

Survey on Data Mining Techniques Used in IT Business Development

M. Mayilvaganan

S. Aparna

Associate Professor: Department of Computer Science, PSG College of Arts and Science, Coimbatore, Tamil Nadu, India

Research Scholar: Department of Computer Science, Karpagam University, Coimbatore, Tamil Nadu, India
mayilvaganan.psgcas@gmail.com

Abstract— Information Technology is a part and of every organization currently. From past two decades, IT has grown *humongously*. All companies collect their day to day transactional business information using Information technology tools and perform analysis on data to run their business smoothly, sustain and grow in the competitive market environment. Information gathering, analysis and prediction is very crucial for organizations profit. Data mining is the area where large analysis of data can be performed and future prediction can be done by statistical techniques. For an IT service company to grow, the human resources, technology, market and innovation play an important role. The following paper explains various researches being done in the factors affecting an IT business growth and how data mining using existing data can help business to grow.

Keywords— Data mining, Business development, Strategic Planning, Customer Relation Management, Project Management, Human Resources Management, Innovation

INTRODUCTION

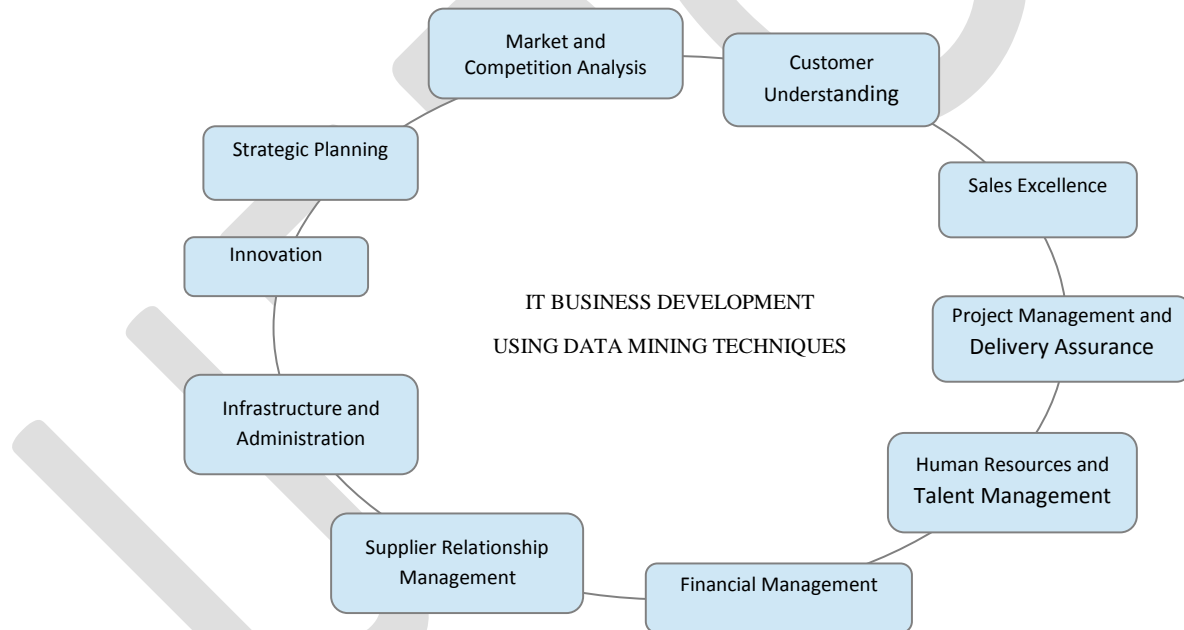


Fig.1 Data mining Techniques Used in IT Business development

From Fig.1 shows the various research organizations have published papers on factors such as Customer relationship management, Sales Excellence, Financial management, Employee Human resources management, partners and vendor management, project management, research, development and innovation.

