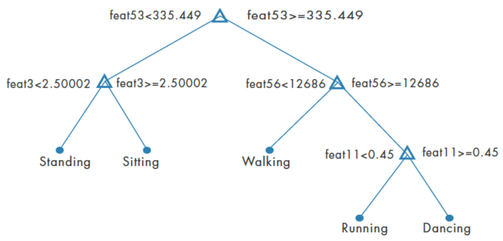
Seismic Facies Classification with Random Forest / Ensemble Method

Yuji Kim*, Kurt J. Marfurt

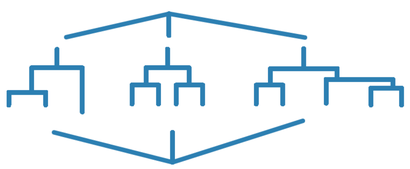
Random forest learning algorithm

Automated seismic facies classification is gradually becoming more common as part of workflows in the E&P industry, and has seen some success in stratigraphic plays (Coléou. et al. 2003). Several algorithms have been developed to automate seismic facies classification. In particular, machine learning provides effective workflows in facies classification and prediction. Machine learning can be classified into several techniques depending on whether the technique is unsupervised or supervised. Unsupervised learning, e.g. *k*-means clustering, principal component analysis (PCA), and self-organizing maps (SOM), were introduced to map and classify seismic data using its waveform or attribute infortmation. In this study, we implement a supervised learning algorithm using seismic attributes to predict seismic facies. Training and test process will be aided by seismic facies estimated from well log and core data.

We apply a random forest algorithm which is an ensemble of deicision trees, generally trained via a bagged or boosted method. The random forest or ensemble methods indicate which trees are considered “strong” learners and “weak” learners and combines them.



- Decision tree: consists of branches and weights which are decided in the training process. In the branching condition, the value of a predictor is compared to a trained weight.

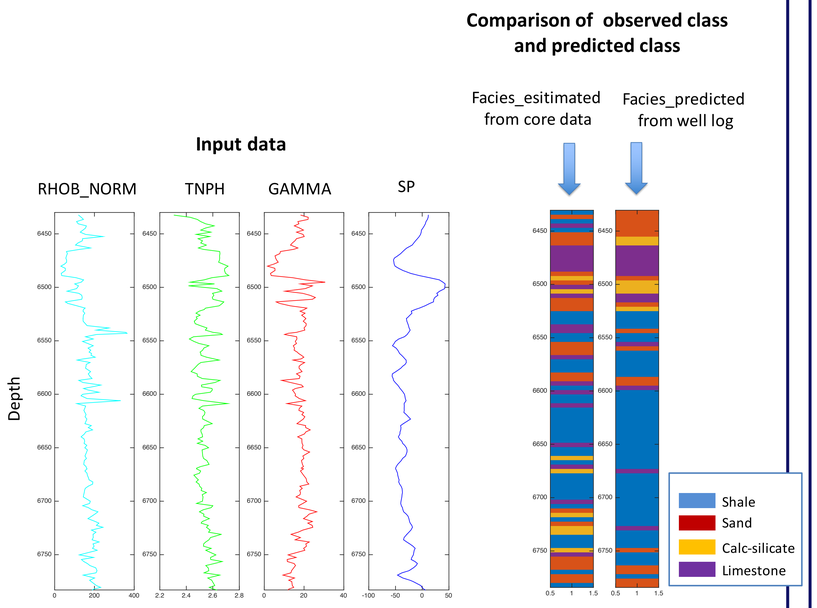


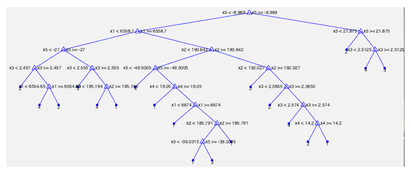
- Random forest/ ensemble methods: A bagged or boosted tree algorithm decides which trees are the “strong” learners and “weak” learners and combines them.

{Applying supervised learning, 2017, Mathworks]

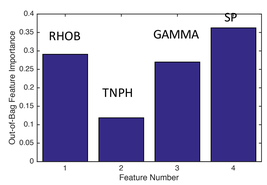
Rock facies classification with well logs

The example below shows a bagging method applied to well log data to predict and classify rock facies. Gamma, SP, RHOB, and NPHI log form the input data. To train data with class labels, rock facies are estimated from core analysis. After repeated training and testing in cross-validation, a fitted model was generated and is later used for prediction.





Graphical view of tree



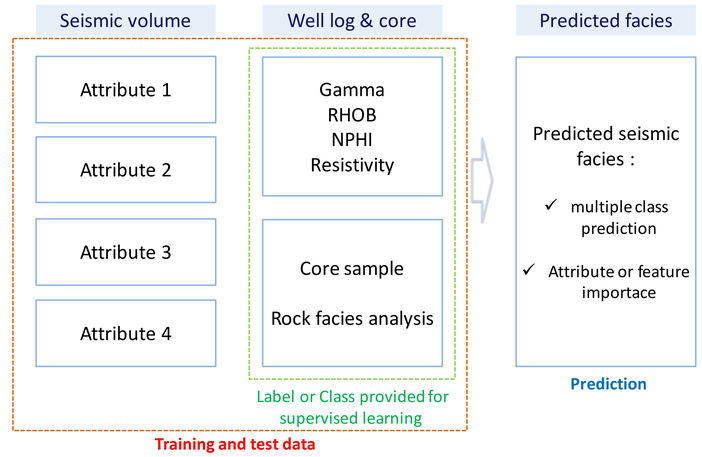
Feature importance

The benefit of a decision tree algorithm is that it gives a quantitative measure of how important or redundant each feature (or attribute) is in the learning process. A decision tree is a white box learning algorithm - the conditions and weights in the classifying process are provided. Classifying into multiple classes is also available.

Random forest methods (and also bagging or boosting methods) take advantage of decision trees, and alleviate overfitting since it integrates trees.

Seismic implementation and mehtod

- **Random forest algorithm applied to seismic data**



We apply the random forest algorithm to seismic facies classification. Several attribute volumes - e.g. coherent energy, peak frequency and magnitude, curvedness and GLCM textures will be used as data input to predict seismic facies aided by well log and core data.

The aim of the study is to use trees with ensemble method, to classify seismic facies and figure out which attribute is more important in classifying seismic facies. The information of feature importance will help to choose the input attribute and reduce the amount of computation.

Reference

- Aurélien G. ,2017, Hands-on maching learning with scikit-learn & tensorflow
- Coléou, T., Poupon, M., and Azbel, K., 2003, Unsupervised seismic facies classification: A review and comparison of techniques and implementation. The Leading Edge, 22(10), 942-953.
- Mathworks, 2017, Applying supervised learning
- Verma, S., Zhao, T., Marfurt, K. J., and Devegowda, D., 2016, Estimation of total organic carbon and brittleness volume. Interpretation.
- Zhao, T., Jayaram, V., Marfurt, K. J., and Zhou, H., 2014, Lithofacies classification in Barnett Shale using proximal support vector machines. In SEG Technical Program Expanded Abstracts 2014, 1491-1495.

For communication email: yuji.kim@ou.edu