+IJESRT

INTERNATIONAL JOURNAL OF ENGINEERING SCIENCES & RESEARCH TECHNOLOGY

AN IMPLEMENTATION OF "DECISION MAKER FOR EDUCATIONAL DOMAIN

Shabina Kausar*, Prof. Kalpana Malpe
* M. Tech. (CSE) GNIET, Nagpur, Maharashtra, India
Assistant Professor GNIET, Nagpur, Maharashtra, India

ABSTRACT

Utilizing a decision support system is a proactive way to use data to manage, operate, and evaluate educational institute in a better way. Depending on the quality and availability of the underlying data, such a system could address a wide range of problems by distilling data from any combination of education records maintenance system. The data mining from data warehouse can be a ready and effective system for the decision makers. Data-driven decision support systems, such as data warehouses can serve the requirement of extraction of information from more than one subject area. Data warehouses standardize the data across the organization so as to have a single view of information. Data warehouses can provide the information required by the decision makers. Developing a data warehouse for educational institute is the less focused area since educational institutes are non-profit and service oriented organizations. In present day scenario where education has been privatized and cut throat competition is prevailing, institutes needs to be more organized and need to take better decisions. Institute's enrollments are increasing as a result of increase in the number of branches and intake. Now a day, any reputed Institute's enrollments count in to thousands. The cost of building a data warehouse is expensive for any educational institution as it requires data warehouse tools for building data warehouse and extracting data using data mining tools from data warehouse. The present study provides an option to build data warehouse and extract useful information using data warehousing and data mining open source tools. In this paper we have explored the need of data warehouse / business intelligence for an educational institute, the operational data of an educational institution has been used for Experimentation. The present study provides an option to build data warehouse and extract useful information using data warehousing and data mining open source tools.

INTRODUCTION

Data warehouse obtains the data from a number of operational data base systems which can be based on RDBMS/ERP package, etc. The data from these sources are converted into a form suitable for data warehouse. This process is called Extraction, Transformation and Loading (ETL). In addition to the target database, there will be another database to store the metadata, called the metadata repository.

This data base contains data about data-description of source data, target data and how the source data has been modified into target data. The client software will be used to generate reports.

In present day scenario where education has been privatized and cut throat competition is prevailing, institutes needs to be more organized and need to take better decisions. Now a day, the educational institutes have to generate funds for their research and other operational activities as the government funding has been limited to aided institutes. By applying data mining on data warehouse is a proactive way to use data to manage, operate, and evaluate educational institute in a better way. Data-driven decision support systems, such as data warehouses can serve the requirement of extraction of information from more than one subject area. Data warehouses standardize the data across the organization so as to have a single view of information. Data warehouses can provide the information required by the decision makers. Developing a data warehouse for educational institute is the less focused area since educational institutes are non-profit and service oriented organizations.

DATAWARE HOUSE ENVIRONMENT

Utilizing a decision support system is a proactive way to use data to manage, operate, and evaluate educational institute in a better way. Depending on the quality and availability of the underlying data, such a system could address a wide range of problems by distilling data from any combination of education records maintenance system. The data mining from data warehouse can be a ready and effective system for the decision makers. The data from reputed engineering college namely R V College of Engineering, Bangalore, Karnataka, India, has been considered for this study. Fig. 1 shows the data warehouse architecture of RV College where source systems are smart campus, asset management server and csv files, the information is spread across diverse platforms, data from different sources is collected and then consolidated to produce required report. ETL activities are performed to extract the data from heterogeneous sources and load into staging and then load the data into dimension and fact tables as per the schedules. We proceed to extract the BI report from data warehouse on demand based on requirement from the management. In an educational institute, main information required will be regarding key components of the education institute, namely students, employees and infrastructure. The purpose of this paper was to investigate current system of information delivery and proposing a better system for timely, accurate, consistent information delivery to the decision makers of the educational institute. The paper has been prepared in order to extend the usage of current available technology in decision making processes of educational institute.
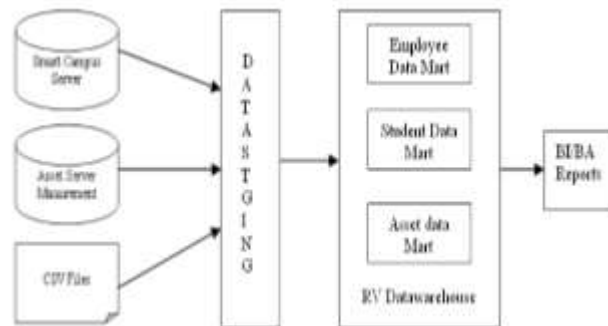

Fig. 1. Engg_Data warehouse architecture

Data warehouse enables the decision makers with benefits listed below.
1) Phenomenal improvements in turnaround time for data access and reporting
2) Standardizing data across the organization so that there will be one view of information.
3) Merging data from various source systems to create a more comprehensive information source.
4) Reduction in costs to create and distribute information and reports.
5) Encouraging and improving fact-based decision making.

