UNCONVENTIONAL RESOURCES TECHNOLOGY CONFERENCE
FUELED BY SPE • AAPG • SEG

URTeC: 2901881

Geological Facies Prediction Using Computed Tomography in a Machine Learning and Deep Learning Environment

Uchenna Odi[*1] and Thomas Nguyen[*1] 1. Devon Energy Corporation.

Abstract

Machine learning and deep learning, a subset of machine learning, have become successful predictive modeling tools. Machine learning applications include classifying DNA sequences and fraud detection. Deep learning applications include language processing and facial recognition. Both methods have shown to be useful in understanding unconventional reservoirs when traditional techniques used to understand conventional reservoirs do not readily transfer. In this work, machine learning and deep learning are used to predict geological facie classifications using computed tomography (CT).

A combined approach is developed to predicted geological facies. The first aspect of this approach utilizes molecular weight, density, and porosity from CT scans on core samples to learn from existing user-defined geological facie classifications. Once the machine understands the relationship between geological facies and CT scan physics, the resulting model is used to predict subsequent geological facies using the CT scan parameters. The second aspect of this approach utilizes CT scan thin section images of core samples to learn from corresponding geological facie classifications. Once a deep learning model learns from these CT scan images, the resulting model is used to predict subsequent geological facies using CT scan images.

Results show misclassification rates at approximately less than 30% which are more favorable to error rates associated with a single person classifying facies. In addition, unsupervised machine learning results show robust clustering of geological facies. Machine learning and deep learning show the value of utilizing CT scan parameters and images in unconventional oil and gas reservoir workflows. While typically used in other areas such as cancer detection, CT scans in conjunction with machine learning and deep learning can drastically reduce the error rate and time it takes to fully categorize an unconventional play.

Current methodologies involve having a team of petrophysicists who characterize core samples at every specified CT scan resolution depth. These workflows can be cumbersome and time-consuming for a team of petrophysicists and at times result in facies classification governed by human bias. The proposed methodology in this work drastically reduces human error and the time to understand geological facies. As a result, business decisions can be made quicker. These models are expected to provide substantially more accurate and consistent geological facie predictions.

Introduction

Core facies analysis workflows are a fundamental step in the characterization of unconventional and conventional reservoirs. Capital investment in the selection and extraction of whole core data is significant. Therefore, rapid discrimination of core data and associated analysis for a return on initial investment, and to improve overall business decisions is critical. Total elapsed time from pulling the core to a finalized core and facies description can take up to, or longer than, a month to complete depending upon formation complexity and laboratory turnaround time. Incorporation of whole-core x-ray CT (computed tomography) imaging into this workflow allows for qualitative,

quantitative, and statistical analysis of core data within days of core arriving at the laboratory. Deliverables from these workflows can be used for sample selection, core slabbing, core description, and business decisions.

Whole-core CT imagery, acquired using modified medical imaging devices, allows for the investigation reservoir textural and compositional variability on basis of density and chemical composition (effective atomic number). Grayscale values assigned to each voxel - or volume pixel - of rendered CT imagery reflects X-ray attenuation, which is function of a materials density and effective atomic number. Compositional heterogeneity lends itself to facies analysis and discrimination of different rock types. Resolution of a typical voxel size of whole-core CT data is approximately 300x300x300 microns. Datasets have a typical sampling rate of approximately 1,000 observations (in Z direction) per foot. In addition to imagery, acquisition scans at multiple energies (high- and low-energy) can be used to calculate petrophysical properties such as bulk density (Rhob) and effective atomic number (Zeff) (Wellington and Vinegar, 1987). The size and high resolution of this data type presents an opportunity for characterization of the reservoir on a statistical basis.

Utilization of whole-core CT imagery occurs early in a core facies analysis workflow, which is shown graphically in Figure 1. 3D computed tomography imaging is non-destructive and can be acquired prior to extrusion when core is still in a capped aluminum barrel. This type of imaging is now a common offering for most laboratory service companies. Scanning continuous core volumes has become common as part of reservoir characterization workflows. Availability of this data type has provided geoscientists with increased data volumes and high-resolution radiodensity models of continuous reservoir intervals. Rapid acquisition and delivery of whole-core CT imagery from the laboratory can be used to identify sample site selection during extrusion and can be leveraged for core description and facies analysis.

