

Applied Statistical Modeling and Big Data Analytics

Course Description

As companies strive to deliver value from “Big Data”, we need to extract as much intelligence as possible from our ever-expanding trove of static and dynamic information to improve operational efficiencies and make better exploration, production, and reservoir management decisions. This 3-day course provides the theoretical background and hands-on problem-solving practice to understand and apply fundamental concepts of classical statistics, as well as emerging concepts from data analytics. Attendees will receive a course notebook and digital reference material.

Course Outline

- 1. Foundational Concepts:** Big data analytics, machine learning and artificial intelligence concepts; Data, statistics, and probability; Distributions; Confidence intervals
- 2. Basic Regression Analysis:** Linear regression; Understanding regression statistics; Non-parametric regression
- 3. Multivariate Statistics:** Dimension reduction; Cluster analysis; Data visualization
- 4. Machine Learning Basics:** Overview of techniques; Evaluating model performance; Variable importance; Model aggregation
- 5. Machine Learning for Regression and Classification:** Classification/regression trees; Random forest; Gradient boosting machine; Support vector machine; Neural networks; Deep learning.
- 6. Miscellaneous Topics and Wrap-up:** Experimental design and response surface analysis; Uncertainty quantification; Selected literature review; Key takeaways and resources; Data analytics dos and don'ts

Who Should Attend

This course is designed for petroleum engineers, geoscientists and other disciplines interested in becoming informed users of statistical modeling and data analytics in the E&P realm, and efficiently interacting with data scientists to develop practical data-driven applications for their projects.

Recipients of Training

This course has been taught over 30 times in the last 5 years in open-enrollment courses, at professional society meetings, and as customized internal courses at companies and universities.

About the Instructor

Dr. Srikanta Mishra is a Senior Research Leader and Discipline Lead (Reservoir Sciences and Data Analytics) for Battelle who has over 30 years of subsurface flow modeling and geoscience data analytics experience. He served as an SPE Distinguished Lecturer on Big Data Analytics during 2018-19 and is the author of a recently published book on statistical modeling and data analytics. He holds a PhD in Petroleum Engineering from Stanford University.