TECHNOLOGY USED

Data Mining

Generally, data mining (sometimes called data or knowledge discovery) is the process of analyzing data from different perspectives and summarizing it into useful information - information that can be used to increase revenue, cuts costs, or both. Data mining software is one of a number of analytical tools for analyzing data. It allows users to analyze data from many different dimensions or angles, categorize it, and summarize the relationships identified. Technically, data mining is the process of finding correlations or patterns among dozens of fields in large relational databases

Pipe Sort :

The Pipe Sort method combines the optimisations share-sorts and smallest-parent to get the minimum total cost. Here a number of statistical procedures for this purpose. The input to the algorithm is the search lattice, which is a graph where a vertex represents a group-by of the cube. A directed edge connects group by i to group-by j whenever j can be generated from i and j has exactly one attribute less than i (i is called the parent of j). Thus, the out-degree of any node with k attributes is k.
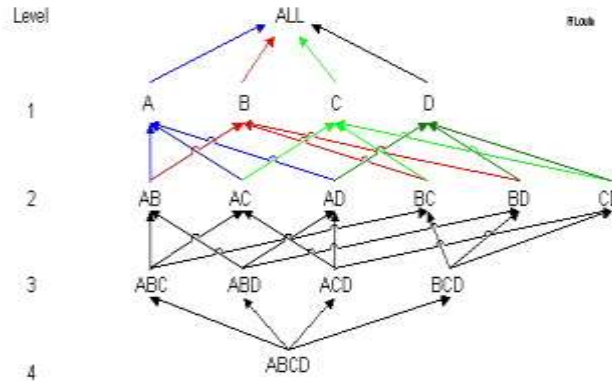
Fig2: Pipe sort algorithm

Algorithm: Pipe Sort
Step1: Input: search lattice with A() and S() edge costs
Step2: For (level k = 0 to N -1 do begin
Generate-Plan ( );
Step3: for (each Group-by g in level k +1) do begin
Fix the sort order of g as the order of the
Step4: Group by connected to g by an A () edge;
Generate-Plan ( )
Make k additional copies of each level k +1 vertex;
Step5: Connect each copy vertex with the same set of level k vertices as the original vertex;
Assign cost A (eij) to edge eij from the original vertex and S(eij) to edge from the copy vertex ;
Find minimum cost matching on the transformed levels.

Viterbi algorithm :
 The Viterbi algorithm is a dynamic programming algorithm for finding the most likely sequence of hidden states – called the Viterbi path – that results in a sequence of observed events, especially in the context of Markov information sources and hidden Markov models.
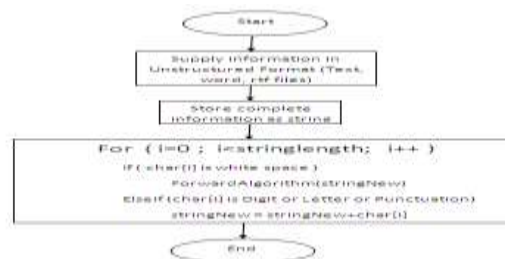

Fig3: Viterbi Algorithm

The algorithm has found universal application in decoding the convolution codes used in both CDMA and GSM digital cellular, dial-up modems, satellite, deep-space communications and 802.11 wireless LANs. It is now also commonly used in speech recognition, speech synthesis, diarization, keyword spotting, computational linguistics, and bioinformatics. For example, in speech-to-text (speech recognition), the acoustic signal is treated as the observed sequence of events and a string of text is considered to be the "hidden cause" of the acoustic signal. The Viterbi algorithm finds the most likely string of text given the acoustic signal.

