January 26, 1999

project title: Customer's Profitability Predictive Model
customer: CDE Mondadori Group
industry: Retail (Publishing, Mail Order)
function: Marketing
project contact: G. Cuzzocrea, F. Ambrogi (NUNATAC)

Customer situation/problem description

The Mondadori Group is a collection of subsidiary companies operating in the publishing industry. One of these, Club degli Editori (CDE - Publisher's Club) was founded in 1960 to sell books by mail order using the book club formula. It is currently a leader in Italy with a turnover of approximately 100 billion lire, and an annual profit of approximately 10 billion lire. CDE has some 900,000 members, to whom it sells some 8 million books annually from all publishers.
The Gruppo Mondadori counts about two thousand staff. CDE is a small unit of less than one hundred people, ten of which are SAS Users. The introduction of SAS System into the company, in 1994, represented quite a turn-around, giving rise to the use of the most advanced statistical techniques.

CDE sends members 13 34-page catalogues per year containing about 100 titles among which 10 new entries and the book of the month, an outstanding work selected by publishing experts. The only obligation imposed on the member is to buy at least three books per year so that his membership isn't suspended; nevertheless, he can cancel his membership at any time.
➤   not suspending low-profitable members for a long time might be expensive.

CDE-Mondadori has a recruitment program for new members during the year. Recruitment is made possible through the acceptance of an offer that, to be capable of capturing the attention of people that generally avoid mail-order purchases, has to be very attractive. This kind of acceptance doesn't allow one to distinguish between a member that is only interested in the initial promotion and one that is really interested in reading. It is probable that only the latter will continue to order from the catalogues he is sent monthly, becoming a good and profitable customer.

Project Description

"CDE operates in a highly competitive market and a harsh environment. So in order to succeed, we are left with only one option: we must increase the performance of our customers." Ivano Maestri, General Manager of CDE SpA.

"The analysis is designed to give us an indication as to whether new customers - those who have been members only for one or two months - are likely to become profitable long-term customers. This is clearly a strikingly important result which will allow us to guide our marketing policy in an extremely efficient manner." Giovanni Lux, Marketing Analysis Service Manager in 1996.

The objective of the analysis is to predict the client's behaviour in the shortest possible time and, consequently, predict the prospective profitability.

Correct indications of this type would allow the company, not only to invest upon the desired group of members, but to apply differentiated marketing strategies. In synthesis the project intends to increase the company's profitability.

> ➢ From an operative point of view, the goal of the analysis is mostly to decide, in advance, the exclusion of non profitable book club members: given that, it is extremely important to carefully define the error measures.

Solution

When the here described project was at the set-up phase, the SAS Data Warehouse project was nearly completed at CDE-Mondadori.
A large amount of data was actually available for the Marketing Analysis Service department. An exaustive and integrated set of data analysis tools was also available.

The opportunity to consider The SAS System's traditional data analysis tools and new tools, such as neural networks, to provide further insights into consumer could allow the company to milk more money from this consolidated market.

The availability of data and software in the SAS environment, the necessity to solve a problem of forecasting and classification having at our disposal a consistent an complex information base lead as to consider the %TNN family of SAS macros. These macros allow one to train MLP (multilayer perceptrons) neural network architectures by specifying all the characteristics or just accepting default parameters.

Extracting data from the Data Warehouse, we built up a denormalized Analysis Data Mart, organized one record for each statistical unit (i.d. for each member). This Data Mart is the input data source for the Mine in cycles step, and the score data table for the Implement in production step.

Project Summary

The project team consisted of:

> ➢ 1 Database Marketing expert and 1 Data Miner from Nunatac,
> ➢ 1 Statistitian, 1 List Manager and 1 Project Leader from Marketing Analysis Service department of CDE.

The project started in early September 1995 and was completed round March 1996.
During that period, the project team was involved part-time, with an average effort of 2 days a week.
This amounts to a total of around 6-8 man-months.

Project Methodology

Define Business Problem:
> ➢ Which is the correct way to measure one member's profitability?
> ➢ When CDE Mondadori can consider a customer definitely profitable?
> ➢ If CDE stops investing on members classified as non profitable by probabilistic rules, how high is the risk of losing good members?
> ➢ How much would the company save by implementing an automatic drop-off of non profitable members?