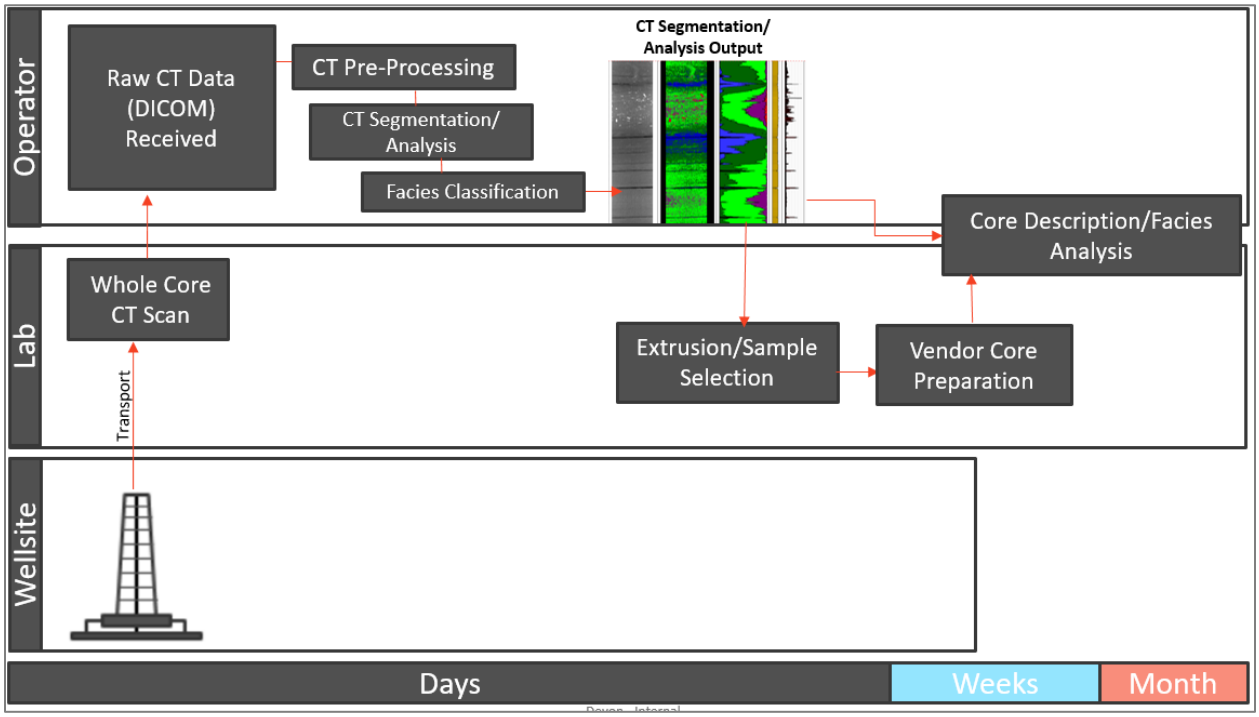

Figure 1: Graphical illustration of utilization of whole-core CT imagery in core and facies analysis workflow. Individual swim lanes show location of analysis and work (i.e., wellsite, lab, Devon Energy). X-axis shows estimated elapsed time increasing to the right.

Supervised and unsupervised machine learning has the potential to greatly help with analysis of large data volumes. This need, recognized and highlighted by Ball *et al.,* (2017), is predicted to drastically change oil and gas operations. Ball *et al*. (2017) outlined that the oil and gas industry has adapted to the volume, velocity, and variety of big data. The current challenge is leveraging big data to remove inefficiencies that slow down the decision-making processes. Workflows presented by Hall and Govert (2016) highlight the application of machine learning algorithms for statistical feature detection in whole-core CT volumes. Facies classification, using machine learning, of wireline datasets has recently received a lot of attention, and was the focus of the 2017 Leading Edge competition. Using

10.15530/urtec-2018-2901881

linear logistic regression, Hoeink and Zambrano (2017) could distinguish between shale and non-shale lithology on the basis of density, gamma ray, resistivity, compressional and shear slowness logs as classifiers. Their results showed successful binary classification of shale and non-shale lithology with correct classifications ranging from 90 percent to 96 percent.

The work presented here evaluates the following machine learning workflows on whole-core CT imagery: 1) supervised machine learning, 2) unsupervised machine learning, and 3) deep learning. In doing so, we evaluate machine learning tools and their potential to minimize the length of time and effort required to analyze a core. Impacts from this work include speeding up and improving decision-making at various levels, which can minimize overall well and field cost. We also hope to de-mystify data and analytics tools and workflows.

Image Processing and Artifact Correction

Whole-core CT data are cropped to remove aluminum core barrel and stitched into continuous core volumes. Artifacts from image acquisition, such as beam hardening and ring artifacts, are corrected. An overview of common CT artifacts and how to correct for them can be found in Boas and Fleischmann (2012). Statistical attributes to be used for machine learning are compiled from CT Imagery on a slice-by-slice basis (in Z direction). These include mean, median, minimum, and maximum CT number, as well as kurtosis and standard deviation. CT data are also binned on basis of the 10^{th}, 25^{th}, 50^{th}, 75^{th}, and 90^{th} percentile of CT number for the entire core. The fraction for each bin is computed on a per-slice basis and used for statistical analysis. Additional variables included in data and analytics include calculated Rhob and PEF values provided by a vendor.

Traditional Supervised Machine Learning Workflow

Traditional supervised machine learning requires the creation of a predictive model based on pre-existing training data. This requires both CT data and a well-constrained and defined facies description to be used to train and create the model. Model outputs are as good as their training set. Therefore, work by a geoscientist is critical for creating a successful model. Once created and validated, a supervised machine learning model can be used either for rapid classification of future CT scans of the same formation or for training a small subset of an existing dataset for classification of the remainder. For the supervised portion of this work, the pre-existing training data was derived from core that had previously been CT scanned and thoroughly described. A subset of assigned facies was compiled as a training set for supervised machine learning. A breakdown of these facies is shown in Figure 2 and Table 1.