Forward algorithm

The forward algorithm in the context of a hidden Markov model is used to calculate a 'belief state': the probability of a state at a certain time given the history of evidence. The process is also known as filtering. The forward algorithm is closely related to but distinct from the Viterbi algorithm.
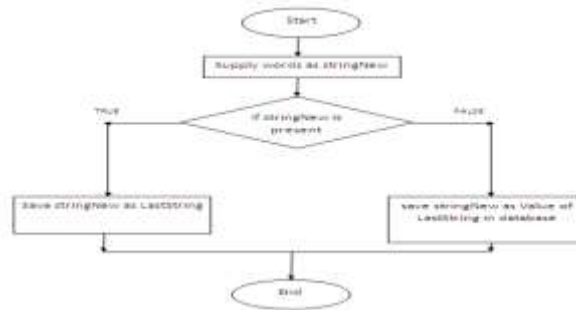

Fig4: Forward algorithm

Need of the system

Today's scenario demands an automotive system that can support an organization to take micro-level decisions in a timely manner without the need to depend on their IT staff. Educational institutes demands for a decision support system which can help to generate report in specific manner and gives information regarding students, employees, assets etc.

The visual lookout of the system is in such a way that it doesn't create any type of confusion for end user due to which to work on the new system is very interesting. The objective is to investigate current system of information delivery and proposing a better system for timely, accurate, consistent information delivery to the decision makers of the educational institute.

The project is proposed in order to extend the usage of current available technology in decision making processes of educational institute.

BI-Reporting

This refers to the variety of capabilities that can be provided to the users to leverage the presentation area for analytic decision making. All data access tools query the data in the data warehouse presentation area. A data access tool can be as simple as an ad hoc query tool or as complex as sophisticated data mining application. The majority of the users use pre-built parameter driven analytic applications to access the data. This enables them to retrieve the required information and analyze hidden pattern in the retrieved data [11].Using suitable data mining techniques, the useful information can be extracted from the data warehouse. Data mining form three main components of the institute, namely Employees, Students and Infrastructure. Employee data mart can provide the users with the information such as career growth and attrition rate. Student mart can provide the information related to the student like best outgoing student considering his academic and non academic activities. Information regarding assets such as the investment in a particular financial year can also be accessed.

Working procedure
- First of all start the software.
- And then start the connect screen.
- And logon to the SQL server as you know such that if you want to connect local SQL server then you have to just select windows authentication and click on connect button, otherwise you have another choice to select SQL authentication in which you have to provide username and password to connect it to server machine SQL server.
- After login the software next click on the link upload database.
- After that there is a button known as browse database click on it we have maximum three database choice.

- After that we can check the tables present in the current database by clicking on tables button below the database.

CONCLUSIONS

The newly generated systems visual lookout becomes very standard and easy for working. Institute can take micro-level decisions in a timely manner without the need to depend on their IT staff. They can perform extensive analysis of stored data to provide answers to the exhaustive queries to the administration cadre. This helps them to formulate strategies and policies for employees and students. This helps students and Employees in making decisions. They are the ultimate beneficiaries of the new policies formulated by the decision makers and policy planner's extensive analysis on student and employee relate data.

REFERENCES

[1] Alex Berson, Data Warehousing Data Mining & OLAP, Computing Mcgraw-Hill, November 5, 1997.
[2] Ralph Kimball, the Data Warehouse ETL Toolkit, Wiley India Pvt Ltd., 2006.
[3] KV. K. K. Prasad, Data warehouse development Tools, Dreamtech Press, 2006.
[4] W. H. Inmon, Building the Data Warehouse. Wiley; 3rd edition March 15, 2002.
[5] McMillen and Randy McBroom.,'what academia can gain from building a data warehouse' no.1,pp.41-46.2008
[6] Channah F. Naiman, Aris M. Ouksel "A Classification of Semantic Conflicts in Heterogeneous Database Systems", Journal of Organizational Computing, vol. 5, 1995.
[7] John Hess, "Dealing With Missing Values In The Data Warehouse" A Report of Stonebridge Technologies, Inc-1998.
[8] Manjunath T.N, Ravindra S Hegadi, Ravikumar G K."Analysis of Data Quality Aspects in DataWarehouse Systems",(IJCSIT) International Journal of Computer Science and Information Technologies, vol. 2 (1) , 2010, 477-485.
[9] Amit Rudra,Emilie Yeo (1999) "Key Issues in Achieving Data Quality and Consistency in Data Warehousing among Large Organizations in Australia", Proceedings of the 32nd Hawaii International Conference on System Sciences – 1999.
[10] Jaideep Srivastava, Ping-Yao Chen,"Warehouse Creation-A Potential Roadblock to Data Warehousing", IEEE Transactions on Knowledge and Data Engineering January/February 1999 (vol. 11, no. 1) pp.118-126.