Information and Knowledge Management
ISSN 2224-5758 (Paper) ISSN 2224-896X (Online)
Vol.4, No.1, 2014
www.iiste.org

Creating Business Intelligence through Machine Learning: An Effective Business Decision Making Tool

Yasir Shafi Reshi* Rafi Ahmad Khan
The Business School, University of Kashmir, Srinagar, Kashmir-190006 (India)
* E-mail of the corresponding author: yasirreshi@gmail.com

Abstract
Growing technological progressions have given rise to many issues concerning the contemporary decision making in business, which is a difficult phenomenon without Business Intelligence/ Machine Learning. The linking of machine learning with business intelligence is not only pivotal for business decision making but also for the business intelligence in totality, owing to the reason that in absence of machine learning, decision making couldn't take place efficaciously. Machines need to learn, re-learn, and then only they can help your learning process. The below paper seeks to make this concept simple/ easy by removing the ambiguities using a general framework. In order to prove the impact of machine learning on business intelligence, we need to forecast the trends, what is going around the world – business has to stay updated, then only it can be a successful endeavour. The paper posits the basic theories and definitions of business intelligence and machine learning. To learn from the past and forecast the future trends, many companies are adopting business intelligence tools and systems. Companies have understood the brilliance of enforcing achievements of the goals defined by their business strategies through business intelligence concepts and with the help of machine learning. It describes the insights on the role and requirement of real time BI by examining the business needs.
Keywords: Business Intelligence (BI); Machine Learning (ML); Artificial Neural Networks (ANN); Self-Organizing Maps (SOM); Data Mining (DM); Data Warehousing (DW).

1. Introduction

Business Intelligence is defined as the process that transforms data into information and then into knowledge. Data are raw facts and figures, information is processed data and knowledge is information analysed, understood and explained. BI deals with gathering high quality and meaning information about the subject matter being researched that will help individuals to analyse the information, draw conclusions and make assumptions. Business Intelligence is an umbrella term that combines architectures, tools, databases, analytical tools, applications, and methodologies (Turban, 2012). BI is the use of high level software intelligence for business applications. The advances which have taken place in the past decade in processing power, connectivity and intelligent technologies have generated an interest in the use of software technologies for business. Customers increasingly interact with the business systems where the high level of effective intelligence affects them directly. In addition to this business systems are increasingly interacting with other business systems. In nut shell the objective of BI is to improve the quality of information and enable managers to understand position of their firms as compare to that of their competitors in a better way.

Business intelligence systems combine operational and historical data with analytical tools to present valuable and competitive information to business planners and decision makers. Business intelligence applications and technologies can help companies to analyse changing trends in market share; changes in customer behaviour and spending patterns; customers' preferences; company capabilities; and market conditions. While the business world is rapidly changing and the business processes are becoming more and more complex making it more difficult for managers to have comprehensive understanding of business environment. The factors of globalization, deregulation, mergers and acquisitions, competition and technological innovation, have forced companies to re-think their business strategies and many large companies have resorted to Business Intelligence (BI) techniques to help them understand and control business processes to gain competitive advantage.

Learning is knowledge or skill that is acquired by instruction or study. It is a process of self-improvement and thus an important feature of intelligent behaviour. Learning is acquiring new or modifying existing knowledge, behaviours, skills, values or preferences and may involve synthesizing different types of information. The cognitive process of acquiring skill or knowledge, human Learning is a combination of many complicated cognitive processes, including induction, deduction, analogy and other special procedures related to observing and analysing situations. (Turban, 2012)

Machine Learning refers to techniques which allow an algorithm to modify itself based on observing its performance such that its performance increases. Machine learning techniques enable computers to acquire knowledge (i.e. learn) from data that reflects the historical happenings. They overcome deficiencies of manual knowledge acquisition techniques by automating the learning process (Turban, 2012). Machine Learning is a part of an emerging Artificial Intelligence technology that over the past few years has been employed by an increasing number of disciplines to automate complex decision making and problem solving tasks. ML is a family of methods that attempt to allow machines to acquire knowledge for problem solving by showing them historical cases. Among the various methods available, Artificial Neural Network (ANN) is the most popular which has been inspired by the biological neural networks of the human brain and started as an attempt to model the learning capabilities of humans. Other techniques include inductive learning, case-based reasoning, genetic algorithms, NLP etc.

Learning forms the base for any sort of intelligence – without appropriate learning infusion, a person cannot possess the genuine skills of intelligence e.g. A child is first taught how to write a letter ‘A’ then only he/ she can apply it, to form a word like A for Apple, in the same way machine learning is the first step, which is subservient to Business intelligence. Without comprehending the nitty-gritty’s of machine learning, business intelligence will be fragile.

2. Review of Literature

2.1 Business Intelligence:

BI is rooted DSS discipline. It has suffered considerable evolution and nowadays is an area of DSS that attracts a great deal of interest from both industry as well as researchers. The first reference to BI was made by (Lunh, 1958), Where other terms such as Executive Information System (EIS) and Management Information System (MIS) were replaced. In 1970’s MIS reporting system was used. However it lacked analytical capabilities. In 1980’s the concept of EIS evolved. This concept expanded the computerized support to top level managers and executives. Some of the capabilities introduced were forecasting, prediction, trend analysis etc. Then same capabilities and some new ones like OLAP, data warehousing, data mining, appeared under the name BI. The literature review on BI reveals few empirical studies. Most of the articles are conceptual. In addition, we recognize the traditional “separation” between technical and managerial aspects, outlining two broad patterns (Table 1).

