IT@Intel White Paper

Intel IT
Advanced Predictive Analytics
August 2013

Using Big Data Predictive Analytics to Optimize Sales

Intel IT developed an advanced predictive analytics solution to prioritize which resellers have the greatest potential for high-volume sales.

Executive Overview

To help the Intel sales organization optimize its account management and increase estimated incremental revenue, Intel IT developed an advanced predictive analytics solution to identify and prioritize which resellers have the greatest potential for high-volume sales.

To create our advanced predictive analytics solution, in 2012 we deployed a proof of concept (PoC) to learn more about how Intel resellers engage with Intel and the sales organization. Since conducting the PoC, we have achieved the following:

- Developed an enterprise-level, end-to-end predictive analytics engine based on what we learned from the PoC
- Produced new estimated incremental revenue increases of USD 3 million in the PoC and we expect an annual increase of USD 20 million when scaled globally
- Formulated a plan to add Intel® Distribution for Apache Hadoop* software to implement the engine globally and expand its capabilities

Using the predictive analytics engine, the sales organization can now analyze real-time data from various sources to determine the potential for each customer to create a new revenue opportunity for Intel.

We plan to add more value to the predictive analytics engine by designing the capability to determine when an account should be contacted and what products and support to offer. This added functionality should enable a more optimized and data-driven sales organization.

Einat Ronen
Advanced Analytics Sales and Marketing Business Line Manager, Intel IT

Kenneth Burns
IT Customer Capability Analyst, Intel IT

Ajay Chandramouly
Big Data Domain Owner, Intel IT

Contents

IT@INTEL

The IT@Intel program connects IT professionals around the world with their peers inside our organization – sharing lessons learned, methods and strategies. Our goal is simple: Share Intel IT best practices that create business value and make IT a competitive advantage. Visit us today at www.intel.com/IT or contact your local Intel representative if you'd like to learn more.

BACKGROUND

As the velocity and complexity of Intel's business accelerates, Intel IT has identified opportunities for business intelligence (BI) and advanced predictive analytics to become a significant competitive differentiator for Intel. Using advanced predictive models provides deeper insight into business patterns that drive operational efficiencies. This enables us to analyze information and make decisions more quickly, helping Intel stay competitive. In one application of these models, we developed an advanced predictive analytics solution that helps the Intel sales organization strategically focus on large-volume resellers to deliver significantly greater revenue.

Intel works with over 140,000 resellers that are members of the Intel® Technology Provider program. These resellers—Intel's customers—specify, design, build, and resell Intel®-based technology products and solutions.

These customers typically purchase through a network of Intel® Authorized Distributors, and Intel's sales organization provides online support, warranty services, and marketing support.

The Intel Technology Provider program started more than 10 years ago. At the beginning, Intel sold components to distribution, distribution sold to resellers, and then resellers built the final product to be sold to end users.

The market trend toward smaller mobile devices has changed channel dynamics. Larger original design manufacturers (ODMs) and original equipment manufacturers (OEMs) are now building the end product, such as a laptop, business Ultrabook™ device, or tablet, and then selling that product to distributors, who in turn sell it to resellers (see Figure 1). The sales organization tracks Intel® components that are sold to the ODMs and OEMs, but little data is available after that. The result is that the sales organization doesn't have the data it needs to support the reseller; specifically, what exactly the reseller is marketing that includes Intel technology.

With a diverse customer base, the sales organization needed assistance prioritizing which customers should receive the most support, determining the optimal time in the customer's buying cycle to contact them, and deciding what products or support to offer. In the past, the sales organization focused on the largest accounts. However, data was limited for some customers, which meant that account size couldn't always be determined.

