Lesson summary

Module 2 Lesson 2: Modeling to Evaluation

Congratulations! You have completed this lesson. At this point in the course, you know:

- The end goal of the Modeling stage is that the data model answers the business question.
- The data modeling process uses a training data set. Data scientists test multiple algorithms on the training set data to determine whether the variables are required and whether the data supports answering the business question. The outcome of those models are either descriptive or predictive.

Modeling to Evaluation

Unveil the Modeling Process and Assess Performance

Modeling and evaluation play a pivotal role in shaping analytical outcomes and refining problem-solving strategies

1. Aims to determine the characteristics and purpose of the modeling process. Models can be descriptive/predictive. They are based on statistical or machine learning approaches.

2. Training set serves as a gauge to calibrate the model. Experiment with algorithms for variable selection is done. Success depends on problem understanding and analytical approach.

3. Model evaluation is iterative, done alongside model building. It occurs before deployment, assessing model quality and alignment with the initial request.

4. Diagnostic measures ensure intended model functionality. Significance testing interprets data. ROC curve aids optimal classification selection by quantifying performance.

IBM DATA SCIENCE METHODOLOGY Skills Network

- The Evaluation phase consists of two stages, the diagnostic measures phase, and the statistical significance phase.
- During the Evaluation stage, data scientists and others assess the quality of the model and determine if the modelanswers the initial Business Understanding question or if the data model needs adjustment.
- The ROC curve, known as the receiver operating characteristic curve, is a useful diagnostic tool for determining the optimal classification model. This curve quantifies how well a binary classification model performs, declassifying the yes and no outcomes when some discrimination criterion is varied.

Author(s)

Dr. Pooja
Patsy R. Kravitz

Changelog

Date	Version	Changed by	Change Description
2023-08-09	0.1	Patsy R. Kravitz	Initial version created
2023-08-14	0.2	Dr. Pooja	Infographic included