Managerial Approach	Technological Approach
Focus on the process of gathering data from Internal and external sources and of analysing themin order to generate relevant information (Liautaud & Hammond, 2000)	Focus on the technological tools that support the process (Kudyba & Hoptroff, 2001)

Table 1: Two approaches to BI

From the managerial approach, BI is seen as a process in which data from inside and outside the company are integrated in order to generate information relevant to the decision making process. The role of BI here is to create an informational environment and process by which operational data gathered from transactional systems and external sources can be analysed and to reveal the “strategic” business dimensions. From this perspective emerge concepts such as “intelligent company”: one that uses BI to make faster and smarter decisions than its competitors (Liautaud & Hammond, 2000). “Intelligence” means reducing a huge volume of data into knowledge through a process of filtering, analysing and reporting information.

The technological approach presents BI as a set of tools that supports the storage and analysis of information. The focus is not on the process itself, but on the technologies that allow the recording, recovering, manipulation and analysis of information. For instance, (Kudyba & Hoptroff, 2001)understand BI as data warehousing (DW); (Scoggins, 1999)classifies data mining (DM) as a BI technique; Hackathorn includes all resources (DW, DM, hypertext analysis and web information) in the creation of a BI system; and finally, linking BI and the Internet, (Giovinazzo, 2002)positioned the integration of DW and customer relationship management (CRM) applications.

Whether managerial or technological, there is a shared idea among all these studies:

- ✓ The core of BI is information gathering, analysis and use.
- ✓ The goal is to support the decision making process, helping the company’s strategy.

BI is an area of Decision Support System (DSS), which is an information system that can be used to support complex decision making, and solving complex, semi-structured, or ill-structured problems (Azevedo & Santos, 2009) . (Golfarelli, Rizzi, & Cella, 2004) Argue that BI is the process that transforms data into information and then into knowledge. (Stackowiak, Rayman, & Greenwald, 2007) Opine that BI is the process of taking large amounts of data, analysing that data, and presenting a high-level set of reports that condense the essence of that data into the basis of business actions, enabling management to make fundamental daily business decisions. (Cui, Damiani, & Leida, 2007) Argue that BI is the way and method of improving business performance by providing powerful assistance to executive decision maker which enables them to have actionable information at hand. BI tools are viewed as technology that enhances the efficiency of business operation by providing an increased value to the enterprise information and hence the way this information is utilized. (Zeng, Xu, Shi, Wang, & Wu, 2007)has put forward the concept of BI as “The process of collection, treatment and diffusion of information, which has an objective, the reduction of uncertainty in the making of all strategic decisions.” While (Wu, Barash, & Bartolini, 2007) argues that BI is a “business management term used to describe applications and technologies which are used to gather, provide access to analysed data and information about an enterprise, in order to help them make better informed business decisions.” (Van Drunen, 1999)has considered Business Intelligence as different as its predecessor, "decision support", in that it is a strategic tool intended to help with planning and performance measurement, rather than with partly operational decisions. BI is the ability of an organization to understand and use information to its gainful operation (Radonic & Curko, 2000) the Enterprise BI is a way that brings synergies to business processes and new efficiencies across business areas (Liautaud & Hammond, 2000).

Taking into account the scarce literature, we looked for other areas that could help us reach a more comprehensive understanding of BI. We find contributions in three distinct topics: information planning, balanced score card and competitive intelligence.

Authors writing about information planning emphasize the importance of identifying few but strategic information (Reich & Benbasat, 2000), which posits a kind of paradox regarding the overload of information we have today: the problem of BI is exactly to reduce quantity into quality. In addition, information relevant to decision making is likely to exist inside the company already or to be clearly defined in the managers’ minds. In a similar vein we have the well-known critical factors of success (CFS) method by which executive objectives, indicators, measures and reports are identified and selected by means of a sequence of interviews with top management (Rockart, 1979)Both areas – information planning and CFS - promote a realist ontology that denies the socially constructed and political process of information “producing” in any organization.

The concept of balanced scorecard (BSC) also offers some contribution to our study of BI as it associates indicators and measures (information) to the monitoring of strategic objectives of the company (Zee & Jong, 1999). BSC can be seen as a set of measures that provide a fast and understandable view of the business to high-level executives (Kaplan & Norton, 1992) its development was motivated by the dissatisfaction with the traditional performance measures that were only concerned with financial metrics and focused in the past and not the future. As a result, the financial measures are complemented with operational, internal process and innovative and organizational learning measures. The main contribution of BSC to our study is this idea of multiple perspectives. Regarding the context of developing countries, such as of Brazilian companies, what alternative and particular perspectives could we aggregate?

Finally, we borrowed insights from a neighbouring area, competitive intelligence (Miller, 2002). It is about “beating your competitors both here and abroad and never being surprised. We found here the same concern with the process of data collection and analysis, and with the distinction between information and intelligence we have found in BI definitions. Information is factual, intelligence is something that can be acted upon; both are contextual.

After analysing all the ideas described above (clearly dominated by a positivistic standpoint) and trying to develop a more critical appreciation, we put forward a distinct definition of BI: a collective and socially constructed process of information gathering, analysis and dissemination where the information is few but strategic, belongs to multiple perspectives, originates internally and externally, and is contextualized. In addition, we believe that such a process should be reflexive, i.e., organizational members should be critical regarding their role, the role of their company in their context and critical regarding the information they help to produce.

