

Applications of Machine Learning for Seismic Quantitative Interpretation

Yexin Liu*, *SoftMirrors Ltd.*; yexinliu@softmirrors.com

Summary

Machine Learning has become popular in regression and classification, especially in image identification and data regressions. In this study, firstly, in order to honour seismic attributes and petrophysical properties, and most importantly to add the domain knowledge to the workflow, a Bayesian-based Support Vector Regression (SVR) workflow for seismic Quantitative Interpretation (QI) is proposed, which uses Bayesian treatment for maximum likelihood estimation to generate highly sparse 'learned' models. And then, if labelled samples (petrophysical data) are not available, the unsupervised Self-Organizing Map (SOM) is applied to reveal geological features from seismic attributes and Quantitative Interpretation results to classify the seismic facies and other geological events. If the labelled samples are available supervised Support Vector Classification (SVC) has the capability to help us classify the lithology, detect hydrocarbon and identify sweet spots. Real data examples have demonstrated that the Machine Learning workflow has the advantages of the data mining for a more accurate seismic Quantitative Interpretation compared to traditional seismic methods.

Introduction

Even though the traditional seismic inversion, like simultaneous inversion and post-stack inversion can provide important information, such as velocity and density, the problem of how to effectively integrate additional information, such as petrophysical data and geological understanding, into the seismic inversion workflow for a more accurate seismic Quantitative Interpretation is still challenging. New technology is required to address these issues. The rapid rise of Machine Learning technology is enormous, which has been widely used for image classification, speech recognition and data regression. The rise is largely due to recent advances in Machine Learning, such as deep Convolutional Neural Network (CNN) and Support Vector Machine (SVM). Although we have developed traditional support vector machine methods for seismic inversion (Liu and Sacchi, 2003, Liu et al., 2012) and gained insights about seismic Quantitative Interpretation in conventional plays, shale plays and carbonate plays, the uncertainties and inconsistencies among different petrophysical properties and seismic attributes prevent us from obtaining a more accurate Quantitative Interpretation. However, other Machine Learning techniques could help us overcome these issues. For example, it is possible to address these issues using unsupervised SOM or supervised SVR/SVC through linking and calibrating seismic properties with petrophysical results and geological data to reveal geological features, such as sweet spots.

Generally Machine Learning, such as SVM and SOM, typically involves four steps: 1) data collection, 2) feature extraction and analysis using the data, 3) train a model using the data and/or the features, 4) make predictions using the trained model on new data.

According to the typical four steps of Machine Learning, in this paper we will address the Bayesian-based SVR for seismic property geostatistical inversion, SVC to detect the hydrocarbon or sweet spots, and SOM to produce a 2D map to identify the seismic facies to overcome seismic uncertainties and inconsistencies.

SVM technology has been gained popularity in small dataset regression and classification. In our study, we propose the Bayesian-based SVR for seismic inversion, which uses Bayesian treatment for maximum likelihood estimation to generate highly sparse ‘learned’ models. In order to speed up SVR convergence, the domain knowledge, such as the traps with small structural closures for traditional hydrocarbon accumulation, can be integrated into our workflow to better infer the hydrocarbon-related petrophysical properties.

Also there are uncertainties and inconsistencies among seismic inversion properties and attributes, which motivate us to develop the SOM and SVC methods to link and cross-validate these properties with petrophysical data and geological understandings in order that we can easily interpret seismic facies and identify sweet spots.

The preliminary results demonstrate that Machine Learning techniques, such as SVM and SOM, can add additional value by reducing the uncertainties of seismic Quantitative Interpretation.

Bayesian-based Support Vector Machine (SVM)

Given a set of input variables, $X = \{x_i \ i = 1, \dots, N\}$ together with corresponding targets, $Y = \{y_i \ i = 1, \dots, N\}$. The Bayesian-based methods build the model and make the prediction based on the following linear combinations of the form (Cortes and Vapnik, 1995, Liu and Sacchi, 2003, Liu et al., 2012):

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

Where $\{\omega_n \ n = 1, \dots, N\}$ are the model weights, and $\{K(x_j, x_n)\}$ denotes the kernel functions (Schölkopf, et al., 1999).

To avoid notational clutter we will re-write equation (1) in matrix form:

$$Y = W\phi \quad (2)$$

Where $\{\phi\}$ denotes the design matrix of kernel functions (Schölkopf, 1999), the vector of targets is given by $\{Y\}$, and $\{W\}$ indicates the vector containing the unknown weights.

