Society of Petrophysicists and Well Log Analysts

Petrophysical Data Driven Analytics

Pseudo Sonic Log Generation With Machine Learning:

A Summary of SPWLA PDDA Machine Learning Contest 2020

Yanxiang Yu, Chicheng Xu, Michael Ashby, Siddharth Misra, Weichang Li (Contest committee)

Wen Pan, Tianqi Deng, Honggeun Jo, and Javier E. Santos (Team UTFE)

Lei Fu and Chengran Wang (Team iwave)

Arkhat Kalbekov and Valeria Suarez (Team RocketAbuser)

Epo Prasetya Kusumah, Mohammad Aviandito, and Yogi Pamadya (Team SedStrat)

Hossein Izadi (Team iPetro)

Background

Machine learning has gained increasing momentum in petrophysical applications in recent years (Xu et al., 2019; Misra et al., 2019). It is imperative to prove the capability of machine learning in solving real petrophysics problems. Started on March 1 and concluded on May 7, 2020, the first machine-learning contest was hosted by the SPWLA PDDA SIG, focusing on synthetic sonic log generation from other "easy-to-acquire" well logs. During the competition, the contest committee received more than 30 submissions from over 50 registered teams globally. Each team was composed of at most five team members, and they were allowed three submissions maximum. The highest score from the three submissions was used for the final ranking. A notebook that detailed the best submission was provided in order to be ranked. The top model beat the performance of the benchmark model by 31% in the root mean squared error (RMSE) score. The complete leaderboard and solutions from each team for the competition are listed on the competition website hosted by Github. In this paper, we will briefly summarize the competition and describe the five solutions submitted by the top winning teams.

	Winner	RMSE Score	Solution	Contact
1st Place	UTFE	12.359	Neural Network	[Wen Pan](wenpan@utexas.edu)
				[Tianqi Deng](tianqizx@utexas.edu)
				[Honggeun Jo](honggeun.jo@utexas.edu)
				[Javier Santos](jesantos@utexas.edu)
2nd Place	iwave	12.551	LSTM	[Lei Fu](lei.fu.rice@gmail.com)
3rd Place	RockAbusers	13.216	Random Forest	[Arkhat Kalbekov](akalbekov@mines.edu)
				[Valeria Suarez](vasuarezbolivar@mymail.mines.edu)
4th Place	StuckAtHome	13.431	Ensemble Trees	
5th Place	SedStrat	13.845	Ensemble Model	[Epo Prasetya Kusumah](epo.kusumah@gmail.com)
				[Mohammad Aviandito](aviandito@gmail.com)
				[Yogi Pamadya](yogipamadya@gmail.com)

Table 1—Top 5 scoring teams and their contact information.

ABSTRACT

Compressional and shear sonic travel time logs (DTC and DTS, respectively) are crucial for subsurface characterization and seismic-well tie. However, these two logs are often missing or incomplete in many oil and gas wells. The competition aims to predict the DTC and DTS logs from seven "easy-to-acquire" conventional logs using machine-learning methods. A tutorial was published in the 2020 March issue of *SPWLA Today* newsletter as a benchmark to the competition (Yu et al., 2020). A total number of 20,525 data points (corresponding to distinct depths) from three wells were collected to train regression models using machine-learning techniques. Each data point had seven features, consisting of the conventional "easy-to-acquire" logs: caliper, neutron porosity, gamma ray (GR), deep resistivity, medium resistivity, photoelectric factor, and bulk density, respectively, as well as two sonic logs as the target. The separate data set of 11,089 samples from a fourth well was then used as the blind test data set. The prediction performance of the model was evaluated using RMSE as the metric. In the tutorial, a random forest regressor model was trained, and an RMSE score of 17.93 was achieved on the test data set.

The top five models, on average, beat the performance of our benchmark model by 27% in the RMSE score. From their submissions, we found that data cleaning and clustering were critical to improving the performance of their models. Different models, including neural network, long-short-term memory (LSTM), and ensemble trees, were used by different teams, achieving impressive performance. In this paper, we have selected five solutions from the top submissions and present their techniques in detail on how they tackled this competition.

TOP AWARD-WINNING SOLUTIONS

1st Place–Team UTFE (Wen Pan, Tianqi Deng, Honggeun Jo, and Javier E. Santos)

Model Selection

In the competition, we used an artificial neural network (ANN) model to predict the sonic logs. The model has two hidden layers: the first layer has 24 neurons, the second has 12, and the output layer has two. The activation function for the first two layers is the Rectified Linear Unit (ReLu), and the output layer's activation function is sigmoid. We compared the ANN model's performance with other models, including LSTM and CNN, and found no obvious difference when validated with the given data. Given that the size of the training set is small, and its simplicity for both conceptualizing and training, we believe a shallow neural network is good enough for this problem.

To achieve better performance, we mainly focused on feature engineering instead of tweaking the hyperparameters. New features were engineered through petrophysics and signal-processing methods, which we believe are very crucial for the success of this method. Three key techniques were applied: 1. Median filters were used for the input logs to alleviate aliasing problems caused by data interpolation and eliminate outliers; 2. Gradients of different logs and logs at adjacent depths were generated in order to take the local heterogeneity into consideration; and 3. MinMax scaling was used to scale both the features and target data. In addition, we separated the training data into different zones based on the gamma ray response and the physical depth of the training data set and trained five different ANN models that were used to predict the test data separately.

Results

Fig. 1 shows the results for the zonation, and the crossplots between the predicted value and true value are shown in Fig. 2.

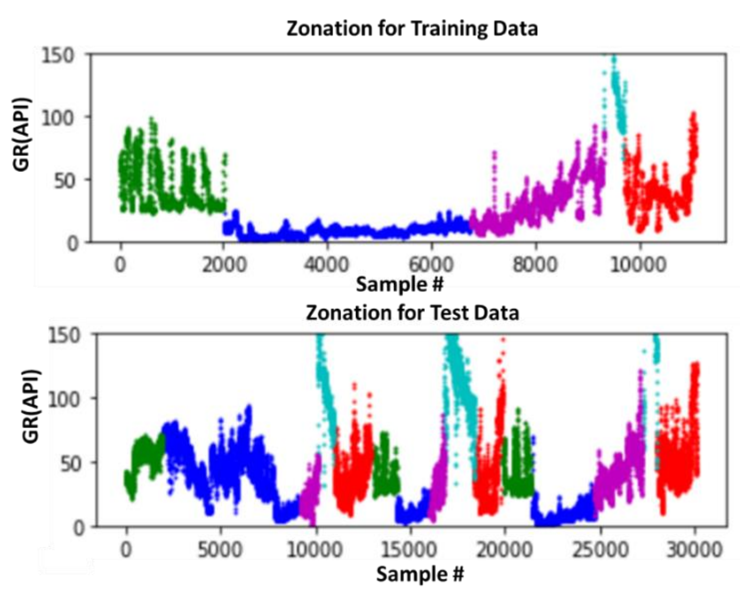

Fig. 1—Zonation performed for different wells based on GR logs.

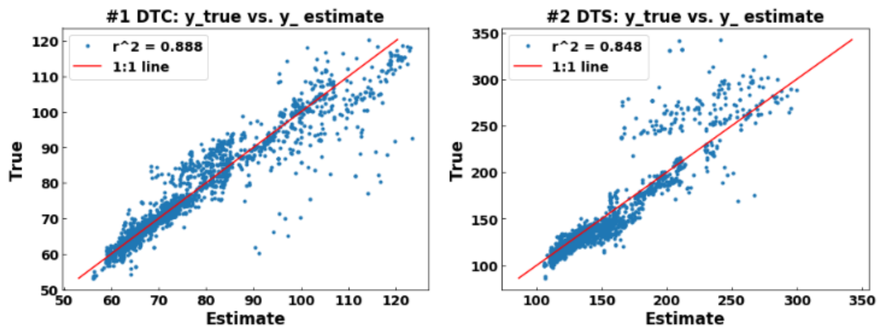

Fig. 2—Cross-plots of estimation error. No strong biases exist in DTC and DTS estimation.

