MACHINE LEARNING FOR DIRECT MAIL MARKETING

Improve Your Direct Mail Results with AI and Machine Learning

AlescoData
An Alesco Company

CIO APPLICATIONS TOP 25 MACHINE LEARNING COMPANIES - 2018

In this new ebook you'll get answers to such questions as:

- Machine learning – What is it and how does it work?
- How does a computer learn?
- How is machine learning different from traditional statistical analysis?
- Why is machine learning superior to other analytic approaches?
- How can machine learning make my Direct Mail campaigns more effective?

INTRODUCTION TO MACHINE LEARNING

> We are drowning in information and starving for knowledge.
>
> - Rutherford D. Rogers

Are you struggling with disappointing results from your direct mail programs? Let's face it, this well-established media channel is not without its challenges, not least of which are ongoing increases in postage and production costs. Unfortunately, traditional tactics like segmentation and statistical regression, while once at the forefront of predictive analytics, are no longer achieving the desired results.

Many clients and prospective clients, possibly like you, have been searching for a game-changing solution capable of significantly improving their direct mail campaigns.

Enter Machine Learning.

Recent advances in cost-affordable, high-speed computing coupled with advanced Machine Learning are changing the game in marketing analytics.

In this ebook we will:

- **Describe machine learning and explain how it works**
- **Highlight the differences between machine learning and traditional analytics**
- **Demonstrate why machine learning offers superior performance**
- **Discuss what machine learning can do for your direct mail marketing campaigns**

Machine Learning: What is it and how does it work?

Over the last few years, tremendous hype has surrounded artificial intelligence (AI) and machine learning (ML). You've probably heard terms like neural networks, recommendation algorithms and deep learning. These are ML techniques that large companies are using to deliver more relevant search results (Google), increase revenue (Amazon) and improve user experience (Facebook). However, most business people probably have a limited understanding of what ML technology is and how it works, and few have ever utilized it in their own business.

> "Humans can typically create one or two good models a week; Machine Learning can create thousands of models a week."
>
> - Thomas H. Davenport, WSJ excerpt

Stanford University defines machine learning as a field of study that gives computers the ability to learn without being explicitly programmed. It is a branch of artificial intelligence based on the idea that computers can learn from data, identify patterns and make decisions with minimal human intervention. For our purposes, Machine learning is a method of data analysis that automates and improves analytical model building.

Machine Learning is the science of getting computers to learn like humans do, and improve their learning over time in autonomous fashion, by feeding them data and information in the form of observations and real-world interactions.

Daniel Faggella,
TechEmergence.com

Rising interest in, and proliferation of, AI and ML is due to rapid advances in computational processing speeds, including massively parallel graphics processing units, coupled with highly affordable data storage. These make it possible to quickly generate more precise mathematical models that analyze larger amounts of increasingly complex data. More precise models mean more profitable marketing.

Within the field of data analytics (specifically predictive analytics), machine learning is a method used to develop complex predictive algorithms that analyze current and historical data to make predictions about future events.

In layman's terms, a machine learning algorithm analyzes a company's internal CRM database and predicts specific outcomes, such as which prospects have the highest likelihood of becoming customers or which offer would most appeal to a prospective customer. In fact, ML algorithms can be effectively applied throughout the customer lifecycle (acquisition, growth, retention and win-back).

IN SUMMARY

Machine Learning...

