

A Practical Note on Predictive Analytics Usage in Marketing Applications

**Arindam Banerjee
Tanushri Banerjee**

**W.P. No. 2016-05-01
May 2016**

The main objective of the working paper series of the IIMA is to help faculty members, research staff and doctoral students to speedily share their research findings with professional colleagues and test their research findings at the pre-publication stage. IIMA is committed to maintain academic freedom. The opinion(s), view(s) and conclusion(s) expressed in the working paper are those of the authors and not that of IIMA.



A Practical Note on Predictive Analytics Usage in Marketing Applications

Arindam Banerjee
Professor of Marketing
Indian Institute of Management Ahmedabad
Email: arindam@iima.ac.in

Tanushri Banerjee
Associate Professor, Information Sciences
Pandit Deendayal Petroleum University Gandhinagar
Email: tanushri.banerjee@spm.pdpu.ac.in

Abstract

Most Predictive Analytics discussions focus on methods that can be used for better quality prediction in a particular context. Realizing that the possibility of perfect prediction is a near impossibility, practitioners looking to support their futuristic initiatives wonder, what is a suitable model for their use.

In other words, if all prediction models are imperfect (have leakage) how much of this imperfection can be tolerated and yet better decisions can be taken with model output. This paper is an attempt to provide a simplified approach to this practical problem of evaluating model performance taking account of the decision context.

Two scenarios are discussed; a) a classification problem often used for profiling customers into segments and, b) a volume forecasting problem. In both cases, the leakage is defined (misclassification or uncertainty band) and their impact (adverse) on the subsequent decision is identified. Contextual dimensions that have an impact on the quality of the decision and the scope to alleviate the problem are also discussed.

A Practical note on Predictive Analytics Usage in Marketing Applications

Why Predictive Analytics?

Prediction being a much discussed and researched area in business decision making for long mainly due to its appeal to make the latter enriching and more importantly, precise. There are many treatises that have been written on this subject matter, however decision makers will agree that it is also necessary to distinguish between methods/tools to achieve good prediction and what should be the realistically a “good” prediction. My suspicion is that a lot of work has been written on the former, with the assumption that the latter is very contextual and may not have sound principles that can be generalized. While that may be true to an extent, it may be still worthwhile to document the generic applications where predictive analytical methods can assist in decision making and the approach that professionals can adopt to evaluate their true value.

This note will be helpful for marketing / business decision makers who would like to adopt predictive models for taking business decisions, but shy away from the doing so because of their perceived lack of technical knowledge and which they think is required. It will also serve as a useful guide to all the modelling technologists, who would want to tweak their model building approach to suit the appetite of their business savvy clients (decision makers). At a broader level, it hopes to build a common platform for the technical and the business factions of the analytics domain to interact productively.

The utopian objective of Prediction and the disappointment

The obvious expectation from prediction models is one of perfect prediction. Every decision maker’s dream would be to come as close to a perfect description of the unknown as possible, if not the perfect one. Unfortunately, many times, as close as possible to perfection is not close to being good as being perfect. Hence, it requires some reflection as to what a predictive model does to support our business systems in terms of taking the right decision.

Most analysts would attribute a less than perfect prediction condition as due to leakages in prediction models. The implication of such leakage to our decision making process needs to be studied. The starting point would be define” leakage” for the readers from them to appreciate its impact (negative mostly) to business performance. The next section does just that.

What is a leakage in a model?

Prediction models are meant for either a) “perfect” classification of respondents to pre-determined categories or, b) in the case predicting continuous outcomes to predict the exact amount. In both instances there can be leakages.

In the case of classification problems, a leakage is defined as a wrong classification by the prediction model. The quality of the prediction output depends on the percentage of wrong classification (inversely dependent). Better models have a lower proportion of wrong

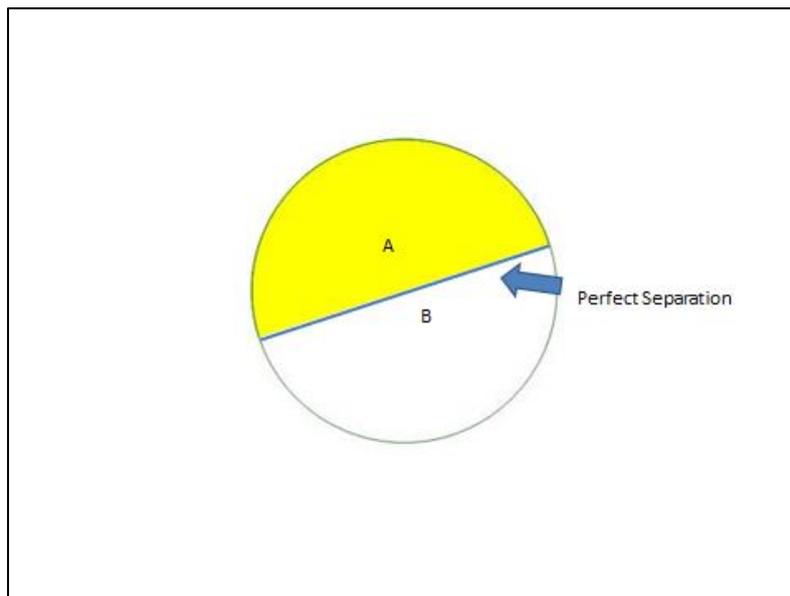
classification. The challenge for most decision makers is to determine what should be the minimum percentage of wrong classification that they can tolerate. This problem has a very contextual answer and is highly dependent on the impact that the wrong classification can have on the business performance.

