Applications and Trends in Data Mining

Assistant professor:
Laith R. Flaih
Computer Science Department

Applications and Trends in Data Mining

- Data mining is defined as a process used to extract usable data from a larger set of any raw data. It implies analyzing data patterns in large batches of data using one or more software
- Data mining has applications in multiple fields, like science and research
- As an application of data mining, businesses can learn more about their customers and develop more effective strategies related to various business functions and in turn leverage resources in a more optimal and insightful manner

Data Mining

- Data mining is the practice of automatically searching large stores of data to discover patterns and trends that go beyond simple analysis.
- Data mining uses sophisticated mathematical algorithms to segment the data and evaluate the probability of future events

DATA WAREHOUSE

- A DATA WAREHOUSE is a subject oriented, integrated,time-varying,Non-Volatilecollection of data in support of the management’s decision-making process.

Data Mining & DATA WAREHOUSE

- Too much data and too little information. There is a need to extract useful information from the data and to interpret the data. It is a process of discovering interesting knowledge from large amounts of data stored either in databases, data warehouses, or other information repositories.

Why Data Mining

- Rapid computerization of businesses produce huge amount of data How to make best use of data?
- A growing realization: knowledge discovered from data can be used for competitive advantage

Data Mining Stages

- Get a clear understanding of the problem to solve, how it impacts organization, and goals for addressing it. Selection
- Review the data, document it, identify data management and data quality issues. Data Understanding
- Get data ready to use for modeling. Data Preparation
- Use mathematical techniques to identify patterns within the data. Modeling
- Review the patterns discovered and assess their potential for use. Evaluation
- Put your discoveries to work in everyday business. Deployment

Market Basket Analysis

- Market basket analysis is a modeling technique based upon a theory that if you buy a certain group of items you are more likely to buy another group of items.
- This technique may allow the retailer to understand the purchase behavior of a buyer.
- This information may help the retailer to know the buyer's needs and change the store's layout accordingly.
- Using differential analysis comparison of results between different stores, between customers in different demographic groups can be done.

Education

- There is a new emerging field, called Educational Data Mining, concerns with developing methods that discover knowledge from data originating from educational Environments.
- The goals of EDM are identified as predicting students' future learning behavior, studying the effects of educational support, and advancing scientific knowledge about learning.
- Data mining can be used by an institution to take accurate decisions and also to predict the results of the student. With the results the institution can focus on what to teach and how to teach.
- Learning pattern of the students can be captured and used to develop techniques to teach them

Manufacturing Engineering

- Knowledge is the best asset a manufacturing enterprise would possess.
- Data mining tools can be very useful to discover patterns in complex manufacturing process.
- Data mining can be used in system-level designing to extract the relationships between product architecture, product portfolio, and customer needs data.
- It can also be used to predict the product development span time, cost, and dependencies among other tasks

Customer Relationship Management

- CRM is all about acquiring and retaining customers, also improving customers' loyalty and implementing customer focused strategies.
- To maintain a proper relationship with a customer a business need to collect data and analyze the information.
- This is where data mining plays its part. With data mining technologies the collected data can be used for analysis

Fraud Detection

- Billions of dollars have been lost to the action of frauds. Traditional methods of fraud detection are time consuming and complex.
- Data mining aids in providing meaningful patterns and turning data into information. Any information that is valid and useful is knowledge.
- A perfect fraud detection system should protect information of all the users.
- A supervised method includes collection of sample records. These records are classified fraudulent or non-fraudulent.
- A model is built using this data and the algorithm is made to identify whether the record is fraudulent or not

Lie Detection

- Apprehending a criminal is easy whereas bringing out the truth from him is difficult.
- Law enforcement can use mining techniques to investigate crimes, monitor communication of suspected terrorists.
- This filed includes text mining also.
- This process seeks to find meaningful patterns in data which is usually unstructured text.
- The data sample collected from previous investigations are compared and a model for lie detection is created. With this model processes can be created according to the necessity.

Financial Banking

- With computerized banking everywhere huge amount of data is supposed to be generated with new transactions.
- Data mining can contribute to solving business problems in banking and finance by finding patterns, causalities, and correlations in business information and market prices that are not immediately apparent to managers because the volume data is too large or is generated too quickly to screen by experts.
- The managers may find these information for better segmenting, targeting, acquiring, retaining and maintaining a profitable customer

Research analysis

- History shows that we have witnessed revolutionary changes in research.
- Data mining is helpful in data cleaning, data pre-processing and integration of databases. The researchers can find any similar data from the database that might bring any change in the research.
- Identification of any co-occurring sequences and the correlation between any activities can be known.
- Data visualization and visual data mining provide us with a clear view of the data

Criminal Investigation

- Criminology is a process that aims to identify crime characteristics.
- Actually crime analysis includes exploring and detecting crimes and their relationships with criminals. The high volume of crime datasets and also the complexity of relationships between these kinds of data have made criminology an appropriate field for applying data mining techniques.
- Text based crime reports can be converted into word processing files. These information can be used to perform crime matching process.

Bio Informatics

- Data Mining approaches seem ideally suited for Bioinformatics, since it is data-rich.
- Mining biological data helps to extract useful knowledge from massive datasets gathered in biology, and in other related life sciences areas such as medicine and neuroscience.
- Applications of data mining to bioinformatics include gene finding, protein function inference, disease diagnosis, disease prognosis, disease treatment optimization, protein and gene interaction network reconstruction, data cleansing, and protein sub-cellular location prediction

Thanks! Any questions?