VIVA-Tech International Journal for Research and Innovation *Volume 1, Issue 5 (2022)*
ISSN(Online): 2581-7280
VIVA Institute of Technology
10th National Conference on Role of Engineers in Nation Building – 2022 (NCRENB-2022)

VIVA-TECH INTERNATIONAL JOURNAL FOR RESEARCH AND INNOVATION
ANNUAL RESEARCH JOURNAL
ISSN(ONLINE): 2581-7280

A Review on Data Mining from Past to the Future

Rohit.U.Desai[1],Priyanshu Yadav[2]
(MCA, VIVA Institute of Technology / Mumbai University, India)
(MCA, VIVA Institute of Technology / Mumbai University, India)

Abstract : *Data Mining is that the data discovery method by analyzing the massive volumes of {information} from numerous views and summarizing it into helpful information. Advancements in Statistics, Machine Learning, AI, Pattern Recognition and Computation capabilities have evolved this day's data processing applications and these applications have enriched the varied fields of human life together with business, education, medical, scientific etc. therefore this paper discusses the varied enhancements within the field of knowledge mining from past to this and explores the longer term trends.*

Keywords *-_Current Trends, Data Discovery in Databases, data processing, Future Trends. Historical Trends, Heterogeneous information.*

I. INTRODUCTION

The process of extracting data to spot patterns, trends, and helpful knowledge that may permit the business to require the data-driven call from Brobdingnagian sets of knowledge is termed data processing. In different words,we will say that data {processing} is that the process of work hidden patterns {of information |of knowledge| of knowledge} to varied views for categorization into helpful data, that is collected and assembled particularly areas like knowledge warehouses, economic analysis, data processing algorithmic program, serving to higher cognitive process and different knowledge demand to eventually cost-cutting and generating revenue.

The data collected from completely different applications need correct mechanism of extracting data /information from massive repositories for higher cognitive process. Data discovery in databases (KDD), typically referred to as data processing, aims at the invention of helpful data from massive collections of knowledge. The core functionalities {of knowledge |of knowledge| of information} mining applying numerous ways and algorithms to get and extract patterns of hold on data. From the last 20 years data processing and data discovery applications have gotten an expensive focus thanks to its significance in higher cognitive process, and it's become a vital element innumerous organizations.

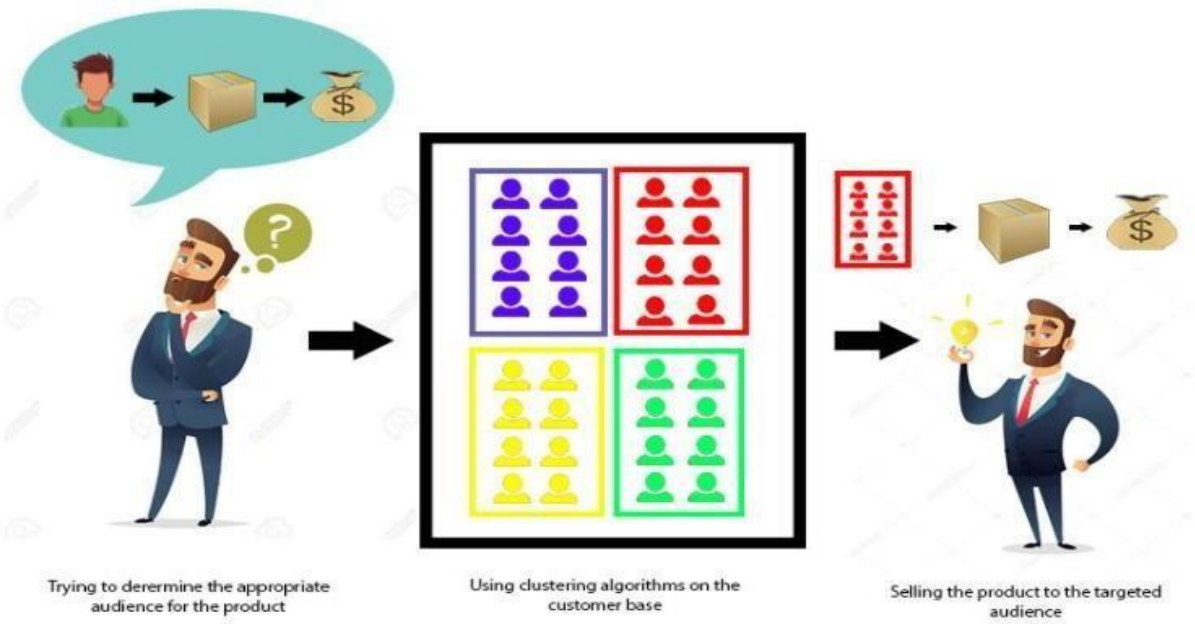
The field of knowledge mining are prospered and posed into new areas of human life with numerous integrations and advancements within the fields of Statistics, Databases, Machine Learning, Pattern Reorganization, AI and Computation capabilities etc.