Summary

To conclude the work:

1. We find the well-log data is redundant, and both the measurement and preprocessing errors need to be carefully handled before they are applied in the machine-learning model. While a deep neural network may not be able to learn the correlation among logs very well and may cause overfitting, simple shallow neural networks with two layers yield good results in this case.

2. Petrophysical and signal-processing-based feature engineering methods help machine-learning models to better learn correlations between the input features and output sonic logs. Domain knowledge from petrophysicists is essential for the successful application between machine-learning methods and well-log analysis. Machine-learning methods are just like any other empirical methods used in petrophysical research, where data calibration and preprocessing are critical.

3. Interwell correlations and different machine-learning models for different zones help to improve the performance of our models. By incorporating spatial continuity and multiwell correlation, machine-learning models may learn better for each zone.

In order to get better performance for prediction the well logs, we think two more things can be done:

1. Take into consideration any other potential features, such as spatial continuity of formation properties or tool configurations, that may cause bias due to different vendors, logging times, and spatial heterogeneity.

2. Generate more features from petrophysical analysis or statistics, such as spectral information or higher-order derivatives.

2nd Place–Team iwave (Lei Fu and Chengran Wang)

Model Selection

In the competition, we used a bidirectional long-short-term memory (LSTM) model to predict the sonic logs. The basic architecture of bidirectional LSTM consists of two hidden layers of opposite directions to the same output. It is well suited to tasks like prediction with sequence data. Because the properties measured by various well-log techniques are influenced by properties from the adjacent depth of both upside and downside, we believe with this form of generative deep learning that this model is a good fit for the well-log prediction problem.

To briefly introduce the algorithm, LSTM can be trained using similar algorithms to a recurrent neural network (RNN) that is capable of learning order dependence in sequence-prediction problems through the chain-like network structure and feedback loops. In addition, it maintains two sets of neurons from two directions that do not have any interactions. An LSTM cell can learn to recognize an important input with an input gate, store it in the long-term state, learn to preserve it for as long as it is needed, that it is maintained in the forget gate, and learn to extract it whenever it is needed. This greatly helps the model to avoid the long-term dependency problems and explains why they have been amazingly successful at capturing long-term patterns in time series, long texts, audio recordings, and more.

In addition, to find the best-fit model, we spent a lot of effort removing anomalies and choosing a good combination of features. We also experimented with different hyperparameters like input sequence length, number of hidden layers, number of nodes, batch size, dropout ration, regularization weighting, learning rate, and early stop (number of epochs).

Results

In Fig. 3, we show the collinearity between DTC and DTS of input data. Blue and red dots represent data points, the first 5,000 data points, and the rest. Different lithological characteristics are observed. Figure 6 shows the prediction results on the training data set, where we can see a great match between the predicted value and true value.

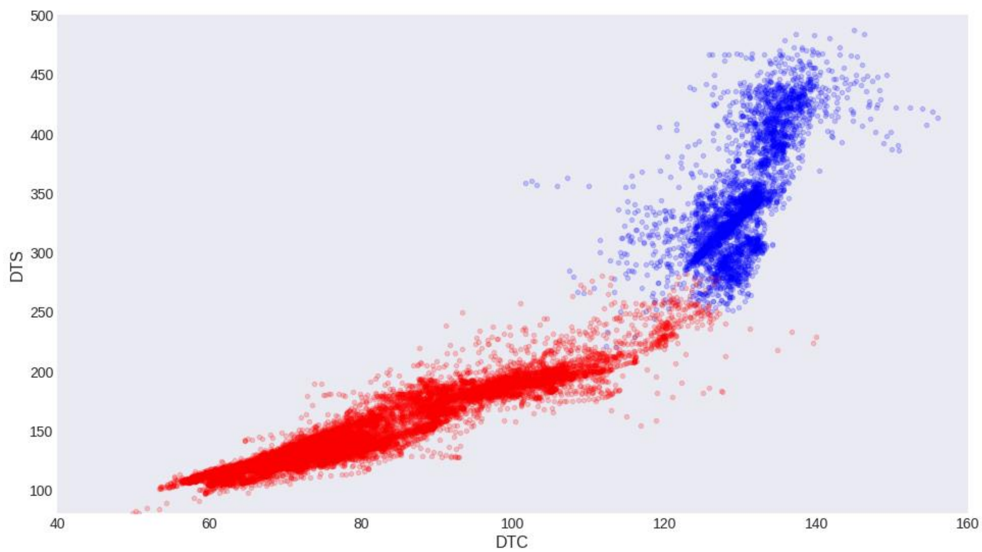

Fig. 3—Collinearity between DTC and DTS of input data.

Summary

To summarize this work, we believe the crucial part is a good understanding of the problem and the data itself by domain knowledge and data analysis. The choice of model depends on the characteristic of the problem. Data quality control and hyperparameter tuning are important.

We believe that the results can be further improved if the depth information could be provided. Depth information may be crucial for our method. We spent a lot of effort trying to separate the input data into different parts, but we gave up because of time limitations.

3rd Place–Team RocketAbuser (Arkhat Kalbekov and Valeria Suarez)

Model Summary

Random forest regressor was our method of choice to develop the prediction model. This method is less sophisticated than neural networks (NN), yet still capable of doing the job as well as NN. Additionally, random forest applies multiple tree-bagging methods that average away the variance, hence decreasing overfitting. Another benefit of using random forest is that it outputs a hierarchy of important features. Obviously, since we were trying to predict the sonic tool results, the other porosity-related tools (density neutron) had the highest importance.

Originally, we generated models for DTC and DTS. However, you will observe that we also created models for the top and bottom sections of the DTS log since those two areas provided the highest error when compared to the real data. Also, we used different feature sets for top and bottom zones as these sections contain some hydrocarbons based on crossovers in density-neutron curves. We decided that resistivity can be contributing more over these zones, thus added them to the features list. The results of the last two models were then placed into the original results for DTS.

Fig. 4 features our workflow to process the data and build the model. Key steps that led to the success of the model include:

1. We defined a function to plot all the logs and used crossplots to analyze the data. Then, we created a crossplot of neutron porosity (CNS) and bulk density (RHOB) and studied the possible outliers. This allowed us to constrain our train data (placed in the variable df) to the following ranges: CNS: (–0.2, 1), RHOB: (1.75, 3), and DTC: (40, 160). Then, the three crossplots were shown comparing these three variables.

2. Transformation for outlier detection, where we use the preprocessing kit from sklearn. With it, we scaled the data using the MinMaxScaler function, and then we transformed it using the fit_transform function. The results from each process were placed in new variables, and then we created new crossplots between CNS and RHOB to ensure the changes to the train data.

3. We applied the EllipticEnvelope algorithm to detect outliers in the train data, particularly for CNS, RHOB, and DTC. The outliers were flagged using the value –1. Three crossplots were displayed in this section that represent the three variables with the outliers flagged in purple.

After several versions of our code, we noticed that our resulting RMSE was higher using this outlier-detection method. Hence, for the data training, we only used the constrained data (as explained previously), and then we reinitialized the train data. Then, we plotted the train data and test data side to side so that we could see the correlation between the data sets. We decided to train the model, taking observations only for zones that were similar to the one in the test data set, as can be seen in Fig. 5.

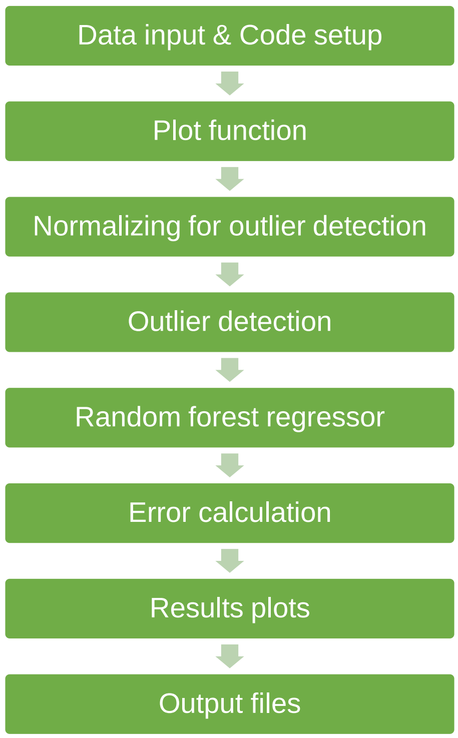

Fig. 4—Workflow of the code.

