2007

Product Development: Knowledge Discovery through Decision Tree Approach

Anza Akram A
University of California, anza@sbcglobal.net

Product Development: Knowledge Discovery through Decision Tree Approach

Anza Akram A.
Lecturer
Anderson School of Management , University of California, Riverside, U.S.A.
anza@sbcglobal.net

Abstract: This paper proposes a conceptual framework for generating an innovation through knowledge discovery from integrated design and analytical systems in monitored and controlled environment by making accurate, timely, precise and profitable decisions. This paper proposes a framework for manufacturing sector to evaluate their product development ideas through integrated technologies to meet the quality and design expectations of the local and international buyers. Previous studies have made contribution in the development of data-driven systems whereas, this study evaluates the internal dimensions of the organizations' tacit and explicit knowledge to generate the idea and make contribution in creating a expert-driven system that take into consideration the factors that effect the organizational decision making process in the development of a new idea.

Keywords: Knowledge Discovery, Expert, Expert-driven Systems, Decision Tree, Profitability, Integrated technologies.

Introduction

Manufacturing sector is an important segment of our economy to generate revenue from export and import of consumer and industrial products and to meet the national and international demand of the market. Presently, due to a geographical shift of plants to overseas, faces challenges to compete the market by overcoming the buyers' and suppliers' concern of quality and price over its competition. The quality can only be controlled through better communication and specification provided by the holding companies. The product development in that environment demands more precise and well thought procedures to produce deliverables like engineering of new ideas, changes to existing product or adding new product line to ensure the quality of the product. The innovation or change mostly a feedback from the market or competition needs an immediate analysis to evaluate the results to make a decision to provide timely gains. The generation of an idea is evaluated by a team of engineers, managers and executives to approve or disapprove the new idea. At present, the manufacturers usually receive the design and manufacturing specs from the holding companies and are in most of the cases, not responsible for the design specifications. The manufacturing follows two principles, production (how the product is manufactured) and quality (what tests to go through to pass the quality testing). Production is done internationally and design and engineering of the products are done domestically due to the marketing challenges among the buyers and suppliers. Its been argued and proved that the country of design (where the product is conceived and engineered) plays an important role in creating a quality perception among purchasing agents [13] and thus, has better chances of passing the quality testing if the initial design and engineering to prepare the specifications are implemented firmly.

Moreover, the rigid environments adds to the management struggle to handle the foreign challenges like management of external facilities, communication, design specs, international tariffs and other regulations. The intricacies of these challenges are more sensitive to financial losses. These risks can be avoided if the distinction between the roles are clear and well defined to enhance the strategical relationships between the foreign organizations.

Literature Review

Presently, the management is trying to find ways to make the process easier to control and understand for both parties in a time efficient and cost effective manner. Large volume of products manufactured in other countries demands a strict procedure (which is still in initial stages) to address this challenge through integrated sophisticated databases and network technologies. In a real world, the data and processes are not perfect and the management responsibility is to find a strategy for changing the information system to identify critical data items and processes in the context of required decision and plans that is cost-effective and time sensitive [3]. Neumann et al. [39] and

Ahituv [2] identified the characteristics of information like timeliness, accuracy, reliability, relevancy and cost as a holistic perception of the users and this multidimensionality play an important role in the analysis process. Ahituv [2] added "utility gained is not a unidimensional function".

Making a decision under uncertainty is a choice between a prospect (outcome of existing known situation) and a gamble (unknown situation) where prospect can be predicted through a decision weight and is risk averse [14]. In case of gamble, the risk is unpredictable and unavoidable. Knowledge discovery is defined as finding answers from already existing knowledge which resides as documents, data sets, facts, figures, images, etc. in databases or data warehouses. Numerous studies in the area of knowledge discovery or data mining strategy have proposed different frameworks. Their research emphasis was on the dimensions of the data, presentation of data, combination of logical and analytical models using decision tree techniques. The studies also contributed in the areas of data-mining integration with data warehousing and some networked all sources of databases. Some of the research is summarized in Table 1.

Table: 1
Research Abstract

Research Summary	**Authors**
1. To extract large and high-dimensional off-line data sets through linked interactive graphical devices	[27]
2. Combined multi-model analysis and decision tree logic	[7], [33]
3. Establishing data analysis from all kinds of sources	[35]
4. For developing product prototypes and discovering the relationships between the customers' preferences and the physical characteristics of product	[29]
5. Data mining framework integrated with data warehousing	[23]
6. Web based infrastructure	[15]
7. Learning from non-linear non-negative multi-regression coefficient of a single element of network	[19]
8. Discovering frequent sequential patterns	[20]
9. Mining three levels for similarity and periodicity patterns to find global patterns from discrete-valued time series set (DTS)	[37]
10. Global learning scheme as a multi-strategy and multi-agent KDD (Knowledge Discovery and Data Mining) system	[26]
11. Foreign key relations	[16]
12. Domain knowledge in evidential data mining	[5]
13. The coverage & overlap of sources for efficient query processing	[24]
14. An Ontology-Based Approach for Knowledge Discovery	[6]