RELATED WORKS:

Strategic Planning

Strategy planning is an organization's process of defining the goals, determining the actions to achieve the goal and mobilizing resources to achieve the goal. The senior management team in the organization is responsible for setting up the strategy goals. Now-a-days each function in an organization has a strategy leader called by various names - Chief Executive Officer (CEO), Chief Operating Officer (COO), Chief Technology Officer (CTO), Chief Information Officer (CIO), Chief Financial Officer (CFO) and Chief Strategy Officer (CSO). Each of these leaders responsible for laying out a plan to achieve the organization's goal. According to Prof. Manjeet Kumar, data mining can be used to analyse the data in different perceptive and summarize an information for increase revenue, cut costs or both by finding correlations or patterns among dozens of fields in large relational databases.

Their five step process involves below activities:

- [6] Mission, Vision and Goal analysis through business understanding
- [7] Environmental analysis through the activities:
 - a) Data understanding
 - b) Data preparation
 - c) Modeling
 - d) Evaluation
 - e) Deployment
- [8] Strategy formulation
- [9] Strategy implementation
- 5. Strategy evaluation

Marketing Intelligence

Marketing intelligence is the process of gathering and analyzing information in order to under the current and future market behavior, needs and preferences and to access the changes needed in the business in order to sustain and grow in the competition. Market analysis, segmentation and penetration by communicating the value of the product or service to clients are part of the marketing intelligence team. Marketing research involves interpreting conducting research into above activities and statistical interpretation of data into information which is used by managers to plan marketing activities, gauge the nature of a firm's marketing environment and attain information from suppliers. Using statistical methods to analyse the strategy of marketing based on quantitative research, qualitative research, hypothesis tests, Chi-squared tests, linear regression, correlations, frequency distributions, poisson distributions, binomial distributions, etc. to interpret their findings and convert data into information. The task of marketing research is to provide management with relevant, accurate, reliable, valid and current information as represent in fig.2. It is also the function of the marketing team to brand the product, showcase demos, intellectual properties and skilled resources to client.

According to Michael J. Shaw [2], knowledge-based marketing can be used for customer profiling like frequency, size, frequency and life time of purchases, deviation analysis like fraud detection and trend analysis like evaluating the performance of products and marketing programs, forecasting future sales.

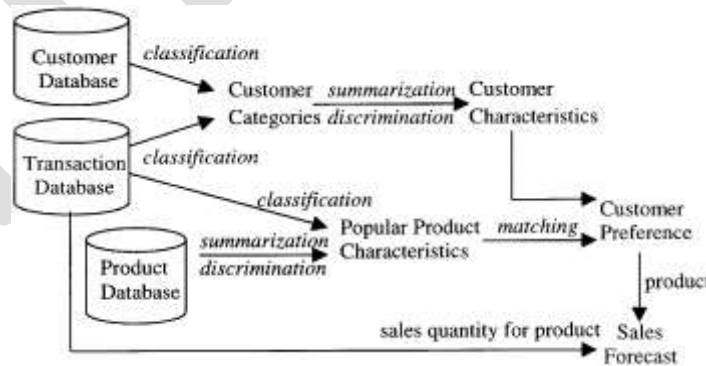


Fig.2 Marketing Research Based on Datamining

Customer Relation Management

Customer relationship management (CRM) is a system for managing a company's interactions with current and future customers. It involves using technology to organize, automate and synchronize sales, marketing, customer service, and technical support.

According to E.W.T. Ngai, dimensions of Customer Relation Management include Customer identification, Customer attraction, Customer retention and Customer development. Data mining techniques can perform one or more of the following types of data modeling: Association, Classification, Clustering, Forecasting, Regression, Sequence Discovery and Visualization using the widely used algorithms of association rule, decision tree, genetic algorithm, neural networks, K-nearest neighbor and Linear/logistic regression.

Sales Excellence

Sales excellence is the process of identifying the right opportunities at the right time. It involves planning for opportunities from existing and new clients, technology and domain and also predicting the future. For example, an organization may be providing production support activities for a client. Penetration can be in the form of obtaining new projects pertaining to enhancements of current product or development of new products not related to the existing support for the client. Currently support may be for reporting, penetration can be in the form of new technologies like reporting on mobile, analytics, enterprise data integration, master data management, web applications development, cloud computing using a different technological stack. Bagging an outsourcing opportunity and acquisitions are all planned activities that can increase the sales of an organization drastically.