Results

We screened test data that contains 11,088 observations to keep only those observations that contained the same rows as in the real data set based on columns CAL, CNC, GR, HRD, HRM, PE, and ZDEN. Then, we removed duplicates and rearranged the real data set so that it was sorted by a column HRD. Then, we assigned the index from screened test data to real data and saved it as a new sorted real results csv. This allowed us to calculate the RMSE more accurately. The notebook cell 47 shows how the screening was done. The rest was done in Excel.

This below plot showcases the crossplots between the predicted and real data for both DTC and DTS logs. It also shows the well logs comparing each data set.

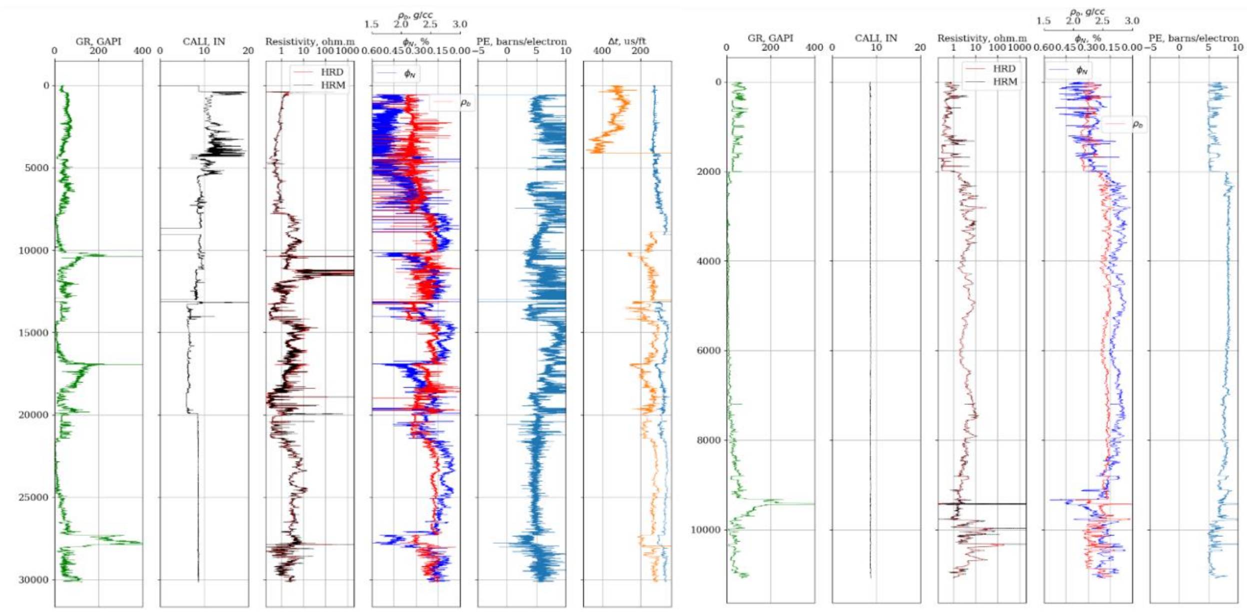

Fig. 5—Comparison of train data (left) with test data (right). DTS for the marked zones was the most inaccurate and problematic. Thus, similar lithology zones were used for predicting DTS. Similar zones are marked with the same colors.

5th Place–Team SedStrat (Epo Prasetya Kusumah, Mohammad Aviandito, and Yogi Pamadya)

Model Summary

The algorithm that we used in this competition was a combination of multiple machine-learning algorithms, also known as an ensemble method. The algorithm did not make the prediction accurately by itself. Data cleanup, feature selection, and feature engineering are as important as algorithms selection.

Fig. 6 illustrates our workflow.

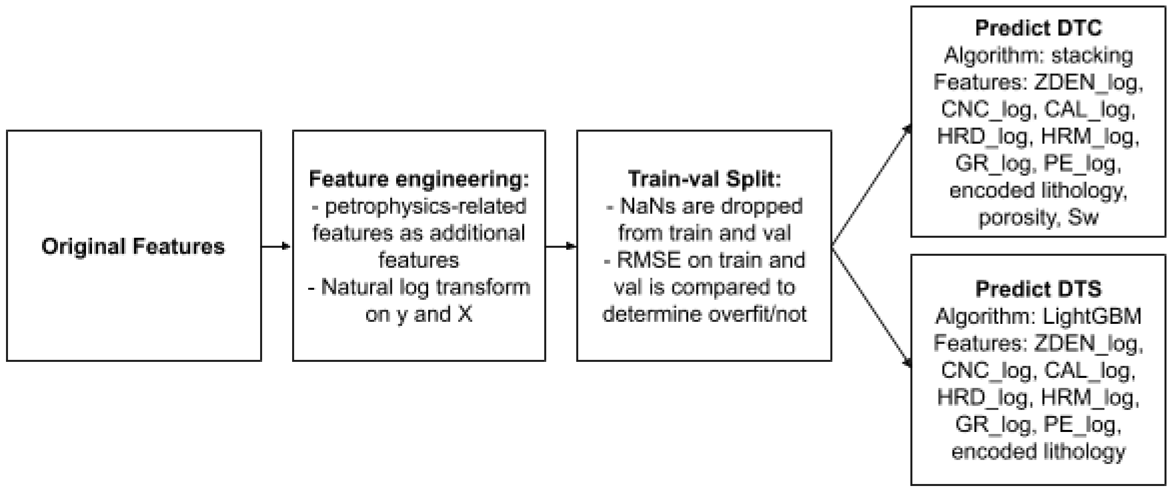

Fig. 6—Workflow for our model generation.

We use two slightly different workflows for the DTC and DTS predictions.

To predict DTC, we used two steps of prediction shown in Fig. 7. The first step is prediction using an ensemble of popular machine-learning algorithms, namely Random Forest, GBR, AdaBoost, ExtraTree, Lasso Regression, Ridge Regression, and Linear Regression. Results from these algorithms will then be used as features for the second-step prediction. The model that was used to wrap the ensemble was LightGBM.

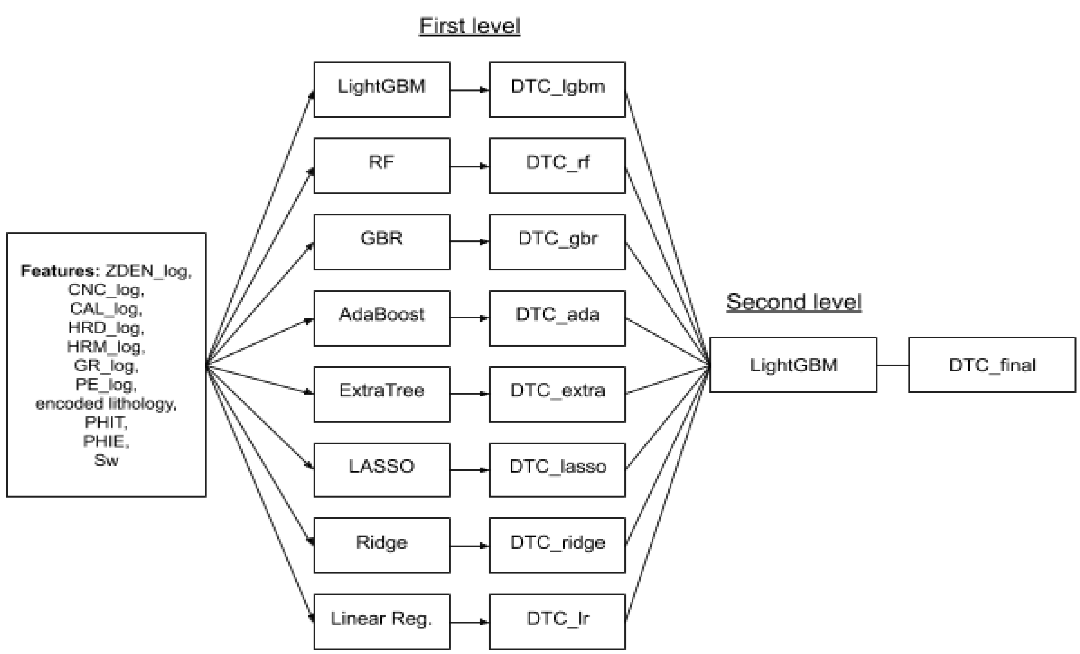

Fig. 7—Workflow for DTC prediction.

