JOURNAL VOL. 34 No. 02

USING MACHINE LEARNING TO ENHANCE PORTFOLIO VALUE

FEATURE

Nelson Henwood,
Partner, Finity Consulting.

"Supervised learning involves the prediction of an outcome given a specific set of inputs."

PORTFOLIO SEGMENTATION IS BECOMING A KEY COMPONENT IN THE MANAGEMENT OF INSURANCE PORTFOLIOS BUT ONE SIZE DOES NOT FIT ALL AND A TARGETED APPROACH TO THE IMPLEMENTATION OF PORTFOLIO MANAGEMENT INITIATIVES IS REQUIRED TO CREATE SUSTAINABLE VALUE. NELSON HENWOOD SUGGESTS THAT THE USE OF MACHINE LEARNING IS AN EFFICIENT AND POWERFUL APPROACH TO DEVELOPING PORTFOLIO SEGMENTATION INSIGHTS.

Introduction

A vital component of successfully managing an insurance portfolio is to adopt a segment-based portfolio management approach. Portfolio segmentation allows management to dissect a portfolio into blocks of business which exhibit similar characteristics. Portfolio segmentation enables the implementation of targeted strategies which can be used to enhance portfolio value.

Machine learning is an efficient and powerful approach to understanding key aspects of portfolio behaviour which are embedded in the portfolio data. Machine learning analysis combined with a commercial overlay based on the skill, experience and judgment of the portfolio manager is a very practical and efficient way of developing an effective portfolio segmentation.

This article introduces machine learning and some of the key types of analyses which can be used for portfolio segmentation. Examples of the results of machine learning analyses of two personal lines portfolios are provided, followed by additional discussion regarding the use of segmentation for portfolio management.

What is machine learning?

Machine learning refers to "computer algorithms that improve automatically from experience"[1]. In this context, the experience is the data used for analysis and the learner is the algorithm. There are two main types of machine learning:

- Supervised learning involves the prediction of an outcome given a specific set of inputs. For example, prediction of a claim frequency for a policy, given policy and customer characteristics.
- Unsupervised learning involves building an understanding of relationships between data components. An example of the use of this approach is for credit card fraud detection. Here, the technique is used to establish a "usual" pattern of transactional behaviour. In application, alerts are triggered when transactions occur which fall outside this pattern.

This article focuses on supervised learning which can be used successfully for portfolio segmentation.

Machine learning is closely related to data mining. Data mining may employ machine learning techniques but other approaches such as summary tables and statistical methods are frequently applied. Machine learning as a discipline is very much focused on the learning algorithms whereas data mining describes a more general process of data handling and analysis.

With machine learning, the generally accepted approach is to split the available data to be used for analysis into two separate sets – a “training” dataset which is used to explore and find segments and a “validation” dataset which is used to verify that the patterns found in the training data repeat in an independent dataset.

Two of the better known machine learning algorithms are Decision Trees and Neural Networks. Other algorithms include Support Vector Machines, Nearest Neighbours and Genetic Algorithms. These algorithms are available in many commercial packages such as SAS and S-Plus. Many are also available in the R programming which is freely available.

Finity use a software suite called Talon for machine learning applications. Talon incorporates some of the most recent developments in the field of machine learning and is tailored for use with insurance data.

We have found that using Talon uncovers signals in insurance data which cannot be detected easily using off-the-shelf package solutions.

Talon incorporates four key families of segmentation analysis focused on the commercial problems of managing an insurance portfolio. These analyses aim to uncover:

1. Segments experiencing differentials with respect to a key metric such as loss ratio, COR, retention or quote conversion.
2. Portfolio segments driving change over time in a key metric such as claim frequency, claim size or loss ratio.
3. Changes occurring over time in portfolio composition. For example, identification of segments growing at a faster or slower rate than the portfolio overall.
4. Talon also enables these types of analyses to be conducted across two dimensions at the same time. For example, it can be used to identify segments with a higher than average loss ratio experiencing higher than average growth.

The portfolio segments identified in these analyses are described in plain English and are also mutually exclusive and exhaustive (i.e. each policy belongs to one, and only one, identified portfolio segment). This means that the results can be readily communicated to the business and are in a form for easy translation for marketing programs, underwriting actions etc.

Examples

In this section, the results from a number of segmentation projects we have undertaken for clients are described. Please note that the results have been modified to protect client confidentiality.

Example 1. Loss Ratio Segmentation – Comprehensive Motor Portfolio

The first example is a portfolio segmentation based on loss ratio for the comprehensive motor portfolio for a mid-sized insurer.

The candidate variables used for the segmentation included the usual policy and customer information as well as a range of external data, including detailed vehicle information sourced from Glasses’ Guide.

The loss ratio calculation was based on the current rates for the portfolio. This enables use of historical loss information in combination with current premiums to draw conclusions about segment profitability at current rate levels. This required recalculation of the premium over historical exposure blocks at current premium rates.

