On the Optimal Allocation of Resources for a Marketing Campaign

Patrick Hosein[a], Shiva Ramoudith[b] and Inzamam Rahaman[c]
Department of Computing and Information Technology, The University of the West Indies, Trinidad and Tobago

Keywords: Resource Allocation, Marketing, Optimization, Data Mining, Machine Learning.

Abstract: Many companies and institutions, such as banks, typically have a wide range of products that they make available to customers. However, such products must be marketed to their customers, especially when the product is new. Phone calls, emails, postal mail, and online advertisements are among the ways companies can market products to specific customers. However, the cost incurred during marketing increases with every contact made. Phone calls are the most personal means of targeted marketing but also the most costly. In telemarketing, a company can make multiple calls to a single customer with each call incurring a human resource cost. Such calls may or may not be able to persuade a customer to subscribe to the service or product. Some customers might subscribe after the first call. Some customers might require several calls to convince them. Other customers might never be persuaded. In light of limited resources, to maximize return, a company would need to determine which customers to contact and how many attempts to make for a customer. In this paper, we present a mathematical model for this problem in which, given a marketing budget of calls, one can determine a policy for selecting customers to target along with the optimal number of calls to use for each selected customer. We illustrate our model using a Portuguese banking dataset and show that our model can achieve significantly higher levels of success performance.

1 INTRODUCTION

Marketing is an essential aspect of modern business. Businesses cannot earn revenue from a product if customers are unaware of said product. There are many media through which a business can market a product. These include email campaigns, online advertisements, social media advertisements, postal mail, and phone calls. Regardless of the medium, a business would want to minimize the cost of conducting a marketing campaign while simultaneously maximizing its efficacy.

Some customers are more likely to purchase certain products over customers. Moreover, some marketing media are more effective with different customers. Making a sound judgement on an individual customer is often impossible. For this reason, marketing often uses features of customers to divide them into customer segments (Loshin and Reifer, 2013). By dividing customers into segments, we can derive more refined marketing strategies that are more likely to entice a particular segment. Furthermore, we can decide if a specific segment is worth targeting.

Phone calls can serve as a useful, personal marketing tool and one potential benefit is that customers can directly signal their intention to purchase or partake in a product. Marketing conducted through the use of phone calls is called telemarketing (Kotler and Keller, 2011). However, as noted by Roach (Roach, 2009), not all segments might react favourably to phone-based marketing campaigns. A corollary of this insight is that not only would calls be wasted on certain segments (Mylonakis, 2008), but such calls can also annoy members of different customer segments. Such irritation can minimize a customer's willingness to consume more products in the future. Given the above, it would be shrewd to design phone-based marketing campaigns to target specific customer segments. The design of such marketing campaigns ought to consider that the budget of contact attempts that can be made is limited.

In this paper, we present a novel formulation of the problem of designing a data-driven phone-based marketing campaign. Given a budget of calls, our model uses historical data to divide a customer base into customer segments, and then using said segments and their computed properties, allocates the number of

[a] https://orcid.org/0000-0003-1729-559X
[b] https://orcid.org/0000-0001-9464-0954
[c] https://orcid.org/0000-0002-3097-8355

Hosein, P., Ramoudith, S. and Rahaman, I.
On the Optimal Allocation of Resources for a Marketing Campaign.
DOI: 10.5220/0010232001690176
In *Proceedings of the 10th International Conference on Operations Research and Enterprise Systems (ICORES 2021)*, pages 169-176
ISBN: 978-989-758-485-5

calls to be assigned across customer segments to yield the highest number of marketing successes. Note that, in targeting a particular segment, we optimize the maximum number of calls to be made for the segment so that sufficient calls are made to entice such customers but not too many are made such that they become frustrated. We validate our method using a dataset used by Moro et al. (Moro et al., 2014) from a Portuguese bank.

2 RELATED WORK

Several papers have examined the problem of designing telemarketing campaigns using computational techniques. Most papers in this space have looked into using machine learning techniques to deciding which customers a business ought to contact during a telemarketing campaign.