The research in the manufacturing sector shows that the change from traditional way of carrying business to using information systems is still evolving and have potential towards technology adaptation for various purposes

like, managing inventory system, reducing working hours, decreasing their response time to market demand, etc. [28], [22]. Previous studies address the factors involved in developing a system which address a product features or a systems that evaluate the data from customers' perspective. Most of the studies address the physical and technical aspects of systems. This study addresses the managerial perspective and take input from multiple integrated systems used by the experts and the decision makers' to introduce and evaluate the idea. The idea generation is not only system dependent but also expert dependent. The framework described is an essential integration of experts with the systems and making choices based upon the input from different levels of expertise. This framework is applicable in any size of manufacturing organization and most of the initial data requirement in this case is from existing technologies or systems and are in use in different areas for various purposes. The literature review confirms that new type of systems are formed by adding features to existing systems [9]. The segment of knowledge discovery if implemented as proposed (Appendix 1) make the decision making and idea evaluation less riskier and more beneficial.

Proposition

"Knowledge Discovery profitability is directly related to integrated Decision making strategy".

Definition of Knowledge Discovery

Knowledge Discovery is a data extraction technique used in developing embedded information or knowledge in decision making interfaces with multiple levels of pattern or models for the professionals involved in various activities related to the process. "The activity uses a large and heterogeneous data sets as a basis for synthesizing new and relevant knowledge" [11]. The knowledge discovery is defined by other authors as a technique of manipulating large volume of data where problem can be categorized by classification (partition of data into disjoint data); association (correlations among data items) [17]; and sequences (sequences among data items) [31]. Discovery is also been done through event pattern discovery from event histories (sequences) and is very similar to finding association rules among a set of items [30].

This study proposes an expert[1]-driven system which is independent of data warehousing repository and assumes that the system can be used with the various types of databases for evaluation purposes due to its flexibility (in most cases) of interaction and interface which is both graphical user interface in web intranet environment or the legacy system (Appendix I). The evaluation is done on both departmental level and on organizational level and are integrated together to share the result of the nature of the idea to develop a new product and each process is consequentially unique thus, the selection of criteria is skill and expertise based. The proposed strategy helps "to avoid the major problem decision makers face when measuring the realistic value of information is the accuracy of the measurement" [1]. The motivation behind developing a framework is to make a best decision based on design and profitability by minimizing the risk, and ignore the flexibility of usage of the system or user friendliness. In other words, the development of a knowledge discovery system whose characteristics can be monitored and controlled [1].

Knowledge discovery techniques have been used in various industries to extract knowledge for making decisions in knowledge generating environment where time and distance is a constraint (short time period, quick response or turnover). The knowledge discovery deals with exploring the ideas or goals the managers come across with on continuous basis and are challenged in producing valuable outcomes. To address the challenges like time, distance, risk and cost to project the future financial decision, the proposed framework for an enterprise model facilitates to answer the sensitive inquiry regarding the change towards a new product or addition to an existing product feature.

Decision Making Strategy

The present study evaluates the decision tree approach as a decision making strategy to discover the knowledge. Decision tree is set of functional models and each model has sub models of the concepts which have multiple plans to a given decision and is a conjunction of tests in the strategy analysis. The system level analysis is class based and in case of multiple paths for a given class form a disjunctive concept description and are mutually exclusive [25]. Decision tree evaluation creates a visualization of the trees and diagnostic measures of their effectiveness [4]. The

[1] Expert word is derived from Latin verb meaning "to put to test." Experts—people with deep knowledge of a subject [8].

analysis is based upon meta decision trees and uses a method of combining multiple models [36]. The levels and association is created by forcing the resultant attribute from the previous decision tree to the next level or tier of the tree which is close to creating networks using Learning Bayesian Networks [21] but the fed node is based upon a fixed and unchangeable node of previous network or databases. The reason it is important to discuss this difference between the nodes of network or databases is to understand the strategy behind the framework of knowledge discovery that may be decentralized or centralized or distributed or hybrid.

The decision tree is traditionally based upon hierarchical system where parent has many children and a node is representing a child or branch whereas in this case a strategy is modified to show a relationship of parent with multiple relationship to single sublevel to multiple relationship. In other words, the decision of each tier is integrated with the other tier and the evaluation of the next tier is not possible in absence of the information from the previous tier. The characteristics of each tier are chosen and are different for each decision and are dependent and changeable as the nature of a given idea changes. For each idea evaluation the criteria selection is independent of previous idea evaluation. The set of characteristics given in the option are all given characteristics of the ideas involved in the previous evaluation or manufacturing and for each new evaluation the chosen criteria defines the characteristics selected for the new idea[2].

