Petrophysical log recovering and core-log calibration using Machine Learning

Yexin Liu
SoftMirrors Ltd.

Summary

When we studied petrophysical logs for comprehensive processing, seismic inversion and seismic velocity building, the logs are crucial to estimate the reservoir properties, such as porosity and saturation, and to generate the seismic synthetic from density and sonic logs. Unfortunately, some logs are often missing or some logs are poor quality or some petrophysical data need to be calibrated. Usually, we have to recover these logs from nearby boreholes or the same borehole with different depth range. Also, if we need to make sure the petrophysical results are reasonable, the core analysis data or other lab measurement data are necessary to be used to calibrate the results. In this study a Bayesian-based Support Vector Machine (SVM), one of Machine Learning methods, will be studied to recover these logs and/or make the calibration using all available information. SVM has proven to successfully perform the regression problem using a small dataset and achieve a global optimization result with small errors between the known data and predicted results if the input data have good quality. Geologically, petrophysical logs can be predicted from other logs in the same wells or nearby borehole logs due to similar response from the same geological formation and stratigraphy. Our studies demonstrated that SVM has the capability of recovering the missing logs and making the calibration using very limited data.

Bayesian-based Support Vector Machine (SVM) method

Due to its excellent performance in dealing with small training dataset, the Bayesian-based Support Vector Machine (SVM) has been widely used in the Machine Learning regression community (Cortes and Vapnik, 1995; Wenzel, 2017). We (Liu and Sacchi, 2003, Liu, 2017, Liu, 2018 and Liu 2019) also presented the use of SVM methods for seismic petrophysical property inversion, such as porosity and density prediction, to incorporate seismic attributes, petrophysical properties and geological knowledge.

Given a set of well log vectors $\{x_{n,}, n = 1, \dots, N\}$ along with the corresponding target petrophysical results $\{t_n, n = 1, \dots, N\}$. The SVM makes the training and prediction of the petrophysical data based on a function of the form:

$$t(x_j) = \sum_{n=1}^{N} \omega_n K(x_j, x_n) + \omega_0 \qquad (1)$$

Where $\{\omega_n \; n = 1, \dots N\}$ are the model weights which can be estimated after SVM training, and $\{K(x_j, x_n\}$ denotes the kernel functions (Schölkopf, 1999).

SVMs are known to achieve the best performance for small dataset regression problems. But if there are large dataset, the method of inducing point Gaussian Process (Wenzel et al, 2017; Liu, 2019) can be used to extend SVM to practically handle the big data issue. Unlike traditional Neural Networks, SVM can effectively handle the over-fitting issue to build a best model or relationship to make the prediction.

Missing log recovering

In the petrophysical study, we often face the challenge to recover the missing logs to fill the gaps within the interest of target formation. For example, in order to generate the seismic synthetic for seismic inversion, the sonic log data are necessary. Unfortunately, the sonic logs are often not available in the target intervals. Based on our study, we found that a workflow can be used to recover the sonic log using a Machine Learning method. The workflow is to build the relationships between sonic log with other available well logs, such as density and gamma ray logs, using the SVM equation 1. In eq.1 $t(x_j)$ is the missing target log and $K(x_j, x_n)$ is the kernel functions to represent the non-linear relationship among the logs x_j. ω_n is the model weights estimated after SVM training. In our case study, we selected the well logs within several selected intervals in the same borehole. For example, the well logs within the two-color horizontal bars (green and red) (figure 1a) will be used for training and the target missing well log within another color (blue) horizontal bar (figure 1a) will be predicted using the training model from the two-color horizontal bars in figure 1a. In this case study, the training attributes of well logs consist of **gamma ray, density, neutron and measured depth** and the training target is the **sonic log**.