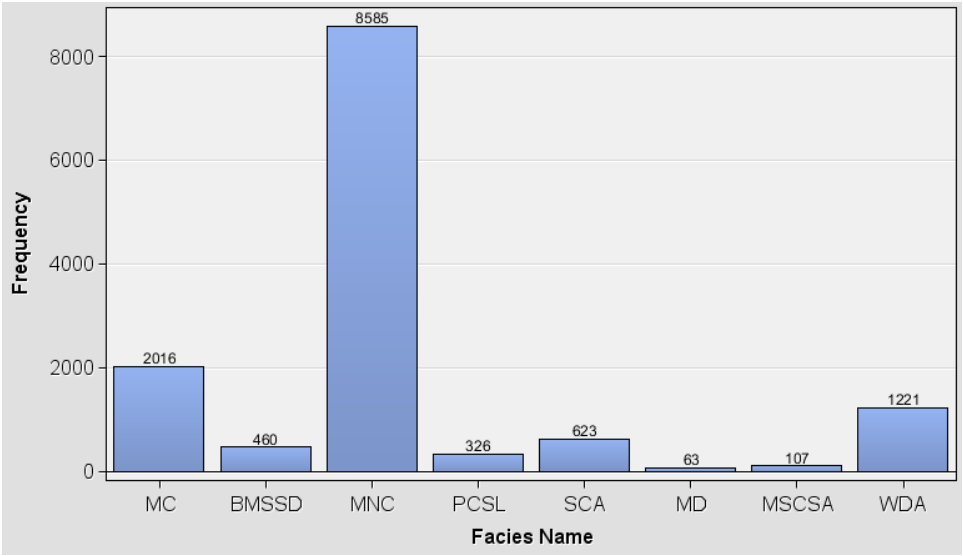

Figure 2: Distribution of assigned core facies of full dataset used for supervised machine learning

Table 1: Facies Distribution Summary of Training Set (13,401 CT slices from 58.5 ft of described core)

Facies	Facies Abbreviation	Facies Percentage
Mudstone – noncalcareous	MNC	64.1%
Mudstone – calcareous	MC	15.1%
Wackestone or dolostone	WDA	9.11%
Siltstone – calcareous, argillaceous	SCA	4.65%
Bioturbated mudstone	BMSSD	3.43%
Packstone – very calcareous siltstone, or massive limestone	PCSL	2.43%
Mudstone – silty	MSCSA	0.80%
Mudstone – dolomitic	MD	0.47%

Classifiers used for the supervised machine learning include calculated outputs from a dual-energy whole-core CT scan, such as calculated density (Rhob) calculated porosity (PHI), calculated photoelectric factor (PEF), and calculated UCS. Prior to implementing supervised machine learning it is critical to understand the relationships between each classifier. This can help identify issues with the data such as multi-collinearity. A matrix plot of these classifiers and their corresponding relation to one another is shown in Figure 3. Classifiers are correlated (Figure 3) and are approximately normally distributed with no extreme outliers (Figure 4).