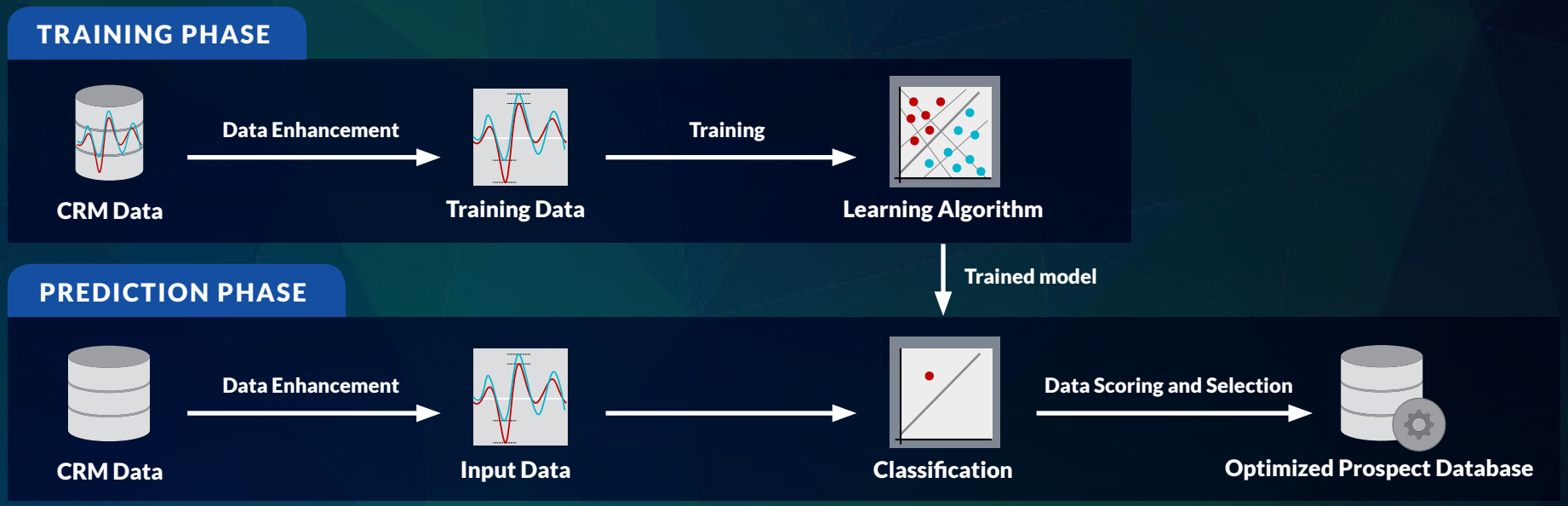
- **...is a type of artificial intelligence that gives computers the ability to learn without being explicitly programmed.**
- **...automates analysis of Big Data.**
- **...allows computers to create algorithms than can learn from, and make predictions on, data.**
- **...can analyze thousands of data points in parallel to identify trends and clusters.**
- **...overcomes the limitations of scale and refinement inherent in traditional model building.**
- **...focuses on the development of computer programs that change when exposed to new data.**
- **...can be used to optimize the entire customer lifecycle.**

But what is “learning” and how can a machine learn?

In machine learning, systems are developed not by codifying rules but rather by feeding the system data and letting the system evolve based on that data. This is where the learning occurs. Just like humans, these systems learn from previous experience (i.e., training data). The machine is programmed to create a model on the basis of the training data it receives, progressively improving the accuracy of its output and adapting to changing patterns in the data. The biggest difference between human learning and ML is the speed at which machine learning can happen.

The algorithm looks for patterns in data by utilizing training data. In direct mail applications, training data include contact details (name and address) along with RFM information typically contained in a CRM database. To this RFM information we add over 4,400 columns of additional demographic, psychographic and socio-economic information from our vast data repository. Once the data is loaded into the platform, the machines utilize parallel processing to run through millions of concurrent simulations, analyzing every combination of data attributes available on each record in order to create a predictive model.

Once the model is built, we can feed new data into the platform, enabling it to score and rank the new data based on a chosen outcome. For example, we can feed prospect data into the platform and it will score and rank the prospects by likelihood of response or conversion. It can also provide estimated response rates, average sale amounts or other KPIs based on whatever measure is of interest.



As noted earlier, advances in computational power and the availability of inexpensive data storage have supercharged the Machine Learning process. Our proprietary ML platform combines massively parallel processing (MPP) with a GPU/CPU infrastructure running at **speeds of over 100 teraflops with a distributed storage system collectively approaching petabyte scale**. The system is a marriage of Big Data assets coupled with a purpose-built machine learning infrastructure optimized specifically for direct marketing applications.

