## Vision

It would do us some good to research what exactly a term deposit is and how a bank makes money from selling it. A term deposit is a fixed-term deposit (customers cannot withdraw the money for a set period of time without paying fees) held at a financial institution that a bank can use as capital to fund loans to other borrowers.[[1]](#footnote-1) The bank then compensates customers who purchase a term deposit with a portion of the interest generated from loans. We find that people are more likely to purchase term deposits during periods of high interest rates, since they are likely to earn more on their deposited funds, which is contrary to current market conditions – interest rates are decreasing.

If people are less likely to purchase term deposits when interests rates are low, and interest rates are currently on a downward trend due to various economic factors, it is even more important for us to reach out to customers sooner than later before people become less interested in term deposits. It is equally important for us to consult with stakeholders to fully understand the breadth of impact that changing market conditions will have on term deposit sales, and how fast we can expect to observe the impact. Furthermore, with lower interest rates on the horizon, the bank needs to prepare another marketing campaign to sell its mortgage products. If the bank can maximize the revenue generated from the current campaign, there will be more funds available to fund the mortgage marketing campaign which will likely generate more revenue for the bank than term deposits

We know that the bank is constrained by a $50K budget so we are limited in the number of customers we can contact. But how many customers does $50K equate to? This depends on the average call per customer which we can roughly estimate using hourly employee costs and the estimated length of sales calls (see the Costs sheet of the Problem Model Excel workbook). Assuming an average cost of $8 per call, the bank can contact 6,557 customers in a population of 100,000 – the number of remaining customers who have not been contacted as part of the current campaign. We can maximize the value of the $50K budget by selecting the 6,557 customers who are most likely to purchase the term deposit product rather than targeting customers randomly.

Luckily, the marketing team has already executed a pilot campaign, contacting 41,188 customers which provides us with a dataset that includes customer, customer relationship, and socioeconomic attributes since the data has already been aggregated by the engineering team. To select the 6,557 customers that are most likely to invest in a term deposit, we can use the dataset to build a model that ranks customers by the probability that they will invest. We can start with a simple model based on stakeholder domain knowledge to use as a baseline and, since this model is easy to create, we can follow up with a more complex model that leverages machine learning to improve the results.

Before we proceed with a model, we need define our success criteria. As we stated earlier, we are interested in selecting the 6,557 customers we have budget to contact such that more of them are likely to invest in a term deposit than would be expected under business-as-usual conditions. With more conversions, we will generate a higher ROI on our marketing spend, maximizing our marketing investment. Therefore, we can state our success criteria as a lift in conversion rate for the targeted customers compared to what we would expect based on existing campaign results. Since we already have a fairly large dataset of campaign result data for our current campaign and the previous campaign, we can estimate our expected conversion rate.

After reviewing the pilot campaign dataset, we find that the conversion rate for the current campaign is 11% while the conversion rate for the previous campaign was 24%. With such a large gap, there appears to be room for improvement. After informing stakeholders of the different conversion rates, they set the ambitions goal of doubling the current conversion rate for the targeted customers to align with previous campaign results – a lift of ~2.2 or 220%. From our perspective, we would rather model the problem using an uplift model[[2]](#footnote-2) since customers may have purchased a term deposit anyway through different channels such as in-store and online but, due to time constraints, we do not have time to develop such a model.

While we are using conversion rate as our success KPI, we are also interested in understanding the financial impact of our targeted marketing efforts. Conversion rates are associated with financial impact through a profit equation which can be stated as *ConversionRate \* NumberOfCustomersTargeted \* Revenue - (1 - ConversionRate) \* Cost*. We already have an estimate for cost but we still need an estimate of revenue. Observing three different groups of deposit sizes in historical data, we calculate an unweighted average revenue of $1,083 per customer. Using our cost, revenue, and conversion rate estimates, we find that a model that delivers a lift of 2.2 is poised to deliver an additional $2.01M in profit (before project costs). All of these figures are available in the Problem Model Excel workbook.

With our definition of success, we discuss with bank stakeholders what they feel is the best indicator to predict conversion in order to establish a baseline model. It turns out that a customer who has responded to a previous marketing bank campaign is likely to convert in the next one so that will be our baseline model. Ideally, we would use some variation of a customer lifetime value estimates to build an additional baseline model but this data is not available. We will judge the success of follow-up modelling efforts to be the improvement over this baseline model.

Since we’ll be developing a machine learning model to try and improve over the baseline model, and since estimating the generalization performance of a machine learning model requires us to train it on a dataset that is independent from an evaluation or test set, we’ll build the baseline model using the same training set. Once the baseline model is built, we move onto to developing the machine learning model which brings up the questions of which metric we’ll use to optimize it and our experimental design.

Since we’re interested in identifying a bounded subset of the population that will maximize the conversion rate and we’d ideally like to tune the model to accept more false positives (selecting customers to target even though they don’t end up converting) if costs are marginal, we can evaluate performance using an estimation of cumulative gain which equivalent to a ROC curve. Since we need a score to compare multiple models in the modelling process and need a single value to do so, we can use AUC which can be interpreted as the probability that a randomly chosen positive instance will be ranked ahead of a randomly chosen negative instance.

