SAP Predictive Analysis
How Predictive Analysis Works.

Scott Leaver, Global Solution Manager
Today’s Agenda

- Predictive Analytics – Strategic Investment Area for SAP
- Introduction to SAP Predictive Analysis
  - Demonstration
- The Integration of Predictive Analysis with Industry / Line of Business applications and SAP BI clients
- Predictive Analysis roadmap at SAP
- Ramp Up Program/Resources
Predictive Analytics – Strategic Investment 2012 and beyond

Solution Overview
- Predictive Analytics is an important requirement for our Industry and LOB applications, and BI solutions. It is a significant market - $2.0B 2014 (IDC).
- To establish a leadership position in this space, our product strategy is to provide -
  - A modern UX/UI to support the definition of predictive analysis processes and their visualization
  - In-Database Predictive Analysis within SAP HANA for real time and large data volume data analysis
  - Open Source “R” integration to provide a very comprehensive range of predictive algorithms
  - Embed Predictive Analysis in our Industry / LOBs / BI client tools.

Customer Perspective
- Value prop:
  - Forrester - “Predictive power makes all the difference in business success”
  - TDWI - “Among BI disciplines, prediction provides the most business value”
- Target Industries/LOB:
  - Retail, Health Care, Banking, Utilities, Manufacturing, HCM, CRM
What do we mean by ‘Predictive’?

“Predictive analysis helps connect data to effective action by drawing reliable conclusions about current conditions and future events.” ~ Gareth Herschel, Research Director, Gartner Group

Note: Data Mining and Statistics are part of the Industry’s view of Predictive Analytics.
Realize significant returns.

*Firms that personalize offers based on historical behavior benefit in terms of redemptions as well as new customer acquisition.*

**Example:**

“A ‘big-box’ retailer used predictive analytics…targeting offers extended to individual customers based on their past purchase patterns, the retailer was able to *increase redemption rates* and *increase membership* in the first six months of rollout."

-February 2012, © *Forrester Research, Inc.*
What is SAP BusinessObjects Predictive Analysis?

A statistical analysis and data mining solution that enables you to…

- Intuitively design complex predictive models
  - Read and write from data stored in HANA, Universes, IQ, and other sources
  - Drag-and-drop visual interface for data selection, preparation, and processing
  - Predictive library with R integration

- Visualize, discover, and share hidden insights
- Unleash Big Data with SAP HANA’s power
Common Myths about Predictive Analysis

Myth #1: “You can’t start until a data warehouse is in place.”

Myth #2: “Predictive analytics requires a Ph.D. or math degree.”

Myth #3: “There is a long time-to-value with predictive analytics.”

Myth #4: “It can do more harm than good unless I have it down to an exact science.”

Myth #5: “The results are often incomprehensible!”
Target Audience for Predictive Analysis
Who is it for?

- **Business Analysts (“New User”)**
  - May have little data mining experience
  - Want good answers easily (no potholes!)
  - “Ease-of-Use” - Automation capabilities

- **Data mining experts**
  - Long time DM users / Expert Analysts
  - Want more time to focus on in depth analysis
  - “High performance,” Rich set of tools; Stats integration
Intuitively, easily, design predictive models (from simple to complex)

- Read and write from data stored in SAP HANA, BO Universes, Oracle, SQLServer, MySQL, CSV file, Excel, and other sources
- Drag-and-drop visual interface for data selection, preparation, and processing
- Unleash HANA’s in-database predictive algorithms (PAL)
  - K-Means
  - Multi-linear regression
- Leverage open-source predictive algorithms via R integration
- Natively implemented algorithms

SAP Predictive Analysis Features
SAP Predictive Analysis
Visualization

- Visualize, discover, and share hidden insights
  - Advanced visualization designed where you’d expect it – natively from within the modelling tool
  - Share insights via PMML and with other analytic tools
Predictive Analysis Process

