PetroSkills

Supervised Machine Learning

MODULE

About the Skill Module

This skill module introduces supervised learning as a key type of machine learning that drives data-driven analysis across economic sectors and impacts the experiences of consumers. The skill module focuses on the emerging uses of supervised learning in the oil and gas industry as an important complement to other forms of analysis, as well as subject matter expertise, in solving diverse problems and providing reliable data streams.

The skill module begins by placing supervised learning among the three forms of machine learning and explains its distinguishing qualities. The two key forms of supervised learning, regression and classification, are examined in detail with diverse examples from daily life to the technical work in the oilfield. The skill module discusses how supervised learning models are trained to fit data and subsequently validated for deployment. The validation procedure discusses procedures to balance model complexity and model predictability to avoid overfitting and to obtain optimal model performance. The skill module discusses data pre-processing steps including exploratory data analysis, scaling, and an assessment of correlation. Significant emphasis on the appropriate choice of performance metrics for regression and classification problems is also provided. Finally, the skill module reviews emerging uses of supervised learning in the oilfield. A case study approach shows basic and more complex applications, including studies from leading experts in the field.

Target Audience

Exploration geologists, geophysicists, engineers, and geoscience managers

You Will Learn

Participants will learn how to:

- Distinguish between two forms of supervised learning: regression and classification
- Recognize use cases for regression and classification
- Identify why an iterative approach is essential in supervised learning
- Recognize covariance and correlation as key aspects of data pre-processing, and track their importance for supervised learning
- Recognize a generalized workflow for supervised learning
- Identify and explain the steps involved in exploratory data analysis
- Recognize the need for, and some of the nuances involved in, handling outliers
- Recognize the purpose of scaling
- Distinguish between Standard and Min-Max scaling
- Identify how to apply both the Standard and Min-Max methods
- Recognize how performance metrics are used to evaluate regression models
- Build awareness of the need to fit models and metrics to the specifics of datasets and data problems
- Identify approaches to more complex cases of model evaluation
- Recognize that there is no universal algorithm that can be effectively used to evaluate machine learning models
- Describe how to determine training and testing sets from a single dataset
- Examine method for overcoming overfitting
- Examine validation techniques including 3-fold cross-validation and K-fold cross-validation
- Follow and explain an end-to-end workflow for regression
- Identify the purpose of non-parametric regression
- Follow and explain an end-to-end workflow for classification problems
- Follow the workflow for several use cases involving supervised learning in the oilfield

Product Details

Categories: Upstream

Disciplines: Data Management, Science and Analytics

Levels: Basic

Product Type: Individual Skill Module

Format: On-Demand

Duration: 4 hours (approx.)

$395.00