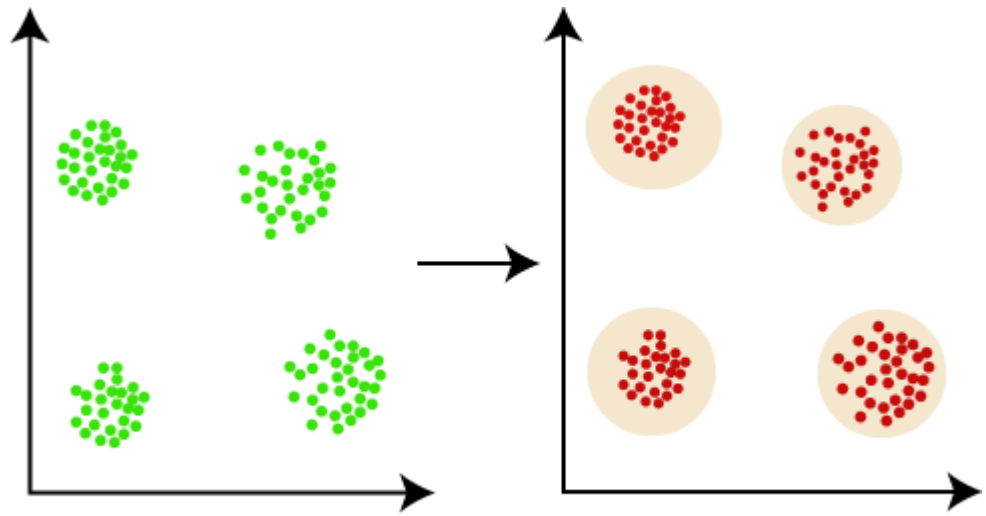
II. HISTORICAL TRENDS OF DATA MINING

The building blocks of knowledge mining is that the evolution of a field with the confluences of varied disciplines, which has management systems (DBMS), Statistics, computer science (AI), and Machine Learning (ML). The time of knowledge mining applications was planned within the year1980 primarily by research-driven tools
focused on single tasks [3]. The first day's data processing trends area unit as underneathData Trends:
In initial days, data processing algorithms work best for numerical knowledge collected from one knowledge base, and numerous data processing techniques have evolved for flat files, ancient and relative databases wherever the info is keep in tabular illustration.
Computing Trends:
The field of knowledge mining has been greatly influenced by the event of fourth generation programming languages and numerous connected computing techniques. In, youth of knowledge mining most of the algorithms utilized solely applied math techniques.

III. CLUSTERING IN DATA MINING

Clustering is AN unsupervised Machine Learning-based algorithmic program that contains cluster|a gaggle|a bunch} of information points into clusters in order that the objects belong to an equivalent group.
Clustering helps to splits knowledge into many subsets. every of those subsets contains knowledge almost like oneanother, and these subsets ar known as clusters.





Showing four clusters formed from the set of unlabeled data

What is clustering in Data Mining?

