Five Ways to Lower Costs with Predictive Analytics
Instructor: Eric Siegel, Ph.D.

Program Chair: Predictive Analytics World

President: Prediction Impact, Inc.

Instructor: predictive analytics training

Former computer science professor: Columbia University

Cofounded software companies: User profiling and data mining
Risk Management

Before:
After:

Risk Management
Response Modeling

Who would have purchased anyway?
Agenda

- What is predictive analytics?
- How does it provide value?
- 5 ways to lower costs with predictive analytics
- Hot methods in predictive modeling
Predictive Analytics:

Business intelligence technology that produces a predictive score for each customer or prospect

© Copyright 2009. All Rights Reserved.
Predictive Analytics:

Your organization learns from its collective experience.
Many Small Decisions Add Up

IMPACT OF INDIVIDUAL DECISION

High value

Strategic

Low value

Tactical

Operational

DECISION VOLUME

Low volume
High volume

Decision automation via predictive analytics

From Smart (Enough) Systems by J. Taylor and N. Raden
Business Applications of Predictive Analytics

- Response modeling
- Customer retention with churn modeling
- Product recommendations
- Behavior-based advertising
- Application processing
- Insurance pricing
- Credit scoring
- Leads scoring
- Forecasting
- Fraud detection

Tracks are by industry, but “look at the app.”
Success Stories

- Aflac
- Amway
- The Coca-Cola Company
- Citizens Bank
- Financial Times
- Hewlett-Packard
- Infinity Insurance
- IRS
- Lifeline Screening
- The NRA
- The New York Times

- Optus (Australian telecom)
- PREMIER Bankcard
- Reed Elsevier
- Sprint-Nextel
- Sunrise Communications (Switzerland)
- Target
- US Bank
- U.S. Department of Defense
- Walden University
- Zurich

Plus: Busch, Disney, HP, HSBC, Pfizer

useR Meeting: DC R Users Group
7:30pm-10:00pm
Magnolia room
## Verticals

- Banking
- Financial services
- E-commerce
- Education
- Healthcare
- Consumer services
- Government
- High technology
- Insurance
- Non-profit
- Publishing
- Retail
- Telecom
Objectives

- Strengthen the business impact delivered by predictive analytics
- Establish new opportunities with predictive analytics

- 30% Managers and Directors - Analytics & BI
- 25% Managers and Directors - Other
- 16% CXOs, Presidents, VPs, Principals
- 14% Analysts
- 12% Practitioners, statisticians, scientists
If you would be wealthy, think of saving as well as getting.

Money is better than poverty, if only for financial reasons.
Lowering Costs with Predictive Analytics

1. Response modeling
2. Response uplift modeling
3. Churn modeling
4. Churn uplift modeling
5. Risk modeling

COST CUTTERS:
Lower operational spend

Brand name anecdotes
1. Response Modeling
Lifeline Screening: Response up 38%, cost down 20%, 62k more customers annually

PREMIER Bankcard: Direct mail response up 3-5%

COST CUTTER #1: Don't contact those who won't respond.
1A. Content Selection

Select the message, product, creative the customer's most likely to respond to – “Dynamic AB Selection”

Target:
15-20% lift over traditional direct mail targeting

US Bank:
Costs down 40%, lift up 2 times, and cross-sell ROI up 5 times

COST CUTTER #1A:
Don't send “A” if the customer's more likely to respond to “B”.

A ➔?
B ➔?
C ➔?
1B. Improve Paid Search Ads

Predict ad bounce rate, based on:

- Ad text
- Landing page

For more:

COST CUTTER #1B:
Don't pay for clicks that will bounce.

Google:
Improved bounce rate prediction 5-10%
Beyond Response Modeling

Who would have purchased anyway?
Uplift Modeling

Some customers will buy anyway.

