

Using Genetic Algorithms in an Electronic Market to Predict Future Best-Selling Products

¹H. Motameni and ²N. Mohsenian

¹Department of Computer Engineering, Sari Branch, Islamic Azad University, Sari, Iran

²Young Research Club, Sari Branch, Islamic Azad University, Sari, Iran

Abstract: Electronic commerce is rapidly expanding and is looking for organizations to develop and improve their online experiences. In this context, data mining tools for discovering new patterns in the data are used. The importance of data mining for e-commerce and these companies with the help of knowledge obtained from data mining can be in marketing, maintaining customer and improve profitability with tools and technologies of data warehousing, data mining and other CRM techniques. Also new methods provide new opportunities for trade. Therefore, through collecting data about consumer behavior and customer, decisions based on past patterns extracted from the hidden relationships among data-by-data mining tools, customer demands can be fulfilled by its axis and the profits of the seller constructions would be maximized. The concepts of this paper predicted selling products in the future, using genetic algorithms to maximize the dealer profits. The ease of selecting a buyer was evaluated.

Key words: Data mining % Genetic algorithms % Selling products % Customer demands % Dealer profits

INTRODUCTION

Progress has been attained in business culture in recent years. Economic relations with its customers, according to the basic and fundamental methods, are changing. Companies, in order to monitor such changes need solutions. The emergence and the rise of the Internet in order to change the center have had an important role in marketing. If the online information is available to make more awareness and consciousness, then there will be more customers that would be offered and those with best demand would be recommended. To manage this there must be some systems that can accurately respond to the customers' needs. Statistics, which are collected from customer data and the main purpose of their behavior, makes it possible and accurate. This type of targeting is a higher plan when creating a tough competition and identifying potential customers when new products release. In this context, data mining or knowledge discovery in databases [1] are a relatively new science that is meant to discover hidden knowledge and reliable information in the databases. On the other hand, with the advent of the new effects of artificial intelligence and ideas such as ants logarithm, genetic algorithm and market

mechanisms, new perspectives were created to improve the performance of software agents including the application of market mechanism in the implementation of electronic commerce simulator.

Ecommerce was used in the form of the exchange of electronic data during a few last decades especially in big companies and organizations [2]. Internet and web have changed the form of e-commerce during a few years ago. E-commerce has evolved in the field of the product sale through different models and generations [3]. Agent-based electronic markets are one of the outcomes of the use of agent IT in e-commerce [4]. The mechanism of the design of private agents as a buyer and seller has been the favorite topic of many researchers in university departments and industry because of the development of agent-based electronic markets [5]. Because of the importance of the role of buyer and seller agents in increasing the satisfaction of its users and for the maximum use of market conditions, the relevant researchers focused their attention on making buyer and seller agents more intelligent. Yet A lot of Dynamic Price Mark Model For Internet are Provided that Included: *metro Paris* Model [6] was provided Different Service Price for various classes. The main advantage of this

method is simplicity and distinctive service is provided. In [7] of the buyer and seller, agents in electronic factors based on market are used. Model presented in [8] using reinforcement learning, credits the seller agents to help the buyer agents will be modeling to avoid buying low-quality goods.

In this paper a model is proposed in which the seller based on customer information in database that is available, with buyer interest calculation model makes the best-selling products to scope the sales process in the future in order to maximize their sales by making offers to new buyers.

Data Mining in E-Commerce: Nowadays, huge volumes of data are facing. To use the tools they need knowledge discovery. Data mining as an advanced ability in data analysis and knowledge discovery is used. Data mining in science (astronomy) in business (advertising and CRM) on the web (search engines) in public issues (anti-terrorist activities is used). Indeed purpose of creating data mining models for decision-making is. Future behavior based on these models predicts the last analysis. The application of data mining as a lever to prepare the data and complete data storage capabilities, the best position to gain competitive advantages is created. Database systems, plays a key role in systems management and data warehousing,. A database system includes database files and database management systems data [9].

