Project Progress Report
Identifying Sweet Spots Using Well Log Datasets

CSCI 470: Introduction to Machine Learning

Team Name: Oil Digger

Nadima Dwihusna

Xiaoyu (Rosie) Zhu

Mohamed Ibrahim Mohamed

Original Timeline

Week Starting	Nadima Dwihusna	Xiaoyu (Rosie) Zhu	Mohamed Mohamed
9/17	Obtain Data Sets	Obtain Data Sets and Research	Research on ML Methods
9/24	Data Processing	Data Visualization	Research on ML Methods
10/1	Data Processing	Data Visualization	Research on ML Methods
10/8	Data Processing	Data Visualization	ML Method Selection
10/15		All: Progress Report (Due Oct18)	
10/22	Data Training	Data Prediction	Help Team Members
10/29	Data Training	Data Prediction	Help Team Members
11/5	ML Method 1	ML Method 2	ML Method 3
11/12	ML Method 1	ML Method 2	ML Method 3
11/19	ML Method 1	ML Method 2	ML Method 3
11/26		All: Visualize Results with Seismic	
12/3		All: Prepare Final Presentation	

Objective:
The main goal of this project is to use Supervised Machine Learning Algorithms to identify the formation bearing fluids (Oil, Gas, and Water). Identifying these bearing fluids has been common in the oil and gas industry. However, this has been done manually for the past decades. Manual identification of sweet spots can result in many errors as human or user errors and moreover manual identification overlooks the instrumental errors, environmental disturbance and units inconsistency. The code could be further develop to determine compaction and depth trends per facies by analyzing larger scale heterogeneous datasets. The end result will be useful for rock physics and inversion workflow to improve reservoir characterization.

Data:
Database is provided by the Reservoir Characterization Project (RCP), geophysics research group at the Colorado School of Mines. The well log datasets are in the Wattenberg Field, Colorado located south of Greeley and north of Brighton, Colorado. All data is labeled and the machine learning algorithm in this project aims to inherent structure from the input data. The four well log datasets to be used are:

- Badding USX W
- Warner 16-14
- Badding 15-26
- Badding 20-26

Features Used:
The well log datasets contain several records throughout depth. Each record includes the well Name, as well as the Depth (ft) of the readings, and the following features:

- **Gamma Ray (api):** measure the natural radioactivity of a rock formation to distinguish between shale and non-shale rock formations useful to indicate top of oil reservoir.
- **Resistivity (siemens/m):** measures how much sediments in the formation opposes electric current useful to indicate porosity of formation, water saturation, and presence of hydrocarbons.
- **Neutron Porosity (percent):** indicates the porosity or the void fraction in the formation; typically hydrocarbons are present in higher porosity zones.
- **Density (g/cc):** measures the bulk density of the formation which includes minerals and fluid.
- **Sonic P and Sonic S:** measures wave velocity for seismic profiles.

Current Model:
For this progress report, we performed data processing and visualization. We label the data with 0, 1, 2 stand for oil, water and gas, and plot the 5 features and target of all 4 wells. In this project, we decided to use supervised learning to get our target which is the Hydrocarbons types; (Oil, water or gas). As we have 5 features in the data, we are also going to perform autoencoder or PCA to extract the features. We used three algorithms to predict our labels. First we used Support vector Machine, where we also iterated to obtain the best parameters for the model naming gamma and C (Cross Validation). Second we used Random Forest Classifier and finally we used K nearest neighbor where we also iterated the to get the best k values. Below the well log datasets are visualized along with the labeled interpreted formation bearing fluids.