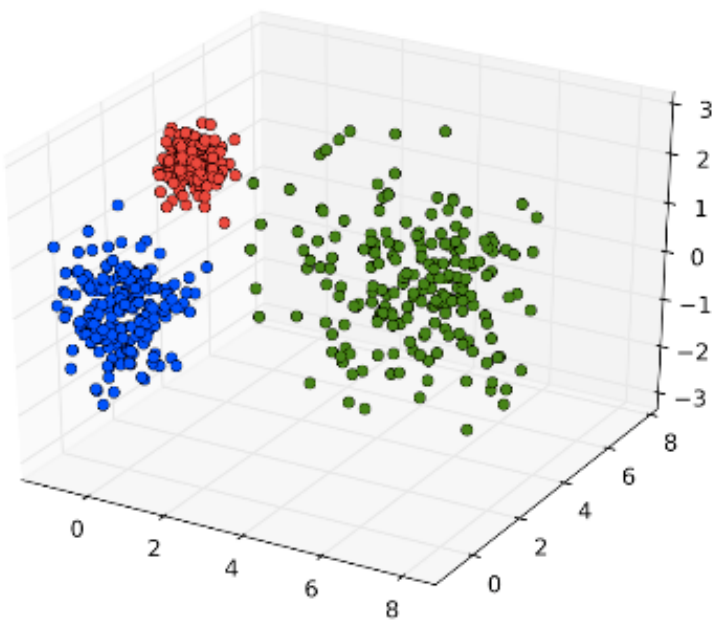
Built to handle wide, sparse and noisy data sets (much like you find in most CRM databases), the platform is coupled with over **half a billion consumer records** containing hundreds of demographic, psychographic and socio-economic attributes used in the development of the learning algorithms.

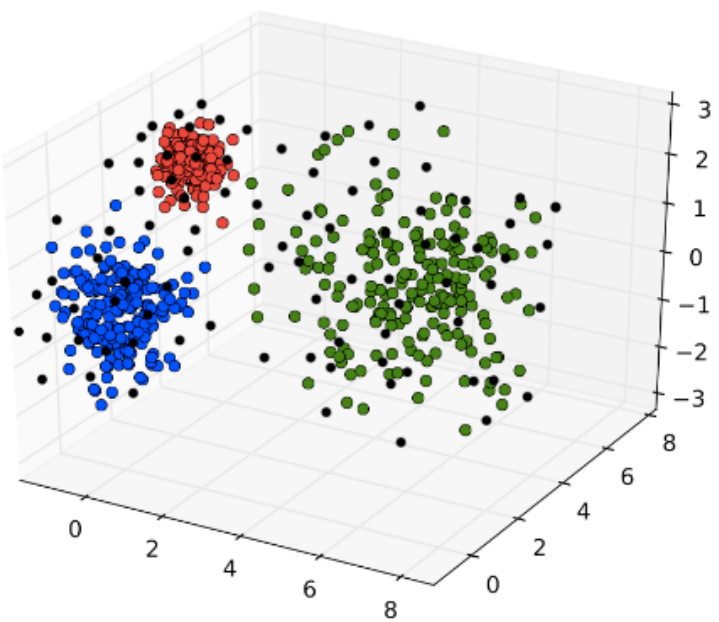
By utilizing this state-of-the-art computer horsepower, the machines are able to **"see" data points in multi-dimensional space** by utilizing what are called vectors. A vector is a multi-dimensional point in a space of numerical features that represents an object or, in our case, a customer or prospect. Essentially, all of the attributes that define each customer or prospect are transformed into a numerical representation, or vector, which facilitates the building of a classification algorithm.

How is machine learning different from traditional statistical analysis?

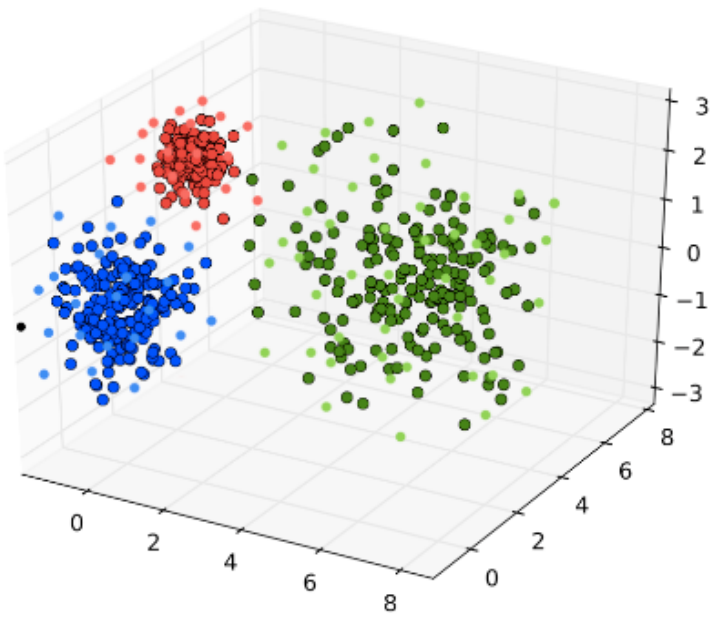
What follows is an illustration of the machine learning process in action. The visual representation below is the result of a ML classification algorithm which was created by analyzing a client's CRM database coupled with the attributes from our data repository. These comprise the algorithms' training data. The red dots indicate high-value customers, the blue dots represent low-value customers and the green dots represent non-customers (non-buyers, cancellations, etc).