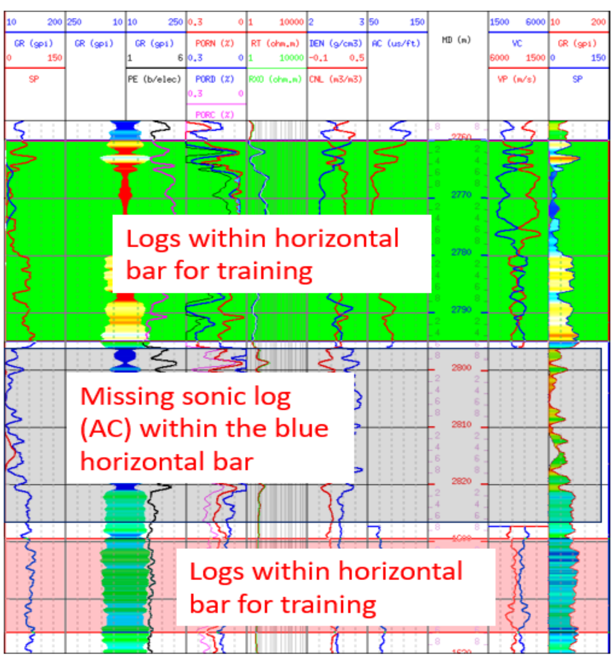

Figure 1a: Training horizontal bars and missing sonic log

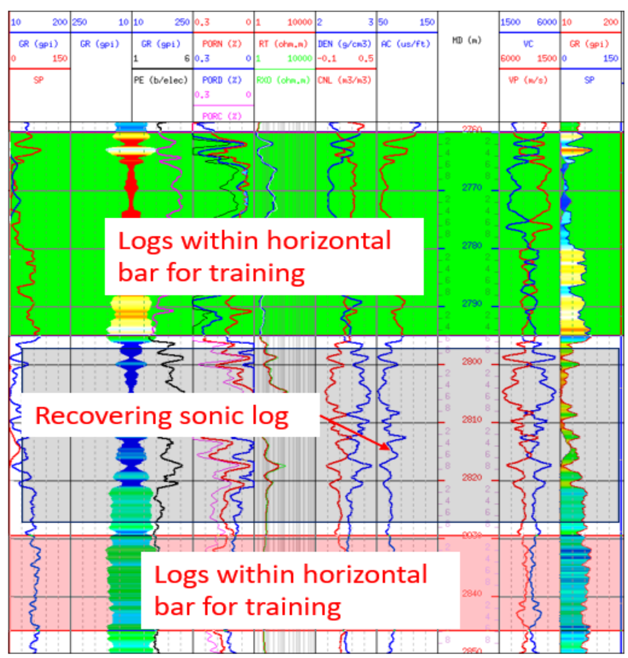

Figure 1b: Recovering the missing sonic log using SVM

After SVM training, we can build a high dimension non-linear relationship between these well logs and sonic log. In the predicted phase, we can apply the relationship to recover the missing sonic log. The figure 1b is the recovering result to fill the gap with the missing sonic log from 2796m to 2826m.

If the training logs are not available in current study borehole, it is possible to generate the training samples from nearby borehole, but it is best to use the same geological formation data due to the similar well logging response.

Core-log calibration

Based on petrophysical study, we can estimate the porosity and saturation during traditional well evaluation. In unconventional shale plays, it is possible to mathematically calculate Total Organic Content (TOC) from well logs. But all these results are necessary to make the calibration using the core data or other lab measurement data in order that the results can be geological meaningful or geological reasonable. Because of limited core data, the SVM method is one of best solutions to build the relationship between well logs and core data to make the calibration. We can select well logs from well evaluation results as the training input samples and the associated training target data is core data, such as porosity, saturation or TOC. During the training phase, a best relationship between well logs and core data (eq.1) can be generated using SVM and then we can make the calibration for the target data using the relationships. There is difference between core-log calibration and missing log recovering. Unlike recovering the missing log, in the calibration workflow, the target data, such as porosity and TOC, can be selected as the input attributes and the new target data calibrated using core data will be outputted. In the figure 2a, the track 1 and 2 are gamma ray (GR), track 3 is density, sonic and neutron porosity and track 4 is resistivity. The track 5 and 6 are density (DEN), neutron (CNL) and sonic (AC). The Track 9 is the water saturation (SW, green) evaluated from well logs and the red stars represent the core saturation analysis. It sounds there are difference between the core data and well evaluation results. In this case for water saturation calibration, the input training attributes for SVM include **porosity (PORT), volume of shale (VCL) and water saturation from well evaluation** and the training target is the **water saturation from core analysis** and **water saturation from well evaluation** if needed.

