EXCERPTED FROM

DECEMBER 2012

CHARITYFINANCE

civilsociety.co.uk

The voice of civil society

Social finance

First steps towards maturity

CHARITY AUDIT SURVEY 2012

interview with William Shawcross • reactivating lapsed donors • working in consortia • insolvency

Rubies in the dust

Ian Harris and Mary O'Callaghan explain how predictive analytics can boost returns from lapsed donors in your fundraising database.

CHARITY FINANCE professionals often need to predict future results, and also attribute those results to the activities that caused them.

There are plenty of techniques available to help predict and attribute: portfolio theory, Monte Carlo simulation and linear programming are good examples.

We tend to use these techniques in charities to solve problems such as cost-benefit optimisation of charitable activity, rebalancing a charity's investment portfolio, or setting reserve-level ranges.

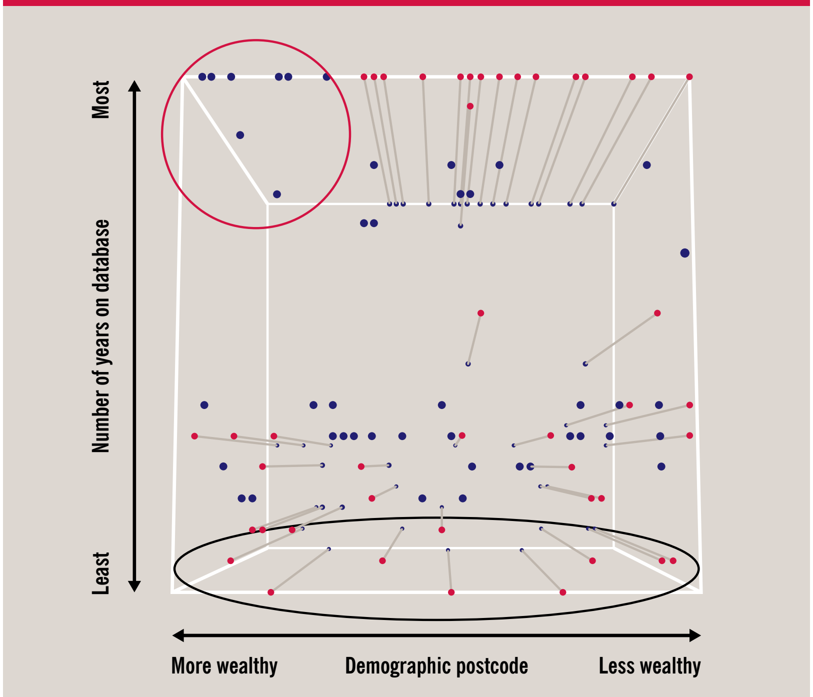
Modelling data

In recent years we have been using statistical learning models, known as support vector machines (SVMs), to predict and classify individual items of information. In simple terms, SVMs are learning algorithms, especially suited when you need to model data in order to classify items and predict outcomes.

> "The predictive analysis produced a revenue boost of over 10 per cent"

It is a sophisticated form of regression analysis, the key point being the remarkable accuracy of its results. Outside the civil society sector, uses of SVMs include predicting television viewing figures from programme and transmission data, and detecting individual anomalous transactions in financial services data.

figure 1: Reactivating lapsed donors – some predictive drivers



Boost

Within the civil society sector, we have been using SVMs to predict the effectiveness of specific grant applications; to boost the effectiveness of fundraising campaigns, by finding individuals who are surprisingly likely to give; and to retain and reclaim supporters who would otherwise leave.

You don't need to be a massive civil society organisation to benefit from these techniques; we have used one or more of the above with various charities, including Action for Blind People, City Parochial Foundation, and RNIB, plus other civil society organisations such as the Marine Stewardship Council, trades unions, professional bodies and religious organisations.

We believe that charities increasingly will be deploying these types of analytical techniques. Even at the best of times, charities need to make the most of their donor and membership lists. In today's difficult climate, that need is even more urgent.

Also, as more charities use analytical techniques to understand their donors and members better, weaker charities will be left behind.

We predict that charities will increasingly 'get down and dirty' with their data, using modern techniques such as SVMs to classify and predict down to the level of an individual donor.

On direct fundraising, such as direct mail from warm lists, we describe the analysis as 'rubies in the dust', ie finding good donors among those people who have given in the past but who now appear to have lost interest in the charity concerned.

Case study

A national charity with annual voluntary income of close to £12m per annum worked with us to see if our own SVM, which we call PropheZy, could improve its direct-mail fundraising performance by finding rubies in the dust of its 400,000-strong donor database.

Our first step was to model historic data, to see whether we could 'predict' donors from data on campaigns that had already run.

Such a step is always needed when using a SVM, as the technique requires a data set, known as a 'training set', to generate the algorithms that will be used on the actual data to be analysed – the 'trial sets'.

It was quickly clear that the data was suitably predictive, but we also used this stage of the experiment to enhance PropheZy's performance. In this case by:

- Varying the amount of donor history assessed by PropheZy;
- Boosting the proportion of givers in the initial training set; and
- Banding donors at differing levels of probability of giving.

At this point it is worth pointing out a couple of unusual, useful characteristics of SVMs. Firstly, SVMs can cope with gaps in the data, as long as the 'gappy variable' has a reasonable amount of data in it.

