discounts and bonuses to promote products and attract customers.
- **Association analysis** may disclose which items are likely to be purchased together with the items on sale, especially in comparison with the sales before or after the campaign.

Customer retention—analysis of customer loyalty

- With customer loyalty card information, one can register sequences of purchases of particular customers.
- Customer loyalty and purchase trends can be analyzed systematically.
- **Sequential pattern mining** - used to investigate changes in customer consumption or loyalty and suggest adjustments on the pricing.

Product recommendation and cross-referencing of items

- By mining associations from sales records, one may discover that a customer who buys a digital camera is likely to buy another set of items.

Collaborative recommender systems

- Make personalized product recommendations during live customer transactions, based on the opinions of other customers.

- Product recommendations can also be advertised on sales receipts, on the Web to help improve customer service, aid customers in selecting items, and increase sales.

- Information such as “hot items this week” or attractive deals can be displayed together with the associative information in order to promote sales.

Data Mining & Big Data in Retail Industry

OBAMA's Win Predicted using Data Mining Techniques

Data Mining for the Telecommunication Industry

Multidimensional analysis of telecommunication data

- Telecommunication data are intrinsically multidimensional, with dimensions such as calling-time, duration, location of caller, location of callee, and type of call.

- It is often useful to consolidate telecommunication data into large data warehouses and routinely perform multidimensional analysis using OLAP(Online Analytical Processing) and visualization tools.

Fraudulent pattern analysis and the identification of unusual patterns

- Fraudulent activity costs the telecommunication industry millions of dollars per year.
- Identify potentially fraudulent users and their atypical **usage patterns**
- Detect attempts to **gain fraudulent entry** to customer accounts.
- Discover **unusual patterns** that may need special attention.
- Eg: busy-hour frustrated call attempts, switch and route congestion patterns, and periodic calls from automatic dial-out equipment (like fax machines) that have been improperly programmed.

- These patterns can be discovered by **multidimensional analysis, cluster analysis, and outlier analysis**.

Multidimensional association and sequential pattern analysis

Mobile telecommunication services

- One important feature of mobile telecommunication data is its association with spatiotemporal information. Spatiotemporal data mining may become essential for finding certain patterns.
- Eg: unusually busy mobile phone traffic at certain locations may indicate something abnormal happening in these locations.

Use of visualization tools in telecommunication data analysis

- Tools for OLAP visualization, linkage visualization, association visualization, clustering, and outlier visualization have been shown to be very useful for telecommunication data analysis.

Data Mining Concepts and Applications

Data mining applications

- Marketing
- Banking
- Retailing and sales
- Manufacturing and production
- Brokerage and securities trading
- Insurance
- Computer hardware and software
- Government and defense
- Airlines
- Health care
- Broadcasting
- Police
- Homeland security

Machine Learning

- An Introduction to Machine Learning Theory and Its Applications: A Visual Tutorial with Example. BY **NICK MCCREA** - SOFTWARE ENGINEER @ TOPTAL

- https://www.toptal.com/machine-learning/machine-learning-theory-an-introductory-primer

Big Data

- Introduction to BIG DATA: Types, Characteristics & Benefits
- https://www.guru99.com/what-is-big-data.html#4
- https://www.planetdata.eu/sites/default/files/presentations/Big_Data_Tutorial_part4.pdf

Cloud Computing

- Cloud computing by Chris Woodford

- http://www.explainthatstuff.com/cloud-computing-introduction.html