2.2 MACHINE LEARNING:

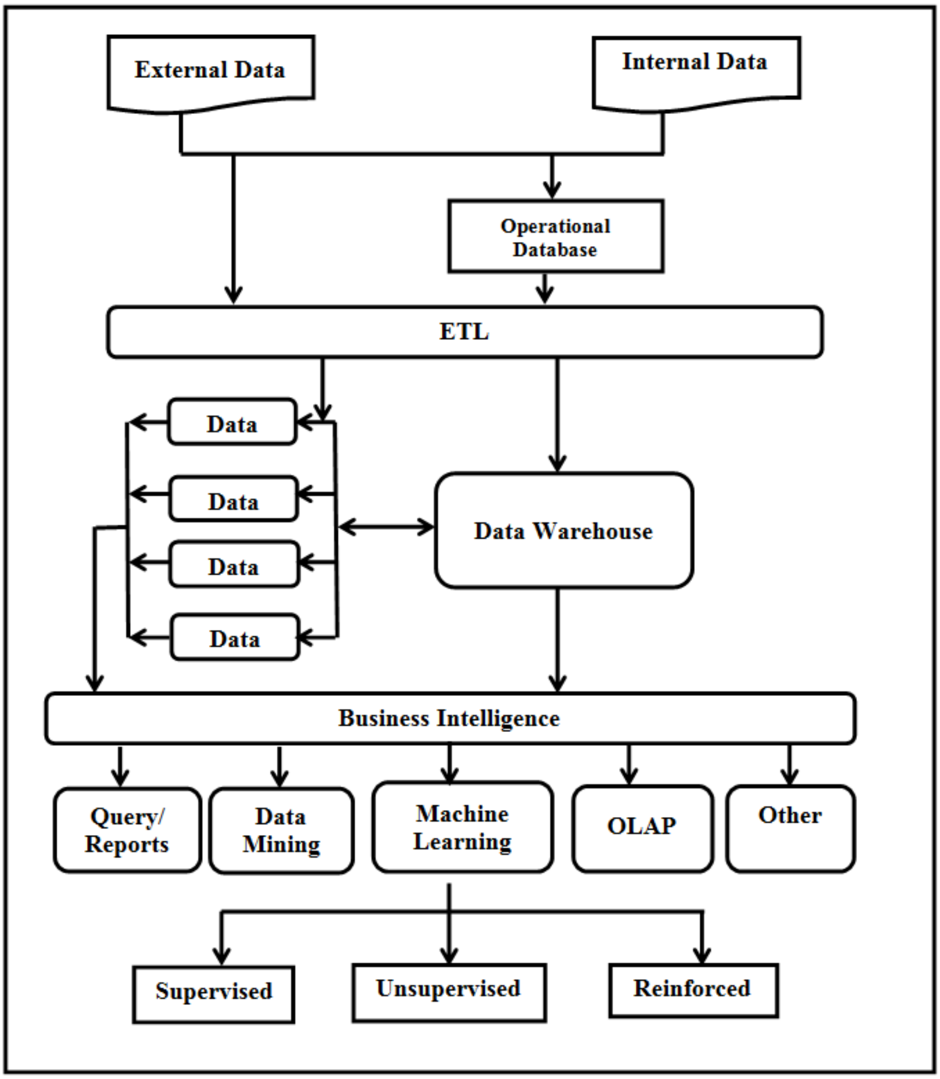
1950’s - 1960’s: An exploratory period when many general techniques were born, early research was inspired by biology and psychology, (Rosenblatt, 1957) Developed the “perceptron” which was modelled after neurons. It was the precursor to later work in neural networks. In (Minsky & Papert, 1969)a seminal paper proving that the perceptron was inherently limited, which discouraged research in this area for almost 20 years. The History of machine learning dates back to the 1950's during the AI and cognitive science days.

1970’s: This period was characterized by the development of more practical algorithms, often using symbolic techniques. (Royce, 1970)Important work on learning in blocks-world domain, “knowledge acquisition bottleneck”, META-DENDRAL learned mass-spectrometry prediction rules.

1980’s: Explosion of different techniques & increased emphasis on evaluation, the notion of “version spaces” was developed by (Mitchell, 1982), and the concept of “inductive bias” (with Utgoff), (Quinlan, 1983) Created the ID3 decision tree algorithm. Valiant in 1984 defined the idea of “probably approximately correct” (PAC) learning, still the dominant theoretical paradigm. The back propagation algorithm (Rumelhart, 1986)overcame many of the limitations of the perceptron. Explanation-based learning, speedup learning, and case-based learning became very popular, continued progress on decision-tree and rule learning. Explanation-based learning and speedup learning are exemplified by utility problem, analogy, resurgence of connectionism (PDP, ANN), PAC

learning, experimental evaluation.
1990's: The machine learning (ML) community has become increasingly concerned with evaluation. ML systems must be formally evaluated to ensure that the results are meaningful. Community data sets have been made available for comparing algorithms. Data mining has emerged as an important application for many business activities. Some new directions included Statistical comparisons of algorithms, Theoretical analyses of algorithms, Successful applications, Multi-relational learning, Ensemble and Kernel Methods. Current Uses of Machine Learning: Reinforcement learning, Bayesian learning, automatic bias selection, inductive logic programming (mostly in Europe), and applications to robotics, adaptive software agents, voting, bagging, boosting, and stacking.
2000 & Beyond: It is an Interaction between symbolic machine learning, computational learning theory, neural networks, statistics, and pattern recognition. New applications for ML techniques: knowledge discovery in databases, language processing, robot control, combinatorial optimization. To improve accuracy by learning ensembles, Scaling up supervised learning algorithms, learning complex stochastic models (Hierarchical Mixture of Experts, Hidden Markov Model, and Dynamic Probabilistic Network).

3. Framework of Business Intelligence through Machine Learning
The below is the framework of Business Intelligence, showing various components/ tools which help in the acquisition of knowledge, pivotal for decision making vis-à-vis Machine Learning (Figure 01):



Data is collected from internal and external sources. Internal sources of data are organizations operational database and external source of data includes data from customers, suppliers, government agencies, competitors, internet etc. The collected data is stored in data warehouse after extract transform and load (ETL). Finally data stored in data warehouse is analyzed for decision making.

