ISSN: 0976-3031

Available Online at http://www.recentscientific.com

CODEN: IJRSFP (USA)

International Journal of Recent Scientific Research
Vol. 9, Issue, 10(B), pp. 29066-29073, October, 2018

International Journal of Recent Scientific Research

DOI: 10.24327/IJRSR

Research Article

WHY ONLY DATA MINING? A PILOT STUDY ON INADEQUACY AND DOMINATION OF DATA MINING TECHNOLOGY

Sohail MN[1]*., Ren Jiadong[1]., Uba, MM[1]., Irshad, MI[1]., Musavir Bilal[1]., Abir, SI[1]., Wasim Iqbal[2]., Usman Akbar[2]., Tahir Rizwan[3] and Anthony. JV[4]

[1]Department of Information Science and Technology, Yanshan University, Qinhuangdao, China
[2]Department of Economics and Management, Yanshan University, Qinhuangdao, China
[3]Department of Controls and Engineering, Shanghai Jiao Tong University, Shanghai, China
[4]Department of Hotel Management, American Hotel and Lodging Association, New York, United States of America

DOI: http://dx.doi.org/10.24327/ijrsr.2018.0910.2787

ARTICLE INFO

Article History:
Received 4th July, 2018
Received in revised form 25th August, 2018
Accepted 18th September, 2018
Published online 28th October, 2018

Key Words:
Data mining, knowledge discovery, limitations, and domination

ABSTRACT

Data mining is the process of discovering the knowledge by analysing and extracting the data from various highly repositories and the result bases on the useful and functional information's for the user. By the passage of time data mining is growing very vastly and became the famous technology by analysing and extraction of knowledge. We are standing at the point where life can have a better understanding of the problems. There was a time to start an active research on data mining but the limitation of this technology is under predictions as; is this technology has any limits for the future or it is limitless towards the growing world? Why only data mining technology is involves in the refining process of data? How efficiently the future relies on this technology? Is this new technology is so capable of being popular and more powerful in all respective fields? What are its limitations and how it is dominating the future? In this paper we have revealed the facts of growing fields with this manifesto and how it is affecting anonymously and how reliable the future is on this technology? And why this technology is so important for future?

INTRODUCTION

Definition

According to (Frawley *et.al*, 1992[1]) "data mining can be viewed as the non trivial extraction of implicit, previously unknown and potentially useful information's from data". Furthermore, "the process of exploring and analysis of large quantities of data in order to discover meaningful patterns and rules", mentioned by (Berry *et.al*, 2000[2]). In the eye of various other scholars and researcher, data mining has defined as:

"Data mining is the process of exploration and analysis, by automatic or semiautomatic means, of large quantities of data in order to discover meaningful patterns and rules" said by (Chen *et.al*, 2008[3]). According to (Wang *et.al*, 2018)"Data mining is the application of statistics in the form of exploratory data analysis and predictive models to reveal patterns and trends in very large data sets"[4]."Data mining is the nontrivial process of identifying the valid, novel, potentially useful and ultimately understandable patterns in data", said by (Fayyad *et.al*, 1996[5]).(Rygielski *et.al*, 2002) narrates as "Data mining is a set of methods used in the knowledge discovery process to distinguish previously unknown relationships and patterns within the data"[6].The most dynamic definition of data mining according to the (Prabakaran *et.al*, 2018[7]), "data mining is the process of Decision Tree, Neural Networks, Rule Inductions, Neatest Neighbours and Genetic Algorithms".

Introduction

Data mining is the new technology of extracting the hidden and potential information's from large-scale databases also known as the knowledge discovery of databasesand process of discovering new data for the potential growth in business platforms(Sohail *et.al*, 2019[8]). This technology holds the greatest potential for different companies on various platforms, whose focus goes on to the growth of needs towards future(Pangrazio *et.al*, 2018[9]). Data mining is not the program but process and series of tools that help to analyse the data and information's, which has collected to create knowledge in order to support decisions-making(Grover *et.al*, 2018[10]). According

Corresponding author:* **Sohail MN
Department of Information Science and Technology, Yanshan University, Qinhuangdao, China

to the researchers(Alonso *et.al*, 2018[11]), the analysis step of knowledge discovery in data bases is called data mining and it uses different tools that allow user to analyse data from different dimensions or angles to categorize and summarize it in the identified relationships. Technically, data mining is the process of finding the patterns among dozens of fields in large relational databases(Silva *et.al*, 2018[12]). Such patterns can be used to increase the revenue, cut costs and sometimes both. In other words, the overall goal of data mining process is to find the information's from a dataset and transform it into an understandable structural format for the future growth of business in various fields(Prabhu *et.al*, 2018[13]).

