Data Mining Issues and Artificial Intelligence for Modern world Problem

Sakura Yua*

Department of Information Sciences and Technologies, University of Tokyo, Tokyo, Japan

DESCRIPTION

The data mining is refers to extracting or "mining" knowledge from huge amount of knowledge. The team is really a misnomer. That the mining of gold from rocks or sand is verified to as gold mining instead of rock or sand mining. Thus the info mining has been highly appropriately named "knowledge mining from data," which is unfortunately somewhat long. It's a term might not reflect the stress on mining from high amount of knowledge. The info mining process breaks down into five steps. First, organizations collect data and cargo it into their data warehouses. Next, they store and manage the info, either on in-house servers or the cloud. Business analysts, management teams and knowledge technology professionals access the info and determine how they need to arrange it. Then, application software sorts the info supported the user's results, and eventually, the end-user presents the info in an easy-to-share format, like a graph or table.

MINING METHODOLOGY and USER INTERACTION

Mining different sorts of knowledge in databases: Different users could also be curious about different sorts of knowledge. Therefore it's necessary for data processing to hide a broad range of data discovery task.

Interactive mining of data at multiple levels of abstraction: the info mining process must be interactive because it allows users to focus the look for patterns, providing and refining data processing requests supported the returned results.

Incorporation of background: To guide discovery process and to precise the discovered patterns, the background are often used. Background could also be wont to express the discovered patterns not only in concise terms but at multiple levels of abstraction.

Data mining query languages and unplanned data processing: data processing command language that permits the user to explain unplanned mining tasks, should be integrated with a knowledge warehouse command language and optimized for efficient and versatile data processing.

Presentation and visualization of knowledge mining results: Once the patterns are discovered it must be expressed in high level languages, and visual representations. These representations should be easily understandable.

Handling noisy or incomplete data: The info cleaning methods are required to handle the noise and incomplete objects while mining the info regularities. If the info cleaning methods aren't there then the accuracy of the discovered patterns are going to be poor.

Pattern evaluation: The patterns discovered should be interesting because either they represent public knowledge or lack novelty.

Efficiency and scalability of knowledge mining algorithms: so as to effectively extract the knowledge from huge amount of knowledge in databases, data processing algorithm must be efficient and scalable.

Parallel, distributed, and incremental mining algorithms: The factors like huge size of databases, wide distribution of knowledge, and complexity of knowledge mining methods motivate the event of parallel and distributed data processing algorithms.

Handling of relational and sophisticated sorts of data: The database may contain complex data objects, multimedia data objects, spatial data, temporal data etc. it's impossible for one system to mine of these quite data.

Mining information from heterogeneous databases and global information systems: The info is out there at different data sources on LAN or WAN. These data source could also be structured, semi structured or unstructured. Therefore mining the knowledge from them adds challenges to data processing.

Correspondence to: Sakura Yua, Director, Department of Information Sciences and Technologies, University of Tokyo, Tokyo, Japan, E-mail: yuasak@uot-jpn.com