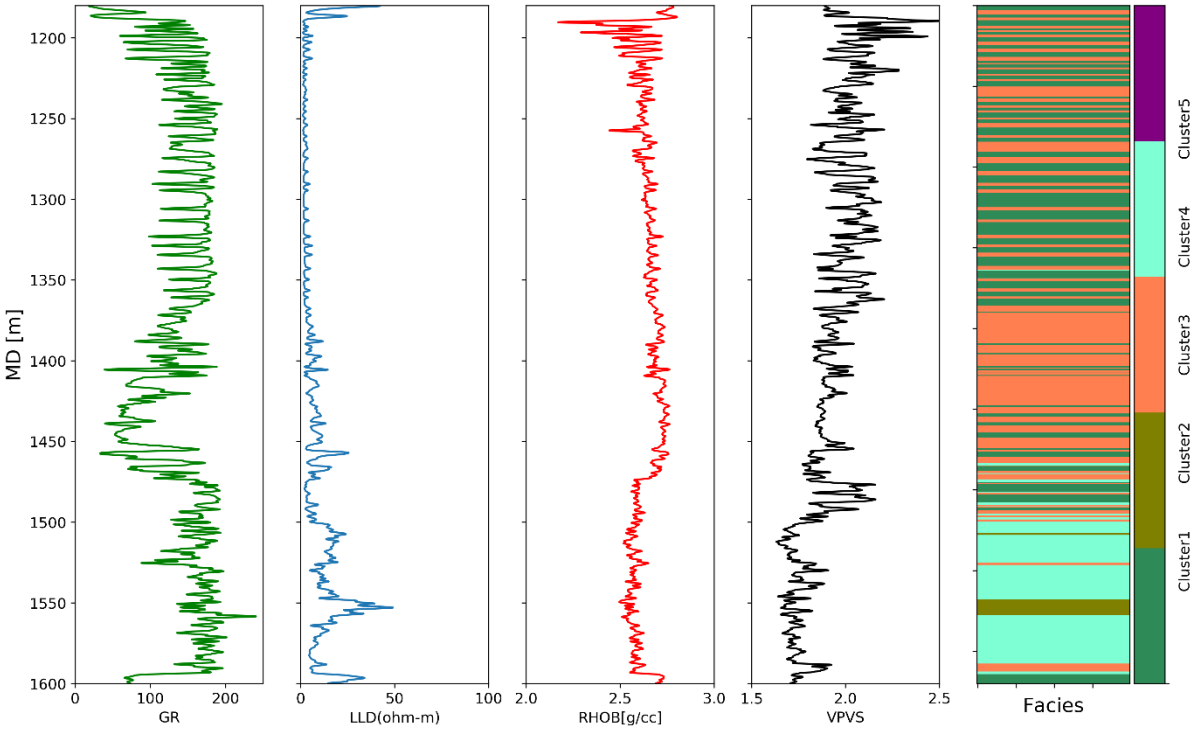
ML based facies prediction from wireline logs

Definition and scope of work

Winter Semester | 2024-25

Instructor – Prof Partha Pratim Mandal



Introduction

Welcome fellow students. I will be interacting with you over the next 14 weeks in the classroom when delivering Well logging & Electrofacies analysis unit. In part of learning this unit, you will be going to work in a group format case study to define continuous electrofacies profile of subsurface rock using machine learning based classification technique. Both supervised and unsupervised approach will be tested and validated. Either python programming or Matlab skills are essential to complete the task. Details about the workflow and scope is defined below.

Case study description

The dataset of this group case study came from a class exercise from The University of Kansas on Neural Networks and Fuzzy Systems. This exercise is based on a consortium project to use machine learning techniques to create a reservoir model of the largest gas fields in North America, the Hugoton and Panoma Fields. For more info on the origin of the data, see Bohling and Dubois (2003) and Dubois et al. (2007).

The dataset consists of wireline logs from nine wells that have been labelled with a facies type based on observation of core. Purpose is to use this well logs data to build continuous electrofacies then compare with defined lithofacies. This can be done in different ways such as statistical distribution of dataset (mean, median, ranges), or available classification techniques.

Later, the group will be going to use few machine learning models that can be trained on available data samples to perform supervised and unsupervised classification tasks. You are free to choose any existing algorithm from various platform, for example scikit-learn.

First, the team will explore the dataset. Then load the training dataset from 9 wells and find relevant input features. You should plot the data from a couple wells and create cross plots to look at the variation within the data.

Install python package via anaconda platform.

https://www.anaconda.com/products/individual

https://scikit-learn.org/stable/index.html

Following steps may be followed to deliver the outcome. You can bring your own idea during implementation of the project workflow.

Step 1: Explore the dataset – create histogram of each log curves, statistical distribution, identify distribution of data, find out possible outlier, etc. Most important is to investigate the key wireline logs from each well and their relationships. Bring your own thinking to describe and visualize it.

Step 2: Generate missing logs which may be beneficial to identify rock types, for example in this caser PE log is not available in three wells.

Step 3: Prepare input wireline log curves for supervised as well as unsupervised classification. For supervised approach, you need to compare predicted facies with interpreted lithofacies. Sometime number of input logs are in number which may not require to define data groups. It is possible to

reduce number of input logs through dimensionality reduction. K-means clustering is a popular clustering technique. Implement K-means clustering with available logs such as GR, resistivity, density, porosity, and PE to find optimum Electro-facies.

Step 4: Visualize electro-facies logs from both supervised and unsupervised classification methods and find any relevance with already supplied lithofacies across wells. Make your observation and findings.

Step 5: Create 12-15 slide presentations and present the outcome. **Date:** 2nd week of April 2025 (12th April from 10 – 12 noon; viva: 4 to 6 pm)

Case Study Presentation Guidelines

Each group is required to prepare a professional PowerPoint presentation. The presentation should be 15 minutes in length (each student ~4 minutes) and 5 minutes for QA. All group participants should take an active role in this case study.

Marking scheme for group presentations

Content: (50%)

- Is the group presentation well-organized and interesting?
- Has the group answered the case study questions and present them clearly?
- Have they given relevant concepts/theories/ideas and knowledge to the questions?

Presentation Style: (10%)

- Is the presentation lively and interesting?
- Is it structured effectively with a clear introduction, results, discussion, conclusion supported by evidence?
- Have the group shown initiative and creativity in the design of the presentation?
- How well do the presenters present themselves? Voice projection, eye contact, confident delivery, and interactions?
- How well prepared are the group to answer or pose questions that are relevant to the topic?

Group skills: (10%)

- Is the presentation clearly an integrated group effort as opposed to individual contributions?
- How well have they co-ordinated their activity and planned their presentation?

Individual viva: (30%)

- Individual viva to understand concept on the case study and their contribution

Marking of the presentation (Total: 30, converted to 20 scale)

Marking total -20	Content - 15	Presentation style - 5	Group skills - 5	Viva - 10
Group-1				
Group-2				
Group-3				
Group-4				
Group-5				