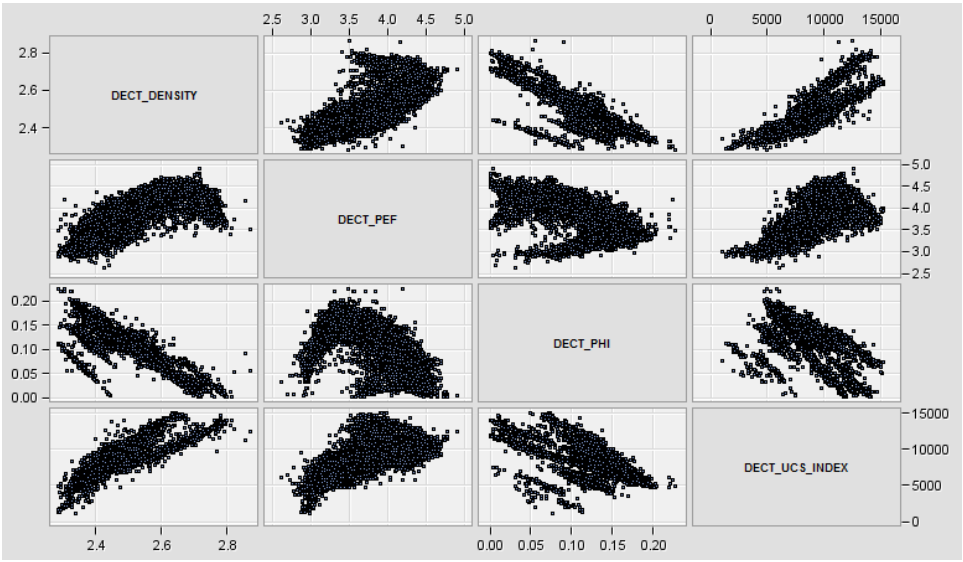

Figure 3: Matrix plot of classifiers used for supervised machine learning

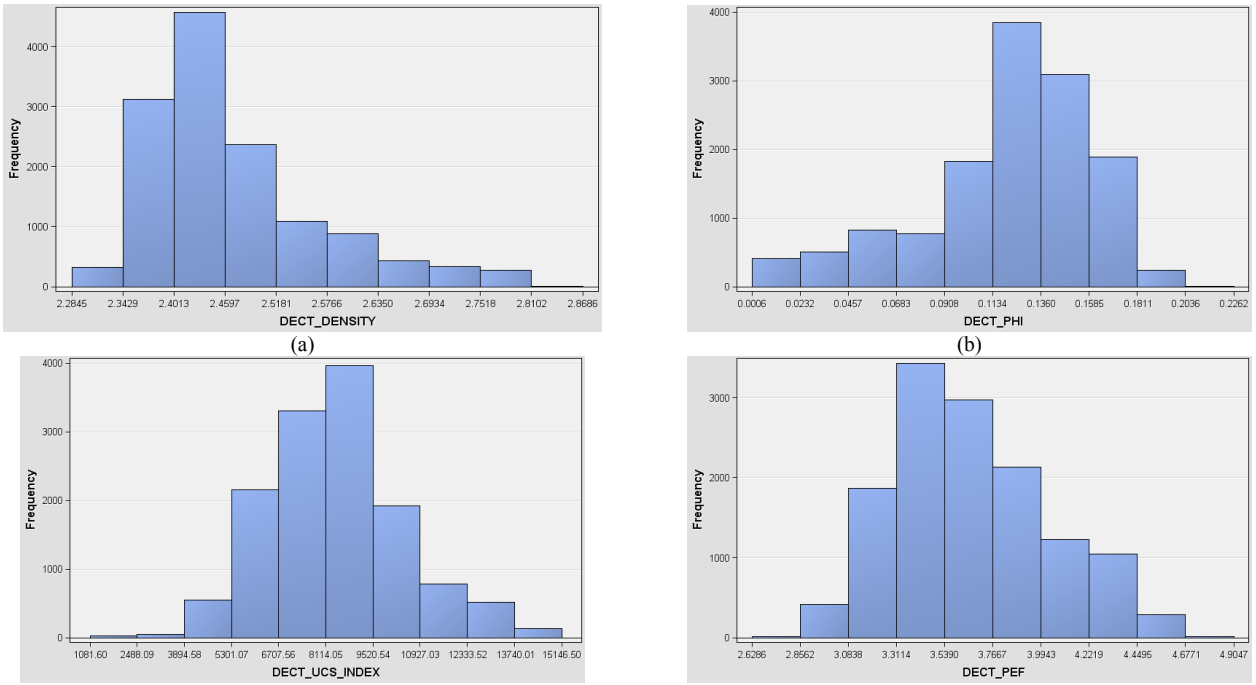

Figure 4: Histograms of classifiers used for unsupervised machine learning: a) calculated Rhob, b) calculated Phi, c) calculated UCS, d) calculated PEF

Prior to supervised machine learning, 75 percent of the data are sectioned off as a sample set and the remaining 25 percent preserved as the "out of bag" test set. The "out of bag" test set was not be used for model creation, but was instead used for validation. From the sample set, 80 percent of the data was used as the training set and the remaining 20 percent of the sample set was used as the validation set. The sample set was utilized in a random forest machine learning algorithm, which is a collection of decision trees merged together to get a more accurate and stable prediction (Donges, 2018). The random forest machine learning algorithm was chosen after a model competition of other machine learning models which compared validation misclassification. Other models considered were the gradient boost, decision tree, neural network, and linear regression models that resulted in validation misclassification rates of 30.8 percent, 30.9 percent, 30.6 percent, and 35.1 percent respectively. The random forest model was the model champion because of training set misclassification of 5.15 percent and a validation set misclassification of 27.1 percent.

The number of decision trees used in random forest construction can be set to minimize misclassification. Figure 5 illustrates the misclassification rate as a function of the number of trees in the random forest algorithm for the training set, validation set, and out-of-bag set (randomly selected subset of data that is not used for training). Predictive power of the random forest increased as the number of decision trees was added. 100 trees were sufficient to minimize the misclassification of the supervised machine learning model.

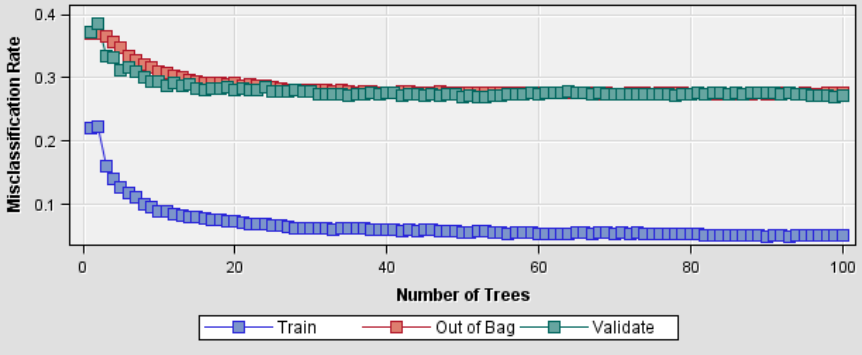

Figure 5: Misclassification rate as a function of the number of trees in Random Forest machine learning algorithm for training, out of bag, and validation datasets

The relative error for each facies from the training set, with a 5.15 percent misclassification rate, is shown graphically in Figure 6. The bar shown next to each facies shows the total number of observations of that facies type in the training dataset. For example, 1209 mudstone-calcareous (MC) were in the training set. Each bar is filled and color-coded with corresponding predicted facies results from supervised machine learning. This facilitates a direct comparison between predicted results and actual facies in the training set and highlights cases of misclassification.