The likelihood function of the dataset can be written as:

$$p(t|w, \sigma^2) = (2\pi\sigma^2)^{-\frac{N}{2}} \exp\left\{-\frac{1}{2\sigma^2} \|Y - W\phi\|^2\right\} \quad (3)$$

The maximum likelihood estimation will be applied and the posterior over the weights is then obtained from Bayes' rule.

In the training phase, the unknown weights can be estimated using the known target pairs and then in the mapping phase, the trained models can be applied to new data to make the predictions.

The SVM can analyze data for classification (SVC) if petrophysical data are labelled and for regression (SVR) if the input pairs are available.

Self-Organizing Map (SOM)

Unlike the SVM, the labelled samples or input pairs are not necessary for the unsupervised SOM learning. The SOM can better reveal geological features from seismic attributes (Roden, et al., 2015).

Given a set of input variables, $X = \{x_i \ i = 1, \dots, N\}$, the connection weights $W = \{w_{j,i}, j = 1, \dots, M\}$ between the input variables (i) and the neurons (j). M is the total number of neurons.

SOM methods build the discriminant functions using the Euclidean distance between the input variables and the connection weights for each neuron (j):

$$d_j(X) = \left(\sum_{i=1}^N (x_i - w_{j,i})^2\right)^{1/2} \quad (4)$$

In the training phase, the neuron whose weight ($w_{j,i}$) adjusts itself until the weight of the neuron comes closest to the input variables is declared the winner (j). The weight update equation is:

$$\Delta w_{j,i}(n) = \eta(n) T_{j,k}(n) (x_i - w_{j,i}(n)) \quad (5)$$

The n is the iteration step, η is learning rate and T is neighbor topological decay function, which is given:

$$T_{j,k}(n) = \exp\left(-\frac{d_{j,k}^2}{2\sigma^2(n)}\right) \quad (6)$$

The σ is the neighborhood distance, which needs to decrease with times. $d_{j,k}$ is the distance between winning neuron j and nearby neuron k on the neuron topology.

Machine Learning workflow for seismic quantitative interpretation

According to typical four steps of Machine Learning, the workflow (Figure 1) is proposed for our seismic quantitative interpretation.

Data Collection:

At this stage one needs to collect the seismic data, petrophysical data and geological data. Usually the seismic data include stack, offset gathers, azimuthal gathers. Petrophysical data should consist of well logs, core analysis, and production. In order to speed up the training, geological data, such as geological structure and stratigraphic interpretation, are also important for more accurate seismic inversion and interpretation.

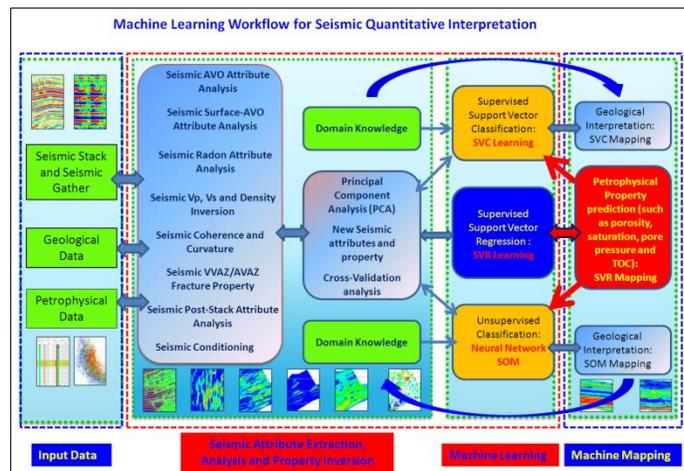


Figure 1: Machine learning for seismic quantitative interpretation

Feature extraction and analysis:

Feature extraction and analysis find the common and hidden patterns that will make it easier to distinguish between the geological events or improve our ability to model the relationships between seismic and geological properties. Usually there are two types of features: 1) the geological meaningful features, such as compressional velocity, shear velocity and density. Simultaneous inversion has the capability to obtain the velocity and density from seismic gathers. This type of feature is easy to understand and interpret their geological significance. 2) The mathematical meaningful attributes, such as average energy and frequency. This type of features is generated using mathematical algorithms. It can enhance the value for Machine Learning, but it is normally difficult to interpret from a geological viewpoint.

Machine Learning to train a model

The training features and/or data are used to generate the model. In seismic inversion, Bayesian-based SVR methods are selected to train the model using the data pairs of input seismic attributes and petrophysical properties. In previous sections, the methods have been discussed in detail. After the training, the weights of the model can be estimated. Then seismic supervised classifications, SVC methods, can be applied to identify the facies and sweet spots, and the unsupervised classifications, the SOM methods, are able to reveal geological features to identify the seismic facies and identify the presence of hydrocarbons after cross-validating against the geological data and petrophysical data.

