

DATA MINING AND BUSINESS ANALYTICS SERVICES

The Case for Data Mining

The capacity of digital data storage worldwide has doubled every nine months for at least a decade, twice the rate predicted by *Moore's Law* for the growth of computing power during the same period. This growth disparity is called the *Storage Law* and is a motivating reason for the increasing importance of data mining.

As a result of the Storage Law, our ability to capture and store business data has far outpaced our ability to process and utilize it. This growing challenge has produced massive, ever growing data tombs, or data warehouses that are effectively write-only; data is deposited to merely rest in peace, since in all likelihood it will never be accessed again. This situation is a significant waste of potential business intelligence (BI). Data warehouses represent the data assets of a business enterprise, and to leave this information untapped is a serious lapse in productivity.

Deployments of e-commerce websites serve as a good example of how companies overlook the valuable nature of data assets. An integral part of many e-commerce sites is the data warehouse that parallels the transactional database upon which operational use of the e-commerce site is based. In some cases, these sites include basic online analytical processing (OLAP) functionality, but true data mining capabilities are not typically exploited.

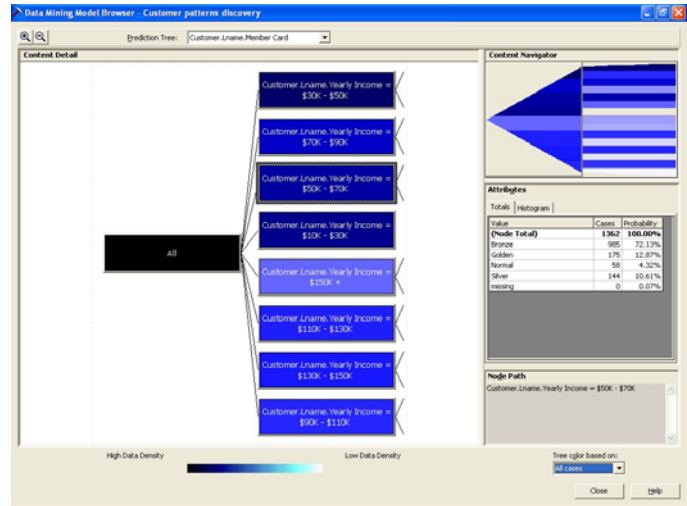
Data Mining

Data mining is defined as the identification of interesting structure in data. Structure designates patterns, statistical or predictive models of the data, and relationships among parts of the data. It is the predictive characteristic of data mining that separates it from OLAP. OLAP works with views the data warehouse to yield analysis of historical data. Data mining takes an important step further by providing predictive capabilities.

AMULET knowledge engineers utilize leading-edge analytical tools in order to apply data mining technology to extract valuable BI from your data. Here are some examples of the business intelligence you can achieve:

- What products do customers purchase together?
- What are the products young female customers like to buy?
- What other products should be recommended to a customer that bought a given item?

The most important facet of AMULET's data mining services is our ability to design, develop, and deploy custom data mining algorithms to suit your specific BI needs. Through years of research experience, AMULET staff is well versed in the cross discipline nature of data mining in fields such as mathematical statistics, computer science, and visualization.



Decision Tree Data Mining Algorithm Analysis

The lack of data mining means that valuable BI is not available for a profit center that is growing in importance to many companies. Furthermore, significant knowledge about customer buying patterns, association rules for shopping cart analysis, and classification rules remain unexplored and unutilized.

Recognizing these deficiencies, AMULET Development Corp. offers specialized data mining services for new and existing e-commerce deployments.

- Utilize Decision Tree and Clustering algorithms.
- Develop/integrate custom data mining algorithms.
- Provide professional data mining consulting services.

Business Analytics

Traditional business analytics is an important precursor to a successful data mining effort. Here are some typical steps that yield the proper data environment for data mining:

- Data cleansing to insure consistency and integrity in the data warehouse.
- Extract, transform, and load (ETL) to populate the data warehouse from distributed transactional systems.
- Define and create OLAP cubes for analysis.
- Use the *Pivot Table Service* for embedding cubes in Microsoft Excel.

With the above preparations in place, AMULET provides you with valuable analytics based on your corporate data assets. We use an iterative approach for successful data mining where our knowledge engineers consult with your staff to understand the goals of the knowledge discovery project, design a business analytics strategy unique to your business, develop an OLAP solution for business analytics analytics, and deploy/tune appropriate data mining algorithms to deliver BI to help your business grow and operate more efficiently. The process repeats over time in order to refine the results.

