

INTERNATIONAL JOURNAL OF ENGINEERING SCIENCES & MANAGEMENT A SURVEY OF BUSINESS INTELLIGENCE USING DATA MINING, WEB MINING AND TEXT MINING

Prateek Rokadiya^{*1}, Siddharth Sharma², Sandesh Joshi³, Akshita Pamecha⁴, Tapan Akolekar⁵,
Sajag Jain⁶

ABSTRACT

Information and technology has taken new changes in the field of communication. This new world is the world of digital with the large amount of data is available on the internet. Lot of peoples is accessing the social sites like facebook, government sites also having the data in huge variety. Banks now maintain all their data related to customer and employees on internet. We required finding new ways and technologies through which we can search the closer observations on this huge data and generate knowledge, i.e. why mining technologies comes in to existence. Mining techniques basically used some automated tools to achieve business intelligence. Using these tools, we will able In this paper, we will talk about mining techniques, their implementations and we will also talk about how these mining techniques will be useful to achieve Business Intelligence.

Keywords- Data mining techniques, web mining, business intelligence, information and knowledge.

I. INTRODUCTION

With the help of new techniques and technologies, peoples are living their lives in the data world. Information is now available to the every next door. The latest trends and techniques finds the new means to the internet and now the data is in the huge amount and in different form to be available on the internet. These new revolution in the field of information and communication provided us new ways to communicate and ease to access any information either through internet or mobile communication. To generate the business intelligence information from these raw data is extremely wearisome task for the enterprises. Many of the enterprises actually rely on the automated mining tools that able to discover the knowledge which can be benifial to the business. The competition between the rivals in the market is increasing day by day and they are able to find some new means so that they can discover the knowledge from these huge raw materials.

Data Mining (DM) is defined as the process of analyzing large databases, usually data warehouses or internet, to discover new information, hidden patterns and behaviors. It's an automated process of analyzing huge amounts of data to discover hidden traits, patterns and to predict future trends and forecast possible opportunities. DM analyses datasets of rational databases. One of the main purposes of using data mining in business intelligence is to risk management.

Text mining is the again interesting field which works on the textual information rather than the record. Text mining is somehow related to data mining but it uses different methodology or approach to discover the knowledge. The data is in the huge amount on the internet, text mining techniques will able to find the hidden information by uses Natural language processing techniques.

Web mining techniques will collect all the information from different servers and discover the knowledge and hidden information from different activities on the internet. It is very helpful for the enterprises to provide different offers to the customer from the knowledge generated by the web mining techniques and they can better their relationships in the market. The major use of web mining techniques is in the network forensics.

II. BACKGROUND DETAILS

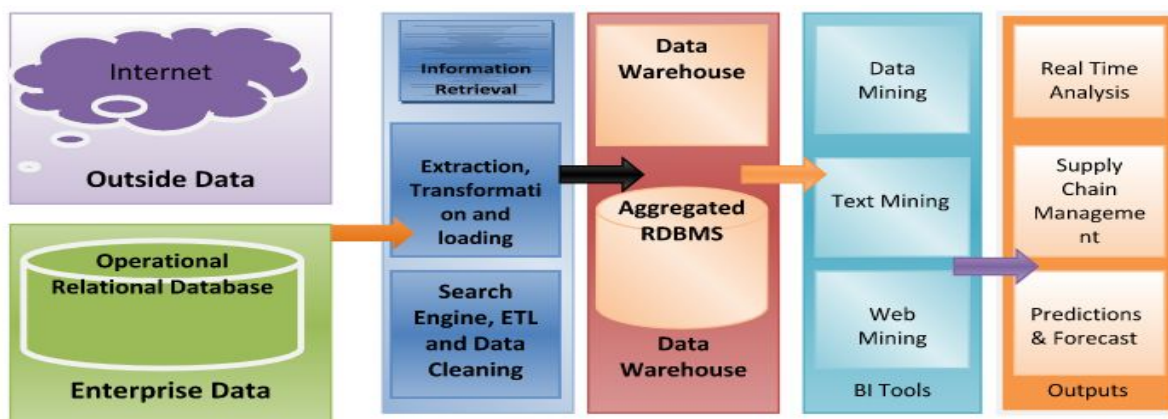
The incredible encroachment made in the mining tools and technologies has modified notion from data compilation to knowledge discovery and repository. With today's prevailing and reasonably economical hardware and network communications, coordinated with highly developed software for mining, business enterprises are acclimatizing mining technologies as fundamental business processes. In addition, the Internet has an primary responsibility as network and interactions are omnipresent today, mining is carried over the world through the network of databases. The gigantic amount of knowledge is not only accumulated at the top senior management level but at all the other levels of an enterprise as well. Today mining software employs complex algorithms for searching, pattern recognition, and forecasting complex stock market changes. IBM and Microsoft are on an epic race to produce best DM software to date. As modern as it is today, WM has become an increasingly adopted business process as well. WM is suited more for ecommerce than DM and TM.

