Maximizing Sales of a Finance Company: A Genetic Algorithm Approach for Customer Identification

Varun Parmal
Assistant Professor,
Malwa Institute of Science & Technology
Indore, India

Snigdha Banerjeeb
Professor,
School of Statistics,
Devi Ahilya University,
Indore, India

Abstract

This paper deals with the problem of optimization of sales of a company involved in selling various financial products. Based on the data available from the company, identification of target category of customers for specific financial products is attempted. Genetic algorithm technique is applied for obtaining near optimal combinations of the customer category and product category which result in good sales. It is shown that Life Insurance for the average ticket size of Rs 1 to 5 lakhs should be targeted at those customers who belong to the age group 20-30 yrs and are in the salary group Rs. 25,000-45,000. Similarly, it is shown that debenture sold by the finance company under consideration is a hot seller in the age group 40-50 yrs for the persons having monthly salary in the range Rs. 45,000-60,000.

Keywords: Genetic algorithm, finance company, insurance, debenture, mutual fund.

Introduction

The problem of variation and chaotic behaviour of customers in the face of lack of sufficient information is a challenge to many financial organizations, which may result in many missed corporate business opportunities. In order to embrace more business opportunities and build up and enhance competitiveness, discovery of unexplored patterns and extraction of useful rules from large databases generally leads the business to a vantage point. Adopting proper marketing strategies generally results in enhanced profits of any finance company.

In particular, for any finance company involved in selling various finance products, there are many criteria which affect the tendency of customer to purchase a product. Such criteria have to be identified based on the historical past sales records of the company. The product purchasing behaviour of the customers will generally depend on the customer's age, previous purchasing experience, income etc. and can be extracted in terms of probabilities or proportions from the past sales records of the firm. This enables the firm to evolve marketing guidelines and sketch a roadmap for success. Thus, it is very imperative for any finance firm to study customer's profile and then target the product based on the guidelines systematically evolved from past behaviour of similar customers. As a result of such a study, company agents can recommend the right product with the correct

investment value to a customer in a specific category. It will then be easy to convince the customer to purchase the suggested product without spending much time and effort. This will help in maintaining the growth of the company in terms of wide market visibility and improving its sales.

Genetic Algorithm (GA) is a heuristic search technique for optimization in situations where it is not possible to analytically establish the extreme values of the function. It employs a strategy based on the theory of natural selection to obtain iterative refinement of a population of potential solutions. It has been applied to diverse fields in problems like Travelling salesman problem (TSP), Marketing, finance etc.

In this work, we address the problem of optimization of sales of a company involved in selling various financial products like life insurance, mutual funds and debentures of various ticket sizes. Based on the data available from the company, identification of target category of customers for specific financial products is attempted.

Genetic Algorithm(GA) is a heuristic search technique for optimization, where it is not possible to analytically establish the extreme of the function. It employs a strategy based on the theory of natural selection to obtain iterative refinement of a population of potential solutions. It has been applied to diverse fields in problems like Travelling salesman problem (TSP), Marketing, finance etc.

In this work, we address the problem of optimization of sales of a company involved in selling various financial products like life insurance, mutual funds and debentures of various ticket sizes. Based on the data available from the company, identification of target category of customers for specific financial products is attempted. Genetic algorithm technique is applied for obtaining near optimal combinations of the customer category and product category which result in good sales. For this, we first study the factors which affect customer behaviour towards purchase of the product as suggested by the company analysts - in terms of variation in their age, income and ticket size of their previous purchase if any, from the same finance company. Using Genetic Algorithm, we have tried to find the solution of the major problem of any finance company selling various financial products like insurance policies, debentures etc. - namely, which customer segment should be the major target for each given product category (product, ticket size), so as to increase the sales of the company significantly.

We have addressed the problem of deciding upon the target customer category on whom the company should focus upon selling a particular product of a specified ticket size, in terms of customer's age group and his income group in both conditions whether he/she has experience of prior businesses with the company or not. We have tried to use GA in a new framework, wherein we have used probabilities of occurrences of different groups of customers in evaluating the study at each step through fitness function. This, to the best of our knowledge and belief, is the first attempt at using GA for such a study, especially, in a probabilistic framework.

The remaining part of this paper is organised as follows: section 2 presents literature review. section 3 explains the methodology of GA with an example; section 4 presents the model developed in this work. section 5 explains result and analysis, section 6 explains the limitations and difficulties faced by us in use of the GA and presents the conclusions.

Literature Review

Genetic Algorithm(GA) has been used by many researchers for solving the Travelling Salesman Problem(TSP) for example, Bryant et al (2000), Abeysundara et al (2005) etc.. Another problem solved using Genetic Algorithm by Huang et al (2005) is that of a supply chain model.