Country	State Province	City	Unit Sales	Store Cost	Store Sales
Canada	All Customers Total		296,773.00	225,627.23	9501
Canada	Canada Total				
Mexico	Mexico Total				
USA	USA Total		296,773.00	225,627.23	9501
	CA Total		74,746.00	63,530.43	1151
	+ Alameda		2,674.00	2,279.71	85
	+ Alameda		2,440.00	2,046.70	85
	+ Bellflower		3,106.00	2,638.47	95
	+ Berkeley		136.00	129.29	
	+ Beverly Hills		2,407.00	2,479.46	\$6,194.37
	+ Burbank		3,056.00	2,623.04	\$6,577.33
	+ Burlingame		198.00	160.42	\$407.38
	+ Chula Vista		2,999.00	2,492.91	\$6,294.30
	+ Colton		126.00	114.88	\$267.76
	+ Concord		105.00	86.48	\$213.77
	+ Coronado		2,391.00	2,031.17	\$5,051.15
	+ Daly City		329.00	112.46	\$771.60
	+ Downey		3,440.00	2,921.76	\$7,367.06
	+ El Cajon		2,543.00	2,173.71	\$5,460.42
	+ Fremont		163.00	141.17	\$350.22
	+ Glendale		3,294.00	2,879.12	\$7,082.91
	+ Grossmont		2,133.00	1,798.32	\$4,458.60
	+ Imperial Beach		1,616.00	1,340.29	\$3,409.34
	+ La Jolla		1,938.00	1,640.61	\$4,081.37
	+ La Mesa		1,634.00	1,565.21	\$3,908.26
	+ Lakewood		2,497.00	2,033.70	\$5,174.12
	+ Lemon Grove		2,651.00	2,243.36	\$5,536.02
	+ Lincoln Acres		2,176.00	1,867.95	\$4,691.94
	+ Long Beach		2,578.00	2,578.35	\$6,422.37
	+ Los Angeles		2,009.00	1,799.89	\$4,312.99

OLAP Cube for Sales Analysis

With the business analytics solutions established by AMULET, you can easily slice-and-dice through your data warehouse to obtain strategic insights – extract a single item (slice) and compare items in a cross-tabulated table (dice).

We employ the following advanced data mining algorithms in order to exact the proper results for your BI needs: decision trees, clustering, naïve Bayes classification, sequence clustering, time series, and association rules.

Case Studies

Descriptive and predictive data mining can be applied to a wide variety of knowledge discovery problems in a wide range of industries. AMULET knowledge engineers work to understand your organization's goals, and develop a data mining strategy to apply the best statistical methods to produce useful results. The deliverable of a typical data mining project is a research report that summarizes the findings of the study. Here is a representative series of case studies for successful data mining projects:

Manufacturer –

A manufacturer of consumer household products requires an analysis of their product registration database containing nine years of customer profile information: age range, income level, education, credit rating, and region of country for residence. The goal of the data mining project is to predict whether a customer will purchase a particular appliance product.

The approach taken to solve this BI problem is to use the decision tree data mining algorithm to classify customers into two distinct groups: likely to purchase the appliance, and unlikely to purchase the appliance. The resulting decision tree yields useful classification rules that are used to drive a successful direct mail campaign. Only customers deemed likely to purchase the appliance receive the mailer, thus significantly lowering the cost of the campaign, and improving sales conversions.

Online Retailer –

A retailer of apparel on the Internet needs an integrated BI solution for adding strategic cross-sell capabilities to their website's product catalog and shopping cart. Based on specific items placed in the cart, strategically selected cross-sell items will be offered to customers.

The solution to this market basket analysis problem centers around the use of the association rule data mining algorithm that is able to discover buying patterns and provide insight into which items are frequently purchased together by customers. Through data mining, the online retailer experiences increased sales and enhanced customer satisfaction.

Automobile Insurer –

An auto insurance company desires to better understand characteristics of its customers; specifically it wants to highlight traits that might identify policyholders with a high average claim cost.

The solution to this problem entails the use of a clustering data mining algorithm which is able to group policy holders into classes having high similarity in comparison to one another, but are very dissimilar to members in other classes. The process is able to identify clusters differentiated by residence location, make/model of vehicle, and income level.

Who is AMULET Development Corp?

AMULET Development Corp. is a premiere Web database and e-commerce integration firm founded in 1995 to provide quality technology solutions for businesses in a broad range of industries. Specializing in Microsoft server, Web, and database technologies, we've built many high profile e-business websites. Our current focus is business intelligence, analytics, and data mining using contemporary technology to help enterprises better utilize valuable data assets.

For information about AMULET's Data Mining and Business Analytics services please contact us:

Sonya Franklin 1-877-722-7393

info@amuletc.com

www.amuletc.com

AMULET maintains a Microsoft Competency in *Data Management Solutions* with a *Business Intelligence* (BI) specialization.