A.K.A.:  
- Net lift modeling  
- Incremental modeling  
- Impact modeling  
- Differential response modeling

Analytical methods:  
*Model on the residual*  
*Dual recursive partitioning*
2. Response Uplift Modeling

<table>
<thead>
<tr>
<th>Buy if do receive an offer</th>
<th>Do-Not-Disturbs</th>
<th>Lost Causes</th>
<th>Persuadables</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>Sure Things</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>Lost Causes</td>
<td>Persuadables</td>
<td></td>
</tr>
</tbody>
</table>

**COST CUTTER #2:** Don't contact those who'd respond anyway.

**Leading financial institution:**
- Incremental conversion up 0.02% to 0.43%
- Revenue per contact up by over 20 times

**Target:** Challenges in uplift modeling
Schrödinger's cat
3. Retention with Churn Modeling

“A customer saved is a customer earned”

Acquisition: New customers

Customer base

Defection: Lost customers

COST CUTTER #3: Don't waste expensive retention offers on those who’ll stay anyway.

PREMIER Bankcard: $8 million est. retained

Reed Elsevier's Caterer & Hotelkeeper
Reduced churn by 16%
Retention ROI up by 10%

Optus (Australian telecom):
Doubled churn model performance with social data

Orange Labs (France Telecom):
Rapid scoring
Churn Baby Churn:

Fast Scoring on Large Telecom Dataset

- What's lift got to do with it?
- Bad, Bad Lee-ROI Brown
- 99 Luft Models
- I am the Very Model of a Modern Major General Electric
- Data Mining... Can You Dig It?
From FORRESTER whitepaper: *Optimizing Customer Retention Programs*, by Suresh Vittal

"Leave well enough alone."
"If it ain't broke, don't fix it."
"Do not disturb!"
"Let sleeping dogs lie."
COST CUTTER #4:
Don't trigger those who'd otherwise stay.

Telenor:
Reduced churn 36%
Cost-of-contact down 40%
Campaign ROI up 11-fold

US Bank:
Costs down 40%, lift up 2 times, and cross-sell ROI up 5 times
How to Get Fired from Harrah’s Casino:

1. Stealing
2. Sexual harassment
3. Failure to use a control set
Control Group

Customers:

Deployment

1. Prove value

2. Monitor value

3. Improve value

Control group
5. Risk Management

Insurance:

Don't charge too little for high-risk applicants.

Credit:

Don't give credit to a likely defaulter.

Zurich Insurance:
Loss ratio down 1/2 point
Savings of almost $50 mill

COST CUTTER #5:
Don't acquire “loss customers”.
6. Leads Scoring

High Tech

Sun Microsystems
Doubled number of leads per phone call

Consumer

Bella Pictures
Scores the brides

Walden University
Scores the students

COST CUTTER #6:
Don't expend sales resources on leads that won’t convert.

Every bride needs a mail solicitor.
7. Fraud Detection

Transactions include:
- Invoices
- Credit card purchases
- Online activities
- Tax returns
- Insurance claims
- Telecom (calls)
- Checks
- Clicks (paid ads)

COST CUTTER #7A: Don't let thieves get away with it.

COST CUTTER #7B: Expend fewer staff resources on fraud

Citizens Bank: Loss prevention up 20% - or - Prevention staff down 30%

Defense Finance and Accounting Service: Detected 97% of known fraud cases
Application processing: benefits, jobs, insurance, ...

8. Don't spend as much time on likely denials.

Forecasting

9. Don't acquire/produce too many.

Human resources

10. Don't hire unreliable employees.
Earn More Money

"A penny earned is a penny earned"

a) Increase response  
b) Acquire better customers  
c) Retain more  
d) Cross-sell  
e) More clicks  
f) Charge more  
g) Collections

h) Improve lift on current deployments
Ensemble Models

Bagging, boosting, bootstrapping, random forests, meta-learning, blending

Doesn’t this violate K.I.S.S.?

Somehow, they’re more robust!
Dean Abbott: How to Improve Customer Acquisition Models with Ensembles

Case study: Non-profit organization

National Rifle Association
Netflix Prize: Product Recommendations

BellKor's Pragmatic Chaos

Grand Prize

Meta-learning

BellKor in BigChaos

Progress Prize

Meta-learning

PragmaticTheory

BellKor

BigChaos

© Copyright 2009. All Rights Reserved.
Netflix Prize

Second Place (close!) - GPT
Grand Prize Team

... 7 others...

