iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 4 **Issue:** IV **Month of publication:** April 2016

DOI:

www.ijraset.com

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

International Journal for Research in Applied Science & Engineering Technology (IJRASET)

The Survey of Data Mining Applications and Feature Scope

Poonam B. Patthe
Student, B.E Final year Computer Science and Engineering
H.V.P.M C.O.E.T Amravati, India

Abstract— In this paper we have focused a variety of techniques, approaches and different areas of the research which are helpful and marked as the important field of data mining Technologies. Each place of operation may generate large volumes of data. Corporate decision makers require access from all such sources take strategic decisions .The data warehouse is used in the significant business value by improving the effectiveness of managerial decision-making. In uncertain and highly competitive business environment, the value of strategic information systems such as these are easily recognized however in today's business environment, efficiency or speed is not the only key for competitiveness. This type of huge amount of data's is available and drastically changed in the areas of science and engineering. To analyze, manage and make a decision of such type of huge amount of data we need techniques called the data mining which will transforming in many fields. This paper imparts more number of applications of the data mining and also focuses scope of the data mining which will helpful in the further research.
Keyword: Data mining task, Data mining life cycle, Data mining applications.

I. INTRODUCTION

In the 21st century the human beings are used in the different technologies to adequate in the society. Each and every day the human beings are using the vast data and these data are in the different fields .It may be in the form of documents, may be graphical formats ,may be the video ,may be record .As the data are available in the different formats so that the proper action to be taken. Not only to analyze these data but also take a good decision and maintain the data .As and when the customer will required the data should be retrieved from the database and make the better decision .This technique is actually we called as a data mining or simply KDD(Knowledge Discovery Process).The important reason that attracted a great deal of attention in information technology the discovery of useful information from large collections of data industry towards field of "Data mining" is due to the perception of *"we are data rich but information poor"*. There is huge volume of data but we hardly able to turn them in to useful information and knowledge for managerial decision making in business. To generate information it requires massive collection of data. It may be different formats like audio/video, numbers, text, figures, hypertext formats. To take complete advantage of data; the data retrieval is simply not enough, it requires a tool for automatic summarization of data, extraction of the essence of information stored, and the discovery of patterns in raw data. With the enormous amount of data stored in files, databases, and other repositories, Data mining tools predict future trends and behaviors, helps organizations to make proactive knowledge-driven decisions. Data mining, popularly known as Knowledge Discovery in Databases (KDD), it is the nontrivial extraction of implicit, previously unknown and potentially useful information from data in databases. It is actually the process of finding the hidden information/pattern of the repositories.

II. THE DATA MINING TASK

The data mining tasks are of different types depending on the use of data mining result the data mining tasks are classified as:

A. Exploratory Data Analysis

In the repositories vast amount of information's are available. This data mining task will serve the two purposes
Without the knowledge for what the customer searching ,then
It analyze the data these techniques are interactive and visual to the customer.

B. Descriptive Modeling

It describe all the data, it includes models for overall probability distribution of the data, partitioning into groups and models describing the `relationships between the variables.

C. Predictive Modeling

©IJRASET: All Rights are Reserved

This model permits the value of one variable to be predicted From the known values of other variables.

D. Discovering Patterns And Rules

This task is primarily used to find the hidden pattern as well as to discover the pattern in the cluster. In a cluster a number of patterns of different size and clusters are available. The aim of this task is "how best we will detect the patterns" .This can be accomplished by using rule induction.

III. TYPES OF DATA MINING SYSTEM

Data mining systems can be categorized according to various criteria the classification is as follows:

A. Classification Of Data Mining Systems According To The Type Of Data Source Mined

In an organization a huge amount of data's are available where we need to classify these data but these are available most of times in a similar fashion. We need to classify these data according to its type (maybe audio/video, text format etc)

B. Classification Of Data Mining Systems According To The Data Model

There are so many number of data mining models (Relational data model, Object Oriented data Model etc.) are available each and every model we are using the different data. According to these data model the data mining system classify the data in the model.

C. Classification Of Data Mining Systems According To The Kind Of Knowledge Discovered

This classification based on the kind of knowledge discovered or data mining functionalities, such as characterization, classification clustering, etc.

D. Classification Of Data Mining Systems According To Mining Techniques Used

This classification is according to the data analysis approach Used such as data warehouse-oriented, etc.

IV. DATA MINING LIFE CYCLE

The life cycle of a data mining project consists of six phases. The sequence of the phases is not rigid. Moving back and forth between different phases is always required. It depends on the outcome of each phase. The main phases are:

A. Business Understanding

This phase focuses on understanding the project objectives and requirements from a business perspective, then converting this knowledge into a data mining problem definition and a preliminary plan designed to achieve the objectives.

B. Data Understanding

It starts with an initial data collection, to get familiar with the data, to identify data quality problems, to discover first insights into the data or to detect interesting subsets to form hypotheses for hidden information.

C. Data Preparation

In this stage, it collects all the different data sets and constructs the varieties of the activities basing on the initial raw data

D. Modeling

In this phase, various modeling techniques are selected and applied and their parameters are calibrated to optimal values.