Karim and Rahman (Karim and Rahman, 2013) examined the problem from the perspective of binary classification. Using Moro et al.'s (Moro et al., 2014) dataset, they sought to use customer features to predict whether or not they would purchase the term deposit being marketed and compared the use of the C4.5 Decision Tree algorithm against Naive Bayes and found that the C4.5 Decision Tree algorithm produced more accurate results. However, Karim and Rahman (Karim and Rahman, 2013) did not take into account the number of calls made to customers.

Similarly, Lawi et al. (Lawi et al., 2017) also approached the problem of determining which customers to call by framing the problem as a binary classification problem. However, instead of decision trees and Naive Bayes, Lawi et al. (Lawi et al., 2017) compared SVMs against Ada-boosted SVMs. They performed grid searches to determine the best combination of hyper-parameters for their models. Their Ada-boosted SVMs performed better than regular SVMs.

Neural Network models have also been examined and compared in some previous work. Puteri et al. (Puteri et al., 2019) compared the use of radial basis functions (RBF) as activation functions against Sigmoidal activation functions for determining customers to call in the framework of binary classification. Puteri et al. (Puteri et al., 2019) found that a network with RBF activations in the hidden layer performed better than Sigmoidal activations in the hidden layer.

Aside from leveraging machine learning models for bank telemarketing, Moro et al. (Moro et al., 2015) also examined feature engineering in the context of bank telemarketing. In particular, they used sliding windows to compute measures of customer lifetime value (LTV). Customer LTV is a proxy for a customer's value over time based on projected future interactions (Dwyer, 1997). They were able to use sensitivity analysis to derive explanations for which LTV measures were most important and demonstrated that LTV measures computed from historical data are useful in predicting future behaviour, thereby obviating the need for acquiring more information about customers.

Bertsimas and Mersereau (Bertsimas and Mersereau, 2007) developed a dynamic programming formulation for allocating messages to multiple customer segments. In their paper, they also propose a Lagrangian relaxation of their initial dynamic programming problem and show that their Lagrangian relaxation performs well in practice. They assumed that customer segments are known and so do not provide a procedure for the extraction of customer segments from data.

3 MATHEMATICAL MODEL

We assume that we have a set of customers and each customer has a set of features. We can make several calls to a customer to convince them to purchase a particular service or financial product (e.g., a loan). The customer may, at some point, accept the product in which case we do not call them again for the duration of the campaign. After a specified number of calls we remove the customer from the pool of potential customers.

We assume that customers with similar features (classified as a customer segment) have the same probability distribution for acceptance of the product. Later we will demonstrate how this distribution can be estimated. Our objective is to determine which customers to target first and also how many times we should contact them before giving up. Note that excessive calling can lead to customer irritation, and this should be avoided. We first consider a mathematical formulation of the problem and later describe, using an example, the application of the formulation in practice. Let us first consider a single customer segment, and later we will consider how to allocate calls among customer segments.

Suppose that we have N customers each with features identical to those in the concerned customer segment. Furthermore, assume that each of these customers accepts the product with probability p_i on the ith call. If a maximum of k calls are made to each customer then let s_k denote the expected number of successes and let c_k denote the expected number of calls made. The expected number of success is given by

$$s_k = N\left(1 - \prod_{i=1}^{k}(1-p_i)\right) \tag{1}$$

and the expected number of calls is given by

$$c_k = N\left(k\prod_{j=1}^{k}(1-p_j) + p_1 + \sum_{i=2}^{k} i p_i \prod_{j=1}^{i-1}(1-p_j)\right) \tag{2}$$

Next consider the function s_k versus c_k as k varies. The gradient of this function at some k is given by

$$\frac{s_{k+1}-s_k}{c_{k+1}-c_k} = \frac{p_{k+1}\prod_{j=1}^{k}(1-p_j)}{\prod_{j=1}^{k}(1-p_j)} = p_{k+1} \tag{3}$$

Note that, in practice, p_k decreases with increasing k since a customer is less likely to purchase in successive attempts. Consider the piece-wise linear function with values at points (c_k, s_k). The gradients of the successive linear components will be decreasing and hence this is a piece-wise linear concave function. Note that we assume here that, once we increase the maximum number of calls for one customer in this customer segment it is increased for all. This convexity property is important since it now implies that when we optimize over customer segments we will be solving a convex optimization problem. Please note as both s_k and c_k varies by N then the plot does not change with N but the axes are scaled accordingly.