On the other hand, the prediction for DTS was more straightforward, as shown in Fig. 8. We did not use an ensemble as it yields worse results compared to a straightforward prediction by using LightGBM.

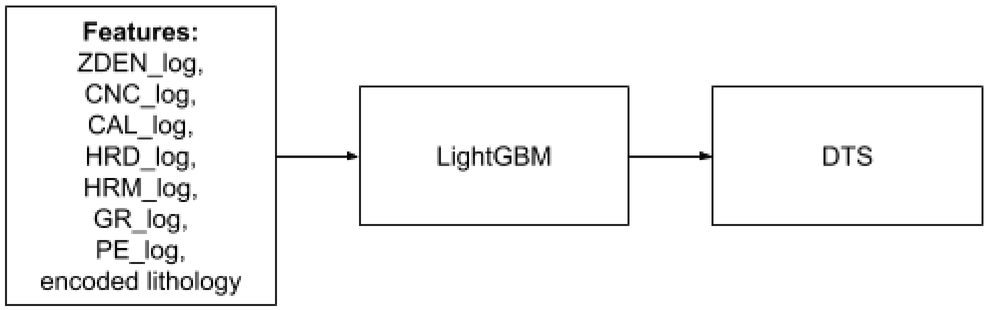

Fig. 8—Workflow for DTS prediction.

During this competition, we did not build the ensemble algorithm right away from the beginning. We did try some simpler methods in the beginning and had it scored on the leaderboard. Once it gave a benchmark score, we then tried to increase the complexity of our algorithm one step at a time. Starting small and simple is always a good mantra.

We found that each algorithm worked well in certain conditions. Regression models and tree-based models have their own strengths and weaknesses. In our opinion, the most reasonable way to make acceptable results in making predictions in this complex situation is by making a consensus from many predictions.

One interesting yet important thing to note from this workflow is that the ensemble method only gave a marginal improvement on the scoreboard over straightforward prediction by using LightGBM. We still think that our data preprocessing and feature engineering is also an important factor in our prediction.

Results

Figure 9 shows our DTC and DTS prediction values against the 20% test data set that was provided by the committee. The black diagonal line shows a perfect fit between actual and prediction data. As we can see visually in Fig. 9, the result from machine-learning prediction is quite good, and this competition can open a whole discussion on the practicality of using synthetic DTC and DTS in actual use cases, e.g., advanced petrophysics, geomechanics, etc., in the absence of sonic log data.

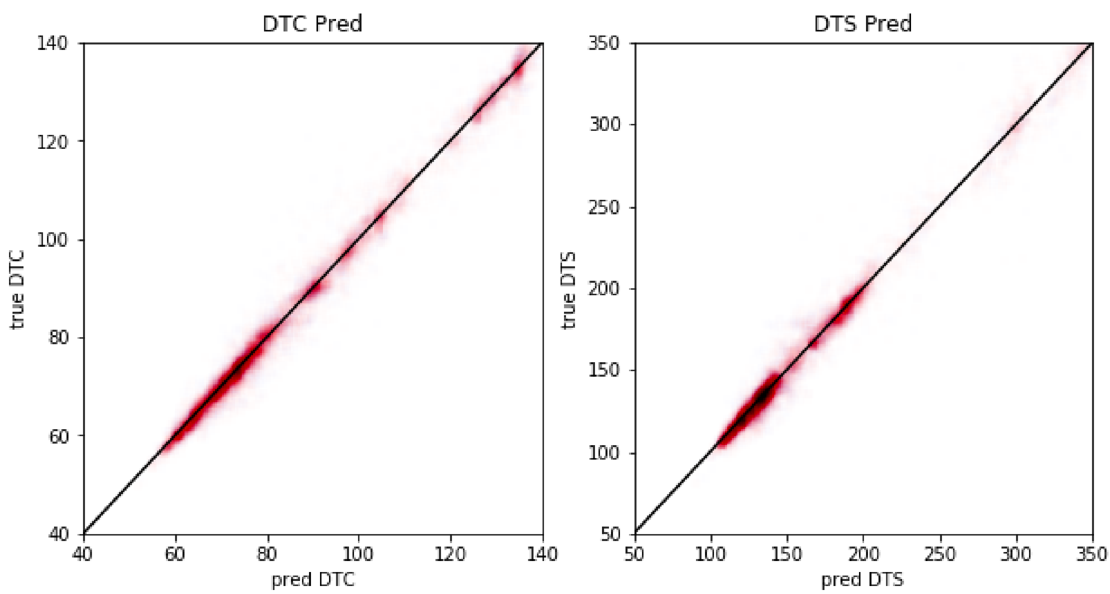

Fig. 9—DTC and DTS prediction vs. true value.

The RMSE scores for DTC and DTS predictions are 0.051 and 7.64, respectively. This contrast difference shows that our algorithm has more difficulties in predicting the DTS.

Our team recognized that data cleanup, feature engineering, and feature selection hold an important role in the prediction. The features that we created as part of feature engineering includes lithological identification using NPHI-RHOB plot, fluid identification, and simple petrophysical analysis. We did some trial and error scoring for every feature we added. The most dramatic increase in the score along this competition is when we introduce lithological information as features for prediction as part of the feature engineering process. We also found that something as simple as a natural logarithm transformation can improve the prediction dramatically, as it creates a more "normal, algorithm-friendly" data distribution to numerical data.

7th Place–Team iPetro (Hossein Izadi)

Model Summary

In this contest, an Elman neural network was used with a Levenbreg-Marquardt training algorithm for Vp and Vs prediction. The network consists of an input layer, a hidden layer, and an output layer. Six inputs, including GR, LLD, NPHI, RHOB, PEF with and without Vp, are assigned to the first layer for Vs and Vp prediction, respectively. In addition, the sigmoid function is used as the activation function. A nonlinear combination of neutron-GR and neutron-Vp are added to the output data for Vp and Vs prediction, respectively. This improves learning the ability of the network. My RMSE for Vp and Vs prediction was 15.3872.

The detailed workflow includes: In the preprocessing step, initially, with the comparison of the test database, I have completed the train database in those intervals in which there are no invalid data points. Afterward, I have used sensitive analysis approaches to determine the most important logs that should be considered for the prediction and different clustering approaches to split the database into different sets. In the training step, different structures of neural networks were trained, and eventually, different methods of assigning inputs and outputs to the network were accomplished.

The proposed workflow consists of three main parts: the preprocessing part, the clustering part, and the prediction part. In Fig. 10, the different parts of the proposed workflow are presented. Some key techniques include:

1. Preprocessing: Based on my reservoir engineering knowledge, I have removed the intervals corresponding to neutron values more than 0.7. A better approach is correcting them based on some available charts and prior knowledge we may have based on neighboring wells; however, in this contest, there was no such information. In addition, completing the no-measurement intervals is another part of the preprocessing step. In order to complete the no-measurement intervals in the train database, an Elman neural network was trained based on neutron, gamma ray, deep resistivity, photoelectric factor, and density logs. Besides, for Vp prediction, Vs was also used, and for Vs prediction, Vp was used as input data. The RMSE for Vp and Vs prediction in this part are 2.9379 and 12.2374, respectively.

