

Optimization of Business Intelligence using Data Digitalization and Various Data Mining Techniques

Deepti Sindhu[#], Anupma Sangwan^{*}

[#]M.D.U., Rohtak, ^{}GJUS&T, Hisar,
Haryana, India.*

Abstract

Data mining is a technique to extract valuable data from huge volume of data. The main aim of data mining is to discover knowledge. Methods of Business Intelligence include retrieval of information, mining data, analysing statistical data and visualisation of data. Statistical analysis of digital data is very much fast. Large amount of data is accessed in different formats from heterogeneous sources. Collected data is consolidated and analysed that is key of business intelligence. Digital data can be correlated very easily within databases of companies. Management and processing of all kinds of documentations has been automated which has reduced paper work. Now business data is stored in electronic format that can be processed, analysed and discovered very easily. Data mining techniques in real are applicable on only digital data. Business Intelligence can be aided very efficiently using data mining techniques and tools. Data mining reduces the workload and makes business intelligence faster and digital data fulfil data mining objectives. Data mining accomplishes the objectives of business intelligence by classification, pattern recognition, clustering, prediction and decision making. Today in every area data has become unlimited. It is impossible to store, process and manage it manually without digitalization. Various application areas of digitalization are banking, commerce, marketing and etc. This paper introduces the concept of Data Digitalization, Data Mining, DM Techniques, Knowledge Discovery and Business Intelligence.

Keywords: Digitalization of Data, Data Mining, DM Techniques, Business Intelligence, KDD Process

1. INTRODUCTION

Data is used in digital form for processing and analysing in every organization. Either digital data is accessed from various sources and processed or if found in another form converted into digital format and processed for various industrial and organizational applications. Data processing includes data mining [1], transformation and knowledge discovery [12]. Regularly huge amount of data is processed in number of organizations in government and private sectors. Infact, today currency is saved and spent through electronic wallet. No company in the market can survive without digital currency. India is also a developing digital country like many other countries. Almost all industries, businesses, markets and organizations from small level to large are doing implementation of digitalization. This has increased the profit in Indian market.

2. DATA DIGITALIZATION

Digitalization must lead to data usage in meaningful quality of a product can be determined more easily using feedback loop created by digitalized data. A feedback loop is worthwhile for influencing digital inputs to be made in future. The main purpose of Data Digitalization [4] is to focus on useful data. A large number of technologies are available today for collecting data and connectivity to previously stored data like smart phones, smart computers and other smart wearable. These devices containing smart technology have completely changed the way of business. Using big data, business can be analysed on broad level.

2.1. Digital Commerce:

Usage of digital channels by consumers is increasing very rapidly. Today more than 50% of the businesses are present online and 70% people purchase products online which have increased e-market. Products with their details are uploaded on business sights. All payments have become e-payments [7] in E-commerce. According to a study online retail increased very rapidly due to large screen size of smart phones and tablets. It has made purchasing and buying very easy.

2.2. Data Mining:

Data mining [1] is a process of mining the hidden patterns from large amount of domain specific data. It is a useful process of knowledge discovery. Various techniques [2] of data mining are classified as follows:

- **Classification:** It is a process of classifying new tuples on the basis of training data into classes. It is a predictive technique [9][10]. Using known values of one variable, unknown value of many other identical variables can be

predicted. It is a supervised technique. Various techniques available for classification are decision tree, Support Vector Machine (SVM), Bayesian classification, neural networks, induction rules and etc.

- **Clustering:** It is an unsupervised technique to find outliers in a given set of values. It is helpful in finding noisy data. Various techniques of clustering are K-Means, Density-based clustering, Distribution based and Centroid-based clustering [10][11][12]. Clustering is a descriptive technique.
- **Prediction:** Prediction [9][10][12] is a technique to predict the value of continuous data. For smart analysis in Business Intelligence historical data can be combined with predicted data.
- **Association:** It is used to study relationships between different variables of a problem [10]. For example, to find the sales opportunities for a business retailer based on historical transaction and purchase.

All techniques discussed above are helpful in data analysis and decision making. They are also important from business point of view.

3. BUSINESS INTELLIGENCE

Business Intelligence also called BI [6] is process that is derived by technology for data analysis and presents some actionable information. This information help corporate and business managers and executives to make business related decisions. Organizations can run queries using data and view useful information for their business. Reports can be generated in various suitable visualised forms. Figure 1 shows the trends for BI in India.