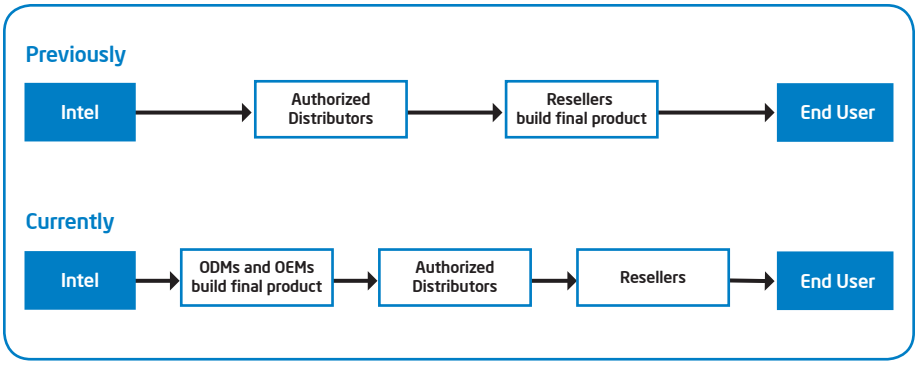

Figure 1. Channel dynamics have changed since the Intel® Technology Provider program began more than 10 years ago.

The sales organization contacted Intel IT to help achieve their goals. To develop a solution, we considered each aspect of their goals:

- **Who.** Create a ranking system that would help ensure the sales organization focused on the customers that had the greatest potential for increasing revenue.
- **When.** Use the unstructured data to trigger alerts for the sales organization to contact accounts at the optimal time in their buying cycle.
- **What.** Help the sales organization know what products or support should be discussed with each account.

Intel IT partnered with the sales organization to develop a predictive analytics engine and re-engineer business processes to create the ranking system: Identify the customers with the greatest potential for high-volume sales that the sales teams should create better engagements with to help them grow their business. Adding the capabilities of determining when to contact the customers and what products and support to offer would be addressed later.

SOLUTION

The Intel sales organization needed a tool to help focus on customers with the greatest potential for increasing sales volume. Intel IT created a predictive analytics engine that could rank a subset of the 140,000 resellers of the Intel Technology Provider program. As our first step in developing a solution, in Q4 2012 we deployed a proof of concept (PoC) in Intel's Asia-Pacific online sales center.

For the PoC, we used the following approach:

- **Understand the business domain.** We needed to learn how Intel resellers engaged with Intel and the sales organization.
- **Devise a solution for the PoC.** Identify the available data sets, develop a predictive analytics engine, and rank the top 300 customers the sales organization should contact.
- **Optimize the engine.** During the PoC, we incorporated a feedback mechanism that enabled us to improve the model with more accurate data instead of static business rules.

Understanding the Business Domain

Before we could start mining data to fuel a predictive analytics engine, we needed to better understand the relationship between Intel, resellers, and the sales organization.

One focus of business growth for the sales organization is the Intel Technology Provider program and its resellers. The sales organization supports the resellers to help them grow their business. However, given the change in channel dynamics and the fact that resellers purchase directly from distributors, the sales organization has limited data to determine the Intel technology purchased by the resellers. Additionally, the sales organization can interact with only a portion of the resellers, so it is constantly re-assessing on which accounts to focus. The sales organization's goal is to concentrate primarily on the accounts with the greatest potential for high-volume sales and recalibrate quarterly to adjust the outreach strategy.

Existing processes to classify accounts involved manually utilizing limited data sets from existing accounts. However, with little data available on many Intel Technology Provider program members, some accounts were being passed over because the data to classify them was insufficient.

As we began to understand the business domain, we realized we needed to find hidden patterns and deciding factors in data that were not available in the sales database (for example, number of employees per Intel Technology Provider member). The solution also had to be automated to help the sales organization most efficiently use the Intel Technology Provider program database each quarter. To accomplish this automation, we needed to use predictive techniques to overcome data limitations that existed in the Intel Technology Provider program membership. This would allow more data to be added as the sales organization gathered feedback from each account.

Devising a Solution

We used a three-phase approach to build the solution. In phase one, we determined what data sets could be used for mining. For phase two, we built the predictive analytics engine that modeled and learned from the mined data. Finally, in phase three, we optimized the engine to include feedback from the sales organization in place of static business rules to help establish initial account rankings.