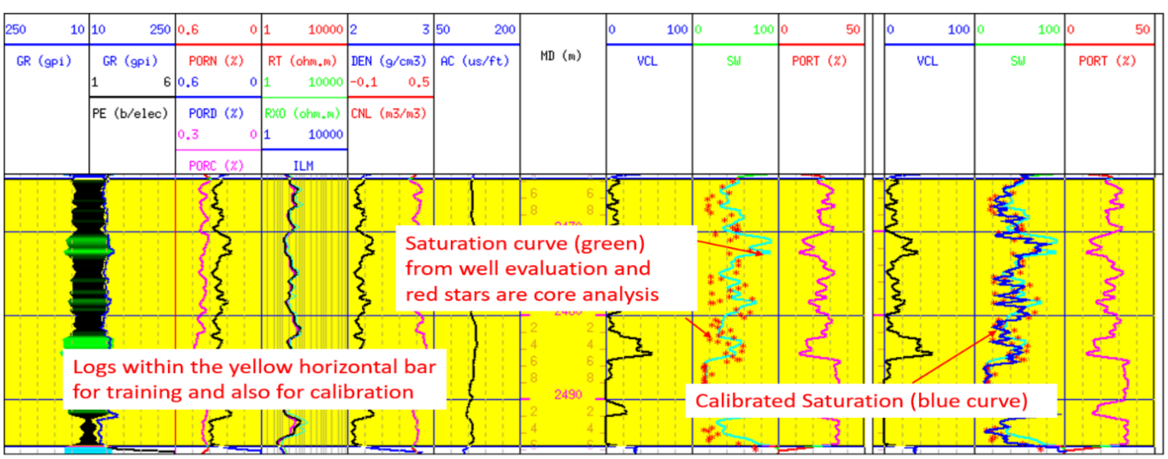

Figure 2a: Training well logs within horizontal bar (yellow)

Figure 2b: calibrated saturation (blue curve)

During the predicted phase, the predicted target output will be the **calibrated water saturation.** Using both core data and well logs from the horizontal bar intervals (yellow) for SVM training, then we can make the calibration for saturation. The track 2 in the figure 2b demonstrated the new calibrated saturation (blue curve) and the original saturation (green curve) before calibration. The red stars are the core saturation results. It sounds the calibrated saturations are consistent with the core data and well evaluation results.

Conclusions

Bayesian-based SVM has been studied to recover the missing logs from available well logs and to make the calibration of the well evaluation results from core data. Based on the case studies, it is possible to use SVM to fill the gaps to recover the missing logs, such as sonic and density logs. If the core data are available it can be used to calibrate the well evaluation results, such as porosity, saturation and TOC, so as that the evaluation results are geological meaningful and consistent with the core and other lab measurement data. SVM is one of best Machine Learning methods to perform regression problems to achieve best performance with a small dataset for training and prediction compared with other methods, such as Neural networks and Deep Learning.

References

Cortes, C. and Vapnik, V., 1995, Support-vector networks, Machine Learning, 20, 273-297

Liu, Y., 2019, A comparison of Machine Learning methods for seismic inversion to estimate velocity and density , Geoconvention 2019, Canada

Liu, Y., 2018, Machine learning to enhance the vertical resolution of seismic geostatistical inversion**,** Geoconvention 2018, Canada

Liu, Y., 2017, Applications of Machine Learning for Seismic Quantitative Interpretation, Geoconvention 2017, Canada

Liu, Y., Sacchi, M., 2003, Propagation of borehole derived properties via a Support Vector Machine (SVM), CSEG Recorder, Canada, December 2003, 54-58

Schölkopf, B., Burges J., Smola, A., 1999, Advances in kernel methods: Support vector learning: MIT Press.

Wenzel, F., Galy-Fajou, T., Deutsch, M., and Kloft, M., 2017, Bayesian nonlinear support vector machines for big data, in Proceedings of the European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases, 2017.