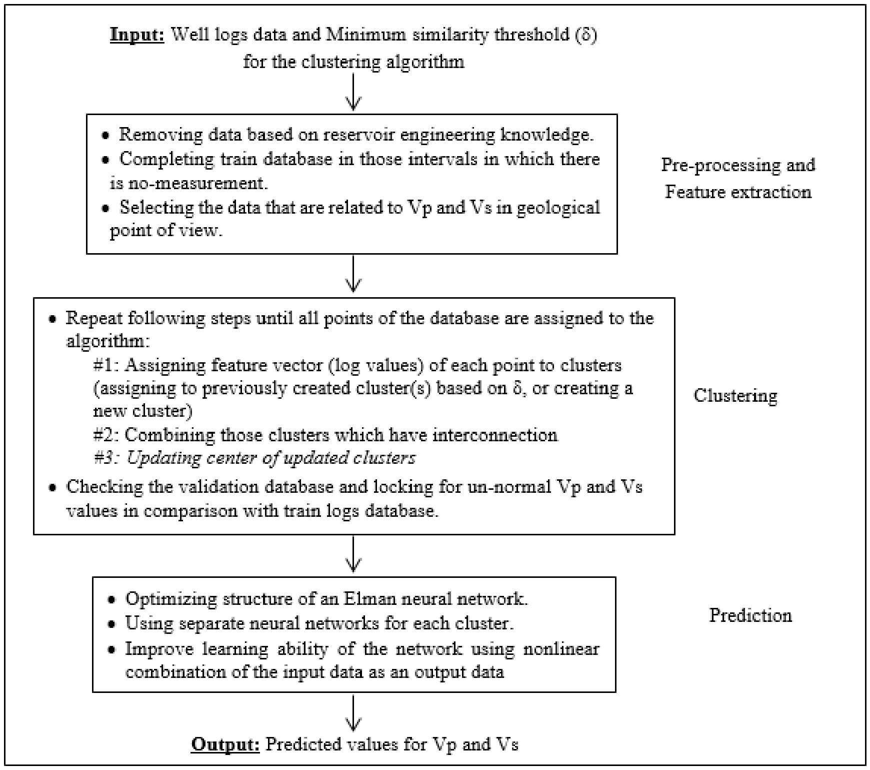

Fig. 10—The proposed workflow for Vp and Vs prediction.

2. Clustering: Clustering is always recommended in geosciences and reservoir engineering purposes (Izadi et al., 2020). This is because of the nature of uncertainty that exists in geosciences and reservoir engineering. One of the most important things is knowing how many clusters may exist in our database. Really, it is difficult to anticipate the number of clusters in a geological database. In this contest, I have used two different clustering concepts: incremental clustering and clustering based on prediction results of validation databases, respectively. The two clustering results can be found in Fig. 11 and Fig. 12, respectively.

3. Elman Neural Network: The Elman neural network contains a time feedback corresponding to recurrent connections between layers, and the recurrent connection allows the Elman network to both detect and generate time-varying patterns (Mehrgini et al., 2019). Therefore, the subsequent behavior can be shaped by previous responses, shown in Fig. 11. These recurrent

connections provide a memorize network (Elman, 1990). For more information about an Elman network, please see Mehrgini et al. (2019).

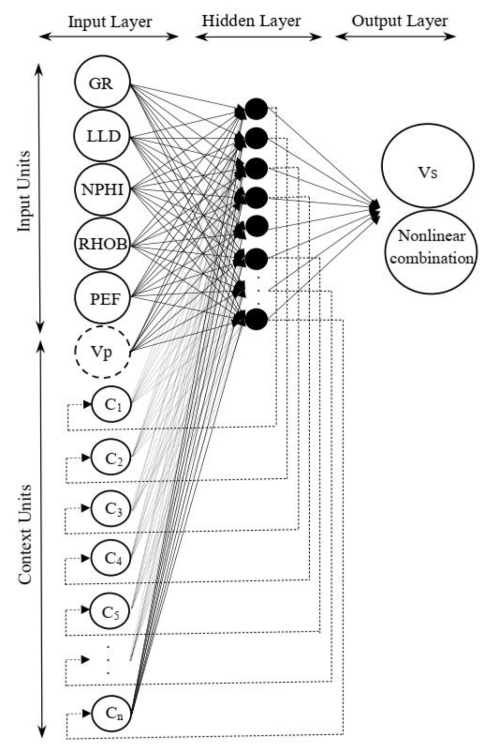

Fig. 11—The structure of the proposed Elman network for Vp and Vs prediction.

Results

The clustering results using the incremental clustering algorithm is plotted in Fig. 12 below; however, the Vs prediction did not meet the expectation under this clustering method. The crossplot for Vs prediction in the validation database is presented in Fig. 13. There is an underestimation for intervals corresponding to high GR and neutron values. The database can be divided into two clusters: less than 230 and more than 230. My final result reported an RMSE equal to 15.3872.

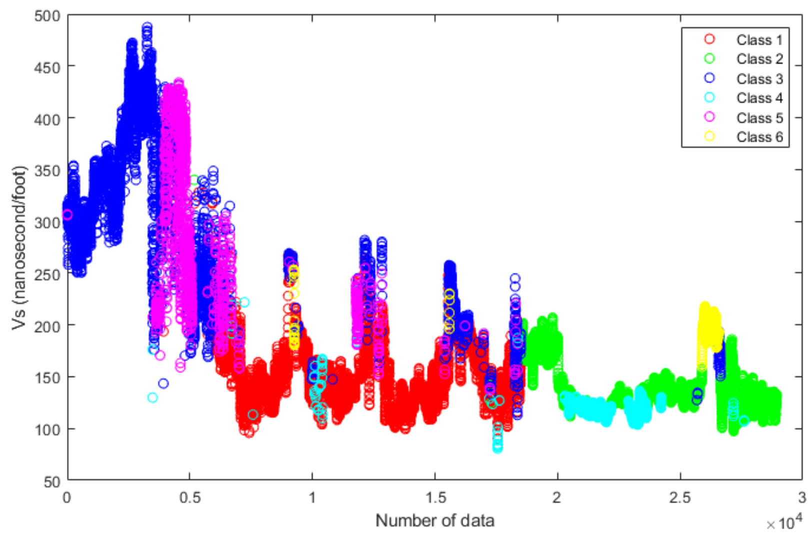

Fig. 12—Clustering results using an incremental clustering algorithm. Based on these classes, Vs prediction failed.

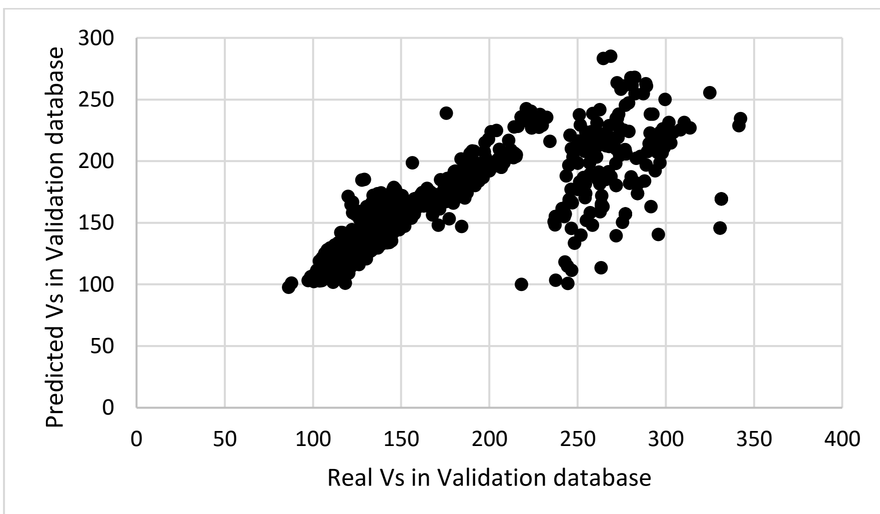

Fig. 13—Crossplot for Vs prediction in the validation database. There is an underestimation for intervals corresponding to high GR and neutron values. The database can be divided into two clusters: less than 230 and more than 230.

SUMMARY

This contest demonstrated a distributed and collaborative technology development effort. Over a two-month span, teams from all over the world worked diligently on improving their models. We saw some great results and a significant improvement in their models' performance compared to the benchmark model. Figs. 14 to 19 show more details about the comparison of the performance of different models in generating pseudo-compressional and shear sonic logs.