The decision tree evaluation is based upon hierarchical sequence of idea in evaluating a decision. This paper follows the decision tree idea and thus follows the process level integration. The projection of innovative idea's profitability is time consuming and costly if done in the absence of technology as data gathering needs tremendous amount of manpower and services to develop the prototype of a product to make a final decision. Even in the presence of information system where the systems are not integrated as in most cases of manufacturing organization the collection of data set to reach an effective decision can be complex and confusing due to the overlapping and lack of expertise in the area as the databases are just repositories and the resided data sets are bits or digital information. Ahituv [1] describes that "all these measures cannot even be considered as informational units because there is not necessarily any relation between the size and the value of a data set". Developing feasibility is cost and time consuming and requires a detailed analysis from previous data sets to reach a final decision that consists of product attributes, design, make-over, attributes identifications, time evaluation, risks evaluation and final yes or no.

Decision Making Tree and its Attributes

The multi-tier decision making tree consists of computer aided manufacturing and analysis models that are defined at each level and are dependent upon the data set of that level. Each level is independent of each other and the datasets are unique to the process and cannot be used in-lieu of other dataset. Each decision node has three alternate branches connected to the option available in a given time period. The alternatives a1, a2 and a3 are time dependent and the outcome is based upon the best choice or option available with minimum cost and maximum return in a short period of time, that is usage of models, functions or formulas. The decision can also be chosen by setting conditions to different alternatives, that is, satisfying certain statements, for example, part A is equal to ASTM A 526 (Material commercial quality), the cost or profit is less than c or greater than r (revenue in dollars) and the time is less than t_n (number of months) (Appendix 2**)**.

Profitability

Profitability of an idea clearly identifies the decision making technique as a success or failure. The cost of the process determines the financial burdens that decision places on the overall performance of the company and can be accepted or rejected depending upon the contribution it makes to evaluate the economical utility we gain from the decision. The profitability analysis is based upon the expected returns, expected returns portfolio and turn over ratio. The profitability is positive if its value is greater than the value of cost. In this case the value of evaluating the innovation is greater than the cost involved to analyze the idea. The profitability, R, and the cost, C, and the Time, t, used to evaluate the value of given process in a given period of time, all three factors play important role in the overall equation. The equation is stated as follows:

---

[2] The idea is a generation of market evaluation through integration of internal environment and thus, the external environment measures are excluded from the framework.

R > C where ECAA > ECMA
And
tc < tm

where

Profitability, R, is said to have a positive value when it is greater than cost, C and the time tc involved to complete the integrated systems evaluation process is less than time tm involved to complete the analysis when the systems and subsystems are not integrated. ECAA represent the expected cost of computerized analysis and ECMA is expected cost of primitive system analysis (paper work, meetings, legacy systems, phone calls, people, etc.) and the value of ECAA is always greater than ECMA to accept the value of integrated process. While the time values are separated based upon the time the analysis is started and finished to bring the idea to life without inviting any other external financial threats in the equation. The time spend to produce results is less by using the integrated systems than the time spend to complete the process through primitive systems.
where

$$tc = totaltimespendtocompleteAA$$
$$tm = totaltimespendtocompleteMA$$

The Cost of Automated Analysis consists of information systems, databases, models/ decision trees, communication and decision making.

$$CAA = IS + DB + C + DT + DM$$

The cost of primitive system (CMA) consists of paper work, meetings, legacy systems, phone calls and people (engineers, management and senior management

$$CMA = \Pr imitiveSystem\left[PW + MT + LS + PC + P\left(E + M + SM\right)\right]$$

If we break down the process of decision making, DM, then the following statements are true if the other factors that effect the equations are kept constant,
DM Process = deciding to make investment.
Deciding = Taking Action.

Action = Investments "committing to financial burdens" for higher returns
DM = Action (Investment in a given period of time)
DM = Action(Time)

$$C\left(AA + CA + CDA\right) < Action\left(Tm + R\right)$$
$$C\left(AA + CA + CDA\right) < C\left(MA\right)$$

Most of the organization realized that the manufacturing of products overseas is riskier than domestically as the operational function is transferred to the foreign entity and no longer is controlled by them.
Thus, decision to produce a new product consists of:

R = Profitability
Cost = Manpower + material + Design + marketing + selling
Time = design time + Manufacturing time + Marketing time + selling time
Risk[3] *= β sensitivity to Δ*
β = risk of accepting or rejecting the idea
Quality = Qi of new Product in a given time

Expected Return

[3] Elton and Gruber 1995

The expected return is based on profitability after the cost of producing the idea and the cost of risk is subtracted from the total gross return.

$$E(R) = Max \sum_{i=0}^{t} [R_i - (C_i + C_i \beta)]$$

Expected Return Portfolio
Portfolio Expected Return is calculated by calculating the percentage value of each idea's expected return on total portfolio.

$$E(Rp) = \sum_{i=0}^{t} RiIp$$

Turnover Ratio[4]

$$Tt = Ci + Ci \text{ â/ } Ti$$

Tt = Turnover Time Ratio
Ti = Total time to complete the AA and A/DA
CA = Computerized Analysis
AA/DA = Automated Approval of Decision or Disapproval.

