iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 10 **Issue:** VI **Month of publication:** June 2022

DOI: https://doi.org/10.22214/ijraset.2022.43782

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 10 Issue VI June 2022- Available at www.ijraset.com

Role of Data Mining Technique: A Boon to Society

Ashwini D N[1], Soumya Dass B[2]
[1, 2]*Department of Computer Science, Bengaluru city University*

Abstract: Datamining is a method of finding interested patterns from huge volume of data. Datamining techniques helps to make business decisions. It analyses information from multiple sources like DataMart, databases. In this paper, we are focussing on datamining tasks and its variety of applications in different fields, which is boon to the society.
Keywords: KDD, Decision Tree, OLAP servers, Cube API, ODBC, Frequent patterns

I. INTRODUCTION

Every day huge data is required and stored on storage devices. This is fed in to computer networks for various purposes like business, medicine, science so on. So, data mining is technique used to retrieve information from huge data. It is also referred as knowledge mining from data (KDD). It helps to find interesting patterns from data source.

A. Types Of Data That Be Mined

1) Database data: collection of related data that is stored and accessed through database management system (DBMS). Example: An organization's sales, finance, marketing database.
2) Data Warehouse: it is a storage structure that is formed from multiple sources like data mart and databases.

II. TASKS OF DATAMINING

According to Kamber and Han, Data mining has four major tasks:

A. Data Characterization and Data Discrimination

1) Data Characterization: The characterization of data is a common features of objects in a target class. A SQL query is used to collect data from user specified class the output of characterization is depicted using Bar charts, curves, and multidimensional cube.
2) Data Discrimination: User specifies two types of classes, target and contrast class. Comparing both classes' general features is data discrimination. It is retrieved using SQL query.

B. Prediction

It classifies one of the attribute values into possible classes. It predicts continuous values. Regression, Decision Tree are the methods to prediction problem. Example: algorithm attempts to predict customer's expenditure.

C. Classification

Classification is used to construct a model to differentiate data classes. It can be used to predict unknown class labels. It is used to predict discrete values. Different methods to represent classification model are Decision Tree, IF-THEN rules, Neural Networks. Example: Loan Manager can decide loan applicant's status is safe, risky.

D. Mining frequent pattern & Association

The items which are purchased together frequently are frequent patterns. Example: milk and curd. To know frequently brought pattern, Association rules are generated. Association rules are rejected as infrequent, if they do not satisfy minimum support and minimum confidence.

E. Outlier Analysis

Data objects that lie outside cluster are called as outliers. They are often referred to as noise or exceptions which exhibit indifferent behaviour. They play important role in few applications as fraud detection. Example: credit card fraud detection.

©IJRASET: All Rights are Reserved

F. Cluster Analysis

Clustering is group of data which forms a cluster. Unlike classification, however, class labels are undefined in clustering and it is up to the clustering algorithm to find suitable classes. Clustering is often called unsupervised classification since provided class labels do not execute the classification. Many clustering is based on the maximizing intra-class similarity and minimizing [1].

III. DATAMINING INTERFACE

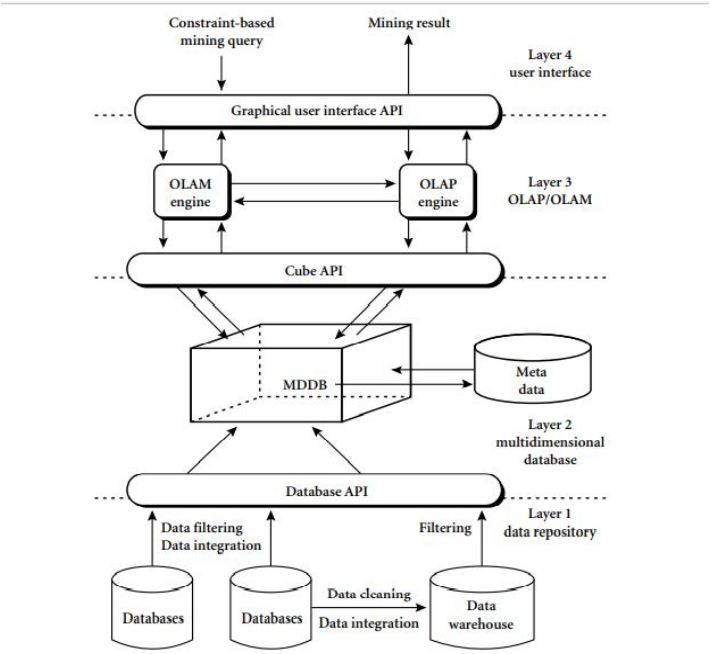


Fig .1 OLAM System

OLAM is integrated with OLP servers to accept user queries through graphical user interface. GUI functions communicates with cube API. Data cube is formed by accessing data from multiple sources. An ODBC connection is established to store data in data warehouse.

