International Journal for Research in Engineering Application & Management (IJREAM)
ISSN : 2454-9150 Vol-04, Issue-06, Sep 2018

Analysis of Data mining Applications and Technologies

T. SHRAVYA, G. Narayanamma Institute of Technology and science for women, M.Tech (CSE), shravyarao64@gmail.com.

Mr. T. RAJESH, Asst. Prof, Dept. of CSE, G. Narayanamma Institute of Technology and science for women, rajesht531@gmail.com

Abstract: Day by day with the increase in number of devices data storage also increases automatically. These devices generate data stored in different storage areas like databases, big data, and cloud. Different devices or user's stores' data in storage areas with huge volume, from this storage place, extracting useful information is called data mining. This paper talks about different uses of information mining and different innovations utilized as a part of information mining.

Key words: Data Mining, Big Data, Data Base, Cloud, Data volume, Data Mining Technologies.

I. INTRODUCTION

Data mining is used to extract the learning or data from a lot of information which is stored in different heterogeneous information base [4]. In data mining data is collected from various resources like databases, big data etc. and it produces meaningful patterns to the users. This is also called as the knowledge mining or knowledge discovery process. Knowledge discovery process consists of the following stages a) Selection b) Preprocessing c) Transformation d) Data mining e) Interpretation or Evaluation.

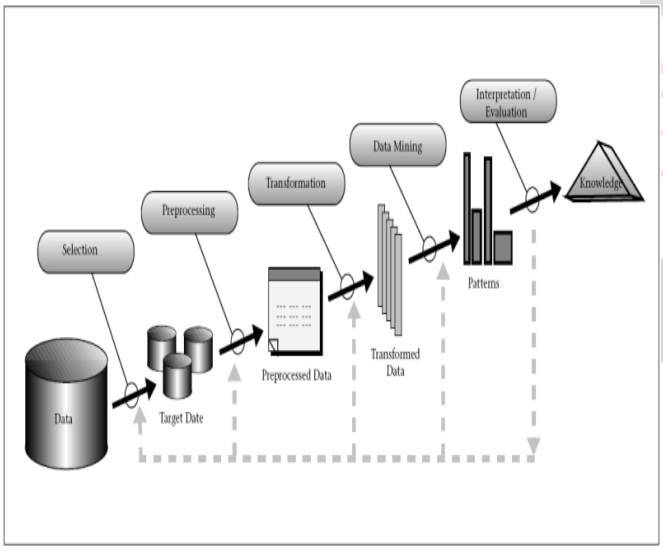

Figure 1: Process of KDD

Data is selected from various data sources where extraction process required is done in the selection process. From this stage the target data (required data) is taken and is prepossessed in the second stage i.e. this is a data cleaning process where unwanted data is removed and the needed data is provided. This preprocessed data is transferred into new patterns and on these new patterns data mining techniques are applied. These techniques are used to obtain the required results. The last step is used to interpret the results to get desired meaningful information.

II. APPLICATIONS OF DATA MINING

Data mining methods are connected to different fields like Broadcast communications, Educational organizations, Marketing, Fraud recognition process, Medical segment, the Share trading system etc. Here we clarified these zones where it's connected [3].

Banking sector: Banking sector or financial sector uses this data mining process extensively. Data mining is applied to extract the customer information, transaction details, debit and credit card frauds, analyzing risk and to estimate the profits etc.

Telecommunications: Telecommunications have more number of customers.The data set has to be updated frequently as the size is large. Data mining is used to improve the network communication, detect frauds, increase the business, produce quality data to the customer.

Data mining in agriculture: Data mining in agriculture is used to predict the weather reports, yearly crop yield analysis, crop production. We use techniques like KNN (K Nearest neighbor), K-Means and artificial networks so as to acquire the desired data.

Decision Tree: A decision tree follows a tree like structure where it consists of parent nodes and sub nodes and it comes under predictive data model [11]. In this a node represents the test on attribute and the sub nodes represent the outcome of those test nodes. Each node attribute takes a decision like true or false and based on this the traversal takes place.

 DOI : 10.18231/2454-9150.2018.0782

Education sector: Data mining in education sector is used to evaluate the student's performance, allotting student information, result analysis, allotting ranks for the educational institutions, providing online data and online classes according to student's choice.

Data Mining in Market Basket Analysis: These procedures in perspective of shopping database. A definitive objective of market bin examination is finding the items that clients purchase together most of the times. The stores can utilize this data by putting these items in nearness of each other and making them more noticeable and open for clients at the season of shopping.

III. DATA MINING TECHNOLOGIES

TECHNIQUES OF DATA MINING:

A data mining process is used to create either descriptive model or predictive model. These models can be accomplished by different kinds of information mining strategies. The predictive models have different algorithms like a) Classification b) Regression c) Time series analysis d) Prediction. Descriptive models have different algorithms like a) Clustering b) Summarization c) Association Rules d) Sequence Discovery [1]. Figure 2 demonstrates the data mining Predictive and Descriptive strategies.