Risk on Expected Returns [footnote 2]

$$Variance\ (ó)^2 = \sum_{i=0}^{t} (\text{Probabilit y}) (\text{Possible Return - Expected Return})^2$$

$$= \sum_{i-1}^{t} (\text{Pi})[Ri - E(Ri)]^2]$$

Standard Deviation= square root of variance

Risk value is based upon the value assign to the condition and is dependent upon the return. Risk of rejecting an idea (risk aversion) is set to a condition where the value of return is less than 0, the assign value is -1 or less. Risk of accepting an idea (risk preference) the value is greater than 0 and is 1 or more. Third alternative, risk of indifference (Risk neutrality), the value equals 0. In real world example the indifference alternative has positive effect on the profit and value can exceed the value of 0 under uncertainty.

Discussion

Simon's [34] idea of Holism is interaction of complex systems with their properties and other subsystems and reducing "whole" in terms of mechanism. Whole is adaptive goal- oriented systems and also an artificial intelligence philosophical element of exploration [34]. The reduction of complex system into subsystem is based upon their properties and their relationship of sets that exists in each tier of the complex system. The exact division of the tier is model driven and each model is pertained to the area or part of analysis it belongs to that is derived and laid out by selected criteria by the experts of that field. The tiers are department biased or process biased and thus, may only be restricted to be performed by their experts due to the nature of complex data sets involved in the selection of model at each level. The dependency of the resultant product and independency of the analysis at each level separates the idea and its complex analysis as described in the framework (Appendix 1). The process of evaluating the under observation inquiry is related and the process can be repeated dependent upon the final decision made by the decision maker and the experts involved. The accepted decision is passed through the system to the other segments that generate the physical aspects of the idea.

[4] Reilly and Chance 1994

"Computer Mediated Communication Systems (CMCS) complicate the information exchange process since it may involve explaining the meaning and significance of ambiguous information to the other group members rather than simply conveying facts" [38]. The enforced structure of analysis helps in decreasing the complexity of information been analyzed by distributing the data to the entities who best defines their attributes and the value of the attribute, usually defined by the experts and thus, save time and money. The scope of the system is knowledge discovery and avoids an effort from the audience to understand the meaning before extracting the knowledge they are discovering. This aspect makes this study unique and different from the existing studies which proposed the systems that evaluate the data resides in the system and not the actual idea.

The technical approach proposed in this paper is decision tree analysis.The decision tree consists of seven decision analysis nodes, each node has three alternative and the three resultant value of the options. The final decision for each analysis is linked through relation between the common set of data in each level, in case there is no common data-set the value is carried forward by creating a link between the two entities and thus, a new relationship is emerged from the previous node. This make the technique an open ended where new setsof data can be included. The data can be included from the existing systems or defined by the experts depending upon the complexities involved in generating the idea. This aspect makes the system a knowledge oriented than a data oriented. The three alternatives are the values to accept the decision ($\theta 1$), reject the decision ($\theta 2$) or indecisive value ($\theta 3$) of the decision is carried forward if the other two values are equal to each other or by selecting any of the other two values have no impact on the final outcome of the decision and the third value have a greater impact on the final outcome. The indecisive decision parameters are calculated set and are selected when the available option's value is any value where changing the attributes satisfies one of the value, time or returns, but the single value exceeds the value of the other two alternatives and still provides a best available decision. The control of the decisions are equally distributed between the multi-tier and each tier has an equal computer aided control which in other cases does not exist and thus, depends upon the tier's political influence on the other higher level tier or vice versa (to kill the idea or not to kill the idea).

The success or failure of the idea is measured against the profitability of the idea. The profitability is then measured against the innovation, cost and time. The formula itself considers the cost and time involved to generate the idea using primitive system and the cost and time involved in generating the idea using automated system. The acceptance and rejection is based upon the produced results. The paper assumes that the success of the innovative idea is dependent upon the strategic information processing system and the value of the outcome is always greater to be true.

The developed framework looks at the decision making process within the organizations as a most important factor that plays a major role in gaining the market share and addresses the challenge of valid decision making under stressed environment by the experts through information systems. The continuous process is also sensitive to the amount of time it takes to complete a cycle of decision making by means of strategically optimal investment.