DATA ACQUISITION: It consists of systems that have interface with the operational systems to load data into data ware house. Data is first entered by a daily business process based on online transaction processing (OLTP) and is stored in operational database such as Oracle, DB2, Informix, SQL Server, etc. Before data is loaded from operational and external sources into data warehouse it undergoes following stages:

a) *Extraction and cleanse:* - data is acquired from multiple sources. The selected data is filtered out from various forms of pollution. Data cleansing validates and cleans up the extracted data to correct inconsistent, missing or invalid values.
b) *Transform:* - Here data is integrated into standard formats and business rules are applied that map data to the warehouse schema.
c) *Load:* - Data loading loads the cleansed data into the data warehouse.

DATA STORAGE: After ETL data is stored in data warehouse or data marts for future analysis.

a) *Data Warehousing:* - Data warehouse is a collection of corporate information derived directly from operational system and some external data sources (Liautaud & Hammond, 2000) (Lunh, 1958) (Negash, 2004). Its specific purpose is to support business decisions. Inmon who coined the term "data warehouse" in 1990, argues that a data warehouse is a subject oriented, integrated, time-variant, non-volatile collection of data that is used primarily in organizational decision making. Before designing data warehouse the requirements of an organization should be taken into consideration.
b) *Data Marts:* - These are also a kind of data warehouses but small in size typically created by individual departments to facilitate their own decision support activities. It's small and easy to develop. One of the purposes to build a data mart is to get prototype as soon as possible without waiting for a larger corporate data warehouse.
c) *Meta Data:* - Its data about data. To understand and locate data in the data ware house users need information about the data warehousing system and its contents. This information is known as Meta data and helps business users to understand what is available, how to access it and what it means.

DATA ACCESS AND ANALYSIS: It is referred to as front end. It consists of access tools and techniques that provide a user with direct, interactive, access to data and hides the technical complexity of data retrieval. It's friendly enough for use even by a non-technical person. This is accomplished by use of various tools

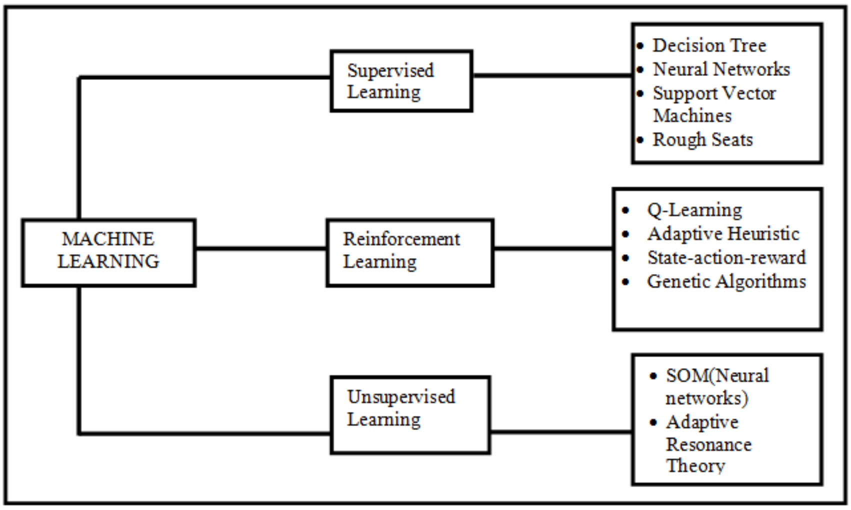
a) Sophisticated data analysis tools like OLAP.
b) Data mining tools.
c) Machine learning tools.

4. Categories/Classification of Machine Learning

Machine learning has three basic categories-Supervised learning, unsupervised learning and Reinforcement learning. The simplistic taxonomy of machine learning with exemplary methods listed under each category is as follows:-

6. Limitations of Text Mining

The fundamental limitations of text mining include:



4.1 Supervised Learning: It is a process of inducing knowledge from a set of observations whose outcomes are known.

Example: Say we induce a set of rules from historical loan evaluation data, because the decision on these loan cases are known, we can test how the induced model performs when it is applied to those historical cases.

Decision Tree: A decision tree outlines several options or choices, the probability outcomes and also costs. A decision tree is defined as " a predictive modelling technique from the field of machine learning and statistics that builds a simple tree-like structure to model the underlying pattern [of data]" (Prediction Works, 2011). Decision trees are one example of a classification algorithm. Classification is a data mining technique that assigns objects to one of several predefined categories (Tan, Steinbach, & Kumar, 2005).

Classification algorithms (also called classifiers) have helped solve problems ranging from credit card theft detection (Chan, Fan, Prodromidis, & Stolfo, 1999) to diagnosing patients with heart problems by recognizing distinctive patterns in a dataset and classifying activity based on this information. From an intrusion detection perspective, classification algorithms can characterize network data as malicious, benign, scanning, or any other category of interest using information like source/destination ports, IP addresses, and the number of bytes sent during a connection.

Neural Networks: The first artificial neuron was produced in 1943 by the neurophysiologist Warren McCulloch and the logician Walter Pits. But the technology available at that time did not allow them to do too much.

An Artificial Neural Network (ANN) is an information processing paradigm that is inspired by the way biological nervous systems, such as the brain, process information. The key element of this paradigm is the novel structure of the information processing system. It is composed of a large number of highly interconnected processing elements (neurones) working in unison to solve specific problems. ANNs, like people, learn by example. An ANN is configured for a specific application, such as pattern recognition or data classification, through a learning process. Learning in biological systems involves adjustments to the synaptic connections that exist between the neurones. This is true of ANNs as well. Neural networks, with their remarkable ability to derive meaning from complicated or imprecise data, can be used to extract patterns and detect trends that are too complex to be noticed by either humans or other computer techniques. A trained neural network can be thought of as an "expert" in the category of information it has been given to analyse. This expert can then be used to provide projections given new situations of interest and answer "what if" questions (Stergiou & Sigsnos, Neural Networks, 1996).

