

## DATA MINING TECHNIQUES AND ITS APPLICATIONS

<sup>1</sup>B.Sundarraaj, <sup>2</sup>Dr.K.P.Kaliyamurthie

<sup>1</sup>Asst.Professor, Dept. of CSE, BIST, BIHER, Bharath University ,  
Chennai, Tamil Nadu, India,

<sup>2</sup>Professor and Head, Dept. of CSE, BIST, BIHER Bharath University ,  
Chennai, Tamil Nadu, India,

Email: <sup>1</sup>sundarraajboobalan@gmail.com, <sup>2</sup>kpkailamurthie@gmail.com

**Abstract:** Data mining is the process of finding useful information from the database. Data mining tool is used to predict future. Data mining also known as Knowledge Data Discovery .Data mining techniques can be used in industry,education,commerce and in the field of agriculture. The paper deals with Data mining and its various techniques related to that.

**Keywords:** Data Mining, Knowledge Discovery in Databases(KDD), Information, Large Datasets

### 1. Introduction

Data mining [1]is the process of finding useful information from the huge amounts of data and to find the pattern using mathematical and statistical techniques. It uses the various techniques from neural networks ,machine learning, genetic algorithm and pattern recognition.

**Data Mining Issues:**There are various issues [2,3,4] related with Data Mining

**Human interaction:** Data mining needs technical experts to interpret the data.[5-7]

**Over-fitting:** It has to consider whether the model is fit for future database.[8-12]

**Outliers:** If the model includes outliers it will not work properly for the database.[13]

**Data integrity:** Data mining consists of integrated data. [14]

**Interpretation of results:** Data mining requires experts to interpret the results.[15]

**Visualization of results:** Visualization techniques are useful to visualize the results.[16]

**Large datasets:** The model created for data mining problem for small data sets should be designed to manage large data sets.

**High dimensionality:** All the application doesn't require entire attributes from the database. It needs some of the attributes from the database. So it is necessary to consider the issue of attribute reduction .Reducing number of attributes for analysis is known as Dimensionality reduction.

**Multimedia data:** Data mining algorithm should support multimedia data.While generating the model this issue needs to be considered.[17-19]

### Relational or Multidimensional databases:

One of the technical issue is whether it uses relational or multidimensional database.In relational database the data is stored in the form of tables where as in multidimensional data model the data is stored in the form of cubes.

**Noisy data:** The wrong data or incorrect data should be corrected before applying the data mining algorithms.

**Irrelevant data:** Irrelevant data from the data base for the data mining analysis need to be removed.

**Missing data:** Missing data is one of the issue need to be considered before applying the model.The missing data is estimated based on the estimation.

**Changing data:** While generating data mining model it has to consider dynamic data because data may not be always static .It may change.

## 2. Application

Data mining applications is an important issue. While generating the Data mining model it has to consider where the output of the data mining model has to be applied.

### Data Mining Techniques:

Data mining techniques includes the following

- Classification
- Clustering
- Prediction
- Association Rule Mining

### Classification

Classify the data into one of the predefined classes or groups.

### Clustering

Clustering is the process of grouping similar objects into one group.

### Prediction:

It is used to predict the future and to find the relationship between the dependent and independent variables.

## 3. Data Mining Applications

Data Mining applications include Finance and Banking to analyze Customer activities. Data mining can be also used in medical care and insurance applications. Data mining can also be used in transportation to analyze the distribution. But the privacy, security and misuse of information is a disadvantage of Data mining.

## 4. Conclusion

Data mining is the process of finding useful information from the data base. While constructing data mining model the above issues need to be considered. This paper discusses about the concepts of data mining and its various applications.

