Supervised and Unsupervised Learning: How Machines Can Assist Quantitative Seismic Interpretation
Tao Zhao*, Sumit Verma and Kurt J. Marfurt, University of Oklahoma

Summary
In this study, we use an example from the Barnett Shale to demonstrate how supervised and unsupervised machine learning techniques provide the right leverages for seismic interpreters. Combining automatic machine analysis and human interpretation helps understanding the extensive data interpreters have.

Introduction
Frackability is a key parameter of recovering unconventional shale reservoirs which can be measured by the brittleness index (Bi) of reservoir rocks. In this study we used both unsupervised learning technique (self-organizing map or SOM) and supervised learning technique (proximal support vector machine or PSVM) to estimate Bi using five petrophysical attributes. For SOM, we cannot directly get Bi from the inputs, but can have clustered lithofacies, which need to be further interpreted based on well logs. For PSVM, we can directly calculate Bi on seismic data based on the relation obtained from a training well. Because Gamma Ray is a good indicator of TOC as well as clay minerals which generally make a rock ductile, we also compared the estimated Bi with a Gamma Ray volume estimated from artificial neural network (ANN).

Workflow
Seismic attributes vs. TOC training and input SOM
Well logs
SOM PSVM Frackability and lithofacies with Bi volume interpretation ANN estimated Gamma Ray volume
Lithofacies

Discussion
By using the same input data for SOM and PSVM, we generated a SOM cluster volume and a Bi volume, respectively. Being an unsupervised learning algorithm, SOM provides clusters which need to be further interpreted using other data, whereas PSVM gives us determined products, which in this study is Bi. There is a virtual correlation between SOM clusters and Bi; however, one cannot conclude the clusters represent Bi values. In fact, SOM clusters contain all the information from five input volumes and recover natural relations within the data other than linking the inputs to a specific output (Bi).

Conclusions and Future Work
Supervised and unsupervised machine learning techniques provide human guided classification as well as data driven clustering which help us better understand the data. Interpretation of unsupervised clusters requires correlating with other data and expert insight to interpret. Supervised training needs carefully chosen input data to insure its geologic meaningfulness.

Acknowledgements
Thanks to Devon Energy for providing the data, all sponsors of AASPI consortium group for their generous sponsorship, and colleagues for their valuable suggestions.

Figure 1. Five input seismic attribute volumes at line AA': (a) P-impedance; (b) S-impedance; (c) Lambda-Rho; (d) Mu-Rho; and (e) Ratio between Young’s modulus/ Poisson’s ratio. All these volumes are either inverted or further calculated from prestack simultaneous inversion.