The objective of the problem is to determine how many calls to assign to customers in each customer segment. Let x_j denote the number of calls that are assigned to customer segment j. Let $S_j(x_j)$ denote the expected number of successes if x_j calls are allocated to customer segment $j = 1 \dots M$. This is the piece-wise linear, concave function that we derived above scaled by the number of members in j. The optimization problem becomes

$$\begin{aligned} V = \max_{\vec{x}} &\sum_{j=1}^{M} S_j(x_j) \\ \text{s.t.} &\sum_{j=1}^{M} x_j = T \\ &\vec{x} \in \{0,\dots,T\}^M \end{aligned} \tag{4}$$

Since S_j is piece-linear and concave one can show that a greedy approach can find a near-optimal solution to this problem. Find the customer segment with the largest initial gradient, assign as many calls as needed to get to the next break-point, update the gradient for that customer segment to that of the next linear segment and repeat until all T calls have been allocated. If the last assignment brings the total calls for the customer segment to the next break-point, then the solution is optimal, and so the solution is typically quite close to being optimal. The pseudo-code provided in Figure 1 can be used to determine the optimal allocation.

Require: T = Total number of calls to be allocated
Require: M = Total number of customer segments
Require: $s_j(k)$ = Expected #successes for a maximum of k calls
Require: $c_j(k)$ = Expected #calls for a maximum of k calls
Require: N_j = Number of customers in customer segment j
Require: $k_j = 0$ Initial maximum call value for customer segment j
Require: $g_j(k_j) = s_j(1)/c_j(1)$ set initial gradient for customer segment j
  while $T > 0$ **do**
    $j^* = \text{argmax}_j\{g_j(k_j)\}$
    $k_{j^*} \leftarrow k_{j^*} + 1$
    $g_{j^*} = \frac{s_{j^*}(k_{j^*}+1) - s_{j^*}(k_{j^*})}{c_{j^*}(k_{j^*}+1) - c_{j^*}(k_{j^*})}$
    $T = T - N_{j^*}$
  end while
  for $j = 1 : M$ **do**
    $x_j = N_j k_j$
  end for
  return $\vec{x}$

Figure 1: Call Allocation Optimization Pseudo-code.

Let us illustrate with a simple example. Assume that we have two segments, $j = 1,2$. For the first customer segment, suppose that the probability of success for each of the first five call attempts is 0.10, 0.07, 0.02, 0.01 and 0, respectively. Hence 20% of customers eventually accept, and the rest reject the offer. For the second customer segment, we assume that the corresponding probabilities are 0.06, 0.03, 0.02,0.01 and 0. The success versus call attempt plot is shown in Figure 2.

Now suppose that we have 500 new customers with 200 being in customer segment 1 (i.e., $N = 200$ for this customer segment) and 300 in customer segment 2. The initial gradients are 0.10 and 0.06, respectively and hence the first set of customers are chosen from customer segment 1. 200, one for each customer in the customer segment are necessary to get to the first break-point. The gradient for customer segment 1 drops to 0.07, so it is chosen again, but this time 180 calls are needed to get to the next break-point. At this break-point, the gradient drops to 0.02, so customer segment 2 is chosen next and 300 calls are required to get to the first break-point. Hence, if

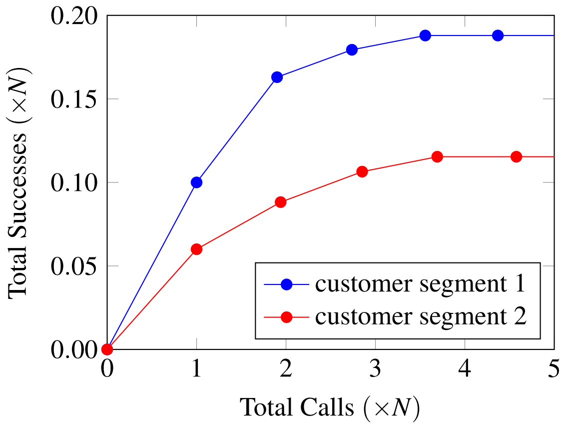

Figure 2: Variation of Successes with Calls.

we assumed a budget of 680 calls, then it is optimal to call every customer in segment 1 no more than twice and to call every customer in segment 2 once.