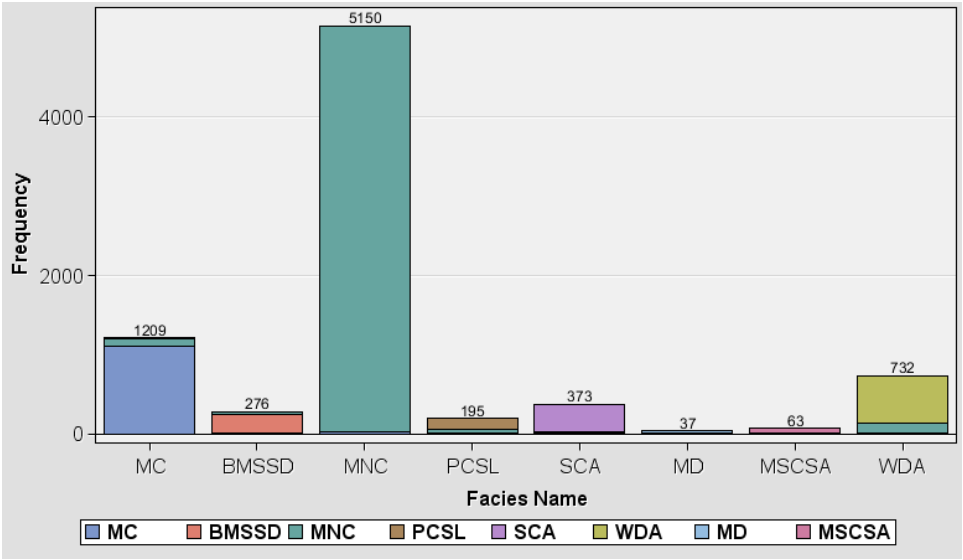

Figure 6: Supervised machine learning facies prediction results on training set with a 5.15 percent misclassification for Mudstone – calcareous (MC), Bioturbated mudstone, siltstone, and sandstone (BMSSD), Mudstone – noncalcareous (MNC), Packstone - very calcareous siltstone, or massive limestone (PCSL), Siltstone – calcareous, argillaceous (SCA), Wackestone or dolostone – argillaceous (WDA), Mudstone – dolomitic (MD), Mudstone – silty; massive, faintly laminated, slightly calcareous (MSCSA)

Supervised machine learning yielded a misclassification rate of less than 30 percent for the validation set. In some instances, misclassification can be explained geologically where beds are dipping as a result in classifier values averaged across multiple facies types. Thin beds and laminations are unrealistically difficult and time-consuming to classify on the millimeter scale, and may contribute to misclassification.

Traditional Unsupervised Machine Learning Workflow

Traditional unsupervised machine learning creates a predictive model that determines the optimal number of groups from the training data and classifies data into the identified number of groups. Creation of an unsupervised machine learning model can occur early in a core analysis workflow, once data are received, stitched, cropped, and corrected for artifacts. Although models of this type do not assign specific rock types, they do cluster or group rocks of similar

whole-core CT signature. Assigned clusters can be used during core extrusion and help ensure a representative sampling selection strategy. Predicted clusters aid with core description and provide a quick look at variability in rock type early in core description workflows and aid in business decisions.

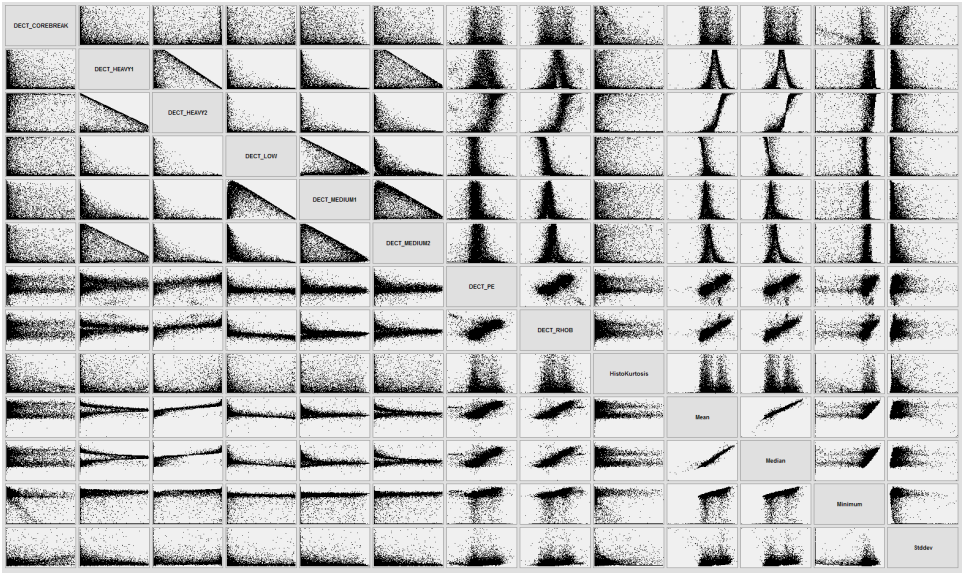

Figure 7: Matrix plot of classifiers for unsupervised machine learning data set

The dataset used for unsupervised machine learning includes 10,423 observations from 738.00 ft of core. Classifiers used in the unsupervised model are shown in matrix plot (Figure 7). Data are filtered on a basis of high standard deviation values and low CT numbers to remove breaks in core/anomalous outliers. This was shown to improve clustering of unique rock types. Hierarchical clustering, using the average distance between observations, was implemented on the unsupervised machine learning data set.

Assigning the ideal number of clusters, or groups, to given dataset is a critical precursor to performing unsupervised machine learning. Moreover, estimating the number of clusters for a given dataset provides insight into internal data structure and the number of lithologic groups that can be identified based on CT data. Estimating the number of clusters can be accomplished using the following clustering statistics: 1) clustering criterion (CCC), the 2) pseudo-F statistic, and 3) pseudo-T squared. Each clustering statistic is summarized in Table 2. Overall, the ideal number of clusters for a given dataset is observed when the CCC statistic is maximized (for positive cases), pseudo-F statistic is maximized, and pseudo-T squared statistic is minimized after a large peak.