A field of e-commerce is essential for data mining. Considering the success factors (Critical Success Factors) related to the successful implementation of data mining, e-commerce provides an ideal platform because all the details required are readily available: the data records in abundance, many features of each case, high reliability of data because of automatic and electronic data collection, the capability of easily operationally insight (Actionable domain) and the ability to easily calculate the return on investment [10]. Most business or strategic decisions need to adopt new policies to better serve the customers. For example, stores store layout to create a greater desire to purchase or re-design and airlines have special facilities for customers to consider their flights frequently. These two examples require data on past customer behavior to determine consumption patterns through data mining. Based on these patterns the necessary decisions would be taken. In fact, data mining tools, data and takes the form of an image of reality make the model, this model ties in the data descriptions [11].

The Process of Data Mining in CRM: Data mining is one of the elements of CRM and move towards customer-oriented companies can help. Data are often bulky and it isn't used alone, but lies in the knowledge data can be used. So the power of data mining process is to identify patterns and models and associated elements in the database. Knowledge discovery lies in the data and finally, data conversion information. This is clearly understood that this type using data mining can shop smart in holding the festival selling goods to customers and how to provide help. So smart companies are able by customers based on their identified interests. It is also possible to identify high-risk customers and jobber based on some criteria such as age, income status, residence, education, jobs, etc.

CRM is a business process that all aspects of characteristics makes addressing customer. Customer knowledge creates relationships with the shape gives customer and withdraws them from the products or services. [10] CRM considers what products or services, what customers, at what time and through what channel is supplied. The management consists of several components. Before the process begins, the company's customer information must be provided [12]. Some of these companies includes large data banks data marketing, human resources and financial possess. Therefore, investment in data warehouse is one of the vital components of CRM strategy [13]. Customer relationship will change with time and if the business and customer know more about each other this relationship would be developed. Input in data mining, database tells what information is available at the output section, what is probably interesting and what decisions should be taken.

Genetics Algorithm: Genetic algorithms works based on Darwinian principle of natural selection by optimal formula for predicting or uses pattern matching. Genetic algorithms are used in research and optimization problems. Genetic algorithm works based on the initial input called population. A desired solution to the problem is calculated with a list of parameters shown (Chromosome) Or (Genome). Chromosomes generally are simple string. The data are displayed, although other types of data structures can be used. The first generation is created from the first population using mutation and crossover. During each generation, each attribute is assessed with Fitness value (fitness) measured by the fitness function. then Some Of best hypothesis is selected

and new population is generated. A GA algorithm has Parameters as following [14]: GA (Fitness, Fitness threshold, p, r, m), that (Fitness = function for assessment hypothesis, Fitness threshold = Threshold value, P = Count hypothesis, R= crossover probability and finally M= mutation rate).

Model Proposed: Our genetic algorithm was provided with the following parameters: GA (Fitness, 65, 100, 20 and 10).

Initialization: Number of members of the population of 100 chromosomes in the first generation based on the highest benefit to give the dealer has been selected. In genetics algorithm, usually hypothesis is a string A Of bits show the data. Gene's values are considered as parameters to the following sales. Much interest as one of the parameters mentioned because that might have made a small lot sales while profit is low so that repeats according to the fitness sharing, the value of a Zvdr Fitness members the same if repeated too much low-income exist in the population decreases. We proposed to calculate the duration once in a month or a season. For example, in a store, following amounts are considered in a week (Table 1).

Fitness Function Evaluation: Evaluation the function fitness criterion is done for next generation population selection. Here, based on the fitness function of each gene in different chromosome the percentage ratio was found. Then, to evaluate fitness of each chromosome, genes divided the total percentages on the number of genes (Table 2).

New Population Creation: Select count (1-r) p hypothesis of choice between P And Add to Ps .

$$Ps = Ps + 20$$

This Roulette Wheel selection method was used. Likely select a hypothesis of hi Between P: P (hi) = Fitness (hi) / Gj Fitness (hj)

C 10 hypotheses among 100 couples selected hypothesis for crossover operator. The children would be created and would be added to the to Ps .

