MVAS Adoption Likelihood Model

- Sagar Gupta
- Raghuraman Chandrasekhar
- Yash Chandwani
- Sudhanshu Dharmadhikari
- Saurabh Choudhary

Predicting the likelihood of using Mobile Value added service (MVAS) by telecom customers

Situation

- The client, a leading telecom player , has several million customers
- The client wants to identify growth opportunities, and wants to cross sell MVAS to its customer portfolio
- Client wanted to rank survey respondents' likelihood to respond to MVAS

Challenges

- The proportion of non MVAS customers to the total number of customers is very small
- For existing MVAS customers, we do not know the repetitive behavior, and hence could not put a financial model in place

Key Questions

- What factors drive the MVAS adoption amongst customers
- What is the strength of the above mentioned variables in driving adoption of MVAS

The model was developed following a rigorous analytical process with the following phases

PHASE I

Project Requirements
Data specification & objectives

- **Understand project needs & expectations**
- Study the available data and prepare hypothesis on the business needs
- Define 'dependent' variable

PHASE II

Data Preparation
Data audit & new variable creation

- Study & analyze description for each value for the variables
- Identify variables useful for further analysis
- Check for missing values & duplication
- Create 'Independent' variables

PHASE III

Model Development
Exploratory data analysis & model dev. & refinement

- Univariate analysis
- Bivariate & Correlation analysis
- Logistic regression /K-NN approach
  - Selecting relevant predictors by subjective analysis
  - Binning relevant predictors into buckets
- Model validation and re-running the model on different data-sets

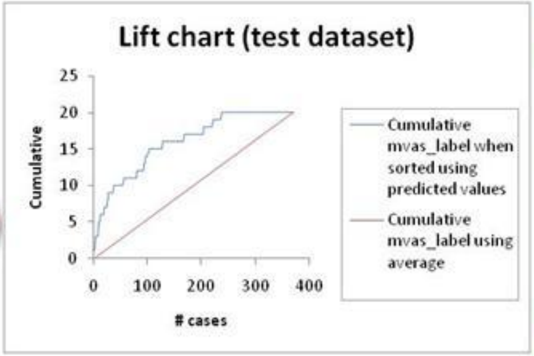
PHASE IV

Post Model Analysis
Insights and recommendations

- **Analyze model results and present the findings**
- **Study the error rate and lift charts**

Two competing modeling approaches were pursued to arrive at an effective model. Top 3 deciles capture more non MVAS subscribers in logistic as compared to KNN

1 Logistic Regression



PROS

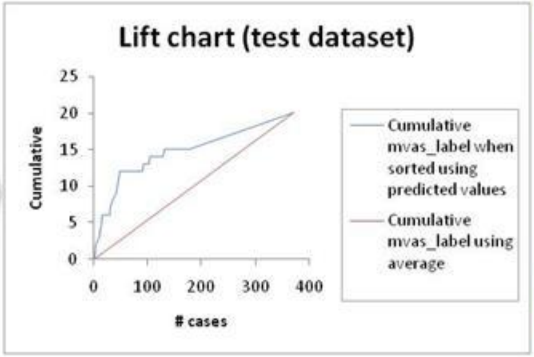
- Superior "goodness" of fit when predicting very low response rates
- Easy to understand model and quantify impact of drivers
- Can adapt cut-off to cater to various business scenarios

CONS

- Lengthy model building process – Data Preparation is very key
- Non-linearities and interactions need to be handled manually

---

2 KNN Model



PROS

- Minimal data preparation enable faster model development
- Simple implementation

CONS

- Sensitiveness to noisy or irrelevant attributes

Impact of different variables on the model provides enhanced understanding of the characteristics captured by those variables

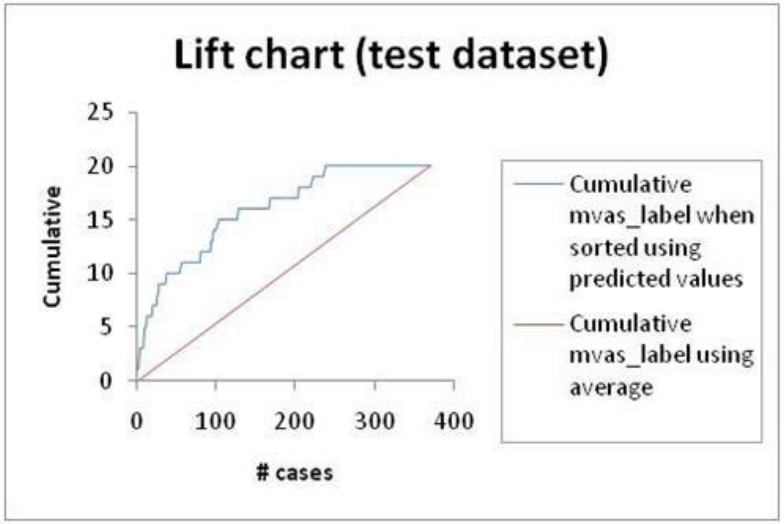
Equation of the model built with β coefficients

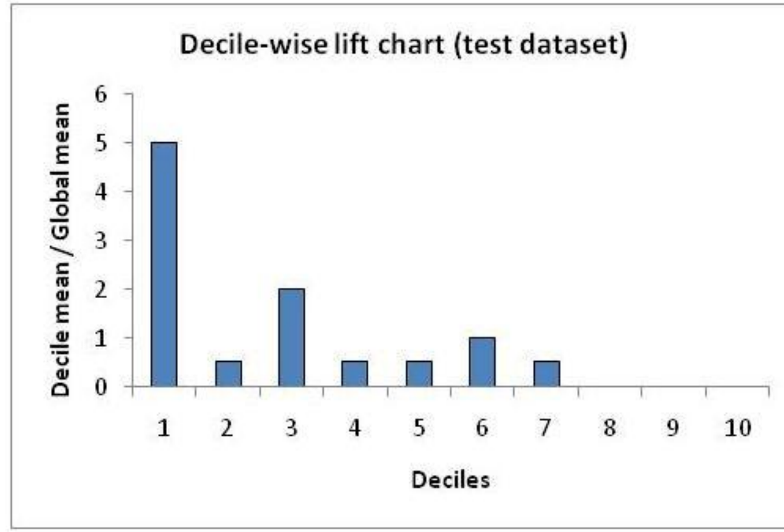
Y= 3.64 - 0.51*X1 – 0.9*X2 + 0.023*X3 - 0.37*X4 + 0.4*X5 – 0.75*X6 – 0.13*X7 – 0.79*X8 -0.19*X9 – 1.48*X10 + 0.02*X11 + 0.38*X12

#	Var.	Positive Drivers
1	X3	Network duration
2	X5	Prefer Apple
3	X11	Age
4	X12	Yearly household income

#	Var.	Negative Drivers
1	X1	Number of mobiles
2	X2	Primary Mobile Type
3	X4	Last Handset purchased
4	X6	Current Handset Brand
5	X7	Frequency of changing handsets
6	X8	Monthly expenditure on mobile
	X9	Usual top up size
	X10	Average SMS per day

75% of the non MVAS subscribers can be identified when 30% of the cases are flagged





Classification Confusion Matrix		
	Predicted Class	
Actual Class	Non-MVAS	MVAS
Non-MVAS	6	24
MVAS	24	504

Recommendations

Sell the service to customers who have:

- ***A smart phone***
- ***Spend at least 750 on monthly expenditure***