Similarly, in continuous outcome determination (predicting volume of sales for instance), most (or practically all) prediction come with a band of uncertainty, also known as the forecasting error band. The challenge for most planners who have to forecast business size in the future, is that a band of uncertainty around the estimated value computed by a forecasting model is more critical than the estimate. Wider is the band, less is the importance of the estimate since the actual may be anywhere within this wide band of uncertainty (or it is expected to be). Hence the value of the information provided by the model output is less. The theory of probability defines expected (or estimated) value as the one that is obtained on an average if the prediction is done many times over. Decision makers on the other hand cannot use an estimate as a good measure of forecast (especially when the band of error is wide) since the actual outcome is usually a single observation rather than an average over many occasions. Unfortunately they have to get a handle on every single outcome in the future and have to forecast it right every time. Hence, a good understanding of the leakage is very essential in ascertaining its role in business decisions. Let us take some examples of these leakages to reiterate this important point.

Leakage in Classification Models:

Let us take the case of a population with two distinct segments “A” and “B”. These segments could be based on similarity in individuals’ characteristics (demographics, psychographics) or attitude or behaviour. A predictive model application in this context would be to use separately a set of exogenous variables associated to these individuals (and known a priori) and build a model that would use these variables (their values) to predict the segment in which each individual belongs (“A” or “B”). Ideally, the expectation from a model is that each individual is perfectly classified into their true segment (see figure 1).

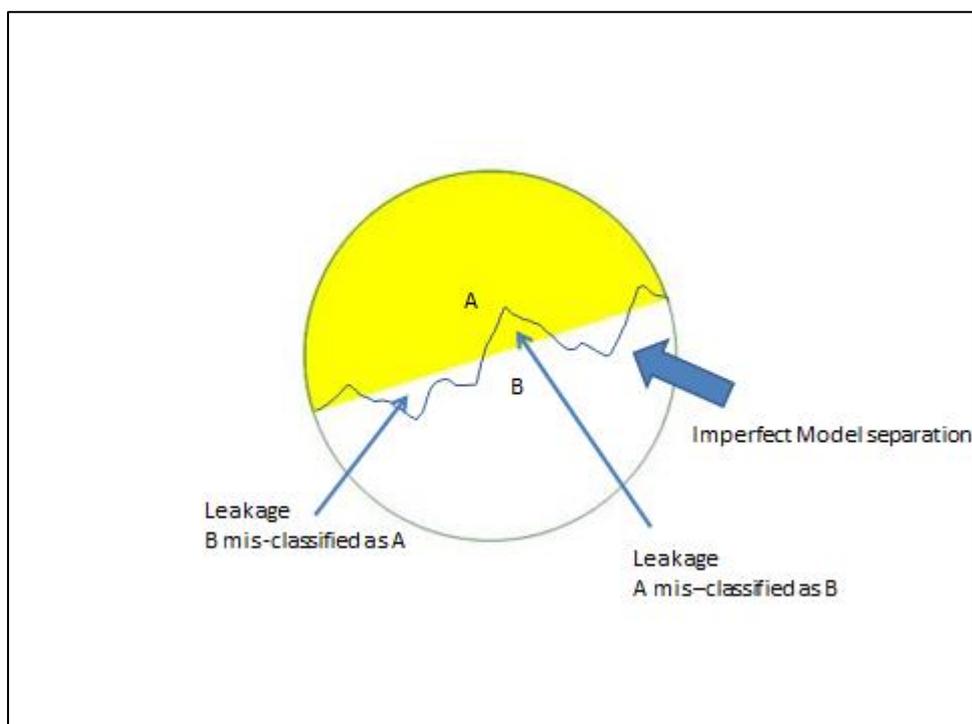
Figure 1



The separation is usually identified by a mathematical model using antecedent variables (X).