Crossover operator used two parent strings and created children. To do this part of the bits is copy children of parents in bits. The bits in two parents would

Table 1: Products sold

Type	Gender	Age Group	Size	Color	Quality	Delivery time (days)	Profit%
Mantoux	Woman	Young	Great	Red	Good	2	5
Pants	Woman	Teenager	Great	Green	Higher	2	10
Pants	Woman	Teenager	Great	Blue	Good	4	5
Shirt	Man	Teenager	Average	Red	Higher	4	10
Shirt	Man	Young	Great	Blue	Good	2	10
Mantoux	Woman	Young	Small	Green	Higher	3	5
Mantoux	Woman	Young	Small	Red	Good	4	5
Shirt	Woman	Young	Great	Red	Good	2	10
Shirt	Man	Young	Average	Green	Good	2	5
Pants	Woman	Young	Average	Green	Good	3	5
Shirt	Woman	Young	Great	Green	Good	3	5
Mantoux	Woman	Young	Great	Green	Good	4	5

Table 2: Calculate eligibility and the possibility of choosing any product

Type	Gender	Age Group	Size	Color	Quality	Delivery time (days)	Profit %	Percent Competency	Percent probability
Mantoux	Woman	Young	Great	Red	Good	2	5	48.75	10
Pants	Woman	Teenager	Great	Green	Higher	2	10	48.75	10
Pants	Woman	Teenager	Great	Blue	Good	4	5	48.625	10
Shirt	Man	Teenager	Average	Red	Higher	4	10	47.625	10
Shirt	Man	Young	Great	Blue	Good	2	10	42.625	9
Mantoux	Woman	Young	Small	Green	Higher	3	5	41.375	8
Mantoux	Woman	Young	Small	Red	Good	4	5	40.25	8
Shirt	Woman	Young	Great	Red	Good	2	10	38.375	8
Shirt	Man	Young	Average	Green	Good	2	5	38.375	8
Pants	Woman	Young	Average	Green	Good	3	5	37.25	8
Shirt	Woman	Young	Great	Green	Good	3	5	36.25	7
Mantoux	Woman	Young	Great	Green	Good	4	5	26.875	5

Table 3: Select products deserve more than 65% after computing

Mantoux	Woman	Young	Small	Red	Good	4days	5 percent	65.125
Shirt	Woman	Young	Great	Red	Good	2days	10 percent	65.125
Pants	Woman	Young	Great	Green	Good	4days	5 percent	65.125

be exchanged, using single-point crossover, two-point crossover or uniform crossover with strings crossover mask, combined and new generation would be produced.

Here is recommended that only the mutation operator be used. Users combine two parent chromosomes when the cost is high.

C Mutation operator is used to create a child with only one parent. In this issue, we selected an elite community whose members were the first level of the population form the fitness function. Not also, selection probability of chromosomes with a higher score has more selection rate but also for all chromosomes, even for chromosomes with the lowest points have.

Total 10% of Members of the Ps selected and the amount of bits that repeats the percentage is less than 30%, choice and value bit in the other chromosome has a maximum percentage change finds.

C Updated selected generation is added to the previous generation and the process is repeated again. This process is repeated until the chromosomes produced chromosome that is likely the answer. Until $65\% > \max_h \text{Fitness}(h)$ a new population is created.

C At this time, a hypothesis, which has $\max_h \text{Fitness}(h)$, we give them back (Table 3).