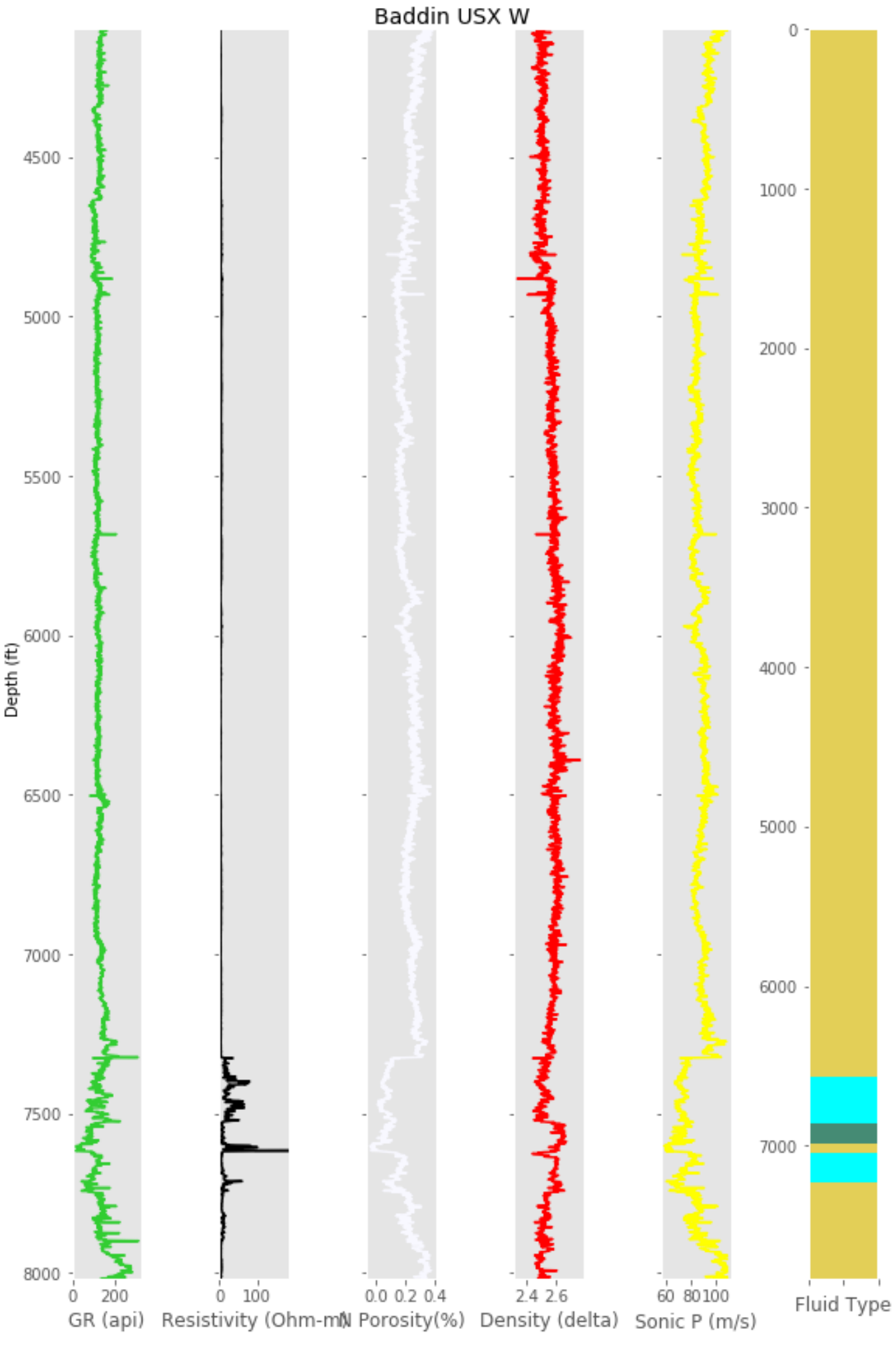
Actual Timeline As of now, Oil Digger group is on track with the original proposed timeline. The Wattenberg Field well log datasets were obtained, processed, and visualized preliminary research regarding the geology of datasets were analyzed, and the different Unsupervised Machine Learning Algorithms to be used were determined. The group will continue on with the timeline, training the data and performing the different algorithms.