These were the key questions the project team was asked to answer. The Finance and Administration department was consulted and economic formulas generally applied to an entire list of prospects, to a recruitment channel (press, mailing, TV advertising) or to a geographical area had to be converted in per-capita equations.

As prospective profitability was of interest and not short term member behaviour, the Finance and Administration department was also asked to suggest the lifetime period necessary to definitely assess one member's profitability.

- The success of the project should be measured in terms of reducing the costs because the total number of mail messages to be sent decrease, and increasing the sales' return because there is a little number of people seeking to exploit only the first convenient offer.
- Neural networks were considered as a potential powerful analysis tool from the very beginning.
- The matter of how to implement in production the Data Mining results was also considered from the very beginning and lead to the idea of building up a Analysis Data Mart.

Evaluate Environment:
About the evaluation of organizational readiness, some considerations have to be mentioned:
for
- database marketing and targeting were considered critical factors for business advantage;
- there was no lack of needed resources;
- considerable investments in the IT department had already been made and a Data Warehouse environment was available;

against
- top management was not yet confident in Data Mining tools reliability;
- the marketing function was used to take decisions according to traditional and conservative rules;
- the Marketing Analysis Service was considered both an operative and a research department, but not a strategic one where new business rules were established;
- implementing in production the solution was not considered a must.

Make data available:
The availability of a Data Warehouse environment was definitively a great advantage. Nevertheless this environment was still in a testing release so the project team had to experiment:
- difficulties in extracting data from the Warehouse because user requests had low priority with respect to data administrators;
- data were not yet certified and great effort had to be invested in data cleaning and data validation;
- missing values were another pain in the neck and we considered mostly business knowledge and base statistical techniques to solve the problem.

It is important to highlight the screening data activity was extremely useful and offered great benefits to the Data Warehousing project team.

Because of the previous statements, if you consider the project time-frame as a whole, the Make data available activity was responsible of 40% of the total efforts.

The main deliverable of this step of the project was a denormalized data table (Analysis Data Mart, 1 record vs. 1 customer), considering data from the warehouse detail tables described in the next figure:

S

Fig. 1   The informartion base
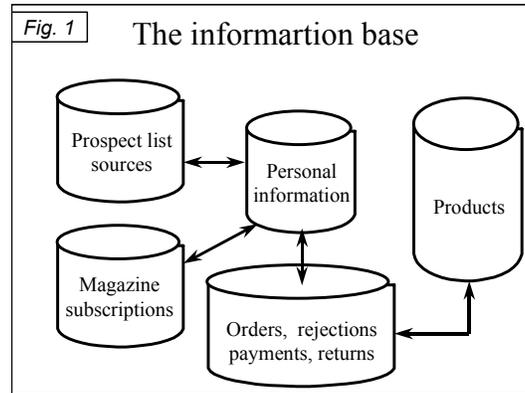
Mine in cycles:

Sampling
The problem of client classification with respect to prospective profitability imposes the analysis of homogeneous cohorts of members: all members in the sample have to have had the same opportunities to purchase, i.e. they have to have received the same number of catalogues.
Only members that joined the club between 1/1/94 and 30/6/94 and had received 8 catalogues by 31/3/95 were considered: In this way as at 31/8/95 (the starting point of the analysis) no payments were pending.
From this population a random sample was extracted at a rate of 1 / 10.

The target variable definition
The profitability of each member is evaluated considering a cost/gain balance determined by payments vs. mailing costs and returns/non payments.
The dependent variable is the profitability cumulated over a period of 8 catalogues.
Even though profitability is a continuous variable, the aim of the analysis is to predict in which of four profitability classes a member might fall. Therefore, depending on the neural network architecture, the dependent variable has been introduced either as continuous or categorical.
The four profitability classes have been constructed on the basis of economic considerations set out by the CDE Mondadori's marketing management.

Explore
Preliminary insight on data matrix, univariate distributions, missing value and outlier detection were extremely important and produced useful feed-backs for the main Data Warehouse certification phase.