The teams were able to demonstrate their machine-learning workflow on a practical petrophysical problem: data set preparation and quality assurance, feature engineering with outlier handling and clustering, training and testing a regression model, and finally, blind-testing (similar to the real-world deployment) the model on the hidden data set. Various models have been adopted by the different teams, and we found that for this particular petrophysical problem, where the data set was relatively small (with a training data set of ~20,000 samples from only three wells), the model itself might not be the key to the success of predicting on the new data set, but rather many other methods that are applied to improve the performance and stability of the model, such as making special treatments for the anomalies and outliers, train different models for zones that show a very distinct DTC/DTS range, train multiple regression models, and/or combine them.

We observed that some modeling inconsistency could also be due to borehole quality and poor raw measurements in both the training and testing data. These can result in erroneous well-log responses that do not represent the formations. Being able to predict these logging errors does not aid in formation evaluation. Caution must be exercised to minimize these artifacts before starting training and testing. As petrophysicists, we want to be able to use the data to accurately describe the rocks, pores, and fluids.

The proprietary nature of the oil and gas industry, in general, limited many machine-learning methods to be adopted in the petrophysical domain. With open data sets becoming more readily available, we hope this contest provides an example of the enthusiasm and talent to help build up a shared knowledge base of the industry.

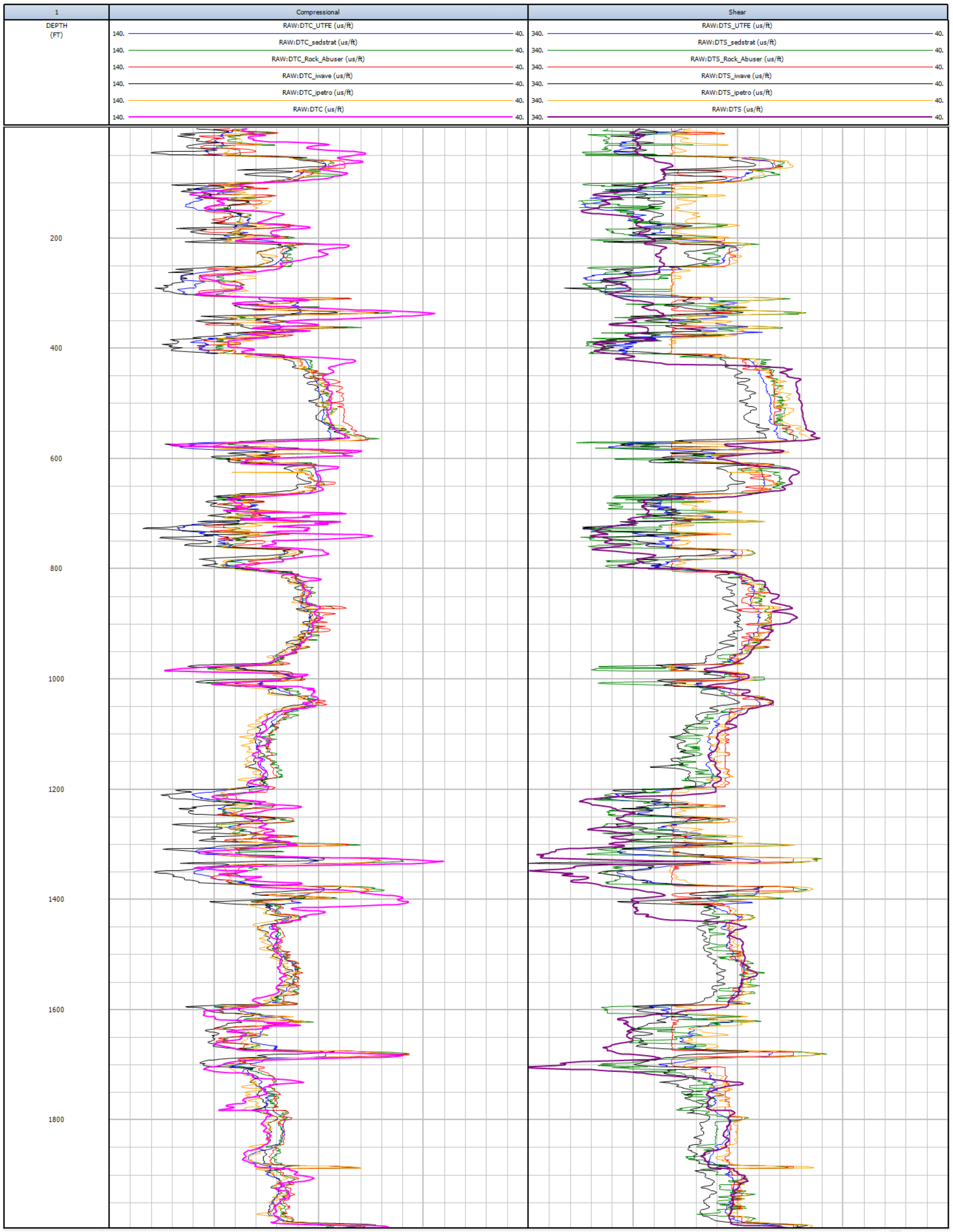

Fig. 14—Comparison of all winning teams in a log profile over the shallow interval DTC and DTS.

Compressional track: DTC predictions UTFE: Blue; SedStrat: Green; Rock Abuser: Red; iwave: Black; ipetro: Orange; and Original: Fuchsia.

Shear track: DTS predictions UTFE: Blue; SedStrat: Green; Rock Abuser: Red; iwave: Black; ipetro: Orange; and Original: Purple.

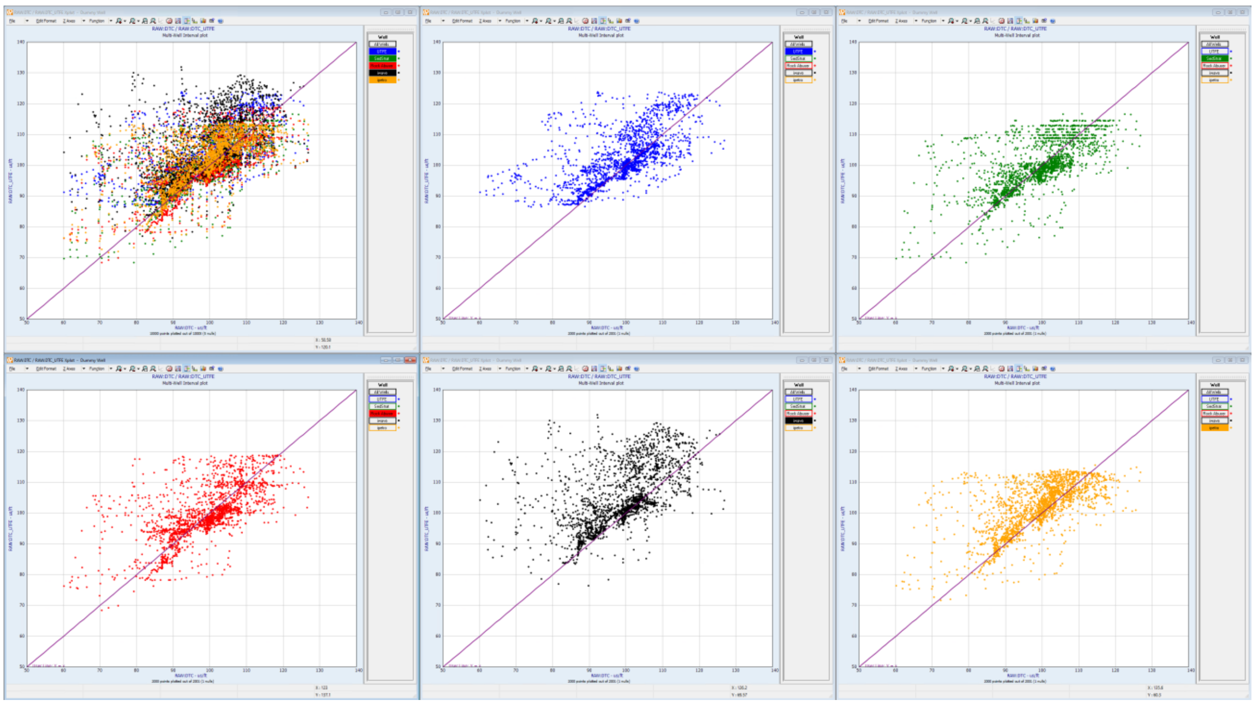

Fig. 15—DTC comparison of all winning teams in the crossplot space over the shallow interval.