In past, we had a question, what happens? But with data mining' we discover that what will happen and why? (Nakasima *et.al* 2018 [14]). Data mining integrates the various technologies like statistics, machine learning and databases(Irshad *et.al*, 2018[15]). It has applications in different disciplines like medical, financial, defence, intelligence and so on(Sohail *et.al*, 2019[8]). The tools of data mining include clustering, classifications, associations and detections (Muhammad *et.al*, 2017[16]). From decades, data mining have developed in many ways regarding techniques, which includes extracting association, neural networks, logic programming, rough sets and decision trees(Zhu *et.al*, 2018[17]). Furthermore, data mining has gone beyond limits like the relational databases to the text mining and multimedia data(Ristoski *et.al* 2018[18]); also it's involved in the information security and detections(Santis *et.al*, 2018[19]). After so many developments, companies are still facing some challenges like Scalability (Sohail *et.al*, 2017[8]), but till far data mining is working on the massive quality of datasets and also engaged in working for the Terabyte sizes(Najjar *et.al*, 2018[20]). By the enormous growth of data in different disciplines' the question arise, Can this technology fulfilthe needs in extraction of Petabyte size data? This comes in the limitation and domination of data mining technology (on which this paper is focusing). As data mining involves some algorithms, it is important to understand the limitation of data mining algorithms and tools. Which requires the time and space for the complexity? In example: can these algorithms be completed in time? If the problem is decided, what is the complexity? For future predictions, we need to find out more about the complexity of markets and business(Alonso *et.al*, 2018[11]) and how data mining is shining in the financial platforms.

During the process of research, we have analysed 360 papers from 1996 to 2018 with the Key search "data mining limitations", "data mining future", data mining growth" and "data mining influences" in IEEE and Web of Science databases. And came up with a single study on complete limitations and dominations of data mining in a single platform for the researchers.

This paper is divided into 7sections. Section 3 will define the roots of this technology arises in the world. Section 4, will hold the various components, process flow structure and will talks about' how this technology is affecting the modern business and needs? Section 5 and 6 will list about data mining phase of facing and further it talks about how this vast growing technology is dominating the entire world? And how far the future of this technology is shining?

Data Mining Roots

"The study of data mining is infancy"as, pointed out by (Chow *et.al*, 2018[21]). Generally by going through from different conferences and articles, the foundation of data mining is as a big topic. But still many questions are arising time by time. Firstly it is necessary to talk about the roots of data mining to clear the main concept by (Arias *et.al*, 2018[22]). Data mining involves different techniques, which are the result of long process research and product development. Data mining came from the three roots as mentioned in (Sohail *et.al*, 2019[8]), which involve statistics, machine learning and databases management.

Statistics

Without statistics there would be no data mining causes this is the core foundation of various technologies. From the perspectives of statistics, data mining is the knowledge of computer programmed and investigative data analysis of complex datasets. This technology has major impacts on Finance and science. Even with the connections in-between data mining and statistics analysis of data, so far many methods have used from statistics to other fields (Muhammad *et.al* 2017[16]).

Machine learning

It is the second largest family line of data mining, where machine learning focuses on designing of algorithms that are capable of making predictions on the datasets (Irshad *et.al* 2018[15]). Machine learning involves supervised and unsupervised learning, basically it blends Artificial intelligence with advance statistical methods to roll up the computer programs for better learning and apply that knowledge to datasets, which has further get extracted with data mining tools and methods.

Databases

The third key root of data mining is databases, which has introduced to store and manage the high amount of data. In past, data was stored and managed through manual file systems, which considered records and fields than various models came up by the passage of time and relational models served well to fulfil the need of storage. But in data mining the volume of data is high, so specialized servers has introduced for it. Which simply called as data warehouses (Wordsworth *et.al*, 2018[23]). Data warehouses also support the OLAP operation to support the decision-makings.

The time business data was firstly stored in computer, the revolution of data mining had begun. It continues with the improvement in data access and recently generated technologies that allow users to navigate through their data in real times. Data mining takes this evolutionary process beyond the retrospectives data access and navigates it to prospective ways and delivers proactive information's. In business community, the data mining is supported by three main matured technologies that are (Wordsworth *et.al*, 2018[23]):

1. Powerful multi processor computers.
2. Massive data collections.
3. Data mining algorithms

By the survey of Meta group's of data on the data ware houses project (Ranjan *et.al*, 2018 [24])has found that 19% of

respondents are beyond the 50 GB levels and 59% expect to be there by second quarter of 1996 but in the retail industries the numbers are much larger than expected. The need to improve the computational engines can now be met in cost-effective manners with some parallel multiprocessor computer technologies. Data mining techniques were been present since the last couple of years but recently it been came up with the maturities and reliabilities that consistently performing the old statically methods(Wordsworth *et.al*, 2018 [23]), which are now getting improved. The "Table 1"summarizes the evolution of data mining in the platform of database developments according to the researchers, which is taking the business industries to the peaks in accurate manners(Usai *et.al*, 2018, Yongmin *et.al*, 2018, Carnerio *et.al*, 2000, Hashem *et.al*, 2015, Marston *et.al*, 2011, Chen *et.al*, 2012 [25-30]).In the revolutionary process of business discipline from data to information, many steps have been taken (Chow *et.al*, 2018[21]). Also in vast research areas of statistics, machine learning and artificial intelligence, the majority of mature techniques are been introduced with high performance relational database engines and with broad data integrations (Nakasima *et.al*, 2018[14]).