**Data Loading**
1. Understand the business and identify issues
2. Load the SAP and non-SAP data into HANA or other RDBMS

**Data Preparation**
1. Visualize and examine the data
2. Sample, filter, merge, append, apply formulas

**Data Processing**
1. Define the model via clustering, classification, association, time series, etc.
2. Run the model

**Data Visualization and Sharing**
1. Visualize the model for better understanding
2. Store the model and result back to HANA or other source
3. Share results via PMML and with other BI client tools

**Step 1**
Data Loading

**Step 2**
Data Preparation

**Step 3**
Data Processing (Define model)

**Step 4**
Data Visualization and Sharing
Options to use Predictive Analysis:
Leverage HANA’s Predictive Capabilities

**PA + HANA**

Win7 Desktop OS

- BO Predictive Analysis
  - Native Algorithms

HANA (PAL)

- Data

**PA “STAND-ALONE”**

Win7 Desktop OS

- BO Predictive Analysis
  - Native Algorithms

R

- Data

SAP Enterprise BI Platform – BI Clients

SAP Applications, CRM, LoB solutions
Algorithms Map to Specific Business Usage with data

- **Clustering**
  - Understand Customers
  - K-Means

- **Forecast**
  - Be Prepared
  - C4.5 Decision Tree
  - Linear Regression

- **Classification**
  - Manage Products
  - KNN
  - ABC Classification

- **Association Rules**
  - Uncover Patterns
  - Apriori

- **Alternative Analysis**
  - Make Correct Decision
  - Weighted Score Tables
PRODUCT DEMONSTRATION

High-level solution overview
SAP Predictive Analysis
Integration Scenarios
Embedding Predictive Capabilities - BI clients

- Bringing BI to more and more users
- Embedding predictive analysis for trend analysis, forecasting, outlier detection etc.
SAP HANA Predictive Applications

- **Smart Meter Analysis** - high performance analytics for the utility industry
- **Manufacturing** - Demand driven collaborative scheduling for just-in-time manufacturing
- **Sales forecasting** - determine the probability of success
- **Retail** - performance management and out-of-stock predictive analysis, Affinity insights
- **CRM** – predictive customer segmentation and retention analysis
- **Offer to Cash** - cross-sell & up-sell opportunities, dynamic pricing strategies
- **Fraud management**
- **Price optimization**
- **Utilities/Oil & Gas** - Load demand forecasting for utilities
- **Banking** - Campaign management
Customer Analytics Overview

Customer Revenue Performance Management
- Target Users: Sales Manager, Product Manager, Marketing Manager
- Empowers Sales and Marketing to invest the right resources into the right customers, products and channels.
- Combination of front office interactions and back office financials for a comprehensive performance view
- Beyond just analysis to what-if simulation & advice

Account Intelligence
- Target User: Account Executive
- Empowers Account Executives with real-time and complete view of all Customer activities
- Across channels
- Along with exception identification and management tools to improve sales success.
- Comprises Sales, Financials, CRM, other sources like customer satisfaction

Predictive Customer Segmentation
- Target User: Marketing Manager
- Empowers Sales & Marketing professional to rapidly and easily segment large customer populations
- Manage Target Groups for insight to action
- Trigger initiatives like customized offers for each segment and channel
High Performance Analysis for Smart Meter Data

Challenge

• Identify energy consumption pattern that can be used for Customer Segmentation
• Smart meter data volume is huge

Solution

Using SAP HANA, K Means algorithm, Multiple Linear Regression, Variable Selection, Significance tests, Normality Test using L and IMSL and SQL Script.