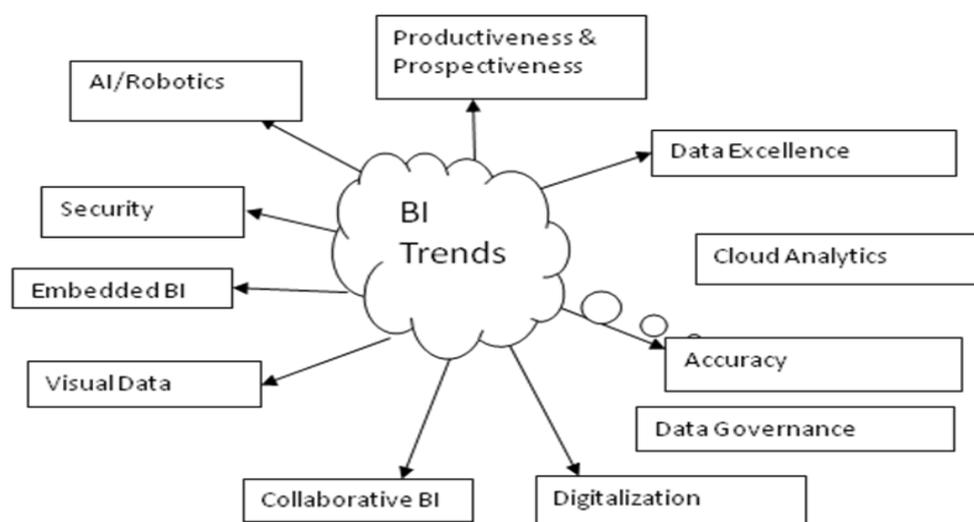


Figure 1: Business Intelligence Trends

3.1. Data Mining in context of Business Intelligence:

Business intelligence includes many relative activity such as data mining, OLAP (online analytical processing), query and data visualization (reporting). Figure 2 shows how digital innovation improved productivity of data mining. Today miner's focus is to improve the productivity. A Breakthrough is coming in the potential to achieve the productivity performance of mining due to digital technology.

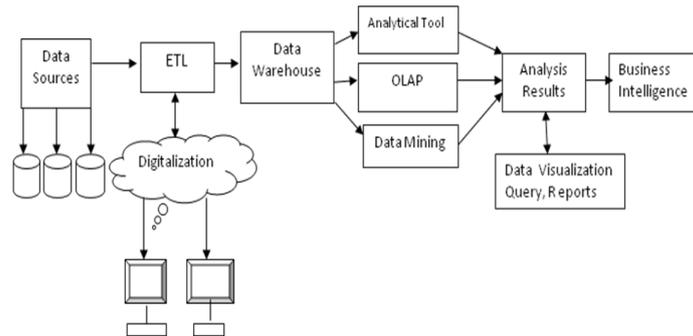


Figure 2: Data Mining in context of Business Intelligence

Data digitalization provides many innovative technologies for data mining. Artificial Intelligence has increased and made advancement in data mining. These applications provide optimization in BI [6].

- Improved data computational power.
- Improve connectivity.
- Easy and comprehensive resource understanding of materials and equipments
- Improvement in failure anticipation
- Optimized mechanization via automation
- Improved performance of monitoring in real time
- Improved time consumption of data analysis
- Improved accessibility
- Maximised human machine interaction

These all are opportunities which have really potential outcomes. Digitalized data of global level industries can be mined for analysis, knowledge discovery and decision making.

3.2. Digital Business:

Before many years, huge amount of time was being wasted on various data discrepancies. Large number of reporting errors was found daily and weekly. Both business organizations and business client were in same situation. They used to

analyse data either manually or large stream of spread sheet, but digital business converted such business into intelligent business.

3.3. Challenges for business before Digitalization:

Following are some challenges in front of a good business:

- Reports making was also one of the main challenges which are important part of decision making.
- Inconsistency in various reports used to be.
- Insufficient resources for testing of reports and data.
- Much time consumption in processing all the requests of data analysis.
- Knowledge discovery was difficult.
- Missing or noise oriented data reports.

Digital market and BI are two sides of a coin. Both work for and work with each other both are developed by data analysing [5] and strong decision making and optimizing investment and returns. Search engine marketing is growing now to show analytical positions of organisations. Social media is also playing role in BI in digital marketing.

3.4. Knowledge Discovery: Knowledge [8] is the most important term for analysis and intelligence in business. The main aim of both is to discover meaningful data that becomes knowledge in near future. Data is stored in databases then knowledge is discovered using KDD or Knowledge Discovery in Databases [10][11]. The knowledge discovery is mandatory part of Data Mining and Business Intelligence. Data stored in form of raw is converted into useful, understandable format by passing through various stages of KDD process. The main task of mining is ETL means Extract, Transform and Load. Figure 3 shows KDD process which is divided into five important parts.