Table 2: Clustering statistic descriptions and optimal cluster criteria (Hare, 2010 and Sas, 1983)

Clustering Statistic	Description	Optimal Cluster Criteria
CCC	Estimates the error between the hierarchical clustering algorithm and a cubic reference distribution.	CCC is maximized (for positive cases)
Pseudo F	Measures the separation among clusters	Pseudo F is maximized
Pseudo T squared	Measures the separation between two clusters most recently joined	Pseudo T squared is minimized before a large peak (read right to left on Pseudo T squared plot versus number of clusters)

Clustering statistic results, which can be used to estimate number of clusters on the unsupervised dataset, are shown in Figure 8 where assigned number of clusters are shown on the x-axis. Cubic clustering criterion statistic did not reveal a suitable cluster size. However, there is strong agreement in the pseudo-F and the pseudo-T squared statistics at 10 clusters. This tool provides validation for a number of facies assigned to core in well-studied

formations, and can be used with new data in new formations as a means to statistically identify the number of facies.

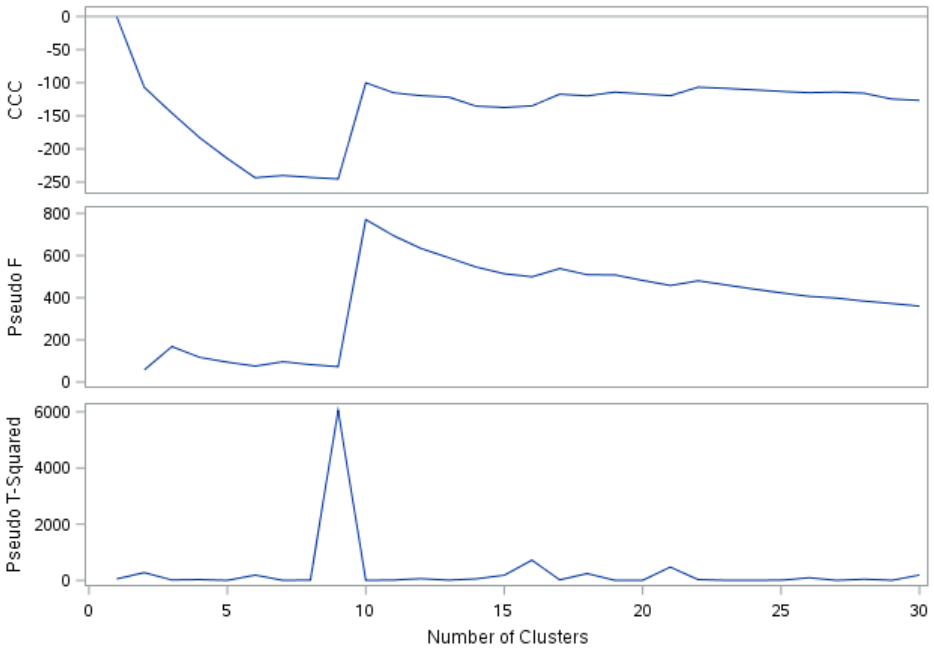

Figure 8: Results of hierarchical clustering on unsupervised machine learning dataset

The resulting relative distribution of the 10 clusters system is illustrated in a Figure 9. Cluster similarity is shown graphically by plotting the distance to the nearest facie versus the facie or cluster number (Figure 10). Figure 10 illustrates that the 5th and 6th facie are the most separated from the other facies. This is an indication of the rarity of these facies as compared with the other facies. In total, these facies account for less than 2 percent of all of the facies.

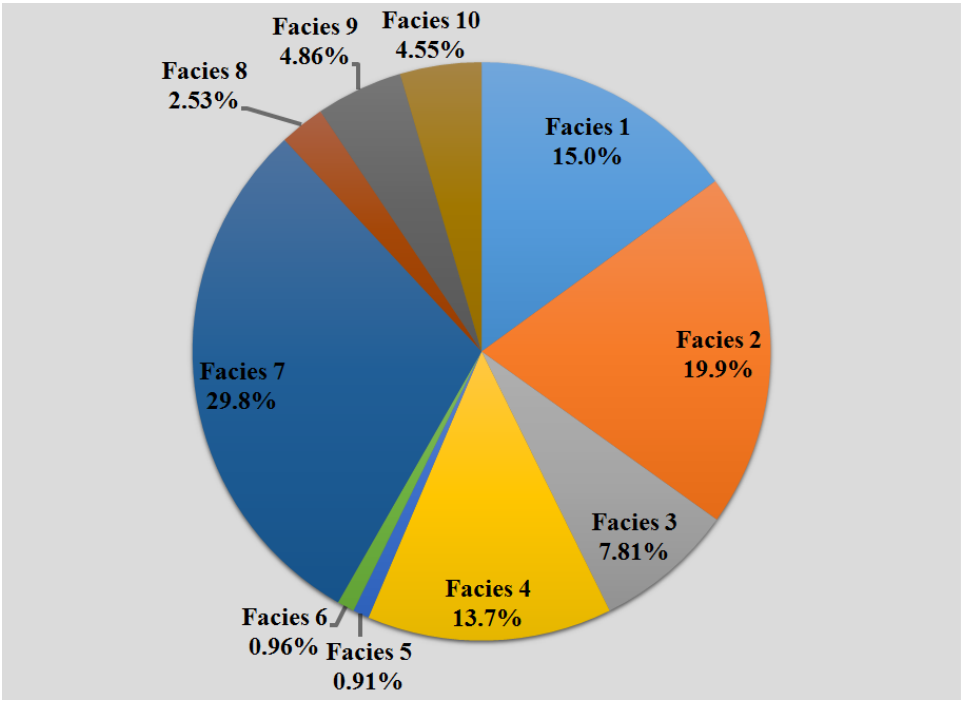

Figure 9: Pie diagram showing distribution of clusters from unsupervised classification. Number adjacent to each slice identifies cluster and percentage of abundance within the dataset.