4 NUMERICAL RESULTS

In this section we provide numerical results to illustrate the benefit of the proposed approach in a real environment. We first describe the dataset used and then we provide details of our segment selection process. Next, for each customer segment, we determine the number of successes as a function of the number of calls made to customers with features in the customer segment. For some customers, after a few calls, it might be best to give up on them (and avoid their resentment), while for other customers it may take more calls before they can be convinced. Finally, we illustrate its performance improvement.

4.1 Data Description

We used the Bank Marketing Dataset collected by Moro et al. (Moro et al., 2014). The dataset is available on the UCI Machine Learning Repository. The dataset contains customers from a Portuguese banking institution and the direct marketing campaigns used to encourage them to subscribe to a term-deposit product. The dataset contains 45211 records, each representing a customer contacted during the marketing campaign. The chart in Figure 3 shows the frequency of contacts for various call ranges. We consider 99% of all calls made (we ignore customers contacted greater than 34 times) to filter potentially anomalous data points from our analysis. Table 3 lists the features from the dataset considered for our analysis.

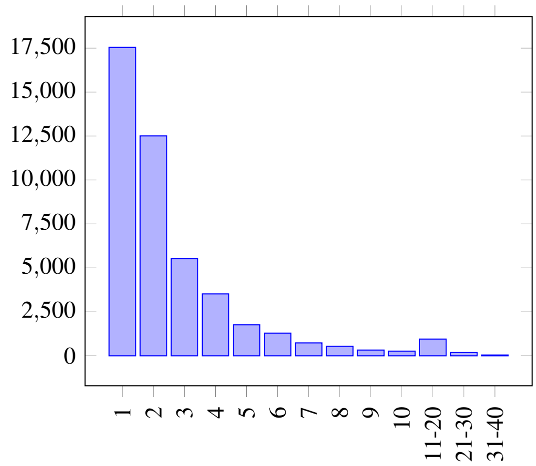

Figure 3: Number of Customers vs. Number of Calls made.

Table 1: Customer Features.

Attribute	Values
Age	Customer's Age
Balance	Customer's Average Yearly Balance (€)
Job	Administrator, Blue-collar, Entrepreneur, Housemaid, Management, Retired, Self-employed, Services, Student, Technician, Unemployed or Unknown
Marital Status	Married, Single, Divorced or Unknown
Education	Primary, Secondary, Tertiary
Risk	Yes, No or Unknown
Housing Loan	Yes, No or Unknown
Personal Loan	Yes, No or Unknown
Number of Calls	1-63
Success of Offer	Yes or No

Given new customers and a budget of calls, our objective is to determine which users to call and how often they should be called. We use historical data to determine customer segments, and for each customer segment, we determine the number of calls to assign to a customer in that segment. Therefore, we first investigate how to derive customer segments (feature selection), and then for each customer segment, we compute the success as a function of calls made. Given a new set of customers, we can then determine their customer segments, and based on this classification, decide whom to call and the maximum number of calls to be made to them.

4.2 Feature Selection

Given the chosen performance metric (success per call rate), we can now use feature selection to determine which features influence this metric. We selected features that influenced the metric to determine which customers to approach and the number of calls to make to them. Note that we could use feature extraction, but we chose feature selection to better understand which information of a customer is most important.

Given a new user, we can use their features to assign them to the appropriate customer segment. Before doing this, we need to define the list of possible customer segments. Our approach starts by computing the success rate of each value for a particular feature in isolation. Values were then aggregated based on their success rates. For some features, we used K-Means clustering to determine the aggregation of the values. We then use Silhouette analysis (Rousseeuw, 1987) (with Euclidean distance) to determine the optimal number of clusters. Using this analysis, we obtained the optimal grouping of values for a particular feature. Next, we give an example of this approach with the job, marital status and education features based on a random sample of users from the dataset.

