

Higher Business ROI with Optimized Prediction

Yottamine's Unique and Powerful Solution

Forward thinking businesses are starting to use predictive analytics to predict which future business events will bring them positive outcomes, thereby investing their limited resources on those events to improve ROI. Many see good results but also find their current tools make this an error-prone process. Yottamine Analytics offers a unique predictive analytics solution that enables businesses to select and rank future business events on the basis of their individual predicted profitability in one single, highly automated step. Benchmarks show that Yottamine can increase profit by more than 50% and ROI by more than 100% over the most popular open source machine learning algorithms used by leading data scientists. It makes 10% more profit and 30% better ROI than models generated by teams of highly skilled data scientists using commercial software. In addition to being very fast and highly scalable, Yottamine is also highly automated, eliminating the need for IT skills and trial-and-error parameter tuning.

Real-World Business Predictive Modeling

Companies of all types and sizes are increasingly looking to use predictive analytics to improve business decision making in Marketing, Risk Management, Fraud Detection, Network Security, Product Quality and many other areas.

Predictive Modeling and Machine Learning solutions are now abundant, but they all share three key weaknesses that make them nearly impossible to align analytic models with business objectives.

Valuable events are rare, but the tools are biased to the majority, common events.

The events that a business wishes to predict for the greatest return are often rare. For instance, respondents to a direct marketing campaign will be a tiny fraction of the targeted recipients. Such data is said to be highly unbalanced.

Conventional predictive modeling tools often have bias toward the majority events to minimize the overall error rate, working against the business goal of predicting rare, highly rewarding events.

The value of the event varies widely, but the tools treat all events as equal in value.

The outcome of each individual business event contributes differently to the overall return of the correctly predicted events. For example, in a marketing campaign, one positive response will bring more revenue than another. Each action the business invests in will normally have some baseline underlying cost associated with it. For instance, the effort involved in checking a fraudulent transaction has its individual underlying cost associated with it. Since it is often impossible for the business to take action on all the possible events, prioritization becomes important. Businesses will only act on the events that have the highest potential to bring positive outcome.

Conventional tools often treat each outcome equally, rather than considering its individual cost or reward. This produces sub-optimal prediction result and consequently sub-optimal business return.

Big data modeling is compute- intensive, but the tools do not scale accordingly.

Big Data means more data types, more data points, more features and more noise. That is what makes predictive modeling more complex and more compute-intensive now more than ever.

What is Optimized Prediction?

Predictive models with the greatest business value are often those which can predict both an outcome and the measurable value of that outcome. Such models are found across many industries including Insurance, Financial Services, Telecommunication Services, as well as across a variety of applications including Fraud Detection, Risk Management, and Marketing.

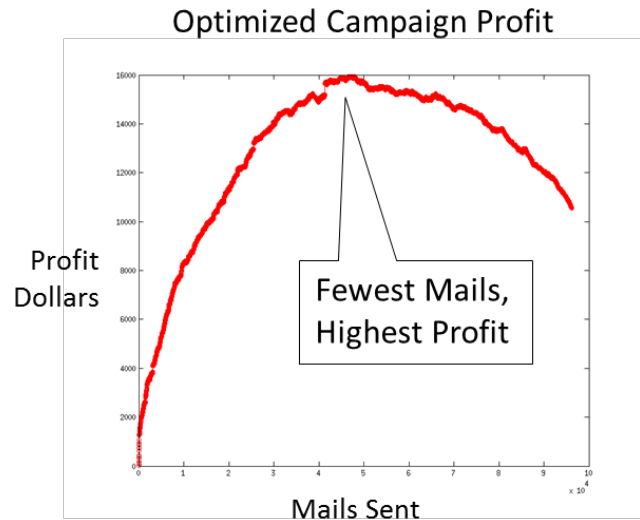
Optimized Prediction enables a bank to avoid the costliest bad loans, an insurance company to predict the fraudulent claims with the highest losses, or a marketer to better target the most profitable customers.

The figure on the bottom right illustrates how Optimized Prediction can find the point of highest business return for a direct mailing marketing campaign.

