

“Technically Research on Framework For Data Mining - Multidimensional Database”

Research Scholars - TiruveedulaGopi Krishna, Dr.MohamedAbdeldaim, Mr.SabaAldeen,
Faculty of Arts and Science,Department of Computer Science,Hoon,Al-jufra,Sirt
University,Libya

Abstract

Multidimensional database and OLAP(On Line Analytical Processing) tools provide an efficient framework for data mining and have led to the so-called OLAP mining architecture. Besides, data from real world are often imperfect, either because they are uncertain or because they are imprecise. Moreover, the use of fuzzy set theory in data mining systems enhances the understandability of the discovered knowledge when considering numerical attributes and leads to more generalizable rules. This is reason why we have proposed an approach to perform OLAP-based mining using multidimensional databases and fuzzy decision trees.

In recent works an extension of multidimensional database has been defined in order to handle imperfect information and flexible multidimensional queries.

numerical descriptions, leading to more understandable knowledge.

However, many fuzzy data mining systems extract knowledge from that file. They do not take advantage of the performances of database management systems, and even less of the benefits from the multidimensional model.

Thus, we propose to integrate multidimensional databases with fuzzy learning systems, for this purpose, the multidimensional model has been extended to the management of imperfect data and flexible queries in for the definition of fuzzy multidimensional databases. In this paper, we show that this model provides the needed operators for its integration in OLAP mining architecture.

This solution has been implemented and tested using the OLAP software from Oracle: Oracle Express Server. We strongly insist on the fact that this model generalizes the classical one, and provides a framework to handle all classical crisp cases, since fuzzy theory provides means to handle imperfect and classical data

Introduction

Learning systems provide efficient data mining tools when they are integrated with multidimensional databases. OLAP tools and multidimensional databases provide indeed a framework for hierarchies and enhance the calculus of aggregates.

Besides, the introduction of fuzziness in the data mining process is very interesting to deal with numerical attributes. In this framework, fuzzy set theory provides tools to manipulate data with words instead of using

- 1. Data warehouse and OLAP technology:**
Data warehouse generalize and consolidate data in multidimensional space. The construction of data warehouse involves data cleaning, data

integration and data transformation and can be viewed as an important preprocessing step for the data mining. Moreover, data warehouses provide on-line analytical processing (OLAP) tools for the interactive analysis of multidimensional data of varied granularities, which facilitates effective data generalization and data mining. Many other data mining functions, such as association, classification, prediction and clustering can be integrated with OLAP operations to enhance interactive mining of knowledge at multiple levels of abstraction. Hence, the data warehouses has become an increasingly important platform for data analysis and on-line analytical processing and will provide an effective platform for data mining. Therefore data warehousing and OLAP form an essential step in the knowledge discovery process.

2. A Multidimensional Data Model

Data warehouses and OLAP tools are based on a multidimensional data model. This model views data in the form of data cube. In this a data cube allows data to be modeled and viewed in multiple dimensions. It is defined by dimensions and facts.

3. OLAP operations in the multidimensional Data Model

In multidimensional model, data are organized into multiple dimensions, and each dimension contains multiple levels of abstraction defined by concept hierarchies. This organization provides users with the flexibility to view data from different perspectives. A number

of OLAP data cube operations exist to materialize these different views, allowing interactive querying and analysis of the data at hand. Hence, OLAP provides a user-friendly environment for interactive data analysis.

4. OLAP systems versus Statistical Databases.

Many of the characteristics of OLAP systems, such as the use of a multidimensional data model and concept hierarchies, the association of measures with dimensions, and the notions of roll-up and drill-down, also exist in earlier work on statistical database(SDBs). A statistical database is a database system that is designed to support statistical applications.

5. Fuzzy Multidimensional Databases

Many models have been proposed for multidimensional databases. Roughly speaking, a multidimensional database is set of hypercube (hereafter cubes), defined on dimensions, which may be organized hierarchically. One dimension of particular interest is chosen as the measure whose values are stored in cube cells. Operations are defined to visualize and manipulate cubes (e.g rotation, selection by slice nod dice, roll-up drill down).

Conclusion

In this paper, a new approach for data mining is described, doubling fuzzy multidimensional databases and fuzzy data mining systems and achieving knowledge discovery from imperfect

References:

1. R.Agrawal,T.Imielinski, and A.Swami, Mining association rules between sets of items in large databases. In proc.of AGM SIGMOD conf.,
2. E.F Codd, S.B.Codd, and C.T.salley. providing OLAP to user –analysis: an IT mandate. White thesis, 1993.
3. B.Bouchon-Meunier,M.Rifqi, and S.Bothorel. Fuzzy sets of Systems, 85: 143-153,1996.

data. It uses these kinds of data repositories to extract relevant knowledge from large data sets from the real world.

IJERT