Machine Learning Driven Exploration Workflow Accelerator Applied at Well and 3D Seismic Scales - Leading to Cost-Efficient New Barrels to Existing Infra-Structure

Executive Summary

Executive Summary

This project was sponsored by the Norwegian Petroleum Directorate (NPD) and performed by Earth Science Analytics. The aim of the project was to identify oil/gas pay intervals using Machine-Learning (ML) methods on the EarthNET platform, in 1D for 545 wells, and in 3D using seismic in Block 35/11). Identified pay intervals were characterized, classified and ranked using different reservoir property cut off ranges.

It is acknowledged that the predicted results contain errors and that the data provides a starting point for further investigation. In order to gain confidence in the predictions the data should be placed in the context of the geological setting and with other datasets to form a consistent geologically orientated story.

The ability to predict and characterize hydrocarbon accumulations in the subsurface using well and seismic data has been the holy grail for all explorationists worldwide. We show here results from a Machine Learning (ML) driven workflow (Figure 1) to predict properties such as lithology, porosity and importantly water saturation (Sw) at well and seismic scales leading to a much needed accelerator to the exploration workflow.

In essence the workflow was performed within the cloud-native EarthNET platform enabling wide access to ML technology to all stakeholders from domain knowledge experts to exploration geologists, geophysicists, and decision makers.

Critical to the success of ML studies is the ability to integrate datasets of differing scales (seismic to pore space). EarthNET has the functionality to enable seamless integration of diverse datasets. Facilitating access to all stakeholders in the project ensures that critical knowledge is leveraged, geological sense check is performed, and the creation of improved models by recycling of workflows. By combining these elements with an understanding

of the geological setting will undoubtedly lead to exploration workflow acceleration and to identifying new exploration opportunities.

The adoption of this ML workflow allows the learnings from many hundreds of wells across the NCS to be condensed into an ensemble of models and their resulting predictions. These property models can then be iteratively refined, new data added, and subsequently applied to other areas saving time and acting as an exploration accelerator.

The workflow outlined here can be easily embraced by explorationists and specialists alike forming part of a company's exploration strategy.

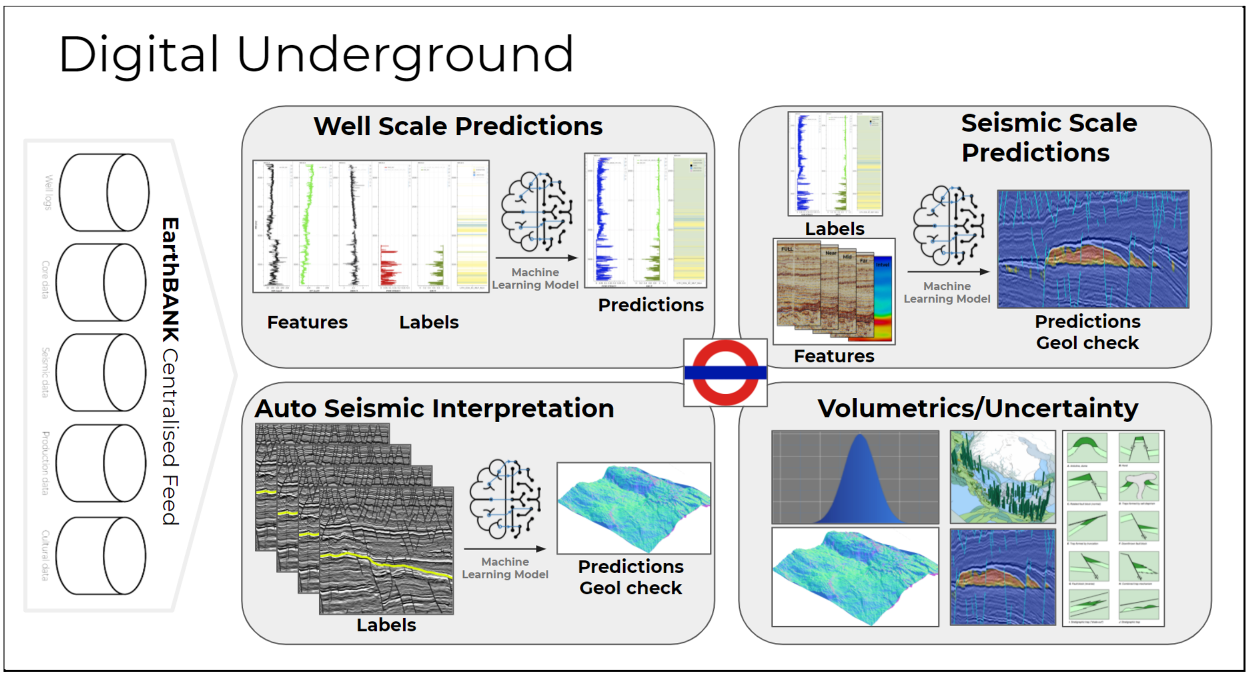

Figure 1 shows the ML workflow adopted to accelerate hydrocarbon exploration.