Data mining techniques is applied on parameters pertaining to obtaining new projects from existing clients, getting new clients, outsourcing offers, acquisitions and predicting future needs and growth. Organizations set up strategic goals based on the sales growth visualized.

Michael J.A. Berry and Gordon S. Linoff [4] state that the data mining methodology has 11 steps:

1. Translate the business problem into a data mining problem.
2. Select appropriate data.
3. Get to know the data.
4. Create a model set.
5. Fix problems with the data.
6. Transform data to bring information to the surface.
7. Build models.
8. Asses models.
9. Deploy models.
10. Assess results.
11. Begin again.

By defining parameters relating to sales like client business, budget, spending potential, growth and government policies and applying data mining techniques to increase growth, an organization can achieve the expected sales, revenue and profit.

Project Management and Delivery Assurance

Project management is the process of initiating, planning, executing, monitoring and controlling and closure of software projects. From fig.3 represents for project management activities include estimating effort, size, schedule, cost, resources and infrastructure and ensuring that project deliverables meet organization and clients standards. Delivery assurance ensures that reviews and testing is conducted as per plan through audits.

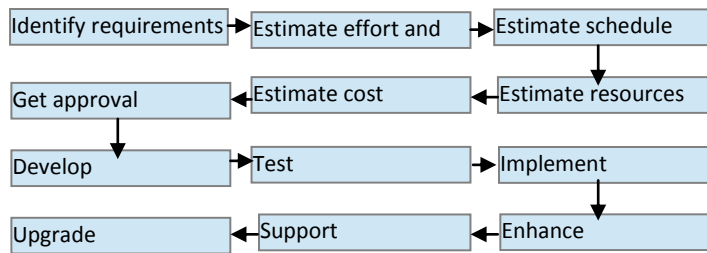


Fig.3 Project Management and Delivery Assurance

Metrics creation and submission is a very important activity of project management. There are lot of project management failures like unrealistic goals of metrics, inaccurate estimates, badly defined system requirements, poor status reporting, unmanaged risks and issues, poor communication among customers, developers and business users. Data mining is used to analyze the existing data on various parameters and arrive at a prediction formula on what factors positively and negatively impact project performance. Data mining techniques can also be used in project proposal creation, project selection, accurate estimation of time and cost to project completion, resolving risk and issues. A company's business development lies largely on delivering successful projects.

PROJECT LIFE CYCLE:

According to Roger Atkinson [5], success factors of delivering code with optimal cost, time and quality, called as Iron Triangle of an Information system that is Maintainable, Reliable and Valid results in organization benefits of improved efficiency and effectiveness, Increased profits, Organizational learning and Reduced waste and Stakeholder benefits of Satisfied Users, Social and Environmental Impact, Personal Development, Contracts Profit, Capital Suppliers.

Human Resources and Talent Management

Human Resource management involves acquiring resources for project execution, laying performance goals for various streams, setting up salary bands according to industrial standards, defining rating criteria and appraise employees of the organization. Human resource management also involves satisfying employees through rewards and reorganization programs. Human resources management also requires identifying training requirements and executing the same for the organization's growth. Talent management focuses in retaining talent within the organization and improving the same. There are various factors that affect employee performance and productivity. Factors like age, sex, marital status, number of kids, university type, specialization, degree, grade, country, years of experience, prior companies worked, job title, rank, service period, technological and domain expertise, certifications, process knowledge play a crucial role in an organization's development. Sangita Gupta and Suma V [6] have arrived at a 5-step process to arrive at selection of personnel for projects.

Step 1: Training data is obtained based on parameters related to software project personnel like General Percentile aggregate, Domain knowledge assessment, Programming skills, General proficiency, Communication skills, Time efficient, reasoning skills, Performance.

Step 2: An Attribute Selection Measurement Function (ASMF) is calculated based on the parameters.

Step 3: An average maximum occurrence for each attribute is calculated as $\Sigma ai/CiD$.

Step 4: Basis of sorted values of ASMF, given training set is divided into subsets and move to another level of tree.