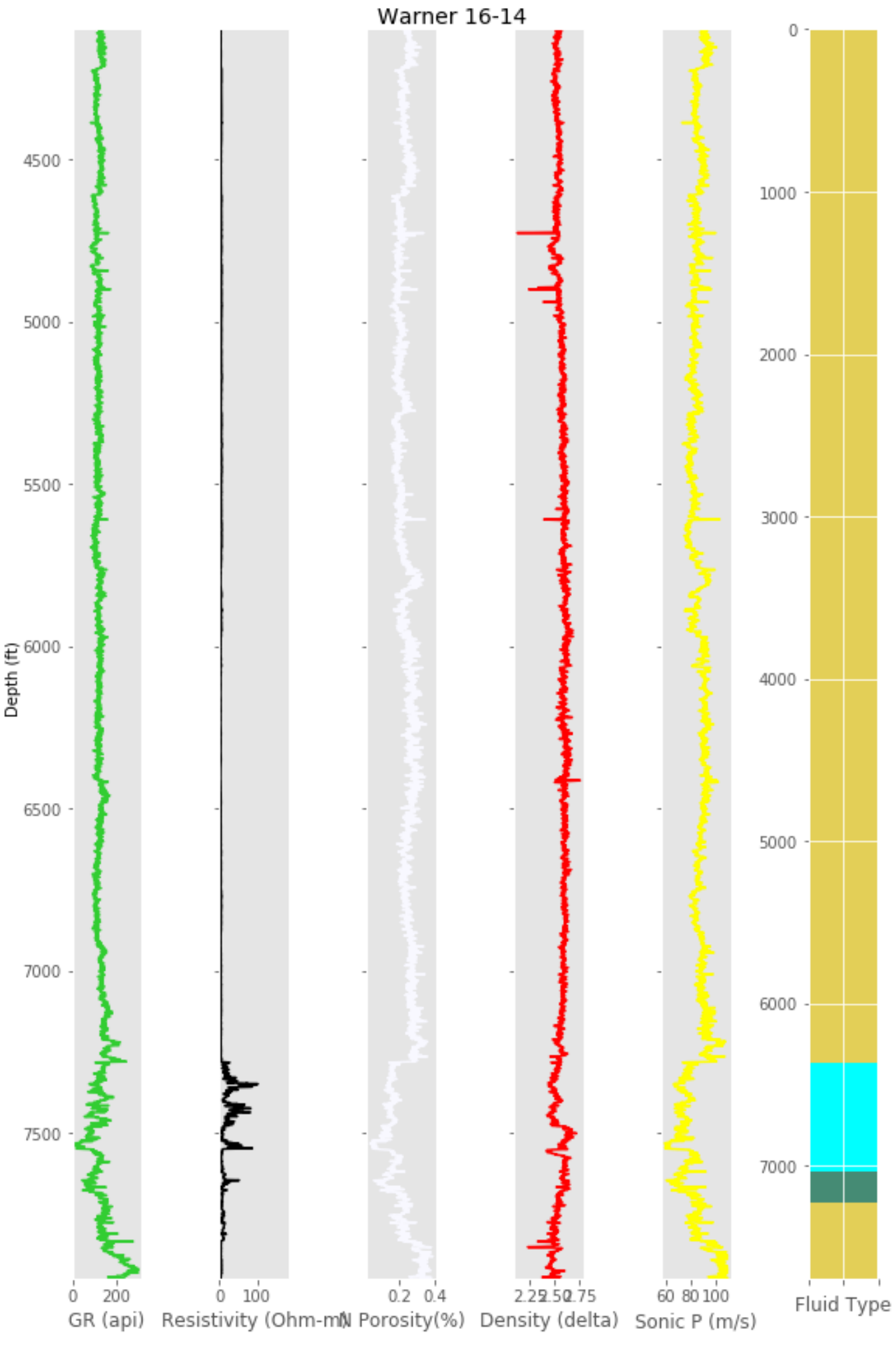
Week Starting	Nadima Dwihusna	Xiaoyu (Rosie) Zhu	Mohamed Mohamed
9/17	Obtain Data Sets	Obtain Data Sets and Research	Research on ML Methods
9/24	Data Processing	Data Visualization	Research on ML Methods
10/1	Data Processing	Data Visualization	Research on ML Methods
10/8	Data Processing	Data Visualization	ML Method Selection
10/15	All: Progress Report (Due Oct18)		
10/22	Data Training	Data Prediction	Help Team Members
10/29	Data Training	Data Prediction	Help Team Members
11/5	ML Method 1	ML Method 2	ML Method 3
11/12	ML Method 1	ML Method 2	ML Method 3
11/19	ML Method 1	ML Method 2	ML Method 3
11/26	All: Visualize Results with Seismic		
12/3	All: Prepare Final Presentation		