The horizontal axis shows the number of actions taken by the business. Too many actions will increase the cost/investment required for the business and reduce the overall return. Optimized Prediction can find the most optimal spot for the business, where the least number of actions produce the highest profit. This translates to higher ROI.

Provable Improvement in Business Result

To illustrate the power of this new algorithm, we conducted a test on a real-world business problem. We compared our results against a field of well-known machine learning algorithms in open source software, as well as results generate by teams of experienced expert data scientists using commercial software products such as SAS.



Problem Setting

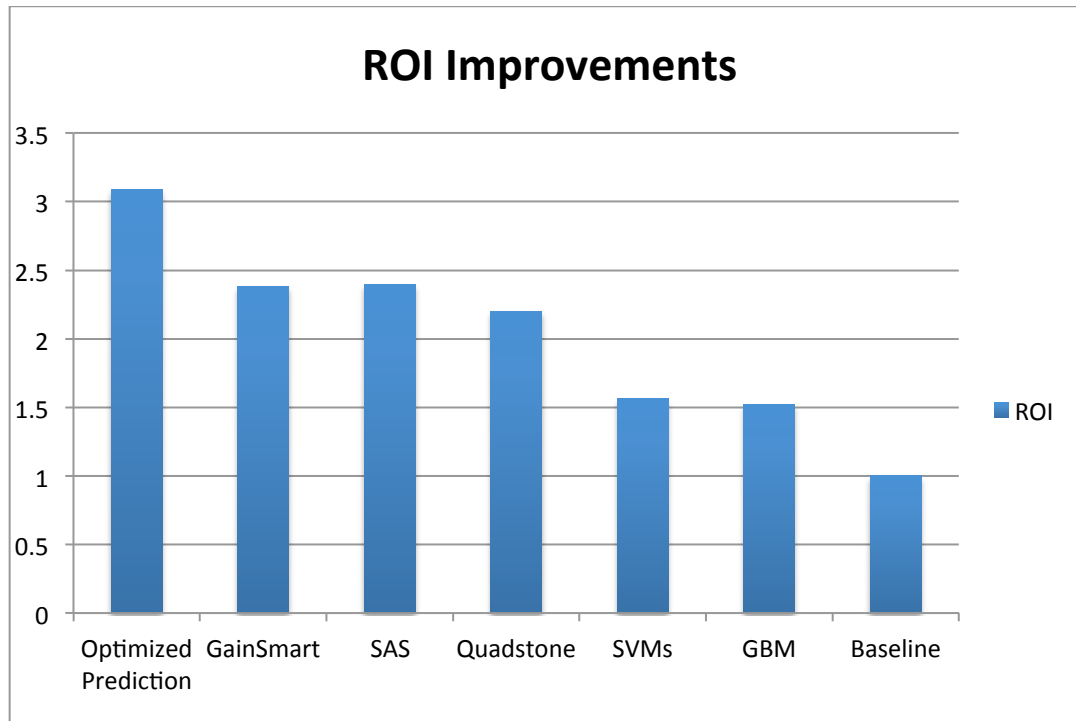
In this real-world business problem, the business wishes to optimize its business return by predicting correctly the outcome of each upcoming business event. The outcome of each event has either a reward or cost associated with it. Each event can either generate revenue or incur a loss for the business. The cost associated with each outcome is coming from the actions or the investments a business is required to explore the opportunity. This could be the mailing cost associated with a direct mailing campaign or the resource required to determine if a transaction is fraudulent or not. In this case, the event that generated the highest return is 600 times larger than the associated cost/investment required to explore it. There is a huge difference in the ratio between the reward and the cost. What makes the problem even more challenging is the fact that only 5% of all the events will generate a positive outcome. This makes building an accurate predictive model which provides the highest return very difficult. This situation occurs very often in commercial applications such as a propensity model, churn analysis, risk rating and fraud detection, as the outcome of each event has different costs associated with it.

Other Solutions Approaches

In the ROI improvement and profit charts, the ROI and profit generated by each different software product/algorithm is presented. To establish the baseline, we used the case where a business has all the resources available to perform an action (i.e. investment) on all potential events. To compare with other machine learning algorithms, we used Gradient Boosting Machine (GBM) in R and Support Vector Machines. GBM is one of the most popular methods that help many data scientists win competitions in

this type of problem. When comparing with GBM and SVMs, all algorithms use the same set of features. Optimized Prediction doesn't require complex preprocessing to achieve a state-of-the-art result.