IDENTIFYING AVAILABLE DATA SETS

Once we understood the business problem we needed to solve, we started looking for data sources to mine. The Intel Technology Provider program provided the registration data we needed. Then, we populated the enterprise data warehouse (EDW) with demographic, training, and sales data, which enabled us to create customer profiles. Once we outlined the profiles, we added information that would help us determine each customer's revenue potential.

THE PREDICTIVE ANALYTICS ENGINE

After we identified the sources to mine, we filled out and scored the profiles of all Intel Technology Provider program members using two data mining techniques: unsupervised clustering and supervised classification. We then used the scores to rank the accounts and provide a list of the greatest potential high-volume accounts to the Asia-Pacific online sales center manager.

Unsupervised Clustering

Unsupervised clustering refers to the process of finding hidden patterns in unlabeled data. We used an algorithm to consider all mined data characteristics for the profiles we created. As patterns emerged, accounts with similar characteristics were grouped into a cluster. Once the clusters were defined, we began the supervised classification process.

Supervised Classification

For supervised classification, we added static business rules—which we created based on characteristics of known high-volume customers—to the clusters to make the classification more intelligent. By using a random forest algorithm to analyze the unsupervised clusters and static business rules, we determined similarities between known and unknown high-volume customers.

Ranking

The unsupervised clustering and supervised classification efforts yielded a ranking of 300 accounts. We included probability data (an assigned number between zero and one) for each account to help rank the customers from 1 to 300 (highest volume potential to lowest volume potential). Figure 2 shows the predictive analytics engine using unsupervised clustering and supervised classification to rank customers.

OPTIMIZING THE ENGINE

When we first ran the predictive analytics engine in the PoC, the initial rankings were based on static business rules. This approach was necessary to provide the sales organization with an initial list of accounts to engage with. After the sales organization engaged with the ranked accounts, they collected and integrated feedback to be included with the most impacting static business rules. This activity helped produce more accurate, up-to-date rankings.