Genetic Algorithm has also been applied in many problems in stock market and other finance fields. Lettau(1997) studies portfolio decisions of agents in a financial market. Shin and Lee(2002) illustrate how GA can be applied to bankruptcy prediction modeling. Lin et al (2004) apply GA for selection of values for all parameters and their combinations so as to attain success of a trading rule by finding a near optimal value in a reasonable time. Lin et al (2005) use genetic algorithm to automatically search for improved trading models. Tsai et al (2011) consider the problem of deciding the signal for buying or selling funds and invoke a genetic algorithm to dynamically select funds according to their past performances (profitability). Bonde and Khaled (2012) use genetic algorithm to predict whether the highest prices of the stocks of companies are going to increase or decrease on the next day on the basis of six attributes. Kozeny(2015) evaluates the predictive performance of different fitness functions used by genetic algorithms in credit scoring. Aguilar-Rivera et al (2015) present a review of the application of evolutionary computation methods including GA to solving financial problems.In the present work, to the best of our knowledge, for the first time attempt has been to use Genetic Algorithm in a probabilistic frame work.

Finance Company Problem

In a company that sells financial products, several factors contribute to growth of the firm. While firms try to improve the services to their clients, in order to achieve growth of the company in terms of improved sales, it is required that the company studies the behaviour of its existing and potential customer. This step can help in finding such aspects of customers which affect the sales.

While there is a wide range of analysis techniques to customize the investment portfolio for the customer, there is still a need to develop an intelligent method for extracting customer behaviour in the financial industry, so as to increase the availability and utilisation of decision support data and hence increase customer satisfaction. It is to be noted that targeted customers are the better customers. Therefore extracting useful rules for targeting specific customer categories can be provided through an insight into customer's buying behaviour.

In this work, we use probabilities based on data collected from a financial company. We aim at using Genetic Algorithm method to identify such factors which will not only improve the cash inflow of the company, but also deepen understanding of investment behaviour and identify the category of customers who will turn out to be good investors for various product categories.

Research Methodology:

This technique of Genetic Algorithm requires utmost care in building the model, especially the fitness function and chromosomes. Once the model is developed, computer intensive methods give near optimal solutions.Our study relates to only three types of products which are debenture, Life Insurance policies and Mutual funds where we further classify the debentures as well as Life Insurance into five different ticket sizes D1- D5 and L1- L5 respectively.

Table 1: Different Groups of the factors for the construction of chromosomes

Product				*Customer*			
Life Insurance		*Debenture*		*Income Group*		*Age Group*	
Code	*Average Ticket size in Thousand Rupees*	*Code*	*Ticket size in Thousand Rupees*	*Code*	*Monthly Incomein Thousand Rupees*	*Code*	*Age Group*
L1	25	*D1*	25	1	Below 10	1	Below 10
L2	75	*D2*	75	2	10-25	2	10-20
L3	200	*D3*	200	3	25-45	3	20-30
L4	400	*D4*	400	4	45-60	4	30-40
L5	750	*D5*	750	5	Above 60	5	40-50
						6	50-60
						7	60-70
						8	Above 70

Table 1 presents the categories for various factors under consideration

The Genetic Algorithm (GA) is a relatively new optimization technique to find the solution of the various problems. This technique does not ensure that the solution obtained is optimal. However it provides good approximation to optimal solution in a reasonable time period. (Deb (2004)) Based on Darwin's principle of "survival-of-the-fittest", this algorithm is quite flexible. The principle used in GA is that if an offspring produced by genetic processing is above-average, it is more likely to survive longer than an average individual. If, on the other hand, if a below-average offspring is created, it is not likely to survive long.

Beginning from an existing population of solutions called chromosomes and a fitness function defined depending upon the problem we make use of the GA operators – Reproduction, Crossover and Mutation to generate new populations (solutions). In the new population, we apply the law of survival of the fittest and select the chromosomes that are fit enough to remain in the population and delete the unfit chromosomes. By continuing this process, at each step, we get an improved population of fit chromosomes and find the solution i.e. identify the right product for the respective customer categories.

We now present some details of these three operators.

Reproduction: Reproduction (or Selection) is usually the first operator applied to a population. This operator selects good (fit) chromosomes in the population for generating an offspring. The commonly used reproduction operator is proportionate selection operator, where a chromosome in the current population is selected with probability proportional to the chromosome's fitness.

Crossover:Crossover operator is next applied. This operator picks two chromosomes randomly from the population and exchanges some portion of the chromosomes. It is interesting to note from this construction that good offsprings from either parent chromosome can be combined to form a better offspring.