Data Review and QC

A structured data bank was created within EarthNET consisting of data from 545 wells. The following data types were imported into the data bank (see below). The data coverage of each log was determined and listed in brackets. The log and CPI data was provided for Earth Science Analytics (ESA) by the Norwegian Petroleum Directorate (NPD).

- Bit Size (806 km)
- Caliper log (963 km)
- Gamma Ray log (1738 km)
- Density log (772 km)
- Sonic log (1291 km)
- Sonic shear log (253 km)
- Deep resistivity logs (1620 km)
- Medium resistivity logs (1526 km)
- Neutron log (625 km)
- CPI logs (porosity and Sw curves) ((71 km) and (35 km))

In total ca.10,000 kms of log data was used in the project. The aerial extent of the wells used in the project are shown in Figure 2.

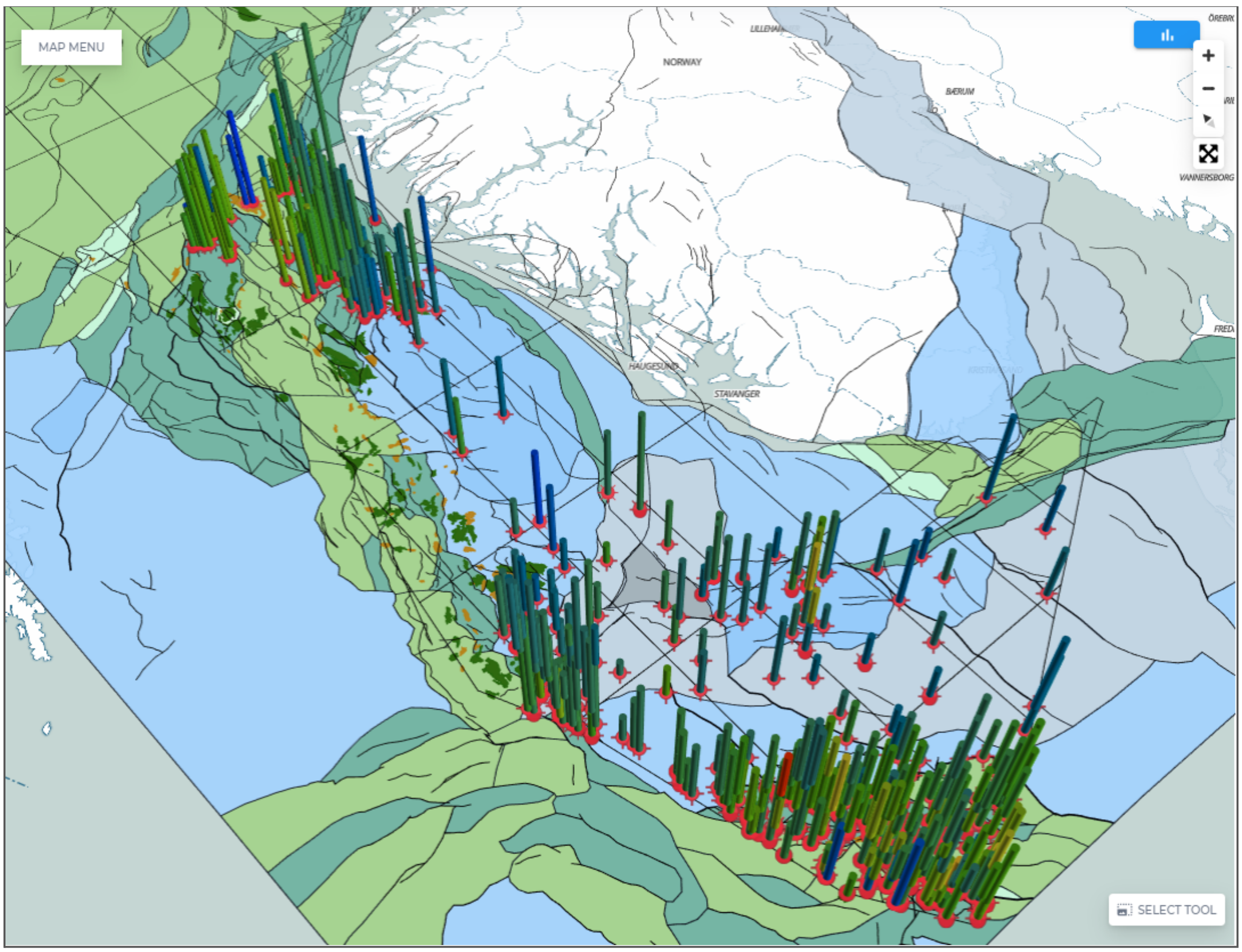

Figure 2 shows the distribution of the wells throughout the Norwegian North Sea used in the project.

