Mining Marketing Data

Andrea Ahlemeyer-Stubbe
Vice President
European Center of Database Marketing
Member of Pro-ENBIS and ENBIS

1:1 Marketing or CRM

The basis of 1:1 marketing is share of customer not just market share. So Instead of selling as many products as possible to who ever will buy them. We try to sell to one customer as many products as possible over the lifetime of that customer.

Marketing Challenges

- Many products, each with different Customer characteristics, acquisition costs, churn rates, etc
- increased diversification
- greater geographic range

Marketing Challenges

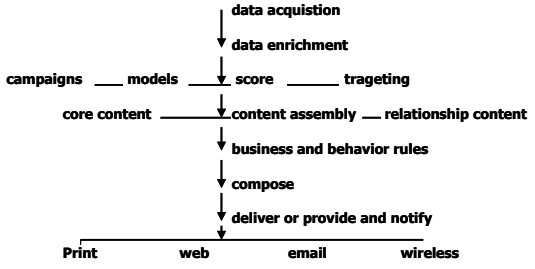
- number of customers
- number of types of relationships
- number of relationships with each Customer
- variety of channels for interaction
  - personal
  - telephone
  - web
  - e-mail

Therefore Businesses need to know:

- Who their customers are
- What their needs are
- How to communicate with them
- What is Their potential value

Which leads to the Business need for analysis, planning,recording and delivery mechanisms

Campaign Data Flow



Data Mining

Data Mining can be seen as the analysis of large quantities of data in order to discover meaningful trend and patterns.

Data Mining tasks in Marketing includes:

- Prediction
- Estimation
- Classification
- Description

Mining Marketing Data = Data Farming

- Preparing the field
  - **Data preparation**
- Seeding the crops
  - **Appending external data**
- Cultivating the soil
  - **Model development**
- Harvesting the crops
  - **Applying the model**
- Ploughing back the field
  - **Learning from results**

Who is useing Data Mining for Marketing and Why?

Industry

- Telecommunications
- Banking / Financial Services
- Insurance
- Mail Order / Catalogues
- Travel & Tourism
- Charity / Non Profits
- Publishing

Applications

- Response
- Cross-sell
- Attrition / Churn
- Valuation
- LTV Life Time Value

Customer Life Cycle Management

Data Mining is key to managing all phases of the Customer life cycle

- Customer acquisition
  - increase the number of relationships
- Value over time
  - increase the profitability of relationships
- Retention and attrition
  - Increase the duration of Relationship

Can we use known information about a customer to predict their future behaviour?

Information Sources

- Customer Databases
- External / Lifestyle Databases
- Census of Population
- Market Research

Geographic Data

- Country
- State / County
- City
- Postcodes / Zip Codes
- Population Density
- Store catchment Areas
- ...

Demographic Data

- Age
- Gender
- Family Status
- Family Size
- Income
- Occupation
- Home Ownership
- ...

Psychographic Data

- Lifestyle
- Personality
- Attitudes
- ...

Behavioral Data

- Purchase History
- Response History
- Communication channels
- Payment
- ...

What Can You do with Data?

- describe
- predict
- prove causality

Descriptions

- Descriptions summarise characteristics of data
- What are the important characteristics to be described?
  - Traditional business drivers are used to guide descriptions
- Descriptions are often generalised to make (possibly unsupported) predictions

Data Description Questions

- How much of a product, in this period, at this place?
- How effective was a cross selling campaign?
  - Revenue produced, cost, ROI
  - Conversion rate
  - pair-wise vs.. Overall effectiveness
- How effective was a promotion?
  - Revenue produced,cost,ROI
  - Conversion rate
- Recently, frequency and monetary (RFM)

Profiling

Description
- What are the characteristics of people for whom a promotion / cross sell was successful or unsuccessful
- What are the characteristics of churner and non-churners.