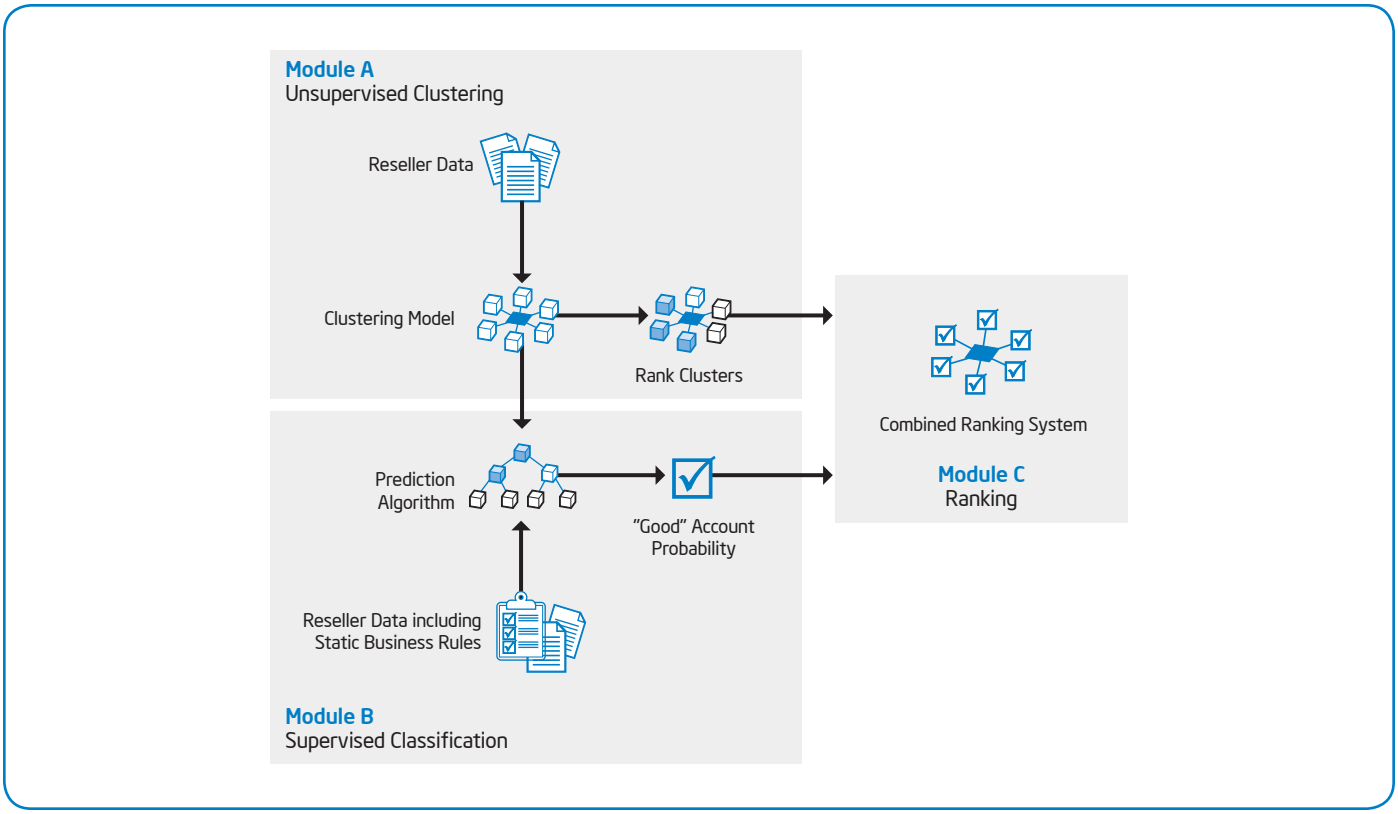

Figure 2. The predictive analytics engine uses unsupervised clustering and supervised classification to rank customers with the greatest potential for high-volume sales.

How the Predictive Analytics Engine Works

The predictive analytics engine that arose from the PoC features three primary elements: the EDW, the predictive analytics platform (PAP), and a user interface (UI). Figure 3 illustrates the cyclical nature of the engine.

EDW

The data warehouse contains most of the data we use for modeling. It is composed of internal databases and sales feedback from customer engagements following prior reseller ranking reports. The data warehouse is continually updated with new data from the various sources.

PAP

The sales organization assigns new accounts quarterly. Therefore, we run the predictive analytics engine each quarter, based on the most recent data available in the data warehouse. The recommendation engine ranks all active accounts and then exports a report to the sales UI. The sales team can then predict which accounts are expected to produce the highest volume and should therefore be prioritized for engagement.

UI

The online sales center manager takes the ranking report and assigns each account to a sales agent. Each sales agent works with his or her assigned accounts to help the account increase its sales volume. In the process, the sales agent gathers feedback from the customer about Intel® products, OEM relationships, sales volume, and other factors that the engine will use for the next quarter's report.

RESULTS

We measured the success of the predictive analytics engine primarily by the potential incremental revenue. We worked with Intel Finance to determine a fair percentage of these estimates that could be realistically attributed to the engine's rankings. An increase in revenue indicated success.

For the PoC, incremental revenue estimates increased by USD 3 million as a result of the Intel sales organization's efforts using the ranked list developed by our predictive analytics engine. Since the PoC was conducted in the Asia-Pacific online sales center in Q4 2012, the predictive analytics engine has also been rolled out in the Europe, Middle East, Africa, and North America regions.[1] The annual estimates for increased incremental revenue for the Asia-Pacific and Europe, Middle East, and Africa regions equal approximately USD 14 million. When the engine is implemented globally, we expect an annual increase in incremental revenue of more than USD 20 million.

The predictive analytics engine identified three times as many potential high-volume resellers substantially faster in the Asia-Pacific region compared to what the sales organization could identify using manual methods. Automation is making the sales organization more efficient and strengthening customer relationships. Until now, customers have been dealing directly with only with OEMs and ODMs, which are understandably more focused on the finished product. Customers appreciate the additional support they receive for the Intel components inside the finished product, as it helps them add more value to how they sell the product.