VanDelay Industries

... 5 others...

© Copyright 2009. All Rights Reserved.
Lowering Costs with Predictive Analytics

Response modeling
A) Content selection
B) Ad bounce prediction
Response uplift modeling
Churn modeling
Churn uplift modeling
Risk modeling
Lead scoring
Fraud detection

Target campaigns
Send the right message
Pay for effective clicks
Suppress “sure things”
Target retention
Let sleeping dogs lie
Avert risk
Target sales resources
Avert theft
## Business Applications of Predictive Analytics

<table>
<thead>
<tr>
<th>Business application</th>
<th>What is predicted:</th>
<th>Driven business decision:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Direct marketing</td>
<td>customer response</td>
<td>Expend cost of contact?</td>
</tr>
<tr>
<td>Customer retention</td>
<td>customer churn/attrition</td>
<td>Expend retention offer?</td>
</tr>
<tr>
<td>Product recommendations</td>
<td>what each customer wants</td>
<td>Which product to recommend?</td>
</tr>
<tr>
<td>Content/ad selection</td>
<td>which one they’ll respond to</td>
<td>Which content to display?</td>
</tr>
<tr>
<td>Lifetime Value</td>
<td>how profitable will be</td>
<td>How much to invest?</td>
</tr>
<tr>
<td>Fraud detection</td>
<td>whether transaction is legit</td>
<td>Whether to investigate?</td>
</tr>
<tr>
<td>Credit scoring</td>
<td>debtor risk</td>
<td>Whether to grant loan/credit</td>
</tr>
<tr>
<td>Fundraising for nonprofits</td>
<td>donation amount</td>
<td>Expend cost of solicitation?</td>
</tr>
<tr>
<td>Application processing</td>
<td>application approval</td>
<td>Whether to fast-track/deny?</td>
</tr>
<tr>
<td>Leads scoring</td>
<td>conversion</td>
<td>Expend sales resources/time?</td>
</tr>
<tr>
<td>Insurance pricing &amp; selection</td>
<td>insured risk</td>
<td>Whether to cover/how to price?</td>
</tr>
<tr>
<td>Forecasting</td>
<td>aggregate profit</td>
<td>(Strategic decisions)</td>
</tr>
</tbody>
</table>
## Business Applications of Predictive Analytics

<table>
<thead>
<tr>
<th>Business application</th>
<th>What is predicted:</th>
<th>Case studies PAW Oct-09</th>
</tr>
</thead>
<tbody>
<tr>
<td>Direct marketing</td>
<td>customer response</td>
<td>Lifeline Screening, NRA, PREMIER Bankcard</td>
</tr>
<tr>
<td>Customer retention</td>
<td>customer churn/attrition</td>
<td>Aflac, Optus (telecom), Orange, PREMIER, Reed Elsevier</td>
</tr>
<tr>
<td>Product recommendations</td>
<td>what each customer wants</td>
<td>US Bank</td>
</tr>
<tr>
<td>Content/ad selection</td>
<td>which one they’ll respond to</td>
<td>Target</td>
</tr>
<tr>
<td>Lifetime Value</td>
<td>how profitable will be</td>
<td>Amway</td>
</tr>
<tr>
<td>Fraud detection</td>
<td>whether transaction is legit</td>
<td>Citizens Bank, DoD, HP, IRS</td>
</tr>
<tr>
<td>Credit scoring</td>
<td>debtor risk</td>
<td>Infinity Insurance</td>
</tr>
<tr>
<td>Fundraising for nonprofits</td>
<td>donation amount</td>
<td>Walden University</td>
</tr>
<tr>
<td>Application processing</td>
<td>application approval</td>
<td>Zurich Insurance</td>
</tr>
<tr>
<td>Leads scoring</td>
<td>conversion</td>
<td>Coke, Financial Times, New York Times, Sprint-Nextel</td>
</tr>
<tr>
<td>Insurance pricing &amp; selection</td>
<td>insured risk</td>
<td></td>
</tr>
<tr>
<td>Forecasting</td>
<td>aggregate profit</td>
<td></td>
</tr>
</tbody>
</table>
Matthew Finlay
Paul Gillis
Fanny Jeanmougin
Take-Aways