Support Vector Machine (SVM): Support Vector Machines are linear functions of the form $f(x) = (w \cdot x) + b$, where $(w \cdot x)$ is the inner product between the weight vector 'w' and the input vector 'x'. The SVM can be used as a classifier by setting the class to 1 if $f(x) > 0$ and to -1 otherwise. The main idea of SVM is to select a hyper-plane that separates the positive and negative examples while maximizing the minimum margin, where the margin for example x_i is $y_i f(x_i)$ and $y_i \in -\{ 1,1\}$ is the target output. This corresponds to minimizing $(w \cdot w)$ subject to $y_i[(w \cdot x) + b] \geq 1$ for alli. Large margin classifiers are known to have good generalization properties. (Cristianini & Shawe-Taylor, 2000)

4.2 Un-Supervised Learning: It is used to discover knowledge from a set of data whose outcomes are unknown. A typical application is to classify customers into several different profiles and lifestyles. Before the classification, we do not know how many different kinds of profiles and lifestyles are available, nor do we know which customer belongs to a particular profile or lifestyle.

Self-organizing map (SOM): The Self-Organizing map (SOM) is a widely used artificial neural network model. In the SOM, learning process is unsupervised: no a priori classifications for the input examples are needed. The learning process is based on similarity comparisons in a continuous space. The result is a system that associated similar inputs close to each other in the two-dimensional grid called the map. The input may be highly complex multidimensional numerical data. Recently, the SOM has also been used for the analysis and visualization of symbolic and text data (Honkela, Leinonen, Lonka, & Raike, 2000).

Adaptive Resonance Theory (ART): The Adaptive Resonance Theory (ART) was introduced as a theory for human cognitive information processing by Stephen Grossberg and Gail Carpenter. The theory has led to neural models for pattern recognition and unsupervised learning. These models are capable of learning stable recognition categories. ART systems have been used to explain a variety of cognitive and brain data.

The Adaptive Resonance Theory addresses the plasticity-elasticity dilemma of a system that asks how learning can be preceded in response to significant input patterns and at the same time not to lose the stability for irrelevant patterns. Besides that, the plasticity-elasticity dilemma is concerned about how a system can learn new information while keeping what was learned before. For such task, a feedback mechanism is added among the ART neural network layers. If an appropriate pattern is developed, the resonance is reached then, weight adaptation can happen during this period. (Vieira & Lee, 2007)

4.3 Reinforcement Learning: It is not as popular as the other two types due to the fact that it is not as mature and its applicability is limited to a small set of real world situations. This type of learning is successfully applied to learning to play backgammon, autonomous search robots, and controlling the flight of helicopters. This type of learning is also called trial-and-error learning.

Q-learning (Watkins, 1989)*:* It is a form of model-free reinforcement learning. It uses the information observed to approximate the optimal function, from which one can construct the optimal policy. The conditions have to do with the exploration policy and the learning rate. For the exploration one needs to require that each state action be performed infinitely often. The learning rate control shows fast we modify our estimates. One expects to start with a high learning rate, which allows fast changes, and lowers the learning rate as time progresses. The basic conditions are that the sum of the learning rates goes to infinity (so that any value could be reached) and that the sum of the squares of the learning rates is finite (which is required to show that the convergence is with probability one). (Even-Dar & Mansour, 2003)

Adaptive Heuristic: The adaptive heuristic critic algorithm is an adaptive version of policy iteration (Barto, Sutton, & Anderson, 1983) in which the value-function computation is no longer implemented by solving a set of linear equations, but is instead computed by an algorithm called temporal difference (TD) learning. TD is related to dynamic programming techniques because it approximates its current estimate based on previously learned estimate. (Kaelbling, Littman, & Moore, 1996)

SARSA (State-Action-Reward-State-Action): The SARSA algorithm was first explored by Rummery and Niranjan in 1994, who called it modified Q-learning. The name "SARSA" was introduced by Sutton in 1996. Like Q-learning, this method has been proved to converge to the optimal action values, provided that the actions asymptotically approach a greedy policy, because it eliminates the use of them min operator over the actions, this method is faster than Q-learning for situations where the action set has high cardinality. (Ribeiro, 2000)

Genetic Algorithm: It is a form of evolutionary computing which is a collective name for problem solving techniques based on the principles of biological evolution like natural selection. Genetic algorithm approach is based on three main processes- crossovers, mutation and selection of individuals. Initially many individual solutions are gathered together to make a randomly generated population. Genetic algorithms are based upon the Darwin theory of "The survival of the Fittest" depending upon the fitness function the best possible solutions are selected from the pool of individuals. The fitter individuals have greater chances of its selection and higher the probability that its genetic information will be passed over to future generations. Once selection is over new individuals have to be formed. These new individuals are formed either through crossover or mutation. In the process of crossover, combining the genetic makeup of two solution candidates (producing a child out of two parents) creates new individuals. Whereas in mutation, we alter some individuals, which means that some randomly chosen parts of genetic information is changed to obtain a new individual. The process of generation doesn't stop until one of the conditions like minimum criteria is met or the desired fitness level is attained or a specified number of generations are reached or any combination of the above. (Singh, Bhatia, & Sangwan, 2007)

5. Machine learning and its applications reformative in Business

One measure of progress in Machine Learning is its significant real-world applications, such as those listed below. Although we now take many of these applications for granted, it is worth noting that as late as 1985 there were almost no commercial applications of machine learning. The following are the applications of machine learning:

Speech recognition: Currently available commercial systems for speech recognition all use machine learning in one fashion or another to train the system to recognize speech. The reason is simple: the speech recognition accuracy is greater if one trains the system, than if one attempts to program it by hand. In fact, many commercial speech recognition systems involve two distinct learning phases: one before the software is shipped (training the general system in a speaker-independent fashion), and a second phase after the user purchases the software (to achieve greater accuracy by training in a speaker-dependent fashion).

