Predictive Analytics Applied: Marketing and Web

Brought to you by Prediction Impact and World Organization of Webmasters (WOW)

Eric Siegel, Ph.D.
Prediction Impact, Inc.
eric@predictionimpact.com
(415) 683 - 1146
1. Introduction
2. How Predictive Analytics Works
3. App: Customer Retention
4. App: Targeting Ads
5. Summary and take-aways
About the speaker: Eric Siegel, Ph.D.

– President of Prediction Impact, Inc.
– Training and services in predictive analytics
– Former computer science professor at Columbia University
– Before Prediction Impact, cofounded two software companies

Eric Siegel, Ph.D.
Prediction Impact, Inc.
eric@predictionimpact.com
(415) 683 - 1146
Business intelligence technology that produces a predictive score for each customer or prospect
“At a time when companies in many industries offer similar products and use comparable technology, high-performance business processes are among the last remaining points of differentiation.”

- Tom Davenport
Data is a core strategic asset – it encodes your business’ collective experience.

It is imperative to:

Learn *from your data*.

Learn *as much as possible about your customers*.

Learn *how to treat each customer individually*. 
Predictive Analytics:

Your organization learns from its collective experience

*Collective experience:* Sales records, customer profiles, …

*Learn:* Discover strategic insights and business intelligence
  Discover something that makes business decisions automatically
  Discover business rules

…and puts this knowledge to action.
How Predictive Analytics Works

Building andDeploying Predictive Models
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
- Predicted response

- Often performed by modeling software
- Possibly online
Predicted response

Business logic

- Usually not performed by modeling software
- Possibly online

Business actions, such as:
- Mail a solicitation
- Suggest a cross-sell option
- Retain with a promotion
Predictive models are created automatically \textit{from your data}

So, they’re generated according to your:
- Product
- Prediction goal
- Business model
- Customer base

This means customer intelligence \textit{specialized for your business}
- \textit{Customized business rules}
- \textit{Unique, proprietary mailing lists}
- \textit{Insights only your organization could possibly gain}
Predictors: Building Blocks for Models

• *recency* – How recent was the last purchase?
• *personal income* – How much to spend?

Combine predictors for better rankings:

\[
\text{recency} + \text{personal income}
\]

\[
2 \times \text{recency} + \text{personal income}
\]

- If the customer is rural, and her monthly usage is high, then the customer will probably renew.
- If the customer is urban, and new feature exploration is high, then the customer will probably *not* renew.
This is the required form for training data; one record per customer.

<table>
<thead>
<tr>
<th>Number purchases:</th>
<th>Last purchase:</th>
<th>Gender:</th>
<th>Income:</th>
<th>Likely to buy:</th>
</tr>
</thead>
<tbody>
<tr>
<td>7</td>
<td>shoes</td>
<td>M</td>
<td>high</td>
<td>gloves</td>
</tr>
<tr>
<td>3</td>
<td>gloves</td>
<td>F</td>
<td>medium</td>
<td>hat</td>
</tr>
<tr>
<td>1</td>
<td>hat</td>
<td>M</td>
<td>high</td>
<td>shoes</td>
</tr>
<tr>
<td>46</td>
<td>gloves</td>
<td>F</td>
<td>low</td>
<td>piano</td>
</tr>
</tbody>
</table>
Why Not Memorize The Training Examples?

To learn from history is to remember history.
So, we could just keep a lookup table and compare new cases to old ones.

**Answer:** There are too many possible rows of data.
That is, *each customer is unique!*
If they bought shoes and they’re not female then sell gloves.
If they didn’t buy shoes and they’re high income then sell shoes.
Lower campaign costs and increase response rates

- Make campaigns more targeted and more selective
- Identify segments five or more times as responsive
- Achieve 80% of responses with just 40% of mailing
- Increase campaign ROI & profitability
Predictive Models Score Each Customer

- Each customer is scored with a response probability
- The customers are listed in order of prediction score
- The highest-scoring customers are targeted first

<table>
<thead>
<tr>
<th>ID number:</th>
<th>Name:</th>
<th>Score:</th>
<th>Response:</th>
</tr>
</thead>
<tbody>
<tr>
<td>429</td>
<td>E. Siegel</td>
<td>35%</td>
<td>Yes 😊</td>
</tr>
<tr>
<td>528</td>
<td>D. Leong</td>
<td>31%</td>
<td>No 😞</td>
</tr>
<tr>
<td>256</td>
<td>G. Clooney</td>
<td>22%</td>
<td>Yes 😊</td>
</tr>
<tr>
<td>674</td>
<td>T. Mitchel</td>
<td>14%</td>
<td>No 😞</td>
</tr>
</tbody>
</table>
Profit vs. Percent of Customers Contacted