Figure 1 – Segment Loss Ratios – Training vs Validation

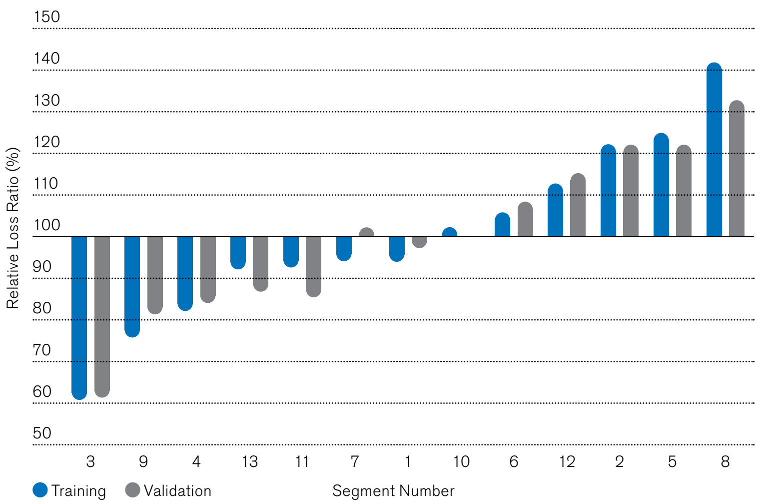
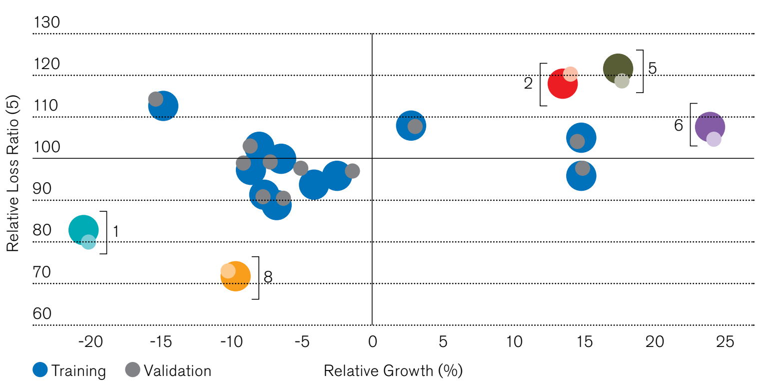


Figure 2 – Relative Loss Ratio vs Relative Growth – Training vs Validation



Based on the parameters selected for the analysis, 13 portfolio segments were identified. Each segment contained at least six per cent of the portfolio exposure.

Figure 1 left illustrates the results of the analysis. The bars illustrate the loss ratio for each segment (expressed relative to the portfolio average). The segments have been ordered from best (i.e. lowest loss ratio) to worst (i.e. highest loss ratio). The results over both the training (blue bar) and validation data (grey bar) are shown.

The segmentation results validate well in that there is a high correlation between the loss ratios for the training versus validation data across the segments. The measured correlation is over 98 per cent.

The segmentation also achieves a good degree of separation with loss ratios ranging from around sixty per cent of the portfolio average for the best segment to over 135 per cent of the portfolio average for the worst segment.

In other words, there are readily identifiable and sizeable parts of the book where current premiums are thirty to forty per cent out of line with the average portfolio loss ratio experience.

Table 1 below gives further information about a selection of the segments at the extremes of the range of profitability. The table details the proportion of exposure, loss ratio relativity and variables incorporated into the rules defining the segments. Each segment is defined as a combination of three to four variables and thus segment membership for a policy is easily resolved, making the segments readily actionable by the business. The loss ratio relativities given in the table have been calculated across the total data (ie including training and validation).

The segmentation is achieved through rules involving current rating variables and customer-based attributes which do not form part of the current rating algorithm. The implications of this are as follows:

- It is not possible (and may not be desirable) to fully eliminate cross subsidies by adjustments to pricing using the current rating structure.
- The analysis undertaken can signal the priorities for changes to the rating algorithm to ensure that unprofitable business can be addressed, and that opportunities for writing profitable business are not constrained.
- The knowledge of profitable/unprofitable segments can be used to direct marketing activities or to implement other customer management strategies aimed at changing the mix of business in the portfolio.

Example 2. Combined Profit/Growth Segmentation – Comprehensive Motor Portfolio

The combined profit/growth segmentation is a two-way segmentation of the portfolio focusing on identification of:

- loss making segments which are growing as a proportion of exposure and
- profitable segments which are diminishing as a proportion of exposure in the portfolio.

Each of these scenarios represents an adverse trend which will ultimately manifest in poorer performance for the portfolio. Remedial steps are required in both cases.

The algorithm works by looking for segments before and after a user specified point in time which show a consistent loss ratio before and after this point combined with higher or lower growth than average. The results are shown in Figure 2 left.

