



Nine Common Types of Data Mining Techniques Used in Predictive Analytics

By Laura Patterson, President, VisionEdge Marketing

Predictive analytics enable you to develop mathematical models to help better understand the variables driving success. Predictive analytics relies on formulas that compare past examples of success and failure, and then use these formulas to predict future outcomes. Predictive analytics, pattern recognition and classification problems are not new. Long used in the financial services and insurance industries, predictive analytics is about using statistics, data mining and game theory to analyze current and historical facts in order to make predictions about future events.

The value of predictive analytics is relatively obvious. The more you understand customer behavior and motivations the more effective your marketing. The more you know why some customers are loyal, how to attract and retain different customer segments, the more you can develop relevant compelling messages and offers. Predicting customer buying and product preferences and habits requires an analytical framework that enables you to discover meaningful patterns and relationships within customer data in order to do better message targeting, drive customer value and loyalty.

Predictive models have been used in business to assess the risk or potential associated with a particular set of conditions as a way to guide decision making. Predictive models analyze past performance to assess how likely a customer is to exhibit a specific behavior in the future in order to improve marketing effectiveness. Marketing and sales professionals are beginning to capture and analyze many different types of customer data: attitudinal, behavioral, and

transactional related to purchasing and product preferences in order to make predictions about future buying behavior.

Today's challenging environment is forcing more organizations to explore predictive analytics. While commonly used by market researchers when analyzing survey data, predictive analytics can actually be applied in real-time scenarios, such as personalizing offers to customers or improving an online customer experience. There are a variety of ways to approach predictive analytics and most approaches depend on clean databases, the ability to mine data in order to look for patterns or create classifications. It is important to understand these various approaches so you know when to use which one.

This article provides a quick explanation of the nine most common types of data mining techniques used in predictive analytics. Becoming familiar with these common approaches and techniques will go a long way toward enabling you to recognize patterns in customer preferences and buying behavior.

a) Regression analysis - Regression models are the mainstay of predictive analytics. The linear regression model analyzes the relationship between the response or dependent variable and a set of independent or predictor variables. This relationship is expressed as an equation that predicts the response variable as a linear function of the parameters.

b) Choice modeling - Choice modeling is an accurate and general purpose tool for making probabilistic predictions about human decision making behavior. It behooves every organization to target their marketing efforts at customers who have the highest probabilities of purchase. Choice models are used to identify the most important factors driving customer choices. Typically, the choice model enables the firm to compute an individual's likelihood of purchase, or some other behavioral response, based on variables that the firm has in its database, such as geo-demographics, past purchase behavior for similar products, attitudes or psychographics.

c) Rule Induction - Rule induction involved developing formal rules that are extracted from a set of observations. The rules extracted may represent a scientific model of the data, or just local patterns in the data. One major rule induction paradigm is association rule. Association Rules is about discovering interesting relationships between variables in large databases. It is a technique applied in data mining and uses rules to discover regularities between products. For example, if someone buys peanut butter and jelly they are also likely to buy bread. The idea behind association rules is to understand when a customer does X, they will most likely also do Y. Understanding these kinds of relationships can help with forecasting sales, promotional pricing or product placements.

d) Network/Link Analysis - This is another technique for associating like records. Link analysis is a subset of network analysis, explores relationships and associations between very many objects of different types that are not apparent from isolated pieces of information. It is commonly used for fraud detection and by law enforcement. You may be familiar with link analysis since a number of web search ranking algorithms use this technique.

e) Clustering /Ensembles - Cluster analysis or clustering is a way to categorize a collection of "objects" such as survey respondents, into groups or clusters in order to look for patterns. Ensemble Analysis is a newer approach that leverages multiple cluster solutions (an ensemble of potential solutions). There are a variety of ways to cluster or create ensembles. Regardless of the method, the purpose is generally the same- to use cluster analysis to partition into a group of segments and target markets in order to better understand and predict the behaviors and preferences of the segments. Clustering is a valuable predictive analytics approach when it comes to product positioning, new product development, usage habits, product requirements, and selecting test markets.

f) Neural networks - Neural networks were designed to mimic how the brain learns and analyzes information. Organizations develop and apply artificial neural networks to predictive analytics in order to create a single framework. The idea is that a neural network is much more efficient and accurate in circumstances where complex predictive analytics is required because

neural networks are composed of a series of interconnected calculating nodes that are designed to map a set of inputs into one or more output signals. Neural networks are ideal for deriving meaning from complicated or imprecise data and can be used to extract patterns and detect trends that are too complex to be noticed by either humans or other computer techniques. Marketing organizations find neural networks useful for predicting customer demand and customer segmentation.

g) Memory based/Case-based reasoning - This technique has results similar to neural network but goes about it differently. MBR looks for "neighbor" kind of data, rather than patterns. This approach uses a process of solving new problems based on the solutions of similar past problems. Memory-based reasoning is an empirical classification method and operates by comparing new unclassified records with known examples and patterns.

h) Decision trees - Decision trees use real data mining algorithms. Decision trees help with classification. A decision tree process will generate the rules followed in a process. Decision trees are useful for helping you to choose between several courses of action and enable you to explore the possible outcomes for various options in order to assess the risk and rewards for each potential course of action. This type of analysis is useful when you need to choose between different strategies or investment opportunities, especially when you have limited resources.

i) Uplift modeling, also known as net response modeling or incremental response modeling directly models the incremental impact of targeting marketing activities. The uplift of a marketing campaign is usually defined as the difference in response rate between a treated group and a randomized control group. Uplift modeling uses a randomized scientific control to measure the effectiveness of a marketing action and to build a model that predicts the incremental response to the marketing action.

These types of approaches and the use of predictive analytics takes you beyond the traditional slicing and dicing of your data so you can be smarter and more agile when it comes to marketing. With predictive analytics you can gain faster insights and optimize programs by

simultaneously testing copy, offers, and creative rather than deploying the more traditional A-B testing methodology which may take longer and therefore delay course adjustments.

Disclaimer: Any VEM information or reference to VEM that is to be used in advertising, press releases or promotional materials requires prior written approval from VEM. For permission requests, contact VEM at 512-681-8800 or info@visionedgemarketing.com. Translation and/or localization of this document requires an additional license from VEM. For more information on VEM, visit www.visionedgemarketing.com.