

Engagement List:

Example Client Projects in Predictive Analytics

Eric Siegel, Prediction Impact
eric@PredictionImpact.com
(415) 385 - 1313

Personalized, Predictive Model-Based Selection of Online Advertisements.

The client is the leading student grant and scholarship search service -- 1 in 3 college-bound high school seniors use the service; more than 28 million students have created accounts. Revenue is generated solely when users opt into promotions, which belong to one of a relevant set of categories, including universities, military recruitment, and student loans. Since users volunteer detailed profiles in order to learn about financial aid opportunities for which they qualify, the data set available for analysis is very rich.

Prediction Impact lead the technical execution and scientific design of the pilot project to implement model-based advertisement selection. Predictive models were generated over **10's of millions of historical web transactions**, data that encodes which users were solicited with which offers, and whether the offer was accepted.

These ***models are used to select which promotional offer is most likely to be accepted by each user.*** With hundreds of programs between which to select, a scalable modeling infrastructure had to be designed and implemented from scratch; there is no industry standard solution to this kind of modeling problem.

Model performance evaluation showed ***promotion acceptance rates will increase by as much as 28%***, proving the theory that the "intelligent," data-driven selection of promotions based on each user's unique profile provides a generous boost in company revenue, and a substantial ROI on such a predictive analytics initiatives. Insights from this analysis also directly informed fundamental design changes to the client's product.

Predictive Analytics for Retention and Acquisition. The client is a media company generating a quickly growing revenue of \$30 million and hosting a leading online dating service. The number of international paying subscribers is roughly 145,000. Since the growth in subscribers results from a mere 20% margin over a high monthly attrition rate (as is expected for this subscription service), a relatively small increase in retention provides a healthy return. For example, a 5% decrease in attrition results in an estimated \$862,500 in customer lifetime values per year, plus a compound effect.

We employed predictive analytics to discover and assess the most effective metrics and indicators to **predict customer attrition**, as well as **customer acquisition** from unpaid to paid subscriptions (i.e., conversion). Both metrics that measure product usage trends, as well as demographic indicators, proved predictive of future customer attrition and conversion events. For example, customers who sign up with "Hotmail" and "Yahoo" email accounts are far less likely (20 - 25% as likely) to convert to a paid subscription than users with email addresses that may be more "permanent," such as ".net" or "EarthLink" email addresses. This could very well turn out to be an important predictor for many other online subscription-based services as well. The **insight speaks directly to business strategy**, such as employing incentives for customers to provide permanent email addresses, or partnering with certain email service providers. Resulting predictive models predicted customer events even more effectively by combining multiple indicators, **enabling tactical deployment**, such as, for example, a retention campaign strategy with a positive estimated ROI.

Acting in a more general business intelligence role, we provided several related services for this client:

- To support the above initiative in predictive analytics, as well as more general analytics goals, we **developed a marketing data mart** that serves as the foundation for a cohesive business intelligence system.
- We discovered a substantial, ongoing loss of subscriptions due to a credit card processing error, and advised as to a tiered business procedure that **may recover an estimated \$1 to 3 million in annual revenue**.
- We produced general marketing reports that analyze attrition and acquisition trends over various customer segments.
- We gathered reporting requirements across multiple departments, designed a customer attrition poll for marketing intelligence, evaluated analytics software and services for purchase, designed and advised as to additional marketing metrics moving forward, and established an ongoing financial reconciliation process for internal audits, quality control, and Sarbanes-Oxley compliance.
- We helped screen candidates for a FTE Director of Business Intelligence.

Spend Forecasting for Global Accounts (project team of 3 analysts). For a Fortune 100 technology company, we produced key business intelligence for global strategy and sales force allocation. The core analytics objective was to forecast the spending of global accounts on principal product groups.

The resulting reports are strategically and tactically central. They predict the number of product units an account will purchase. Therefore, this analysis is central to global strategic decision-making, such as:

- Recognizing and acting upon outstanding opportunities, across global accounts, regions, and product groups
- Assisting with partnering and pricing decisions, regionally and world-wide

These BI deliverables also apply tactically, as in:

- Deploying the right type and variation of sales activity, according to a metrics-based segmentation of accounts, account divisions, and region-specific account sites
- Identifying the strongest sales opportunities within accounts, for each product group
- Adjusting sales quotas in a data-driven manner

Substantial technical challenges were overcome for this project. The spend forecasting model pulls together a range of evidence, clues and indicators that describe each global account. The data sources that fulfilled these demands were disparate, originating across internal business units, as well as external resources such as survey results. These data sources required a heavy dose of discovery, assessment, evaluation, interpretation, cleansing, "crunching," augmentation, unification and merging.

Customer Analytics: Valuation. For a national consumer awards incentives program that is integrated with all major national credit cards, we lead a project to evaluate predictive modeling to distinguish new customers with high value from "hit-and-run" customers. Predictive accuracy was accomplished in part by integrating external geographical data.

Customer Analytics: Misc. We have consulted to assist with miscellaneous predictive analytics projects such as bank attrition, office supply "hit-and-run customer" detection, and online service attrition and acquisition.

Data Mining for Breakthroughs in Science. Corporations and investors alike require objective intelligence that speculates over areas of scientific and technology progress, and forecasts which are most promising. For a cutting-edge startup company, we conducted a data mining initiative for the study of science. This initiative examines the patterns in how published articles cite one another. We implemented and optimized proprietary data mining and clustering methods, applying them to discover key research and technology trends.

Analytics: Teaching. I have conducted various industry training programs in data mining (both software vendor-neutral and product specific). I assisted with sales for various data mining software, for a variety of business applications. I

was a computer science professor and award-winning teacher at Columbia University; I taught 29 graduate, undergraduate and equivalent semesters, including graduate courses in data mining and related technologies

Text Mining and Web Mining technology development projects:

- Grouping written reports and characterizing these groups
- Linguistics-based knowledge extraction from web pages
- Web page topic identification
- Other proprietary web page analysis methods
- Misc. text interpretation tasks (i.e., distinguishing shades of meaning)

Security, Intrusion Detection and Fraud Detection:

- Profiling malicious Internet surveillance -- large dataset of 11 million records
- Lead design, development and delivery of an email "flow" data extraction and analysis tool for a major government agency. Includes fraud and virus detection
- Solved conceptual design problems for general Internet user fraud detection
- Worked with credit rating analysis data (similar to fraud)