Limitation of the Research

The study's limitation is generalization of the findings in the theoretical field. For generalization purpose there is a need of testing the framework against a real data from the manufacturing sector. The replication of the framework would be more profound and useful if the tested data validate the assumptions I made and provides both internal and external validity.

Implications and Future Research

The paper postulates a new knowledge in the field of knowledge discovery and opens areas to study new ways of analyzing the questions or inquiries in the fields of Information Management, Engineering and Computer Science. The proposed inquiry may be studied further using positivist approach to analyse the proposed expert-driven system and the outcomes. Recent studies concentrates on logical and sequential techniques whereas, this study encourages flexible and analytical techniques to find related data from the data repositories and the experts. For the researchers in the area of management, the paper proposes profitability formula urges to conduct a positivist and empirical studies. It provides opportunities to replicate expert-driven system in other areas of management, engineering, and computer science.

Conclusion

The study analyses complex systems and their subsystems in the process of idea generation in product development. The motivation behind the study is to gain the optimum time and investment value from different integrated sources of knowledge and not compromising the value and quality of an idea. The systems are expert dependent and the approach used to reach the final decision is a function of interrelated activities performed in different frames at different time. The integration of systems with other relevant systems provides a sophisticated technology to handle complex knowledge. The derived findings to implement the framework provide maximized outcomes and the process of reaching a final decision is time efficient and cost effective in manufacturing sector.