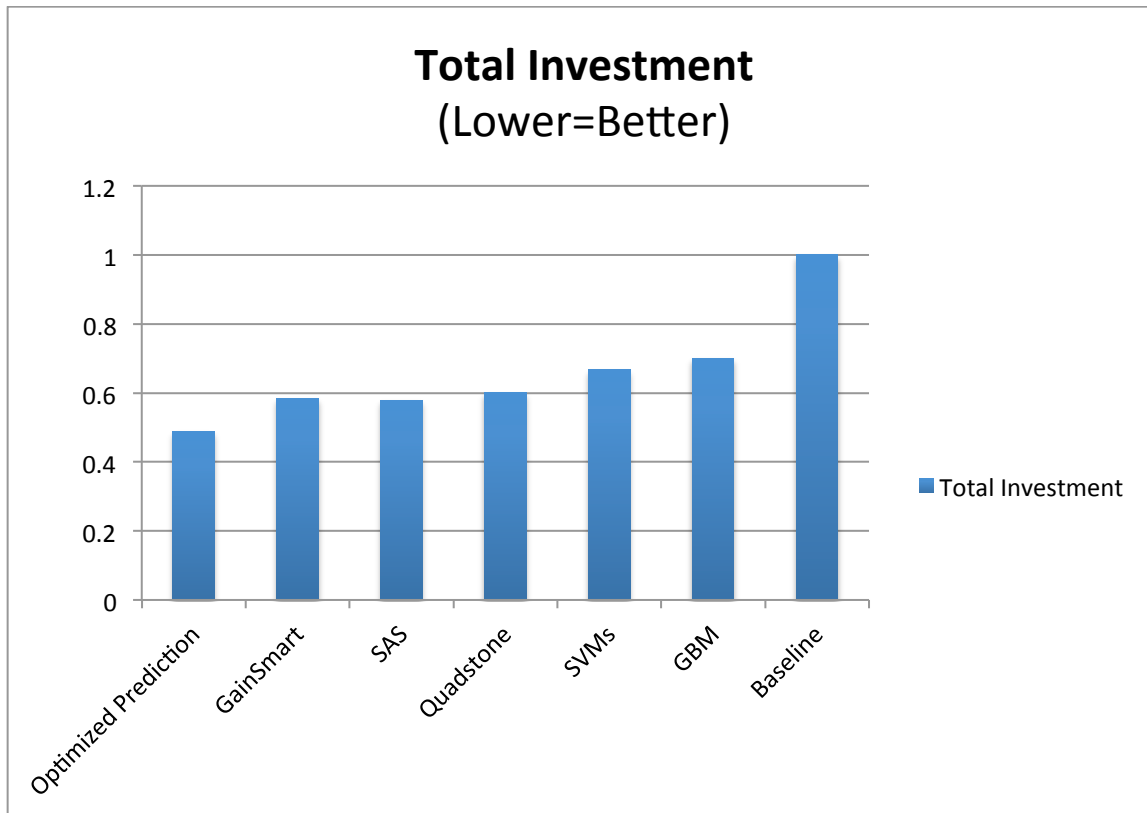
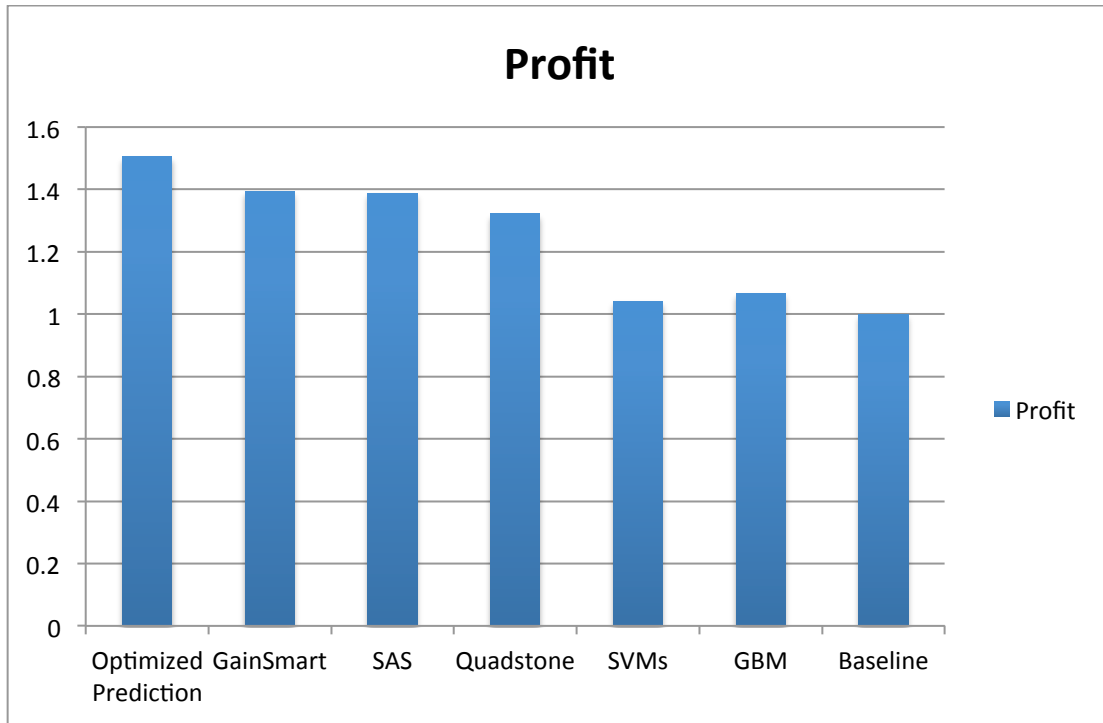
To illustrate how Optimized Prediction compared with a team of elite data scientists and experts on this problem, we also included results obtained by three teams, namely GainSmart, SAS and Quadstone. These teams of scientists applied several techniques together including combining different models, sampling, feature reduction on this problem and they worked on this problem for an extensive period of time. The SAS team is a group of experts using SAS as their main modeling tool.



