Paul Beswick

Machine Learning for Retail

USING MACHINE-LEARNING ALGORITHMS TO PREDICT SALES ON PROMOTIONS

Sales promotions are on the rise in most countries today, and retailers are struggling to make better predictions to control spending and increase returns.

Historically, when trying to predict sales based on different factors, managers have applied business logic based on experience – the quality of a brand, the shelf placement, the promotion, and so on. They typically use a series of linear regressions, plotting known sales volume against the variables, to get a decent forecast for the next promotion. This approach essentially relies on the human brain to select and analyze data.

But machine learning is much more powerful. A machine can look at history to determine which factors are most important, and to find the best way to predict what will occur based on a much larger set of variables.

In the old forecasting world led by the brain, you used one model for just about every category or type of business. In a changing environment, using customized models for each category or type of business increases the accuracy of predictions, because even if two categories are similar, they have underlying intrinsic differences that require customized machine-learning methods to capture.

In the new forecasting world of machine learning, you can build a customized model for every category or sub-category or type of business. Instead of a few decision trees, machine-learning algorithms randomly create thousands of decision trees based on sub-groups of explanatory variables; typically, if there are 20 explanatory variables, the random trees will only use four or five variables at a time (which could easily be handled by any computer). The algorithm then combines the thousands of trees to make a single predictive model that incorporates all the variables. Once "trained," the algorithm is able to automatically predict sales at the product level during any promotion. And it continues to learn as it takes in more data and results.

IT teams are now embedding machine learning algorithms in legacy systems, to run automatically week after week, as the promotions are churned out. Predictions to date have been done at the country level, but there is interest in driving down to the store level. Beyond that, the next wave of innovation will center around customization – offering personal promotions to specific customers for specific products at a specific time. Industrializing the understanding of promotions with artificial intelligence is the first step in making this a reality, as each decision to make a specific promotion will need to be undertaken a million times a day.

CASE STUDY

MULTINATIONAL RETAILER

At OW Labs, we applied a machine-learning model to help a large multinational retailer determine how given products would sell based on its print promotions. Since the retailer does 50,000 to 60,000 promotions a year, even a small increase in predictability would drive a huge increase in sales volume or save tens of thousands in wasted discounts.

Senior leadership wanted to get better at advanced analytical techniques to determine, say, what effect a 10-day promotion on a case of Coke (or shampoo) would have on sales over the next six months. They wanted to understand how strong or weak a product was, how much you had to give away to drive results, and how categories differed.

Merchandisers know that a front-page position will likely increase sales, but the impact is less clear when also considering seasonality. Factoring the type of products into the forecast adds more complexity, as soda and shampoo, say, behave quite differently, making "manual" predictions hit or miss. Add a few more variables, and the problem of accounting for 20 or more factors interacting at the same time quickly becomes unfathomable for the human mind.

In our model, we used variables such as the depth of the discount (the deeper the discount the higher the sales); the duration of the promotion (the longer, the higher the sales); the average sales without promotion (the more popular the product, the higher the sales); the display in the circular (the bigger the photo, the higher the sales); the display and shelf placement in stores; the type of promotion (Buy 2 Get 1 Free, immediate discount, or loyalty points); the type of product (soda, water, shampoo); the promotion elasticity (how much customers react to promotions on a given product) the competitive pressure (other promotions from the competition); and seasonality.

The retailer's forecasting team of six or seven, using a simple linear model to make predictions, and basically inputting data by hand, was predicting results with a 30 percent to 35 percent error rate. Right off the bat, our machine-learning model achieved far higher accuracy, cutting the error rate to 24 percent, a major improvement. In addition, the ML model was automated – and it could be expected to improve over time.

Better predictability had two immediate impacts: One, it prevented the merchandising team from generous promotions that would never deliver an ROI, thus preventing costly mistakes; two, it allowed for more informed discussions with stores on ordering inventory, preventing over- or under-ordering.