- Reduce costs
- Control set
- Uplift modeling
- Ensemble models

- Social data
- Text & speech analytics
Predictive Analytics World Conference

February 16-17, 2010: San Francisco

Bigger wins!
Strengthen the business impact delivered by predictive analytics.

"Predictive Analytics World was probably the best analytics conference I have attended... turned into my new must-go-to conference."
Dennis Mortensen
Director of Data Insights
Yahoo!

February 2009 drew 200 attendees from 13 countries, and included case studies from: Acxiom, Amazon.com, Bella Pictures, Charles Schwab, ClickForensics, Google, The National Rifle Association, Pinnacol Assurance, Reed Elsevier, Sun Microsystems, TaxBrain, Telenor, Wells Fargo, Yahoo! and several more.

www.pawcon.com
Training program:
Predictive Analytics for Business, Marketing and Web

November 11-12, 2009
San Francisco, CA

businessprediction.com

98% of attendees since Oct. 2008 rated this course Excellent or Very Good
**Predictive Analytics and Data Mining**

**Services:**
- Defining analytical goals & sourcing data
- Developing predictive models
- Designing and architecting solutions for model deployment
- "Quick hit" proof-of-concept pilot projects

**Training programs:**
- **Public seminars:** Two days, in San Francisco, Washington DC, and other locations
- **On-site training options:** Flexible, specialized
- **Instructor:** Eric Siegel, Ph.D., President, 16 years of data mining, experienced consultant, award-winning Columbia professor
- **Prior attendees:** Boeing, Corporate Express, Compass Bank, Hewlett-Packard, Liberty Mutual, Merck, MITRE, Monster.com, NASA, Qwest, SAS, U.S. Census Bureau, Yahoo!

*If you predict it, you own it.*
Predictive Analytics and Data Mining

Applications:
- Response modeling for direct marketing
- Product recommendations
- Dynamic content, email and ad selection
- Customer retention
- Strategic segmentation
- Security
  - Fraud discovery
  - Intrusion detection
  - Risk mitigation
  - Malicious user behavior identification
- Cutting-edge research for groundbreaking data mining initiatives

Verticals:
- Online business: Social networks, entertainment, retail, dating, job hunting
- Telecommunications
- Financial organizations
- A fortune 100 technology company
- Non-profits
- High-tech startups
- Direct marketing, catalogue retail

www.PredictionImpact.com

If you predict it, you own it.
Predictive Analytics and Data Mining

Team of several senior consultants:
- Experts in predictive modeling for business and marketing
- Relevant graduate-level degrees
- Communication in business terms
- Complementary analytical specialties and client verticals
- Published in research journals and industrials

Extended network of many more:
- Closely collaborating partner firms
- East coast coverage

Eric Siegel, Ph.D., President
Prediction Impact, Inc.
San Francisco, California
eric@predictionimpact.com
(415) 683 - 1146

For our bi-annual newsletter, click “subscribe”:
www.PredictionImpact.com

To receive notifications of training seminars:
training@predictionimpact.com

If you predict it, you own it.
ADDENDUM

Misc. reference slides follow
Wisdom Gained: A Predictive Model is Built

- Customer profiles
- Customer behavior
- Customer contact logs

Modeling tool:
- Performed by modeling software
- Usually offline

Predictive model
Applying a Model to Score a Customer

- Customer profile
- Customer behavior
- Predictive model
  - Often performed by modeling software
  - Possibly online
- Predicted response
Prediction Drive Operational Decisions

Predicted response

Business logic

Business action
- Solicitation
- Cross-sell
- Retention offer