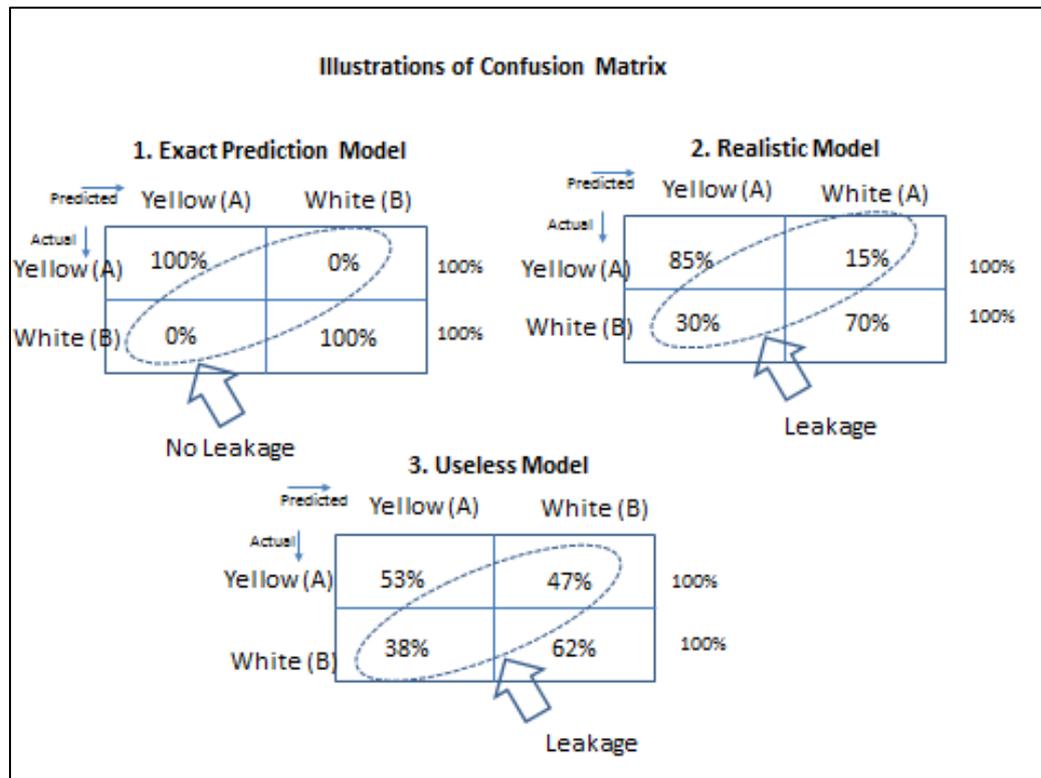
In reality this is not achievable except in rare instances. In normal circumstances there is bound to be some leakages – “A” classified as “B” and vice versa (see figure 2). The technical superiority of the

Figure 2



model is judged by how small are these leakages (also reported as Confusion matrix, Gini index, ROC curve, K-S statistic etc. in technical output). The simplest statistical indicator of the quality of the model, the confusion matrix is given below (figure 3).

Figure 3



The first confusion matrix (Exact Prediction Model) shows no leakage. The total number of “Yellow” (A) are correctly predicted since there are none that are predicted as “White” (B) and vice-versa. The second model output is more realistic where there is some leakage (or mis-classification of both “yellow” and “white” segments). The third model depicts close to equal proportion of correct and wrong classification. This is a case where the leakage is significant enough to render the model output useless.

The business measure of relevance of these models depends upon how much of these leakages (and which one) can we tolerate – so that the model output is still useful for decision making.

Taking a specific example, if the segment of interest to us is “A” for a certain policy implementation but the predicted segment “A” has on an average 30% of “B” included (model output 2 in Figure 3), the obvious question that comes to mind is the implication of the specific policy that is being targeted on “A” when a fairly proportion of “B” is also being inadvertently targeted with the same policy. The adverse impact of this mishit must be account for. Similarly, some members of segment “A” (15%) have been erroneously classified as “B”. Is there an adverse impact of missing out on these “A” since they are

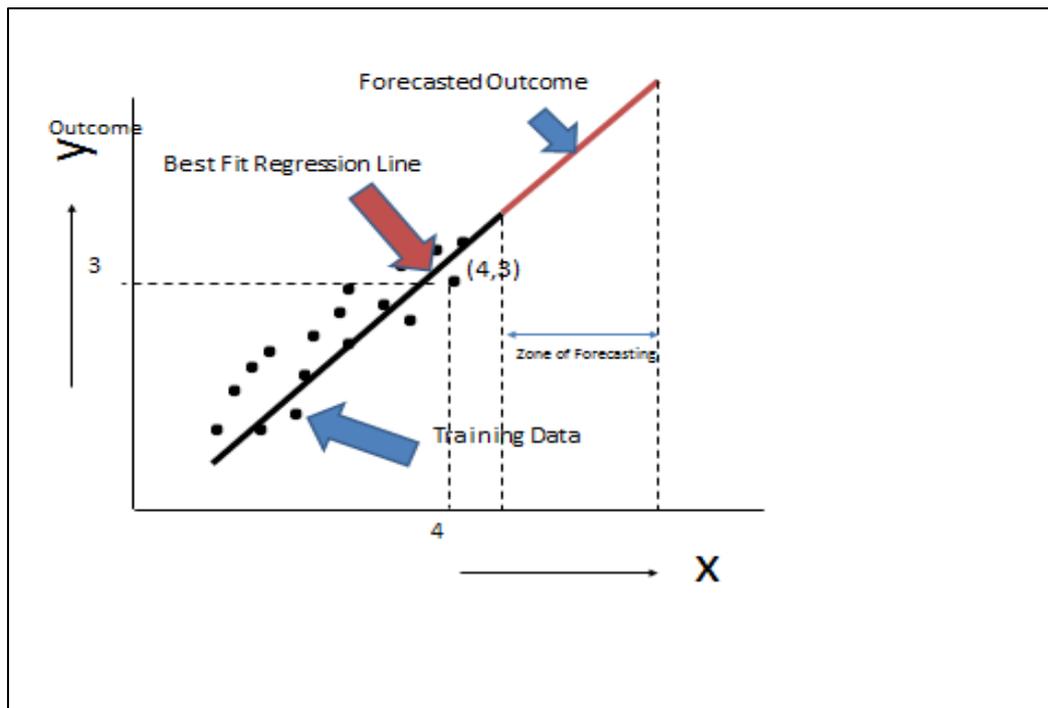
classified as “B”? These impact need to be ascertained before a view is taken about the worth of the prediction model.

