Dynamic Prediction of Online Purchases

Paper 171

Dimitris Bertsimas
Adam J. Mersereau

July 2002

Dynamic Prediction of Online Purchases

Dimitris Bertsimas Adam J. Mersereau

Research supported by E-business Center and Dell Corp

1 User Example

1.1 Visit 1

SLIDE 1

Visit 1: User starts in homepage, views some product info, then bookmarks configuration page for later. Web log records:

	URL	Explanation
1	www.dell.com/us/en/dhs/default.htm	home page
2	.../topics/segtopic_free_cdrw.htm	CD-RW promotion
3	.../products/series_dimen_desktops.htm	desktop info
4	.../products/spec_dimen_4100_desktops.htm	desktop specs
5	.../choose_dim_4100.htm	selection
6	commerce.dell.com/...config.asp	configuration

1.2 Visit 2

SLIDE 2

Visit 2: User resumes configuring a computer, saves the results as a shopping cart.

	URL	Explanation
1	commerce.dell.com/...config.asp	configuration
2	commerce.dell.com/...config.asp	configuration
3	www.dell.com/choose_dim_4100.htm	selection
4	commerce.dell.com/...config.asp	configuration
5	commerce.dell.com/...config.asp	configuration
6	commerce.dell.com/...basket.asp	basket
7	commerce.dell.com/...svcart_save.asp	shopping cart

1.3 Visit 3

SLIDE 3

Visit 3: User finalizes configuration, makes purchase.

	URL	Explanation
1	commerce.dell.com/...svcart_access.asp	shopping cart
2	commerce.dell.com/...config.asp	configuration
3	commerce.dell.com/...basket.asp	basket
4	commerce.dell.com/...chkout1.asp	checkout
5	commerce.dell.com/...chkout2.asp	checkout
6	commerce.dell.com/...chkout3.asp	checkout
7	commerce.dell.com/...chkout4.asp	checkout
8	commerce.dell.com/...chkout5.asp	checkout
9	commerce.dell.com/...chkout6.asp	checkout

2 Data available

SLIDE 4

- URLs from mixture of buyers and non-buyers from all sections of the web site over a 2-month period. 7,000 unique URLs.

- Users were identified using cookie information, and then separated into visits whenever there is a 30-minute break in activity.
- 30,000 users, 200,000 visits, and 2,000,000 clicks.
- Training set (50%), test set (30%), and validation set (20%).

3 Research Objectives

SLIDE 5

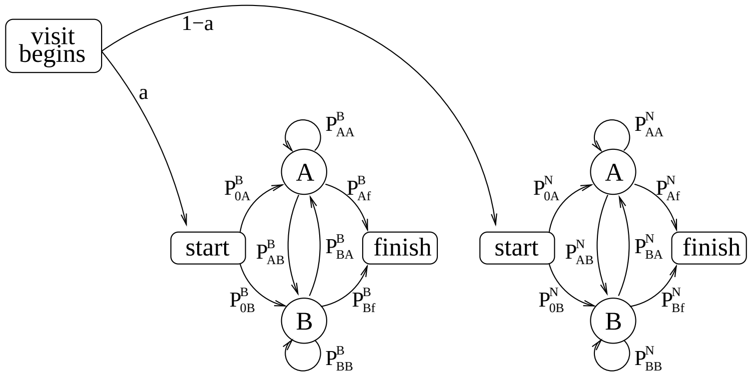
- **Dynamic Prediction**
After each click of a visitor of Dell's website, we want to *dynamically* compute the probability that the visitor will become a buyer.
- **Dynamic Classification**
Having computed the probability that a visitor will become buyer, we want to classify the customer as a "buyer", "non-buyer" or "not able to make a determination."
- **Clustering**
We want to know why and how users are using the website.
- **Dynamic Control**
Having classified the visitor as a "buyer" or a "non-buyer", we want to dynamically change the website, potentially offer a promotion or email.

4 Content Classification

SLIDE 6

- **Content Classification:** Summarize URL data using a small number of categories.
- Why?
  - Explosive number of possible URLs on Dell site.
  - Algorithmic advantages to reducing the space of input data.
  - Allows human visualization of sequences.
- We choose a simple approach based on finding keystrings in the URL addresses.