Step 5: The process is repeated on each subset iteratively and a decision tree is derived. The result and rules obtained can classify project members into three classes of performance- good (should be deployed), average (can be deployed with training) and poor (should not be deployed).

Similar process can be applied to obtain the effectiveness of training programs, campus recruitments, employee retention rates, etc.

Financial Management

Financial Management is the Operational Activity of a business that is responsible for obtaining and effectively utilizing the funds necessary for efficient operation of the organization. The goal of Financial Management is to optimize the cost of IT Services while taking into account quality and risk factors. Budgeting, IT accounting and Charging are 3 major sub-processes in Financial Management. Managing customer contracts, budgeting, project costing, profitability calculation based on global outsourcing, resource grades, account level, geography, industry trends, cost cutting, dollar fluctuations are all functions of financial management. Planning of acquisitions and calculating their profitability, loan management, forecasting stock market and money laundering analysis also form part of financial management.

According to Abhijit A. Sawant and P. M. Chawan [7], Bayes Classification can be used to classify data, Decision Tree can be used for decision making, boosting applied to predictive data mining, bagging and random forest algorithm improves classification of data. Using available data and applying the various data mining algorithms to the available data, it is possible to manage the finances for above said areas.

Supplier Relationship Management

Supplier relationship management is a comprehensive approach to managing an enterprise's interactions with the organizations that supply the goods and services it uses. Suppliers can be vendors who provide human resources and partners with whom the organization has a tie up. Supplier management drives cost accountability and innovation through active management of supplier performance and ensures sustainability through compliance and measurement. They have also reaped significant potential benefit in being able to prepare against any supplier-related performance issues or risks through a deeply cooperative ongoing supplier relationship.

According to Suhong Li, Bhanu Ragu-Nathan, T.S. Ragu-Nathan, S. Subba Rao [8], the five dimensions of Supply Chain Management (SCM) practice are strategic supplier partnership, customer relationship, level of information sharing, quality of information sharing, and postponement. Higher levels of SCM practice can lead to enhanced competitive advantage in terms of price, quality, delivery dependability, product innovation and time to market and improved organizational performance in terms of market and financial performance.

Infrastructure and Administration

Infrastructure for an IT firm includes the building, facilities and hardware equipment. Location of the firm and proximity to resource availability plays a vital role in development of a firm. Special Economic Zone (SEZ) is a geographical region within a country in which tax and investment incentives are implemented to attract foreign businesses and investments. IT Companies are moving to SEZ in order to benefit the exemptions provided by government like exemption from central sales tax, service tax, etc. Servers, Desktop computers, laptops, virtual machines, phones lines, cubicles, video conferencing and meeting rooms, pantry all these need to be planned to suit organization's goals and vision. Administrative services focus on transportation facilities, food and beverages, client visit management, housekeeping, event and hospitality services and security services. Security is a basic necessity for an organization to grow. Security covers security to infrastructure, data security and human security.

According to Leonard Jessup and Joseph Valacich [9], ensuring availability, reliability and security of the facilities infrastructure, managing the human resources and services infrastructure, outsourcing, disaster planning and information system auditing are key for an organization's growth. Based on the organization's vision, data mining techniques can be applied to obtain the right location, infrastructure, services and security for the organization.

Research, Development and Innovation

Research and development focus on identifying new areas for an organization to grow from technical, functional and process perspective. Innovation is identifying new methodologies and products to provide customer delight and satisfaction. According to Andrew Kusiak [10], there are three different types of innovation intermediaries – Invention Capitalist, Innovation Capitalist and Venture Capitalist. Invention Capitalist operates on Lost Cost and High Risk Innovation Cycle. Venture Capitalist operates on High Cost and Low Risk Innovation Cycle. Invention Capitalist connects with companies with ideas but not market ready. Venture Capitalist connects with companies with market-ready products. The innovation capitalist optimizes the trade-off between cost and risk bringing market-ready ideas to companies. Generic process for innovation is to search new methods, evaluate them, refine them, develop prototypes and connect with market. The level of innovation can be expressed as a Function of requirements X , $I = F(X)$. For example, in an IT company, innovation can be a function of requirements like quality deliverables, cost effectiveness and long-term reliable relationship for customer delight. Thus research, development and innovation play a key role in an organization sustenance and growth.

