Print

Explanatory versus Predictive Modeling

Regardless of the statistical model you use, you need to distinguish between explanatory and predictive modeling. In explanatory models, or inferential statistics, you make conclusions or inferences about a population from the analysis of a random sample drawn from that population. So you generalize from the data you observe to the population that you haven't observed. The goal is to develop a model that answers the question, "How is X related to Y?" That is, how does the outcome change as I change the predictor value? In explanatory modeling, you're concerned with accurately estimating model parameters, and you assess this using p-values and confidence intervals. You typically have small sample sizes and few variables.

Predictive modeling, on the other hand, predicts future values of a response variable based on the existing values of predictor variables. It's focused on making accurate predictions. That is, regardless of the parameter estimates, can I still make good model predictions? You assess the prediction's accuracy using a holdout or validation data set, and the model usually has many variables and a large sample size.

In this course, we'll focus predominately on explanatory modeling. However, when you're comfortable with creating these models, we'll show you a few techniques to transition into the world of predictive modeling.

Close