For each occupation, we averaged the success per call rate of all customers with that job title. These averages are provided in the bar chart in Figure 4 for each of the job titles and are used as input for the clustering step. As mentioned before, we used K-Means clustering and computed the Silhouette Score for all possible cluster numbers. In this example, the highest Silhouette score corresponds to 2 clusters, which is what we use. The two clusters are ['unemployed', 'admin.', 'management', 'self-employed', 'technician', 'unknown', 'services', 'housemaid', 'blue-collar', 'entrepreneur'] and ['student', 'retired'].

The success per call rate for the various values of the marital feature are as follows: (Married, 0.037), (Single, 0.059) and (Divorced, 0.048). We repeated the process with these values in conjunction with their success per call rates to determine the resulting groups. They are as follows: [married, divorced] and [single].

Regarding the education feature, our approach revealed that two clusters were ideal: [primary, secondary] and [tertiary, unknown]. The success per call rates for the various values of the education feature are as follows: [(Primary, 0.032), (Secondary, 0.042), (Unknown, 0.052), (Tertiary, 0.055)].

It would not be feasible to use the approach mentioned above with the age and balance features since they are continuous. We chose to discretize each feature by creating a set of contiguous intervals that span the range of the feature's values. We determined the possible intervals by utilizing a Decision Tree algorithm since it is easy to interpret and it determines the optimal splitting points that would determine the contiguous intervals. We utilize the full dataset (age/balance and success of offer features as input) with the Decision Tree algorithm. We then extract the splits at each level. This will be used as the criteria for creating the intervals for the age and balance features. We considered the following hyperparameters: max depth and criterion. We limited max depth to a value of 5 since this has a direct impact on the number of customer segments. There was no difference between either criterion (Gini Index and Entropy) based on our testing. We did not consider other parameters for the Decision Tree since they play an insignificant role when the tree is shallow along with a sizeable dataset. We do not cluster the remaining features since the number of values for those remaining features were already small.

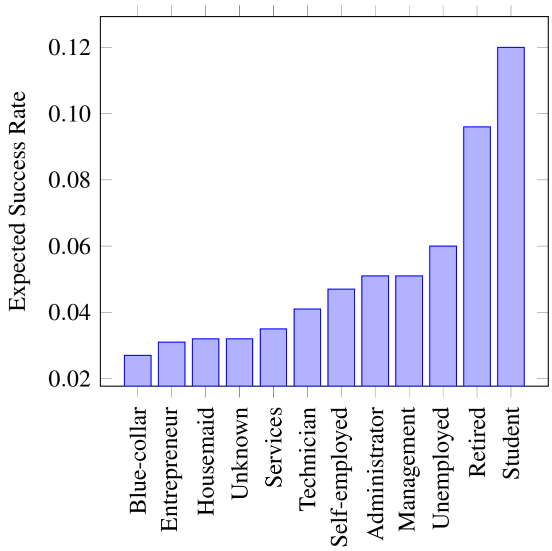

Figure 4: Average Success Rate for each Occupation.

Table 2: Success Rates for Other Features.

Attribute	Expected Success per Call Rate
Risk	Yes (0.023), No (0.045)
Housing Loan	Yes (0.061), No (0.031)
Personal Loan	Yes (0.025), No (0.048)

4.3 Optimal Allocation of Calls

We assume that we are given a set of features for each customer i and that we have historical data on the number of attempts made to each customer and the outcome (success or failure) of each customer. Let v_i denote the number of call attempts made to customer i. Let $q_i = 1$ if attempts to customer i are successful (i.e, success was achieved on call v_i) and $q_i = 0$ otherwise. We assume M customer segments and these are indexed by j.

