N759: Applied Data Science and Machine Learning for Upstream Professionals

Format and Duration
Classroom - 5 Days

Instructor(s): Vitali Molchan

Summary

This course provides Upstream Professionals with a structured and practical introduction to machine learning – combining clear theoretical foundations with hands-on data science skills directly applicable to subsurface workflows. You will build confidence from the ground up, learning each concept in simple, accessible terms before applying it step-by-step to real challenges in reservoir engineering, production engineering, geomodelling, and petrophysics. The skills and code examples you develop can be used immediately in your current and future projects, enhancing your technical workflows, decision-making capabilities, and overall performance.

Learning Outcomes

Participants will learn to:

1. Build confidence in core machine learning and data science concepts, tailored specifically to Exploration, Production, and other Upstream Oil & Gas domains.
2. Evaluate and select the most appropriate analytical workflows and machine learning algorithms for addressing domain-specific challenges.
3. Gain hands-on experience in implementing end-to-end solutions using modern open-source tools and Python libraries relevant to Upstream applications.

Training Method

This is a classroom course comprising a mixture of lectures, discussions, case studies, and practical exercises.

Who Should Attend

A reservoir engineer, geologist, petrophysicist, or production engineer with 5+ years of experience who wants to build strong foundational and practical skills in scientific programming, data science, and machine learning.

Course Content

Day 1

Lectures

- Introduction to Data Science in the Oil & Gas Industry.
- Introduction to Python:
  - Python basics.

  - Data Wrangling (using Pandas and SQL).
  - Data visualisation.

Daily Learning Outcomes

- **Solve upstream challenges with Python** – Confidently apply Python and leading machine learning libraries to tackle problems in reservoir engineering, production analysis, and field optimisation.
- **Build scalable data workflows** – Develop reusable pipelines for analysing production data from both local files and databases, applicable to small and large oil & gas fields.
- **Prepare complex reservoir data efficiently** – Rapidly transform production and pressure data from multi-compartment reservoirs for material balance analysis, formatted for PETEX MBAL.
- **Accelerate reservoir simulation analysis** – Streamline the review of large numbers of simulation results to quickly assess history match quality and forecast reliability.
- **Create impactful visualisations** – Produce high-quality plots of production, pressure, and well log data suitable for technical reviews, presentations, and reporting.

Case Studies and Exercises

- Production data analysis – Norne Field, The North Sea.
- Production data analysis and visualisation.
- Data preparation for material balance calculation (PETEX MBAL).
- Reservoir simulation model QC.

Day 2

Lectures

- Numerical optimisation.
- Exploratory Data Analysis.
- Uncertainty Evaluation and Decision Making.

Daily Learning Outcomes

- **Apply numerical optimisation to complex tasks** - Use a range of optimisation techniques to fit rate-time data and pressure gradients, improving reservoir depletion analysis and enabling consistent initialisation of reservoir simulation models.
- **Upscale static models intelligently** - Convert fine-grid geological models into simulation-ready coarse grids, with full control over upscaling parameters to balance dimensionality reduction and geological details preservation.
- **Estimate volume-in-place** - Perform probabilistic volume-in-place assessments that incorporate input uncertainty, enabling quick volumetric estimates without building a full geological model.

- Optimise water and gas injection allocation - Determine the most efficient distribution of water and gas injection volumes across injection wells to maximise oil recovery, using a minimal number of targeted simulation runs using experimental design and surrogate models.

Case Studies and Exercises

- Decline Curve Analysis (DCA).
- PVT data preparation for reservoir simulation.
- Probabilistic volume-in-place estimation.
- Static model upscaling.
- Waterflood optimization.

Day 3

Lectures

- Introduction to Machine Learning.
- Clustering methods.
- Anomaly detection methods.
- Dimensionality reduction methods.

Daily Learning Outcomes

- Apply ML concepts in practice - Confidently use machine learning terminology and translate business and technical requirements into an actionable project scope.
- Select fit-for-purpose ML techniques - Choose the most appropriate algorithms based on problem type, data availability, data quality, and performance requirements.
- Streamline reservoir model evaluation - Screen static model scenarios to simplify history matching, reduce simulation runs, and quantify the impact of geological uncertainty.
- Optimise electrofacies definition - Determine the optimal number of electrofacies for realistic property distribution in reservoir models.
- Automate production well monitoring - Develop fully automated workflows to detect abnormal behaviours in production wells, helping prevent production losses and reduce environmental risks.

Case Studies and Exercises

- Electrofacies identification based on well log data.
- Screening of static model realisations.
- Detection of abnormal behaviours in production and injection wells.

Day 4

Lectures

- Core machine learning concepts.
- Regression methods.
- Machine learning model turning.

Daily Learning Outcomes

- **Design robust ML studies** - Structure and execute machine learning projects with high reproducibility, clarity, and model quality.
- **Apply essential ML concepts** - Master train/test split, cross-validation, objective functions, bias–variance trade-off, and hyperparameter tuning through hands-on exercises.
- **Predict and optimise well performance** - Forecast new well productivity and optimise completion design in unconventional reservoirs without physics-based simulation.
- **Develop data-driven fluid property models** - Build accurate ML models to predict saturation pressure using available field data, even when full PVT experiments are lacking.
- **Automate ML model hyperparameter tuning** - Use modern tools to efficiently identify optimal model configurations and reduce manual experimentation time.

Case Studies and Exercises

- Production forecasting for unconventional reservoirs.
- Saturation pressure prediction.

Day 5

Lectures

- Classification methods.
- Neural networks and Deep learning.
- Advanced machine learning topics:
    - Imbalanced datasets.
    - ML model interpretability.

Daily Learning Outcomes

- **Communicate results effectively** – Explain machine learning outputs clearly to both technical and business audiences to support QA/QC and decision-making.
- **Classify lithofacies from well logs** – Develop robust ML-based lithofacies classification models to extend interpretation to wells without core data.

Case Studies and Exercises

- Lithofacies identification.
- Production forecasting for unconventional reservoirs.