Original DTC: *x*-axis, Predicted DTC: *y*-axis.

DTC predictions UTFE: Blue; SedStrat: Green; Rock Abuser: Red; iwave: Black; and ipetro: Orange.

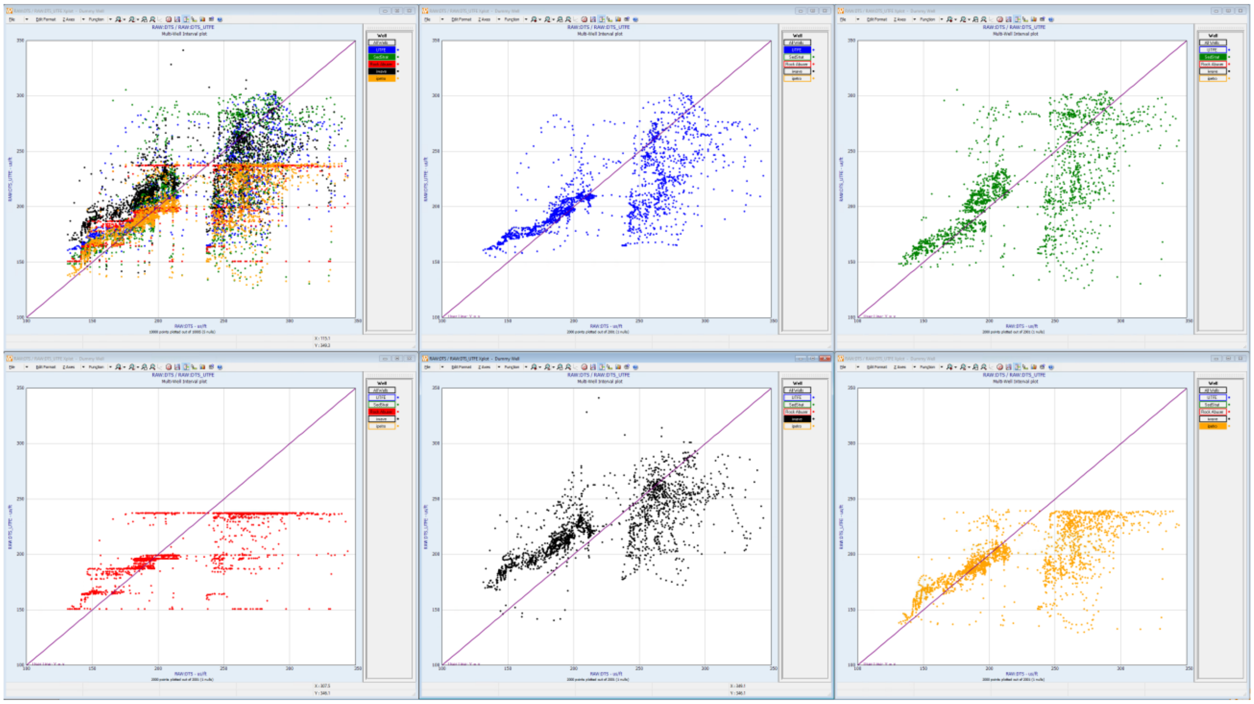

Fig. 16—DTS comparison of all winning teams in the crossplot space over the shallow interval.

Original DTS: *x*-axis, Predicted DTS: *y*-axis.

DTS predictions UTFE: Blue; SedStrat: Green; Rock Abuser: Red; iwave: Black; ipetro: Orange.

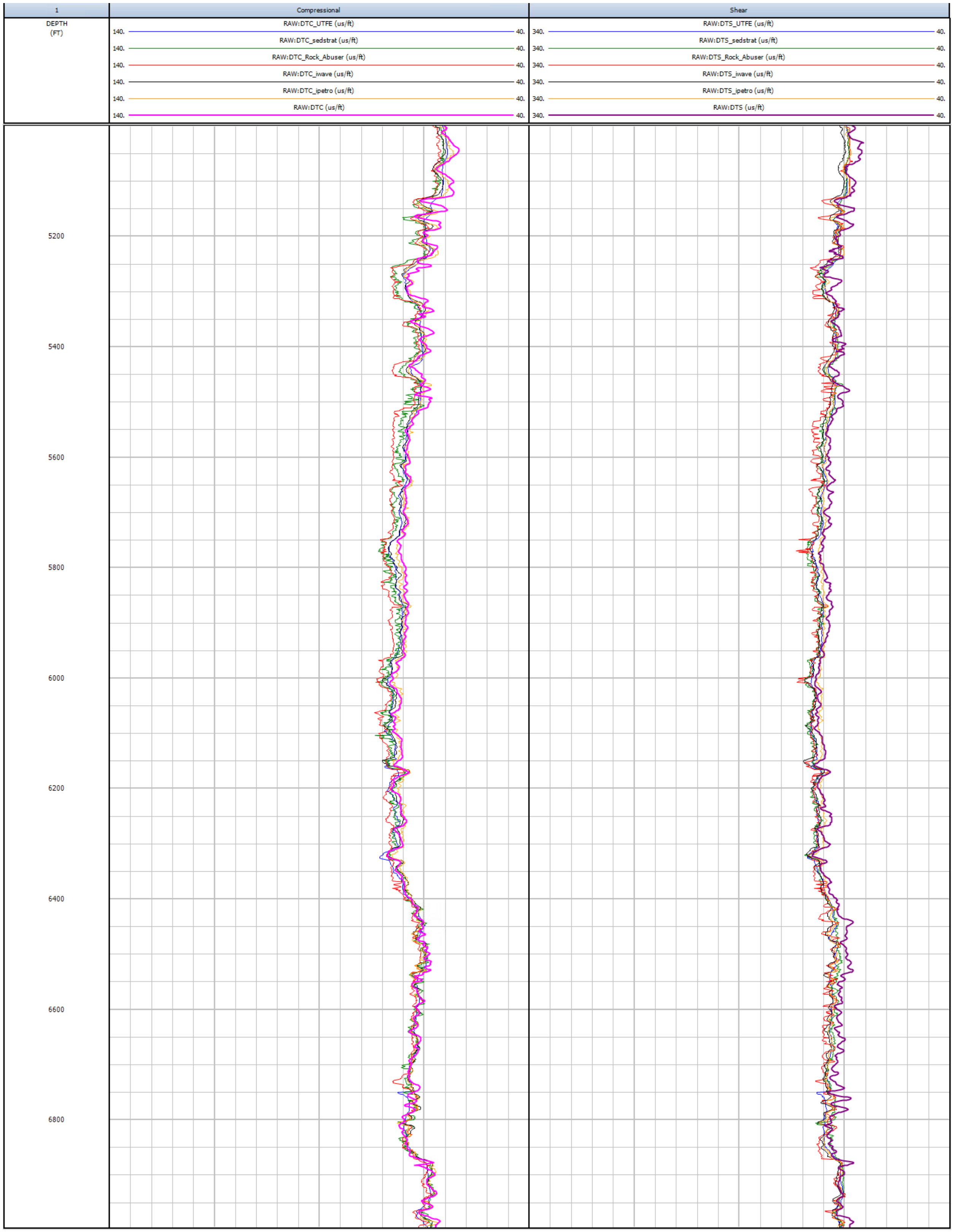

Fig. 17—Comparison of all winning teams in a log profile over the deeper interval DTC and DTS.

Compressional track: DTC predictions UTFE: Blue; SedStrat: Green; Rock Abuser: Red; iwave: Black; ipetro: Orange; Original: Fuchsia.

Shear track: DTS predictions UTFE: Blue; SedStrat: Green; Rock Abuser: Red; iwave: Black; ipetro: Orange; Original: Purple.