Feature Data

Feature data is the input data used for training ML models, i.e. the target variables (labels) are predicted as functions of the feature data.

Data quality control in EarthNET is divided into two procedures.

- Semi-automated approach is used where non-physical values as well as anomalous values are deleted.
- Individual log is reviewed with the intention to remove areas of a log that seem unrealistic, for example straight or diagonal lines that may be induced through interpolation methods.
- Wellbore QC flags were created using caliper and bit size data to ensure training of models on high quality feature datasets.

Label Data

Label data is the target variable data used for training ML models, i.e. the target variables (labels) that are predicted as functions of the feature data.

- The lithology labels were derived from the ESA data library and were supplemented with new labels specifically for this project that concentrated on gleaning labels from chalk, Upper Jurassic sandstone, halite, and anhydrite lithologies.
- Porosity and Sw data were derived from CPI logs available in the study.
- The CPI data (porosity and Sw labels) were also quality controlled in the same manner as the wireline log data listed above. In total 71.2 kms of PHIE logs and 31.3 kms of Sw log were used as labels to create models.

1D Well-Based Predictions

Lithology, Porosity and SW Models

The workflow to generate the best models/property curves is shown in figure 3. In essence multiple models were created based on different feature combinations, each one using only the wells containing the target labels. For the models related to each feature set; hyperparameters were tuned, multiple algorithms tested and a hole condition flag used. Each model was assessed based on its Train and Blind Test Score as well as a geological sense check (comparing the target curve vs the Blind Test curve and its variance away from the target curve). Once the best models for the porosity and Sw target parameters were identified these models were then applied without the hole condition flag to all of the 545 wells in the project. This resulted in 6 curves being created that were then summed to produce a Target_Mean and a Target_SD (SD = Standard Deviation) master curves.

For the lithology predictions the best Blind predictions were seen using the MLP algorithm. A majority voting scheme was applied leading to the creation of a final lithology prediction classification. Priority is given to each of the applied models. High priority was given to models that were created using the greatest number of features. The multi class lithology classification models include the confidence of predicted classes.

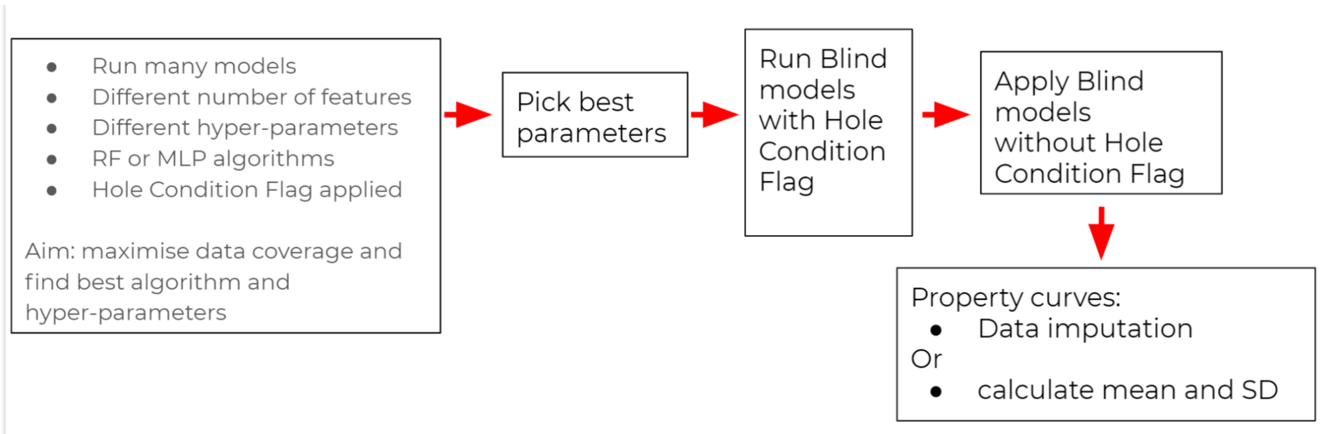

Figure 3 shows a high level workflow describing how the best models were created.