SLIDE 7



Category	Explanation
A	Homepages
B	Software and Accessories
C	"Outlet" Shopping
D	Special Offers
E	General Product Information
F	Specific Product Information
G	Choose and Configure
H	Cart and Quote
I	Checkout
J	Order Status
K	Miscellaneous & Unrecognized

4.1 Example Visits

Slide 8

- Example visits can be expressed compactly as strings.
  - Visit 1: "ABDDFG"
    "www.dell.com/us/en/dhs/default.htm" → A
    "www.dell.com/segtopic_free_cdrw.htm" → B
    ⋮
    "commerce.dell.com/dellstore/config.asp" → G
  - Visit 2: "GGFGGGH"
  - Visit 3: "HGHIIIIII"

5 Dynamic Prediction

5.1 First-order Markov model

SLIDE 9

5.2 Related Markov models

SLIDE 10

- Zeroth-order Markov model: each click is independent from the past.
- Second-order Markov model: category of the next page request in a visit would depend on the previous two page requests.
- Models with additional information: parameters depend on: the amount of time sincelast visit, number of previous visits, frequency of past purchases, the interclick time.

5.3 Prediction

SLIDE 11

- Fitting model parameters amounts to counting transitions in a training data set.
- To classify a new sequence "AAB," use Bayes' Rule:

$$Pr\{\text{Buy}|\text{“AAB”}\} =$$

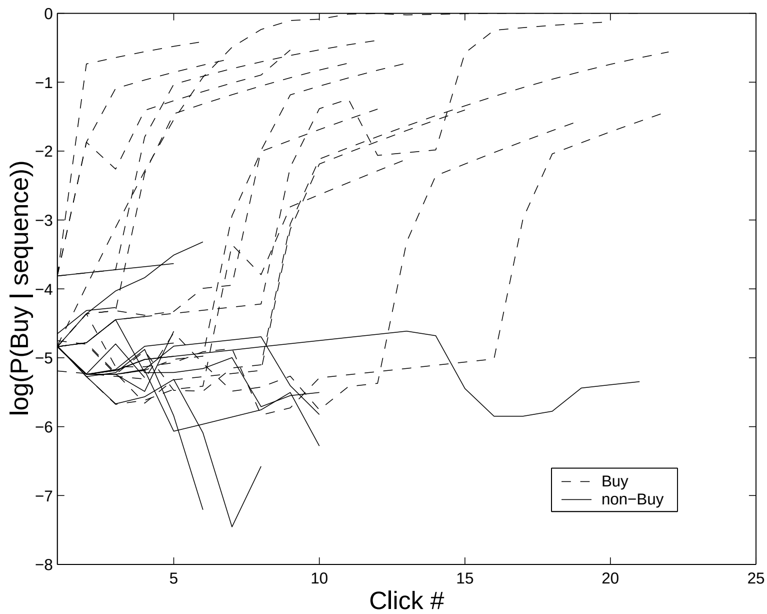
$$\frac{Pr\{\text{“AAB”}|\text{Buy}\}Pr\{\text{Buy}\}}{Pr\{\text{“AAB”}|\text{Buy}\}Pr\{\text{Buy}\} + Pr\{\text{“AAB”}|\text{non-Buy}\}Pr\{\text{non-Buy}\}}$$

  - $Pr\{\text{Buy}\}$ and $Pr\{\text{non-Buy}\}$ estimated apriori.
  - $Pr\{\text{“AAB”}|\text{non-Buy}\} = p(sA)p(AA)p(AB)$
  - $Pr\{\text{“AAB”}|\text{Buy}\} = q(sA)q(AA)q(AB)$