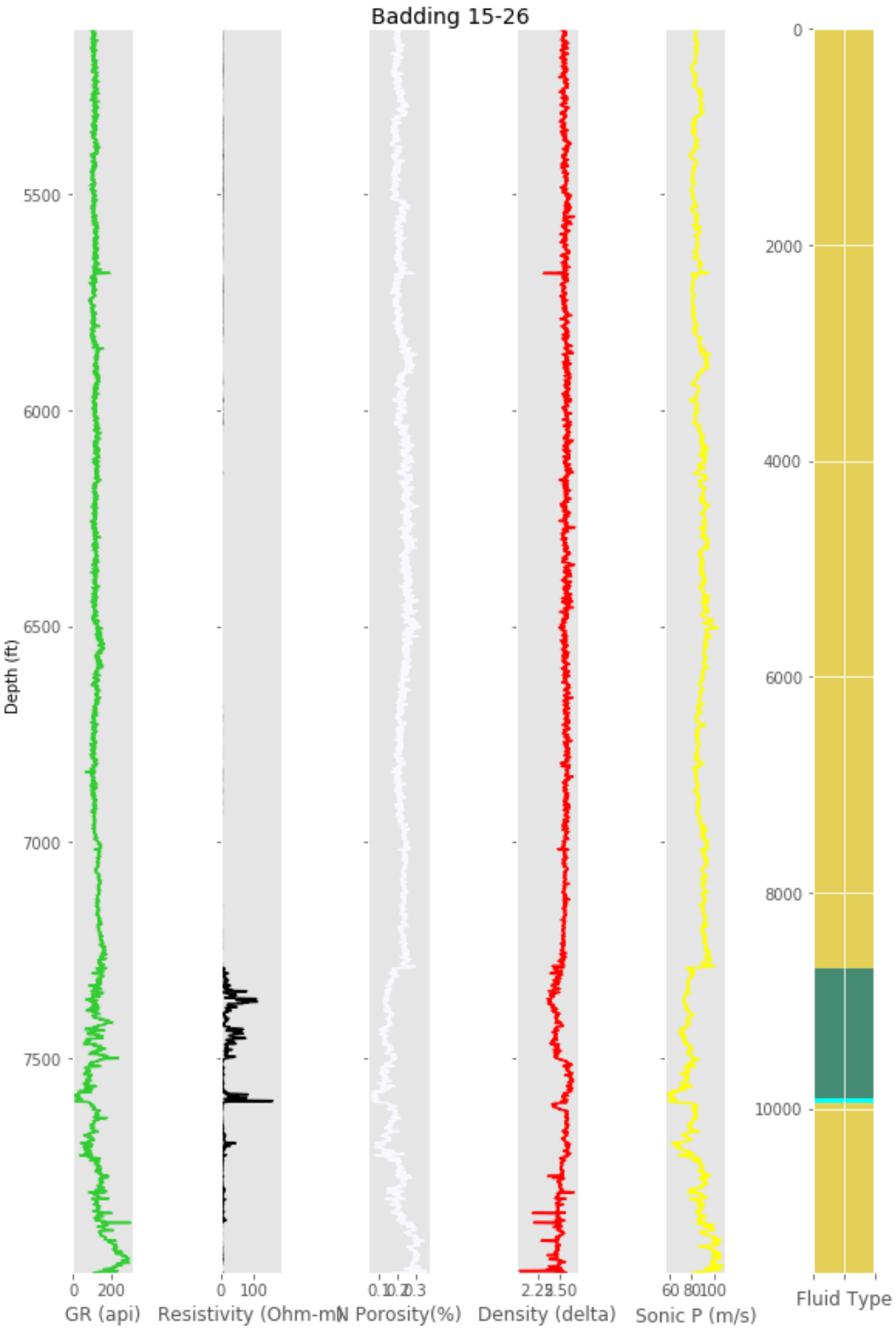
Issues, Constraints & Mitigation Strategies:
There are no issues that came up thus far. For the future, we plan to do further supervised learning algorithms and develop the code even more to us machine learning to identify the formation bearing fluids.