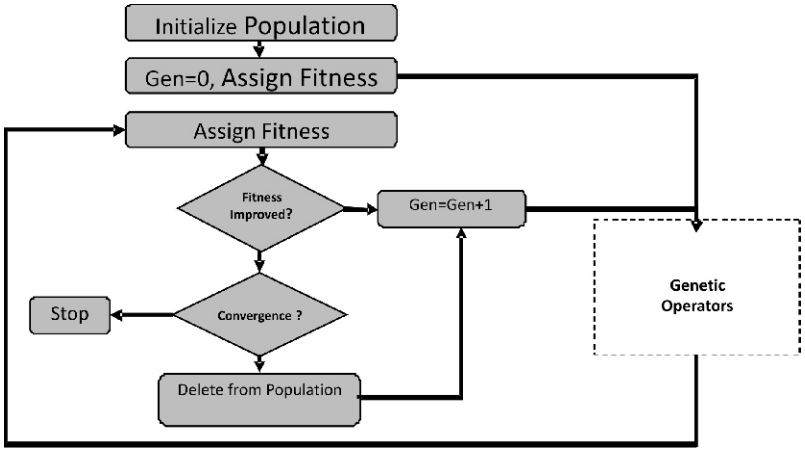
Mutation: The need for mutation is to maintain diversity in the population.The mutation operator changes 1 to a 0 and vice versa with a small mutation probability.

0 0 1 1 1 1 1 1 $\longrightarrow$ 0 0 1 1 1 1 0 1

Here, the seventh bit 1 of the chromosome is changed to 0, thereby creating a new solution.

Figure-1shows, how beginning from initial population of chromosomes successive generations are created using Genetic Algorithm:

Figure – 1: One cycle of Genetic Algorithm

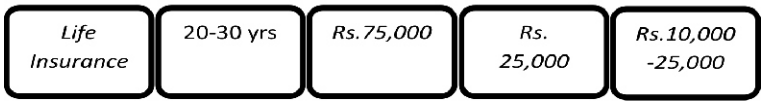


Proposed Model And Methodology Of GA

In this work, as explained by the company experts, we have used the customer purchase behaviour factors to be: customer's age, income group, , history of earlier purchase from this company , product along with its ticket size. We wish to use the past records of customer penchant, we are prompted to develop a probabilistic model for the fitness function of Genetic Algorithm. In order to apply this, we have found the probabilities of the different customer age groups , the conditional probability of the various products with assumed ticket size, previously purchased product and ticket size and also marginal probability of different income groups of the customers. We then construct the five bit chromosomes using these factors. The first bit in chromosome is taken as product, second bit as customer age group, third bit as ticket size, fourth bit as previously purchased product and the last bit as income group. We have taken each of these factors as one bit of chromosomes as shown in figure – 2.

Figure – 2: Representation of Chromosome

A typical chromosome (above chromosome decoded).



Since customer behaviour is a random variable, we have designed the fitness function of the associated problem as a probabilistic function which estimates the expected sales value for the individual chromosomes. As we improve the sales, we maximize the fitness function. We define the fitness function to be maximised as

$$F(x) = AVG\ T_i \times [(1 - repurchase) \times \Pr(customer\ age\ grouop\ j) \times \Pr(purchasing\ a\ ticket\ size\ k | age\ group j) \times \Pr(income\ group\ l) + (repurchase) \times \Pr(customer\ age\ grouop\ j) \times \Pr(previous\ purchasing | age\ group\ jand\ ticket\ size\ k) \times \Pr(purchasing\ a\ ticket\ size\ k | age\ group j) \times \Pr(income\ group\ l)] \times N \quad (1)$$

Where, the notations are as follows:

$AVG\ T_i = Average\ ticket\ value\ for\ a\ particular\ product\ i.$

$N = Total\ Number\ of\ sales\ in\ the\ population$

and

$$\text{Repurchase} = \begin{cases} 1, & if\ customer\ has\ purchased\ earlier \\ 0, & if\ customer\ has\ not\ purchased\ earlier \end{cases}$$

Pr(A|B) = Conditional probability of event A given the event B.

Pr{Event = Probability of the event}

We have assumed that income group is independent of age and repurchases. This is true in current real life scenario where young executives earn much more than older clerks/school teachers.

Equation (1) presents a probabilistic fitness function for model in order to enable the evaluation of the chromosomes,

as we are interested in generating the population of profitable chromosomes. Further analysis provides us with good combinations of customer factors for each product and ticket size. For practical reason, we have fixed the product with individual ticket size, so that we could find the particular age group, income group and earlier purchase information for each individual ticket size of respective products.