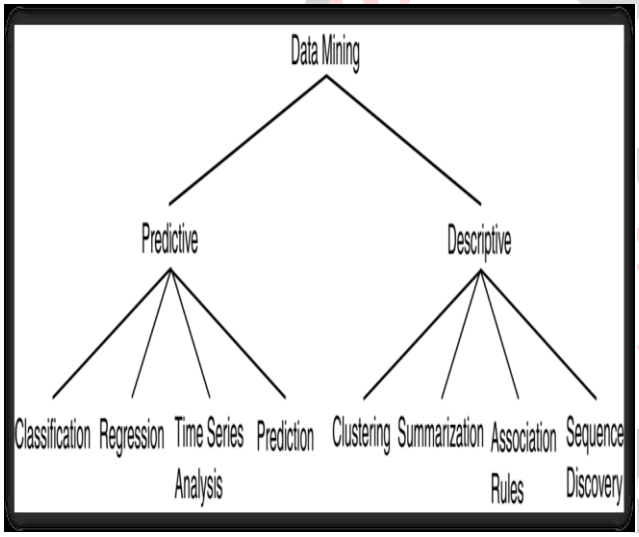

Figure 2 Data mining techniques predictive and descriptive.

Technologies: We use various technologies in data mining to improve the data efficiency [2]. These technologies are like artificial neural networks, Decision Tree, Genetic Algorithm, Classification and Clustering.

Clustering: Clustering is a way of isolating the items into a group of comparative articles. In clustering we group the data based on the similarity and assigning labels to them. Clustering has numerous application zones like market examine, design acknowledgment, information examination, and picture handling. It has numerous strategies to make bunches like Parceling Method, Hierarchical Technique, Density-based Method, Grid-Based Method, Model-Based Method, and Limitation based Method. We also have different types of clustering methods such as "grid based method, partitioning methods, model based methods [9].

Genetic Algorithm: Genetic Algorithms are used to produce child chromosomes from a set parent chromosomes. It uses three operations ie., selection, cross over, mutation [10,8]. It has a population which is selected based on the selection process then the cross over is applied to select the child populations, from these selected child populations mutation points are applied. Then based on these every new child chromosome is generated and processed. Genetic algorithms are used to predict the results by using rule induction process and K-nearest neighbor algorithms.

Classification: Classification is applied to allot labels to the unclassified data. We use various methods to classify data ie., supervised and unsupervised classification techniques. We have many other classification techniques in data minin like "classification by neural network, classification based on associations, SVM (support vector mechines), Bayesian classification, classification by decision tree [7]".

Association method: This method is used to find the frequent data items from large data set. We have few types of association rules those are "multi level, multi dimensional, quantitative association rules [6,7]. These rules are mainly used in business applications where a decision is to be taken.

Neural Networks: A neural network is set of neurons (input and output units) connected to each other with weight as a unit of each connection [5]. The main use of neural network is to derive meaningful data from very complicated data. This will produce continuous inputs with continuous outputs.

IV. CONCLUSION

Data mining is a process of taking out the useful patterns from bulk storage of data. In this process whatever the data we have collected is called as the raw data where this raw data is preprocessed and grouped into clusters such that it is useful to the user.The raw data is taken as the input and the processed patterns is produced as the output to the user. This paper contemplates in insight about the uses of data mining and the different sorts of information mining advancements we utilize. Here we have clarified about the predictive data mining techniques and descriptive data mining methods.

REFERENCES

[1] Sang Jun Lee, Keng Siau. (2001). A review of data mining techniques. Industrial Management & Data Systems. 101 (1), 41-46.

 DOI : 10.18231/2454-9150.2018.0782

[2] Sadiq Hussain. (2017). Survey on Current Trends and Techniques of Data Mining Research. London Journal of Research in Computer Science and Technology. 17 (1), 1-15.

[3] Deepashri.K.S, Ashwini Kamath. (2017). Survey on Techniques of Data Mining and its Applications. International Journal of Emerging Research in Management &Technology. 6 (2), 198-201.

[4] Mohammed J. Zaki Wagner Meira Jr.. (2013). Data Mining and Analysis: Fundamental Concepts and Algorithms. Cambridge University Press.

[5] Priyanka Gaur. (2017). Neural Networks in Data Mining . International Journal of Electronics and Computer Science Engineering . 1 (3), 1449-1453.

[6] Irina Tudor . (2008). Association Rule Mining as a Data Mining Technique . BULETINUL Universităţii Petrol – Gaze din Ploieşti . LX (1), 49-56.

[7] Bing Liu Wynne Hsu Yiming Ma. (1998). Integrating Classification and Association Rule Mining. KDD-98 Proceedings.

[8] Siew Mooi Lim, Abu Bakar Md. Sultan, Md. Nasir Sulaiman, Aida Mustapha, and K. Y. Leong. (2017). Crossover and Mutation Operators of Genetic Algorithms. International Journal of Machine Learning and Computing. 7 (1), 9-12.

[9] Dr. Sankar Rajagopal . (2011). CUSTOMER DATA CLUSTERING USING DATA MINING TECHNIQUE. International Journal of Database Management Systems (IJDMS). 3 (4), 1-11.

[10] Robert E. Marmelstein. (1997). Application of Genetic Algorithms to Data Mining. :MAICS-97 Proceedings. 0 (0), 53-57.

[11] Nikita Jain1, Vishal Srivastava2. (2013). DATA MINING TECHNIQUES: A SURVEY PAPER. IJRET: International Journal of Research in Engineering and Technology. 2 (11), 116-119.

 DOI : 10.18231/2454-9150.2018.0782