III. BUSINESS INTELLIGENCE THROUGH DATA, TEXT AND WEB MINING

Business intelligence is now one of the crucial factors to run the business in today’s market. To discover the knowledge and information through mining is known as business intelligence. Data mining tools and techniques will become one of the essential parts of any enterprise. In this digital world the rivals are increasing day by day and to run a business is becoming more tedious and cumbersome. Every organization is finding and generating the new ways to run their business in profitable manner. Tremendous amount of data is generated from different web sites and to maintain the information of those data will be difficult problem for the organizations. Consider an example of credit card transactions. There are many customers using credit card. The numbers of customers are in the range of millions and to maintain their related and essential information is extremely hard. The mobile users in the market are using the different freely downloaded apps like whats up and many more and billions of messages will be sent via the mobile on the daily basis. The social websites will closer the people but it also manages the huge amount of data. Data can be structured, unstructured and semi structured but very useful to generate the information so that the organizations are able to discover the hidden patters.

Business intelligence has influenced many sectors like banking, government, insurance, retail and etc. As we know that today maximum customers are doing online shopping. From the small thing to the big thing, all the buying will be take place on internet. The regular customer who will buy the all kinds of electronic equipments, clothes and many more things are essential to the enterprises like Amazon, Flipkart, craftvilla and many more. They need to maintain their relationship with the customer. The customer who buys the things using these sites needs to maintain the customers buying information’s so that they can provide them the handsome offers in the future. All these are the rivals to each other and always try to impress the customers by providing them good offers. Another thing to be considered is that the credit card companies may also tie up with that and provide the extra offers to the customers like for the business oriented customers they have the different credit card schemes, for the government officers, for the people who works in the education sector, and the people who are working in private sectors. For each of them they provide the greedy proposals.

Mining tools provides predictive profiling; this means that using current and historical behaviors of your customers, possible future behaviors of purchase are predicted. The insurance industry is most interested in their customers’ medical records and its history. Stock market predictions are mostly done using mining. The Google Finance web page contains dynamic charts of international stock markets, with references to critical points in the graph directing to web pages that are the service got its information, providing assurance to end users. Mining tools are also used for automatic spam detection, and the defence against fraud, through fraud detection techniques utilizing mining tools [30]. Most major banking and telecommunication companies apply automated fraud detection systems through mining techniques, AT&T, bank of America are examples of such users of fraud detection. In the next subsections we will look at the main aspects were BI through mining tools is used to gain business proficiency. BI and mining tools are used exchangeable in the following text. In addition, mining tools, especially DM, provides market basket analysis that helps the discovery of products that are bought usually together. As modern economies around the world today are driven by information, becoming information and knowledge based economies. BI tools are from the top reasons of development information technologies in business today. BI tools in business today are integrated in most enterprises tools such as Enterprise Resource Planning ERP tools, Customer Relationship Management CRM tools, supply chain management tools, data warehouses, and even RDBMS.



IV. BUSINESS ADVANTAGES THROUGH MINING

- ❖ Every enterprises/ organizations have develop the research group that analyze the market needs by collecting the data sets from different resources and then able to discover the fruitful information. They research the market through mining techniques and will consider the different scenario that, what is the market strategy, what product is the need of every single person, why the product is so popular in the market, they try to find out the unseen factors. For example every online shopping site will provide the mining baskets, the people will collect the different things in that basket, and from that basket they buy many things and the remaining items may be rejected by them. The enterprises then analyze the hidden facts so that they can provide them the better offer in future.
- ❖ Data mining tools are also useful to analyze the risk management. Like in banking sector total loss predictions can be analyzed so that the bank will do the needful.
- ❖ In manufacturing industries Management Information system tool is used to automate the work but with the help of such tools we cannot monitor the process of manufacturing nor guide them for better performance.
- ❖ Improved customer relationship.
- ❖ Better logistics and inventory management.
- ❖ Fraud detection