GENERAL FRAMEWORK:

For every business process below would be the framework on implementing data mining techniques to improve business by representing in fig.4. As a rule of thumb, identify success and failure factors in each business process, implement the success factors, take corrective and preventive action for failure factors and bring the business to success.

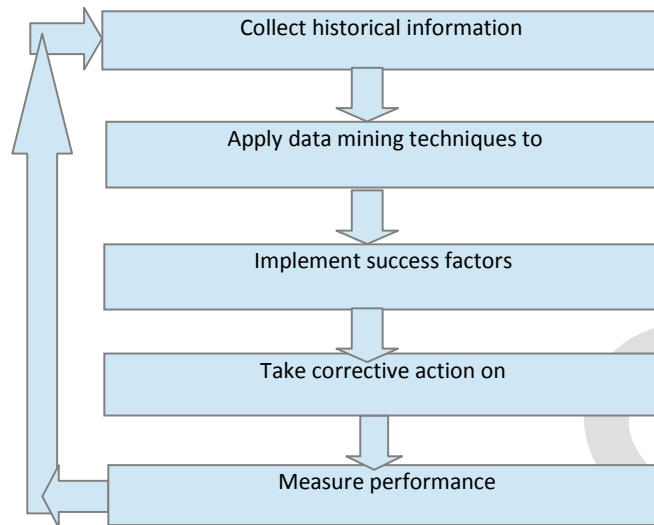


Fig. 4 General framework

CONCLUSION

In this paper the research can be focused on issues that companies face in sustenance and growth, best practices that growing and grown companies have adopted with respect to technology, management, process and employees. In future work, it will be presenting the relation between the various factors and how each and every function needs to perform in the overall organization growth.

REFERENCES:

- [1] Data mining-the essence of strategic management, prof. Manjeet kumar, Assistant Professor, its, Mohan Nagar, Ghaziabad, Ifrsa International Journal of Data Warehousing & Mining |vol1|issue 1| Aug 2011.
- Knowledge management and data mining for marketing - Michael J. Shaw a,b,c, Chandrasekar Subramaniama, gek woo tana, Michael e. Welge b - Decision Support Systems 31 2001 127-13
- Application of data mining techniques in customer relationship management: a literature review and classification - e.w.t. Ngai a*, li xiu b, d.c.k. Chau a - a Department of Management and Marketing, The Hong Kong Polytechnic University, Hong Kong, pr china, b Department of Automation, Tsinghua University, Beijing, pr China
- Data mining techniques for Marketing, Sales and Customer Relationship Management, Second edition, Michael J.A. Berry and Gordon s. Linoff, wiley publishing inc., 2004
- Project_management_cost_time_and_quality - roger atkinson, International Journal of Project Management vol. 17, no. 6, pp. 337-342, 1999
- Empirical study on selection of team members for software projects – Data Mining Approach - Sangita Gupta¹, suma. V.2, 1 Jain University, Bangalore, 2 Dayanada Sagar Institute, Bangalore, India
- Study of data mining techniques used for financial data analysis, Abhijit A. Sawant and pramila. M. Chawan Department of Computer Technology, Vjti, Mumbai, India, International Journal of Engineering Science and Innovative Technology (ijesit) volume 2, issue 3, may 2013 503
- The impact of supply chain management practices on competitive advantage and organizational performance, Suhong Li^{a,*}, Bhanu Ragu-Nathan^b, T.S. Ragu-nathan^b, s. Subba Rao^b, Omega, The International Journal of Management Science, Omega 34 (2006) 107 – 124.
- Information Systems Today - Leonard M. Jessup and Joseph s. Valacich, Managing the Information Systems Infrastructure, http://www.prenhall.com/behindthebook/0132335069/pdf/jessup_ch04.pdf
- Data mining and Innovation Science - Andrew Kusiak - Mechanical and Industrial Engineering, 3131 Sc, The University of Iowa