Prediction
- To whom should I offer a product
- Who will churn
- Who is likely to declare bankruptcy

Predictions

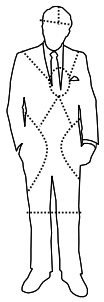
- Of my prospects, who is likely to become a good Customer?
- Which cross selling appeal should I make to which customers?
- Who is likely to chrun and what can I do to reduce the likelihood?
- What response to a Customer or Prospect will produce the best result?
- Which customers will click through to purchase?

Segmentation

Segmentation is a business problem that requires identifying groups with common characteristics
- Clustering is a descriptive technique that identifies potentially meaningful groups
- Classification is a predictive technique that identifies membership rules for known groups

Customer Segmentation

- A Priori
  - Not based on Data analysis
- Unsupervised Classification
  - Alike with respect to several attributes
- Supervised Classification
  - Alike with respect to a target, defined by a set of inputs



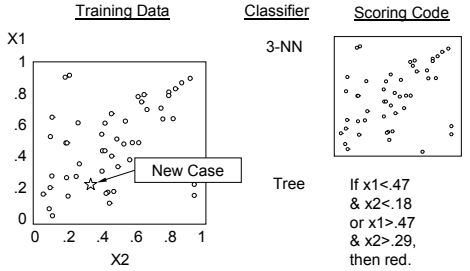
Scoring

Scoring assigns a number to a particular Customer, prospect, manufactured unit, etc.
- ranks by likelihood of a response, e.g. places Customer into a particular decile only order is important
- calculates a probability of response.
- accuracy of probability is important

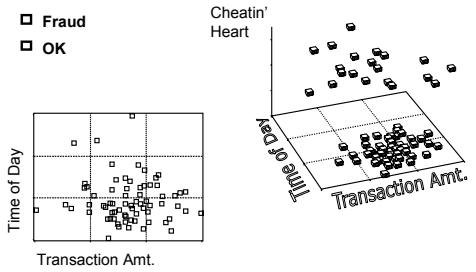
Scoring Recipe

- Model
  - Formula
- Data Modifications
  - Derived inputs
  - Transformations
  - Missing value imputation
- Scoring Code
  ≠ Scored Data
  ≠ Original computation algorithm

Sociability



The Secret to Better Predictions



Target Marketing

- Cases = customers, prospects, suspects, households
- Inputs = geo/demo/psycho-graphics, RFM
- Target = response to a past or test solicitation
- Action = target high-responding segments of customers in future campaigns

Fraud Detection

- Cases = past transaction or claims
- Inputs = particulars and circumstances
- Target = fraud, abuse, deception
- Action = impede or investigate suspicious cases

Attrition Prediction/Defection Detection

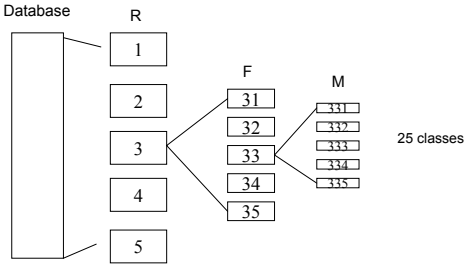
- Cases = existing customers
- Inputs = payment history, product/service usage, demographics
- Target = churn, brand-switching, cancellation, defection
- Action = Customer loyalty promotion

Required Expertise

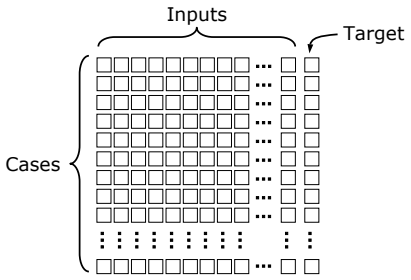
- Domain
- Data
- Analytical Methods

RFM

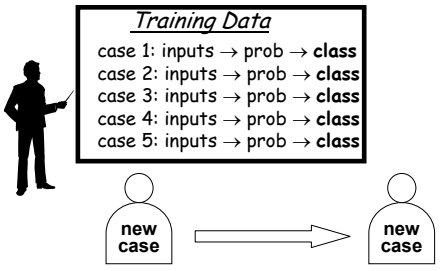
Recently, **Frequency**, **M**onetary Analysis