Pseudocode:

- C Genetic Algorithm
- C Begin
- C Choose initial population
- C Repeat
- C Evaluate the individual fit nesses of a certain proportion of the

population:

{
Calculation the repetition percent for each gene in chromosome;

Total percent of genes in each chromosome divide in genes number for calculation the Fitness (hi);

$$P(h_i) = \text{Fitness}(h_i) / \sum_j \text{Fitness}(h_j);$$

}

- C Select pairs of best-ranking individuals to reproduce
- C Apply crossover operator
- C Apply mutation operator
- C Until terminating condition
- C End

CONCLUSION

Electronic commerce is rapidly expanding and in accordance with its economic ties to customers in basic and fundamental ways is changing. Organizations seek to develop and improve their online experiences. In this context, data mining tools for discovering new patterns in the data are used. Data mining is one of the elements of CRM and can move towards customer-oriented companies. Today's companies through analysis of customer interests have achieved increased customer value. Therefore, collecting data about consumer behavior and customer decisions based on past patterns extracted from the hidden relationships among data-by-data mining tools can be pivotal for fulfilling customer demands. CRM considering what products or services, what customers, at what time and through what channel is supplied. Therefore, through data mining to determine customer value, their future behavior would be predicted and informed and e-commerce would be adopted. In this paper, using genetic algorithms, the best-selling products were calculated. Identifying the best-selling products and advertising these products to customers and increase such products, increase sales and profits. As can be seen in the chart on the normal selling products of large retailers, their income increased slightly, while the genetic algorithms will further increase these benefits.

REFERENCES

1. Mobasher, B., H. Dai, T. Luo and M. Nakagawa, 2002. Discovery and evaluation of aggregate usage profiles for web personalization, Data Min. Knowl. Disc., 6(1): 61-83.

2. West, P.M., D. Ariely, S. Bellman, E. Bradlow, J. Huber, E. Johnson, B. Kahn, J. Little and D. Schkade, 1999. Agent to the Rescue. *Market. Lett.*, 10(3): 285-300.
3. Kalakota R. and A.B. Whinston, 1996. *Electronic Commerce: A Manager's Guide*, 1st ed. Addison-Wesley Professional, pp: 210-217.
4. Fitzpartrick, T., 2001. *Electronic Auction House First To Offer Mobile Agent*. Washington University in St. Luis, pp: 18-20.
5. Chavez, A. and P. Maes, 1996. Kasbah: An Agent Marketplace for Buying and Selling Goods. In *Proceedings of the First International Conference on the Practical Application of Intelligent Agents and Multi-Agent Technol.*, pp: 87-93.
6. Odlyzko, A.M., 1999. Paris metro pricing for the Internet", *Proceedings of the ACM Conference on Electronic Commerce*, pp: 140-147.
7. He, M., N.R. Jennings and H.F. Leung, 2003. On Agent-Mediated Electronic Commerce, *IEEE T. Knowl. Data En.*, 15(4): 985-1003.
8. Howell, M.N. and T.J. Gordon, 2001. Continuous action reinforcement learning automata and their application to adaptive digital filter design, *Engineering Applications of Artificial Intelligence*, pp: 549-561.
9. Kargupta, H., A. Joshi, K. Sivakumar and Y. Yesha, 2005. *Data Mining: Next Generation Challenges and Future Directions*, Prentice Hall of India, pp: 157-219.
10. Dzeroski, S. and H. Blockeel, 2004. *Multi Relational Data Mining 2004: Workshop Report*, *Sigkdd Explor.*, 6(2): 140-141.
11. Kirschner, M., G. Pujol and A Radu, 2002. Oligonucleotide microarray data mining: search for age-dependent gene expression. *Biochem. Biophys. Res. Commun.*, 298(5): 772-778.
12. Liu, P., 2007. Evaluation Model of customer Satisfaction B2C E-Commerce based on Combination of Linguistic Variables and Fuzzy Triangular Numbers. In *eight ACIS international conference on software engineering, artificial intelligence, networking and parallel/ distributed computing*.
13. Zheng, Y., 2005. Customer satisfaction-the most important task of electricity service. *China Qual.*, 3: 65-66.
14. Kuncheva, L.I. and L.C. Jain, 2000. Designing Classifier Fusion Systems by Genetic Algorithms, *IEEE T. Evol. Computation*, 33: 351-373.