Winning Numbers

Optimized Prediction generates a much higher profit (50% or more) than both Support Vector Machines and Gradient Boosting Machines (GBM). It is important to note that all three algorithms use the same features. Optimized Prediction is not based on an ensemble of models, but rather a single model. This result is achieved by using the automatic mode of Yottamine Predictive Platform. This is designed not only for this problem, but also can accommodate a wide range of problems.

When comparing with teams of expert data scientists who develop highly specialized models, Optimized Prediction produces 10% more profit and 30% more ROI than an ensemble of approaches developed specifically for this problem. It is quite significant; as Optimized Prediction is now comparing with results of multiple techniques combined together and parameters that are fine-tuned particular to this set. For example, the next best result by GainSmart is achieved by combining a regression and a classification algorithm together, and also a sophisticated feature reduction scheme. In contrast, Optimized Prediction is a single algorithm and single model computed in a single automated step.



The figure in the previous page shows the amount of cost/investment required by the business in order to generate the corresponding profit. The Optimized Prediction algorithm requires less than 50% of the investment of the baseline to generate 1.5 times the profit. This translates to 300% higher ROI for every dollar invested. When comparing with both machine-learning approaches, Optimized Prediction required 20% less cost/investment, but generated more than 40% profit. As a result Optimized Prediction is roughly two times better in terms of ROI than other competing machine learning approaches.

Free Yottamine Trial

If you have labeled machine learning training data available, you can try Yottamine at no charge. Get started on the [Free Tier](#) page.

On-Premise (Hadoop Cluster)

Yottamine is currently available for your Hadoop Cluster and also on the AWS cloud. Our Hadoop-based machine learning software allows you to build Optimized Prediction model from your big data stored in your on-premise Hadoop and Spark cluster.

On the Cloud (Amazon AWS)

Our cloud based service lets you convert business data to blind training data on-premise and then use the unlimited computing scale of the cloud to build your model.

- No hardware or software to buy, install, configure or manage.
- No limits on the size of the data set or compute resources.
- No vendor lock-in – pay as you go, cancel at any time.

To find out more about Yottamine and what we can do for your company, please [contact us](#).