In a marketing context, often times, a segment is identified for offering additional incentives. However, with leakages of the type described above, offering incentives to some of the wrong target segment can lead to additional costs (of subsidizing wrong target). Such costs need to be factored into the business impact model to ascertain whether the usage of the model output is indeed adding value to the firm’s initiative. Specifically, it needs to be ascertained if the incremental benefit to the firm from the incentives to the target population is large enough to tide over the unnecessary cost of subsidizing (providing incentives to whom it is not meant) the wrong segment constituents due to the imperfection in the prediction model. Another detailed example of this leakage cost in a non-marketing context is described in Appendix 1.

Leakage in volume forecasting models:

These are usually regression (OLS) based models that forecast continuous outcomes such as sales, revenue, cost and other business parameters that vary on a “continuous” scale (interval or ratio scale). A set of exogenous variables are used to predict outcome through a methodology of “best fit” of the mathematical function (of exogenous variables) to represent the outcome variable in a sample of data points, also called the training sample. The same function when used outside the precincts of the training sample can provide estimates of outcome for various combinations of values of exogenous variables in the future (see Figure 4). Along with the estimated outcome, there is also an estimate of the uncertainty band around the estimate which is the zone in which realistically the actual outcome may fall (see Figure 5).

Figure 4

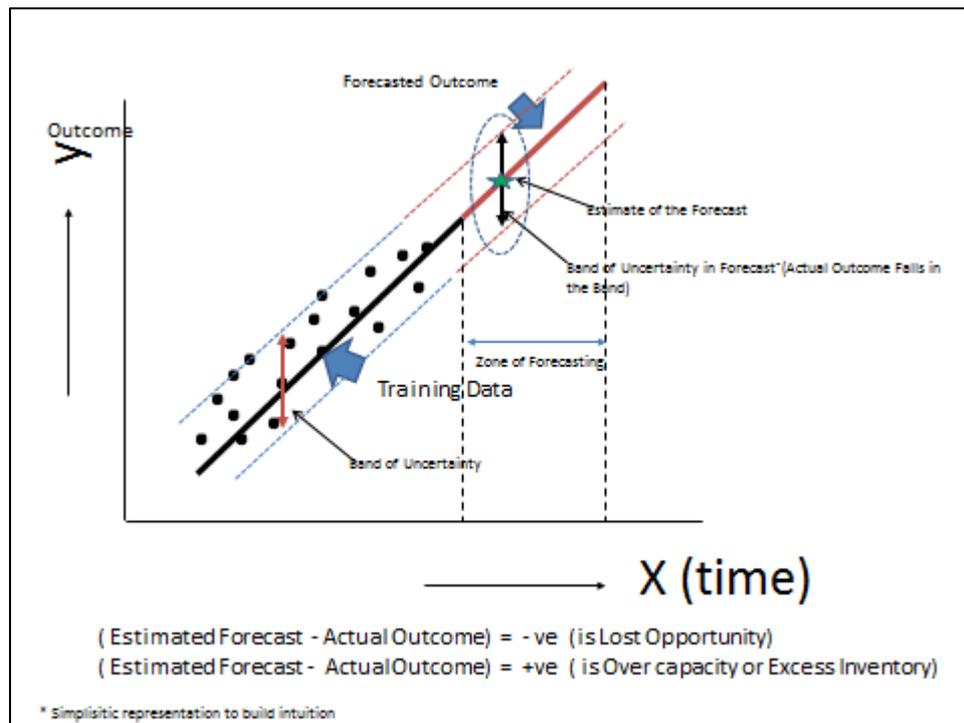


The regression line is extended beyond the range of data to forecast (X may be a time dimension)

In a perfect model the zone of uncertainty reduces to zero and the estimated outcome is indeed the actual outcome. However, with realistic models the zone of uncertainty is nowhere close to zero (figure 5). Better models have narrower bands of uncertainty compared to no-so-good-models. But the question that begs an answer is what is a good model (or a suitable model).