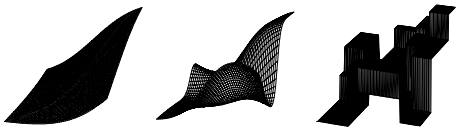
Predictive Modelling



Supervised Classification

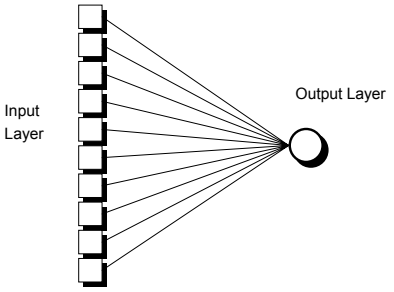


Modelling Methods



Generalized Linear Models | Neural Networks | Decision Trees

Generalised Linear Models



Stepwise Regression Analysis

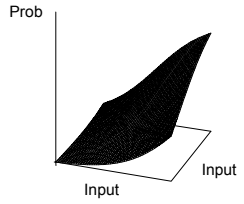
1. **Repo** = α_0 + α_1 **income**
2. **Repo** = β_0 + β_1 **income** + β_2 **age**
3. **Repo** = γ_0 + γ_1 **income** + γ_2 **age** + γ_3 **phone**
4. **Repo** = δ_0 + δ_1 **age** + δ_2 **phone**
5. **Repo** = ε_0 + ε_1 **age** + ε_2 **phone** + ε_3 **%down**

⋮

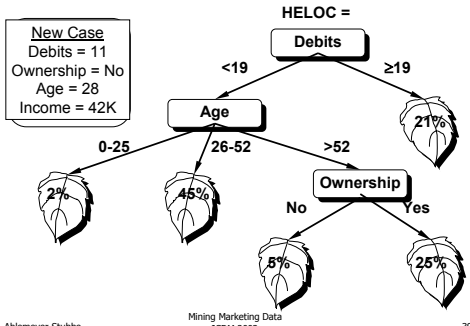
Note: The coefficients are confidential

Logistic Regression/Discrimination

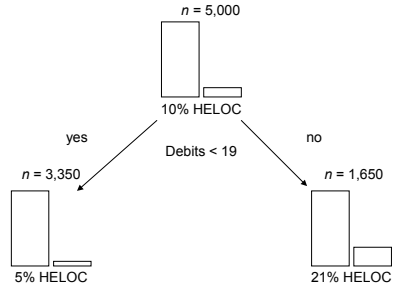
logit(probability of target event) = linear combination of the inputs



Fitted Decision Tree



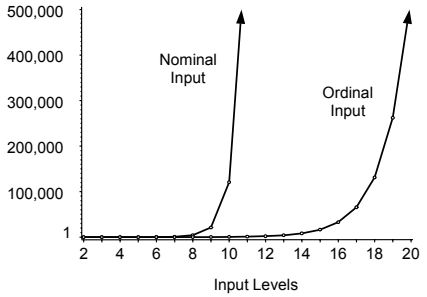
Divide and Conquer



The Cultivation of Trees

- Split Search
  - Which splits are to be considered?
- Splitting Criterion
  - Which split is best?
- Stopping Rule
  - When should the splitting stop?
- Pruning Rule
  - Should some branches be lopped-off?

Possible Splits to Consider



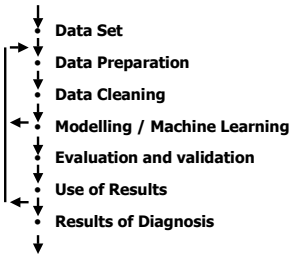
The Scope of Generalisation

- Model Selection and Comparison
  - Which model gives the best prediction?
- Decision/Allocation Rule
  - What actions should be taken on new cases?
- Deployment
  - How can the predictions be applied to new cases?