E. Evaluation

In this stage the model is thoroughly evaluated and reviewed. The steps executed to construct the model to be certain it properly achieves the business objectives. At the end of this phase, a decision on the use of the data mining results should be reached.

F. Deployment

The purpose of the model is to increase knowledge of the data, the knowledge gained will need to be organized and presented in a way that the customer can use it. The deployment phase can be as simple as generating a report or as complex as implementing a repeatable data mining process across the enterprise.

©IJRASET: All Rights are Reserved

V. DATA MINING APPLICATIONS

In this section, we have focused some of the applications of data mining and its techniques are analyzed respectively Order.

A. Data Mining Applications In Healthcare

Data mining applications in health can have tremendous potential and usefulness. However, the success of healthcare data mining hinges on the availability of clean healthcare data. In this respect, it is critical that the healthcare industry look into how data can be better captured, stored, prepared and mined. Possible directions include the standardization of clinical vocabulary and the sharing of data across organizations to enhance the benefits of healthcare data mining applications

B. The Data Mining Is Used An Emerging Trends In The Education System In The Whole World

In Indian culture most of the parents are uneducated .The main aim of in Indian government is the quality education not for quantity. But the day by day the education systems are changed and in the 21st century a huge number of universalities are established. As the numbers of universities are established side by side, each and every day a millennium of students are enrolls across the country. With huge number of higher education aspirants, we believe that data mining technology can help bridging knowledge gap in higher educational systems. The hidden patterns, associations that are discovered by data mining techniques from educational data can improve decision making processes in higher educational systems. This improvement can bring advantages such as maximizing educational system efficiency, decreasing student's drop-out rate, and increasing student's promotion rate, increasing student's retention rate in, increasing student's transition rate, increasing educational improvement ratio, increasing success, increasing student's learning outcome, and reducing the cost of system processes. In this current era we are using the KDD and the data mining tools for extracting the knowledge this knowledge can be used for improving the quality of education. The decisions tree classification is used in this type of applications.

C. Data Mining Is Now Used In Many Different Areas In Manufacturing Engineering

When we retrieve the data from manufacturing system then the customer is to use these data for different purposes like to find the errors in the data ,to enhance the design methodology ,to make the good quality of the data ,how best the data can be supported for making the decision . But most of time the data can be first analyzed then after find the hidden patterns which will be control the manufacturing process which will further enhance the quality of the products .Since the importance of data mining in manufacturing has clearly increased over the last 20 years, it is now appropriate to critically review its history and Application

D. Sports Data Mining

The data mining and its technique is used for an application of Sports center**.** Data mining is not only use in the business purposes but also it used in the sports .In the world, a huge number of games are available where each and every day the national and international games are to be scheduled, where a huge number of data's are to be maintained .The data mining tools are applied to give the information as and when we required. The open source data mining tools like WEKA is used for sport. This means that users can run their data through one of the built-in algorithms, see what results come out, and then run it through a different algorithm to see if anything different stands out. As these programs' are available in the form of open source in nature, that's why the users are frequently to modify the source code, so that other can get the updated information. In the sports world the vast amounts of statistics are collected for each player, team, game, and season. In the game sports the data's are available in the form of statistical form where data mining can be used and discover the patterns, these patterns are often used to predict the future forecast. Data mining can be used for prediction of performance, selection of players, coaching and training and for the strategy planning. The data mining techniques are used to determine the best or the most optimal squad to represent a team in a team sport in a season, tour or game.

VI. THE SCOPE OF DATA MINING

Data mining derives its name from the similarities between searching for valuable business information in a large database given database of sufficient size and quality; data mining technology can generate new business opportunities by providing the capabilities that is automated prediction of trends and behaviors.

VII. CONCLUSION

Most of the previous studies on data mining applications in various fields use the variety of data types range from text to images and stores in variety of databases and data structures. The different methods of data mining are used to extract the patterns and thus the knowledge from this variety databases. Selection of data and methods for data mining is an important task in this