Let us pay attention to the band of uncertainty around the estimate (Figure 5). If the band is wide, both the lower and the upper limit are a concern since this is the range over which the actual outcome may fall and wider is the range lesser is the confidence of the business planner to provide a concrete “view” of the future outcome. For instance, if the plan is made conservative based on this model output (on the lower end of the range from the estimate) there is a higher probability of falling short of actual and vice versa. What then is a “manageable” band of uncertainty (to be expected from a model) if a precise forecast is not available?

Figure 5



Model Output Usability: General Insights

The generic answer to the above question is dependent upon a few dimensions:

- How much of a risk is the decision maker going to bear? Which risk is bearable and which is not?
- How flexible are operations process to address changes in final outcome from the planned?

The risk appetite of a decision maker has a significant influence on the way the model results are interpreted. If the business loss due to “lost opportunity” is lower than the cost of excess inventory caused due to “optimistic planning”, the planned output will be usually calibrated to a level lower than estimate to mitigate losses due to over production/capacity. Also, the uncertainty band on the positive side can be better tolerated than the uncertainty in the negative spectrum. Hence, the mandate for the analytics person would be to minimize the error band on the lower side to the extent possible. The focus of planning is reversed if the cost of “lost opportunity” is much higher than the cost of excess inventory. In this case the error band on the positive range from the estimate requires better calibration (minimized as much as possible).

If a corrective reaction to actual outcome is possible (the ability to calibrate action in real time when the future unravels and the actual outcome is known), planners would be more than comfortable with error bands in forecast that are in synch with this margin of operational flexibility. Therefore, so long as the “quality” of the model output (error range of the

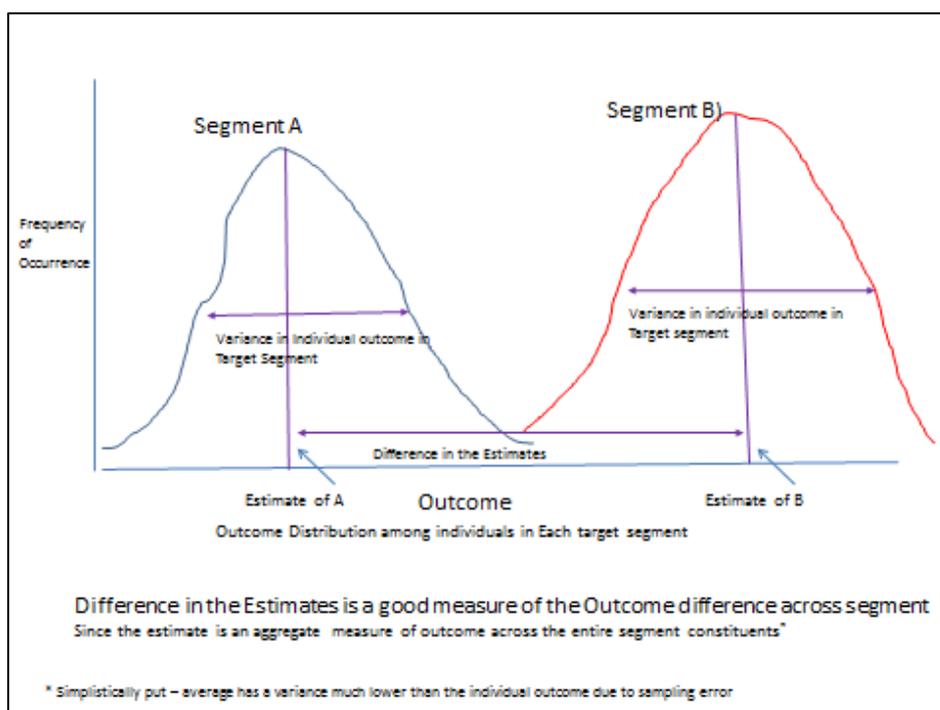
forecast) matches with the range of calibration of the operations process, the estimate obtained is good for decision making. For example, if the error band around the forecast is about $\pm 20\%$ and the flexibility in changing plan to meet actual output is about the same, then the estimated forecast is a good input to the planning process. The intuition here is that irrespective of what the actual outcome is (within the range) actual delivery operations can seamlessly recalibrate without any significant incremental cost. Hence, if the forecast error is within this limit of flexibility, using the estimated forecast is a good planning input.

However, if the flexibility to recalibrate production / delivery is just about $\pm 10\%$, there is a possibility that the adjustment needed from plan may be much higher than what can be realistically achieved, both in case of a positive and a negative deviation. Here, the planner may want to decide which actual “unmet” deviation may cost her (him) more and accordingly readjust the plan in a way to reduce the cost.