Secondly, SVMs will ignore a variable completely if the data within it is too sparse or not predictive. Note that we are saying 'ignore' here, so the predictive quality doesn't deteriorate with poor data, the SVM simply predicts as well as it can from the data it has.

In this case, the age field was very sparsely populated, to such an extent that it had no predictive quality. The gender field had gaps in it, but could still help the predictive performance.

We also learned that increasing the amounts of donor history improved predictive performance, until you tried to include more than five previous asks for each donor.

> No charity can afford to ignore the opportunity to increase returns

The data from six or seven asks ago had no effect on the results at all. The other highly predictive field was the gift aid indicator.

Figure 1 illustrates some of the variables, and provides some clues to the predictive drivers.

The X-axis shows the demographic postcode (an indicator of wealth); the Y-axis shows number of years on the database; and the third dimension – the Z-axis – shows the PropheZy prediction: 'will give' in the foreground, 'won't give' in the background.

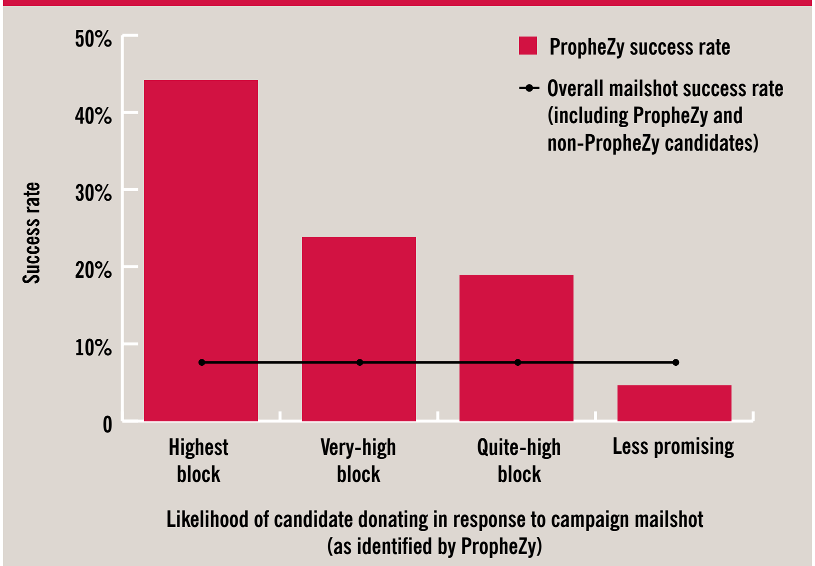
Blue dots show the actual result for an individual; if there is no grey line attached this indicates that PropheZy correctly forecasted the result. Red dots with grey stalks show the PropheZy prediction when it is at variance with the actual results.

People in wealthier demographic postcodes, who had been on the database for a number of years,

figure 2: Predicted likelihood to give

Likelihood block	Potential donors identified by PropheZy	Actual donors in response to campaign mailshot	PropheZy success rate (%)
Highest block	3,722	1,645	44.2
Very-high block	5,837	1,393	23.9
Quite-high block	6,520	1,239	19.0
Sub-total	16,079	4,277	26.6
Less promising	103,566	4,828	4.7
Mailshot total	**119,645**	**9,105**	**7.6**

figure 3: Success rate by likelihood block



tended to be correctly identifiable as likely to give, even if they hadn't given for some time (see the blue dots in the area ringed in red). However, people who had not been on the database for long were much harder to identify as potential donors (see the red dots with grey stalks in the area ringed in black).

Having tuned the model on a number of trials using historic data, we were ready to conduct a trial on a substantial live mailing.

Results

The charity identified the donors it would normally mail for that campaign, while Z/Yen ran the whole dataset against PropheZy to look for rubies in the dust.

PropheZy identified 16,000 donors as highly likely to give. Around 14,000 would have been included in the charity's regular mailshot anyway, but 2,000 of them would not have been chosen, as they seemed to be poor prospects using the charity's regular methods.

The table in figure 2 and the chart in figure 3 illustrate the SVM predictions and the results, divided into four blocks. The top three blocks in figure 2 represent varying levels of predicted high-likelihood-to-give, the final block represents the less-promising prospects.

The successful response rate among the 2,000 'rubies in the dust' of the three promising blocks was 29.6 per cent; around 600 additional donors.

This represented roughly a 15 per cent increase in the number of donors responding to the mailing concerned, and a revenue boost of just over 10 per cent for the campaign.

Valuable

We consider this result to be very significant; the predictive ability of the SVM over and above conventional techniques is extremely promising from this trial.

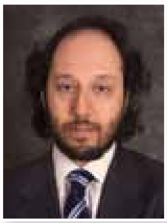
But, financially, the results on a warm direct mail campaign of this sort are good but could not be called spectacular.

However, applying the same technique on higher-value items, for example identifying donors who have a propensity to graduate to regular giving or legacy pledges, could be very valuable indeed.

No charity can afford to ignore such opportunities to increase the returns from its supporter base and to boost voluntary income.

In the next part of this feature, we'll explain the work we did with another client, where the SVM's predictive ability is helping to reclaim lapsing members and helping to retain likely-to-lapse members.

Membership income's recurring nature makes the financial results of that case study both significant and spectacular. ■




Ian Harris is a director, and Mary O'Callaghan is a senior consultant, with Z/Yen Group