Predictive Modelling for Customer Targeting
A Banking Example
Pedro Ecija Serrano
Customer Targeting

• What is it?
• Why should I care?
• How do I do it?
What Is Customer Targeting?

• I want to sell more to my customers. I want to propose only relevant products.
  – Product targeting: “What products may this person want to buy?”

  – Customer targeting: “I have a product, who may want to buy it?”

• Modelling approaches vary (recommender systems vs. propensity modelling)
What Is Customer Targeting?

- Amazon: “Customers who bought this item also bought…”
- Netflix: “Trending Now”, “Because you liked…”
- These are recommender systems – Product Targeting.
What Is Customer Targeting?

Propensity modelling: identify customers likely to buy a specific product.

Would you like to take advantage of sophisticated analytics without a major investment?

Mastercard Propensity Models for Marketing draw from transaction behavior to provide objective information about likely future behaviors in the areas of card use, likely preference and retention.

Account scores help you make more profitable decisions by directing the right messages and offers to the cardholders most likely to respond. The top 10-20% of accounts identified through these models typically represent 50% or more of the desired activity across all accounts, or are 2-3 times more likely than average to engage in the forecasted behavior.

These models are available on consumer debit, consumer credit and commercial portfolios.

Card Expansion
- Cross Border
- Recurring Payments
- Bill Pay
- PayPass® High Activation
- PayPass® High Usage

Cross-Sell
- Small Business
- Product Cross-Sell

Card Activity
- Category Spend
- Early Months on Book (EMOB)
- Likeliness Spend
- Category Expansion
- Overall Card Usage

Retention
- Spend Attraction
- Reactivation
- Balance Attrition

(MasterCard Advisors: http://bit.ly/2gHgkUP)
What Is Customer Targeting?

- Recommender systems are useful (required) when your product range is very large.
  - Have you seen the catalogue in US Netflix?
  - Amazon UK had nearly 490 million items in 2015.

- Propensity modelling is suited to a product-specific model approach.
  - A model per product.
  - Contact the right customers with the right offer through the right channel at the right time.
Why Should I Care?

The Netflix Prize

21st September 2009: Netflix awards a $1 million prize to “BellKor’s Pragmatic Chaos” after three years of ongoing competition. 
https://en.wikipedia.org/wiki/Netflix_Prize

Amazon’s Cross-Selling

Amazon’s many recommender systems are jointly responsible for 35% of the company’s sales.

Why Should I Care?

The Netflix Prize

21st September 2009: Netflix awards a $1 million prize to “BellKor’s Pragmatic Chaos” after three years of ongoing competition. 
https://en.wikipedia.org/wiki/Netflix_Prize

Amazon’s Cross-Selling

Amazon’s many recommender systems are jointly responsible for 35% of the company’s sales.

Why Should I Care?

The Netflix Prize

21st September 2009: Netflix awards a $1 million prize to “BellKor’s Pragmatic Chaos” after three years of ongoing competition. 
https://en.wikipedia.org/wiki/Netflix_Prize

Amazon’s Cross-Selling

Amazon’s many recommender systems are jointly responsible for 35% of the company’s sales.
Why Should I Care?

“Our industry has been quite traditional in terms of marketing and cross-selling, but consumers now expect you to pop up at the most relevant point and in the most relevant way to them, which is not necessarily on your own website.” He clarifies that this targeted approach to customer engagement uses customer analytics and internal or external data.

http://www.theactuary.com/features/2016/12/the-digital-revolutionary/
How Do I Do It?

Banking Telemarketing: selling bank deposits to your customers.

- Acknowledgement: [Moro et al., 2014] S. Moro, P. Cortez and P. Rita. A Data-Driven Approach to Predict the Success of Bank Telemarketing. They made the database public.


- 41,188 customers; 4,530 historical buyers, many customer-related variables (age, education, etc.)
How Do I Do it?

There are a few challenges…

• Define the problem.
• Internal data and external data.
• Class imbalance.
• Model choice.
• Is there an optimal solution?
• What does it look like?
Define The Problem

• Classification vs. Regression
  – Does it actually matter? Do we really need the probability of buying?
  – Ideal model vs. reality. The ideal model might not be good enough. What then?
Internal Data and External Data

• The existing variables were not enough.
• We use external, publicly available data to supplement.

**Internal Data**
- Customer
- Internally available
- Age, education, occupation, etc.

**External Data**
- Environment
- Publicly available
- Interest rates, employment rates, consumer confidence, etc.
Class Imbalance

• 41,188 customers with 4,530 historical buyers (11%). Pause for thought.

• There is a wild suggestion… no one wants our product. This is correct for 89% of customers (100% correct for non-buyers and 0% correct for buyers).

• I challenge you to find a model with better accuracy than 89%!
Class Imbalance

• The problem is the 89:11 proportions. If only it were 50:50…

• We can try the following:

Do nothing

Undersample

Oversample

Balanced sample

Synthetic sample – SMOTE!
Model Choice

• Just a few suggestions:
  – Naïve Bayes, Support Vector Machines, Logistic Regression, Decision Trees, Random Forests, XG Boosting, K-Nearest Neighbours, Neural Networks, Lasso Regression, Elastic Net, etc…

• How did we define the problem? Classification vs. Regression.