[1] The increased incremental revenue estimates for North America are not available at this time.

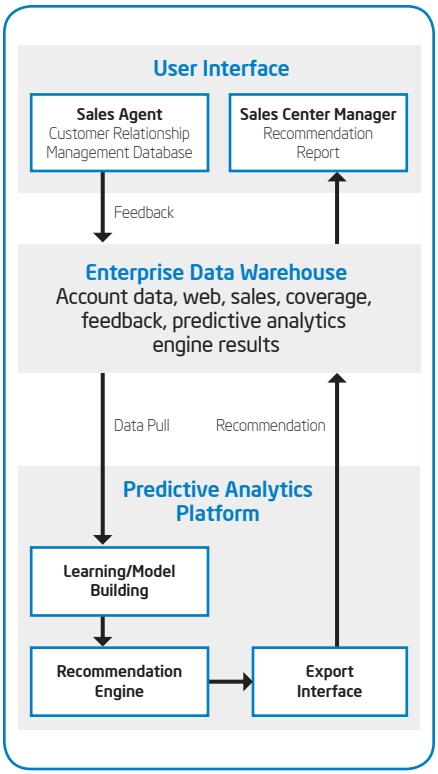

Figure 3. The predictive analytics engine includes an enterprise data warehouse, a predictive analytics platform, and a user interface where the sales organization can provide real-time feedback.

NEXT STEPS

We plan to add more value to the predictive analytics engine by designing it to determine when an account should be contacted and what exactly should be discussed. This added functionality should enable a more data-driven and optimized sales organization.

To enable this functionality, we plan to use the Intel® Distribution for Apache Hadoop* software (Intel® Distribution). This open platform for next-generation analytics more cost-effectively handles large unstructured data sets than traditional approaches.

Determining when an account should be contacted and what should be discussed will require more data sources, such as structured and unstructured social media and web activity data. Processing this additional data will require the stronger, more efficient calculation capabilities offered by the Intel Distribution.

Furthermore, we expect to apply the prediction of potential high-volume accounts to customers other than resellers and Intel Technology Provider program members. This approach will require new data sources, and we will not be able to use the structured registration data we had in the PoC. We will need to rely heavily on the Intel Distribution to mine unstructured data and predict account potential.

CONCLUSION

Intel IT is embracing big data and advanced predictive analytics to help Intel's sales organization drive operational efficiency and competitive differentiation. We developed an enterprise-level, end-to-end predictive analytics engine that is directly responsible for a portion of the sales organization's increase in estimated incremental revenue.

The increased velocity and complexity of Intel's business affects almost every business unit. Keeping up with market changes requires agility and adaptability, and big data predictive analytics solutions such as this one enable the sales organization to act faster based on real data.

As markets change so do our customers. A customer may initially show little potential for high-volume sales. But later—because of new products, new technology, or a merger and acquisition—that same customer could rise in the rankings to become a candidate for increased sales volume. With the predictive analytics engine continuously analyzing real-time data from various sources, the sales organization now has a way to recognize the potential in that customer, which equals new revenue opportunity for Intel.

FOR MORE INFORMATION

Visit www.intel.com/it to find white papers on related topics:

- "Mining Big Data in the Enterprise for Better Business Intelligence"
- "Integrating Apache Hadoop* into Intel's Big Data Environment"

CONTRIBUTORS

Michal Afek, Advanced Analytics
Ronen Aharony, Advanced Analytics
Eiran Bolless, Advanced Analytics
Maor Haber, Advanced Analytics
Abraham Israeli, Advanced Analytics
Lilian Kubail, SMG
Patty Schorch, Customer Capability
Shahar Shpigelman, Advanced Analytics

ACRONYMS

PoC	proof of concept
BI	business intelligence
EDW	enterprise data warehouse
ODM	original design manufacturer
OEM	original equipment manufacturer
PAP	predictive analytics platform
UI	user interface

For more information on Intel IT best practices, visit www.intel.com/it.