5.4 Out-of-Sample Data

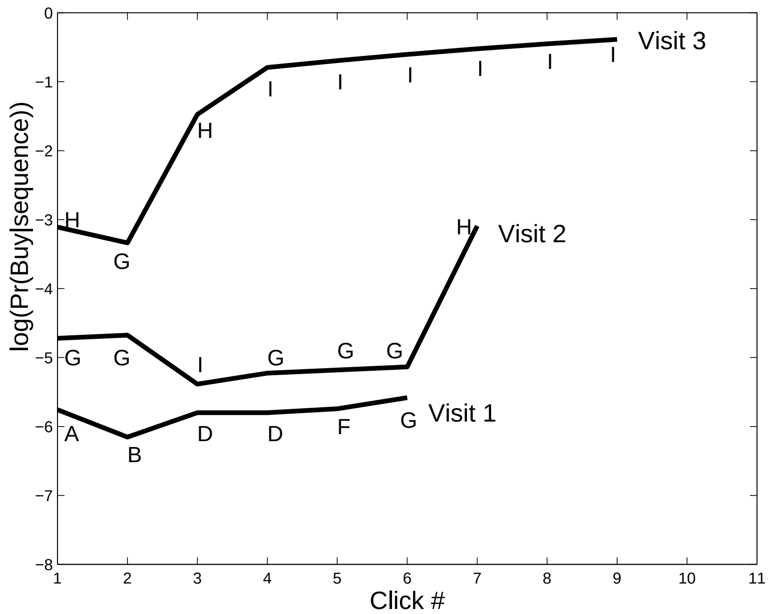
SLIDE 12

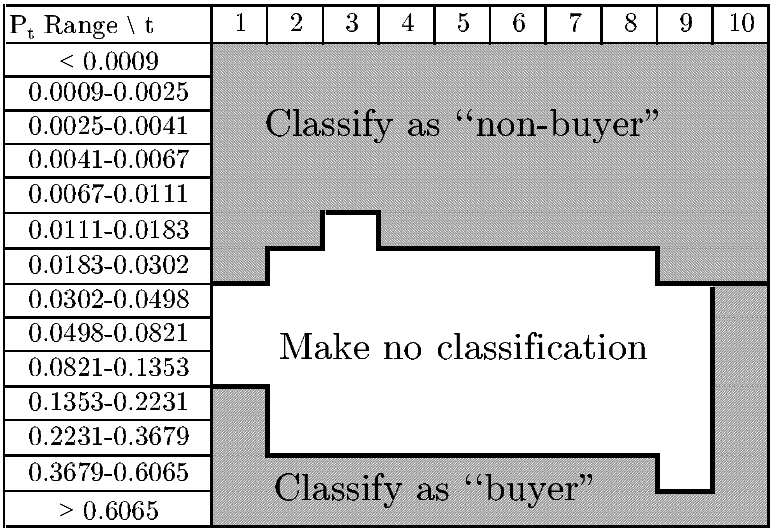
Results for 15 "Buy" visits and 15 "Non-Buy" visits:



5.5 Example

Slide 13





6 Dynamic classification

SLIDE 14

- Immediately after each click, our dynamic prediction models provide a probability that a visit will result in a "buy."
- Intuitively, if this probability is high, then we will classify the visit as a "buy", if it is low, then we will classify the visit as a "non-buy", and if it has intermediate values, we will not classify it.
- Use of dynamic optimization methods to find these dynamic thresholds.

SLIDE 15

6.1 Algorithms

SLIDE 16

- Zero-th order MC
- First order-MC
- Second-order MC
- Covariate: MC using the number of clicks in "Choose and Configure" observed in the previous visit.
- Naive: predict purchase the first time we observe a click in the "Checkout" section of the site. Otherwise, predict non-purchase at the 40th click

