IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

ISO 3297:2007 Certified ⁜ Impact Factor 7.39 ⁜ Vol. 11, Issue 5, May 2022

DOI: 10.17148/IJARCCE.2022.115176

Research on Data Mining

Chandrakant A. Zade, Prof. Vijay Rakhade, Prof. L. Yadav

Dept. of computer Science and Engineering, Sscet, Bhadrawati, India

Abstract: The world is deal with various kind of datal like social media data, medical data, stock market data, environmental data, scientific data, financial data, Mathematical data .So analyzing and summarizing this data manually is impossible because of incredible increase in data due to daily uses of internet and information sharing . This research investigate the fundamental of data mining, scope of data mining and develop new techniques for assimilate uncertainty management in data mining.

Keywords: Data mining , scientific data ,deal , investigate, assimilate

I. INTRODUCTION

Data mining terms refers to extracting useful information from vast amount of data. Many other terms are used to clarify data mining, such as Knowledge extraction, data analysis knowledge mining from database, data dredging, data /pattern analysis and data archaeology, In current era we can easily say that data mining is essential step in the process of knowledge discovery in database. In this paper we coverd lot of topic about data mining.

WHY DATA MINING ?

- The explosive growth of data from terabytes to petabytes
- Data collection and data availability
- Major source of abundant data
- Market analysis and management
- Risk analysis and management

OVERVIEW OF DATA MINING

Due to development of information technology industry has generated large amount of database and huge data in various area. The research in database and it industry has give new approach to store and manipulate this precious data for further decision making.

so the data mining is a process that is used to search a particular data from huge amount of data effectively . the main aim of these technique is to find patterns that were previously unknown .Due to this pattern we can make certain decision for development of their business.

Different Data Mining Methods : There are so many method used for data mining.

- Association
- Classification
- Prediction
- Clustering Analysis
- Pattern Tracking
- Decision Trees
- Outlier Analysis
- Neural Network

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

ISO 3297:2007 Certified ⁜ Impact Factor 7.39 ⁜ Vol. 11, Issue 5, May 2022

DOI: 10.17148/IJARCCE.2022.115176

1) Association : This method is used to find a correlation between two or more item by identifying the hidden pattern in data set and hence also called relation analysis .it also used in market basket analysis to predict the behaviour of the customer.

2) Classification : Classification method is used to distinguish the item in the data set into groups. Classification method is also help to predict the behaviour of entities within the group accurately.

3) Prediction : Prediction method is used to predict the future on basis of past and present trends or data. Its mostly used to combine other mining method such as pattern matching ,classification ,trend analysis and relation.

4) Clustering Analysis :it is also similar to classification ,but cluster are made on similarities of data items. Different groups have unrelated object and its knowns as data segmentation .

5) Pattern tracking : Pattern tracking is used to identify patterns that frequently occur over a certain period of time.

6) Decision Trees : A decision tree is a tree structure .
Where ,

- Each internal node represent a test on the attribute.
- Branch denotes the result of the test.
- Terminal nodes hold class label.
- Top node is root node which has a simple question that has two or more answer.

7) Outlier Analysis : Outlier analysis method identifies the data item that do not comply with the expected pattern and this unexpected data item are considered as outlier or noise

8) Neural Network : Neural network is based on biological neural network.neural network is a collection of neuron like processing units with weighted connection between them.

Data Mining tools :

- MonkeyLearn : it is a machine learning platform that are specializes in text mining.its provide userfriendly interface.
- Rapidminer :it is free open source data science platform that feature hundred of algorithm for data preparation,machine learning,deep learning, text mining etc.
- Oracle Data Mining : it is component of of oracle advance analytic that enables data analyst to build and implement predictive model.
- IBM SPSS Modeler : it is data mining solution ,which allows data scientist to speed up and visualize data mining process.

Three step involed are:

1) Exploration

2) Pattern Identification

3) Deployment

Exploration : In the first step , data is cleaned and transformed into another form, and important variable and then nature of data based on the problem are determined.

Pattern Identification : Once data is explored, refined and defined for the specific variable the second step is to from pattern identification . .

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

ISO 3297:2007 Certified ⁜ Impact Factor 7.39 ⁜ Vol. 11, Issue 5, May 2022

DOI: 10.17148/IJARCCE.2022.115176

Deployment : pattern are deployed for desired outcome.

Data Mining Application

- Research
- Education sector
- Transportation
- Busisness Transaction
- Market Basket Analyst
- Intrusion detection
- Finance and Banking Sector
- Scientific Analysis
- Insurance and Healthcare
- Retail sector

Data mining Advantages

- Marketing : To creat model, marketing companies data mining.
- Finance : Data extraction provide financial institution information on loans and credit report, data can determine good or bad credit by creating a model for historical customers.
- Increases Brand Loyalty : In marketing campaign , mining technique are used to understanding their own coustomers needs and habits.
- To predict future trend : With the help of data mining we can easily detected future trends.

CONCLUSION

Data mining initially generated a great deal of excitement and press coverage and as is common with "new technologies" , overblown expections. However as data mining has begun to nature as a discipline , its method and technique have not only proven to be useful, but have begun to be accepted by the wider community of data analyst.

This paper gives a general introduction of data mining, the process of extracting useful data from huge amount of data. We also discuss about future scope , application and advantages of data mining. It is also shown that data mining technology can be used in many area in real life including biomedical and DNA data analysis.

All knowledge worker in our information society , particularly those need to make informed decision based on data , should have at least a basic familiarities with data mining . future research will involve the development of new techniques for incorporating uncertainty management in data mining.

REFERENCES

1. Fayyd, Usama, Gregory Piatestsky -shapiro, Padhraic smyth, from data mining to knowledge Discovery in Databases,1996.
2. Jiawei Han , Micheline Kamber, Data mining :concept and technique , London :Academic press,5 2001.
3. Kantardzic , Mehmed ,Data mining : concept, model, methods, and Algorithm,New York : john Wiley &sons Inc publishes 2003.
4. Michael chau, Reynold Cheng, Ben Pao, Uncertain Data mining : A New Reasearch Direction , Introduction, 2005.
5. Mika sato ,Yoshiharu sato, L.C. jain , J kacprzyk, fuzzy clustering models and Application . UK:physica -verlag Heidelberg,1999
6. Mr.Vijay M.Rakhade,"Reducing Routing Distraction in IP Network using Cross-Layer Methodology" ICRTEST Volume 5 Issue 1(21-22 January 2017).
7. Lowlesh Nandkishor Yadav, "predictive Acknowledgement using TRE system to reduce cost and Bandwidth" IJRECE VOL. 7 ISSUE 1(JANUARY-MARCH 2019).
8. Hirali Devendra Wadaskar ,"Research on Association Rule Mining Algorithms",IJARCCE Vol.11, Issue 5,May 2022.