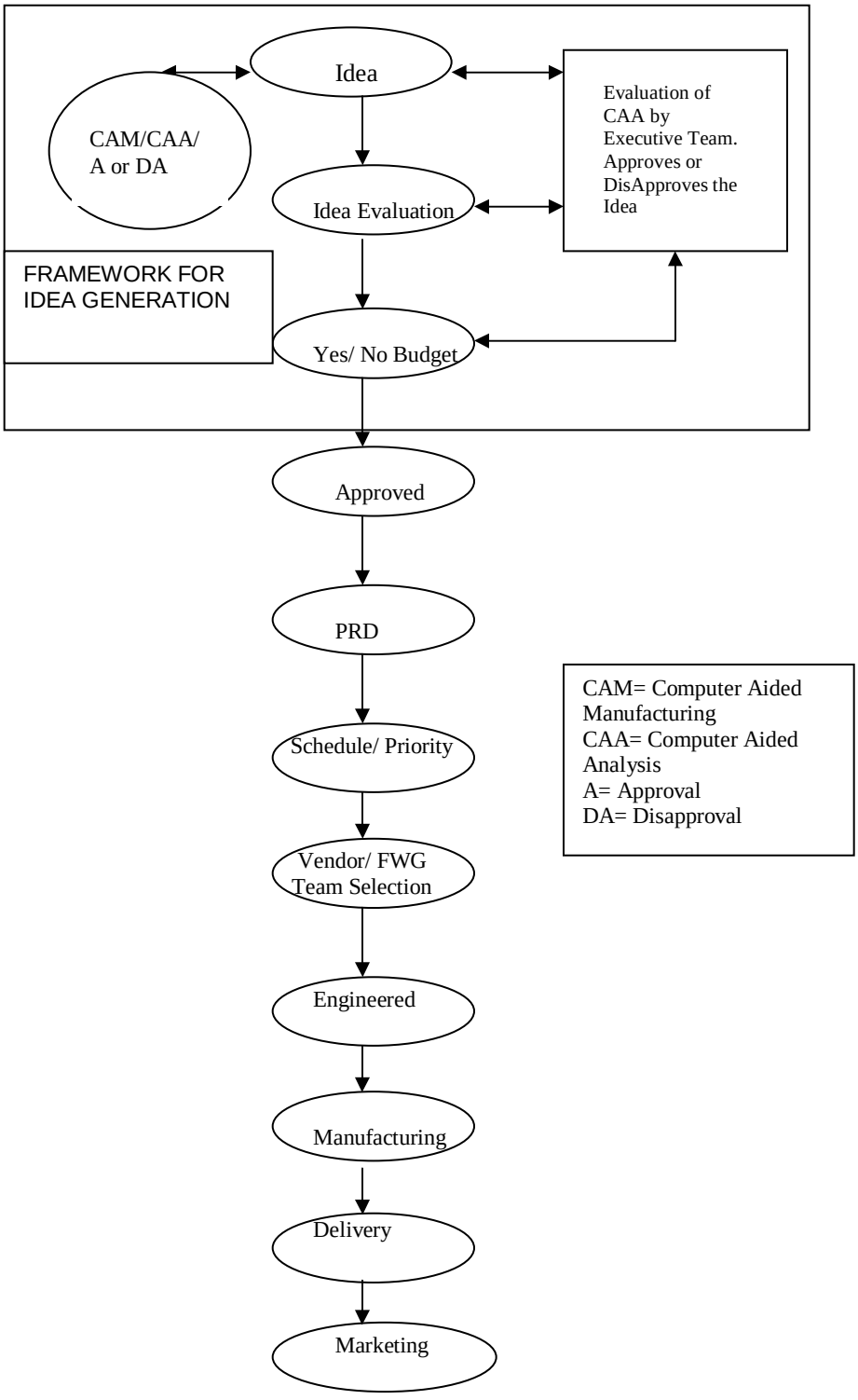
References

1. Ahituv, Niv. "Assessing the value of information: Problems and approaches." Proceedings of ICIS (1989)
2. Ahituv, Niv. "A Systematic Approach Toward Assessing the Value of an Information System." MIS Quarterly, (Dec 1980) 61-75
3. Ballou, Donald P. Pazer, Harold L. "Designing Information Systems to Optimize the Accuracy-timeliness Tradeoff." Information Systems Research, 6, (March 1995) 51-72
4. Barlow, T. Neville P. "Case study: visualization for decision tree analysis in data mining." Proceeding 2001 Information Visualization 2001. INFOVIS 2001 Compu. Soc. Los Alamitos, CA. USA. (2001) 149-52
5. Bell DA. Anand SS. Hughes JG. "The role of domain knowledge in data mining." Proceedings of the 1995 ACM CIKM International Conference on Information and Knowledge Management, NY, USA. (1995) 37-43
6. Berstein, Abraham, Provost, Foster, Hill, Shawndra. "Toward Intelligent Assistance for a Data Mining Process: An Ontology-Based Approach for Cost- Sensitive Classification." IEEE Transactions on Knowledge and Data Engineering, Vol. 17, No. 4, (April 2005) 503-518
7. Chen, Shu-Ching, Shyu, Mei-Ling, Chen, Min, Zhang, Chengcui. "A decision tree-based multimodel data mining framework for soccer goal detection." IEEE International Conference on Multimedia and Expo. Part vol. 1, Piscataway, NJ, USA (2004) 265-8
8. Davenport, Thomas H., Prusak, Laurence. Working Knowledge. Harvard Business School Press, Boston, Masssachusetts (1998) 7
9. Ein-Dor, Phillip, Segev, Eli. "A Classification of Information Systems: Analysis and Interpretation." Information Systems Research,4:2 (June 1993) 166-204.
10. Elton, Edwin J., Gruber, Martin J. Modern Portfolio Theory And Investment Analysis, Fifth Edition, John Wiley & Sons, Inc. (1995) 216
11. Foss, Bjarne A. "Knowledge Discovery – A Control Theory Perspective." Principles of Data Mining and Knowledge Discovery, First Eropean Symposium, PKDD ' 97 Trondheim, Norway, Proceedings (June 1997) 1-3
12. Hitt, Lorin M. Brynjolfsson, Erik. "Productivity, Business Profitability, and Consumer Surplus: Three Different Measures of Information Technology Value." MIS Quarterly (June 1996) 121-142
13. Insch, Gary S. " The impact of Country-of-Origin Effects on Industrial Buyers' Perceptions of Product Quality." Management International Review, vol. 43, No. 3 (2003) 291-310
14. Kahneman, Daniel, Tversky, Amos. "Prospect Theory: An analysis of decision under risk." Econometrica, vol. 47, no.2 (March 1979) 263-291.
15. Kantardzic, M. Kumar, A. "Toward automatic distributed data mining with intelligent Web services." International Conference on Information and Knowledge Engineering- IKE '03. CSREA Press. Part vol.2, Athens, GA, USA (2003) 544-50
16. Knobbe Aj. Adriaans PW. "Discovering foreign key relations in relational databases." Cybernetics and Systems '96. Proceedings of Thirteenth European Meeting on Cbernetics and Systems Research Austrian Soc. Cybernetic Studies. Part vol.2, Vienna, Austria (1996) 961-6
17. Knobbe, Arno, J., Siebes, Arno, Blockeel, Hendrik, Van Der Wallen, Daniel. "Multi-relational Data Mining, Using UML for ILP." Principles of Data Mining and Knowledge Discovery, Fourth European Conference, PKDD 2000, Lyon, France Proceedings (Sept. 2000) 1-12.
18. Leung H. "A data mining framework for semantically distributed data bases." Proceedings of the Seventh IASTED International Conference on Artificial Intelligence and Soft Computing. ACTA Press, Anahiem, CA, USA (July 2003) 301-6