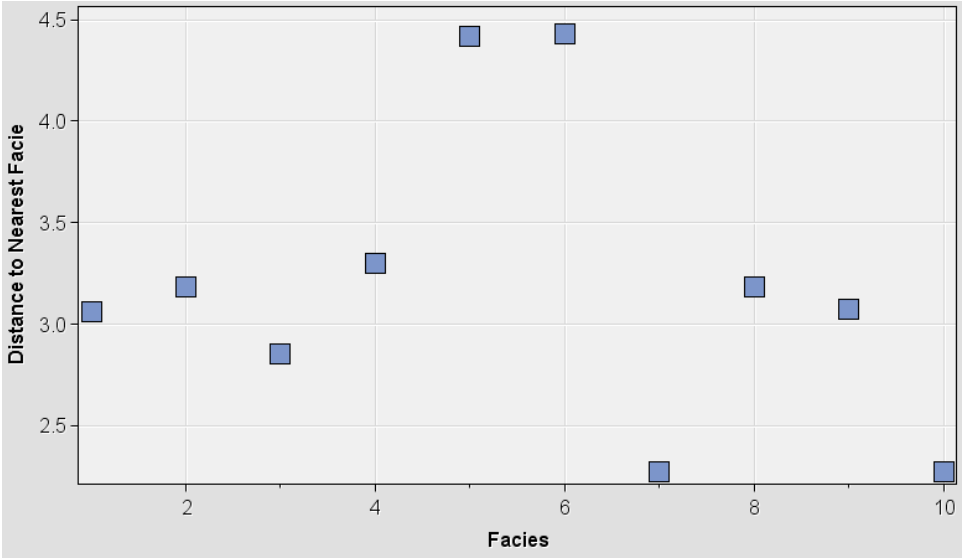

Figure 10: Cluster proximity map for unsupervised machine learning. Axes are determined from multidimensional scaling analysis using a matrix of distances between cluster means as input (SAS, 2016). Clusters with similar whole-core CT statistics will plot near one another.

Assigned cluster IDs and a grayscale XZ and YZ slice of associated CT imagery is shown in Figure 11. Grayscale intensity of whole-core CT imagery is a function of a materials density and effective atomic number. Unsupervised machine learning can distinguish high-density, high effective atomic number facies from facies of low density low effective atomic number on basis of clusters. In some cases, thin beds and sections containing significant core breaks are difficult to discriminate. However, when used over the length of an entire core, unsupervised machine learning is able to group rocks of similar CT signature.

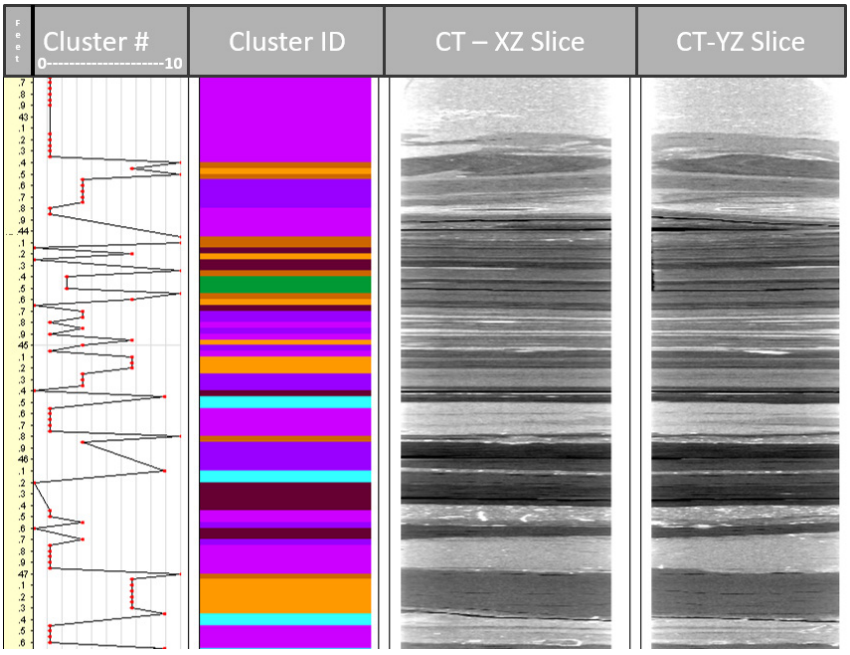

Figure 11: Unsupervised machine learning facie prediction showing cluster number and associated colored ID and grayscale XZ and YZ slice of whole core CT imagery over approximately 4 ft of core

Deep Learning

Deep learning is a subfield of machine learning based on learning pattern representations from data. Modern deep learning techniques have increasingly become more complex involving tens or even hundreds of layers of representations (Chollet, 2017). Recently, deep learning has made significant progress and advancement specifically in the field of image recognition and object detection. Convolutional Neural Networks (CNNs) are a category of Neural Networks that have proven very effective in areas such as image recognition and classification. CNNs have been successful in identifying objects, scenes, faces, and providing vision in robots and autonomous self-driving cars. Recently, CNNs have been trained to predict lung cancer and detect brain tumors on CT scans with high accuracy results.