Table 1 The evaluation of data mining technology for database development phase in business industries

Steps	Business	Expertise	Vendors	Selves
Data pool (1960s)	What was total revenue in previous year?	Computer, Tapes, Disks	IBM, CDC	Static data delivery
Data access (1980s)	What was the unit sale in ABC branch by March?	RDBMS, SQL, ODBC	Oracle, Sybase, Informix, IBM, Microsoft	Dynamic data delivery by recorded time level
Data ware house and decision support	What was the unit sale in ABC branch by March?	OLAP, data warehouse, multidimensional databases	Pilot, Camshare, Arbore, Cognos, Micro strategy	Dynamic data delivery at multiple levels
Data mining (early today)	What is likely to happen in sales of ABC branch and why?	Advance algorithms, multiprocessor, massive databases, computers	Pilot, Lockheed, IBM, SGI, start-up industry	Prospective, proactive information delivery

Why data mining?

There are two main ways of doing mining (Usai *et.al*, 2018[25]). The one comes is Manual mining, which receive data from the competitor sites for time by time analysis and make decisions and other one is Automatic process, which provides the desire accuracy in business and has ability to overcome the headaches of employers. Data mining is moving very fast in present time(Yongmin *et.al*, 2018 [26]). By observing its growth, still some questions come in the mind like:

1. Why only data mining technology is involves in the refining process of data?
2. How efficiently the future relies on this technology?
3. Is this new technology is so capable of being popular and more powerful in all respective fields?
4. What are its limitations towards the future?

This research paper talks about the various perspectives of data mining in different platforms that how does it shines and how far it can go?

MIT (Hashem *et.al*, 2015 [28]), the top ranked research university considered data mining as the top ten technologies for the growth of organizations in future. So many things come in mind regarding the collection of data that is; the data should be collected from the registration databases or the purchasing one (Hashem *et.al*, 2015 [25]). Mean to say is how truly the information's are been collected? And what decision the organization want to take with this information? Does they want to see, how to get members or how should they can sell their product? After all reviews(Santis *et.al*, 2018, Najjar *et.al*, 2018, Marston *et.al*, 2011, Chen *et.al*, 2012, Chow *et.al*, 2018, Hashem *et.al*, 2018 [19,20,29,30,21-28]), the main concept is' what business objective the organization's wants to achieve?

So data mining is the mace to acquire above the intact humanity. Further this part talks about the major reasons that lead the use of data mining to the peaks and that's competition. Competition is high as a result of modern marketing and distribution channels like telecommunication and Internet. In the worldwide competition (Li *et.al*, 2018,[31]), the main key to success of business is the ability to retain the customers and acquire more new. Scholars presented hundred of consulting organizations like SAAS, KD NUGGETS, SPSS, IBM, Microsoft, Oracle, KXEN, Agnos and etc. (Usai *et.al*, 2018, Yongmin *et.al*, 2018, Li *et.al*, 2018[25,26,31]). According to research(Sharma *et.al*, 2019 [32]), SaaS is the largest data mining market share product vendors, who are working in statistics field from decades. SPSS is another statistics working organization but the common thing in them is they have number of data mining products but decision tree is main focus. IBM is also holding sets of algorithms as well as the visualization tools. Further, Microsoft is the top databases vendor to include data mining features in relational databases. On other hand, Oracle is holding data mining algorithms based on the association and Naïve Bayes. Agnos has a knowledge studio that includes the power to build cluster analysis and other likely models. Also, KXEN has models like SUM regression, time series and segmentations(Sharma *et.al*, 2019 [32]).

Key process structure

In past(Yaqoob *et.al*, 2016 [33]), people used to say, data mining has no scope but eventually by the deep thinking of scholars many contributions are heading in this field and the future of data mining is shining accordingly in all respective fields from business to the world of science and technology. Data mining is the process of work(Li *et.al*, 2018 [31]), which makes the business more active, sensible and sustainable. It is the process of optimizing large databases in such short time to make the business vital and intelligent to the growth in organization. It organizes the hidden patterns in such way that organization can process the data sequentially to the success.

High performance in data mining allows the in-depth study of databases by dividing them into more rows/columns and attributes (Chen *et.al*, 2018 [34]), which is the sufficient success of data mining to analyse the business process from different perspectives. In the success of data mining, the most adorable techniques, which are still shining and capable of taking this technology to the peaks, are:

1. Association:(Chen *et.al*, 2018 [34]) In simple words, association means relations, which is the commonly and most simple technique of data mining. The process of this technique occurs by correlation between the two or

main items but often the same type of patterns to be identified.