- Customers ranked with predictive analytics
- Customers not ranked
Lift Chart

9.3 → 30.6
16.4 → 43.5
20.3 → 51.1
32.4 → 70.1
44.0 → 80.2

• 50,000 customers 3 times as likely to buy
• 200,000 customers 2 times as likely to buy
Example Business Rule

New customers who come to the website off organic search results, buy more than $150 on their first transaction, are male, and have an email address that ends with “.net” are three times as likely to be return customers.
App: Customer Retention
Retain Customers by Predicting Who Will Defect
Which is the best way to increase growth?
1. Increase acquisition
2. Decrease defection
Increase retention by 3% and boost growth by 12%
Predictor Variables

- Age of membership
- External acquisition source
- U.S. state of residence
- Willing to receive postal mailing (opt-in)
- Num days since last failed login attempt
- Num days between logins
At-Risk Segment

- SOURCE_COBRAND$ == chat
- SUBSCR_AGE <= 237.819
- LOGIN_DAYSSINCELAST_F > 1.85344
- LOGIN_DAYSSINCELAST_F <= 22.6578

Churn rate: **33.5%** (vs. 20% average)
Loyal Segment

- STATE is one of many, including ME, NE, NH, NS, QC, SK, WY
- SOURCE_COBRAND$ == chat
- SUBSCR_AGE > 237.819
- TIME_SINCE_JOINED <= 535.456
- LOGIN_DAYSSINCEFIRST_F > 99.6539
- LOGIN_DAYSSINCEFIRST_F > 112.003
- LOGIN_DAYSSINCELAST_S > 131.45

Churn rate: **6.5%** (vs. 20% average)

Very loyal

Small segment
Forecasted Profit of Retention Campaign

Cost per mailing: $25. ave profit: $100. num subscribers: 100,000
Loyal Segment: Online Retail

- ITEM_QTY > 2.5
- TAX_SHIP <= 8.78

- **19%** will return (versus 10% average)
App: Targeting Ads

Predictively Modeling Which Promotion Each Customer Will Accept

Eric Siegel, Ph.D.
Prediction Impact, Inc.
eric@predictionimpact.com
(415) 683 - 1146
Learn More From AB Testing: “Dynamic AB Selection”

Target content
- Landing page
- Featured product
- Color of product displayed
- Promotion or advertisement

Based on:
- Customer behavior profile
  - Browsers vs. hunters
- Time of day
- Geographical location
- Pages
- Where they came from
  - Ad that clicked them through
  - Site they came from
  - Search query they were on
  - Profile, when available
Dynamic AB Selection

Trials of A → Modeling tool → Predictive model for A → Predictive score: What’s the chance Mary responds to A?

Trials of B → Modeling tool → Predictive model for B → Predictive score: What’s the chance Mary responds to B?
Selecting Between 291 Sponsored Promotions

**Client:** A leading student grant and scholarship search service

Sponsors include:
- Universities
- Student loans
- Military recruitment
- Other misc.
The Business: Great Potential Gain

• High bounties, up to $25 per acceptance

• High acceptance rates, up to 5%, due to general relevancy of the promotions

• Big Data: *Wide and Long*
  – Rich user profiles volunteered for funds eligibility
  – Over *50 million* training cases: 01/05 - 09/05
Already seen? N/Y

Email opt-in Y/N

Region: S, NE, W, M-W

Fraternity/None/Sorority
Highly Responsive Segment for "Art Institute" Ad

Segment definition included:

- Still in high school
- Expected college graduation date in 2008 or earlier
- Certain military interest
- Never saw this ad before

Segment probability of accepting ad: 13.5%
Average overall probability of accepting ad: 2.7%
Highly Responsive Segment for a "Navy" Ad

Segment definition included:
• Has opted in for emails
• Has not seen this ad before
• Is in college
• SAT verb-to-math ratio is not too low, nor too high
• SAT written is over 480
• ACT score is over 15
• No high school name specified

Segment probability of accepting ad: 2.6%
Average overall probability of accepting ad: 1.6%
A-B test deployment to compare:
   A. Legacy system based on acceptance rates across users
   B. Model-based ad selection

Result:
25% increased acceptance rate
3.6% increased revenue observed; 5% later reported by the client

This comes to almost $1 million per year in additional revenue, given the existing $1.5 million monthly revenue.
Conclusions

Summary and take-aways
Predictive Analytics Initiatives

- **Per-customer predictions**: Unlimited range of business objectives
- **Management**: An organizational process ensures predictions are actionable, driven by business needs; multiple roles and skills
- **Deployment**: Mitigate risk and track performance
• **Data preparation:** An intensive bottleneck, critical to success

• **Modeling methods and tools:** No one method that is always best; compare multiple methods
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, 15 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!

www.PredictionImpact.com

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

If you predict it, you own it.

www.PredictionImpact.com
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.