Table 1 – Segment Details – Example 1

Segment			% of Exposure	Loss Ratio Relativity	Vehicle Age	Sum Insured	Cust Attribute #1	Cust Attribute #2	Cust Attribute #3	Cust Attribute #4
Best	1st	#3	8%	60%	✔		✔	✔		
	2nd	#9	7%	77%	✔	✔	✔	✔		
<….9 segments in between….>										
Worst	2nd	#5	6%	125%			✔	✔	✔	✔
	1st	#8	7%	138%		✔	✔	✔		

Table 2 – Segment Details – Example 2

Segment		% of Exposure	Loss Ratio Relativity	Relative Growth	Power to Weight Ratio	Rated Driver Age	Vehicle Value	Vehicle Factor #1	Rating Factor #1	Protected NCB	Cust vs Driver R'ship	Cust Attribute #1	Cust Attribute #2
Profitable and Declining	#8	6%	72%	-9%	✔			✔			✔		
	#1	5%	79%	-18%	✔			✔					
<….11 segments in between….>													
Loss Making and Growing	#6	8%	107%	22%			✔	✔		✔			✔
	#2	5%	117%	11%		✔	✔	✔		✔			✔
	#5	9%	118%	16%				✔	✔			✔	

In this analysis, 16 portfolio segments were identified, each containing at least five per cent of the portfolio exposure. The graph shows the relative loss ratio and relative growth for each of the segments. The large circles are the training results and the smaller circles are the validation results. The segments of most interest are highlighted in different colours.

Table 2 below provides the segment details for the segments of most interest from the analysis.

A number of segments strongly meet the criteria (faster growth and unprofitable, slower growth or deteriorating in size and profitable). The segments are easily described in English as combinations of two to five variables. In this case the segmentation is dominated by the interaction of vehicle related factors with customer related factors.

Clearly the segments shown in Table 2 warrant further investigation to determine, for example, how the insurer's pricing compares to the market. It may be that the pricing is severely out of line and needs to be addressed as a matter of urgency – by definition, deficiencies in the pricing are causing the range of loss ratios observed across the segments.

Where the pricing response may need to be constrained due to market-related factors or pricing engine restrictions, other remedial measures may need to be used in conjunction. For example, underwriting guidelines may need to be tightened around the poorer performing segments. For the better performing segments, a combination of looser underwriting and customer care initiatives (eg quote and renewal follow-up) may be appropriate. Some guidance in relation to discretionary discounting or loading for segments meeting the criteria may also assist to address the imbalance.

Table 3 – Using segmentation to manage an insurance portfolio

Pricing	The primary focus for pricing would be identification of profitable and unprofitable business in the portfolio for targeting price changes. Complementary analyses would include segmentation based on price elasticity, sales metrics such as renewal rate and quote conversion and joint analyses of profitability and growth.
Underwriting	Profitability segmentation can be used to prioritise business for underwriting scrutiny at new business or renewal time. This would enable prioritisation of underwriting resource to the more troublesome areas whilst potentially profitable business could be processed straight through.
Monitoring	Segmentation focussed on detecting change is particularly relevant for regular portfolio monitoring. Analysis to detect changes in the mix of business can be performed at regular monthly or quarterly intervals. The drivers of change in key metrics such as claim frequency, claim size or loss ratio can be examined on a quarterly or semi-annual basis.
Target marketing	Joint segmentation analysis can be used for example to uncover segments which are profitable and where the insurer's premiums are competitive. Such segments would represent prime targets for focussing marketing efforts.
Claims management	Segmentation based on average claim size can be used as the basis for automation of initial case estimates or early identification of potentially large claims.
Customer management	Segmentation based on renewal propensity could be used to identify business at high probability of lapse for potential pre-renewal customer contact.

Application of segmentation for portfolio management

Portfolio segmentation can be used to underpin initiatives focused on a number of the key portfolio management disciplines as described in Table 3 left.

The potential applications of segmentation to create value in the portfolio are numerous. The key to success is to use the findings across a number of related initiatives. This creates the best prospects of leveraging the results to achieve meaningful change in the portfolio.

Conclusion

Portfolio segmentation is becoming a key component in the management of insurance portfolios. One size does not fit all and a targeted approach to the implementation of portfolio management initiatives is required to create sustainable value.

The use of machine learning is an efficient and powerful approach to developing portfolio segmentation insights. One of the most valuable assets that the insurer has is its portfolio data. The application of machine learning algorithms focused on the right target(s) can help to ensure that the insurer is leveraging this asset to its full potential.

There are a variety of segmentation targets about which insight can prove valuable. The particular approach taken will depend on the prevailing portfolio issues and the data available.

Typically, there are numerous initiatives which could be used to capitalise on the results. Where there are constraints to implementing the segmentation findings using a particular avenue, there are usually alternative approaches which can bring about the desired change in the portfolio. The greatest degree of improvement in portfolio performance will be achieved when a range of related initiatives focussed on achieving a particular outcome are implemented in combination.][

Notes

1 Machine Learning, Tom Mitchell, McGraw Hill, 1997