Computer vision: Many current vision systems, from face recognition systems, to systems that automatically classify microscope images of cells, are developed using machine learning, again because the resulting systems are more accurate than hand-crafted programs. One massive-scale application of computer vision trained using machine learning is its use by the US Post Office to automatically sort letters containing handwritten addresses. Over 85% of handwritten mail in the US is sorted automatically, using handwriting analysis software trained to very high accuracy using machine learning over a very large data set.

Bio-surveillance: A variety of government efforts to detect and track disease outbreaks now use machine learning. For example, the RODS project involves real-time collection of admissions reports to emergency rooms across western Pennsylvania, and the use of machine learning software to learn the profile of typical admissions so that it can detect anomalous patterns of symptoms and their geographical distribution. Current work involves adding in a rich set of additional data, such as retail purchases of over-the-counter medicines to increase the information flow into the system, further increasing the need for automated learning methods given this even more complex data set.

Robot control: Machine learning methods have been successfully used in a number of robot systems.

For example, several researchers have demonstrated the use of machine learning to acquire control strategies for stable helicopter flight and helicopter aerobatics. The recent Darpa-sponsored competition involving a robot driving autonomously for over 100 miles in the desert was won by a robot that used machine learning to refine its ability to detect distant objects (training itself from self-collected data consisting of terrain seen initially in the distance, and seen later up close).

Accelerating empirical sciences: Many data-intensive sciences now make use of machine learning methods to aid in the scientific discovery process. Machine learning is being used to learn models of gene expression in the cell from high-throughput data, to discover unusual astronomical objects from massive data collected by the Sloan sky survey, and to characterize the complex patterns of brain activation that indicate different cognitive states of people in MRI scanners. Machine learning methods are reshaping the practice of many data-intensive empirical sciences, and many of these Sciences now hold workshops on machine learning as part of their field’s conferences.

6. Benefits of BI – When Dovetailed With Machine Learning

Because of the wide applicability of Business Intelligence in both enterprise and extranet deployments, the business benefits are numerous.

Eliminate guesswork: "Running a business shouldn't be like gambling," said Ken Dixon, executive vice president of marketing at Kogent Corporation. Business intelligence provide more accurate historical data, real-time updates, synthesis between departmental data stores, forecasting and trending, and even predictive 'what if?' analysis," thus eliminating the need to guesstimate.

Make you aware of emerging crises: BI software can pace up decision-makers with on-going operations via regular, scheduled reports. In addition, BI software can be configured to send automated alerts to appropriate decision-makers to notify them of incidents that require action – such as warehouse stock quantities dropping below a predefined level**.**

Get key business metrics reports when and where you need them: Today, many business intelligence software vendors are making it possible for users to access key business metrics, reports and dashboards on mobiles devices like their iPhone, iPad, Android or BlackBerry, giving sales and marketing people access to critical business information on the fly.

Get insight into customer behaviour: Identify, track and monitor customer purchase habits to effectively segment your current and future customer base.

- ✓ Help you identify and capitalize on up-sell and cross-sell opportunities
- ✓ Allow you to use customer data to meet and exceed customer expectations

Improving customer satisfaction: With access to information, users can make better decisions faster without having to escalate standard problems up the management hierarchy. This guarantees pragmatic and effective solutions since the people directly involved in the operations make decisions. In addition, users have the increased satisfaction of controlling their own process.

Ingram Micro: This wholesale provider of high-tech goods to technology solutions providers is working to create a new BI extranet in order to deliver advanced information to the company’s suppliers and business partners. Says Ingram Micro CIO Guy Abramo, “Today it’s incumbent on us to provide our partners with sell-through information so that they can see what happened once their PCs hit distribution. That’s critical for them to do inventory planning and manufacturing planning—helping them to understand what products are selling to what segments of the marketplace.”

Learn how to streamline operations: With detailed insights into business performance, organizations can quickly, easily and effectively identify and address areas of operational inefficiency.

Learn what true manufacturing costs are: Business intelligence software can give users greater insight into manufacturing costs and the ability to adjust production on the fly for greater profitability.

Increasing revenue: Leading organizations are using BI to differentiate their product and service offerings from competitors through value added, web-based services. In the past, many departments generated zero revenue, but now with BI extranets, they create a recurring revenue stream by selling information to customers, partners, and

suppliers.
Owens & Minor: The $3 billion medical supplies distributor has signed up 80 hospital accounts and six of its top suppliers, including pharmaceutical giant Johnson & Johnson, for the service. Hospitals pay up to $1,250 per month and suppliers pay $2,000 per month. It is estimated that Owens & Minor will generate at least $2 million in fees next year.
Manage inventory better: Business intelligence software can help you "order the right amount of inventory at the right time so that customers receive their products when they need them," and your business doesn't bear the cost of stocking excess inventory, noted John Krech, the president and founder of e-Phi phony.
See where your business has been, where it is now and where it is going: "Business Intelligence has been very successful at explaining what happened to the business over some defined period of time — for example, how many units were sold, through which store, in which geography, or by which customer segment," said Sid Probstein, CTO of Attivio. "These issues are now well understood, and so the next generation of competitive advantage comes from analysing unstructured content to understand how and why these things happen. From simple content-based metrics to sophisticated sentiment analysis, BI provide a more complete view on customer and competitor experience and opportunities therein," and help executives plan for the future.