Note that one can reduce the number of calls made by reducing the maximum allowed calls per customer, thereby reducing the number of calls made. This reduction will lead to fewer successes, but the net result might be a higher success per call rate. For the given dataset, the maximum number of calls were already chosen. However, we can deduce what would happen if this maximum was reduced. If the maximum is less than the number required to achieve success, then one would be unsuccessful since the required number of calls would not have been made. We use G_j to represent the set of users who have the features specified by customer segment j. Let us introduce the variable k as the specified maximum number of calls. The number of successes achieved for a call maximum of k for customer segment j is given by

$$s_k(j) = \sum_{i \in G_j} q_i[\min\{1, \max\{0, k - v_i + 1\}\}] \quad (5)$$

and the number of call attempts made is given by

$$c_k(j) = \sum_{i \in G_j} \min\{k, v_i\}] \quad (6)$$

We use historical customer data to compute the function $S_j(x)$ for a given customer segment j. This is a piece-wise linear function with endpoints given by $(0,0), (s_1(j)c_1(j)), (s_2(j), c_2(j))$ etc.

Let us illustrate this function for one of the customer segments (segment A in Table 3) obtained through feature selection. This customer segment contains 874 customers. In Figure 5 we plot (in blue) the function $S(x)$ based on the data. However we find that the function is not exactly concave and hence in order to apply the optimization approach previously defined we replace the function $S(x)$ with its Concave Hull denoted by $\tilde{S}(x)$. This hull is depicted in red. Note that the difference in this case is quite small.

4.4 Performance Analysis

We use 5-Fold cross-validation to evaluate the proposed method on the dataset in (Moro et al., 2014). We use 80% of the customers (randomly chosen) for training and the remaining 20% is used for testing.

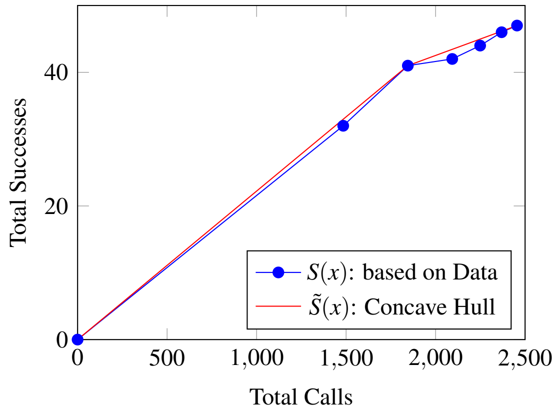

Figure 5: Successes as a function of calls made.

For each cross-validation fold, we use the training set to infer the groupings of education, marital, and job attributes. We form customer segments by all possible combinations of groupings across all features. We then assign customers to customer segments given their characteristics. We then apply the previously defined methods and compare performance among them. We repeat the process for each combination of the various age/balance intervals. We found the optimal intervals (based on our performance metric which is described later on) for the age and balance features respectively to be [(18, 25), (26, 59), (60, 87), (88, 93), (94, 100)] and [(-10000, 60), (61, 1578), (1579, 105000)].

Table 3 provides a breakdown of customer segments, randomly chosen, from best to worst, for one of the cross-validation folds. This fold contains 960 customer segments in total, with 661 of them having no customers. We recognize that customer segments having low success per call rates are associated with customers that have an existing loan. A possible reason is that these customers do not have the financial resources (low account balances) to support both the loan and the term deposit. An exception to this observation is customer segment 2 where persons did have an existing loan but still had the highest success rates. A potential reason is that these customers have high account balances. Customer segment 366 contains customers who defaulted on a previous loan. As a result, this segment had the lowest success rates. These customers are not ideal for targeting. In the dataset, there were only 52 customers who subscribed to the term deposit after defaulting on a loan.

Interestingly enough, 7 out of the top 10 customer segments are associated with customers greater than 59 years old and they represented the minority population in the dataset. These customers could potentially have savings that they are willing to invest in the term deposit. Next, we rank customer segments

Table 3: Some Sample Customer Segments for One of the Cross Validation Folds (ranked by success rate).