Identifying Target Segments Based On Forecasted Outcomes: The Boon Of Outcome In Aggregate

This is one regression based application where the error of forecasting may do much less havoc to a planning process. If market opportunity segments are identified based predicting outcomes using individual customer profiles, an estimate of the outcome (basis of identification) variable is a good input into a market identification process, provided it is significantly different across segments. The reason is that while the outcome is predicted for individual customers, a profile describing a group of similar customers may have different outcomes, but the average of the segment is close to the estimated outcome (as computed by the model using the profile variables). Hence, when the use of the model output is in an aggregated form, the estimate is a reasonably good measure of the forecast (never mind the size of the error in predicting individual outcomes), as long as the estimate is significantly different from other profile outcomes. This is one application where the concept of running statistical experiments “many times over” seems valid and therefore the expected value (mean / estimate) is a good input for a business planning process (see figure 6).

Figure 6



To cite an example, if the segment B on an average has a spending power significantly higher than segment A (outcome is spending power in this case), a decision maker can with relative ease conclude that segment B may be the target segment to focus on for a particular sales initiative to attract customers who can spend more. Of course, since the measure of the difference in spending power between the segments is subject to a sampling error, the actual difference in spending among the segment will be subject to some error band (hopefully much smaller due to the aggregation effect).

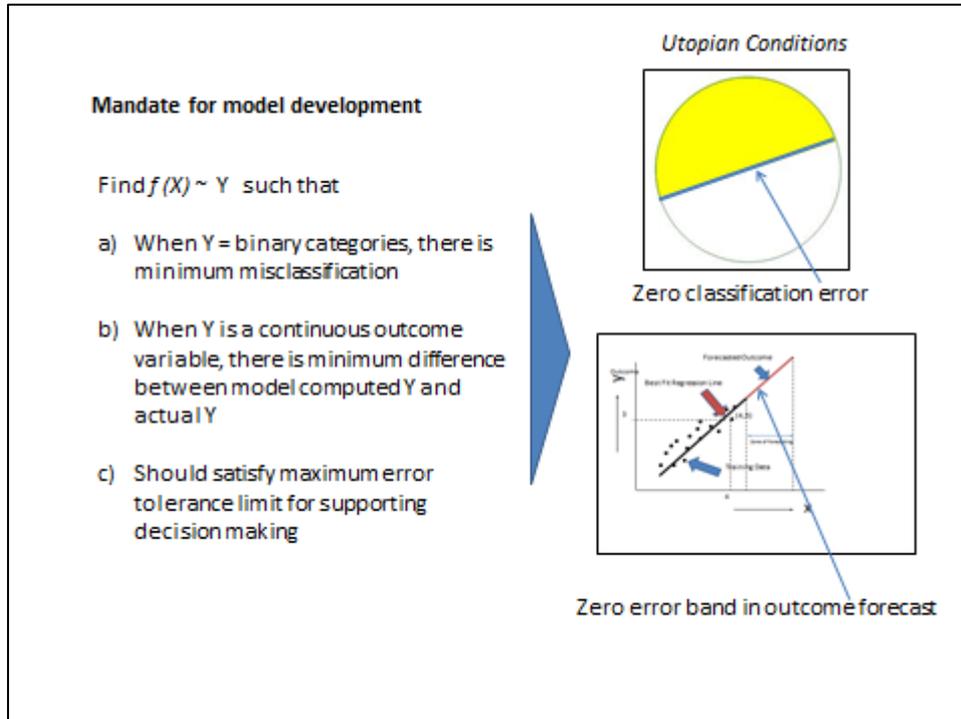
So Are We Ready To Build Models With An Objective To Help Business Decision Making?

Once the desired output of the model is well defined – in terms of, a) maximum tolerable error band in case of output estimation or, b) maximum tolerable mis-classified proportion in our target category in case of a classification problem, the mandate for model development becomes relatively simple. Analysts may now seek any methodology, simple or sophisticated, so long as they are able to achieve the objective criterion of a suitable output that adds value to decision making (see figure 7).

A word of caution on the methods used to develop models. More sophisticated methods, for both types of outcome variables – continuous or classification (neural networks, decision trees, regression splines, support vector machines) provide an opportunity to fit the training data better and hence the performance of the models are better (same or better than simple models). However, what is unclear is whether a better fit in the training samples amounts to reliable outcomes in other sample data (which is the true test of the usability of the model).

Simple models (Ordinary regression, logistic regression, contingency tables) on the other hand may not perform as well as sophisticated models in fitting responses in the training sample, but may at times perform better than sophisticated models in other data sets.

Figure 7



There are automated processes for selecting models (simple or sophisticated) that best fit both training and other data (validation sample). In the process of rote search, analyst may be able to provide the best solution in a given context and compare it with the quality of outcome desired by decision makers.