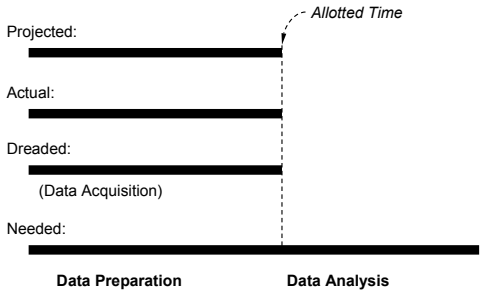
The O^{th} Law of Data Mining

The greater the similarity between the training Data set and the Data set on which you want to make a prediction, the better the prediction will be.

Data Mining Process



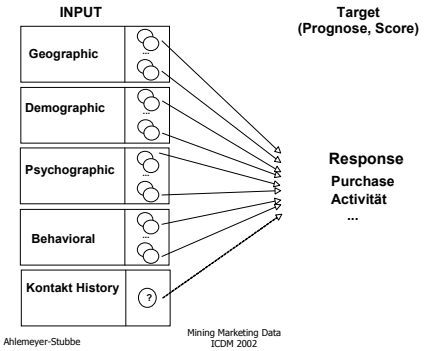
Time Line



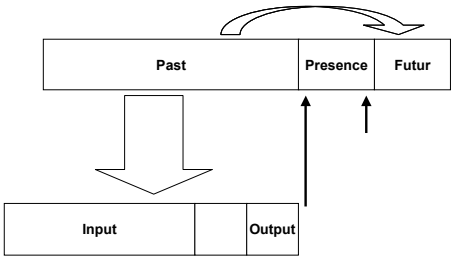
Data Preparation

- Build Data Mining Database
- explore Data
- prepare Data for modelling

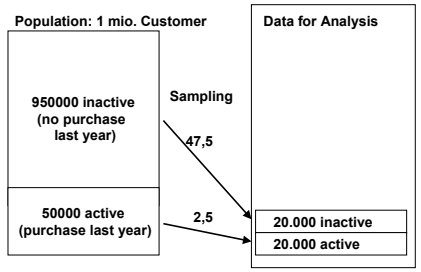
60% to 95% of the time is spent preparing the Data



Timescale

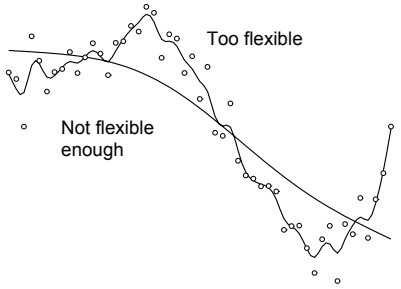


Sampling Example

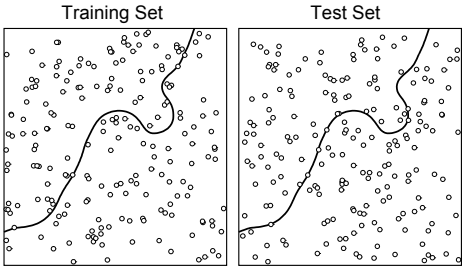


Data Splitting

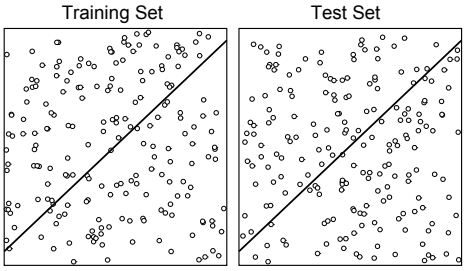
Model Complexity



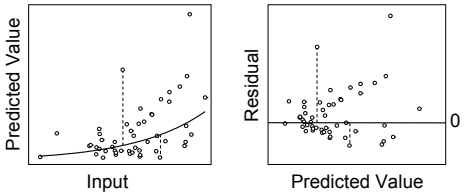
Overfitting



Better Fitting



Residuals



Confusion Matrix

Actual Class	Predicted Class 0	Predicted Class 1
0	**True Negative**	**False Positive**
1	**False Negative**	**True Positive**