2. Classification:(Sharma *et.al*, 2019 [32]) This technique is mostly used to create an idea on base of type of customers, objects and items by specifying multiple attributes to specify the defined class. The simple example is that you can easily identify the cars of different types like sedan, convertible, hatchback, 4/4 and etc. by specifying attributes like space, number of seats, shape and etc. Same thing happens while identifying the customers like behaviour, age, colour, height and etc.
3. Clustering:(Sharma *et.al*, 2019 [32]) The clustering process takes place to grouping the piece of data to form a structure opinion by identifying attributes and classes. Clustering is used to identify different information's because it correlates with various examples to check out the similarities and ranges in-between. In example clustering used to compare the sales data with the customer age and size of sales.
4. Predictions:(Li *et.al*, 2018 [31]) As this is the vast topic and runs to identify the fraud detections and in analysis of company's profit. Prediction works with the different data mining techniques like classification, relations, patterns identifications and etc.
5. Artificial Neural Networks:(Yaqoob *et.al*, 2016 [33]) Through this technique the non-linear models of predictive tasks are being structures into the biological neural forms of network.
6. Decision Tree:(Sharma *et.al*, 2019 [32]) It represents the structure of data in a tree format for better understanding and decisions. Decision tree in data mining includes the classification and regression trees like CART and CHAID (Chi square automatic interaction detection).
7. Nearest Neighbour Methods:(Sharma *et.al*, 2019 [32]) The KNN technique classifies the records in data sets based on the class of K record's, where K behaves as (K*1). For time being this technique is very sufficient in data mining methods for clustering and according to the predictions of scholars this method can takes data mining to the shining future.
8. Genetic Algorithms:(Usai *et.al*, 2018 [25]) Genetic algorithm applies in the optimization techniques for the process of evaluating the concepts by using genetic, mutation and natural selections of data.
9. Rule Induction:(Usai *et.al*, 2018 [25]) This technique involves "if-then" statement, which is getting more importance gradually for extraction of data based on the statically significances.

Many of these techniques are being used from the half past decade in scope of data mining and carry the main structure of progress towards future. These capabilities are approaching vastly and directly to the integration of OLAP platforms(Sharma *et.al*, 2019, Yaqoob *et.al*, 2016, Chen *et.al*, 2018 [32-34]).

Components

The fast growing technology 'data mining' works under three main components of data preparation, data acquisition and result evaluations, which cannot be missed while talking about the work and behaviour of this technology in different platforms especially towards the finance.

1. Preparation of Data:(Li *et.al*, 2018, Sharma *et.al*, 2019 [31,32])Data preparation is the process of checking data integrity and data consistency to deal with incorrect and useless data. The process of data preparation involves selection of data, integration of data and the conversion of data. In data selection, all data regarding business are searched' internally and externally and the suitable data for data mining application are selected to narrow down the scope and enhance the speed and efficiency of data mining. This data is selected and integrated from the databases systems to eliminate the redundancy of data. After that the data is converted for data mining algorithms in suitable formats according to the algorithms, which is the successful key of data mining.
2. Data Acquisition:(Usai *et.al*, 2018, Li *et.al*, 2018, Sharma *et.al*, 2019 [25,31,32])Data collection is the basic terminal in data mining technology. The achievement process of data mining involves four components. In steps, first comes the determination of task type, than selection of data mining techniques base on the task like classification models, which are rottenly used by neural networks, clustering by clustering analysis and correlation analysis by the technology of finding sequences and relations. Third step involves the selection of desire algorithms, which should have ability to find the hidden patterns with the matching algorithms to fulfil the goal of data mining. Finally in data acquisition, the mining process comes in which the selected algorithm used to repeat down the search process in order to collect the new models from the set of data.
3. Result Evaluation:(Sohail *et.al*, 2019 [8]) In some cases, the result of data mining doesn't persuade the users. According to the development of decision, the extracted information's are being analysed and the most precious information's are sent to the decision makers by the sympathetic platforms. If the desire results are not noteworthy than the process of data mining repeated. The analysis results are amalgamating into the business information platforms. By the help of visualization tools user determines the data consistency. This knowledge has been presented in different forms like charts, flow tables and etc. that helps and gives a big support to decision makers of organization's to make efficient decisions.

How business requirements are prepare firstly and process to define the knowledge and aggregate those information's through tools and algorithms for analysis to produce the desire format for decision makings are briefly illustrate in "Table 2".

Table 2 Business potential analysis in phase of data mining workflow

Increasing potential to support business decisions		
Decision-making		End users
Data presentation	Visualization techniques	Business analyst
Data mining	Information discovery	
Data explore	Summary	Data analyst
Data pre-processing/ Integration		
Data sources	Papers, web files, data bases, scientific experiments	

While talks about the different techniques and methods(Chow *et.al*, 2018, Wardsworth *et.al*, 2018, Carnerio *et.al*, 2000, Hashem *et.al*, 2015, Chen *et.al*, 2012, Sharma *et.al*, 2019, Chen *et.al*, 2018 [21,23,27,28,30,32,34]), here comes the "Table 3", which helps in better understanding of what technique or tool should be used by which time constraints.

Table 3 Method and techniques used in data mining workflow with the appropriate phase of class

Algorithms	Examples	Analytical Problems
Decision Tree Naïve Bayes Neural Networks	Credit Card analysis Customer retention	Classification
Clustering	Customer profit analysis	Segmentations
Decision Tree Association	Market base analysis Data exploration	Association
Time series	Forecast sales Predict stocks prices	Time series forecasting
ALL	Insurance rates Predict customer income Fraud detections Network analysis	Predictions Deviation analysis

How data mining is facing the world?

