Abstract

The shear wave velocity (Vs) is widely used as quick, easy to use, and cost-effective means of determining the mechanical properties of formations in the oil and gas industry. However, shear wave logs are only available in a limited number of wells in an oil field due to the high cost of the log acquisition and the associated technical difficulties in the past. For this reason, many attempts have been made in myriad literature to find a correlation between Vs and other petrophysical logs. In this study, a set of log data consisting of depth, neutron porosity (NPHI), density (RHOB), photoelectric (PEF), gamma ray (GR), caliper, true resistivity (RT), VP, and Vs were used to develop models for Vs estimation in five unconventional wells drilled in the tight gas formation and naturally fractured reservoir of the Ahnat Basin, Algeria. The focus of this paper is to apply machine learning algorithms to synthesize the shear velocity log. Five different algorithms were developed and evaluated against the test data, namely: Gradient Boosting regressor, Random Forest Regressor, Linear Regression, AdaBoost Regression, and Support Vector Regression. Overall, the results showed the R² -score varied from -1.24 to 0.97, with the Random Forest regressor outperforming the other algorithms. The work presented in this paper elaborates a roadmap to accurately approximate and synthesize missing data using machine learning algorithms and extend it across the rest of the wells to compensate the lack of data. It is hoped that the results of this study can improve the elaboration of the shear velocity in the wells.
Machine Learning Application for Shear Velocity Prediction in Ahnet Basin, Algeria

Abdeldjalil Latrach¹, Nassim Bouabedallah¹, Almene Aihar²
¹University of Wyoming, Department of Energy and Petroleum Engineering
²University of North Dakota, Department of Petroleum Engineering

Abstract

The shear wave velocity (Vs) is widely used as a quick, easy to use, and cost-effective means of determining the mechanical properties of formations in the oil and gas industry. However, shear wave logs are only available in a limited number of wells in an oil field due to the high cost of the log acquisition and the associated technical difficulties in the past. For this reason, many attempts have been made in myriad literature to find a correlation between Vs and other petrophysical logs. In this study, a set of log data consisting of depth, neutron porosity (NPHI), density (RHOB), photoelectric (PEF), gamma ray (GR), caliper, true resistivity (RT), VP, and Vs were used to develop models for Vs estimation in five unconventional wells drilled in the tight gas formation and naturally fractured reservoir of Ahnet Basin, Algeria. The focus of this paper is to apply machine learning algorithms to synthesize the shear velocity log. Five different algorithms were developed and evaluated against the test data, namely: Gradient Boosting regressor, Random Forest regressor, AdaBoost Regression, and Support Vector Regression. Overall, the results showed the R² score varied from -1.24 to 0.97, with the Random Forest regressor outperforming the other algorithms. The work presented in this paper elaborates a roadmap to accurately approximate and synthesize missing data using machine learning algorithms and extend it across the rest of the wells to compensate the lack of data. It is hoped that the results of this study can improve the elaboration of the shear velocity in the wells.

Objectives

The objectives of this study are as follows:

• To compare the performance of five different machine learning algorithms, including Gradient Boosting regressor, Random Forest regressor, Linear Regression, AdaBoost Regression, and Support Vector regressor.
• To demonstrate the feasibility of using machine learning algorithms to synthesize missing data and extend it across the rest of the wells to compensate for the lack of data.
• To emphasize the potential of this study to improve the elaboration of the shear velocity in the wells, and ultimately enhance the accuracy and efficiency of petroleum exploration and production processes.

Materials and methods

Data collection

The dataset was collected from 5 wells in Ahnet basin, Algeria. The data consisted of a total of 8465 rows, and 5 different input variables, representing different well logs:

- Caliper
- Gamma ray
- Photoelectric factor
- Porosity
- Bulk density

Four of the wells were used for training, with one of the wells was reserved for testing the machine learning models.

Exploratory data analysis

Preliminary data analysis was carried out to provide some insights into the dataset of interest. A correlation analysis of the independent and dependent variables resulted in the following correlation matrix. The variable of interest for the prediction, the shear velocity, is highly correlated with porosity, photoelectric factor and to a lesser degree with Gamma ray, with correlation being calculated using Pearson’s correlation coefficient. This is a good indication that our predictive variables are very indicative of our predicted variable, which is a good sign that machine learning models will have high predictability power.

Conclusion

This work proves the possibility of synthesizing shear velocity log from conventional logs, which eliminates the need for additional and expensive operations to measure the log directly. The best performing model achieved predictions with R² score of 0.97, which are near-perfect predictions.

References