### References

- [1] Agrawal R., Shrikant R., (1994), Fast algorithms for mining association rule. In the proceeding of 20th international conference on VLDB, pp487-499.
- [2] Arun K. Pujari, Data Mining Techniques.
- [3] Ilayaraja, K., Ambica, A., Spatial distribution of groundwater quality between injambakkam-thiruvannamiyur areas, south east coast of India, Nature Environment and Pollution Technology, v-14, i-4, pp-771-776, 2015.
- [4] Gopinath, S., Sundararaj, M., Elangovan, S., Rathakrishnan, E., Mixing characteristics of elliptical and rectangular subsonic jets with swirling co-flow, International Journal of Turbo and Jet Engines, v-32, i-1, pp-73-83, 2015.
- [5] Kerana Hanirex, D., Kaliyamurthie, K.P., Kumaravel, A., Analysis of improved tdt algorithm for mining frequent itemsets using dengue virus type 1 dataset: A combined approach, International Journal of Pharma and Bio Sciences, v-6, i-2, pp-B288-B295, 2015.
- [6] Thooyamani, K.P., Khanaa, V., Udayakumar, R., Efficiently measuring denial of service attacks using appropriate metrics, Middle - East Journal of Scientific Research, v-20, i-12, pp-2464-2470, 2014.
- [7] Thooyamani, K.P., Khanaa, V., Udayakumar, R., Using integrated circuits with low power multi bit flip-flops in different approach, Middle - East Journal of Scientific Research, v-20, i-12, pp-2586-2593, 2014.
- [8] Thooyamani, K.P., Khanaa, V., Udayakumar, R., Partial encryption and partial inference control based disclosure in effective cost cloud, Middle - East Journal of Scientific Research, v-20, i-12, pp-2456-2459, 2014.
- [9] Thooyamani, K.P., Khanaa, V., Udayakumar, R., Virtual instrumentation based process of agriculture by automation, Middle - East Journal of Scientific Research, v-20, i-12, pp-2604-2612, 2014.
- [10] Sundar Raj, M., Saravanan, T., Srinivasan, V., Design of silicon-carbide based cascaded multilevel inverter, Middle - East Journal of Scientific Research, v-20, i-12, pp-1785-1791, 2014.
- [11] Thooyamani, K.P., Khanaa, V., Udayakumar, R., Wide area wireless networks-IETF, Middle - East Journal of Scientific Research, v-20, i-12, pp-2042-2046, 2014.
- [12] Udayakumar, R., Kaliyamurthie, K.P., Khanaa, Thooyamani, K.P., Data mining a boon: Predictive system for university topper women in academia, World Applied Sciences Journal, v-29, i-14, pp-86-90, 2014.
- [13] Lingeswaran, K., Prasad Karamcheti, S.S., Gopikrishnan, M., Ramu, G., Preparation and characterization of chemical bath deposited cds thin film for solar cell, Middle - East Journal of Scientific Research, v-20, i-7, pp-812-814, 2014.
- [14] Premkumar, S., Ramu, G., Gunasekaran, S., Baskar, D., Solar industrial process heating associated with thermal energy storage for feed water heating, Middle - East Journal of Scientific Research, v-20, i-11, pp-1686-1688, 2014.
- [15] Gopalakrishnan, K., Sundeep Aanand, J., Udayakumar, R., Electrical properties of doped

azopolyester, Middle - East Journal of Scientific Research, v-20, i-11, pp-1402-1412, 2014.

[16] Achudhan, M., Prem Jayakumar, M., Mathematical modeling and control of an electrically-heated catalyst, International Journal of Applied Engineering Research, v-9, i-23, pp-23013-, 2014.

[17] Thooyamani, K.P., Khanaa, V., Udayakumar, R., Application of pattern recognition for farsi license plate recognition, Middle - East Journal of Scientific Research, v-18, i-12, pp-1768-1774, 2013.

[18] Jiawei Han, Micheline Kamber, Data Mining: Concepts and Techniques.

[19] Dr. Lokanatha C. Reddy, A Review on Data mining from Past to the Future, International Journal of Computer Applications (0975 – 8887) Volume 15– No.7, February 2011.

