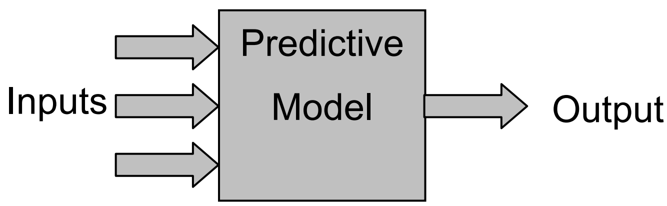
PREDICTIVE DYNAMIX INC

Predictive Modeling Technology

Predictive modeling is concerned with analyzing patterns and trends in historical and operational data in order to transform data into actionable decisions. This is accomplished by analyzing and modeling the dynamics of the application-specific data. In its raw form, this data is of limited value and is mainly used for reporting what has happened. However, when the data is compiled into a compact model, it is a powerful tool for proactively predicting what will happen.

In the abstract, a predictive model is a computational structure that can accurately forecast an outcome of interest (aka, output or dependent variable) when provided with input data (aka, independent variables) that have a measurable causal or coincident relationship to the output.



In order for predictive modeling to be useful in given application, two fundamental principles must hold:

- *Outcomes must have some level of predictability from known data.* That is, similar patterns represented across model inputs should be indicative of similar outputs. There should exist some measurable relationship between the set of known data values that will be used as model inputs and the resulting output value(s) that the model is tasked to approximate.
- *Relationships that existed in the past will continue to hold in the future such that it is reasonable to use past observations to infer future behavior.*

Predictive modeling, then, is a body of technologies capable of approximating the relationship between known input data measures and some resulting output.

Application Classes

There are generally two classes of predictive modeling applications that differ by the type of output the model produces:

- Forecasting: Forecasting models generate outputs that are continuous-valued. That is, the output should be a value ranging from the minimum to the maximum allowed. These models are used in applications such as forecasting/estimating: sales, volumes, costs, yields, rates, temperatures, scores, etc.
- Classification: Classification models generate outputs that are 1-of-n discrete possible outcomes. Often there is a single output that represents a boolean (i.e., yes/no) outcome. These models are used in pattern recognition applications to do fraud detection, target recognition, vote forecasting, prospect classification, churn prediction, bankruptcy prediction, etc.

Predictive Modeling Algorithms

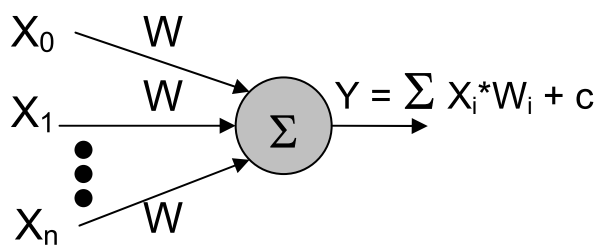
There are generally two classes of modeling architectures: mapping and segmentation. There are a multitude of flavors of each type. Some of the most useful are described below.

Functional Mapping Model Types	Segmentation Model Types
Linear Regression	Decision Tree
Logistic Regression	Cluster-based
Polynomial Regression	Rule-based
MLP Neural Networks	Fuzzy Logic

Mapping Architectures: Functional mapping models generally adjust the coefficients in some continuous equation(s) in order to approximate the relationship between input and output variables. The computational basis can range from simple equations, with a limited number of terms, to complex equation series, with many weights that allow highly non-linear interactions and synergies between variables.

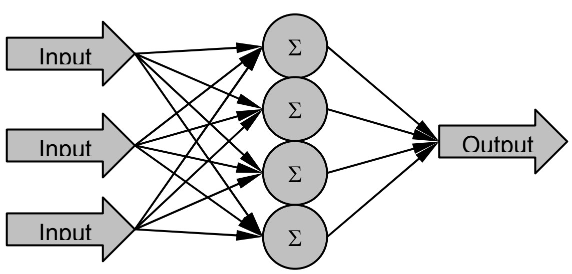
- Regression: Regression is generally the simplest mapping form. Regression models use only one coefficient (i.e., weight term) per input variable.
    - Linear: Linear regression models are widely used. Because there is a single coefficient per input variable they are easily understood. For the

case of a single input variable, the solution is a line (y = X*W + c), for two inputs the solution is a plane (y = X1*W1 + X2*W2 + c), and for a higher number of inputs the solution is an n-dimensional hyperplane (y = X1*W1 + X2*W2 + … + Xn*Wn + c). Linear regression models are adequate for many applicatoins but cannot describe variable interactions and can be highly susceptible to noisy data.



- Logistic: Logistic regression applies the same general equation as linear regression ($y = \Sigma X_i W_i + c$). As a post process, logistic regression applies a logistic function (i.e., s-curve) to the output value of the summation function. This allows the model to better represent binary outputs (logit) and more real-world principles such as 80-20 relationships, diminishing returns, etc.
- Polynomial: Polynomial regression techniques generally pre-process one or more inputs into a quadratic/polynomial form ($y = AX^2 + BX + C$), then solves for the coefficients using linear regression. This form allows inputs to interact and can represent more complex relationships than linear or logistic regression.

- MLP Neural Networks: Multi-layer perceptron (aka, MLP) neural networks have an architecture that allows multiple coefficients per input variable. The processing at each node is functionally equivalent to logistic regression. Multiple nodes (organized into layers) allow the model to represent complex, non-linear, interactions between variables. MLP neural networks have been described as "universal approximators" because they are capable of accurately approximating any functional relationship.