- Clustering is that the methodology of changing a bunch of abstract objects into categories of comparable objects..
- Clustering may be a methodology of partitioning a group of information or objects into a group of

great subclasses known as clusters..

Applications of cluster analysis in data mining:

- In several applications, cluster analysis is wide used, like knowledge analysis, marketing research, pattern recognition, and image process.
- It assists marketers to search out totally different teams in their shopper base and supported the buying patterns. they will characterize their client teams.

clustering used in data mining:

Clustering ANalysis has been an evolving drawback in data processing thanks to its type of applications. the appearance of assorted knowledge cluster tools within the previous couple of years and their comprehensive use ina very broad vary of applications, together with image process, procedure biology, mobile communication, medicine, and political economy, should contribute to the recognition of those algorithms..

1. Scalability:

Scalability in cluster implies that as we tend to boost the quantity of information objects, the time to perform cluster ought to around scale to the complexness order of the algorithmic program..

2. Interpretability:

The outcomes of cluster ought to be explainable, explicable, and usable..

3. Discovery of clusters with attribute shape:

The cluster algorithmic program ought to be able to realize whimsical form clusters. they ought to not be restricted to solely distance measurements that tend to get a spherical cluster of little sizes.

4. Ability to handle differing kinds of attributes:

Algorithms ought to be capable of being applied to any knowledge like knowledge supported intervals (numeric), binary knowledge, and categorical knowledge.

IV. CURRENT TRENDS

The field of information mining has been growing because of its huge success in terms of broad-ranging application achievements and scientific progress, understanding. numerous data processing applications are with success enforced in numerous domains like health care, finance, retail, telecommunication, fraud detection and risk analysis...etc .

1.1 Mining the Heterogeneous data

The following table depicts numerous presently utilized data processing techniques and algorithms to mine the assorted information formats in numerous application areas.

.

1.2 Utilizing the Computing and Networking Resources

Parallel data processing applications have evolved victimisation the Parallel computing, typical parallel data processing applications use the Apriori rule.

1.3 Research and Scientific Computing Trends

The explosion within the quantity information from several scientific disciplines, like physics, remote sensing, bioinformatics, combinatorial chemistry, medical imaging, and experimental physics square measure standardization to numerous data processing techniques, to search out out helpful information.

1.4 Business Trends

Today's business should be additional profitable, react faster and supply top quality services that ever before. With these varieties of expectations and constraints, data processing becomes a elementary technology in facultative customer's transactions additional accurately. data processing techniques of classification, regression, and cluster analysis square measure used for in current business trends [17].

V. FUTURE TRENDS

Due to the large success of varied application areas of knowledge mining, the sector of knowledge mining has been establishing itself because the major discipline of technology and has shown interest potential for the long run developments.

- Standardization of knowledge mining languages
- Data preprocessing
- Complex objects of knowledge
- Computing resources
- Web mining
- Scientific Computing
- Business information

1.Standardization of knowledge mining languages:
There square measure varied data processing tools with completely different syntaxes, thence it's to be standardized for creating convenient of the users. data processing applications needs to concentrate additional in standardization of interaction languages and versatile user interactions.

2.Information Preprocessing:
To identify helpful novel patterns in distributed, large, complicated and temporal information, data processing techniques needs to evolve in varied stages.

3.Complicated object of data:
Data mining goes to penetrate altogether fields of human life, the presently accessible data processing techniques square measure restricted to mine the normal sorts of information solely, and in future there's a potentiality for data processing techniques for complicated information objects like high dimensional, high speed information streams, sequence, noise within the statistic, graph, Multi-instance objects, Multi-represented objects and temporal information.

4.Computing Resources
The up to date developments in high speed property, parallel, distributed, grid and cloud computing has exhibit new challenges for data processing.

5. Internet mining:
The development of World Wide internet and its usage grows, it'll still generate ever additional content, structure, and usage information and also the price of internet mining can keep increasing.