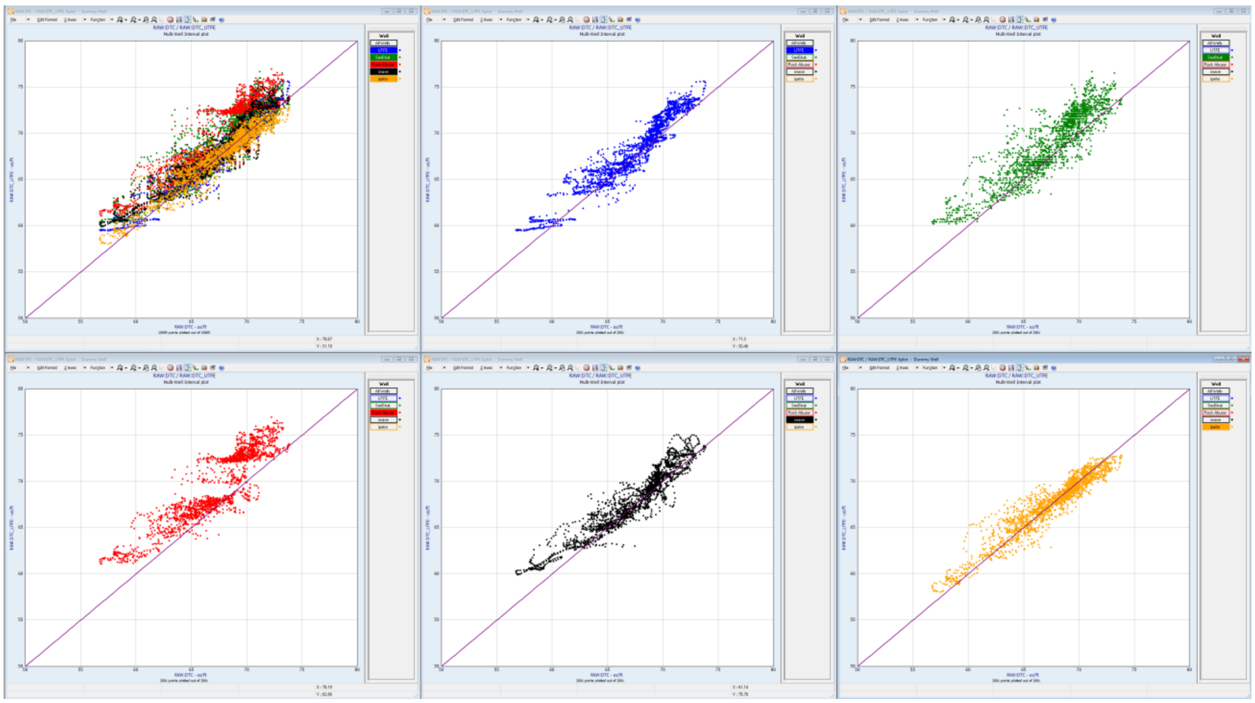

Fig. 18—DTC comparison of all winning teams in the crossplot space over a deeper interval.

Original DTC: *x*-axis, Predicted DTC: *y*-axis.

DTC predictions UTFE: Blue; SedStrat: Green; Rock Abuser: Red; iwave: Black; ipetro: Orange.

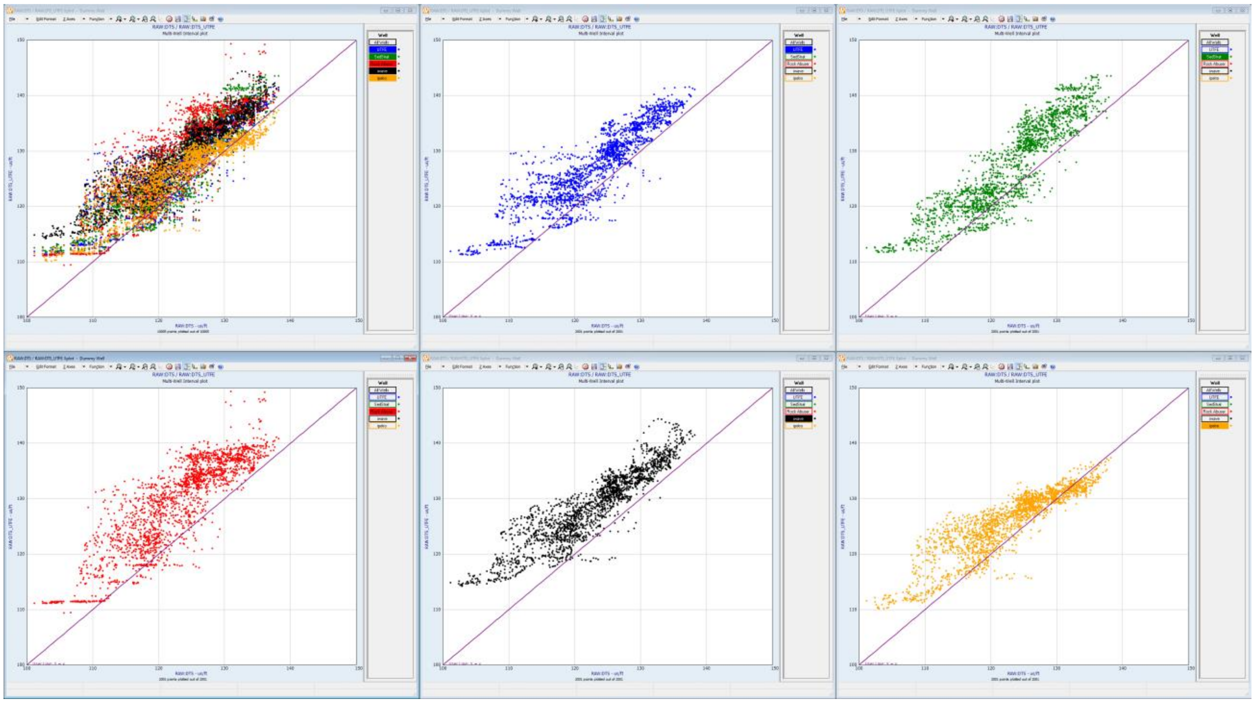

Fig. 19—DTS comparison of all winning teams in the crossplot space over a deeper interval.

Original DTS: *x*-axis, Predicted DTS: *y*-axis.

DTS predictions UTFE: Blue; SedStrat: Green; Rock Abuser: Red; iwave: Black; ipetro: Orange.

ACKNOWLEDGMENTS

A note of thanks goes to SparkCognition for sponsoring the event and to Equinor for releasing the Volve data set. We also thank the other members of the SPWLA PDDA SIG ML Contest Committee: Brendon Hall, Bin Dai, Zheng Gan, and Yan Xu for their contribution.

REFERENCES

Elman, J.L., 1990, Finding Structure in Time, *Cognitive Science*, **14**(2), 179–211. DOI: 10.1016/0364-0213(90)90002-E.

Github website for SPWLA ML Contest 2020: https://github.com/pddasig/Machine-Learning-Competition-2020.

Izadi, H., Sadri, J., Hormozzade, F., and Fattahpour, V., 2020, Altered Mineral Segmentation in Thin Sections Using an Incremental-Dynamic Clustering Algorithm, *Engineering Applications of Artificial Intelligence*, **90**, 103466. DOI: 10.1016/j.engappai.2019.103466.

Mehrgini, B., Izadi, H., and Memarian, H., 2019, Shear Wave Velocity Prediction Using Elman Artificial Neural Network, *Carbonates and Evaporites*, **34**(4), 1281–1291. DOI: 10.1007/s13146-017-0406-x.

Misra, S., Li, H., He, J., 2019, Machine Learning for Subsurface Characterization, 1st Edition, Gulf Professional Publishing.

Xu, C., Misra, S., Srinivasan, P., and Ma, S. (2019, March 15). When Petrophysics Meets Big Data: What can Machine Do?, Paper SPE-195068 presented at the SPE Middle East Oil and Gas Show and Conference, Manama, Bahrain, 18–21 March. DOI: 10.2118/195068-MS.

Yu, Y., Misra, S., Oghenekaro, O., and Xu, C., 2020, Pseudosonic Log Generation With Machine Learning, A Tutorial for the 2020 SPWLA PDDA SIG ML Contest. *SPWLA Today*, **3**(2), 97–101. https://www.spwla.org/Documents/Newsletter/SPWLA_Today_Newsletter_Vol3_No2.pdf.