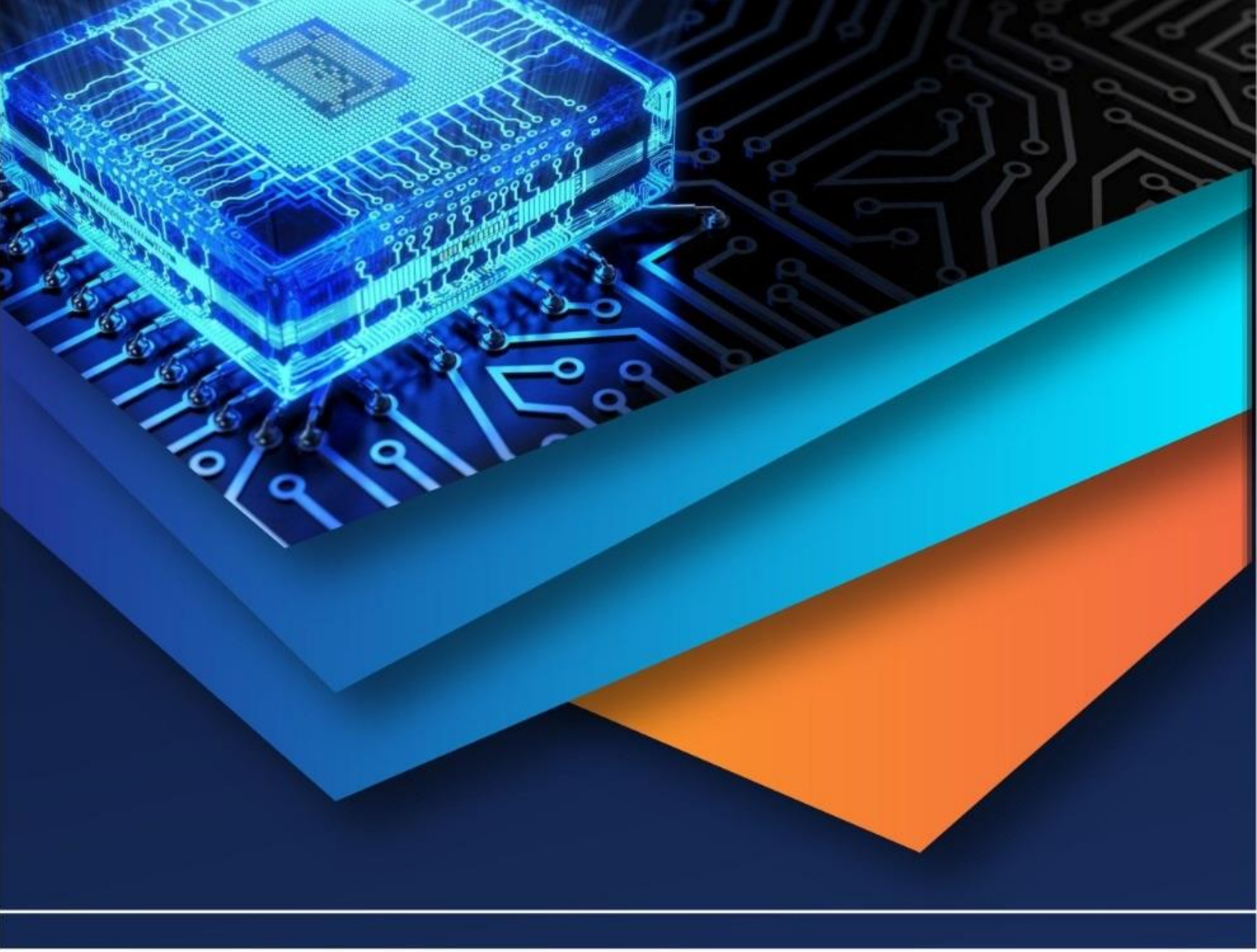
©IJRASET: All Rights are Reserved

process and needs the knowledge of the domain. Several attempts have been made to design and develop the generic data mining system but no system found completely generic. Thus, for every domain the domain expert's assistant is mandatory. The domain experts shall be guided by the system to effectively apply their knowledge for the use of data mining systems to generate required knowledge.

REFERENCES

[1] Neelamdhab padhy, Dr. Pragnyaban Mishra, and Rasmita Panigrahi Asst. Professor, Gandhi Institute of engineering and Technology, GIET, Gunupur.
[2] Associate Professor ,Gandhi Institute of Engineering and Technology, GIET, Gunupur Lecturer, Gandhi Institute of Engineering and Technology, GIET, Gunupur
[3] Introduction to Data Mining and Knowledge Discovery, Third Edition ISBN: 1-892095-02-5, Two Crows Corporation, 10500 Falls Road, Potomac, MD 20854 (U.S.A.), 1999.
[4] Larose, D. T., "Discovering Knowledge in Data: An Introduction to Data Mining", ISBN 0-471-66657-2, ohn Wiley & Sons, Inc, 2005.
[5] Dunham, M. H., Sridhar S., "Data Mining: Introductory and Advanced Topics", Pearson Education, New Delhi, ISBN: 81-7758-785-4, 1st Edition, 2006
[6] Chapman, P., Clinton, J., Kerber, R. Khabaza, T., Reinartz, T., Shearer, C. and Wirth, R... "CRISP-DM Step-by-step data mining guide, NCR Systems Engineering Copenhagen (USA and Denmark), DaimlerChrysler AG (Germany), SPSS Inc. (USA) and OHRA Verzekeringenen Bank Group B.V (The Netherlands), 2000".
[7] Fayyad, U., Piatetsky-Shapiro, G., and Smyth P., "From Data Mining to Knowledge Discovery in Databases Magazine, American Association for Artificial Intelligence, 1996.
[8] Tan Pang-Ning, Steinbach, M., Vipin Kumar. "Introduction to Data Mining", Pearson Education, New International Journal of Computer Science, Engineering and Information Technology (IJCSEIT), Vol.2, No.3, June 2012.

©IJRASET: All Rights are Reserved



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)