Lithology Classification based on Machine Learning

Wang Yingbo

2018-5-18
Overview

1. Introduction
2. Lithology classification based on raw data analysis
3. Lithology Classification using Supervised Learning
4. Lithology Classification using Unsupervised Learning

Goal

To achieve auto-classification of lithology using machine learning
Why do we employ ML to conduct lithology classification

1. The data volume from the field acquisition is ever-increasing due to the more advance acquisition tools and modern acquisition technology.

2. It becomes a challenge for data processor to interpret geologic structure within a limited time.

3. Machine learning can explore the hidden connection among different physical quantities.
Machine Learning (ML)

ML is a field of computer science that gives computers the ability to learn without being explicitly programmed (Samuel, 1959)
Well Log Data

Figure 1. Plot of well log data corresponding to the depth
Lithology classification based on raw data analysis

Since shale usually contains bound water, it is not reliable to use density or porosity to divide it from sand. Therefore, we normalize the Gamma ray data and Vp/Vs. Since shale always shows high Vp/Vs and high gamma ray. The layer (normalized Vp/Vs>0 and normalized Gamma ray>0.5) is judged as shale layer.

Student classification: This is a standard way of log analysis.
k-nearest Neighbors Algorithm (KNN) — Supervised Learning Method

In the classification phase, k is a user-defined constant, and an unlabeled vector (a query or test point) is classified by assigning the label which is most frequent among the k training samples nearest to that query point.
Lithology Classification using Supervised Learning

Therefore, similarly, considering the similar lithology has difference property, we input Vp, Vs, density, porosity, gamma into the model and use 80% labeled data to train the model. (k=5) (Cross-validation)
Parameter study using Supervised Learning

3P
- Vo
- Vs
- Gamma

KNN
- Sand
- Shale

5P
- Vp
- Vs
- Density
- Porosity
- Gamma

KNN
- Sand
- Shale

20% Test Set
80% Training Set
Cross-validation
Lithology Classification using Unsupervised Learning

1. Cluster: Automatic classification by computer
2. Cross-check with student classification
Parameter study using Unsupervised Learning

2P
- **Vo/Vs**
- **Gamma**

Competitive learning

- **Sand**
- **Shale**

3P
- **Vo**
- **Vs**
- **Gamma**

Competitive learning

- **Sand**
- **Shale**

5P
- **Vo**
- **Vs**
- **Density**
- **Porosit**
- **Gamma**

Competitive learning

- **Sand**
- **Shale**
Result

Student classification

Competitive Learning

KNN

Sand

Shale
Manmade

Competitive

KNN

Sand

Shale
Student Classification
Competitive learning ----5P
Competitive learning ---- 3P
Competitive learning ----2P
KNN---5P
KNN----3P
Conclusions

1. The ML supervised learning and unsupervised learning provide similar classification results with human interpretation

2. The parameter study is needed when multiple sets of data are acquired

3. Unsupervised learning shows promises for automatic log analysis
Next step

We have already built a DNN model (Deep Neural Network) using Tensorflow with 3 hidden layer (10, 20, 10)

And we get the data from another well log (with 10000+ usable data)
Thank you!