The field of data mining is growing vastly due to its vast achievements in various platforms(Sharma *et.al*, 2019, Yaqoob *et.al*, 2016 [32,33]). Numerous data mining applications have successfully implemented in healthcare field(Sohail *et.al*, 2019 [8]), as well as finance, fraud detections, telecommunications and risk management etc.(Sharma *et.al*, 2019 [32]). These are the just number of known fields but this technology is holding tight grips in various other domains. Data mining has an involvement in fight against terrorism(Dencik *et.al*, 2018 [35]). There are some intelligence agencies like in USA, who has started a new approach and launched Total Awareness programs with the approach and goal of establishing new databases of all information's about populations. Same thinking is launched in European countries and so on others in the world but this thinking is facing some problems like multimedia, text and audio mining's(Breiger *et.al*, 2018, Dimitrona *et.al*, 2018 [36,37]). Researchers and scholars are overcoming such problems from passage of times and somehow they are successful but still some more have to come in this field to overcome the issues but it all depend on the data sizes. Similarly data mining is also involves in health informatics to fight against the cancer and Aids like diseases(Sohail *et.al*, 2019, Irshad *et.al*, 2018, Muhammad *et.al*, 2017 [8,15,16]). After this Web is the hottest trend now, which is called the Semantic web to help the web to organize in the proper forms? The Resource description Framework is used to analyse the resources of social media like Face book. Data mining is not capable of working for the big size of databases same like the fight against the terrorism agencies are facing problems to mine the big amount of data(Dimitrona *et.al*, 2018 [37]). In health informatics the tools of data mining are not sufficient and scalable to mine the data within the time. Health industry is facing also some issues regarding the time scale process of data mining(Romero *et.al*, 1997 [38]). In part of social media, Face book like industries are having an issues of analysing statements and dealing with the erroneous data in their platform under hands with FOAF, which is the high supporting technology for describing resources (Drown *et.al*, 2018 [39]).

Many industries offer the wide range of banking services, credits and insurance also. According to the survey of present times (Chamoso *et.al*, 2018 [40]), the data collected in the banking industries are more reliable and complete with better quality. Which also involves the data mining technology to analyse and organize it in better ways but such industries are also facing so many issues regarding the size of data and time scales process. Same like banking, retail platforms also collect huge amount of data regarding the sales and customers(Li *et.al*, 2018 [31]). As the different scholars says that retail industries can provide the rich amount of data to mine but they are also facing the problems like scalability(Yongmin *et.al*, 2018 [26]). More over the telecommunication industries, which has rapid amount of data, like voice, text, fax, pager, images and etc. As telecommunication is in all over the world and has many competitors around the globe, they require a great demand of data mining in order to help them to understand the business involvements by identifying the various patterns and catch the fraudulent activities from number of users (Sharma *et.al*, 2019 [32]).

In fact, beside all other organizations Educational industries (Yaqoob *et.al*, 2016 [33])are also having troubles to mine their system by predicting paths for the student and alumni like which student will enrol in a particular programs or course? Who requires any additional assistance? Meanwhile the developers of educational industries are facing such issues and they require new and faster solutions by analysis of times and duration of the courses. Institutions can address their students and alumni in better ways by analysing the data and presentations of data. Until now, only Data mining has such enthusiasm and potentials to overcome such difficulties in various fields by extracting hidden patterns and up to some extend data mining is successfully doing its job but still some more have to be done in this technology for polishing its future.

Business problems for data mining

Data mining is the future of retail industry by analysing the cart bonding rates(Carneiro *et.al*, 2000, Hashem *et.al*, 2015, Marston *et.al*, 2011, Li *et.al*, 2018 [27-29,31]). The knowledge of screen checkouts can be improved or mainly this factor can be used to analyse the cost impacts and checkout decisions by having an eye on the competitor price. Data mining methods can help to improve the optimization of inventory and frequency of mining data because data mining has various techniques to solve many problems but still the vendors are facing difficulties, such as:

1. Cross Selling:(Marston *et.al*, 2011 [29]) This is important business challenge for the retailers that what product customers like to purchase?
2. Fraud Detections:(Hashem *et.al*, 2015 [28]) Insurance companies process thousands of claims in a day. Is this insurance claim is fraudulent?
3. Risk Management:(Carneiro *et.al*, 2000 [27]) In banks, the most common question is. Does this load should be approved for this customer or not?
4. Customer Segmentation:(Marston *et.al*, 2011 [29]) A deep thinking for the retailers is who is my Customer?
5. Sales Forecast:(Marston *et.al*, 2011 [29]) How the sales should be for the next week?

How business is getting competitive edges?

Data mining is being used to solve ton of problems in business and in every platform but still some more grips should be necessary to take deeply(Hashem *et.al*, 2015, Marston *et.al*, 2011, Li *et.al*, 2018, Sharma *et.al*, 2019 [28,29,31,32]). The major benefit that business has is, data mining enables the automatic explore of data and present it in more understandable format by identifying patterns and relationships' those effects on the business growth. More over it uncovers the relationship between data by prediction of different models, which can improve business in term of strategies and planning's. Data mining also helps for the future outcomes of business based on the predictive functions (Yaqoob *et.al*, 2016 [33]). Beside these benefits data mining is playing a vital role in decision management also to make efficient decisions, analysis of business market and in inventory cost system. In business, data mining is effecting on the psychology of customers to be better understood by analysing the market trends and also in psychology of making right sales decision for the better growth. Further more data mining is changing the environment of business in five ways, which are (Sharma *et.al*, 2019 [32]):