As described earlier, traditional machine learning requires a subject matter expert to manually extract features from training data. In a deep learning framework, the algorithm automatically learns useful features directly from the training data such as image, text, patterns, digital signals, and etc.

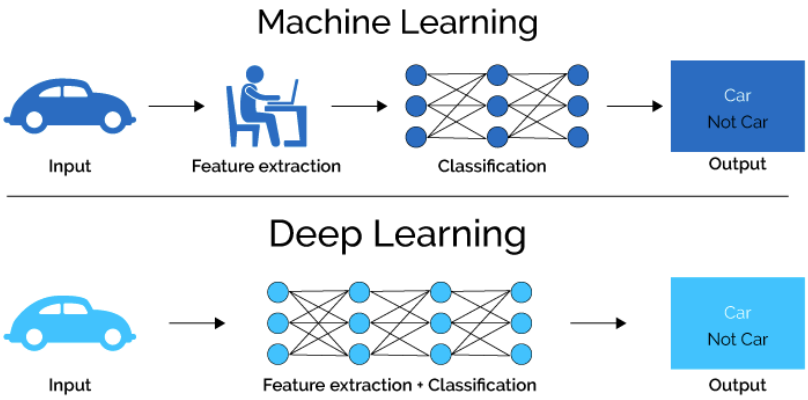

Figure 12: Difference between traditional machine learning and deep learning (Alex, 2017)

As an example of using a CNN, a preliminary approach to implement a basic small CNN model was able to achieve greater than 60% accuracy using a small subset of core CT scan images (approximately 10,000 images). These results are outlined in Figure 13. The performance is expect to be higher as more images are used to train the model.

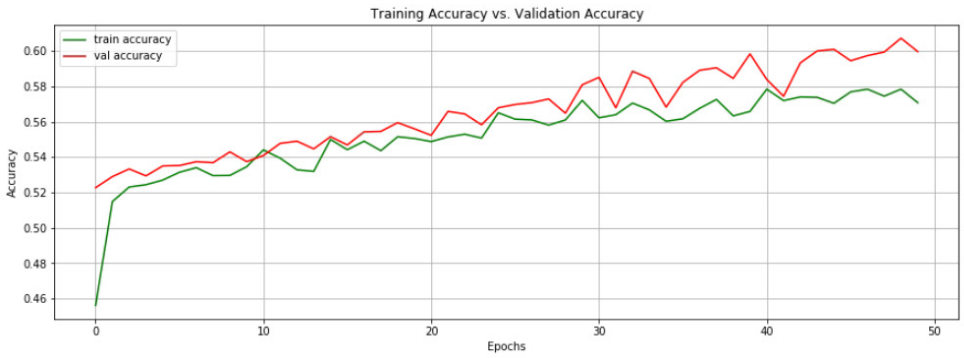

Figure 13: Basic CNN Model Training vs Validation Accuracy

Future work includes training larger CNN models with various parameter tuning methods and implementing 3D CNN to learn the volumetric representation of the core CT imaging.

Conclusions

Machine learning offers the potential for rapid characterization of rock facies. Supervised machine learning based on prior geological classification allows a geologist's facies to be used to classify rocks of the same formation. This can be used for classifying rocks of the same formation if additional cores are taken. However, machine learning is not a replacement for conventional core description and facies analysis workflows, but offers a potentially powerful tool to utilize during core extrusion sample selection workflows, as well as a means to speed up descriptive workflows by identifying rock types that are compositionally similar. Unsupervised machine learning can be used to identify the number of statistically significant rock types on the basis of CT data. Deep learning can be used to automatically classify and predict facies from CT scanned images. Deep learning shows promise in replacing the existing workflows inherent in supervised and unsupervised machine learning.

References

1. Alex, D., 2017: Difference between Machine Learning and Deep Learning. https://artificialintelligencehow.com/2017/10/18/difference-machine-learning-deep-learning/.
2. Ball, K., Arbus, T., Odi, U., and Sneed, J., 2017: The rise of the machines, analytics, and the digital oilfield: Artificial intelligence in the age of machine learning and cognitive analytics. Unconventional Resources Technology Conference.
3. Boas, E and Fleischmann, D., 2012: CT artifacts: Causes and reduction techniques in Imaging Med. 4(2), 229-240.
4. Chollet, F., 2017: Deep Learning with Python, 8, 9.
5. Donges, N., 2018: The Random Forest algorithm, in Towards Data Science. https://towardsdatascience.com/the-random-forest-algorithm-d457d499ffcd.
6. Duliu, O. G., 1999: Computer axial tomography in geosciences: an overview. Earth-Science Reviews, 44, 265-281. http://dx.doi.org/10.1016/S0012-8252(99)00056-2.
7. Hall, M. and Hall, B., 2017: Distributed collaborative prediction: Results of the machine learning contest. The Leading Edge, 36(3), 267–269.
8. Hare, T., 2010: Nearest neighbor based approaches to multivariate data analysis.
9. Hoeink, T. and Zambrano, C., 2017: Shale discrimination with machine learning methods. American Rock Mechanics Association.
10. SAS Enterprise Miner 14.2., 2016: Cluster Node.
11. SAS Institute Inc., 1983: SAS Technical Report A-108, Cubic Clustering Criterion.
12. Wellington, S. L., Vinegar, H. J., 1987: X-ray computerized tomography, *in* Journal of Petroleum Technology, August 1987, pp. 885.