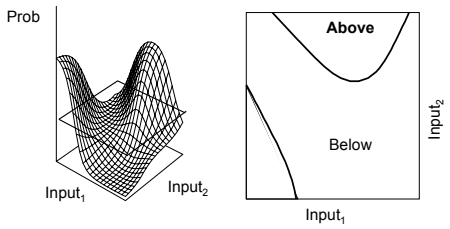
Examples: Modell 1

	Prognose		Validation	
Segment	**Index**	**Index kum.**	**Index**	**Index kum.**
1	246	246	213	213
2	168	207	157	185
3	123	179	112	161
4	101	159	107	147
5	84	144	90	136
6	73	132	79	126
7	67	123	76	119
8	56	115	62	112
9	47	107	54	105
10	36	100	51	100

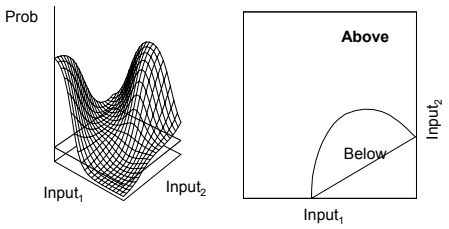
Examples: Modell 2

	Prognose		Validation	
Segment	**Index**	**Index kum.**	**Index**	**Index kum.**
1	234	234	228	228
2	162	198	156	192
3	117	171	115	166
4	106	155	108	152
5	89	142	91	139
6	78	131	80	130
7	67	122	69	121
8	58	114	61	113
9	50	107	51	107
10	39	100	41	100

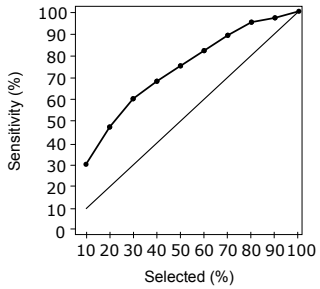
Cut-off Probability



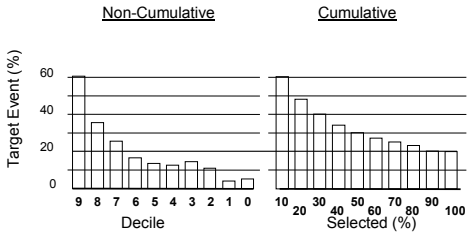
Different Cut-offs



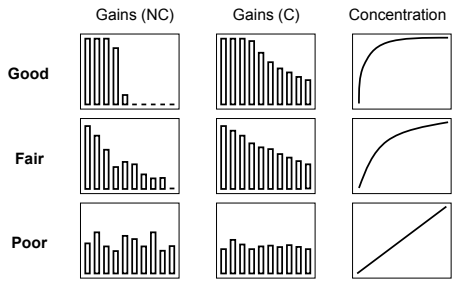
Target Concentration Curve



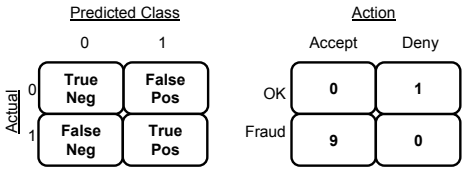
Gains Chart



Predictive Power



Misclassification Costs



Successful Data Mining

The keys to success:

- know the domain
- Formulating the problem
- using the right Data
- flexibility in modelling
- acting on results

Success depends more on the way you mine the Data rather than the specific tool.

Thank You

Andrea Ahlemeyer-Stubbe
DataBase Management
Hauptstr.34
D-77723 Gengenbach
Tel: + 49 (0) 7803 92 93 59
Fax: + 49 (0) 7803 92 93 60
E-mail:ahlemeyer@ahlemeyer-stubbe.de

Member of
Pro – ENBIS www.enbis.org/pro-enbis
ENBIS www.enbis.org
Branta www.branat.de

ECDM
EUROPEAN CENTRE OF DATABASE MARKETING