VI. CONCLUSION

In this paper we tend to shortly reviewed the varied data processing trends from its origination tothe long run. This review would be useful to researchers to target the varied problems with data processing.
The final goal of knowledge mining is that the prediction of human behavior, and is out and away the foremost common business application; but this will simply be sculptural to fulfill the target of detection and deterrence of criminals. the automated analysis of databases and also the recognition of necessary trends and activity patterns.
Sophisticated data processing and AI tools ar currently obtainable to the enforcement communities. These tools ar very powerful, fast, and comparatively simple to use. data processing supports increased deciding and analysis, and could be a powerful tool that may be wont to address the big volume of crime info presently facing all agencies.

ACKNOWLEDGEMENTS

I would like to express my special thanks of gratitude to our Head of Department – Prof Mrs. Chandani Patel Ma'am, research coordinator Prof. Mrs. Neha Lodhe Ma'am and our research guide Prof. Mrs. Sonia Dubey Maam who gave us the opportunity to this wonderful research on the topic A Review on Data Mining from Past to the Future , which also helped me in doing a lot of research and I came to know about so many new things. I am really thankful to them. Secondly, I would like to thank my research group's members in helping me carrying out my tasks at each phase of the development and contributing for the same equally and willingly, I also thank my parents and friends who helped me a lot in finalizing this research within limited time frame.

REFERENCES

1. "Data Mining Curriculum". ACM SIGKDD. 2006-04-30. Retrieved 2014-01-27.
2. Clifton, Christopher (2010). "Encyclopædia Britannica: Definition of Data Mining". Retrieved 2010- 12-09.
3. Hastie, Trevor; Tibshirani, Robert; Friedman, Jerome (2009). Archived from the original on 2009- 11-10. Retrieved 2012-08-07.
4. Han, Jaiwei; Kamber, Micheline; Pei, Jian (2011). Data Mining: Concepts and Techniques (3rd ed.). Morgan Kaufmann. ISBN 978-0-12-381479-1.
5. Fayyad, Usama; Piatetsky-Shapiro, Gregory; Smyth, Padhraic (1996). "From Data Mining to Knowledge Discovery in Databases" (PDF). Retrieved 17 December 2008.
6. Bouckaert, Remco R.; Frank, Eibe; Hall, Mark A.; Holmes, Geoffrey; Pfahringer, Bernhard; Reutemann, Peter; Witten, Ian H. (2010). "WEKA Experiences with a Java open-source project". Journal of Machine Learning Research. 11: 2533–2541.
7. Olson, D. L. (2007). Data mining in business services. Service Business, 1(3), 181–193. doi:10.1007/s11628-006-0014-7
8. Lovell, Michael C. (1983). "Data Mining". The Review of Economics and Statistics. 65 (1): 1–12. doi:10.2307/1924403. JSTOR 1924403.
9. Charemza, Wojciech W.; Deadman, Derek F. (1992). "Data Mining". New Directions in Econometric Practice. Aldershot: Edward Elgar. pp. 14–31. ISBN 1-85278-461-X.
10. Piatetsky-Shapiro, Gregory; Parker, Gary (2011). "Lesson: Data Mining, and Knowledge Discovery: An Introduction". Introduction to Data Mining. KD Nuggets. Retrieved 30 August 2012.
11. Fayyad, Usama (15 June 1999). "First Editorial by Editor-in-Chief". SIGKDD Explorations. 13 (1): 102. doi:10.1145/2207243.2207269. S2CID 13314420. Retrieved 27 December 2010.
12. Coenen, Frans (2011-02-07). "Data mining: past, present and future". The Knowledge Engineering Review. 26 (1): 25–29. doi:10.1017/S0269888910000378. ISSN 0269-8889. S2CID 6487637.
13. Kantardzic, Mehmed (2003). Data Mining: Concepts, Models, Methods, and Algorithms. John Wiley & Sons. ISBN 978-0-471-22852-3. OCLC 50055336.