ESA Applied SW Models

In addition, Sw predictions were gleaned from the application of the pre-trained ESA Sw models.

Sw Label Harvesting and Models

Due to the lack of Sw label data, data harvesting was performed in order to increase the number of Sw label data points. The main goal with SW training lebel harvesting is to build more balanced training data in all lithologies.

3D Property Prediction

The prediction of lithology, porosity and Sw away from the wellbore can give great insights into the hydrocarbon prospectivity of an area. 3D property prediction was performed on Block 35/11 in the Norwegian North Sea.

At a high level figure 4 illustrates the 3D property prediction workflow.

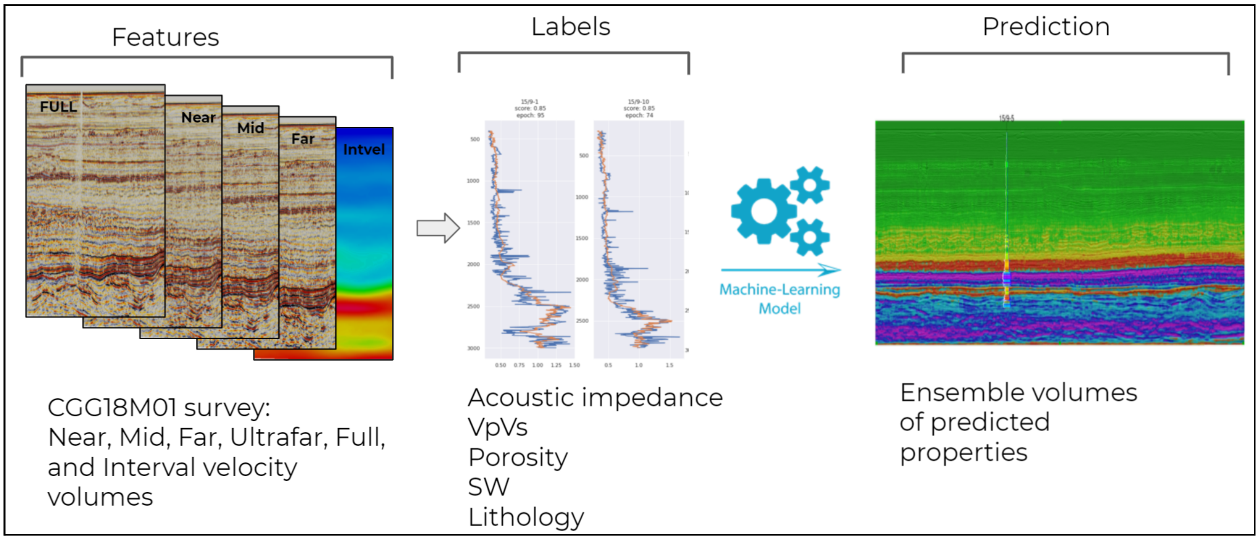

Figure 4 showing at a high level the ML 3D property prediction workflow.

Well Tie

Prior to the creation of models all wells in Block 35/11 were tied to the full stack seismic volume. This is an important step and time and effort was placed on this element of the workflow.

Lithology

The best lithology model was created using 11 wells from Block 35/11. The label set used as the training data was curve LITH_ESA_6C_MLP_MLA from the 1D prediction work. The features used to create the lithology prediction model were;

- CGG18M01_NVG_PSDM_NEAR
- CGG18M01_NVG_PSDM_MID
- CGG18M01_NVG_PSDM_FAR
- CGG18M01_NVG_PSDM_ULTRAFAR
- CGG18M01_NVG_PSDM_FULL
- Interval velocity volume

Porosity

Four different models were produced for the prediction of porosity. The label set used as the training data was curve PHIE_ED_MEAN_CPI_FILL from the 1D prediction work. The features used to create the lithology prediction model were;

- CGG18M01_NVG_PSDM_NEAR
- CGG18M01_NVG_PSDM_MID
- CGG18M01_NVG_PSDM_FAR
- CGG18M01_NVG_PSDM_FULL
- Interval velocity volume

Water Saturation (Sw)

Five different models were produced for the prediction of Sw. The label set used as the training data was curve SW_COMP_ML_MEAN_CPI from the 1D prediction work. The features used to create the lithology prediction model were;

- CGG18M01_NVG_PSDM_NEAR
- CGG18M01_NVG_PSDM_MID
- CGG18M01_NVG_PSDM_FAR
- CGG18M01_NVG_PSDM_FULL
- Interval velocity volume

1D Hydrocarbon Pay Identification

In this work hydrocarbon pay zones were identified at well scale by the prediction of Mean porosity and Mean Sw curves and applying cutoffs. The porosity cutoff being 0.1 and water saturation cutoff being 0.6.