CS	Age	Balance	Education	Status	Job	Risk	Pers	House	Rate
A	26-59	-10000 - 60	(sec, prim)	M	(unemp, admin)	no	no	no	0.0169
2	60-87	1579 - 105000	(sec, prim)	M	(student, retired)	no	yes	no	1.0
16	18-25	61 - 1578	(sec, prim)	(S,D)	(student, retired)	no	no	no	0.2425
46	18-25	-10000 - 60	(tert, unk)	(S,D)	Same as A	no	yes	yes	0.1667
108	26-59	61 - 1578	(tert, unk)	(S,D)	Same as A	no	yes	no	0.0507
366	26-59	-10000 - 60	(sec, prim)	M	(student, retired)	yes	no	yes	0.0

by the average success per call rate, and, for a subset of these segments, we plot, in Figure 6, the average success rate as a function of maximum number of calls allowed. Once a call is successful then no more calls are made (even though the maximum allowed for that segment is continued to be increased in the plot). For the high success rate segments success tends to be achieved in the first call while for the others success tends to come at a later call.

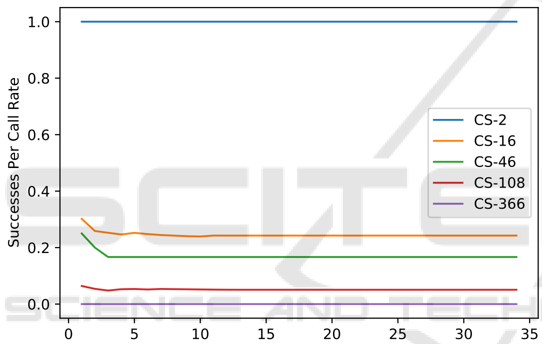

Figure 6: Success Rate vs. Call Limit for Various Segments.

We apply three methods to the test set for each fold. The first method is the present mode of operation, which we call the Baseline (BL) method. Here we assume that each call is provided to the customer who has not yet accepted with the fewest contact attempts made. Note that all customers are called once and then, for those who did not accept, they are called a second time, and this continues until the maximum allowed number of calls is reached. For simplicity, we only compute the cases where all customers who have not accepted received the same number of calls, and linearly interpolate between these points. Note that this approach does not use any information from the training set.

In the second approach, we order customer segments by their average success per call rate. We then exhaustively call all customers in the highest success rate customer segment, then the next highest, etc. In this case, the average success rate of each customer segment is computed from the training set and hence this approach uses some information from the training data to provide segment priority. We evaluate the successes and calls each time a new customer segment is added and plot these points but use a linear interpolation between these points. This approach is called the Greedy Customer Segment (GC) approach.

The final method is called the Gradient Ascent (GA) approach. Here we apply the approach detailed in Section 3 whereby we find the customer segment with the largest gradient, call each customer in that customer segment who can be called, update the gradient for this segment and then repeat. In this case, we incrementally choose customer segments that give the best improvement in success per call and hence will provide a near-optimal solution. The probability distribution for each customer segment is based on the training set data, so here we extract even more details from the training set.

We also determined an upper bound on performance as follows. Suppose that we know the outcomes for all customers(i.e. which customers subscribed and the number of calls needed to get them to subscribe). We can then allocate calls first to those customers who we know will subscribe and, of these, we start with those requiring the least number of calls. This will provide the most successes for a given number of calls and hence is an upper bound which we denote by (UB). Figure 7 shows a plot of successes versus calls for a sample cross-validation fold. The rest of the folds were very similar. As expected, for a given number of calls, GC is better than BL, and GA performs even better. Our implementation is publicly available at (Ramoudith et al., 2020).

4.5 Performance Metric

Note that combining performance results across folds is difficult as each test set contains a different number of calls and successes. Instead, we compute the area under the curve (AUC) for each method (BL, GC, and GA) and use this as a measure of performance. Note that a greater area under the curve indicates better performance, with the Upper Bound having the greatest area. Furthermore, since we are interested in the increase in performance over the baseline, we use the

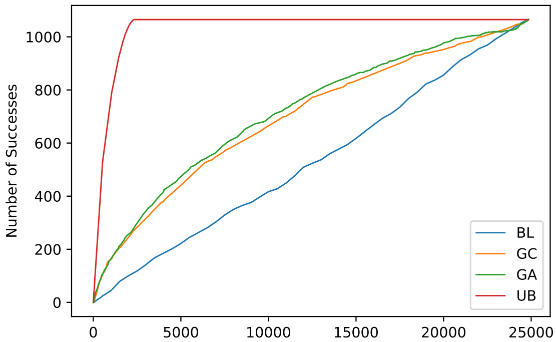

Figure 7: Successes versus Calls for a Sample Fold.