• Are predictions enough or do we also need information on model factors?
  – Black box vs. transparent model.

• Wishful thinking vs. reality: do we really have a choice?

• Let’s focus on two: Naïve Bayes and Support Vector Machines.
Naïve Bayes

• Bayes theorem for conditional probabilities.
• It allocates observations to their most probable class.
• Some assumptions may be hard to meet (normality, non-correlation…)
• However the classifier can be very effective even when assumptions are not met.
Support Vector Machines

• Geometric approach.
• SVM with different kernels:
  – Linear
  – Polynomial
  – Radial
  – Sigmoid
• There is room for error.
• We can play with that room for error.
Naïve Bayes vs. Support Vector Machines

<table>
<thead>
<tr>
<th>Naïve Bayes</th>
<th>SVM</th>
</tr>
</thead>
<tbody>
<tr>
<td>- Regression</td>
<td>- Classification</td>
</tr>
<tr>
<td>- Transparent</td>
<td>- Black box</td>
</tr>
<tr>
<td>- Rigid</td>
<td>- Customisable room for error</td>
</tr>
<tr>
<td>- Fast</td>
<td>- Iterative (i.e.: slow)</td>
</tr>
</tbody>
</table>
Is There an Optimal Solution?

• How do we assess if a solution is good?
  – Training (2/3) and testing (1/3) datasets.
  – Confusion matrix.

<table>
<thead>
<tr>
<th>Truth \ Predictions</th>
<th>Buyer</th>
<th>Non-Buyer</th>
</tr>
</thead>
<tbody>
<tr>
<td>Buyer</td>
<td>True Positive</td>
<td>False Negative</td>
</tr>
<tr>
<td>Non-Buyer</td>
<td>False Positive</td>
<td>True Negative</td>
</tr>
</tbody>
</table>

• We want to maximise True Positives. What about…
  – False positives?
  – False negatives?

• We do not have a cost-benefit matrix to provide an optimal solution.
Is There an Optimal Solution?
What Does an Optimal Solution Look Like?

<table>
<thead>
<tr>
<th>Model</th>
<th>Buyers</th>
<th>Calls</th>
<th>Success Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vanilla SVM Radial</td>
<td>801 (18%)</td>
<td>1,195 (3%)</td>
<td>67%</td>
</tr>
<tr>
<td>Vanilla Tree</td>
<td>1,010 (22%)</td>
<td>1,566 (4%)</td>
<td>64%</td>
</tr>
<tr>
<td>Synthetic Tree</td>
<td>1,956 (43%)</td>
<td>4,325 (11%)</td>
<td>45%</td>
</tr>
<tr>
<td>Synthetic SVM Radial</td>
<td>2,668 (59%)</td>
<td>6,632 (16%)</td>
<td>40%</td>
</tr>
<tr>
<td>Undersampled Naïve Bayes</td>
<td>3,021 (68%)</td>
<td>10,586 (26%)</td>
<td>29%</td>
</tr>
<tr>
<td>Calling all customers</td>
<td>4,530 (100%)</td>
<td>41,188 (100%)</td>
<td>11%</td>
</tr>
</tbody>
</table>
Anything Else?

• An accurate model is good….

• …but a good model that provides information on customers too is better.

• We can use Naïve Bayes’ conditional probabilities and descriptive statistics to learn about buyers.
Colour palette for PowerPoint presentations

- Dark blue: R17 G52  B88
- Gold: R217 G171  B22
- Mid blue: R64 G150  B184

Secondary colour palette

- Light grey: R220 G221  B217
- Pea green: R121 G163  B42
- Forest green: R0 G132  B82
- Bottle green: R17 G179  B162
- Cyan: R0 G156  B200
- Light blue: R124 G179  B225
- Violet: R128 G118  B207
- Purple: R143 G70  B147
- Fuscia: R233 G69  B140
- Red: R200 G30  B69
- Orange: R238 G116  B29

Number of previous contacts

- Age

Previous outcome

Supers vs. Previous outcome

11 September 2017
Anything Else?

The average buyer is:

• An existing customer who has bought our products before.
• Well educated.
• Best contacted by mobile phone.
• Free from family responsibilities (student or retired)
• Of good credit standing (no past loan defaults)

We have also learned that timing is critical. The following factors favour sales:

• Increasing unemployment.
• Low interest rates.
And That’s All!

Now you know:

• What customer targeting is.

• How it adds value to your business.

• How to do it.

In addition:

• You might be inspired to apply these techniques in your actuarial work!
The views expressed in this presentation are those of invited contributors and not necessarily those of the IFoA. The IFoA do not endorse any of the views stated, nor any claims or representations made in this presentation and accept no responsibility or liability to any person for loss or damage suffered as a consequence of their placing reliance upon any view, claim or representation made in this presentation.

The information and expressions of opinion contained in this publication are not intended to be a comprehensive study, nor to provide actuarial advice or advice of any nature and should not be treated as a substitute for specific advice concerning individual situations. On no account may any part of this presentation be reproduced without the written permission of the author.