19. Leung, Kwong-Sak, Wong, Man-Leung, Lam, Wei, Wang, Zhenyuan, Xu, Kebin. "Learning Nonlinear Multiregression Networks based on Evolutionary Computation." IEEE Transactions on Systems, Man, and Cybernetics-Part B: Cybernetics, vol.32, no.5 (Oct 2002) 630-44.
20. Masseglia F. Teisseire M. Poncelet P. "Real time Web usage mining: a heuristic based distributed miner." Proceedings of the Second International Conference on Web Information Systems Engineering. IEEE Comput. Soc. Part vol.1, Los Alamitos, CA, USA (2002) 288-97
21. Masuda, Gou, Yano, Rei, Sakamoto, Norihiro, Ushijima. "Discovering and Visualizing Attribute Associations Using Bayesian Networks and Their Use IN KDD." Principles of Data Mining and Knowledge Discovery. Proceedings of Third European Conference, PKDD '99 Prague, Czech Republic (Sept. 1999) 61-70.
22. Mukhopadhyay, Tridas, Lerch, Javier F., Mangal, Vandana. "Assessing the impact of information technology on labor productivity; A field study." Decision Support Systems 19 (1997) 109-122.
23. Nestorov S. Jukic N. "Ad-hoc association-rule mining within the data warehousing." 36th Hawaii International Conference on Systems Sciences. IEEE Comput. Soc., Los Alamitos, CA, USA (2003) 10
24. Nie, Zaiqing, Kambhampati, Subbarao, Nambiar, Ullas. "Effectively Mining and Using Coverage and Overlap Statistics for Data Integration." IEEE Transactions on Knowledge and Data Engineering, Vol. 17, No. 5 (May 2005) 638-651
25. Nikolaev, Nokolay I., Slavov, Vanio. "Inductive Genetic Programming with Decision Trees." Intelligent Data Analysis, 2 (November 1998) 31-44
26. Ning Zhong, Chunnian Liu, Ohsuga S. "Dynamically organizing KDD processes." World Scientific. International Journal of Pattern Recognition & Artificial Intelligence, Singapore. vol.15, no.3 (May 2001) 451-73
27. Porzio GC. Ragozini G. "Visually mining off-line data for quality improvement." Quality and Reliability Engineering International, UK. vol.19, no.4 (July-Aug 2003) 273-83
28. Ragowsky, Arik, Ahituv, Niv, Neumann, Seev. 1996. "Identifying the value and importance of an information system application." Information & Management, 31 (1996) 89-102
29. Raicu, D.S. "A data mining framework for target marketing." 18th International Conference on Computers and their Applications. Int. Sc. For Computer and their Applications, Cary, NC, USA (2003) 476-9
30. Rajamani, Karthick, Sung, Sam, Cox, Alan. "Extending the Applicability of Association Rules." Proceedings of Methodologies for Knowledge Discovery and data Mining, Third Pacific-Asia Conference, PAKDD-99, Beijing, China (April 1999) 64-73
31. Ramstein, Gerard, Bunelle, Pascal, Jacques, Yannick. "Discovery of Ambiguous Patterns in Sequences Application to Bioinformatics." Principles of Data Mining and Knowledge Discovery. Proceedings of Fourth European Conference, PKDD 2000, Lyon, France (Sept. 2000) 581-586
32. Reilly, Frank K. Chance, Don M. Investment Analysis and Portfolio Management, Fourth Edition, The Dryden Press (1994) 330-333
33. Rui, Ma, Hongwen, Yan. "A new data-mining framework based on decision-tree for competitive bidding assessment in power market." Automation of Electric Power Systems, vol.26, no.15, Publisher: State Power Corp. of China, China (Aug. 2002) 22-6
34. Simon, Herbert A. The Science of the Artificial, Third Edition, The MIT Press (1996) 170-174
35. Tao, Xu, Ren-mu, He, Peng, Wang, Dong-jie, Xu. "Power system transient stability prediction based on data mining." Journal of North China Electric Power University, vol.31, no.4 (July 2004) 1-4.
36. Todorovski, Ljupco, Dzeroski, Saso. "Combining Multiple Models with Meta Decision Trees." Principles of Data Mining and Knowledge Discovery. Proceedings of 4th European Conference, PKDD 2000 Prague, Czech Republic (Sept. 2000) 54-64.
37. Weiqiang Lin. Orgun MA. Williams GJ. "Temporal data mining using hidden Markov-local polynomial models." Advances in Knowledge Discovery and Data Mining. 5th Pacific-Asia Conference. Proceedings. Springer-Verlag, Berlin, Germany (2001) 324-35
38. Hightower, Ross, Sayeed, Lutfus. "Effects of Communication Mode and Prediscussion Information Distribution Characteristics on Information Exchange in Groups." Information Systems Research, vol. 7, No. 4 (Dec 1996) 462-464
39. Neumann, S., E. Segev. "A case study of user evaluation of information characteristics for systems improvements." Information and Management (1979) 276-277