Results

• Clustering of smart meter data → applying the k-means clustering to >20 million x 96-dimensional vectors
• High performance clustering computation
DNA - Sales Forecasting on SAP HANA

**Challenge**

- Up-to-date reports on sales pipeline towards end of fiscal period
- Data Latency
- Aggregation of line items
- Drill-down to single opportunities

**Solution**

- Single SAP HANA 1.0 server connected to SAP CRM
- Live replication of opportunity items into SAP HANA
- Leverage Forecasting Algorithms

**Results**

- 4 TB in transactional system = 600GB SAP HANA
- 1,000,000 opportunity items
- Data latency reduced from 2 hours to seconds
- Enables real-time decisions based on real-time data
Affinity Insight for Retail on SAP HANA

Challenge

- Tailored to the roles of retail category managers and merchandising managers
- Both of whom are concerned with product relationships and their associated financial performance in achieving retail corporate targets.

Solution

- Using SAP HANA and the PAL Apriori algorithm with modification for maximal performance for the target usage

Results

- See affinity relationships easily on a broad scale
- Understand product relationships that contributed to affinity relationships
- Conduct simple “what-if” queries against a demand model
- See store-level performance relative to the product dimension.

### Challenge

Tailored to the roles of retail category managers and merchandising managers. Both of whom are concerned with product relationships and their associated financial performance in achieving retail corporate targets.

### Solution

Using SAP HANA and the PAL Apriori algorithm with modification for maximal performance for the target usage.

### Results

- See affinity relationships easily on a broad scale
- Understand product relationships that contributed to affinity relationships
- Conduct simple “what-if” queries against a demand model
- See store-level performance relative to the product dimension.
Cisco Adopt SAP HANA to Deliver Reporting and Predictive Insight in Near Real-Time

Company
Cisco Inc.

Headquarters
San Jose

Industry
High-Tech Computer Peripheral Manufacturing

Products and Services
High-Tech and Networking Equipment

Employees
71,000

Web Site
www.cisco.com

Objectives
- Allow Cisco sales to have real-time insights and make better decisions
- Better data delivery decisions to improve the business, both in terms of optimization and also to drive topline growth
- Apply predictive analytics to sales performance to better define causality of the drivers behind performance and better understand the seasonality of buying patterns

Why SAP
- Ability of SAP HANA to combine near real-time reporting with predictive analytics to provide new insights
- Support for R integration with SAP HANA to provide flexibility in predictive analysis of seasonality and clustering
- Ability to transform results of statistical analysis into actionable business insights effectively

Benefits
- Better respond to customers and deliver tangible business value
- Reduce time to transform information into insights and improvement in the quality of decision-making on those insights
- Ultimately drive higher profitability and growth

“The HANA platform at Cisco has been used to deliver near real-time insights to our execs, and the integration with R will allow us to combine the predictive algorithms in R with this near-real-time data from HANA. The net impact is that we will be able to take the capability which takes weeks and months to put together, and deliver just-in-time as the business is changing.”

Piyush Bhargava, Distinguished Engineer IT, Cisco Systems

Faster
Reporting and analysis

Better
Understanding of business drivers

Greater
Ability to support evolving business needs
SAP Predictive Analysis and HANA Roadmaps
PAL Algorithm Roadmap

**SP3**
- Cover classical predictive analysis algorithms in each category
  - Clustering
    - K-Means
  - Classification
    - KNN
    - C4.5 Decision Tree
    - Linear Regression
  - Association
    - Apriori
  - Classification
    - ABC Classification
    - Weighted Score Tables

**SP4 & SP5**
- Extend algorithms in each category
- Cover Time Series and some preprocessing algorithms
  - Clustering
    - Anomaly detection
  - Classification
    - Bi-Variate Exponential / Geometric / Logistic / Natural Logarithmic Regression
    - Naive Bayes
    - CHAID Decision Tree
  - Time Series
    - Moving Average
    - Single / Double / Triple exponential smoothing
  - Preprocessing
    - Outlier Detection (Inter-Quartile Range)
    - Correlation calculations
    - Sampling