7. Conclusion

The above paper extensively dealt with the general framework of business intelligence in relation to machine learning, how we can improve our decision making through machine learning and unless and until machine learns, we cannot percolate the advantages down to business personnel. The general framework of business intelligence made it clear for all, that how with the help of machine learning as a tool of Business Intelligence, decisions could be reached upon. Furthermore, the exhaustive framework of machine learning, lucidly explaining the sources from which machine learns, improves and digitizes the entire scenario is in itself presenting a detailed picture of all the sources of machine learning, which indirectly ameliorate the process of business decision making. A lot needs to be done in this scenario, the loopholes are evident in many aspects, but we need to improve upon those ambiguities and come up with a more polished system of decision making, through an effective and efficient use of Machine Learning. In today's highly competitive world, the quality and timeliness of business information for an organization is not just a choice between the benefit and burden; it may be a question of endurance or failure. No business organization can deny the benefit of machine learning and business intelligence. The rapidly changing business environment will increase the need for machine learning as a tool of business intelligence. In this paper an attempt has been made to educate the technological personnel's and academics about the unnerving development and application of business intelligence through machine learning.

Bibliography

Abiteboul, S. B. (1999). *Data on the Web: from relations to semistructured data and XML.* San Francisco,USA: Morgan Kaufmann Publishers Inc.

Alkin, M. (2007). Text Mining: A Burgeonining Quality Improvement Tool. *Ph.D. Proposal, Department of the Computer Sciences, University of Texas at Austin,* .

Azevedo, A., & Santos, M. F. (2009). Business Intelligence: State of the Art, Trends, and Open Issues. *Knowledge Management and Information Sharing* , 296-300.

Barto, A., Sutton, R., & Anderson, C. W. (1983). Neuronlike elements that can solve difficult learning control problems. *IEEE Transactions on Systems, Man, and Cybernetics* , 835-846.

Blumberg, R. A. (2003). *The Problem with Unstructured Data.* Information Management Magazine.

Chan, P., Fan, W., Prodromidis, A., & Stolfo, S. (1999). Distributed data mining in credit card fraud detection. Intelligent Systems and Their Applications. *IEEE* , 67-74.

Chen, H., & Chau, M. (2004). Web Mining: Machine Learning for Web Applications. *Annual Review of Information Science and Technology* , 289-329.

Cristianini, N., & Shawe-Taylor, J. (2000). *An Introduction to Support Vector Machines.* Cambridge: Cambridge University Press.

Cui, Z., Damiani, E., & Leida, M. (2007). Benefits of Ontologies in Real Time Data Access. *Digital EcoSystems and Technologies Conference. Inaugural IEEE-IES* (pp. 392-397). Cairns: IEEE Conference Publications.

Dixon, N. M. (2000). *Common Knowledge-How Companies Thrive by Sharing What They Know.* Harvard Business School Press.

Even-Dar, E., & Mansour, Y. (2003). Learning Rates for Q-learning. *Journal of Machine Learning Research* , 1-25.

Fayyad, U., Piatetsky-Shapiro, G., & Smyth, P. (1996). *From Data Mining to Knowledge Discovery in Databases.* Cambridge Mass: MIT Press/AAAI Press.

Frawley, W. J., Piatetsky-Shapiro, G., & Matheus, C. J. (1992). Knowlegde Discovery in Databases. *An Overview: in Knowledge Discovery in Databases* , 1-27.

Gelfand, B., Wulfekuhler, M., & Punch, W. F. (1998). Automated Concept Extraction from Plain Text.

Giovinazzo, W. (2002). *Internet-Enabled Business Intelligence.* Prentice Hall.

Golfarelli, M., Rizzi, S., & Cella, I. (2004). What`s Next in Business Intelligence. *CIKM-Conference on Information and Knowledge Management* (pp. 1-6). New York: ACM New York.

Guernsey, L. (2003). *Digging for Nuggets of Wisdom.* The New York Times.

Gupta, V., & Lehal, G. S. (2009). A Survey of Text Mining Techniques and Applications. *Journal of Emerging Technologies in Web Intelligence* .

Harmon, R. R. (2003). Marketing Information Systems. *Encyclopedia of Information Systems* , 137-151.

Hearst, M. A. (1999). Untangling Text Data Mining. *In Proceedings of ACL'99: the 37th Annual Meeting of the Association for Computational Linguistics, University of Maryland* .

Honkela, T., Leinonen, T., Lonka, K., & Raike, A. (2000). Self-Organizing Maps and Constructive Learning. *ICEUT* (pp. 339-343). Beijing: International Federation for Information Processing (IFIP).

Kaelbling, L. P., Littman, M. L., & Moore, A. W. (1996). Reinforcement Learning: A Survey. *Journal of Artificial Intelligence Research* , 237-285.

Kaplan, R. S., & Norton, a. D. (1992). The Balanced Scorecard Measures that Drive Performance. *Harvard Business Review* .

Kelley, J. (2002). *Knowledge Nirvana-Achieving The Competitive Advantage Through Enterprise Content Management and Optimizing Team Collaboration.* Xulon Press.

Kudyba, S., & Hoptroff, R. (2001). *Data Mining and Business Intelligence: A Guide to Productivity*. Hershey: Idea Group Publishing.

Liautaud, B., & Hammond, M. (2000). *E-Business Intelligence turning information into knowledge into profit.* New York: McGraw Hill.

Lunh, H. P. (1958). A Business Intelligence System. *IBM Journal of Research and Development* , 314–319.

McDonald, D., & Kelly, U. (2012). The Value and Benefit of Text Mining. *Joint Information systems Committee* .

Miller, J. G. (2002). Bringing culture to basic psychological theory--Beyond individualism and collectivism. *Psychological Bulletin* , 92-109.

Minsky, M., & Papert, S. A. (1969). *An Introduction to Computational Geometry.* Cambridge: MIT Press.

Mitchell, T. R. (1982). Motivation: New directions for theory and research. *Academy of Management Review* , 80-88.