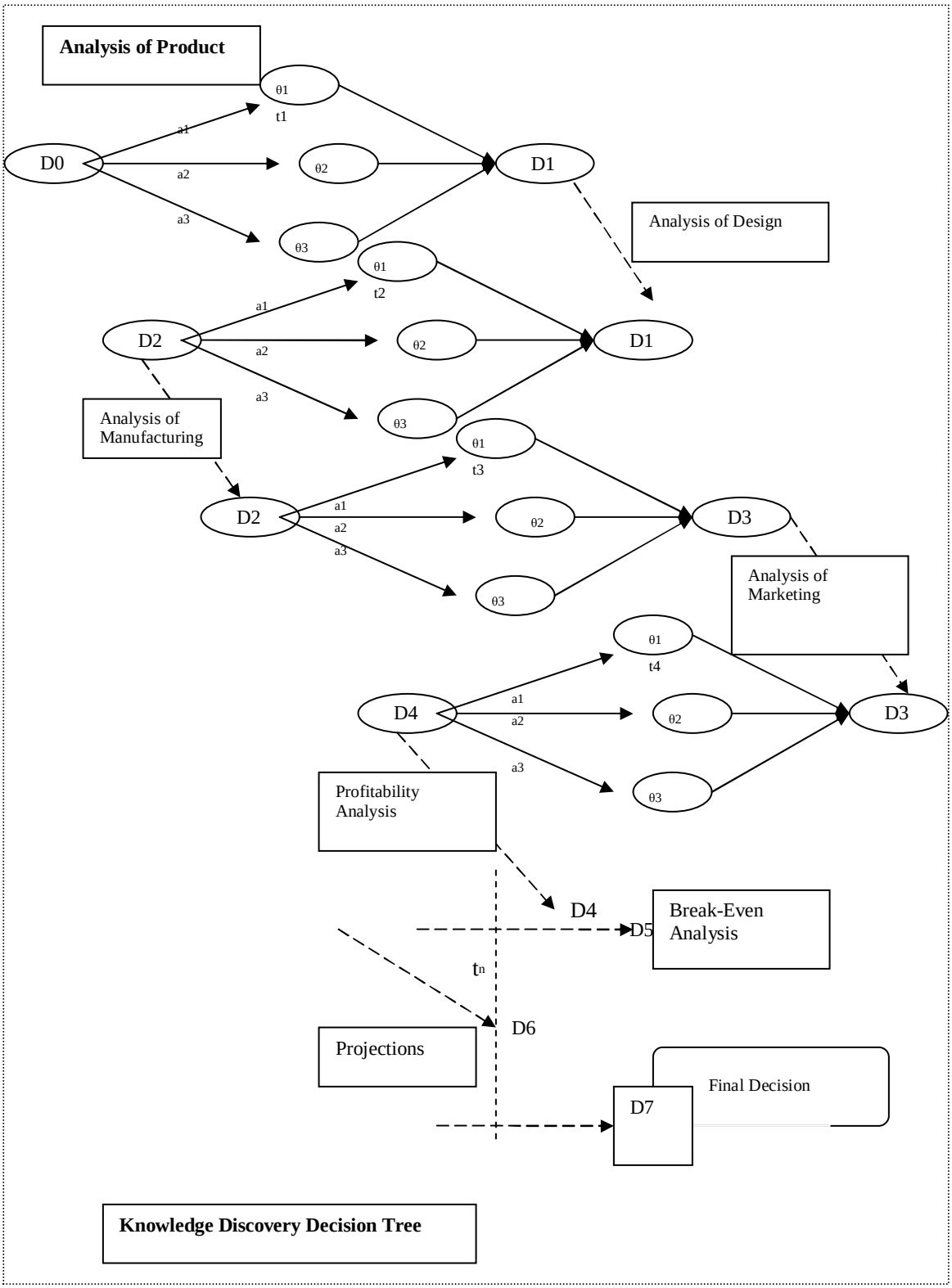
Appendix 1

Business Process of Idea Acceptance and Rejection



CAM= Computer Aided Manufacturing
CAA= Computer Aided Analysis
A= Approval
DA= Disapproval

Appendix 2


Analysis of Product
θ1
t1
a1
D0
a2
θ2
D1
a3
θ3
Analysis of Design
θ1
t2
a1
D2
a2
θ2
D1
a3
θ3
Analysis of Manufacturing
θ1
t3
D2
a1
a2
θ2
D3
a3
θ3
Analysis of Marketing
θ1
t4
D4
a1
a2
θ2
D3
a3
θ3
Profitability Analysis
D4
D5
Break-Even Analysis
tn
D6
Projections
Final Decision
D7
Knowledge Discovery Decision Tree