1. Customer devotion and Possession: A simple user has many interactions with the products before buying it, he checked the product usage and needs etc. so many predictions are being held to devote the buyer at first touch point, which also called a pre- buying point. So the customer possession is real need for long lasting value. To this point data mining is helping business to make the superb initial strategies for buyer in the market. In other words, data mining can able to work effectively with the management system of customers to make sure that customer will continue buying the product after the first purchase.
2. Customer Segmentation: Business always requires their customers to be ideal. Time by time with the experience the marketing strategies comes up. Data mining can helps to cross-reference the records of customers. By customer segmentation, business get benefits to improve the marketing strategy and development of desire products and by data mining it is possible to develop the new segments based on the minor checking's.
3. Inventory Management: The most sufficient way to control the cost of the company in areas of supply chain and manufacturing is by having a good control in inventory. By the good control in inventory, company can save millions of dollars annually. Here in this matter, data mining helps to control the streamline of inventory system by analysing the product demand and patterns of buying products. For the logistic companies, data mining can calculate the freight, shipping and aerial cost by analysing the local and seasonal variations. By this business can accelerate their services to the users in efficient manners.
4. Forecasting: The major part of business success in companies is forecasting because it is totally involved as sciences. In forecasting the sales and expenses should be touches the desire production demands. In the aspects of data mining, forecasting develop the accounts for more knowledge of market factors than ever before. This factor is making huge contributions in the company's growth.
5. Product conception: To observe the future of business, it is important to know what product is performing well today. Data mining only provides the opportunity to well compare the purchasing records.

Data Mining Domination

As the present era is describing about the data mining that it is the most commonly used method to extract data from different platforms and organize that data for the better use while implementing data mining, which has to face some challenges with the rapid growth. It lies in so-called smart technologies to identify the patterns and trends in large datasets(Santis *et.al*, 2018, Ranjan *et.al*, 2018, Careniro *et.al*, 2000, Chen *et.al*, 2012[19-24,27-30]). Recently so many players are been introduced but the most common and reliable player to identify this space is 'Factor Prism'(Tsironis *et.al*, 2018, Hasani *et.al*, 2018, Li *et.al*, 2018 [41-43]); this automatic tool can tease out the meaningful patterns in transactional time series data. Currently many organizations are fascinated with this new player to save time like Microsoft, IBM and etc. It allows to find the statically deviations automatically and remove the statically noise from data but the reviews of some scholars say that it's good for the near future but as the field is more challenging and only through practitioners will last longer. Companies from different platforms are awake to face all the developments. The basic step in data mining is complex algorithms as we studied, which allow data to be identified with different patterns and prediction of probabilities. That what is about to happen and how? But it's all depending on the source of data that even if it is raw in terms of analogy or digital. For that, industries should also be awake for the challenges of all the time (Sharma *et.al*, 2019 [32]). By the study of researcher's(Usai *et.al*, 2018, Yongmin *et.al*, 2018, Li *et.al*, 2018, Yaqoob *et.al*, 2016 [25,26,31-33]), the most common directions, where the future of data mining is devoting are:

1. Distributed data mining: One of the most fascinating areas of data mining is distributed data mining, which is grabbing the main attention for researcher. This often relates to collective data mining holding much attention in regards of future perspectives. In recent times the most of data mining happened on databases and data warehouses. Which being located on single plates but as the fast growing of companies, their data is becoming distributed to different platforms or can say to the different locations. To mine the data from different locations is called the distributed data mining (Niazalizadeh *et.al*, 2018 [44]). So this become the main goal for data mining user's to mine the data effectively, which is located on different locations. According to (Niazalizadeh *et.al*, 2018)"DDM is performing local data analysis for generating partial data models and combine the local data models in order to develop the global model".
2. Ubiquitous data mining: The mining of data from mobile has a rapid growth, as the technology is moving towards the future. This method of mining data is involved in mobile industries to mine the informational data of individuals. But still the challenges are on the way like cost, complexity and privacy. This method holds the grip

of studying the human machine interactions, which is a challenge in this line because the visualizing patterns on portable devices are not easy like classification, association, clustering and etc. This small display is holding a serious challenge for data mining interactive atmosphere. The next level in the world on platform of UDM is the advance data analysis to extract constructive acquaintance.

3. Multimedia data mining: The latest system to catch the data accurately involves multimedia data mining, which is useful to extract the information's from curious type of multimedia sources like videos, images, hypertexts and audio texts etc. the hyper text and hyper media mining is relevant to this platform. Although this platform is new but it' having a much more for the future. The major issue in this platform includes content-based retrieval, multidimensional analysis, classification and prediction analysis and association mining in multimedia data. Scholars has mentioned the few approaches to deal with but the most ordinary is to generate the cubes of multimedia data through which the data can be transformed in to various forms and analysis approach can be occur towards the techniques of data mining, it includes the gauge and facet of texts, shapes and other attributes. Beside this another devoting part is acoustic mining, which includes music mining. The main lead of this mining line is while using techniques it ask users to deliberate on examine patterns, it's possible to alter the patterns in to sounds, Itches, rhymes, melody and tunes to find some unusual and hidden things.
4. Spatial mining: Extracting hidden information's from the geographical data is important and seems to be common in future activities.
5. Time series data mining: This method is most common for the retail business markets to analyse the behaviour of customers and also accesses the buying patterns of customers deeply. By the study, we come across that this type of mining is useful in seasonal and time cycling rends. By practicing, it has ability to discover even random events, which stands outside the normal time series.