With no intention to complicate matters beyond necessary, it is important to point out that actual forecasting processes in organizations require factoring in deviations beyond what the historical data used for prediction is able to identify. If the future scenario is significantly different from the past, the prediction model is not capable to factoring in this pertinent deviation since no information about such scenario is available from the past data. In circumstances such as this, no data driven model output alone will provide the right basis for choosing planning estimate. Information from outside the realm of data and modelling is required to adjust the estimate appropriately. Suffice to say, there are many other manuscripts that have addressed this matter in detail.

What have we achieved so far?

We have covered the generic requirement of predictive models for many marketing / business decision making. We have discussed how the quality of the model output can be ascertained in the context of a specific decision making situation. More importantly, we have separated

the exercise of setting objectives of developing predictive models from the actual process of building them. This element distinguishes the current discussion from many standard description of Predictive Analytics found in many other treatise.

In the process, we have identified a distinction between a) the technical function of building models using data science methods from b) the evaluation of the model performance on business support parameters. Decision makers and sponsors of analytics process may now have a practical approach to set realistic benchmarks for model performance without understanding the mechanics of building models.

Appendix : A case on how leakage can affect the business performance assessment:

MICROLOAN Bank

Finding the “right” credit defaulters to pursue Arbitration Proceedings

Customer Service managers at MICROLOAN bank are at their wit’s end trying to find the best possible way of curtailing bad debt losses which amount to over \$1.8 billion annually, about 10% of the total balance outstanding on their VISA card. Proactive servicing of delinquent customers (customers who stop paying their minimum due amounts on their credit card bills) by telephone calls, letters and others means of motivating them to stay updated on their credit card payments had marginal effect on reducing the amount of bad debts.

The legal cell of MICROLOAN bank proposed initiating arbitration proceedings against delinquent customers in order to expedite the payback process. While telephone calls and letters were milder forms of reprimanding the customer for not fulfilling their part of the obligation when they availed of credit, threatening to take legal action was considered to be more aggressive and presumed to have a big impact on debt collections.

The downside of this measure was the price tag attached to initiating an arbitration procedure. For every customer on whom arbitration procedure was initiated, MICROLOAN bank would spend \$125 as arbitration fee to the National Legal Council that administered the arbitration process. This fee had to be paid regardless of the outcome of the arbitration procedure.

Collection managers at MICROLOAN bank were reluctant to invest an additional amount of money on all defaulters in the hope that arbitration would increase the collection rate. They wanted to know how best to go about setting up the process and to send only the best prospects (customers who were most likely to pay under the threat of a legal suit) to arbitration. What they identified as the need was to find an efficient selection procedure to categorize defaulters as viable/non-viable for arbitration. This would cut down on the total cost associated with arbitration proceedings. The incremental cost would be spent only on those defaulters who were likely to pay under threat and hence the arbitration fees could be recovered from them as well.

A clause in the arbitration laws in the United States required that any arbitration proceedings should be preceded by a one-month notice period, communicated by a letter from the bank, to