Results and Analysis

We adapted the procedure as describe in figure-1and the best solutions are presented in table 2

***Table* – 2:** *Best Solutions for each ticket size with product*

S. No.	Product	Ticket grp code	Age grp	Earlier purchase	Earlier prdct	Avg ticket size in thousand rupees	Monthly Income group in thousand rupees
1.	Debenture	D1	40-50	yes	D1	25	below 10
2.	Debenture	D2	60-70	no		75	25-45
3.	Debenture	D3	60-70	no		200	45-60
4.	Debenture	D4	60-70	no		400	60- 100
5.	Debenture	D5	60-70	no		750	100-250
6.	Life Insurance	L1	40-50	no		25	below 10
7.	Life Insurance	L2	30-40	no		75	25-45
8.	Life Insurance	L3	20-30	no		400	45-60
9.	Life Insurance	L4	30-40	no		400	60- 100
10.	Life Insurance	L5	30-40	no		750	100-250
11.	Mutual Fund	MF	30-40	yes	D1	5	below 10

In order to carry out modelling and obtaining near optimal solution using Genetic Algorithm, first we have taken a simple random sample of the past record of about ten thousand sales. We have found the probabilities of the different age groups of the customer for the proposed model as described in equation (1). The chromosome is taken over the customer identification factors – age, income and previous purchasing status and the company factors – products and ticket size. For all these factors the probabilities are estimated from the sample of the firm (Finance Company) for evaluating the fitness function of the problem

Conclusion

In this work, we have presented near optimal solution of the problem of the finance company for sales of each product, since the company is interested in recognizing the combinations of the customer category and product category which would result in good sales. For this, while we have used the familiar Genetic Algorithm technique, we have added a new dimension to it by considering a probabilistic framework. Excel macros were developed by us for the large amount of data processing. We observed that the LI is good to be sold to those customers who belong to the age group 20-30 yrs with the salary group Rs. 25,000-45,000 for the average ticket size of Rs. 100,000-5,000,000. Similarly, Debenture is a hot seller in the age group 40-50 yrs., with the salary group Rs. 45,000-60,000 for the average ticket size of Rs. 100,000-300,000. We have also identified the individual ticket size for the different products as follows:

As Genetic Algorithm was applied for each of the three products (Life Insurance, Debenture and Mutual Fund) and each of the eleven sub products, conditional, joint and marginal probabilities had to be calculated separately for all these using the past data the fitness function for the GA applications. Each application required about hundreds of runs of the macro before reaching convergence of the fitness functions. However, the method employed is easy once the macros are ready. It is hoped that this methodology and results would benefit many finance companies facing similar problems.

References

Deb, K. (2004). Genetic algorithms for optimization. In D. Kundu and A. Basu (eds.) Statistical Computing: Existing Methods and Recent Developments. New Delhi, India: Narosa Publishing House, (pp. 85-123).

Huang G.Q., Zhang X. and Liang L. Towards integrated optimal configuration of platform products, manufacturing processes, and supply chains. Journal of Operations Management. 2005, 23, 267-290.

Krutika D Joshi, Amit A Pandya. Genetic Algorithms and Their Applications – Traveling Sales Person and Antenna Design. Publication – Department of Information Technology, Faculty of Technology, Dharmsinh Desai University, Nadiad, Gujrat. Sep. 2010.

Kylie Bryant, Arthur Benjamin – Prof of maths, Johns Hopkins University. Genetic Algorithms and the Traveling Salesman Problem. Department of Mathematics, Harvey Mudd College. Publication – Clarement, US. December 2000.

Li Lin, Longbing Cao, Chengqi Zhang, GENETIC ALGORITHMS FOR ROBUST OPTIMIZATION IN FINANCIAL APPLICATIONS, www-staff.it.uts.edu.au/~lbcao/publication/ci2005.pdf.

Li Lin, Longbing Cao, Jiaqi Wang, Chengqi Zhang, The Applications of Genetic Algorithms in Stock Market Data Mining Optimisation, http://www-staff.it.uts.edu.au/~lbcao/publication/DM2004.pdf.

Manuel Valenzuela-Rendon, J.J. Rodriguez-Ortiz, Genetic algorithms and Darwinian approaches in financial applications: A survey Rubén Aguilar-Rivera , Expert Systems with Applications 42 (2015) 7684–7697.

S. Hurley L. Moutinho and N.M. Stephens. Solving marketing optimization problems using genetic algorithms. Department of Computing Mathematics, University of Wales, Cardiff. December 1994.

Sachith Abeysundara, Baladasan Giritharan and Saluka Kodithuwakku. A Genetic Algorithm Approach to Solve the Shortest Path Problem for Road Maps. Proceedings of the International Conference on Information and Automation, December 15-18, 2005, Colombo, Sri Lanka.

Shin K.S. and , Lee Y.J., (2002) "A genetic algorithm application in bankruptcy prediction modelling", Expert System with application ,PP.

Sufal Das & Banani Saha, Data Quality Mining using Genetic Algorithm. International Journal of Computer Science and Security, (IJCSS) Volume (3) : Issue (2) 105-112, 2009.

Sung-Hwan Min, Jumin Lee, Ingoo Han, Hybrid genetic algorithms and support vector machines for bankruptcy prediction, Expert Systems with Applications 31 (2006) 652–660.