CONCLUSION

In conclusion, this paper approaches to the data mining domination towards the future in aspect of businesses industries, where this technology is playing a vital role for business growth. This study talks about the data mining features, work process flow, history dimensions of business in respect to data mining and why this technology is so important towards the business and easy scale for the customers and clients and how this is dominating the future growth.

This study is the combination of several past studies and analysed them on the single platform for the better understanding of data mining in a single eye.

Acknowledgement

We adore prompting our appreciations to "Yanshan University, Qinhuangdao, Hebei, China" for accompanying us in this research.

Author's Contribution

Noman Sohail has written the main manuscript under the supervision of Prof. Ren Jiadong. Uba, Irshad, Musavir, Abir, Wasim, and Usman has contributed in the literature survey process and key finding with guidance of Noman Sohail and Ren Jiadong. Tahir and Anthony have contributed in the design, proof reading and revision process with Noman Sohail. Finally, all authors have contributed equally in this research.

Conflict of interest statement

All authors state that they have no conflict of interest to declare.

Informed consent statement

N/A

Human and Animal Rights statement
N/A

References

1. Frawley, W. J., Piatetsky-Shapiro, G. & Matheus, C. J. Knowledge Discovery in Databases: An Overview. *AI Mag.*13, 57-57 (1992).
2. Berry, M. J. A., Linoff, G. & Berry, M. J. A. *Mastering data mining : the art and science of customer relationship management*. (Wiley Computer Pub, 2000).
3. Chien, C.-F. & Chen, L.-F. Data mining to improve personnel selection and enhance human capital: A case study in high-technology industry. *Expert Syst. Appl.*34, 280-290 (2008).
4. Wang, Y., Kung, L., Wang, W. Y. C. & Cegielski, C. G. An integrated big data analytics-enabled transformation model: Application to health care. *Inf. Manag.*55, 64-79 (2018).
5. Fayyad, U., Piatetsky-Shapiro, G. & Smyth, P. The KDD process for extracting useful knowledge from volumes of data. *Commun. ACM*39, 27-34 (1996).
6. Rygielski, C., Wang, J.-C. & Yen, D. C. Data mining techniques for customer relationship management. *Technol. Soc.*24, 483-502 (2002).
7. Prabakaran, S. & Mitra, S. Survey of Analysis of Crime Detection Techniques Using Data Mining and Machine Learning. *J. Phys. Conf. Ser.*1000, 012046 (2018).
8. Sohail, M. N., Jiadong, R., Uba, M. M. & Irshad, M. A Comprehensive Looks at Data Mining Techniques Contributing to Medical Data Growth: A Survey of Researcher Reviews. in 21-26 (Springer, Singapore, 2019). doi:10.1007/978-981-10-8944-2_3
9. Pangrazio, L. & Selwyn, N. "It's Not Like It's Life or Death or Whatever": Young People's Understandings of Social Media Data. *Soc. Media + Soc.*4, 205630511878780 (2018).
10. Grover, V., Chiang, R. H. L., Liang, T.-P. & Zhang, D. Creating Strategic Business Value from Big Data Analytics: A Research Framework. *J. Manag. Inf. Syst.*35, 388-423 (2018).
11. Alonso, S. G. *et al.* Data Mining Algorithms and Techniques in Mental Health: A Systematic Review. *J. Med. Syst.*42, 161 (2018).
12. Silva Junior, C. A. da *et al.* Object-based image analysis supported by data mining to discriminate large areas of