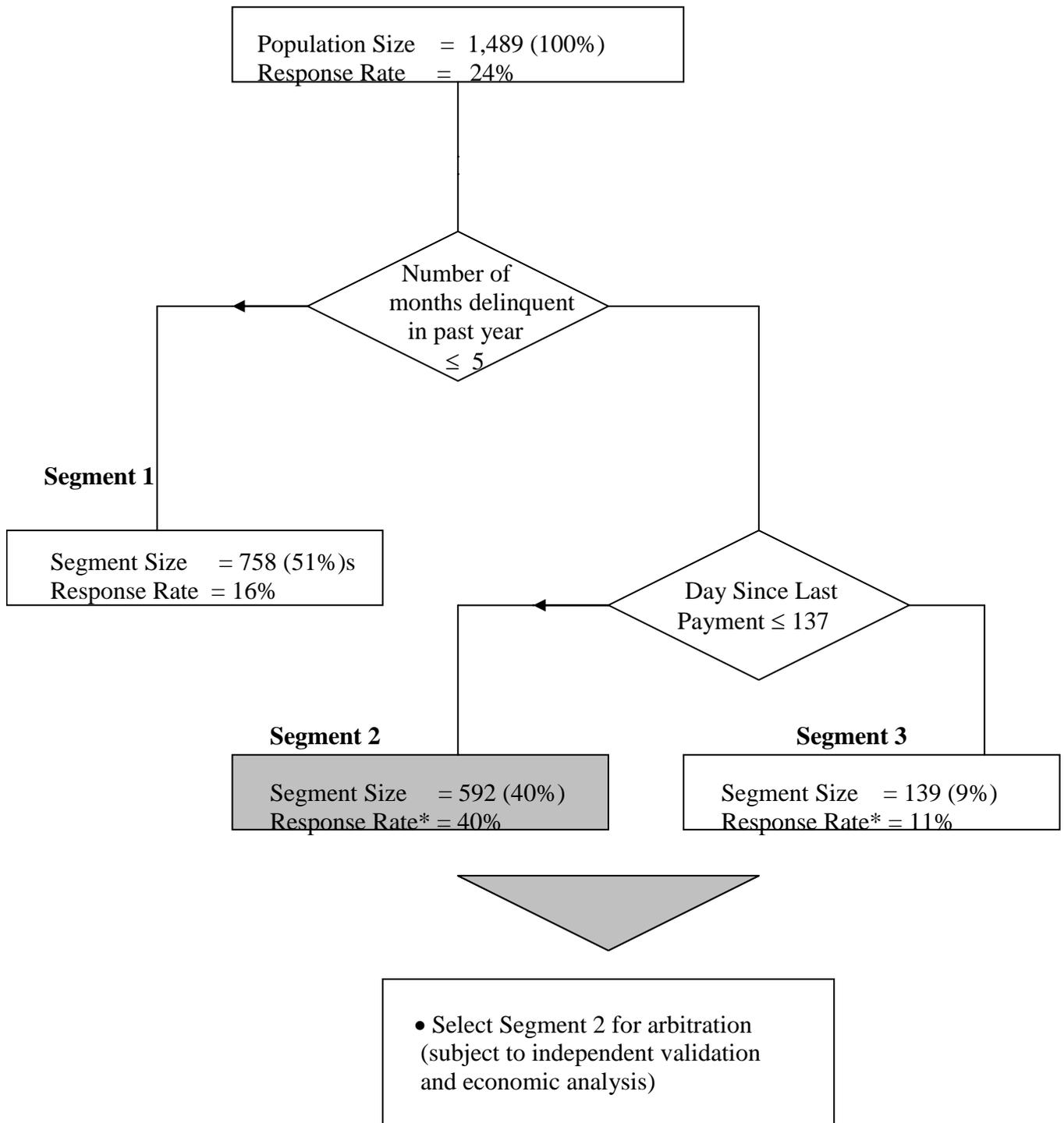
the defaulter about the former's inclination to start arbitration if no response or payment was received from the customer during the notice period. Collection managers perceived this would be sufficient threat to motivate defaulters to act, at least the ones who would pay under threat of arbitration. The only catch to the arbitration clause was that if the letter of intent was sent, arbitration had to be followed through (which meant spending the additional \$125 per case on arbitration) on every case that did not respond and pay up.

MICROLOAN bank decided to launch a pilot study by sending a random sample of defaulters the letter informing them of possible arbitration proceedings if they did not pay up and stay current on their bills. The idea behind the pilot was to identify segments of customers, based on their characteristics available in the internal databases, who responded to the threat. MICROLOAN bank managers believed that identifying a segment of customers with a high proportion of response to the letter would help them develop an efficient selection algorithm for defaulters who are ideal candidates for arbitration. This so called "arbitration group" not only promises a big positive impact on dollars of debt recovered and subsequent reduction in losses due to bad debt, it also guarantees minimal amount of arbitration expenditure since most of the defaulters pay up before any formal arbitration procedures are started.

The data was obtained from running the pilot had the following characteristics. There were about 1500 defaulters, whose responses were tracked over a period of one month after the letters were mailed. The letters were mailed in three separate batches and the results obtained over the one-month period after the mailings were recorded. The data set contains information about who responded to the mailing and sent in at least enough money to stay current on his/her bill payment (coded as 0- non responders, 1- responders). The data set also contains information about each member of the sample collected from various internal databases at the Bank.

Figure A

CART Model for Arbitration Sample



*Response is defined as paid current

Having no hypothesis about what type of payment defaulters respond to the threat of arbitration, the analysts at the bank decided to use all the cardholder information they had, to search which dimensions were able to discriminate between responders and non-responders. The hope was that in this data exploration process some sensible segmentation variables may be identified which could be used systematically to design an efficient selection routine for arbitration (see Figure A).

Based on the predictive model (CART model) on response to arbitration letter, analysts identified Segment 2 (positive response rate of 40%) as the prospects for sending through the arbitration process. The reason for the same is that the response in this segment was significantly higher than the average in the sample (24%).

However, the analysts still had to worry about the cost of sending the letter of intent to the other 60% non-responders in the segment 2 (leakage) in our initiative. The choice was between, a) not sending the letter to anybody (in which case the arbitration procedure was to be put on hold) and, b) sending letters to segment 2 identified by the antecedent variables that CART model provides.

In the latter case, analysts would have to factor the incremental response received in segment 2 to the letter (collections from 40 % of the target), against the cost of sending the other 60% non-responder through the process of arbitration (and pay the application fee) with a relatively smaller possibility of recovering money from them in the long term. This cost needs to be factored to account for the model leakage that we described in the text of the paper. In case of a perfect prediction model (which would identify the 40% responders only), such a leakage cost would not have been factored.

Turns out that in this particular context, after factoring in the cost, the bank would still make a healthy positive collection from the initiative and hence the modeling effort was rendered successful.