

Predictive Scoring Overview

Leveraging machine-learning techniques is crucial to efficiently and effectively understand customer data. Marketers who use TD do not need to be familiar with machine learning and data science.

Our predictive scoring enables you to enjoy machine-learning capability in your day-to-day activities with no technical or theoretical expertise. Marketers can predict profile behavior such as who is likely to churn, purchase, click, or convert in the near future.

A predictive model is a set of rules that makes it possible to predict an unmeasured value from other, known values. The form of the rules is suggested by reviewing the data collected. Training is then used to make some predictions. Predictive modeling uses statistics to predict outcomes.

Predictive modeling is a typically used statistical technique to predict future behavior. Predictive modeling solutions analyze historical and current data and the generated model helps predict future outcomes. In predictive modeling, data is collected, a statistical model is formulated, predictions are made, and the model is validated (or revised) as additional data becomes available. For example, risk models can be created to combine member information in complex ways with demographic and lifestyle information from external sources to improve underwriting accuracy. Predictive models analyze past performance to assess how likely a customer is to exhibit a specific behavior in the future. This category also encompasses models that seek out subtle data patterns to answer questions about customer performance, such as fraud detection models. Predictive models often perform calculations during live transactions—for example, to evaluate the risk or opportunity of a given customer or transaction to guide a decision.

Machine learning algorithms learn from data. They find relationships, develop understanding, make decisions, and evaluate their confidence from the training data they're given. The better the **training population** is, the better the model performs. The quality and quantity of your machine learning training data has as much to do with the success of your data project as the algorithms themselves. Your training population is defined by a segment. For example, to train facial recognition, you need a variety of pictures of people from all countries and with all shades of skin. If you were to use only a small and very homogeneous training population, your facial recognition program is not likely to work when you roll it out to a worldwide population.

Scoring is another method used to help you determine how good your predictive model is.

An attribute or feature where the values correspond to discrete categories. For example, *state* is a categorical attribute with discrete values (CA, NY, MA, etc.). Categorical attributes are one of the following:

- non-ordered (nominal) like state, gender, etc.
- ordered (ordinal) such as high, medium, or low temperatures