Now that the model has been created, it can be used to analyze a prospect data set to identify which prospects have the highest likelihood of becoming high-value customers, low-value customers or non-customers. In the image below, we've ingested the prospect records (represented by the black dots), appended attributes from our data repository and processed them through the machine learning algorithm.





Based on the proximity of each of the black dots to one of the three groups identified from the training data, the algorithm is able to categorize each prospect into one of the three categories with a high degree of accuracy. In this way, marketing campaigns can be optimized to focus on prospects with the highest likelihood of becoming high-value customers.



IN SUMMARY

Our Machine Learning Platform is...

- **...a Marriage of Big Data Assets & Purpose-Built Machine Learning Infrastructure.**
- **...Built and Optimized for Direct Marketing Applications.**
- **...Coupled With Over Half a Billion Consumer Records Containing Hundreds of Attributes.**
- **...Built To Handle Wide, Sparse & Noisy Data Sets.**
- **...Operated by an Experienced Team of Data Scientists.**
- **...Generates Custom, Client-Specific Models Built For Each Application.**

How is Machine Learning different from, and better than, traditional analytic approaches?

Traditional statistical approaches begin with a reduction of the data set. In fact, most statistical regression models only contain a handful of variables. The statistician first manually analyzes the data characteristics to determine which attributes will be candidates for model selection. This is generally done on a univariate or bivariate basis in order to make the process manageable. This initial narrowing of the data set represents a significant difference from machine learning and is the primary reason traditional techniques lack the predictive power of their ML counterparts.

Essentially, the statistician has to look for and find patterns in the data and it helps significantly if they know what they're looking for ahead of time and have a deep understanding of the data with which they are working. In many ways, it's more art than science and helps to explain why there is sometimes significant performance variance between statistical models.

Conversely, in machine learning we seek expansion of the data set. We look to incorporate more, rather than less, relevant data into the analysis. Machine learning enables computers to analyze data for structure, even if we do not have a theory of what that structure looks like. Because machine learning uses an iterative approach to learn from data, the learning process is easily automated and results in a deeper and more layered method of applying iterative math to break down large data sets. Machine learning platforms are able to learn from and analyze billions of observations. Passes are run through the data until a robust pattern is found. The scale and speed at which this is done simply cannot be matched by traditional analytic approaches and raw manpower.

MACHINE LEARNING is ...

a subfield of computer science and AI which deals with building systems that can learn from data, instead of explicitly programmed instructions.

STATISTICAL MODELING is ...

a subfield of mathematics which deals with finding relationships between variables to predict an outcome.

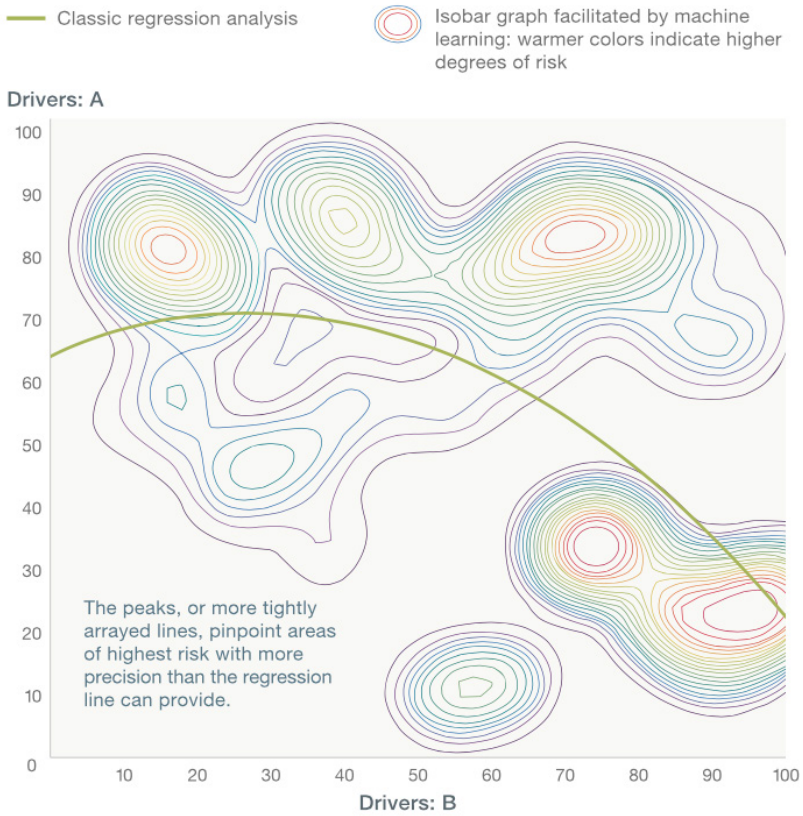
(Analytics Vidhya)