Input variable selection
Both automatic stepwise techniques and business knowledge were considered in input variable selection.

Training and testing
Neural networks may be considered to be iterative algorithms capable of learning to recognize empirical laws on the basis of a vast number of relevant examples. With that in mind it is opportune to separate data into two sub-samples: the first will be the training set and the second the test set upon which the predictive ability of a network is gauged.
After the training phase the network has to be validated over the test set. It is extremely important to verify whether or not the knowledge gained by the network can be satisfactorily generalized. In other words, it is necessary to verify how well the network performs in the prediction of unobserved cases.
In Data Mining, the same approach has been extended to all kinds of modeling techniques.

In this case the input sample was randomly divided into two sub-samples using the SAS/BASE RANUNI (Random Uniform) function.

Model

S

➢ MLP Neural networks and Multiple Linear Regression were considered.

Fit and generalization
The relationship that exists between goodness of fit and the capability to generalize of neural networks corresponds to the concept of robustness in traditional statistical modelling: under equal conditions, the greater the number of variables introduced, the less robust the model.
In the case of neural networks the problem is even more complex: one should not only avoid the introduction of non significant input variables, which only cause noise, but it is equally important to minimize the complexity of the algorithm. If one introduces too many neurons and hidden layers or runs too high a number of iterations to perfectly fit the training data, then the neural network learns the patterns off by hart heart. The neural network becomes less and less capable of generalizing: you can rarely lend a tailor-made suit to a friend.

Measures of the misclassification error
From an operative point of view, the goal of the analysis is mostly to decide, in advance, the exclusion of non profitable book club members: given that, it is extremely important to carefully define the error measures.
The Marketing Analysis Service department wanted to give a correct answer to the following question: if CDE stops investing on members classified as non profitable by the network, how high is the risk of losing good members?
On the other hand, an important piece of information we obtained from our analysis is the percentage of truly profitable members among those who are predicted to be the best by the network.

Simulation trials
Within the customized definition of error measures, we obtained a significant justification of the choice of the network architecture and, at the same time, we could validate the model choice by running numerous simulation trials based on re-sampling techniques. At each trial the training and test sub-samples are randomly reallocated. Then the network is trained and tested, monitoring step by step the error percentages. For each kind of network and regression model we considered the summary results of at least 100 simulation trials.

➢ Both MLP Neural networks and Multiple Linear Regression performed quite well and results were comparable, slightly better in the NN case.

Implement in production:
Some considerations are extremely important to be highlighted:
➢ even a small difference in model performances is determinant when your database contains millions of potential customers. Therefore Neural Networks should have been implemented;
➢ if you have to knock down a cultural wall, then results coming from a Linear Regression Model are easier to understand for non statistitians;
➢ the existence of an Analysis Data Mart (organized per statistical unit) is a must if you want to easily score your customer universe;

Review:
The Marketing Analysis Service department is going on in the transformation of traditional marketing approach into Database Driven Direct Marketing.
The automatic drop-off rule has not definitively replaced traditional criteria. Furthermore, the new tools is being used in a conservative way in order to minimize the error of dropping-off even a small percentage of potentially profitable members.
In a sufficiently long period of simulated operative use, members predicted non-profitable by the tool produced a negative profit of 5.000 Lire per capita, while those predicted profitable a positive one of 40.000 per capita.

S

Lessons learned

➢ when you implement in production a new rule coming from Data Mining, never forget traditional rules: think about parallel processes, control samples, campaign tests and so on;

➢ provide a data environment to measure results and be prepared to sacrifice the cost of targeting less profitable segments so that you can make an accurate cost/benefit analysis ex-post.

Future developments

During the last months of 1998 Mondadori and Bertelsmann Groups created a new company, 'Club del Libro', that will be the result of merging between CDE itself and Euroclub, owned by Bertelsmann and very similar for membership procedures to CDE. The new company will count about 200 people and the use of SAS System will probably increase because of the problems arising from the fusion of the two mail order companies and the great need of advanced analysis in this kind of business.

Additional information

List of SAS System products installed during the project frame:
Base, FSP, Graph, Connect, STAT, IML, OR, AF, EIS.