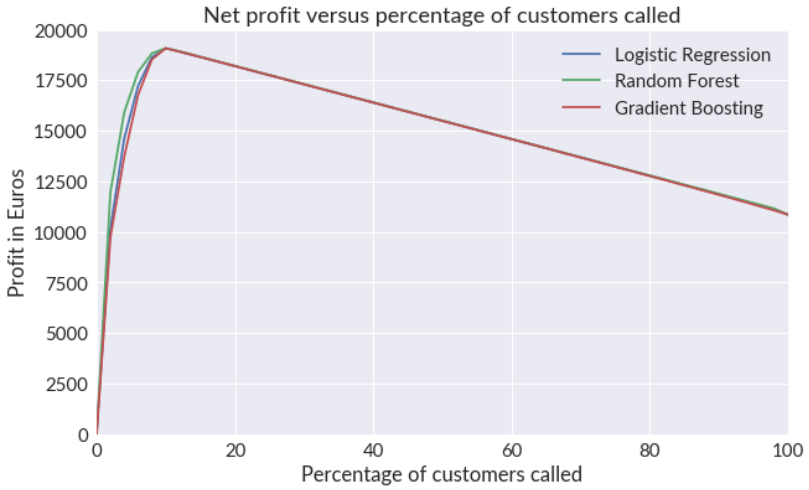
- Given the defined payoff per customer and calling cost, the objective is to maximize the sum of the net profit over all contacted customers
- Net profit is equal to $\sum_{TP}(0.01 \times \text{Balance} - 1) + \sum_{FP}(-1)$
- Need to account for the fact that misclassification cost is asymmetric, i.e. false positives are ‘cheaper’ than false negatives, since in general the expected payoff will be greater than the call charge.
- One possible strategy is to modify the decision function of the classifiers so that we reduce the number of false negatives at the expense of increasing the number of false positives.

Threshold Method



- Calculate profit for each entry in test data
- Build classification models using training data with targets 0/1 as previously
- Calculate the Expected Profit for each customer X: **E(X) = profit(X) * p(X=1)**
- Bin the Expected Profit into 20 bins and calculate the profit in each population bin
- Plot the cumulative profit as a function of the percentage of customers called
- In the test case (9403 customers), the random forest classifier yielded the maximum profit of approximately 18300 Euros when 10% of the customers were called before the cost of unsuccessful calls started to reduce profit

If the total number of customers is 10 million, the expected profit is approximately 20 million Euros