The schematic at right from McKinsey & Co was derived from a churn analysis performed as part of an actual client engagement. It provides an excellent visual representation illustrating the difference between a classic regression model and a machine learning algorithm.

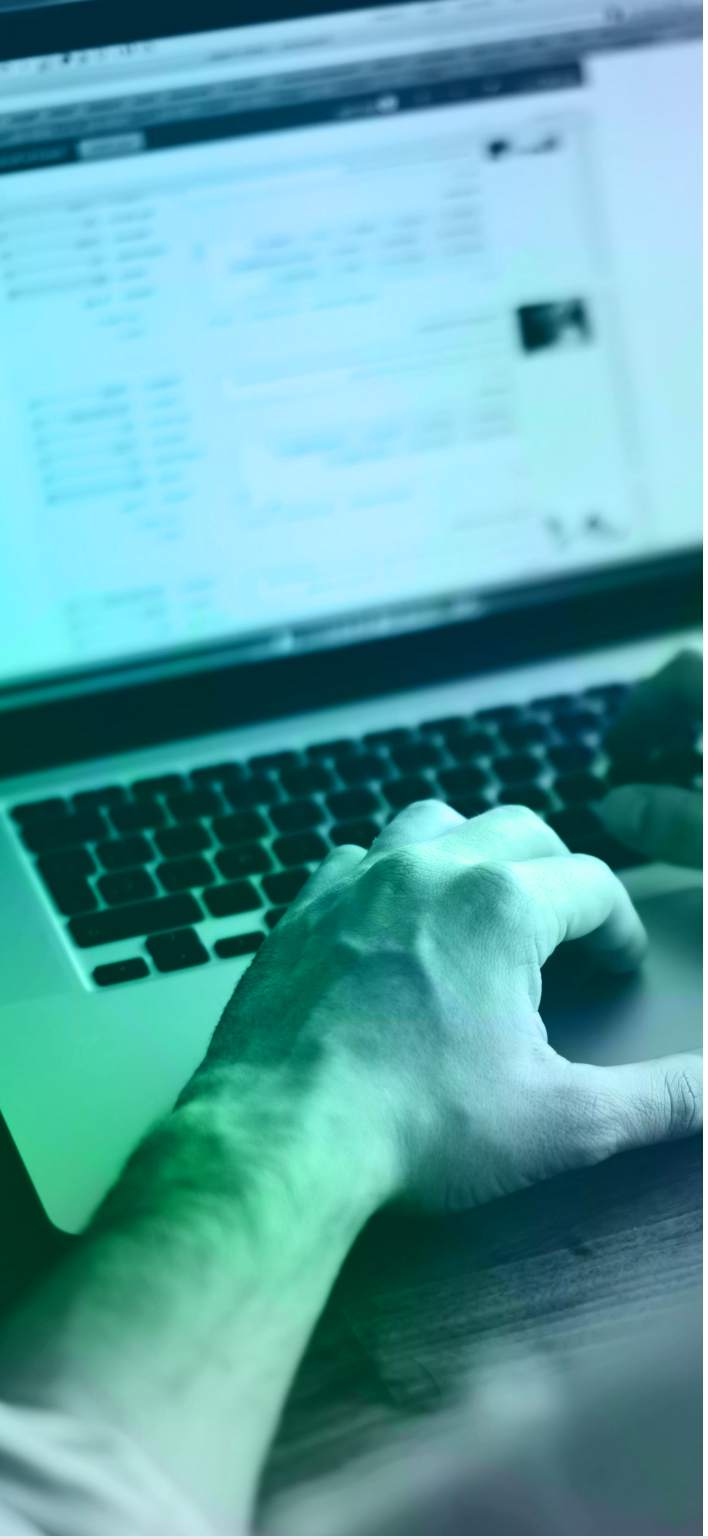
In the example, Driver A and Driver B are two customer metrics that have been intentionally anonymized. We see the correlation that the regression model identified between Drivers A and B in the analysis via the arching green line. Because regression models are either linear (straight) or nonlinear (curved), the models are represented by a line (the regression line) that has been "fitted" to the raw data. In this example, the curved green regression line represents the boundary which separates at-risk customers from non-at-risk customers.