OLAM servers performs many datamining functionalities like classification, association.

IV. APPLICATIONS OF DATAMINING

Data mining is used in our day to day life, today most of the organizations use data mining in many fields for analysing huge data. Here are few applications of data mining that we come across in real life.

A. Telecommunication.

The telecommunications industries generate and store enormous amount of data. They use datamining to design their marketing campaigns and to retain customer. The large amount of data such as customer details, billing information, email, text messages, call details, web data transmissions, and customer service etc. The advent of data mining technology promised solutions to handle these problems and for this reason the telecommunications industry was an early adopter of data mining technology.

B. Retail Sector

Retail data mining helps the supermarket and retail sector owners to know and identify the choices and buying behavior of the customers, finding user shopping patterns and purchase history, enhance the quality of user service, buying preferences of the customers, to achieve better customer retention and satisfaction, increase goods consumption ratios, design more effective goods transportation and decrease the cost of business. The retail sector design the placements of products on shelves and bring out offers on items such as coupons on matching products, and special discounts on some products.

C. Artificial Intelligence

Artificial Intelligence is the study to create an intelligent machine that can work like humans, which does not depend on learning or feedback, rather it has a direct programmed control system. The AI systems come up with the solutions to the problems on their own by calculations. A system or a machine is made artificially intelligent by feeding it with relevant patterns of data. These patterns of data come from data mining outputs. The data mining technique is used by the AI systems for creating solutions. Data mining serves as a foundation for artificial intelligence. The recommender systems use data mining techniques to make personalized recommendations when the customer is interacting with the machines. The AI is used on mined data such as giving product recommendations based on the past purchasing history of the customer in many retail sectors.

D. Ecommerce

Data mining in E-commerce has a vital role of repositioning the e-commerce company for supporting the enterprise with the required information concerning the business. Many E-commerce sites use data mining to offer cross-selling and upselling of their products. The shopping sites such as Amazon, Flipkart show “People also viewed”, “frequently bought together”, “review rating”, “option to compare the products”, based on product viewed by the customers who are interacting with the site, and also recommendations are provided over the purchasing history of the customers of the website.

E. Crime Detection and Prevention

Data mining is used as a tool to detect and prevent criminal behavior. It detects outliers across a vast amount of data. The criminal data records include all the details related the crime that has happened.

Data Mining will study the unexpected patterns, discovery of hidden knowledge, and new rules from large databases to predict future events with better accuracy. It can help in finding out which area in the city is more prone to crime, how much police personnel should be deployed, which age group should be targeted, vehicle numbers to be scrutinized, etc. Advanced technology has allowed to analyze large quantities of data relatively in new field known as crime analyze, which has become an emerging field in law enforcement without standard definitions.

F. Research Analysis

Researchers use Data Mining tools as a process to explore and extract useful and valuable information from large data in data cleaning, data pre-processing and integration of databases. Data mining helps in associations between the parameters under research such as environmental conditions, weather forecasting, spread of diseases, prevention of diseases, growth of organization, educational sectors and many more fields where Identification of any co-occurring sequences and the correlation between any activities can be known for research purpose.

G. Agriculture

Agriculture has been an obvious target for big data. Data mining helps in finding environmental conditions, changes in soil, input levels, combinations of crops, existing crop, soil and climatic change data, commodity pricing and yield of vegetables with the amount of water required by the plants have made it all easier for farmers to use this information to get help to make critical farming decisions.

H. Automation

By using data mining, the computer systems automatically build accurate and interpretable predictive models by recognizing patterns among the parameters which are under comparison. The system will store the patterns that will be useful in the future to achieve goals. Automated data mining tools helps in meeting the targets through machine learning and relieve the burden on data teams by consolidating data sources into one analytics search bar for everyone to use. The user has to just point to the data and it does the rest.