Mooney, R. J., & Bunescu, R. (2005). Mining Knowledge from Text Using Information Extraction. *Special Interest Group on Knowledge Discovery and Data Mining (SIGKDD)* , 1-10.

Mustafa, A., Akbar, A., & Sultan, A. (2009). Knowledge Discovery using Text Mining: A Programmable Implementation on Information Extraction and Categorization. *International Journal of Multimedia and Ubiquitous Engineering* .

Nahm, U. Y., & Mooney, R. J. (2002). Text Mining with Information Extraction. *Spring Symposium on Mining Answers from Texts and Knowledge Bases* .

Nasukawa, T., & Nagano, T. (2001). Text analysis and knowledge Mining system. *IBM Systems Journal* .

Negash, S. (2004). Business Intelligence. *Communications of the Association for Information Systems* , 177-195.

P.Bhargavi, B.Jyothi, S.Jyothi, & K.Sekar. (2008). Knowledge Extraction Using Rule Based Decision Tree . *IJCSNS International Journal of Computer Science and Network Security* .

Phua, C., Lee, V., K.Smith, & Gayler, R. (2005). A Comprehensive Survey of Data Mining-based Fraud Detection Research. *Clayton School of Information Technology, Monash University* .

Quinlan, J. (1983). Learning efficient classification procedures and their application to chess end games. *Machine Learning: An Artificial Intelligence Approach* .

Radonic, D. G., & Curko, D. K. (2000). A Review Of Business Intelligence Approaches To Key Business Factors In Banking. *Journal of Knowledge Management Practice* .

Rajman, M., & Besançon, R. (1998). Text Mining - Knowledge extraction from unstructured textual data. *6th Conference of International Federation of Classification Societies-IFCS* .

Reich, B. H., & Benbasat, I. (2000). Factors that influence the social dimension of alignment between business and information technology objectives. *Management Information Systems Quarterly* , 81-113.

Ribeiro, C. H. (2000). A Tutorial on Reinforcement Learning Techniques. *Proceedings of International Conference on Neural Networks.* Washington, DC: INNS Press.

Rockart, J. (1979). Chief Executives Define Their Own Information Needs. *Harvard Business Review* , 81-93.

Rosenblatt, F. (1957). The Perceptron: A Probabilistic Model for Information Storage and Organization in the Brain. *Psychological Review* , 386-408.

Royce, W. (1970). Managing the Development of Large Software Systems. *Proceedings of IEEE WESCON* , 1-9.

Rumelhart, D. M. (1986). Parallel Distributed Processing: Explorations in the Microstructure of Cognition. *MIT Press* .

S.Nestorov, Abiteboul, S., & Motwani, R. (1998). Extracting Schema from Semistructured Data. *SIGMOD'99* .

Scoggins, J. (1999). *A Practitioner’s View Of Techniques Used In Data Warehousing For Sifting.* Kansas City: Proceedings Of The Eight International Conference On Information And Knowledege Management.

Singh, Y., Bhatia, P. K., & Sangwan, O. (2007). A Review of Studies on Machine Learning Techniques. *International Journal of Computer Science and Security* , 70-84.

Stackowiak, R., Rayman, J., & Greenwald, R. (2007). *Oracle Data Warehousing and Business Intelligence Solutions.* Indianapolis: Wiley Publishing Inc.

Stergiou, C., & Siganos, D. (2009). *Neural Networks.*

Stergiou, C., & Sigsnos, D. (1996). *Neural Networks.* London: Imperial College of Science and Technology.

Stewart, T. (2002). *The Wealth of Knowledge: Intellectual Capital and the Twenty-First Century Organization.* London: Nicholas Brealey Publishing.

Tan, A.-H. (2006). Text Mining:The state of the art and the challenges. *In Proceedings of the Pacific Asia COnference on Knowedge Discovery and Data Mining PAKDD'99 Workshop on Knowledge Discovery from Advanced Databases* , 65-70.

Tan, P., Steinbach, M., & Kumar, V. (2005). *Introduction to data mining* . Boston: Pearson Addison Wesley.

Tiwana, A. (2001). *The Essential Guide to Knowledge Management.* Prentice- Hall.

Turban, E. (2012). *Decision Support System and Business Intelligence.* Pearson Education.

Ultsch, A., & Korus, D. (1995). Automatic acquisition of symbolic knowledge from subsymbolic neural nets. *3rd European Congress on Intelligent Techniques and Soft Computing EUFIT’95* , 326-331.

Van Drunen, H. (1999). Three stages can work better as one process. *Computing Canada* , 19.

Vieira, F. H., & Lee, L. L. (2007). A Neural Architecture Based on the Adaptive Resonant Theory and Recurrent Neural Networks. *International Journal of Computer Science & Applications* , 45-56.

Watkins, C. J. (1989). *Learning from delayed rewards.* Cambridge: Cambridge University.

Wu, L., Barash, G., & Bartolini, C. (2007). *A Service-oriented Architecture for Business Intelligence.* Newport Beach: IEEE Conference Publications.

Zee, J. V., & Jong, B. d. (1999). Alignment is not enough: Integrating business and information technology management with the balanced business scorecard. *Journal of Management Information Systems - Special section: Strategic and competitive information systems* , 137-156.

Zeng, L., Xu, L., Shi, Z., Wang, M., & Wu, W. (2007). Techniques, process, and enterprise solutions of business intelligence. *IEEE Conference on Systems, Man, and Cybernetics* (pp. 4722- 4726). Taiwan: IEEE Conference Publications.

Zhang, D., Zhai, C., Han, J., Srivastava, A., & Oza, N. (2009). Topic Modeling for OLAP on Multidimensional Text Databases:Topic Cube and its Applications. *Statistical Analysis and Data Mining* , 378-395.