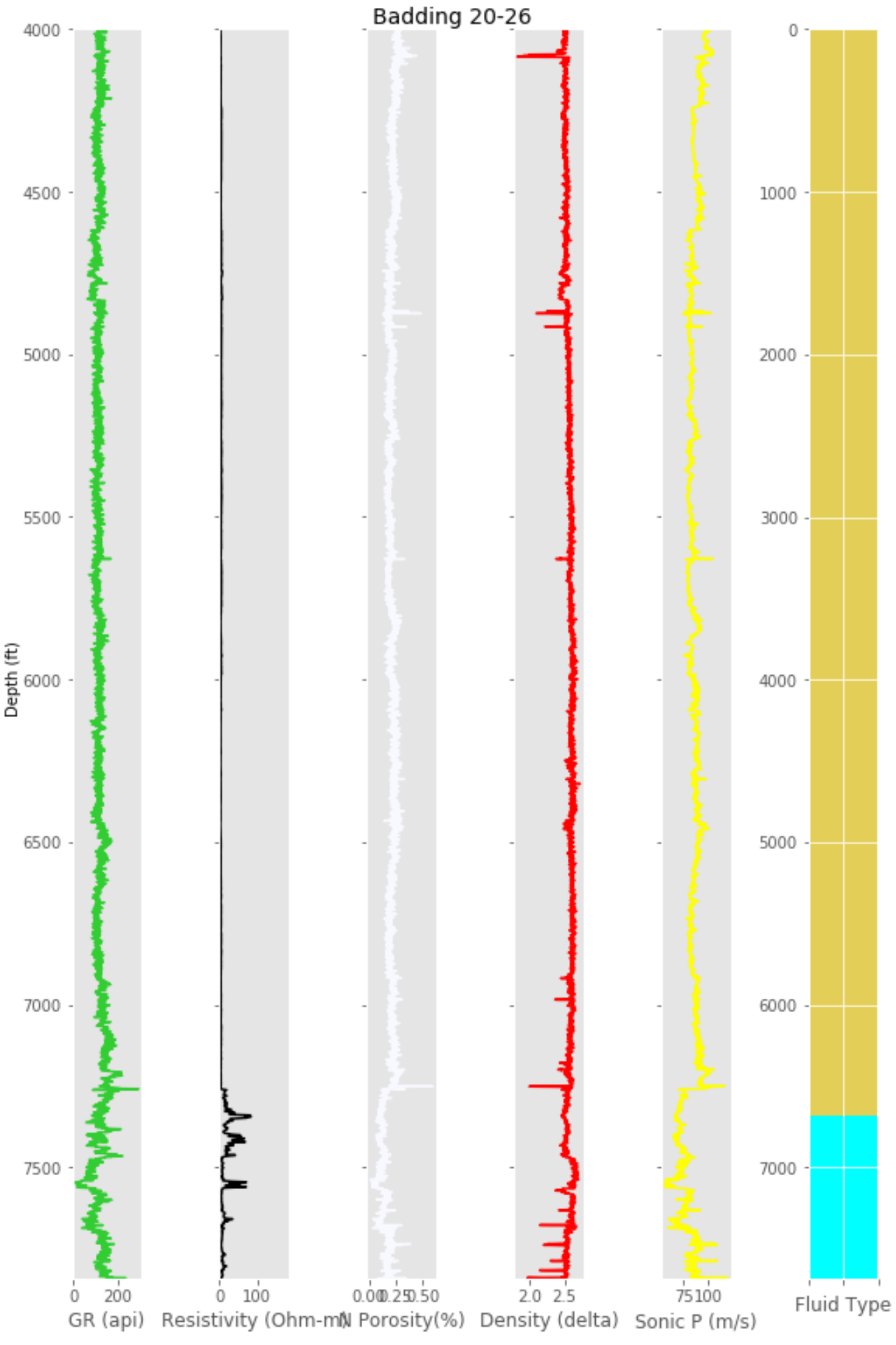
Updated Project Expectation and Future Timeline:
The Oil Digger group will continue on with the original timeline, training the data and performing the different algorithms. The future project expectation will take approximately six more weeks until all the results are obtained and visualized along a possible Seismic Profile for further accurate geological interpretation.

Week Starting	Nadima Dwihusna	Xiaoyu (Rosie) Zhu	Mohamed Mohamed
10/22	Data Training	Data Prediction	Help Team Members
10/29	Data Training	Data Prediction	Help Team Members
11/5	ML Method 1	ML Method 2	ML Method 3
11/12	ML Method 1	ML Method 2	ML Method 3
11/19	ML Method 1	ML Method 2	ML Method 3
11/26	All: Visualize Results with Seismic		
12/3	All: Prepare Final Presentation		


Baddin USX W
Depth (ft)
4500
5000
5500
6000
6500
7000
7500
8000
0 200
GR (api)
0 100
Resistivity (Ohm-m)
0.0 0.2 0.4
N Porosity(%)
2.4 2.6
Density (delta)
60 80 100
Sonic P (m/s)
0
1000
2000
3000
4000
5000
6000
7000
Fluid Type


Warner 16-14
Depth (ft)
4500
5000
5500
6000
6500
7000
7500
0
200
GR (api)
0
100
Resistivity (Ohm-m)
0.2
0.4
N Porosity(%)
2.25
2.50
2.75
Density (delta)
60
80
100
Sonic P (m/s)
0
1000
2000
3000
4000
5000
6000
7000
Fluid Type


Badding 15-26
Depth (ft)
5500
6000
6500
7000
7500
0
200
GR (api)
0
100
Resistivity (Ohm-m)
0.1
0.2
0.3
Porosity(%)
2.25
2.50
Density (delta)
60
80
100
Sonic P (m/s)
0
2000
4000
6000
8000
10000
Fluid Type


Badding 20-26
Depth (ft)
4000
4500
5000
5500
6000
6500
7000
7500
0
200
GR (api)
0
100
Resistivity (Ohm-m)
0.00
0.25
0.50
N Porosity(%)
2.0
2.5
Density (delta)
75
100
Sonic P (m/s)
0
1000
2000
3000
4000
5000
6000
7000
Fluid Type