soybean. *Int. J. Digit. Earth* 1-23 (2018). doi:10.1080/17538947.2017.1421722
13. Prabhu, N. R., Andro-Vasko, J., Bein, D. & Bein, W. Music Genre Classification Using Data Mining and Machine Learning. in *Information technology-Generations* 397-403 (Springer, Cham, 2018). doi:10.1007/978-3-319-77028-4_52
14. Nakasima-López, S., Sanchez, M. A. & Castro, J. R. Big Data and Computational Intelligence: Background, Trends, Challenges, and Opportunities. in *Theory and Applications* 183-196 (Springer, Cham, 2018). doi:10.1007/978-3-319-74060-7_10
15. Irshad, M. *et al.* Li-local: green communication modulations for indoor localization. in *Proceedings of the 2nd International Conference on Future Networks and Distributed Systems - ICFNDS '18* 1-6 (ACM Press, 2018). doi:10.1145/3231053.3231118
16. M U Muhammad, O E Asiribo & Sohail Muhammad Noman. Application of Logistic Regression Modeling Using Fractional Polynomials of Grouped Continuous Covariates. *Niger. Stat. Soc. Ed. Proc. 1st Int. Conf.*1, 144-147 (2017).
17. Zhu, Y., Imamura, M., Nikovski, D. & Keogh, E. Introducing time series chains: a new primitive for time series data mining. *Knowl. Inf. Syst.* 1-27 (2018). doi:10.1007/s10115-018-1224-8
18. Ristoski, P. Exploiting semantic web knowledge graphs indata mining. (Universität Mannheim, 2018).
19. De Santis, F. & Presti, C. The relationship between intellectual capital and big data: a review. *Meditari Account. Res.*26, 361-380 (2018).
20. Najjar, A., Reinharz, D., Girouard, C. & Gagné, C. A two-step approach for mining patient treatment pathways in administrative healthcare databases. *Artif. Intell. Med.*87, 34-48 (2018).
21. Chow, T. E., Schuermann, R. T., Ngu, A. H. & Dahal, K. R. Spatial mining of migration patterns from web demographics. *Int. J. Geogr. Inf. Sci.*32, 1977-1998 (2018).
22. Arias, M., Saavedra, R., Marques, M. R., Munoz-Gama, J. & Sepúlveda, M. Human resource allocation in business process management and process mining. *Manag. Decis.*56, 376-405 (2018).
23. Wordsworth, S. *et al.* Using "Big Data" in the Cost-Effectiveness Analysis of Next-Generation Sequencing Technologies: Challenges and Potential Solutions. *Value Heal.*21, 1048-1053 (2018).
24. Ranjan, R. *et al.* The Next Grand Challenges: Integrating the Internet of Things and Data Science. *IEEE Cloud Comput.*5, 12-26 (2018).
25. Usai, A., Pironti, M., Mital, M. & Aouina Mejri, C. Knowledge discovery out of text data: a systematic review via text mining. *J. Knowl. Manag.*22, 1471-1488 (2018).
26. Yongmin Quan, Z. J. & Jin*, M. Application Research of Text Mining in Big Data Environment about North Korean News Public Opinion Analysis. *researchpublication*3, No 70 to 78-No 70 to 78 (2018).
27. Carneiro, A. How does knowledge management influence innovation and competitiveness? *J. Knowl. Manag.*4, 87-98 (2000).
28. Hashem, I. A. T. *et al.* The rise of "big data" on cloud computing: Review and open research issues. *Inf. Syst.*47, 98-115 (2015).
29. Marston, S., Li, Z., Bandyopadhyay, S., Zhang, J. & Ghalsasi, A. Cloud computing-The business perspective. *Decis. Support Syst.*51, 176-189 (2011).
30. Chen, H., Chiang, R. H. L. & Storey, V. C. Business Intelligence and Analytics: From Big Data to Big Impact. *MIS Quarterly*36, 1165-1188 (2012).
31. Li, L., Chi, T., Hao, T. & Yu, T. Customer demand analysis of the electronic commerce supply chain using Big Data. *Ann. Oper. Res.*268, 113-128 (2018).
32. Sharma, N. & Shamkuwar, M. Big Data Analysis in Cloud and Machine Learning. in 51-85 (Springer, Singapore, 2019). doi:10.1007/978-981-13-0550-4_3
33. Yaqoob, I. *et al.* Big data: From beginning to future. *Int. J. Inf. Manage.*36, 1231-1247 (2016).
34. Chen, G. *et al.* From Learners to Earners: Enabling MOOC Learners to Apply Their Skills and Earn Money in an Online Market Place. *IEEE Trans. Learn. Technol.*11, 264-274 (2018).
35. Dencik, L., Hintz, A. & Carey, Z. Prediction, pre-emption and limits to dissent: Social media and big data uses for policing protests in the United Kingdom. *New Media Soc.*20, 1433-1450 (2018).
36. Breiger, R. L., Wagner-Pacifici, R. & Mohr, J. W. Capturing distinctions while mining text data: Toward low-tech formalization for text analysis. *Poetics*68, 104-119 (2018).
37. Dimitrova, A. & Brkan, M. Balancing National Security and Data Protection: The Role of EU and US Policy-Makers and Courts before and after the NSA Affair. *JCMS J. Common Mark. Stud.*56, 751-767 (2018).
38. Romero, P., Obradovic, Z. & Dunker, A. K. Sequence Data Analysis for Long Disordered Regions Prediction in the Calcineurin Family. *Genome Informatics*8, 110-124 (1997).
39. Brown, A. W., Kaiser, K. A. & Allison, D. B. Issues with data and analyses: Errors, underlying themes, and potential solutions. *Proc. Natl. Acad. Sci. U. S. A.*115, 2563-2570 (2018).
40. Chamoso, P., González-Briones, A., Rodríguez, S. & Corchado, J. M. Tendencies of Technologies and Platforms in Smart Cities: A State-of-the-Art Review. *Wirel. Commun. Mob. Comput.*2018, 1-17 (2018).
41. Tsironis, L. K. Quality improvement calls data mining: the case of the seven new quality tools. *Benchmarking An Int. J.*25, 47-75 (2018).
42. Hasani, H., Jalali, S. M. J., Rezaei, D. & Maleki, M. A data mining framework for classification of organisational performance based on rough set theory. *Asian J. Manag. Sci. Appl.*3, 156 (2018).
43. Li, R. *et al.* Meningococcal conjugate vaccine safety surveillance in the Vaccine Safety Datalink using a tree-temporal scan data mining method. *Pharmacoepidemiol. Drug Saf.*27, 391-397 (2018).
44. Niazalizadeh Moghadam, A. & Ravanmehr, R. Multi-agent distributed data mining approach for classifying meteorology data: case study on Iran's synoptic weather stations. *Int. J. Environ. Sci. Technol.*15, 149-158 (2018).