Five pay classes were defined based on cutoff values from the SW_MEAN and POR_MEAN curves (Figure 5). A NON-PAY class was defined as having 0.9-1 Sw and porosity values from 0 to 0.5.

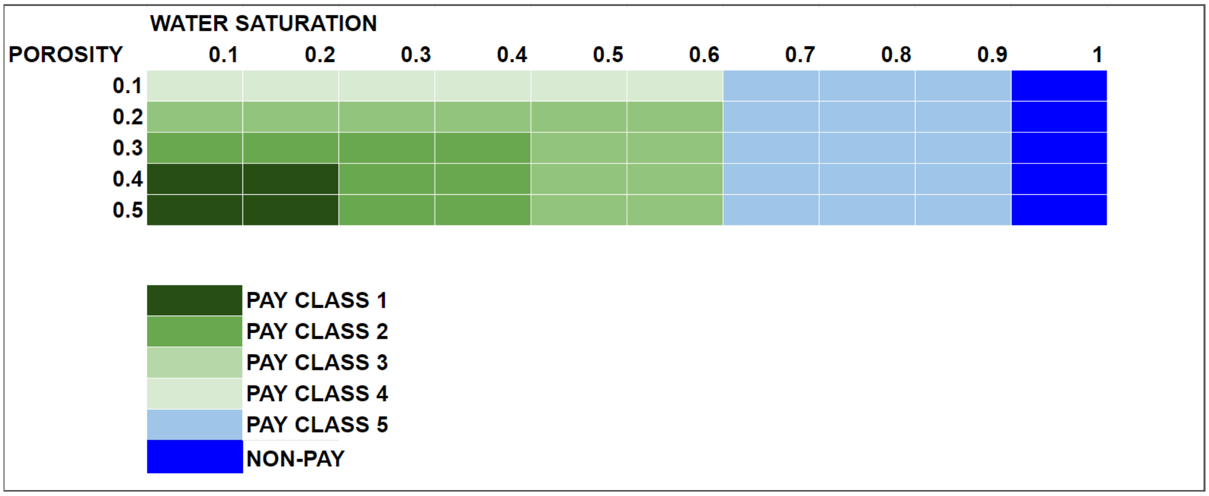

Figure 5 showing the defined pay classes based on cutoffs of Sw and porosity.

Wireline and predicted logs used in the project

- BS - Bit Size
- BS_CALI - Bit size/Caliper Log
- BS_ED - Bit Size edited
- CALI - Caliper Log
- CALI_ED - Caliper edited
- Casing Diameter - Casing Diameter
- DEN - Density
- DEN_ED - Density edited
- DT - Sonic Log
- DT_ED - Sonic Log edited
- DTS - Shear Sonic
- DTS_ED - Shear Sonic edited
- GR - Gamma Ray
- GR_ED - Gamma Ray edited
- Hole_Condition_Flag - Based on BS and Caliper Log. Bad hole defined as caliper log being 1.6 inches larger than the bit size
- LITHOLOGY - Lithology Prediction Classes
- LITHOLOGY_CONFIDENCE - Lithology prediction confidence curve. Where 1 is high confidence and 0 is low confidence. Included is a lithology.csv file lookup table defining the lithology type vs number in the .LAS file

- NPHI - Neutron Log
- NPHI_ED - Neutron Log edited
- Pay_Class - Pay classes defined from cutoff values from the Porosity_MEAN and Sw_MEAN curves. Included is a PAY_CLASS.csv file lookup table defining the pay class type vs number in the .LAS file
- PHIE - Porosity CPI
- PHIE_ED - Porosity CPI edited.
- POR_MEAN - Porosity mean curve
- POR_MEAN_SD - Standard deviation away from the POR_MEAN curve
- RDEP - Deep Resistivity
- RDEP_ED - Deep Resistivity edited
- RMED - Medium Resistivity
- RMED_ED - Medium Resistivity edited
- Sw - Water Saturation CPI
- SW_ED - Water Saturation CPI edited
- SW_ED_COMP - Water Saturation CPI plus harvested Sw
- SW_MEAN - Summed 4 Sw predicted curves into a MEAN curve
- SW_MEAN_SD - Standard deviation away from the SW_MEAN curve.
- VSH - Vshale curve