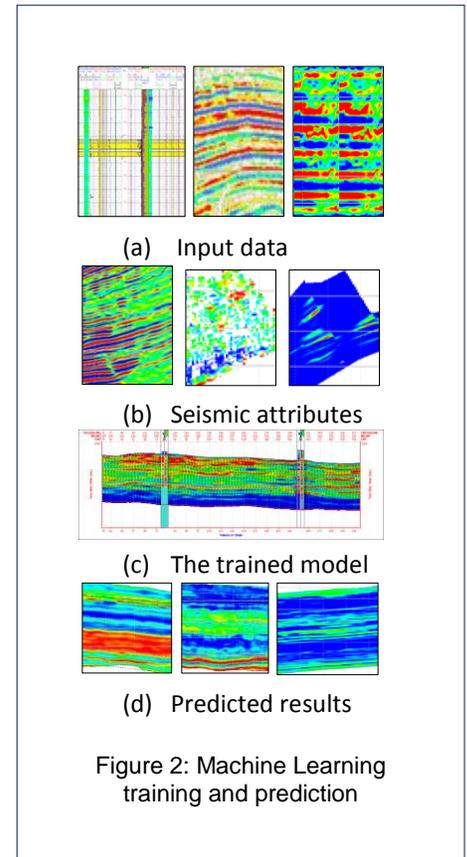
Machine Learning to prediction using the trained model

The first two steps of the Machine Learning workflow are used in the third step to create a trained model which is acceptable after cross-validating against the known data. The new data can then be fed into the trained model to make the predictions. The models can be improved after new drilling boreholes are available.

Examples

The Machine Learning workflow has been applied to the real data examples. In the first step, the petrophysical data, such as sonic, density, porosity and saturation are necessary. Also one needs to collect seismic stack and gathers to extract seismic attributes, such as AVO attributes. In the second step, seismic velocity and density can be inverted using simultaneous inversion. For shale and carbonate plays, fracture anisotropy intensity is also crucial as the input for our Machine Learning workflow. In the third step, for the facies classification, the supervised SVC can directly classify the petrophysical facies in seismic space and the unsupervised SOM can reveal the seismic facies after calibration with petrophysical data. For the reservoir properties study (Liu, etc., 2012), the SVR will be trained using the input pairs of petrophysical data and seismic attributes. In the last step, one can make predictions of petrophysical properties based on the trained models on new seismic space.

Figure 2 is an example. (a) is the petrophysical and seismic data, (b) is the example of AVO and VVAZ attribute (Liu, 2014) extracted from seismic gathers, (c) is the trained models (SVR) displayed around the well trajectory after the training phase, which demonstrated that the models match very well with borehole information. (d) is the inversion and facies results based on the trained models. We achieve results which are a better fit to petrophysical and seismic data because the domain knowledge and geological experience can be integrated in our workflow.



Conclusions

We present a Machine Learning workflow for seismic Quantitative Interpretation, which includes data collection, attribute extraction and analysis, model training and prediction using the trained model. In the model training, the supervised SVC and unsupervised SOM methods can be selected for the seismic facies analysis. For the reservoir petrophysical property study, the SVR methods can be applied to make the prediction of petrophysical properties within seismic space. The Machine Learning methods have the advantages of incorporating the seismic attributes and petrophysical data, and most importantly, the domain knowledge, such as geological understanding, to infer the petrophysical properties and identify the seismic facies. The real data examples demonstrate that the Machine Learning workflow has the capability to reduce the uncertainties and inconsistencies of seismic Quantitative Interpretation.

Acknowledgements

The author would like to thank SoftMirrors to support the research of Machine Learning applications and also thank Brian Link for useful suggestions and review of the paper.

References

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

Liu, Y., AVAZ and VVAZ practical analysis to estimate anisotropic properties, *GeoConvention 2014*, Canada

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

Liu, Y., Wang, L; Shuai, Y., Zhang, Y., Liu, Z., Chen, Z., 2012, Seismic methods to detect biogenic gas accumulations, example from the Qaidam Basin, China , *Society of Exploration Geophysicists, Annual International Meeting*, doi:10.1190/segam2012-0211.1

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

Roden, R., Smith, T., and Sacrey, D., 2015, Geologic pattern recognition from seismic attributes: Principal component analysis and self-organizing maps, *Interpretation*, Vol. 3, No. 4, p. SAE59–SAE83