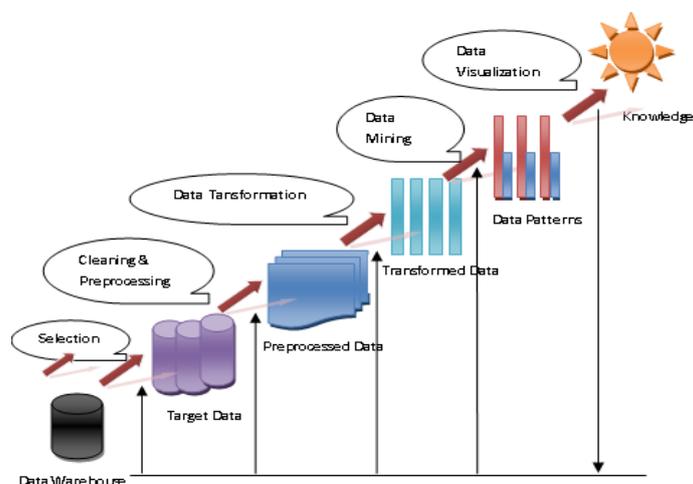


Figure 3: KDD Process to Discover Knowledge

1. Selection of data from Data Warehouse: Data is collected from heterogeneous sources and integrated.
2. Data Cleaning and Pre-processing: The unformatted raw data is pre-processed and cleaned. There can be missing data and noisy data in fields of databases.
3. Data Transformation: Data is transformed and reduced in required form after pre-processing to achieve goals.
4. Data Mining: Transformed data is processed for generating patterns using data mining techniques.
5. Data Visualization: After data mining patterns are generated. Data is visualized to show results to interpret easily. Data can be evaluated very efficiently.

4. CONCLUSION

Above discussion evident that digitalization of data has really improved tasks of data mining and business intelligence. But now data mining is being optimized using digitalization that is also improving the business analytics [5] [6] and business benefits. There are very minor differences in BI and BA. Business analytics is an umbrella like structure, business intelligence comes as a part under it. Both Business Intelligence and Business Analytics can't be achieved without Knowledge Discovery. It can be concluded that Data Mining and knowledge discovery can be achieved very efficiently on electronic data which are key to Business Intelligence

REFERENCES

- [1] Smita M. Nirkhil, R.V.Dharaskar, V.M.Thakre, "Data Mining: A Prospective Approach For Digital Forensics" Proceedings of the International Journal of Data Mining & Knowledge Management Process (IJDMP) Vol.2, No.6, November 2012.
- [2] B.Sangameshwari, P.Uma, "A Survey on Data Mining Techniques in Business Intelligence" International Journal Of Engineering And Computer Science ISSN: 2319-7242, Volume 3 Issue 10, October 2014, Pp. 8575-8582.
- [3] Yunita, M. Galinium and Lukas, "Integrating data mining technique and AHP in market analysis to propose new product development in real estate" IOP Conference Series: Materials Science and Engineering, Volume 166, Conference1.
- [4] Pratiksha R. Deshmukh, "Big Data Digital Reference Technology", International Journal of Advanced Research in Computer Engineering & Technology (IJARCET), Volume 5 Issue 3, March 2016, 663 ISSN: 2278 – 1323.

- [5] Kuldeep A. Maurya, Vidyashri Shinde, “Review On Business Analytics & Business Intelligence”, ISSN:2278–1323, International Journal of Advanced Research in Computer Engineering & Technology (IJARCET), Volume 5, Issue 6, June 2016.
- [6] Nur Hani, Zulkifli Abai, Jamaiah Yahaya & Aziz Deraman, “Business Intelligence and Analytics in Managing Organizational Performance: The Requirement Analysis Model”, Journal of Advances in Information Technology Journal of Advances in Information Technology Vol. 7, No. 3, August 2016.
- [7] B.I.Asiligwa, E.I.Omwenga, “A roadmap for the Adoption of Government E-payments in Kenya”, International Journal of computer Applications”, Vol.144, No.1, June 2016, pp.08-18.
- [8] Farzad Tarhani1 & Omid Zare Ameli, “Business Intelligence Application Model in Hedge Funds Supporting Knowledge-Based Companies”, Modern Applied Science; Vol. 10, No.12; 2016, ISSN: 1913-1844, E-ISSN: 1913-1852, Published by Canadian Center of Science and Education.
- [9] Deoshree Diwathe, Snehlata S.Dongare, “Classification Model Using Optimization Technique: A Review”, International Journal of Computer Science and Network, Volume 6, Issue 1, February 2017, ISSN (Online): 2277-5420, pp. 42-48.
- [10] Pallavi D. Bagul, K. C. Waghmare, “A Survey of Data Mining Techniques for Improvement of Prediction Accuracy”, International Journal of Innovative Research in Computer and Communication Engineering, (An ISO 3297: 2007 Certified Organization) Vol. 5, Issue 3, March 2017, ISSN(Online): 2320-9801, pp 3822-3825.
- [11] Arpit Bansal, Mayur Sharma, Shalini Goel, “Improved K-mean Clustering Algorithm for Prediction Analysis using Classification Technique in Data Mining”, International Journal of Computer Applications (0975 – 8887) Volume 157, No 6, January 2017, pp.33-40.
- [12] Aishwarya Pawar, Varsha Bavane, Bhagyashree Nigade, Siddhant Navaratna, Sampada Kulkarni, “A Survey of Data Mining Techniques for Improvement of Prediction Accuracy”, International Journal of Innovative Research in Computer and Communication Engineering, ISSN(Online): 2320-9801, Vol. 5, Issue 2, Feb 2017, pp 2229-2233.

