CEGA
Canadian Energy Geoscience Association

Short Course

Data Science for Geoscientists

Instructor: Ryan Mardani, Data Energy

Location: CEGA Classroom, +15 level, 540-5 avenue SW, Calgary

June 27-28, 2023, | 8:00am-4:00pm (MST)

This course requires a laptop, windows operating system preferred, though MAC is accepted. Free software download and support will be available one week prior to the course.

COURSE OVERVIEW

Introduction to Python

- Variables and Data Types
- Operations
- Python Data Collections and Structures
- Loops and Functions
- Generate Synthetic Seismogram

Data preparation with Pandas

- Data cleaning
- Null elaboration
- Merging datasets

Data Analysis with Pandas

- Data statistics and correlation
- Data subsetting, filtering, aggregation, and pivoting
- Data visualization, plot histograms and KDE, box plots, …

Machine learning (Model Construction and Evaluation)

- Well log and lithology Data preparation for ML algorithms
- Feature engineering
- Classification and Regression problems
- Select and train model with cross validation
- Fine tune model with hyperparameter adjustments
- Evaluation metrics and model performance analysis

In this workshop, examples will be mainly from geoscience like rock facies and well logs prediction.

COURSE OBJECTIVE

Python is a popular programming language, easy to learn, has a readable code structure, and comes with a lot of powerful libraries. You will use Python to elaborate datasets with oil and gas origins. Data analysis and visualization will be implemented with Pandas library.

CEGA
Canadian Energy Geoscience Association

This course is designed for geoscientists and subsurface engineers to develop machine learning algorithms to learn from data then estimate your favorable variables like lithology or missing logs.

By the end of the course, learners will be able to:

- Implement and analyze learning algorithms for classification and regression
- Implement practical steps: data preprocessing, learning, regularization, and model selection
- Compare and contrast different paradigms for learning
- Design experiments to evaluate and compare different machine learning techniques on geoscientific problems like lithofacies prediction from well logs.

WHO SHOULD ATTEND

This is an interdisciplinary course introduces data science application on geoscientific problems. This course is ideal for professionals and researchers of oil and gas domain with minimum background knowledge of geoscience. We encourage geologists, geophysicists, petrophysicists, reservoir engineers, and managers to attend the course to experience digital transformation on geoscience domain data. You don't need to have solid knowledge of Python programming. We will provide codes, scripts and datasets required for the course.

BIOGRAPHY

Ryan A. Mardani is a Geo Data Scientist who has over 18 years of industry experience as a geoscience interpreter and software trainer and developer. He has worked on conventional and unconventional resources in the Middle East, Africa, and Canada. In addition to conventional integrated reservoir studies, he helped the oil and gas, energy, and mining companies to apply machine learning and data science techniques to put their valuable data to more efficient use.

Ryan is currently the Senior Consultant for Data Energy where he automates geoscience tasks, analyze relational and non-relational datasets and extract business values by combining conventional approaches with machine learning algorithms. He builds intelligent machines to learn from GeoData to discover patterns and trends embedded in vast volumes of data to improve data driven decisions.

He is also adjunct instructor, teaching database and data analytics courses in Southern Alberta Institute of Technology (SAIT) for corporate division.