Business Aspect	Business Intelligence Advantage	Benefits
Competitive Advantage	<ul style="list-style-type: none"> • Market Research • Risk Management • Manufacturing Optimization 	<ul style="list-style-type: none"> • Finding Elements of Market Dominance • Bankruptcy Prediction, Better Investments • Better material usage, shipments, scheduling
Customer Relationship Management	<ul style="list-style-type: none"> • Customers’ targeting • Pricing Discrimination • Market Baskets • Customers Satisfaction 	<ul style="list-style-type: none"> • Target specific customers with the right products • Dynamic pricing • Better Marketing and Advertisements • Find the reasons and the costs of switching, chum, and satisfactory levels
Logistic and Supply Chain Management	<ul style="list-style-type: none"> • Production Managements • Scheduling Supply Chain • Dynamic Reactions • Forecasting 	<ul style="list-style-type: none"> • Prevent overproduction and underproduction • Help dynamically manage the supplies during their move through the chain • React immediately to changes to help sustain supply • Forecast the demand for production
Anomalies and Fraud Detection	<ul style="list-style-type: none"> • Fraud Detection • Anomaly Detection 	<ul style="list-style-type: none"> • Help find fraudulence transactions, fraudsters, hackers, and possible counterfeiting • Find what data to leave out, why such anomalies happened, and avoid considering them.

V. CONCLUSION

Competitiveness today is driven through BI. Companies achieving high competitiveness are the companies utilizing BI tools. Mining technologies had come a long way. The software development, along with hardware developments made possible of more commercially available mining tools. Revolution of information brought in by the Internet and the telecommunications technologies, made them a huge source of information, sometimes for free even.

BI utilizing this vast amounts of data can help in achieve competitive advantages, better customers’ relationships, effective resource planning, and fraudulence detection. As BI tools implement AI techniques, decision trees, NLP, and SM technologies, they are considered as sophisticated and highly specialized tools. Many challenges hinder the further developments of such tools. The challenges are technological, ethical, and legislative. As more enterprises and governments are more dependent on such tools, we think that some obstacles have to go in order to progress.

REFERENCES

[1] Doug Alexander, (2011) “Data Mining”, dea@tracor.com
 [2] Ralf Mikut, and Markus Reischl, (2011) “Data mining tools” Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, Vol. 1, Issue 5.
 [3] Clifton Phua, Vincent Lee, Kate Smith, Ross Gayler, (2010) “A Comprehensive Survey of Data

Mining-based Fraud Detection Research” Cornell University library, CoRR.

- [4] Thiagarajan Ramakrishnan, Mary C. Jones, Anna Sidorova, (2011) “Factors Influencing Business Intelligence and Data Collection Strategies: An empirical investigation”, *Decision Support Systems*.
- [5] Surajit Chaudhuri, Vivek Narasayya, (2011) “New Frontiers in Business Intelligence” *The 37th International Conference on Very Large Data Bases, Seattle, Washington, Vol. 4, No. 12, VLDB*.
- [6] Dien D. Phan, Douglas R. Vogel, (2010) “A Model of Customer Relationship Management and Business Intelligence Systems for Catalogue and Online Retailers”, *Information & Management, Vol. 47, Issue 2, Pages 69-77*.
- [7] Christian Thomsen, Torben Bach Pedersen (2009) “A Survey of Open Source Tools for Business Intelligence” *International Journal of Data Warehousing and Mining, Vol. 5, Issue 3, IGI Global*.
- [8] Meryem Duygun Fethi, Fotios Pasiouras (2010) “Assessing Bank Efficiency and Performance with Operational Research and Artificial Intelligence Techniques: A survey” *European Journal of Operational Research, pp. 189–198, Elsevier*.
- [9] Rafael Berlanga, Oscar Romero, Alkis Simitsis, Victoria Nebot, Torben Bach Pedersen, Alberto Abelló, María José Aramburu (2012) “Semantic Web Technologies for Business Intelligence” *IGI*.
- [10] Manuel Mejía-Lavalle, Ricardo Sosa R., Nemorio González M., and Liliana Argotte R. (2009) “Survey of Business Intelligence for Energy Markets” *E. Corchado et al. (Eds.): HAIS, LNAI 5572, pp. 235–243, Springer-Verlag Berlin Heidelberg*.