Fuzzy-Rule Based Adaptive Data Warehouse: An Extension of Data Warehouse as Knowledge Warehouse

Rajdev Tiwari, ABES Institute of Technology, India
Anubhav Tiwari, Amity University, India
Manu Pratap Singh, Dr. B. R. Ambedkar University, India

ABSTRACT

Data Warehouses (DWs) are aimed to empower the knowledge workers with information and knowledge which helps them in decision making. Technically, the DW is a large reservoir of integrated data that does not provide the intelligence or the knowledge demanded by users. The burden of data analysis and extraction of information and knowledge from integrated data still lies upon the analyst’s shoulder. The overhead of analysts can be taken off by architecting a new generation data warehouses systems those shall be capable of capturing, organizing and representing knowledge along with the data and information in it. This new generation DW may be called as Knowledge Warehouse (KW) shall exhibit decision making capabilities themselves and can also supplement the Decision Support Systems (DSS) in making decisions quickly and effortlessly. This paper proposes and simulates a fuzzy-rule based adaptive knowledge warehouse with capabilities to learn and represent implicit knowledge by means of adaptive neuro fuzzy inference system (ANFIS).

Keywords: Adaptive Neuro Fuzzy Inference System (ANFIS), Artificial Neural Networks, Data Warehouse, Decision Support Systems (DSS), Fuzzy Logic, Intelligent Data Warehouse, Knowledge Warehouse (KW), Rules, Soft Computing

INTRODUCTION

The traditional Data Warehouse supports the cleaning, transformation and loading of corporate data into libraries. The corporate data contained in warehouse may be of multiple granule size to make it capable of handling queries of versatile nature.

According to Husemann, Lechtenbörger, and Vossen (2000), a data warehouse design is commonly supported by a conceptual data model called multidimensional model by which users could view data from different dimensions necessary for analysis purposes. In multidimensional model, data are represented in terms of facts and dimensions where each fact is associated to multiple dimensions. In this manner, facts are the focus of interest by...
which they are analyzed through the quantifying context stored in measures and the qualifying context determined through dimension levels. Categorizing data along dimensions is a mean to organize them into hierarchical levels so that data can be viewed from their finer to coarser granularities as per Agrawal, El Abbadi, Singh, and Yurek (1997). Data warehouses store data and can be a source of knowledge but do not store knowledge directly. Knowledge is presented in the DW in the form of analysis reports, stored statically, the value and importance of which may vary from time to time. To make this knowledge available, Metadata is defined which describes data attributes, transformations and aggregation levels. Metadata helps socialize the Data Warehouse to the knowledge workers so that they can discover information contained and often hidden within the data.

Applications of artificial intelligence (AI) technology in the form of knowledge-based systems within the context of database design have been extensively researched particularly to provide support within the conceptual design phase can be found in Phipps and Davis (2002), Hahn, Sapia, and Blaschka (2000), and Sitompul and Noah (2003, 2005). However, a similar approach to the task of data warehouse design has yet to be seriously initiated. A design methodology is proposed by Sitompul and Noah (2006) for conceptual data warehouse design called the transformation-oriented methodology, which transforms an Entity-Relationship (ER) model into a multidimensional model based on a series of transformation and analysis rules.

The next generation Data Warehouse systems are constructed as Intelligent Data Warehouses (IDW). Bramblett (2002) says that IDW’s are Data Warehouses that are managed as active data sources by a Knowledge Warehouse (KW). The Knowledge Warehouse applies knowledge objects that are created and controlled by software engines using expert system models. Knowledge objects supply the rules, methods and procedures in a reusable way and are activated to not only link Data Warehouse data with a specific business process, but to re-package data so that the IDW can manage the business. The existing enterprise-wide information delivery systems provided in a data warehouse can be leveraged and extended to create a knowledge warehouse detail can be found in Nemati, Steiger, Iyer, and Herschel (2002). This warehouse can be used as a clearinghouse of knowledge to be used throughout the organization by the knowledge workers to support their knowledge intensive decision-making activities. The KW can also evolve over time by enhancing the knowledge it contains.

The Data Warehousing System, which began as a low volatility system, is now a system that may integrate DSS, batch and OLTP processing, and that therefore may incorporate considerable volatility. According to Firestone (2000), the new generation Data Warehousing raises the following issues.

- Achieve dynamic integration.
- Comprehensively integrate and support knowledge production.
- Store knowledge for high capability decision support.
- Efficiently deliver tactical decision support using volatile data stores.
- Integrate ERP systems.
- Integrate increasingly varied business process engines.

Therefore, Firestone (2000) says that data warehouse systems need an integrative component with the capabilities of the Artificial Knowledge Manager (AKM) to successfully resolve these issues. Till now, Knowledge Warehouse or Intelligent Data Warehouse is either just a concept or have been implemented partially. The word knowledge warehouse and intelligent data warehouse is used interchangeably in the literature, hence in this paper also.

This paper suggests a model of knowledge warehouse and demonstrates how the Dynamic Knowledge Pool (DKP) shall be maintained as part of data warehouse using soft computing techniques to make them adaptive in nature. In this proposed method neuro-fuzzy system is used for extracting and representing the
Natural Language Intelligences: The Virtual or Digital Aristotle
[www.igi-global.com/chapter/natural-language-intelligences/28318?camid=4v1a](www.igi-global.com/chapter/natural-language-intelligences/28318?camid=4v1a)

Innovation Strategies of New Product Development (NPD)
[www.igi-global.com/article/innovation-strategies-of-new-product-development-npd/129084?camid=4v1a](www.igi-global.com/article/innovation-strategies-of-new-product-development-npd/129084?camid=4v1a)