ratio of the AUC for GC and the AUC for BL as the performance metric for GC and similarly for GA. We average these ratios over all of the folds to estimate performance. When computed, these ratios were 1.34 for the GC approach and 1.38 for the GA approach. Note that this ratio for the upper bound case, UB, is approximately 2. Therefore, on average, we experience a call success rate gain of approximately 34% for GC and 38% for GA when compared to the approach used by the Bank. We can translate this into cost savings. Note that simply optimizing the allocation of calls to segments based on the average success rates provides most of the benefit. The Gradient Ascent algorithm provided a small additional increase of 4% in performance but at the cost of additional complexity. When deployed, the approach will work as follows. We will apply the method periodically (e.g., one week) to all eligible customers given the available number of possible calls. One would then contact the chosen customers based on their allowed call limit. Every two months, one may also use samples obtained over the prior two months to update customer segments and parameter estimates.

We are currently conducting a thorough evaluation of our methodology against machine learning approaches.

5 CONCLUSION

Our results indicate that implementing the proposed method would increase the success of telemarketing campaigns with a limited budget of calls. Additionally, a firm can use the computed customer segments to ameliorate other marketing decisions. In the future, we will repeat the analysis using additional features from the dataset and will also deploy a prototype to investigate our method's performance in practice. Note that, as more customer outcomes are collected, we can improve the accuracy of the estimated probability distribution of each customer segment and hence improve performance.

REFERENCES

Bertsimas, D. and Mersereau, A. J. (2007). A learning approach for interactive marketing to a customer segment. *Operations Research*, 55(6):1120–1135.

Dwyer, F. R. (1997). Customer lifetime valuation to support marketing decision making. *Journal of Direct Marketing*, 11(4):6–13.

Karim, M. and Rahman, R. M. (2013). Decision tree and naive bayes algorithm for classification and generation of actionable knowledge for direct marketing. *Journal of Software Engineering and Applications*, 6(4):196–206.

Kotler, P. and Keller, K. (2011). *A Framework for Marketing Management*. Prentice Hall, Upper Saddle River, NJ, 5th edition.

Lawi, A., Velayaty, A. A., and Zainuddin, Z. (2017). On identifying potential direct marketing consumers using adaptive boosted support vector machine. In *2017 4th International Conference on Computer Applications and Information Processing Technology (CAIPT)*, pages 1–4, Kuta Bali, Indonesia. IEEE.

Loshin, D. and Reifer, A. (2013). *Using Information to Develop a Culture of Customer Centricity: Customer Centricity, Analytics, and Information Utilization*. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA.

Moro, S., Cortez, P., and Rita, P. (2014). A data-driven approach to predict the success of bank telemarketing. *Decision Support Systems*, 62:22–31.

Moro, S., Cortez, P., and Rita, P. (2015). Using customer lifetime value and neural networks to improve the prediction of bank deposit subscription in telemarketing campaigns. *Neural Computing and Applications*, 26(1):131–139.

Mylonakis, J. (2008). The influence of banking advertising on bank customers: an examination of greek bank customers' choices. *Banks and Bank Systems*, 3(4).

Puteri, A. N., Tahir, Z., et al. (2019). Comparison of potential telemarketing customers predictions with a data mining approach using the mlpnn and rbfnn methods. In *2019 International Conference on Information and Communications Technology (ICOIACT)*, pages 383–387, Yogyakarta, Indonesia. IEEE, IEEE.

Ramoudith, S., Rahaman, I., and Hosein, P. (2020). Bank marketing approach - implementation. GitHub. https://github.com/shiv1994/BankMarketing.

Roach, G. (2009). Consumer perceptions of mobile phone marketing: a direct marketing innovation. *Direct marketing: an international journal*, 3(2):124–138.

Rousseeuw, P. J. (1987). Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. *Journal of computational and applied mathematics*, 20:53–65.