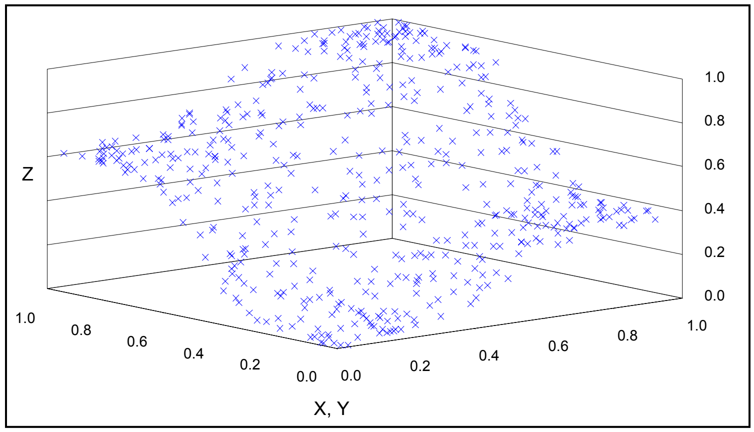
Segmentation Architectures: Segmentation architectures use a variety of methods to segment or "break-up" the solution space into discrete groups of cases.

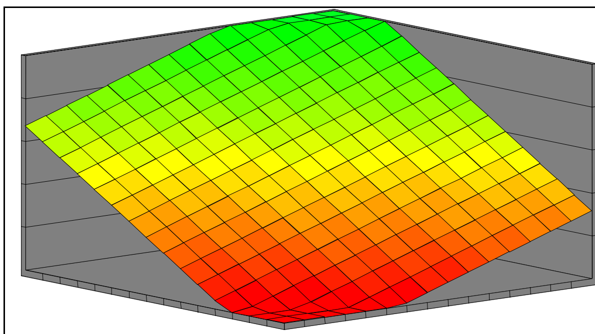
- Decision Tree: Decision trees use hierarchical, joint variable conditions to break-up a solution space into subspaces. These conditional sub-spaces can then be used to classify input patterns. They are called trees due to the hierarchical node/link flowchart graph that is often used to depict the various conditional decision paths.
- Clustering: Cluster models segment a solution space by distributing some number of prototype vectors or cluster centroids across the solution space in order to group similar cases together. An input pattern is classified by mapping it to the cluster centriod that is closest to it. The model's output is inferred based on the known cases that were assigned to the cluster during training.
- Rules: Like decision trees, rule-based methods can break-up a solution space by using If-Then conditional tests. Rule-based methods do not use decision tree's hierarchical conditional checking, but rather incorporate independent If-Then rules to infer an output result.
- Fuzzy Logic: Fuzzy logic segments each variable's range into overlapping membership functions/sets. If-Then rules are then used to test joint variable membership conditions in order to describe model logic. When a fuzzy model is fired, all rules are fired, however, each rule's conclusion is asserted only to the degree that its premise is true. Because membership functions can overlap, fuzzy models can resolve simultaneously conflicting rules.

PREDICTIVE DYNAMIX INC

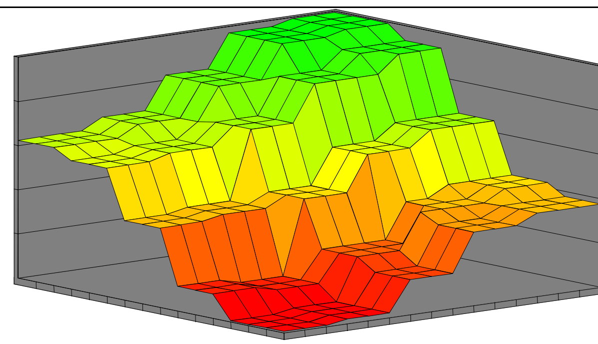
Comparing Modeling Technologies

The following example shows how different model types learn the dynamics of the dataset on a simple 2-input / 1-output problem. The raw data used to build the model is shown in this 3-D scatter plot. The X & Y variables are model inputs and the Z variable is the output value that the model is tasked to predict.

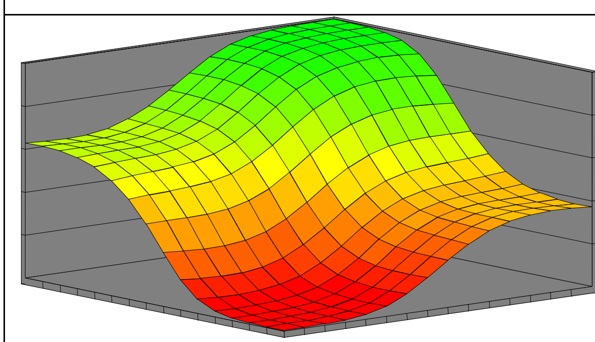




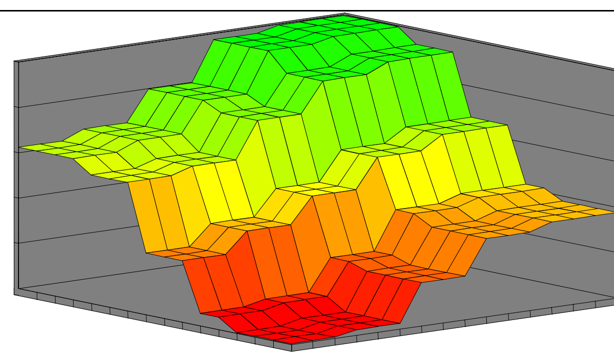
Linear Regression: The model's fit is limited to a plane through the data.



Cluster-based Model (Self-Organizing Map): Cluster centroids can be seen as plateaus on the surface.



Neural Network (MLP): The model is highly accurate due to its ability to represent non-linear, interacting variable relationships.



Decision Tree (CHAID): Rules (i.e., tree leaf nodes) can be seen as plateaus on the response surface.

It is interesting to note how the two segmentation methods (Decision Tree & Clustering) produce very similar results on this dataset.