However, when we look at the results of the Machine Learning algorithm overlaid on the same graphic, we see that there are "pockets" of at-risk customers within multiple, distinct groups all along the spectrum of Driver A and Driver B values. And, we see that there is at least one group of at-risk customers that the regression model misses completely (bottom middle of the graphic). In other words, there is no continuity of boundary as assumed by the regression analysis.

This real-world example shows the stark contrast between the abilities of traditional statistical analysis and machine learning. The ML algorithm was able to more finely discern underlying patterns within the data, providing a more accurate and highly refined identification of at-risk customers. This, in turn, allows the company to more efficiently focus its efforts on the correct customers to reduce churn.



Value at risk from customer churn, telecom example, by McKinsey & Co.



How can Machine Learning make my Direct Mail campaigns more effective?

Many clients need a solution to help them better align marketing investments with potential customer value to maximize profitability. As we've shown, machine learning algorithms excel in their predictive capabilities and their ability to identify the best, and worst, candidates for a direct mail campaign. They can be used effectively to optimize direct mail campaigns throughout the entire customer lifecycle and can maximize any KPI or combination of KPIs.

Machine Learning can optimize your lists and the databases you use for acquisition mailings, boosting response rates by 20%-30% or more. It can also be used to optimize data derived from a previously built statistical model, generating double-digit lifts in ROI.

Machine learning can enable you to enhance your upsell/cross-sell efforts and so identify additional revenue-generating opportunities. Once customers are lost, a win-back algorithm can be developed to identify those previous customers having the highest likelihood of returning when presented with a compelling offer or series of offers. It can even be used to successfully reactivate archived names when coupled with our historical name and address file containing over 40 years of address histories of over 100 million households.

All campaigns can be optimized based on any outcome desired, including response rate, cost-per-lead, cost-per-conversion, average donation or any combination thereof.

In more advanced scenarios, termed prescriptive analytics, Machine Learning can analyze timing, frequency, offer, creative, packaging, list input, list cost and more. Once the analysis is completed, our platform can tell you whom to mail, when to mail them, what to mail them and which offer to make, creating the most optimized and data-driven direct mail solution available anywhere.

Traditional tools of the trade are no longer producing the results they once did. If you're in need of a data-driven solution that outperforms traditional approaches, look no further than Alesco's machine learning service—the next generation of predictive analytics.

IN SUMMARY

Machine Learning can...

- **...optimize prospect list selection—whether yours or ours.**
- **...optimize the results of a previously built statistical model.**
- **...identify upsell/cross sell opportunities.**
- **...optimize win-back campaigns.**
- **...optimize response rate, CPL, cost per conversion, average donation.**
- **...determine who, when, what and which offer should be mailed.**

Click on the following links to download our Machine Learning Case Studies:

CASE STUDY: Home Security Company Realizes 26% Increase in Direct Mail Response Rates

CASE STUDY: Regional Non-Profit Sees 98% Increase in Direct Mail Response Rate

ABOUT ALESCO

Alesco **combines omni channel Big Data assets with powerful Machine Learning algorithms** that drive client's new customer acquisition programs. Our machine learning platform was purpose-built for direct marketing applications and is generating results that are **30%-50% greater than traditional segmentation models**. Our people-based, offline and online data assets enable you to reach prospects through the mailbox, inbox, desktop, and device.

If you need to improve your direct mail results but aren't sure where to start, reach out to the team at Alesco. We've helped countless clients meet and exceed growth goals by incorporating machine learning into their direct mail marketing efforts. Contact Paul Theriot at **Paul@alescodata.com** or via phone at **626-437-9902**.

AlescoData
An Alesco Company

CIO APPLICATIONS TOP 25 MACHINE LEARNING COMPANIES - 2018