An Overview of Data Mining: Proposed four more Methods Exalt Performance

Dr.G.Srinaganya[1]

Assistant Professor, Department of Computer Science
National College (Autonomous), Trichy, Tamil Nadu, India.

M.Gayathri[2]

Research Scholar, Department of Computer Science
National College (Autonomous), Trichy, Tamil Nadu, India

Abstract: ***The term Data mining refer to elicit information to identify patterns, trends, and useful data it can be allow the business to take the data directed decision from immense set of data is called data mining. Data mining is take advantage of complicated mathematical problems or methods for part and assesses the probability of future events. It's also called Knowledge Discovery of Data (KDD).Data mining providing investigating hide pattern of guidance to various outlook and classification into useful data, which is gathered in specific areas like as data warehouses, efficient analysis, data mining algorithm, helping decision making and other data requirement to eventually cost-cutting and generating revenue. This paper aims to overview the basic characteristics of Data Mining and how it will work and we proposed four more Methods to exalt performance.***

Keywords: Data Mining :Data Mining Overview; Characteristics of Data Mining; Real Time applications use of Data Mining ; Sources of Data Mining; Advanced methods of Data Mining.

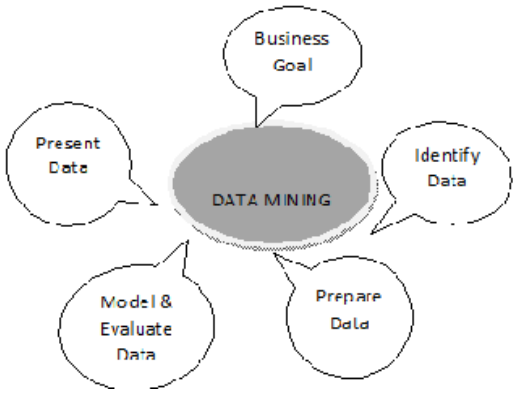
1. Introduction

The state of the art, data mining is everywhere, all the business sector are computerized, so each and every seconds they produce large amount of data, these data's can help in decision making and problem solving, which its requires more advanced storage and processing systems to extra knowledge from those gathered data. So all the establishment has the greatest challenges for managing these huge set of data. The statistical beginnings of data mining were set into motion by Bayes' Theorem in 1763 and discovery of regression analysis in 1805. Through the Turing Universal Machine (1936), the discovery of Neural Networks (1943), the development of databases (1970s)and genetic algorithms (1975), and Knowledge Discovery in Databases (1989), the stage was set for our modern understanding of what data mining is today. And, as the growth of computer processors, data storage, and technology exploded during the 1990s and 2000s, data mining became not only more powerful, but also more prolific in all kinds of situations. In 2003, the book Money ball introduced data mining to a much broader audience through the story of a professional baseball team's analytics-driven approach to roster building. Now, with companies employing big data solutions in a growing variety of situations, data mining plays a critical role in countless industries. In this paper we are going to see how the Data Mining will handle this vast amount of data and we proposed four Methods Performance.

2. Sources of Data

Files, Binary form with a structure that can be easily extracted by data mining algorithms. ...

- Relational Databases. ...
- Data Warehouse. ...
- Transactional Databases. ...
- Multimedia Databases. ...
- Spatial Database. ...



- Time-series Databases. ...
- WWW.

3. Types of Data Mining

It has two basic parts that are as following:

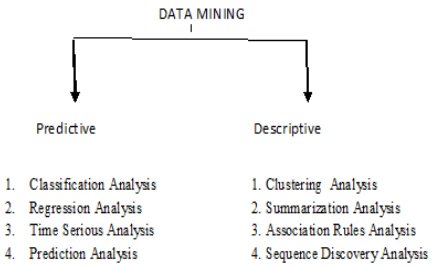
1. Predictive Data Mining Analysis
2. Descriptive Data Mining Analysis

Predictive Data Mining Analysis: It analysis works on the data that may help in business. Predictive Data-Mining can also divided into four categories that are listed below:

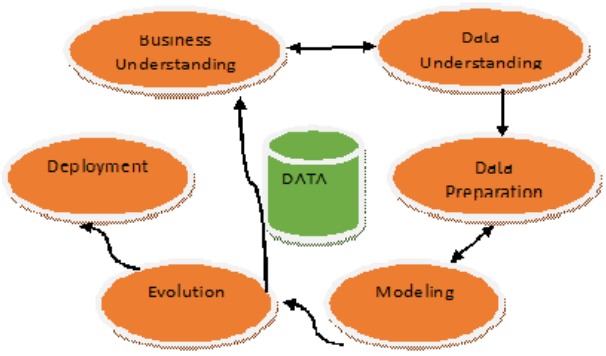
- Classification Analysis
- Regression Analysis
- Time Serious Analysis
- Prediction Analysis

Descriptive Data Mining Analysis: To summarize or turn given data into relevant information. It can also divide into four types that are as follows:

- Clustering Analysis
- Summarization Analysis
- Association Rules Analysis
- Sequence Discovery Analysis



4. Overview of Data Mining



The data mining process is usually the following steps.

4.1 Understand the Business

This stage makes a understanding of the project at your hand, the current business metric for success.

4.2 Understanding the data

Once if you have cleared format the project and business goals, next it will be comes all the relevant information data that will be to solve all the problems. This all data is collected from all the database sources like cloud storage and soils.

4.3 Preparing the data

We can collect all data from the sources; next it will be prepared the data. The main objective is to bring all the data is standardized format. Its carry out future processes.

4.4 Modeling

All the information is used to various develop and behavior model. Like marketing, customer, and modeling data. These consumer lines are used for modeling campaign. It's used to artificial intelligence.

4.5 Evaluation

It is evaluate the proper information is proceed, actually what they understand this model. There are various tools used. Like charts, bar charts, scatter plots.

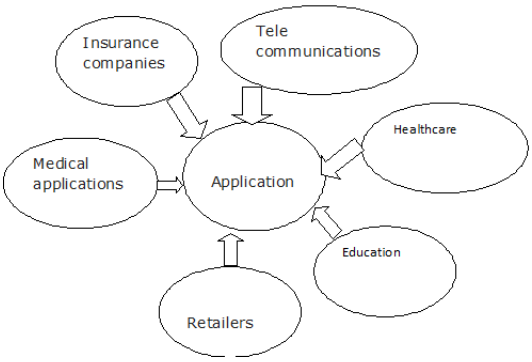
4.6 Deployment

The deployment is applied for new information to previous cause. These model for used in sequence manner. Finally we get new models for getting results in future.

5. Data Mining Applications

- Telecommunications and credit card companies.
- Insurance companies/stock exchanges – apply data-mining techniques to reduce fraud
- Medical applications – to predict the effectiveness of surgical procedures, medical tests or medications.
- Retailers – data mining helps in to identify which promotion and coupon to be applied and which product to be stored.

- Pharmaceutical firms



6. Characteristics of Data Mining

1. Increased quantities of data
2. Provides incomplete data
3. Complicated data structure

- **Large quantities of data**: The huge volume of data is to be analyzed by technically .e.g. satellite information, credit card transactions etc.
- **Incomplete data:** Infinite data
- Various data stored in consequence systems**.**

7. Data Mining Techniques

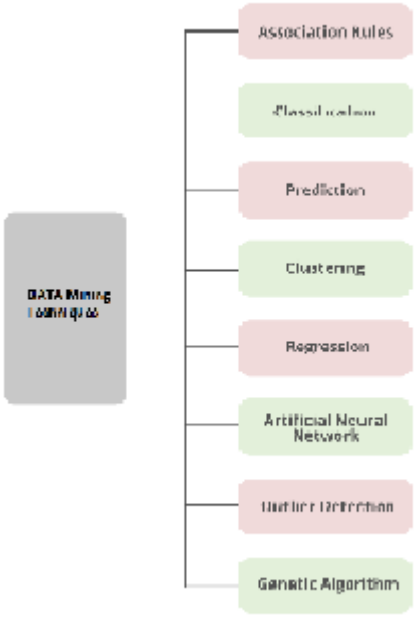
Data mining is used to various algorithms and techniques to convert huge amount of data and its output.

7.1 The association rules:

The association rules for company's sales history to see the products are purchased simultaneously. This information can be stored to plan, promote and forecast.

7.2 Classification:

Its predefined classes to assign to objects. These classes are explaining about the characteristics o items .this data mining techniques for understanding the summarized data.



7.3 Clustering

It's similar to classification. While classification may be result in groups. Such as "shampoo", "conditioner", "soap", and "toothpaste", the clustering may identify groups such as "hair care" and "dental health".