**Further Release**
- Cover more complicated algorithms
  - Clustering
    - Hierarchical Agglomeration
    - Kohonen SOMs
  - Classification
    - Polynomial / Stepwise / Cox Regression
    - C & RT Decision Tree
    - Neural Networks
    - Support Vector Machine (SVM)
  - Social Network Analysis
    - Sequential / Link Analysis
  - Optimization
    - Linear programming
    - Monte Carlo Method
    - Generic Algorithm

- Nov 2011
- May - Nov 2012
- N/A
**Predictive Product Releases 2012**

Please note that these are planned dates, not commitments

<table>
<thead>
<tr>
<th>Current release</th>
</tr>
</thead>
<tbody>
<tr>
<td>Predictive Workbench</td>
</tr>
<tr>
<td>SBOP Predictive Analysis 1.0</td>
</tr>
<tr>
<td>SBOP Predictive Analysis 1.1</td>
</tr>
<tr>
<td>HANA Predictive Analysis Library SP03</td>
</tr>
<tr>
<td>HANA Predictive Analysis Library SP04</td>
</tr>
<tr>
<td>HANA Predictive Analysis Library SP05</td>
</tr>
<tr>
<td>HANA R Integration</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>2011</th>
<th>2012</th>
</tr>
</thead>
<tbody>
<tr>
<td>O</td>
<td>N</td>
</tr>
<tr>
<td>PW 14.1</td>
<td>SBOP PA V1.0</td>
</tr>
<tr>
<td></td>
<td>SBOP PA V1.1</td>
</tr>
<tr>
<td></td>
<td>PAL V1</td>
</tr>
<tr>
<td></td>
<td>PAL V2</td>
</tr>
<tr>
<td></td>
<td>PAL V3</td>
</tr>
<tr>
<td></td>
<td>Embedded &amp; Standalone</td>
</tr>
</tbody>
</table>

This presentation and SAP's strategy and possible future developments are subject to change and may be changed by SAP at any time for any reason without notice. This document is provided without a warranty of any kind, either express or implied, including but not limited to, the implied warranties of merchantability, fitness for a particular purpose, or non-infringement.
Predictive Analysis – Release Roadmap Plan

**SBOP PA 1.0**
- Modern UX/UI
- Universe UNX as a data source
- Universe UNV XI3 and BI 4, JDBC, XLS, CSV data sources
- Outlier detection, regression, exp. smoothing...
- In-database HANA PAL support in SP1
- R integration of 13 algorithms
- Predictive analysis visualizations
- PMML support

**SBOP PA 1.1**
- Universe UNIX as a data source
- SAP BW as data source
- PAL SPS03 & SPS04 Revisions
- Multi Language support
- Algorithm comparison
- Data Preparation Techniques – Binning, Normalization & Data Types
- Business Rules Integration
- PMML Reader
- Visualization - Large data volume, More Visualizations, More Interactivity
- SBOP PA SDK for internal consumption
- Invoking R scripts from SBOP PA
- Export Analyses and models to HANA
- Improve BI Clients Integration
- Support for HANA R Integration

**SBOP PA 1.2**
- Ensemble modeling, bagging, boosting, auto modeling
- Visual Numerics integration
- More algorithms from HANA PAL
- Predictive Analysis wizard
- Automated data discovery
- SBOP PA SDK available for external consumption
- Visualization – more & sharing & reusability options

Q1 2012
Q2 2012
Q3 2013
Q4 2012
SAP Predictive Analysis Ramp Up Program
Please familiarize yourself with the SAP Predictive Analysis Service 1.0 Ramp-Up program by visiting the Ramp Up home page on Service Marketplace or http://bit.ly/PARAMPUP

For further details about the nomination process, please contact the regional Ramp-Up Owners: (first.last@sap.com)

APA: Allen Xiao
EMEA: Nicole Fuchs
Japan: Bobby Borromeo
Latin America: Andrio Garcia
North America: Krishna Mohan Mamidipaka

Also visit SCN.SAP.com for further details about Predictive Analysis!
THANK YOU

Scott Leaver
Scott.leaver@sap.com
Twitter: @scottleaver