After performing a brief exploratory analysis of the pilot campaign dataset, we found that there were several attributes that are predictive of conversion in the current marketing campaign. Sticking with the idea of doing the simplest thing that could possibly work we’ll use the attributes as is, with some preprocessing so that they are compatible with machine learning models. If the initial model we produce is insufficient or we have time to experiment with additional features then we will do so in a follow-up experiment.

A successful model will identify a subset of customers to target for which the conversion rate will be higher than the conversion rate of the entire population (11%) and will ideally minimize the number of false positives so as to reduce campaign costs. Since our test set contains 8,238 customers and our target population contains 100,000 customers, we’ll keep the targeted group in the test set proportional to the 6,557 customers (~6.6%) that we can afford to target in a full-scale deployment. With a test set size of 8,238, this works out to 543 targeted customers.

With a vision for a model, our next step is designing an experiment test our hypothesis – that we can improve the conversion rate of customers buying term deposits by targeting the most likely customers. Formally, our hypothesis is as follows:

* Null Hypothesis – There is no difference in the conversion rate between choosing customers to target at random and the model conversion rate or
* Alternative Hypothesis – The model conversion rate is higher than the random conversion rate or .

The first step in the experimental design is figuring out our deployment strategy. Do we want to first test the efficacy of the model by deploying it to a random sample of the customers, before deploying it to the full population? Do we have time to do this? Do we want more evidence to support our conclusions, thus requiring a larger sample size? These questions classically require us to choose between exploration and exploitation.

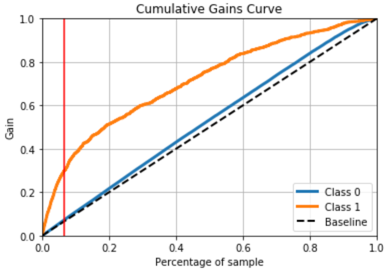
We could choose to explore, scoring a random sample of the population, contacting the top 6.6% then comparing that group’s conversion rate with the conversion rate of the remaining 94.4% of the population, or we may trust our model enough to contact only the 6,600 customers in the full population that the model predicts are most likely to convert. Ideally, we would deploy the model to a subset of the population to confirm that it delivers the results we expect (a lift in conversion rate) before deploying to the full population but that would require more time which we don’t have, since the director of marketing wants to move forward ASAP with the new mortgage campaign.

The next step in our experimental design is deciding how we’re going to assign customers to treatments. Since the purpose of a targeting model is to select only those customers with the highest model scores to target, we necessarily need to measure the response rate of the customers with the highest scores and so they will be assigned to the treatment group. We do, however, need to holdout a proportion of the high-scoring customers, assigning them to the control group, to get an accurate estimation of response rate of the “general” population (including high-scoring customers).

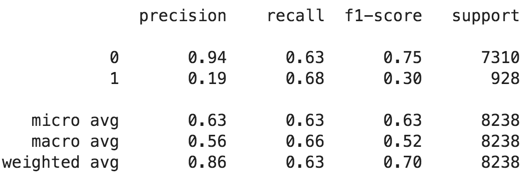
## Arguments

Using recent and historical results of marketing campaigns, we have the ability to predict which customers in the remaining population will be more likely to convert so we can target them and maximize our campaign spend. After reviewing results from a recent term deposit pilot campaign executed by the marketing team and prepared by the engineering team, we estimate that the expected conversion rate to be 11% if it were deployed to remaining population (100,000 customers). We also know, from historical data, that a previous marketing campaign achieved a conversion rate of 24%.

By simply choosing to target customers who converted in the last campaign and choosing customers at random when we run out of customers who previously converted, we found that we could improve the conversion rate by 254% (to 39%) in a sample size of 8,238, targeting 543 customers (6.6% of the sample). While this is an impressive improvement, we found that a basic machine learning model performs even better at a lift of 354% (to 51%). As shown below, the model identifies 30% of campaign respondents in the top 6.6% of the population, ranked by probability to convert.



The classifier’s precision is low at 20%, meaning that 80% of the targeted instances are false positives and won’t convert even though they were predicted to do so. Given the large profit margin of contacting a customer by phone, as estimated by our costs and benefits, we can accept the low precision.



We believe our results will generalize to the remaining population of customers to be contacted since we used an independent training set to train the machine learning model and evaluated its performance on a separate test set. We were also sure to select a proportion of customers to target that is consistent with what our budget will allow for in the remaining population.

There’s a possibility that our training distribution does not align with the distribution of the remaining 100,000 customers we need to select from which is why a formal experiment is critical to measuring the impact of our efforts. Another consideration is that the marketing team did not design a formal experiment that randomly assigned customers to call-centre agents. We do not see this as being a major concern, however, in the estimation of the campaign’s conversion rate since customers were still randomly assigned to agents.

TODO: Arguments for the experiment.

1. https://www.investopedia.com/terms/t/termdeposit.asp [↑](#footnote-ref-1)
2. https://link.springer.com/article/10.1057/jma.2014.18 [↑](#footnote-ref-2)