7.4 Neural networks

This process is used to nodes. These nodes are compressed the data (similar to how the human brain is interconnected)

7.5 Regression Analysis

It's the historical information to build the mathematical models to future outcomes. This techniques main aim is to supporting an unknown figure in future on current data.

8. Benefits of Data Mining

- How the problems are identify, and that problem are related how to apply the solution. The data mining helpful to business and more beneficial and efficient.
- The data mining overall process is used new applications.
- Data mining techniques can help company to detect and check cheating.

- ❖ Data mining techniques can help to improve their customer service and increase customer satisfaction.

Data mining techniques also have some disadvantages,

- ❖ This technique can be very difficult but it has special knowledge and experience need implement.
- ❖ The data mining software is work to difficult and it needed training to work.so to select correct data mining tool is very difficult one.
- ❖ This techniques not accurate , so this techniques cause is consequences in conditions

9. How Data Mining Works

This techniques are involves to analyzing large block of information to draw patterns. It's used a various way, like marketing, credit management, fraud, spam email. The data mining process into four steps.

- Start with historical data
- Analyze the data
- Write rules
- Apply the rules

10. Future Scope

Finally the data mining understand - computing, neural networks, machine learning- so more powerful, and less cost to use. Overall data mining's future is it will be improving more and more, kinds of business to used

11. Acknowledgments

We would like to thank to our Chairman, Principal, Vice-Principal and our Colleagues.

12. References

[1] An introduction to data mining Published on November 15th, 2021, By Future Learn.

[2] Definition data mining By Craig Stedman, Industry Editor.

[3] Data Mining – Working, Characteristics, Types, Applications & Advantages, **By** Nidhi Sethi

[4] M. Y. N. Attari, B. Ejlaly, H. Heidarpour and A. Ala, "Application of Data Mining Techniques for the Investigation of Factors Affecting Transportation Enterprises," in *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 7, pp. 9184-9199, July 2022, doi: 10.1109/TITS.2021.3092348.

[5] S. Aggarwal, A. Verma and J. Singh, "Application Based Categorization of Datasets for Implementing Data Mining Techniques," *2021 2nd Global Conference for Advancement in Technology (GCAT)*, Bangalore, India, 2021, pp. 1-7, doi: 10.1109/GCAT52182.2021.9587622

[6] Shruti, R. Soumya, M. Rakhra, N. S. Chintagunti, B. Singh and D. Singh, "Modern Data Mining Approach to Handle Multivariate Data and to Implement Best Saving Services for Potential Investor," *2022 International Mobile and Embedded Technology Conference (MECON)*, Noida, India, 2022, pp. 611-616, doi: 10.1109/MECON53876.2022.9752101.

[7] Data Mining Techniques, Difficulty Level: Medium Last Updated: 02 Feb, 2023.

[8] What Is Data Mining? How It Works, Benefits, Techniques, and Examples By ALEXANDRA TWIN Updated August 02, 2022

[9] Educational data mining and learning analytics: An updated survey Cristobal Romero, Sebastian Ventura First published: 13 January 2020

[10] D. A. Shafiq, M. Marjani, R. A. A. Habeeb and D. Asirvatham, "Student Retention Using Educational Data Mining and Predictive Analytics: A Systematic Literature Review," in *IEEE Access*, vol. 10, pp. 72480-72503, 2022, doi: 10.1109/ACCESS.2022.3188767.

[11] A. Abazari, M. Zadsar, M. Ghafouri, R. Atallah and C. Assi, "A Data Mining/ANFIS and Adaptive Control for Detection and Mitigation of Attacks on DC MGs," in *IEEE Transactions on Smart Grid*, 2022, doi: 10.1109/TSG.2022.3223279.

Authors Profile

1.

2. M.Gayathri awarded the BSc degree from Dhanalakshmi Srinivasan College of Arts & Science for womens in 2003 to 2006, B.Ed degree from Dhanalakshmi Srinivasan College Education in 2008 to 2009. M.Sc degree from Annamalai University in 2009 to 2011.M.Phil degree from Roever College in 2011 to 2012.and finally M.Ed degree from PRIST University in 2012 to 2013. Dec 2013 sewved as an Assistant Professor in the department of Education at Sri Ramakrishna college of Education. June 2015 served as an a Assistant Professor in the department of Computer Applications at Srinivasan College of Arts and Science(co-eed), Perambalur.