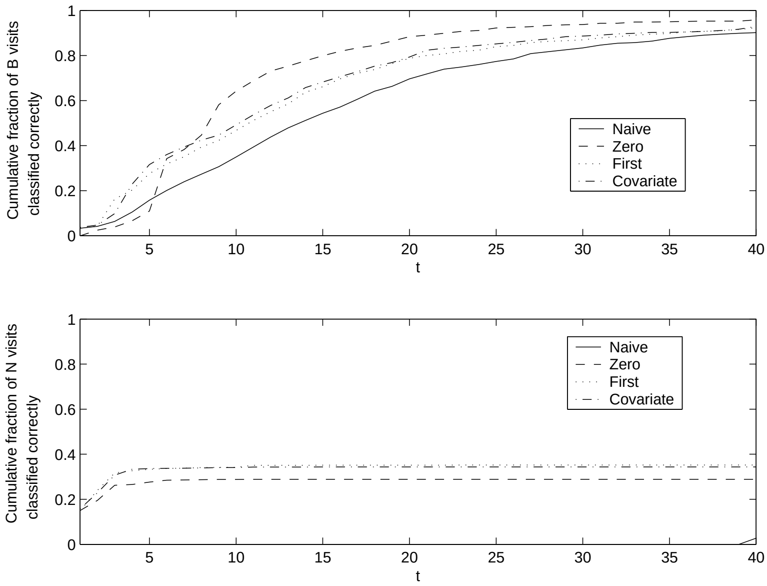
6.2 Results

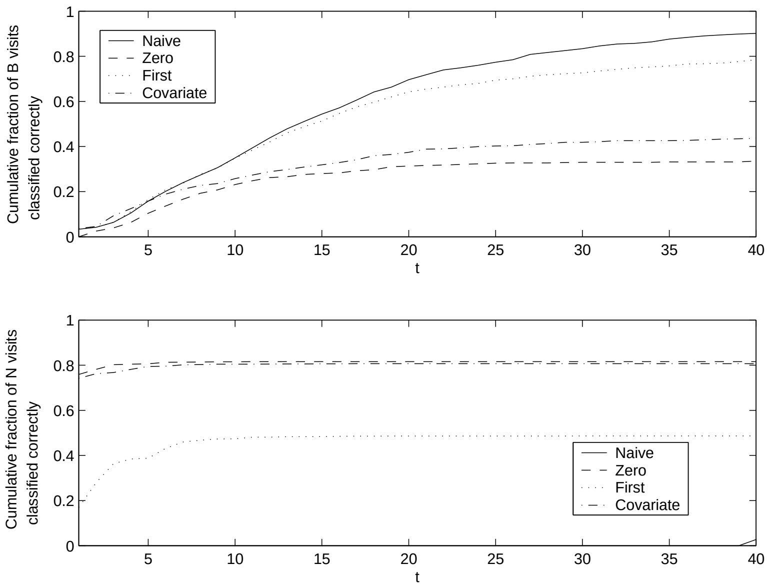
SLIDE 17

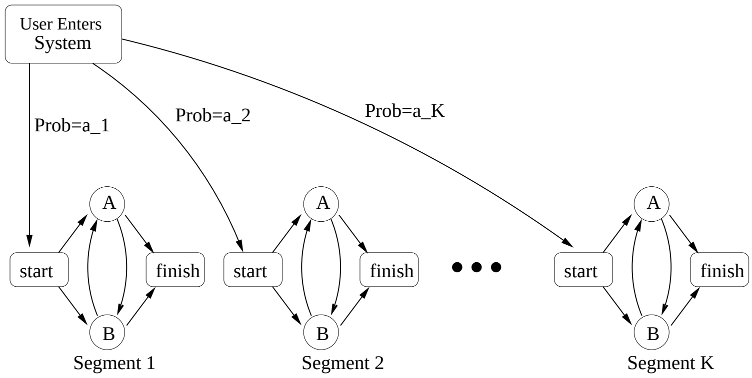
6.3 Insights

SLIDE 18
SLIDE 19

- The performance of the various Markov chain-based methods is comparable. Stronger than naive.


Cumulative fraction of B visits classified correctly
Cumulative fraction of N visits classified correctly
t
Naive
Zero
First
Covariate


Cumulative fraction of B visits classified correctly
Cumulative fraction of N visits classified correctly
t
Naive
Zero
First
Covariate



- Our proposed models can be tuned to detect more accurately buyers or non-buyers.
- Adding additional information did not improve the accuracy of the models.
- All methods can be implemented in a real time environment.

7 Clustering

SLIDE 20

SLIDE 21

- Users are clustered into a small number (3-5) distinct behaviors.
- Choose parameters of model such that the likelihood of the model is maximized.
- EM algorithm used

7.1 Results

SLIDE 22

8 Conclusions

SLIDE 23

- Dynamic prediction methods
- Dynamic Classification methods
- Clustering with interpretable results.

9 Future work

SLIDE 24

- Dynamic reconfiguration and personalization of the website in order to maximize the probability of purchase.
- Experimental design

Probability (a)	%A	%B	%C	%D	%E	Interpretation
$K = 1$						
1.00	0.14	0.12	0.38	0.24	0.11	
$K = 2$						
0.57	0.14	0.11	0.44	0.28	0.03	
0.43	0.15	0.14	0.33	0.20	0.19	
$K = 3$						
0.49	0.19	0.14	0.53	0.09	0.05	Configure
0.26	0.10	0.09	0.33	0.44	0.04	Checkout
0.25	0.13	0.15	0.30	0.19	0.24	Misc. clickers
$K = 4$						
0.48	0.19	0.14	0.54	0.09	0.05	Configure
0.26	0.10	0.09	0.33	0.44	0.04	Checkout
0.25	0.12	0.15	0.30	0.18	0.24	Misc. clickers
0.01	0.52	0.07	0.19	0.14	0.08	Home clickers
$K = 5$						
0.32	0.18	0.14	0.50	0.10	0.08	Configure
0.31	0.14	0.11	0.44	0.29	0.02	Config/Checkout
0.19	0.10	0.10	0.31	0.40	0.10	Checkout
0.18	0.14	0.18	0.31	0.09	0.29	Misc. clickers
0.01	0.53	0.07	0.19	0.13	0.08	Home clickers

- Stochastic optimization to balance the tradeoff between experimenting with the website in order to learn the customers better and maximizing short term purchases.