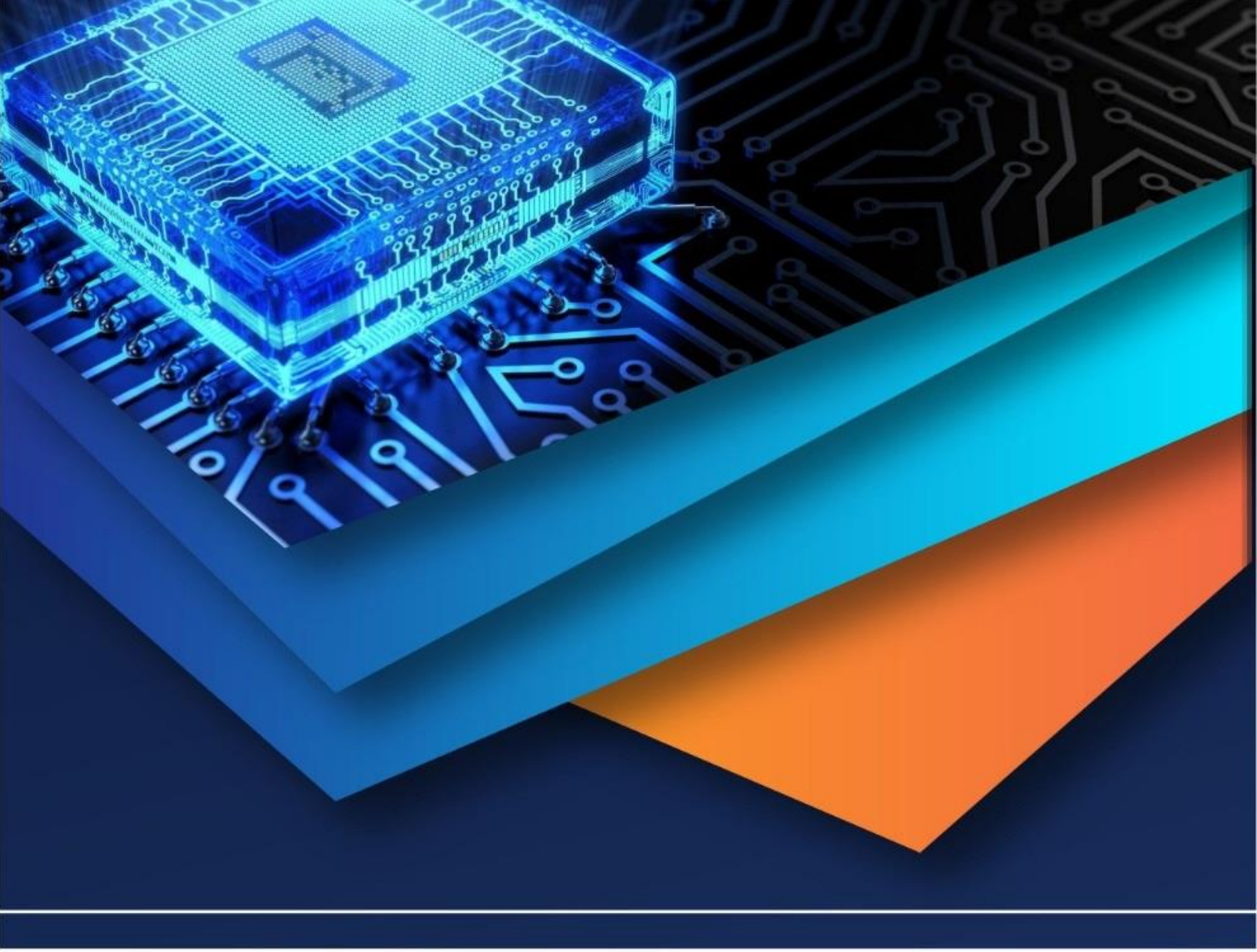
V. CONCLUSIONS

Datamining is a process of transforming data into information. The functionalities of datamining enable it to play vital role in different real-time applications. The paper gives us a brief idea about the applications in different fields which help humans in day-to-day life making their work easy and quick solutions. This paper has focused on the how datamining techniques are implemented and used. Hence datamining has become a key-part of the society.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

REFERENCES

[1] Jiawei Han and Michael Kamber "Datamining Concepts & Techniques", 3rd Edition
[2] Michael Steinbach, Pang-Ning Tan, and Vipin Kumar "Introduction to Data Mining", 2005
[3] Margaret H. Dunham "Data mining introductory and advanced topics", Pearson Education, 2013
[4] Johina, and Vikas Kamra "A Review: Datamining Techniques used in education Sector", (IJCSIT) International Journal of Computer Science and Information Technologies, Vol. 6 (3) , 2015, 2928-2930
[5] Bharati M. Ramageri "Data Mining Techniques And Applications "Indian Journal of Computer Science and Engineering Vol. 1 No. 4 301-305



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)